IST-2003-507581 WINNER

D2.7 ver 1.0

Assessment of Advanced Beamforming and MIMO Technologies

Contractual Date of Delivery to the CEC:	28.02.2005
Actual Date of Delivery to the CEC:	28.02.2005
Editor:	Martin Döttling
Author(s):	Angeliki Alexiou, David Astély, Karsten Brüninghaus, Pedro Coronel, Martin Döttling, Martin Fuchs, Jochen Giese, Martin Haardt, Marko Hennhöfer, Tao Jiang, Eduard Jorswieck, Ateet Kapur, Zexian Li, Abdelkader Medles, Tim Mousley, Magnus Olsson, Keith Roberts, Stéphanie Rouquette-Léveil, Wolfgang Schott, Aydin Sezgin, Per Skillermark, Veljko Stankovic, Patrick Svedman, Samuli Visuri, Dafei Wang, Thorsten Wild, Ernesto Zimmermann
Participant(s):	*ACL, EAB, ETHZ, IBM, KTH, LUK, MOT, NOK, PRL, SM, TUD, TUI, UOulu*
Workpackage:	*WP2 – Radio Interface*
Estimated person months:	*110*
Security:	**Restricted**
Nature:	**R**
Version:	1.0
Total number of pages:	191

Abstract: This document contains a first assessment of spatial transmission schemes regarding performance and complexity. In order to assess advanced beamforming and MIMO technologies, mathematical models, as well as methodologies for simulation and assessment have been developed. Conceptual work and simulation campaigns have been conducted. A baseline WINNER multi-antenna concept is proposed for further investigation based on multi-user spatial domain link adaptation consisting of the following basic components: linear dispersion codes, directive transmission (beamforming), multi-user precoding, and per stream rate control.

Keyword list: beamforming, spatial diversity, spatial multiplexing, vector/matrix modulation, LDC, space time coding, opportunistic beamforming, SIMO, MISO, MIMO, switched beams, grid of beams, linear and non-linear precoding, link level simulations, system-level simulations, link-to-system level interface, spatial mode, multi-antenna concept.

Disclaimer:

Executive Summary

Advanced multi-antenna solutions are an integral part of the overall WINNER system. In this document, initial investigations and assessment of adaptive spatial processing schemes are performed considering system performance and complexity. The evaluation is based on a *review* of state-of-the-art techniques, development of new and adapted *methodologies* for simulation and assessment in the WINNER context, *simulative investigations*, as well as *conceptual work*. Based on the findings, a first proposal for the *WINNER multi-antenna concept* is elaborated and presented in the form of a generic transmitter processing chain that will serve as the baseline for further investigation and system integration.

The WINNER overall system requirements mandate a single ubiquitous radio access system concept that is able to adapt to a comprehensive range of mobile communication scenarios with scalability in complexity. Under the constraint of low deployment effort and cost, it aims at complete coverage, while at the same time providing significant performance enhancement compared to legacy systems and their evolutions [D7.1]. Spatial processing has the potential to contribute significantly to this goal by:

- Improved spectral efficiency and increased user peak data rate
- Increased range or coverage in a cost-efficient manner
- Enhanced interference management
- Adaptivity to scenario and channel conditions
- Concurrent support of adaptive and non-adaptive transmissions
- Support of different terminal types (including single-antenna terminals)
- Reduced terminal power consumption to increase talk and stand-by time
- Management and reduction of human exposure to electromagnetic fields

Major requirements for the *WINNER multi-antenna concept* therefore include high performance, robustness, adaptability to a wide range of scenarios and terminal classes, as well as efficient interworking with the basic transmission and network concept defined in other tasks and work packages.

Based on a review of state-of-the-art spatial processing algorithms, an initial selection of techniques for further consideration has been identified first. Then a generic framework that allows a uniform description and mathematical modelling for most of the proposed techniques has been developed.

Due to the manifold world of spatial processing techniques it is not a trivial task to *compare and assess* the identified techniques with respect to performance, complexity, scalability, application scenarios, or required signalling and network functionality. The fact that no common simulation environment is available adds further difficulty. Therefore, a *methodology for assessment and evaluation* in the WINNER context has been developed. Assessment criteria include *user-oriented* and *system-oriented* physical layer *criteria*, like user throughput, cell throughput, and spectral efficiency. In addition, *deployment aspects* (coverage) and *robustness* are considered. A major novel aspect of this document is that in contrast to the vast amount of theoretical literature on spatial processing a strong emphasis is put on practical questions related to the *implementation* of such techniques in a radio access network. Therefore, questions concerning system complexity, measurements, feedback, control overhead, enabling system functions, robustness against impairments, flexibility and scalability are addressed in the assessment process and criteria. Due to constraints in the available traffic models and the simulation capabilities, some of the assessment criteria proposed by WP7 have been adapted for single-site, snapshot system-level simulations without dedicated modelling of packets and data traffic. Additionally, a method to obtain a first estimate of control overhead in the absence of a defined control channel structure and other relevant system parameters has been established.

For a comprehensive assessment of multi-antenna techniques, it is mandatory to consider the performance at system level, since many effects of spatial processing, like multi-user precoding, the impact of spatially-coloured interference, and the benefits of interference management techniques are not tractable at link level. Therefore considerable effort has been dedicated to the development and verification of *simulation methodologies*, in particular to the link-to-system-level interface definition. Major challenges included the support of multi-antenna techniques in a multi-state OFDM system like WINNER, while keeping the complexity of the interface within feasible limits. The chosen link-to-system interface (see Chapter 4) allows predicting the BLER in system-level simulations without performing modulation,

interleaving and coding explicitly. It is based on mapping the large number of different *SINR* values of the multi-state channel within one codeword to *one* effective *SINR* based on a specific averaging function in the mutual-information domain. This one-dimensional compression function reduces complexity considerably, while at the same time maintaining high accuracy, as has been shown by extensive validation. This link-to-system level interface has been adopted as WINNER standard for WP2 and serves as a baseline for other WPs, who adapt it to their particular needs.

Based on this methodology the partners have started to develop or adapt simulators for investigation of individual multi-antenna techniques. First *simulation results* and comparison of techniques are provided, predominantly for the downlink of the wide-area and short-range scenarios, but also for MIMO relaying. Although the results obtained until now are preliminary and suffer from considerable constraints as detailed above, several conclusions have been drawn and are reflected in the current proposal for the WINNER multi-antenna concept. The most general and important findings are shortly stated in the sequel.

In propagation conditions with small angular spread (like wide-area or rural scenarios), results prove that introducing directivity by means of beamforming is a robust and efficient means to improve coverage, user and cell throughput. In the single-site and single-cell case, for a given number of users, cell throughput scales nearly linearly with the number of beams up to a maximum that is given by the achievable beam width in combination with the required sidelobe suppression and available transmit power per beam. When using a fixed grid of beams, it is important to carefully design the grid and the beam pattern, since performance is sensitive to the inter-beam interference. Means to reduce this sensitivity to inter-beam interference are strong channel coding and link adaptation. If beamforming is combined with adaptive resource assignment based on short-term CQI, considerable performance enhancements can be obtained due to the exploitation of multi-user diversity, especially for high load and increasing angular spread in the propagation channel. In the wide-area scenario, it seems better to allocate degrees of freedom in spatial processing to intercell interference rejection compared to spatial multiplexing, since reuse of resource elements between cells is more efficient than reuse by spatial multiplexing, which in average suffers from higher cross-talk.

For the space-time-frequency coding and modulation, linear dispersion codes (matrix modulation) are considered as baseline for further investigations. Regarding their applicability to different channel conditions diagonal ABBA is more robust than DABBA or vector modulation. Linear precoding adds significant gain to pure matrix modulation (STBC). Spatial mode selection has proven substantial gain within each scenario, and therefore it is a must considering that the WINNER multi-antenna concept must support many different scenarios. Promising input parameters for spatial mode selection are received SNR and the second condition number. Conceptually, these two parameters could also serve as a simplified link-to-system-level interface for investigations that require low computational complexity in the physical layer modelling.

Initial comparisons between single-site and network system-level simulations show that a realistic modelling of inter-cell interference has dramatic impact on cell and user throughput. Single-cell or single-site evaluations may overestimate the relative benefits of spatial multiplexing and spatial division multiple access (SDMA) and underestimate the benefits of multiple antennas at the terminals. Therefore the upgrade of the simulators to full network simulations is of highest priority for future work.

One important focus of future research will be the quest for the optimum amount of adaptivity. At the example of limited CSI spatial multiplexing (PARC, PSRC), it is shown that the control overhead grows linearly with number of transmitted streams or antennas, and depending on the degree of link adaptation and flexibility in channel-dependent scheduling, this may result in unacceptably high overhead. This indicates clearly that after a certain breakpoint further adaptivity will merely increase complexity and overhead without providing corresponding performance gain. However, another associated tradeoff is that higher effort in link adaptation combined with simpler space-time-frequency codes might enable to concentrate the complexity at the network side compared to a reduced-adaptivity technique combined with spatial processing that involves higher decoding complexity.

Relaying and multi-hop communication is a concept with high potential. It gains additional performance if spatial processing is applied on the corresponding links. Even using the simple 2x1 Alamouti scheme higher throughput can be guaranteed over an extended area compared to SISO relaying. Spatial multi-

plexing combined with relaying also leads to improvements over SISO relaying in terms of cell coverage.

The multi-user multi-antenna signal design problem has been tackled both from a communication theory and a simulative perspective and the following conclusions are drawn: Multi-user precoding techniques that will be further investigated for the wide-area scenario are restricted to linear precoding (e.g., SMMSE or BD based on long-term statistics). Techniques like BD provide performance improvements primarily for multi-antenna mobile terminals. SMMSE is especially attractive since it does not suffer from restrictions considering the number of antennas in the system, provides good performance for various terminal classes, and can exploit either long-term or short-term CSI at the transmitter to perform the precoding. For the short-range scenario precoding techniques for further consideration include also non-linear techniques that rely only on the short-term CSI at the transmitter. SO THP (non-linear), SMMSE THP (non-linear), and SMMSE (linear) are of interest in realistic scenarios with a mixture of terminal classes. However, it is noted that, particularly for the techniques that rely on short-term CSI at the transmitter, further studies regarding robustness and the feasibility of exploiting reciprocity in TDD systems are required. Further investigations are also required once a baseline WINNER channel code is available, since coding reduces the impairments due to channel estimation errors significantly.

For the baseline WINNER multi-antenna concept proposal, the baseband complexity involved in spatial processing for signal detection or precoding (when using linear or SIC/THP based approaches) can be considered to be negligible compared to other building blocks, like the decoding algorithm. However, a substantial amount of processing is required for some of the preprocessing algorithms and will significantly outweigh complexity directly related to precoding/detection whenever the channel coherence time and/or the burst length are short. For space-time turbo coded modulation, where initial results have also been presented, complexity should be investigated in more detail. While substantial performance gains can be obtained by using MIMO detector-decoder iterations, it should be kept in mind that using such iterative techniques requires repeated execution of the decoding algorithm, which contributes substantially to the total baseband system complexity.

Since the available channel knowledge at the transmitter strongly depends on the terminal velocities, which will be different for each scenario, and in particular since the spatial channel properties are scenario-dependent, the most promising spatial processing techniques for different scenarios are in general different. Baseline spatial processing modes have been identified for the wide-area scenario, short-range scenario, relay links, and peer-to-peer communication. Apart from the scenario, the spatial mode will depend on details of the channel conditions and system load. For example, for low load and high angular-spread, spatial multiplexing will be an important component, whereas for highly loaded cells and smaller angular-spread beamforming and exploitation of multi-user diversity is capacity-achieving.

For wide-area scenarios only long-term channel state information (CSI) seems to be reasonable for spatial processing in the majority of cases, most favourably combined with short-term channel quality information (CQI) for link adaptation. In all other scenarios, we consider short-term CSI, either due to reduced mobility or even fixed point-to-(multi)point connections. Apart from the fallback mode based solely on spatial diversity, beamforming (including interference suppression and precoding) is considered as integral part in all scenarios. A certain amount of spatial diversity is also beneficial in all scenarios, except for the fixed relay links, where due to the quasi-deterministic fixed link we should use the spatial degrees of freedom preferably for increasing spectral efficiency by means of spatial multiplexing and SDMA. In general, a flexible combination of SDMA and spatial multiplexing is required, since the first is capacity-achieving for highly loaded cells, whereas the latter is optimum if only few users are served concurrently and channel conditions allow it. A promising candidate to implement spatial diversity and multiplexing is using dual-polarised antenna set-ups.

To ensure a versatile and future-proof WINNER multi-antenna concept a high degree of flexibility and generalisation is required. Due to the superior characteristics with respect to scalability and flexibility, we propose to use a *multi-user spatial domain link adaptation concept* as the baseline WINNER multi-antenna concept for further investigation. This concept consists of the following basic components:

- Linear dispersion codes (LDC)
- Directive transmission (beamforming)
- Multi-user precoding

Note that linear precoding is equivalent to multi-user beamforming, and in the single-user case it can also implement typical single-user beamforming techniques, like eigenbeamforming.

For the *multi-user spatial domain link adaptation concept*, a transmitter structure has been developed that includes multiplexing, channel coding, chunk layer multiplexing, and spatial processing on chunk layer basis. The chunk layer multiplexing allows flexible bit stream distribution and therefore to serve users with adaptive transmissions or non-adaptive transmissions that relies on diversity and interleaving concurrently with short-term adaptation of the individual resources. In the spatial chunk layer processing, space-time-frequency coding and modulation is adopted based on power and bit loading, and LDCs. LDCs are considered as a general spatial processing option that allows a flexible tradeoff between spatial diversity and multiplexing and includes several classes of space-time codes, like orthogonal STBC, vector modulation, etc. A definitive decision for a particular LDC will depend on further investigations and input from other tasks. In particular, the optimal split between complexity and overhead in the link adaptation and receiver needs to be investigated: one basic tradeoff is between strong link adaptation combined with simple space-time coding (i.e., low receiver complexity), or reduced overhead due to limited link adaptation combined with high-rate, non-orthogonal space-time codes. A beamforming block allows implementation of directional transmission and linear multi-user precoding. Non-linear precoding will also be further considered for short-range applications and therefore a corresponding block is included in the proposed transmission block diagram.

To operate the proposed *multi-user spatial domain link adaptation concept*, forward- and return-link control information are required. Further investigations will focus on input parameters for *spatial mode selection,* smooth transistions between modes and on details of the spatial processing techniques involved. Additionally the robustness of different spatial processing techniques with respect to imperfections encountered in an operational system (e.g., channel estimation errors, quantisation, and delay) will be investigated in more detail based on the latest status of the WINNER air interface concept. Of major importance is also the development of a joint framework and optimisation of scheduling and radio resource management in frequency, time, and spatial domains in cooperation with other tasks of WP2.

Authors

Partner	Name	Phone/Fax/e-mail
ACL	Thorsten Wild	Phone: +49 711 821 35762 Fax: +49 711 821 32185 e-mail: thorsten.wild@alcatel.de
EAB	David Astély	Phone: +46 8 585 301 49 Fax: +56 8 585 314 80 e-mail: david.astely@ericsson.com
EAB	Magnus Olsson	Phone: +46 8 585 307 74 Fax: +46 8 585 314 80 e-mail: magnus.a.olsson@ericsson.com
EAB	Per Skillermark	Phone: +46 8 585 319 22 Fax: +46 8 757 57 20 e-mail: per.skillermark@ericsson.com
ETHZ	Ateet Kapur	Phone: +41 1 632 2808 Fax: +41 1 632 1209 e-mail: ateet@ee.ethz.ch
IBM	Pedro Coronel	Phone: +41 1 724 8532 Fax: +41 1 724 8955 email: pco@zurich.ibm.com
IBM	Wolfgang Schott	Phone: +41 1 724 8476 Fax: +41 1 724 8955 email: sct@zurich.ibm.com
KTH	Jochen Giese	Phone: +46 8 790 6577 Fax: +46 8 790 7260 e-mail: giese@s3.kth.se
KTH	Tao Jiang	Phone: +46 8 790 8634 Fax: +46 8 790 7260 e-mail: taoj@s3.kth.se
KTH	Patrick Svedman	Phone: +46 8 790 7411 Fax: +46 8 790 7260 e-mail: patricks@s3.kth.se
LUK	Angeliki Alexiou	Phone: +44 1793 776620 Fax: +44 1793 776725 e-mail: alexiou@lucent.com
LUK	Abdelkader Medles	Phone: +44 1793 776783 Fax: +44 1793 776725 e-mail: medles@lucent.com
MOT	Stéphanie Rouquette-Léveil	Phone: +33 (0)1 69 35 48 24 Fax: +33 (0)1 69 35 77 01 e-mail: stephanie.rouquette@motorola.com
NOK	Samuli Visuri	Phone: +358 50 486 8219 Fax: +358 7180 36857 e-mail: samuli.visuri@nokia.com

PRL	Tim Moulsley	Phone: +44 1293 815717 Fax: +44 1293 815024 e-mail: tim.moulsley@philips.com
PRL	Keith Roberts	Phone: +44 1293 815754 Fax: +44 1293 815024 e-mail: keith.roberts@philips.com
SM	Karsten Brüninghaus	Phone: +49 2871 91 1742 Fax: +49 2871 91 3387 e-mail: karsten.brueninghaus@siemens.com
SM	Martin Döttling	Phone: +49 89 722 58703 Fax: +49 89 722 913433 e-mail: martin.doettling@siemens.com
SM (TUB)	Eduard Jorswieck	Phone: +49 30 31002 868 Fax: +49 30 31002 863 e-mail: eduard.jorswieck@hhi.fhg.de
SM (TUB)	Aydin Sezgin	Phone: +49 30 31002 868 Fax: +49 30 31002 863 e-mail: aydin.sezgin@hhi.fhg.de
SM (TUB)	Dafei Wang	Phone: +49 30 31002 852 Fax: +49 30 31002 863 e-mail: Dafei.Wang@mk.tu-berlin.de
TUD	Ernesto Zimmermann	Phone: +49 351 463 34958 Fax: +49 351 463 37255 e-mail: zimmere@ifn.et.tu-dresden.de
TUI	Martin Fuchs	Phone: +49 3677 69 1156 Fax: +49 3677 69 1195 e-mail: martin.fuchs@tu-ilmenau.de
TUI	Martin Haardt	Phone: +49 3677 69 2613 Fax: +49 3677 69 1195 e-mail: martin.haardt@tu-ilmenau.de
TUI	Marko Hennhöfer	Phone: +49 3677 69 2615 Fax: +49 3677 69 1195 e-mail: marko.hennhoefer@tu-ilmenau.de
TUI	Veljko Stankovic	Phone: +49 3677 69 1156 Fax: +49 3677 69 1195 e-mail: veljko.stankovic@tu-ilmenau.de
UOulu	Zexian Li	Phone: +358-407-089-219 Fax: +358 8 553 2845 e-mail: zexian.li@ee.oulu.fi

Table of Contents

List of Acronyms and Abbreviations

3GPP	3rd Generation Partnership Project
AF	Amplify-and-Forward
AM	Alamouti
AMC	Adaptive Modulation and Coding
AoA	Angle of Arrival
AoD	Angle of Departure
AP	Access Point
AWGN	Additive White Gaussian Noise
BD	Block Diagonalisation
BF	Beamforming
BICM	Bit-Interleaved Coded Modulation
BLAST	Bell Labs Layered Space Time
BLEP	Block Error Probability
BLER	Block Error Rate
BS	Base Station
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CESM	Capacity Effective SINR Mapping
CPM	Compression and Mapping to PER/BLER
CQI	Channel Quality Indicator
CSI	Channel State Information
D-BLAST	Diagonal - Bell Labs Layered Space Time
DET	Dominant Eigenmode Transmission
DF	Decode-and-Forward
DS-CDMA	Direct Sequence – Code Division Multiple Access
DT	Diversity Technique
EBF	Eigenbeamforming
EESM	Exponential Effective SINR Mapping
FDD	Frequency Division Duplex
FEC	Forward Error Coding
FER	Frame Error Rate
FFT	Fast Fourier Transformation
FSK	Frequency Shift Keying
ftp	file transfer protocol
GMC	Generalised Multicarrier
GoB	Grid of Beams
HARQ	Hybrid Automatic Repeat Request
HPA	High-Power Amplifier
HPBW	Half-Power Beamwidth
ISI	Inter-Symbol Interference
L2S	Link-to-System
LDC	Linear Dispersion Code
LESM	Logarithmic Effective SINR Mapping
LNA	Low-Noise Amplifier
LOS	Line-Of-Sight
LR	Lattice Reduction
LT	Long-Term
MAC	Medium Access Control
MCS	Modulation and Coding Scheme
MDS	Multiple Data Streams
MGF	Moment Generating Function
MIESM	Mutual Information Effective SINR Mapping
MIMO	Multiple Input Multiple Output

MISO	Multiple Input Single Output
MMSE	Minimum Mean -Square-Error
MS	Mobile Station
MSE	Mean Square Error
MU	Multi-User
MUI	Multi-User Interference
MUX	Multiplexing
NL	Non-linear
NLP	Non-linear Precoding
OBF	Opportunistic Beamforming
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PARC	Per Antenna Rate Control
PDF	Probability Density Function
PDP	Power Delay Profile
PEP	Pairwise Error Probability
PER	Packet Error Rate
PHY	Physical Layer
PLL	Phase-Locked Loop
PPM	Pulse Position Modulation
PSK	Phase Shift Keying
PSRC	Per Stream Rate Control
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RF	Radio Frequency
RRM	Radio Resource Management
RS	Relay Station
Rx	Receiver
SCCH	Scheduled Control Channel
SCM	Spatial Channel Model
SDCH	Scheduled Data Channel
SDMA	Space Division Multiple Access
SFC	Space-Frequency Code
SIC	Successive Interference Cancellation
SIMO	Single Input Multiple Output
SINR	Signal-to-Interference-plus-Noise Ratio
SISO	Single Input Single Output
SMMSE	Successive Minimum Mean Square Error
SMUX	Spatial Multiplexing
SO	Successive Optimisation
ST	Space Time
STBC	Space Time Block Codes
STCM	Space Time Coding and Modulation
STFCM	Space-Time-Frequency Coding and Modulation
STTC	Space Time Trellis Code
STTuC	Space Time Turbo Code
SVD	Singular Value Decomposition
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
THP	Tomlinson-Harashima Precoding
TPC	Transmit Power Control
Tx	Transmitter
UCA	Uniform Circular Array
ULA	Uniform Linear Array
V-BLAST	Vertical - Bell Labs Layered Space Time
VASS	Virtual Antenna Stream Summation
WCDMA	Wideband Code Division Multiple Access
WINNER	Wireless World Initiative New Radio

WP	Work Package
WF	Water Filling
ZMCSCG	Zero Mean Circularly Symmetric Complex Gaussian

List of Mathematical Expressions

k	user index
$a_{c,q}$	symbol stream of layer q of chunk c after STFCM
$\mathbf{B}$	feedback matrix
$b_{k,s}$	substream s of user k after segmentation
c	chunk index, one chunk spans n_{frame} OFDM symbols, n_{sub} subcarrier, and $Q(c)$ spatial layers
$\mathbf{D}$	demodulation matrix
$d_{c,q}$	bitstream of layer q of chunk c
f	beamforming weight
$\mathbf{F}$	precoding matrix
FEC_k	forward error coding of subflow k
f_k	bit stream of user k before FEC
K	number of scheduled users
$M(c,q)$	modulation of bitstream $d_{c,q}$
M_r	number of receive antennas
M_t	number of transmit antennas
M_{va}	number of virtual antenna streams
N_c	total number of subcarriers
n_{frame}	OFDM symbols per chunk
n_{sub}	number of subcarriers per chunk
$P(c,q)$	power assignment of bitstream $d_{c,q}$
$Q(c)$	maximum number of spatial layers in chunk c
$s_c(n_a, n_t, n_f)$	virtual antenna signal of virtual antenna n_a at time n_t and subcarrier n_f
$\mathbf{V}$	beamforming matrix
v_{site}	average normalised throughput in bps/Hz/site
$v_{site,\ 98\%}$	area spectral efficiency
$\mathbf{H}$	channel matrix
Π_k	interleaving of subflow k
$\gamma(k,s)$	chunk mapping function
$\sigma(k,s)$	layer mapping function
$\|\mathbf{X}\|$	Frobenius norm of $\mathbf{X}$

1. Introduction

Advanced multi-antenna solutions are considered an integral part of the overall WINNER system. Major requirements for such a *WINNER multi-antenna concept* include high performance, robustness, adaptability to a wide range of scenarios and terminal classes, as well as efficient interworking with the basic transmission and network concept defined in other tasks and work packages.

In order to identify such a versatile WINNER multi-antenna concept, different spatial processing techniques need to be investigated and compared with respect to their performance in different usage scenarios, and their compatibility to form a component of the overall WINNER multi-antenna concept. Based on this analysis a first proposal for the overall WINNER multi-antenna concept will be elaborated.

Due to the manifold world of spatial processing techniques with different underlying ideas, objectives, and use cases, it is not a trivial task to *compare and assess* the techniques with respect to performance, complexity, scalability, application scenarios, or required signalling and network functionality. The fact that no common simulation environment is available adds further difficulty. In order to tackle this challenging task the following prerequisites have been identified:

- A simulation methodology suited for multi-antenna system-level simulations must be developed
- The corresponding simulation assumptions and parameters need to be clearly identified and kept identical as far as possible to ensure comparability
- Assessment criteria need to be defined and adapted to the simulation assumptions
- The system complexity, required support functions (like pilots, channel estimation, protocols) and control overhead (feedback specific to the spatial processing) need to be compiled and integrated in the assessment process

Since the *available channel knowledge at the transmitter* strongly depends on the terminal velocities, which will be different in different scenarios, and in particular since the spatial channel properties are scenario dependent, the most promising spatial processing techniques for *wide-area coverage* and *short-range scenarios* are in general different. Therefore, these two scenarios are often distinguished within this document. However, ideally, a *single overall concept* is able to adapt to both of these scenarios (e.g., via an adaptive use of channel knowledge) and may therefore be regarded as *the* preferred concept for a wide range of scenarios.

Some of these multi-antenna techniques have been developed with primarily link-level improvements in mind (e.g., spatial diversity techniques, like space-time coding). Others focus on improved system capacity (e.g., improved cell throughput via SDMA or multi-user precoding) and cannot be analysed at link level. Furthermore, multi-user interference has significant impact on the overall performance. Hence, *link-level results* will be used solely for explanation, information and as input to the corresponding system-level simulations. The assessment and comparison, however, is based on system-level performance measures like throughput and spectral efficiency. Due to the early deadline of this document, simplifying assumptions needed to be adapted for such system-level simulations.

A major novel aspect of this document is that, in contrast to the vast number of theoretical literature on spatial processing, a strong emphasis is put on practical questions related to the implementation of such techniques in a radio access network. Therefore questions concerning system complexity, measurements, feedback, control overhead, robustness against impairments, and flexibility/scalability (with respect to the terminal capability, number of transmit and receive antennas, and propagation conditions) are addressed and reflected in the assessment process and criteria.

Chapter 2 provides an introduction, overview, review and classification of multi-antenna techniques. Important terms are defined and a unified view of spatial processing techniques is given along with a summary of the main features of the different multi-antenna techniques investigated. Based on this review, a candidate set of promising candidates for the wide-area and short-range scenario are identified and a generic transmitter block diagram and a unified mathematical model are developed. A detailed description of the investigated multi-antenna techniques is presented in Chapter 3. Apart from the description of the link and multi-user signal design, the required supporting functions, control channels, and protocols are explained.

To allow thorough investigations of a MIMO-OFDM system using fast radio resource management and link adaptation, considerable effort has been spent on developing the required simulation methodology, and in particular a suitable link-to-system-level interface for multi-antenna techniques in the WINNER context. The corresponding information and the selected link-to-system-level interface are presented in Chapter 4. Due to the absence of a common WINNER simulation tool, comparability of results is of major concern. The simulation assumptions and parameters used for this document are given in Chapter 5.

Due to the early deadline, the system-level simulations in this document focus on snapshot single-cell multi-user simulations primarily for the downlink. A full buffer traffic model without dedicated modelling of packets is adopted. As a consequence the assessment criteria of WP7 are not directly applicable. The adaptation of those assessment criteria to this kind of simulations is the content of Chapter 6. The basic idea has been to maintain as much of the assessment criteria of WP7 as possible and if alterations have been unavoidable to minimise them and go along the spirit of the original criteria.

Chapter 7 provides the simulation results and assessment for the wide-area scenario, the short-range scenario, and for MIMO relaying. Note that the simulation results in Chapter 7 are based on the WINNER simulation methodology and the agreed simulation assumptions and parameters. Further specific results, e.g., for precoding techniques, can be found in Chapter 3 along with the description of the corresponding techniques. System and terminal complexity is addressed in Chapter 8, focusing on implications of spatial processing on pilot design, on different ways to obtain short-term channel knowledge at the transmitter, and on baseband and RF complexity of the spatial processing algorithms.

Based on this analysis, initial proposals for the overall WINNER multi-antenna concept are derived in Chapter 9. After an outlook on the further evolution of proposed investigation in WINNER Phase I in Chapter 10, conclusions are drawn in Chapter 11.

2. Overview on Spatial Processing

This chapter provides basic terms and definitions, as well as an overview of gains that can be obtained by spatial processing. A classification of multi-antenna techniques is given and a generic transmitter architecture is introduced that serves as a framework for the further description of the multi-antenna techniques in the following chapters.

2.1 Multi-antenna Methods, Techniques and the Concept

For clarity, the following text distinguishes multi-antenna *methods*, multi-antenna *techniques*, and the WINNER multi-antenna *concept.* A multi-antenna method is the pure algorithm to foster specific benefits of spatial processing (e.g., gains of beamforming, spatial multiplexing, spatial diversity, etc.). A multi-antenna technique is a multi-antenna method integrated in a radio access network. This means that required measurements, signalling, procedures, protocols, and supporting radio interface functions must be provided in addition to the pure spatial processing algorithm. We will refer to the method together with its "enabling requirements" as a multi-antenna technique. This is illustrated in Figure 2.1. While the majority of publications in open literature only investigate spatial processing algorithms, few investigations are available on implementation aspects. Therefore, the major focus of this document and the multi-antenna task in WINNER, is the evolution of pure *methods* into *techniques* suitable for implementation in a radio access network.

Different multi-antenna methods exploit different properties of the radio channel and have different *performance objectives* (e.g., optimisation of cell throughput, peak data rate, delay, etc.). The ubiquitous WINNER radio access concept and therefore also the WINNER multi-antenna concept shall be able to operate under a wide range of radio access conditions, characterised by different deployments (e.g., cell radius, access point antenna configuration, etc.), and usage scenarios (with respect to number of users, terminal capabilities, user velocity, service type). Since the properties of the channels are expected to be different, the WINNER multi-antenna concept will consist of a set of multi-antenna techniques as illustrated in Figure 2.1. Therefore, one important task of this deliverable is to identify a set of multi-antenna *methods* that provides flexibility and high performance in many different conditions, while at the same time minimizing the overall required enablers, complexity and overhead. Flexibility and the option of future evolution is also an important objective when defining the WINNER multi-antenna concept, since research and development on methods is expected to continue even after a possible system deployment, as has been the case for second and third generation cellular systems.

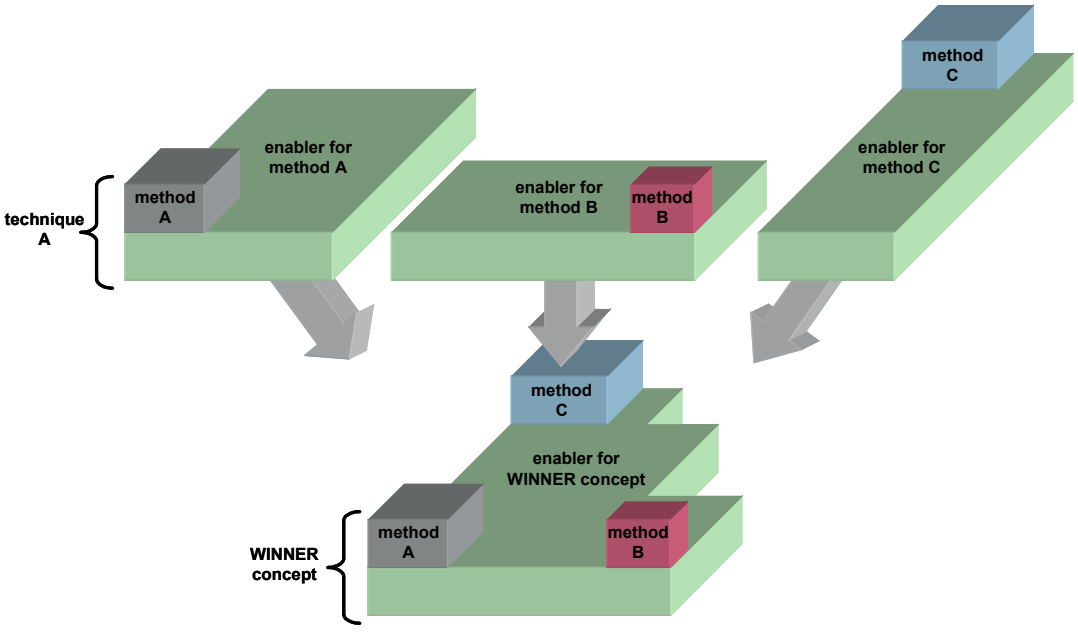


Figure 2.1: Basic terms, definitions, and components of the WINNER multi-antenna concept

The final WINNER multi-antenna concept should provide the means to ensure:

- Observability
  Adequate feedback of measurements from the receivers in addition to measurements done on a return link to facilitate scheduling, link adaptation, and selection of transmit parameters controlling the spatial processing. This channel knowledge makes it possible for the transmitter to assess the suitability of different resource allocations in terms of users, streams, and spatial processing techniques in the multi-user setting. Examples from WCDMA include feedback bits for closed-loop transmit diversity and channel quality measurements for fast scheduling and link adaptation.
- Controllability
  Adequate signalling, procedures, and architecture to ensure that the transmitter can adapt the transmission and resource allocation according to the current conditions. Thus, multiplexing of dedicated and common channels, the control of phase references for demodulation, the procedures for mobility management, and the mechanisms for (re)configuring the transmission format of the radio links, should not prevent the use of efficient spatial processing.

Observability ensures that the right information is available at the right place and on the right time in the network, whereas controllability makes it possible to use this information.

In short, the WINNER multi-antenna concept should make it possible to acquire channel knowledge, use it in a sensible and robust way and be able to adapt the transmit technique, in particular the spatial processing component, to the current conditions.

2.2 Principles of Spatial Processing

As stated above, different multi-antenna methods exploit different properties of the radio channel in order to accomplish performance improvements. Typically, this exploitation is realised by the spatial processing component of the multi-antenna method that leverages, more or less, one or more of the following basic gains:

- Array gain
- Diversity gain
- Spatial multiplexing gain
- Interference suppression gain.

It should be understood that there is in general a trade-off between these types of gains. In addition, specific processing at the transmitter and/or receiver is needed in order to leverage them, and the exact link gains depend critically on the properties of the radio channel and the amount of channel knowledge available at the receiver and transmitter.

Channel knowledge is typically described with two sorts of measures; channel state information (CSI) and channel quality indicators (CQIs). The term CSI usually refers to knowledge of the (complex valued) radio channel, not only for the link of interest but also for other interfering links, such as a quantised version of the instantaneous realisation, its second order statistics, or main directions. A CQI, on the other hand, is rather a (real valued) measure of the quality of the channel, for example an SINR after receiver processing that may be used to adapt the code rate, modulation order, and spreading at the transmitter. In terms of multi-antenna techniques, it can be envisioned that CSI typically is used to control the directivity of the transmission, i.e., in which directions to spread energy, whereas CQI could be used for scheduling and link adaptation, including the control of spreading in space, time, and frequency. Obviously, CSI and CQI are related, and CQI may be a viewed as a special, quantised form of CSI.

The amount of channel knowledge dictates which methods are applicable and the potential benefits of spatial processing techniques. Further details will be given in the detailed descriptions of the techniques.

The array or beamforming gain is the average increase in SINR at the receiver that arises from the coherent combining effect of multiple antennas at the receiver or transmitter, or both. To extract receive (or transmit) array gain the receiver (or transmitter) needs to have accurate CSI.

Diversity mitigates the fading in the radio channel so that the variability of received signal power is

reduced. The basic idea of diversity is to provide the receiver with multiple replicas of the data. The degree to which these replicas are independently faded reflects the diversity gain. In the following, the focus is on diversity gains that can be achieved with spatial processing (multiple antennas with possibly different polarisation [SAT03]). However, diversity can also be provided in other forms such as multipath diversity (frequency selectivity) and time diversity (time selectivity).

Receive antenna diversity can be extracted in a SIMO system by combining signals from several receive antennas so that the resulting signal exhibits considerably reduced amplitude variability compared to a single-receive antenna system. Similarly, transmit diversity can be applied to MISO systems independently of whether the transmitter has CSI or not. For the latter case, space-time (ST) coding is a suitable design that codes the data across the transmit antennas. More generally, diversity in a MIMO system implies a combination of receive and transmit diversity. In frequency flat channels, the maximum achievable diversity gain is the product of the number of transmit antennas and the number of receive antennas. This maximum diversity gain can be achieved when the correlation matrix of the vectorised channel has full rank, for example, when the channel coefficients between each transmit-receive antenna pair fade independently. The drawback of transmit diversity is that only a limited amount of the channel capacity is exploited.

ST(F) coding relies on spreading transmitted symbols in space, time, and/or frequency. This dispersion can be done linearly using an orthogonal block structure, which has the benefit of considerably simplifying the decoding. A more general framework is matrix modulation where the symbols are dispersed by means of matrices that can be optimised with respect to various criteria. Matrix modulation allows for significant flexibility in the design of dispersion matrices; linear vs. nonlinear mapping, orthogonal vs. quasi- or non-orthogonal designs, etc. This flexibility is the key to avoid structural limitations of simpler ST codes and to achieve high rates without sacrificing diversity. The price to pay, however, is often an increase in receiver complexity. Finally, we note that dispersion can also rely on trellis codes and corresponding sequence detection, but this approach quickly becomes prohibitively complex as the system grows (number of antennas, constraint length, etc.).

Spatial multiplexing increases the spectral efficiency of a multi-antenna system. The increase is linear in the number of transmit-receive antenna pairs, and it requires no additional power expenditure. Specifically, in spatial multiplexing the multi-antenna channel is viewed as parallel spatial “pipes” through which independent data streams can be transmitted. Under favourable channel conditions, the spatial signatures induced by these streams are well separated. In this case, the receiver, which has CSI, can differentiate and extract each stream. Spatial multiplexing also applies to a multi-user context, also known as SDMA, in which case the streams possibly originate from or are transmitted to different users. The drawback of spatial multiplexing is that the diversity gain, and hence the reliability is reduced; individual streams are not exposed to multiple, independent fading channels.

There are several encoder and receiver structures used for spatial multiplexing. We focus here on the single-user case. In vertical encoding (also known as V-BLAST), the input data stream is demultiplexed into as many streams as transmit antennas. Each resulting substream undergoes independent temporal encoding, interleaving and symbol mapping, and only then is it fed to a single transmit antenna. This structure simplifies receiver design, but is suboptimal in terms of diversity order (see the previous comment about reliability). The receiver can decode streams independently (using ZF or MMSE) or successively (in conjunction with ZF and MMSE), and hence there is no need for joint decoding. In diagonal encoding (also known as D-BLAST), the input data stream undergoes horizontal encoding after which each codeword is split into frames/slots. These frames pass through a stream rotator that rotates the frames in a round-robin fashion so that the mapping from bit streams to antennas is periodically altered. Such a structure can potentially achieve full diversity.

Spatial interference suppression (sometimes referred to as interference avoidance) is achieved by setting the antenna weights based on some sort of knowledge about the interference situation. In this way, it is possible to shape the effective antenna beam, and transmit and/or receive in different directions. The result is that the interference spread to/from users in other directions is reduced, i.e., a spatial interference suppression, or interference avoidance, is realised.

2.3 Summary of Multi-Antenna Methods

This section summarises the basic ideas behind well-known and promising multi-antenna methods. In particular, we detail how these methods leverage the spatial dimension to achieve the MIMO gains described in Section 2.2. The methods that we simulate in Chapter 7 are mentioned here, and the corresponding techniques are described in detail in Chapter 3. Our goal in this section is not to go into details, but to give a general overview of multi-antenna methods and reference the relevant literature. Detailed reviews of state-of-the-art MIMO techniques can also be found in [FLOD14], [METD3.2].

2.3.1 Beamforming Techniques

Beamforming techniques use multiple antennas to focus beams in certain spatial directions to leverage array and interference rejection gains. In the single-link case, the transmitter and/or receiver adapts its antenna weights according to the amount of channel knowledge and can thus achieve array gain. In the multi-user case, beamforming enables spatial selectivity by allocation different antenna weights to different users, thereby allowing for SDMA and achieving interference rejection and multi-user diversity gains. These techniques are described in detail in Section 3.2.1 to 3.2.4 and investigated in Chapter 7.

2.3.2 Diversity Techniques

As discussed in Section 2.2, diversity techniques aim at increasing link reliability, but this increase often comes at the expense of transmission rate.

2.3.2.1 Receive Diversity

A receiver equipped with multiple antennas receives multiple copies of the transmitted signal (or signals) that it can combine to increase reliability. This combination forms the basis of *receive* diversity and it is achieved by summing up the weighting output of each receive antenna. The optimum weights are chosen to maximise the output SINR, and they depend on the *instantaneous* fading state when the receiver knows this state, or on the receiver correlation matrix when it only has statistical channel information. In particular, *maximum ratio combining* (MRC) is optimal for spatially white noise plus interference, and MMSE combining for the spatially coloured case. See [LR99] and [PP97] for more details on receive diversity.

2.3.2.2 Transmit Diversity

Conversely, a transmitter equipped with multiple antennas can achieve *transmit* diversity by spreading the transmitted symbols over time and space (ST coding). The classical theory of space-time coding as developed in [TSC98] specifies the so-called rank and determinant criteria to construct ST codes: the amount of transmit diversity is equal to the smallest rank of codeword-difference matrices, and the coding gain is related to the product of the non-zero singular values of these matrices. The specific design and implementation of the spreading depends on the level of channel knowledge at the transmitter and receiver.

2.3.2.2.1 Open-Loop Techniques

Space-Time Trellis Codes (STTC) were originally proposed by Tarokh *et al.* in [TSC98] based explicitly on the above two criteria. In fact, this construction combines trellis coding and modulation design (including the mapping of the encoded data) with a redundancy expansion in both the signal and antenna space. In this sense, STTCs can be viewed as a multi-antenna extension of classical trellis coded modulation (TCM) [BMM94]. The basic approach in [TSC98], however, has several drawbacks. The decoding complexity quickly gets prohibitive with increasing constraint length, there is no systematic procedure for reducing the cardinality of the searched class of codes, and the codes don't demonstrate robustness in practical applications characterised by extremely variable Doppler frequencies. Alternative approaches are presented hereafter that combine the coding gain benefits of turbo coding with the diversity advantage of ST coding techniques and/or the bandwidth efficiency of coded modulation.

STTCs can be logically extended in the same way as TCM is generalised in single-antenna systems to turbo coded modulation and multiple-TCM. The first technique leads to parallel concatenated space-time turbo coded modulation (STTuCM) introduced in [Tuj00]: it integrates code concatenation into a random-like ST coding approach by means of an information interleaver that provides both large equivalent constraint-lengths and randomness. This new ST coding framework applies to any STTC from the litera-

ture and achieves improved spectral efficiency. In particular, punctured recursive constituent codes (CCs) were designed in [Tuj00] based on the handcrafted STTCs from [TSC98], and later extended to higher spectral efficiencies in [TJL02]. Similar schemes were studied in [Nar99], [FCV+01] and [CH00], and other ST Turbo code proposals [SD99] [SG01] [LF99] differ by the single antenna turbo coded modulations (for instance, [SD99] can be seen as a multi-antenna extension of the bit-interleaved turbo coded modulation). The second generalisation, namely multiple trellis coded modulation with space-time constellations, combines matrix modulation schemes with trellis-coding by generalizing Divsalar's multiple-TCM to multidimensional space-time. Examples of such codes are given in [SWX] [Ion03] [IMY+01] [JS03].

Space-Time Block Codes (STBCs), on the other hand, spread the symbols in time and space in a block-by-block fashion. Among the class of STBCs, the so-called orthogonal STBCs, which are designed based on an orthogonal structure, are especially attractive because they dramatically reduce decoding complexity: via linear processing, joint maximum-likelihood decoding of all the transmitted symbols decouples into symbol-by-symbol decoding. The first orthogonal STBC, proposed by Alamouti for 2 transmit antennas [Ala98] , spreads the data (symbols s_1 and s_2) in time and space as

$$\mathbf{X}_{\text{Alamouti}} = \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix},$$

where rows and columns represent transmit antennas and symbol intervals, respectively. This simple code yields full diversity $2M_R$ and code rate equal 1, so that rate is preserved compared to the single antenna case. However, from a capacity point of view, this scheme is optimal only when there is a single receive-antenna, since in this case there is only one effective channel [Sam03]. The basic construction of Alamouti generalises in the case of real-valued data symbols to arbitrary number of transmit antennas, and still retains full diversity and full rate equal 1 [TJC99]. In the case of complex-valued data symbols, however, there do not exist full-rate codes for more than two transmit antennas. Allowing for non-square ST matrices (with more columns than rows), we can construct from the Alamouti code some full-diversity orthogonal STBCs with rate 1/2 for any number of transmit antennas [TJC99]. For instance, a four-element array can transmit 4 symbols over 8 symbol intervals according to:

$$\mathbf{X} = \begin{bmatrix} s_1 & -s_2 & -s_3 & -s_4 & s_1^* & -s_2^* & -s_3^* & -s_4^* \\ s_2 & s_1 & s_4 & -s_3 & s_2^* & s_1^* & s_4^* & -s_3^* \\ s_3 & -s_4 & s_1 & s_2 & s_3^* & -s_4^* & s_1^* & s_2^* \\ s_4 & s_3 & -s_2 & s_1 & s_4^* & s_3^* & -s_2^* & s_1^* \end{bmatrix}.$$

Some better rates (inferior to 1) can even be achieved from generalised linear processing orthogonal designs. For 3 or 4 transmit antennas, several codes [TJC99] [GS00] [TH00] [HMP99] have been presented with an overall coding rate of 3/4 that corresponds to the maximum achievable data rate [TH00].

To achieve higher rates, one can sacrifice full-diversity, linearity or orthogonality. Sacrificing full-diversity gain was considered in [JRK+99] where transmit-diversity 2 and rate 1 orthogonal STBCs were proposed for three and four antennas by applying a Walsh-Hadamard transformation to Alamouti codes. In the case of four transmit antennas, the corresponding code matrix is:

$$\mathbf{X}_{\text{Jalloul}} = \begin{bmatrix} s_1 & -s_2^* & s_1 & -s_2^* \\ s_2 & s_1^* & s_2 & s_1^* \\ s_3 & -s_4^* & -s_3 & s_4^* \\ s_4 & s_3^* & -s_4 & -s_3^* \end{bmatrix}.$$

Sacrificing linearity was first considered in [HM00], in the context of unitary space-time modulations. The basic idea is to select a set of constellation points that lie in the hyper-surface of a unitary group manifold. Another approach is to use linearly dependent modulation matrices [GMF00], but in contrast to the linear case, these modulation matrices do not have a simple interpretation in terms of symbols. Another construction proposed in [SPP02] concatenates a linear, unitary STBC with cyclic redundancy check. These non-linear constructions suffer from suboptimal Euclidean distance properties and increased

detection complexity.

Finally, sacrificing orthogonality, which of course creates self-interference, has been studied in, e.g., [HE00] [TBH00] [Jaf01][HH01] [[PF01]] [HBP01] [TH01] under a variety of names (non-orthogonal or layered space-time block codes [TBH00] [TH01], quasi-orthogonal STBCs [Jaf01], linear dispersion codes (LDC) [HH01]). Different heuristic design rules for advanced matrix modulation schemes can be derived based on Frobenius orthogonality (traceless self-interference), minimal self-interference, symbol homogeneity, maximal symbol-wise diversity, or maximal mutual information (see [HTW03] for details). These rules can often be complemented by other rules such as the well-known rank [GFB+96] and determinant [TNS+99] criteria, minimizing the union bound of bit-error probabilities [HBP01], or constellation rotation [TH01] (also called precoding). The detection complexity (due to the non-orthogonal structure) can be limited by ensuring that parts of the symbols are mutually orthogonal.

Specific examples of high-rate, non-orthogonal, linear matrix modulations are Twisted Space-Time Transmit Diversity (TSTTD), which for a two-transmit antenna system achieves the same diversity as the Alamouti code, but with twice the rate (2 instead of 1) and a higher coding gain. It is constructed by combining two transposed Alamouti codes combined with a Clifford-basis twisting operation:

$$\mathbf{X}_{TSTTD} = \begin{bmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{bmatrix} + \begin{bmatrix} s_3 & s_4 \\ s_4^* & -s_3^* \end{bmatrix} \mathbf{U}, \qquad \text{where} \quad \mathbf{U} = \frac{1}{\sqrt{7}} \begin{bmatrix} 1+j & -1+2j \\ 1+2j & 1-j \end{bmatrix}.$$

The ABBA code proposed in [TBH00] for 4 transmit antennas (equivalent constructions are given in [Jaf01] and [PF01]) achieves rate 1 and full diversity by permuting two Alamouti codes:

$$\mathbf{X}_{\text{ABBA}} = \begin{bmatrix} s_1 & -s_2^* & s_3 & -s_4^* \\ s_2 & s_1^* & s_4 & s_3^* \\ s_3 & -s_4^* & s_1 & -s_2^* \\ s_4 & s_3^* & s_2 & s_1^* \end{bmatrix}.$$

Higher rate codes with full diversity can be constructed by simply layering the ABBA code in combination with precoding (constellation rotation), but the self-interference and hence the detection complexity increase with the number of layers.

2.3.2.2.2 Closed-Loop Techniques

When both the transmitter and the receiver have (full or partial) instantaneous channel knowledge, the optimum technique is *dominant eigenmode* transmission. This technique converts the MIMO channel into a single pipe with maximum power, and can be viewed as short-term beamforming. It achieves full diversity and higher array gain than open-loop techniques.

2.3.2.2.3 Partial-CSI Techniques

High channel correlation can significantly degrade performance [SFG+00] [BP00]. Linear precoding techniques [SP02] assume knowledge of the (slowly varying) antenna correlations at the transmitter, and can thus force transmission on the nonzero eigenmodes of the transmit antenna correlation matrix [BHN00] . Such a technique is illustrated in Figure 2.2.

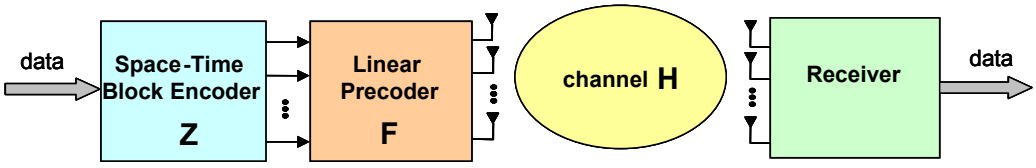


Figure 2.2: Reconfigurable transmission scheme combining space-time block codes and a linear transformation designed w.r.t. channel knowledge available at the transmitter

The linear transformation **F** is determined so as to optimise a given criterion, such as minimise an upper bound on the pairwise error probability as in [JSO02] or [SP02], which focuses specifically on correlation-based CSI, or maximise the expected SNR at the receiver. The optimal design in the latter case is to transmit in the direction of the strongest correlation matrix eigenvector, i.e., an invariant beamformer pointing to one direction. We discuss this method in detail in Section 3.1.3.

Further performance improvements can be leveraged if, in addition to correlation-based CSI, instantaneous and quantised CSI is fed back from the receiver as illustrated in Figure 2.3. By building upon techniques developed in, e.g., [JSO02] [JB02] [JBS04], a joint method for designing the quantiser and the precoding matrices so as to minimise a certain upper bound on the symbol error probability can be derived for orthogonal and quasi-orthogonal STBCs. The computation of the optimal precoder is not tractable in real-time anymore, but it can be easily done off-line since the CSI is quantised.

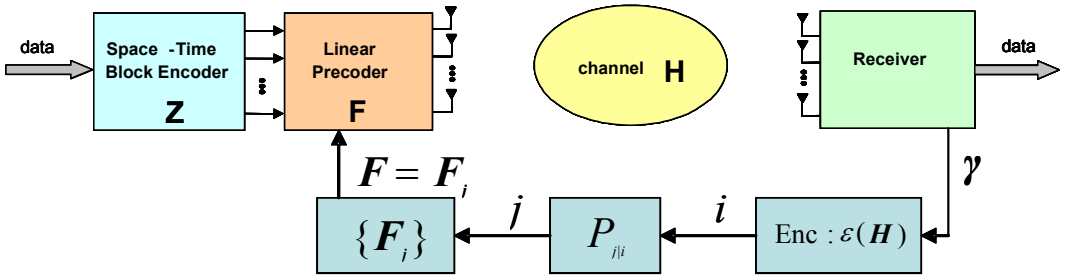


Figure 2.3: Weighted space-time coding where quantised CSI from a feedback link is used to choose a suitable precoder matrix.

2.3.2.2.4 Noncoherent MIMO Transmission

In certain scenarios (e.g., high mobility) channel parameters can often not be acquired even at the receiver. This is especially the case in frequency-selective channels, where the number of channel parameters to estimate can be very large. For such cases, space-frequency code (SFC) designs for the noncoherent MIMO-OFDM channel can be based on a fully noncoherent approach [BB04] or a differential coding one [GS03] [GS04]. The case for noncoherent SFCs specifically tailored to frequency-selective channels is demonstrated by the existence of codes designed for noncoherent frequency-flat channels whose performance breaks down in a frequency-selective environment (this is in stark contrast to the coherent case).

For the fully noncoherent approach, ML detection is applicable when the receiver knows the power delay profile (PDP), and it yields a code design criterion based on the PEP (if the PDP is unknown, GLRT-based detection can be used) for unitary constellations [BB04]. The maximum achievable diversity order is $M_T M_R L$ as in the coherent case, and it can be achieved by a code construction inspired from the minimum-correlation criterion proposed for the frequency-flat fading case in [HMR+00] suitably adapted to the frequency-selective case [BB04]. A generalised construction is also proposed, but its analytical performance does not appear tractable. Finally, these new design criteria can be used to draw fundamental conclusions on the impact of channel delay spread and PDP on SFC performance.

The drawback of the fully noncoherent approach is an increased design and detection complexity. Differential modulation can be used to avoid this complexity increase at the expense of performance by adapting the flat-fading channels differential space-time modulation in [HS00] and [Hug00] to the frequency-selective case: differential coding is then done over subcarriers rather than time [GS03] [GS04] (since detection is required within a single OFDM frame). Thus, it exploits channel correlation across adjacent subcarriers while realizing space and frequency diversity gain. In addition, several receivers apply that have increasing performance but also complexity: symbol-wise differential detection (as in [HS00]), symbol-wise detection with decision feedback (as in [LS01] [SL02]), and joint detection over d symbols with decision feedback.

2.3.3 Spatial Multiplexing

The simultaneous use of multiple transmit and receive antennas promises, at least theoretically, a large increase in data rates and hence spectral efficiency. This increase is achieved by Spatial Division Multiplexing (SMUX) techniques that transmit simultaneously independent data streams from different antennas. This architecture, illustrated for 4 transmit and M_R receive antennas in Figure 2.4, effectively creates *spatial data pipes* between transmitter and receiver.

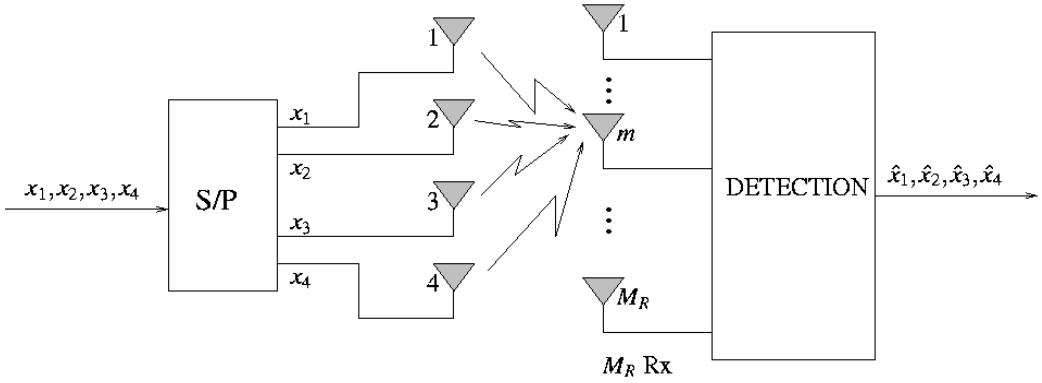


Figure 2.4: Spatial Division Multiplexing Scheme

The input data streams are spatially multiplexed and, under favourable channel conditions, the spatial signatures of these transmitted streams are well separated. The receiver, having knowledge of the channel, can differentiate between these co-channel streams and extract each. The condition for spatial signature separation is that the channel matrix should have full rank, in which case the increase in transmission rate is linear in the number of transmit-antenna pairs $Q=\min(M_T,M_R)$. A rank-deficient channel matrix, such as would arise in limited scattering environment, reduces the effective number of data pipes Q.

At the receiver, several approaches can be used to perform detection with different levels of complexity and thus error performance. We discuss them in some detail in Section 2.3.5.

2.3.3.1 Open-Loop SMUX

When the transmitter does not have channel state information, it distributes rate and power equally among the transmit antennas. This architecture is sufficient to provide full spatial multiplexing gain.

2.3.3.2 Closed-Loop SMUX

With full knowledge of the channel (and of the noise plus interference covariance matrix), the transmitter may apply pre-coding to effectively create Q orthogonal sub-channels. In addition, it can apply different coding and modulation schemes on each data-stream [GC98]. This 'bit-loading' is done to reflect the differing capacities of each sub-channel. Similarly, the available transmit power may be distributed unequally amongst the sub-channels in order to improve throughput. Typically this is done according to the sub-channel singular values by using a 'water-filling' (WF) algorithm [GJJ+03]. Closed-loop SMUX does not increase the effective number of data pipes (since open-loop SMUX already achieves this number), but it can improve the throughput on these data pipes.

The optimal signalling strategy for closed-loop SMUX requires theoretically a Gaussian codebook with continuous rate allocation among eigenmodes, an approach that is not feasible in practice. Therefore, a near optimal solution is required that takes into account the finite rate allocation granularity imposed by the limited number of modulation/coding schemes. The most accepted amongst such algorithms is Hughes-Hartogs scheme explained in [Bin90]. Its reduced complexity versions [CCB95][FH96][YC01] mainly rely on the fact that it is more important to avoid using the most attenuated sub-channels than to exactly find the optimum amount of power to be allocated to the strong sub-channels. For example, it has been shown in [TC04] that an on/off power distribution, as long as it uses nearly the same transmission sub-channels as WF, exhibits negligible capacity loss with respect to the exact WF shape. A simple logarithm-free, bit/power loading algorithm that requires low signalling overhead is proposed in [CTL04] for MIMO-OFDM systems. It maintains constant FER by controlling the instantaneous transmitted power in

such a way that the average SNR at the equaliser output is kept constant. This strategy results in negligible throughput degradation compared to the universally accepted Hughes-Hartogs algorithm.

2.3.3.3 Partial CSI SMUX

Closed-loop SMUX as described above requires instantaneous channel knowledge at the transmitter and usually complex bit/power loading algorithms. When only limited CSI is available, simpler schemes have been proposed to improve over the basic throughput gains of open-loop spatial multiplexing. These include Per-Antenna Rate Control (PARC) and Selective-PARC. We describe these techniques in detail in Section 3.1.4.

Instead of antennas also beams can be selected (without the receiver being aware of the use of antennas or beams). In case of a single beam, this corresponds to a fixed beam approach – in case of antenna set selection combined with vector modulation (BLAST) this corresponds to spatial multiplexing in the beam domain. Selection of beams is for instance suggested in [MZ04] (FFT over the antenna elements = orthogonal beams), showing a superior performance in spatially correlated channels compared to antenna selection.

2.3.4 Diversity-Multiplexing Trade-off Techniques

The diversity and SMUX techniques described above leverage competing MIMO gains that cannot be fully achieved simultaneously. Transmit diversity increases link reliability by introducing redundancy in multiple dimensions (space, time, frequency,…) and hence it often comes at the price of sending a unique data stream (or reduced number of streams) over all the available channels. SMUX techniques achieve high spectral efficiency by transmitting independent symbol streams, and hence do not directly provide reliability. We discuss here techniques that trade-off between these two types of MIMO gain.

Open-Loop Hybrid schemes [ZVR+03], also called Double Space-Time Transmit Diversity (DSTTD) [ODS02], offer a very simple such trade-off by transmitting several streams simultaneously, each stream being encoded accordingly to an STBC to benefit from transmit diversity. Detection is then being performed by applying one of the techniques already proposed for SMUX schemes (see Section 2.3.3) to the new system. This technique is described in detail in Section 3.1.2.

A natural trade-off between diversity and multiplexing arises in the context of frequency-selective fading MIMO channels as described in [BBP03]. For such channels, space-frequency coded OFDM is capable of realizing both spatial and frequency-diversity gains. The presence of multipath propagation and hence the availability of frequency-diversity generally allows to increase the “spatial signalling rate”, i.e., to use the multiple antennas to multiplex independent data streams rather than to provide spatial diversity. It is therefore desirable to be able to allocate the channel's degrees of freedom in space and frequency in a flexible way so as to achieve a variable allocation of the channel's degrees of freedom to multiplexing and diversity transmission modes. A simple construction is proposed in [BBP03] that consists of a linear outer code that guarantees a certain amount of frequency-diversity, and an inner (independently designed) space-time code that realises spatial diversity. The amount of SMUX is controlled by the inner code, whereas the outer code guarantees a prescribed diversity order. In general, the amount of spatial multiplexing available should be fed back from the receiver.

“Hard switching” between transmitter modes, also called spatial domain link adaptation, can trade off between the two gains by exploiting channel conditions fed back by the receiver. This technique is described in detail in Section 3.1.1.

To conclude this section, we note that a powerful theoretical framework has recently been developed in [ZT03] to fully characterise the trade-off between diversity and spatial multiplexing in MIMO systems. Practical techniques that optimally achieve this trade-off are still being investigated.

2.3.5 Receiver Architecture for (Coded) MIMO

The main classes of receiver architecture for MIMO techniques are:

- *Full search (or joint) maximum likelihood* (ML) receivers
- *Sphere decoder* as a low-complexity implementation of a (near) ML receiver
- *Linear ZF/MMSE receivers* (ZF: zero forcing, MMSE: minimum mean square error)

- (Ordered) *successive interference cancellation* (SIC) type receivers (e.g., V-BLAST)

Joint ML detection yields best performance at the cost of a high computational cost that can indeed get prohibitive in the MIMO context. Sphere decoding (or Fincke-Pohst algorithm) may be a computationally attractive alternative with close to ML performance [JO04]. The idea is to estimate the transmitted symbols by searching only over those symbols that belong to a hyper-dimensional sphere centred at the received signal. This algorithm has been adapted to decode lattice codes for the Gaussian and Rayleigh channels in [VB99], and extended to the MIMO case in [DCB00]. List sphere decoding algorithm is proposed in [HT03] to further reduce the complexity of the receiver.

Less expensive techniques are based on linear detection using the ZF or MMSE criteria to perform the detection of the transmitted signals. ZF linear processing nulls out completely unwanted signal components disregarding noise and results in noise enhancement. MMSE processing takes into account noise components and is usually preferred in cases when significant unwanted signal components are to be suppressed. The performance of linear receivers depends on the dimensionality of the system, i.e., the number of transmit and receive antennas, and on the correlation characteristics of the received signals.

The performance of linear receivers can be improved by combining them with successive cancellation. The key idea is layer peeling where the individual data streams are successively decoded and stripped away layer by layer. The algorithm starts by detecting an arbitrarily chosen data symbol (using ZF or MMSE front-end) assuming that the other symbols are interference. After the detection of the chosen symbol, its contribution from the received vector is subtracted and the procedure is repeated until all symbols are detected. To minimise effects of error propagation, it is advantageous to perform successive cancelling from the "strongest" to the "weakest" signal as proposed for V-BLAST [WFG+98] (specifically, the rule is to detect streams in the decreasing order of their post-processing SNR).

A similar family of techniques suited to coded transmission uses prior knowledge provided by a soft-in-soft-out decoder of the utilised channel code. Such "Turbo" approach considers the transmission as serial concatenation of a channel code (outer code) and the channel (inner code), usually separated by an interleaver. The soft-in-soft-out processing of the MIMO channel must usually be sub-optimal due to the dimensionality of the channel. One such algorithm relies on interference cancellation and MMSE filtering, which can be applied either as cancellation followed by filtering [WP99] [RW00] or with a feedforward-feedback filter pair [DKF+]. For low-dimensional cases reduced-complexity ML schemes can be utilised [AVZ02] [VB99] [DCB00] [BB03] [HT03].

2.3.6 Multi-user Techniques

Some topics discussed in this section are strongly related to task T2.4 on "Multiple Access" schemes. On the other hand, these general investigations (by means of information theory or simplified, analytical, or semi-analytical tools) can give additional insight and hints on the appropriate MIMO signal design for the various scenarios under consideration in WINNER. An important class of multi-user MIMO techniques is transmit precoding, which is especially suited to downlink access, i.e., from the BS/AP equipped with multiple antennas to users potentially also equipped with multiple antennas. Precoding is done based on some amount of channel knowledge at the Tx, and it can be linear or non-linear. We discuss these techniques in detail in Section 3.2.5-3.2.7 and 3.3 and investigate them in Chapter 7.

2.3.6.1 Antenna Hopping for Multi-User MISO/MIMO Downlink

Beamforming with random beamforming weights in combination with channel-dependent scheduling, often referred to as opportunistic beamforming has been proposed in [VTL02]. Here, we consider a multi-user downlink hopping system that operates similar to such opportunistic beamforming, with the difference that hopping operates in the *antenna* rather than beam domain. The gain is achieved by multi-user scheduling while hopping over the transmission antennas. This operation is particularly attractive in case of a large angular spread (i.e., rich scattering) when thinking of spatial separated antennas because beamforming is less beneficial. Another application is in the polarisation domain. Applying multi-user antenna hopping over cross-polarised antennas can be even combined with traditional opportunistic beamforming.

2.3.6.2 MIMO Gains without User Cooperation in Large Adhoc Networks

We consider here capacity scaling laws for large adhoc wireless networks as developed in the pioneering approach of [GK00] [GV02] [GT02]. It is proved in [BN04] that even when users don't cooperate, these networks exhibit MIMO-style gains: namely, multi-user multiplexing, distributed diversity, array, and interference cancelling gains are realised. Namely, for a fixed number of source-destination pairs L, and given the one-hop and two time slot communication model described in [BN04], the sum-capacity of the adhoc network in the large relay limit $K \to \infty$ is given by

$$C = \frac{L}{2}\log(K) + O(1) \text{ bps/Hz}.$$

This result follows from an upper bound that relies on the cut-set theorem and from a lower bound that relies on a simple relaying strategy described in [BN04]. We can conclude that the entire network behaves like a MIMO link with spatial multiplexing gain $L/2$, and with each of the multiplexed streams experiencing an array gain of K. The factor 1/2 penalty in the multiplexing gain comes from the fact that communication takes place over two time slots.

2.3.6.3 Ergodic Capacity and Outage Properties of CDMA

We describe here the ergodic and outage properties derived in [GB04] for a CDMA system where K users (with no CSI) communicate with the base station (that has full CSI) over a frequency-selective fading channel of length L. The users employ spreading sequences and OFDM, where the number of tones is chosen equal to the spreading gain N. These results extend the conclusions of Rupf and Massey [RM94] for time-invariant, deterministic CDMA systems to this more general channel model. Since spreading allows varying the amount of collision in signal space, it is interesting to understand how (if at all) the fundamental results in [RM94] change in the fading case. And because fading also gives rise to the notions of ergodic capacity and outage capacity [BPS98], which are not encountered in AWGN channels, it is important to understand the properties of spreading sequences that govern the ergodic capacity and the outage performance.

The conclusions given here focus on the single-antenna case, but the extension to multiple antennas is straightforward for ergodic capacity. For this measure, the so-called Welch Bound Equality (WBE) spreading sequences, as defined in [RM94], are optimal for $K >> N$, and they achieve the ergodic sum capacity of the underlying fading MAC. For the overspread case ($N \geq KL$), shift-orthogonal sequence sets maximise the ergodic sum capacity achieved by CDMA, which is strictly smaller than its counterpart for the underlying fading MAC. With respect to the high-SNR outage probability and to achievable diversity order of CDMA, WBE sequence sets are found to be optimal in the under-spread case ($K \geq N$) and strictly inferior to shift-orthogonal sequences in the overspread case. In the low-SNR regime, we show that neither WBE nor shift-orthogonal sequence sets yield optimality in terms of outage probability.

2.3.6.4 Sum Capacity Optimisation of MIMO Multiple Access Channel

The sum capacity of a MIMO MAC is given by

$$C(\mathbf{Q},\mathbf{H}) = \log\det\left(\mathbf{I} + \mathrm{SNR}\sum_{k=1}^{K}\mathbf{H}_k\mathbf{Q}_k\mathbf{H}_k^H\right),$$

where $\mathbf{H}_k$ is the channel matrix and $\mathbf{Q}_k$ the covariance matrix of the data of user k. Here, we have assumed that the receiver knows the channel realisations perfectly. If in addition the transmitters have full channel state information, we can optimise the sum capacity over the covariance matrices $\mathbf{Q}_k$ given the following sum power constraint:

$$\sum_{k=1}^{K} p_k = \sum_{k=1}^{K} trace(\mathbf{Q}_k) \leq P.$$

The sum power constraint is important if power can be distributed across the users in one cell but the cell sum power is limited in order to control the intercell interference.

The optimisation algorithm is described and analyzed in [BJ02]. It is divided into two parts, namely (i)

power allocation and (ii) transmit covariance matrix optimisation for fixed power allocation. The outer loop is between power allocation and covariance matrix optimisation under individual power constraints. The covariance matrix optimisation can be decomposed into an inner loop in which single-user covariance matrix optimisation with respect to the effective channel is performed. The inner single-user water-filling algorithm can be derived in closed-form [YRB+04]. Then, the covariance matrix optimisation corresponds to iterative water filling. This iterative structure can be further simplified: instead of complete iterative water-filling in each step, one single iteration is computed.

Interestingly, this algorithm yields that for small SNR only one user is active, in which case signal processing at the receiver is easy and the scheduling is very simple. Therefore, the analysis of this single-user range is important.

2.3.6.5 Worst Case Noise plus Interference Analysis for Multi-User MIMO

Interference from inter- or intra-cell users degrades the performance of a point-to-point MIMO link. If some information about this interference is known, it can be mitigated (even avoided) [RUY02], often via optimised joint user transmission [BJ02] [YRB+04]. When this is not possible, we can quantify this degradation in terms of capacity (in bits per channel use) by modelling the link as corrupted by a worst-case coloured noise plus interference (n+i) term [BJ03]. The desired signal and the n+i term are modelled according to the capacity achieving distribution (i.e., Gaussian) with covariance matrices **Q** and **Z**, respectively. This choice of distribution for the n+i term is motivated by the law of large numbers and it yields the worst-case interference, but we additionally need to solve for the worst n+i covariance matrix **Z** [BJ03].

The spatial signature of the interfering users determines the matrix **Z** whose eigenvectors are the directions of the interfering signals and whose eigenvalues are the average powers. We distinguish three cases: (i) fixed sum noise power, (ii) fixed interference power, and (iii) fixed per-receive-antenna noise power. The results in [BJ03] are as follows:

Case (i): the vector MIMO channel with perfect CSI at the transmitter transforms into a MIMO channel without CSI and white Gaussian noise as illustrated in Figure 2.5.

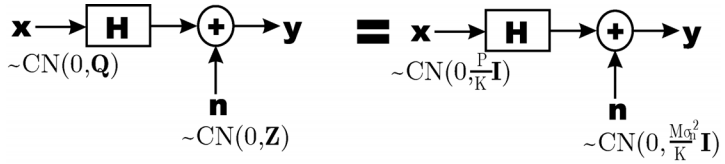


Figure 2.5: Vector MIMO channel capacity with worst case noise with trace constraint equals capacity of open-loop MIMO system and white noise (no interference)

Case (ii): the worst case noise directions are the right eigenvectors of the channel matrix times an arbitrary permutation matrix. In other words, with perfect CSI the worst-case n+i decomposes the MIMO link into a set of parallel channels as shown in Figure 2.6.

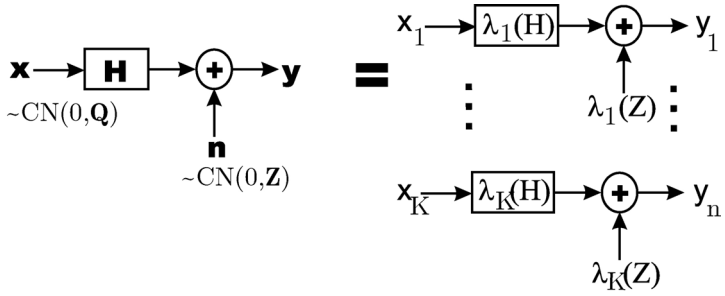


Figure 2.6: Worst Case Noise Directions: Vector MIMO channel capacity equals capacity of orthogonalised parallel channels with different noise powers

Case (iii): The worst-case noise colour reduces the achievable performance of a MIMO system with n cooperating transmit antennas to n users who perform only power control as shown in Figure 2.7.

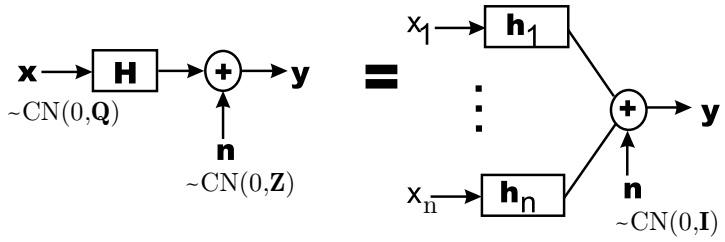


Figure 2.7: Worst Case Coloured Noise: Vector MIMO channel capacity equals sum capacity of corresponding SIMO MAC

2.3.7 Characterisation of Multi-Antenna Elements

We have presented *basic* beamforming (BF), spatial diversity, spatial multiplexing (SMUX), more *advanced* link-level MIMO techniques (e.g., matrix modulation), *combinations* of schemes, and *multi-user signal design*. We summarise in Table 2.1 the *basic characteristics*, *advantages* (+), and *disadvantages* (–) of the concepts/techniques that have been described here and in Chapter 3.

Based on this detailed review of multi-antenna methods, Task 2.5 of WINNER needs to reduce its scope for D2.7 and Phase I of the project. The selection of (the most promising) candidate multi-antenna techniques that can be investigated in detail is given here as a function of the scenario (wide-area or short-range). Indeed, the scenario critically determines the amount of CSI available and hence the techniques that are applicable, tough a single overall concept is ideally able to adapt to these scenarios.

The prevailing propagation conditions in the wide-area scenario are low angular-spread at the BS, large angular-spread at the mobile terminal (except for line-of-sight conditions), and (relatively) high velocities. Therefore, it is more realistic to assume that only long-term CSI and/or highly quantised short-term CSI (e.g., single CQI) can be made available to the transmitter. This essentially precludes closed-loop techniques that require very accurate short-term CSI. It allows, however, for techniques that limit the instantaneous knowledge to a reduced set of quantities (e.g., PARC, grid-of-beams BF, opportunistic BF, antenna selection), closed-loop techniques that rely only on long-term channel knowledge (e.g., long-term eigenbeamforming, multi-user beamforming based on 2nd order statistics for all users), open-loop techniques that do not require any channel knowledge apart from maybe a low-rate feedback to switch between transmission modes (e.g., open-loop Tx diversity, general open-loop matrix modulation, or antenna hopping), and noncoherent transmission. In addition, proper combination of techniques that exploit long-term CSI with open-loop signal design (e.g., long-term eigenbeamforming plus STBC, vector modulation, or matrix modulation across beams and/or polarisations) is interesting.

By contrast to the wide-area case, propagation conditions in short-range scenarios are characterised by an increased angular spread at the BS (therefore beamforming can be less effective, but SMUX and/or diversity is applicable), increased temporal dispersion (which allows for frequency diversity), and reduced time variability of the channel. Therefore, schemes that require high-quality short-term channel knowledge may be feasible in this case, whereas open-loop signal and noncoherent designs are less relevant. The list of promising techniques for hotspot scenarios includes multi-user linear (e.g., ZF, block diagonalisation) and nonlinear precoding (e.g., THP, sphere encoding), linear/non-linear closed-loop MIMO (e.g., SVD-based SMUX), open-loop vector/matrix modulation for above-pedestrian velocities, and general techniques discussed for wide-area coverage but with scheduling or link adaptation taking into account the spatial/temporal channel properties and Doppler spread typical for hotspot scenarios.

In order to provide a unified framework to represent and compare the identified techniques, we define the important notion of *spatial mode* in Section 2.4 and we introduce a generic multi-antenna transmission model in Section 2.5. In addition, since the final choice of techniques obviously depends on other air interface functions, it is increasingly important that Task 2.5 closely follows the results of other WP2 tasks and updates its simulation parameters and simulations accordingly.

		channel knowledge					kind of gain						complexity/flexibility/applications						
		no CSI at Tx	long-term CSI at Tx (maybe via feedback)	mode, channel rank, and/or CQI feedback	short-term CSI at Tx (maybe via feedback)	CSI required at Rx	Tx array gain	Tx spatial diversity	SMUX/SDMA gain	Rx array gain	Rx spatial diversity	IF suppression	Tx complexity	Rx complexity	signalling effort	flexibility	for wide-area	for short-range	....
focus on link-level processing	space time block codes (STBC)	✓	×	×	×	✓	×	✓	×	?	?	?	+	+	+	o	+	–	
	space time trellis codes (STTC or STTrC)	✓	×	×	×	✓	×	✓	×	?	?	?	+	–	+	o	+	–	
	space time Turbo codes (STTuC)	✓	×	×	×	✓	×	✓	×	?	?	?	+	– –	+	o	+	–	
	open-loop vector modulation (BLAST)	✓	×	(✓)	×	✓	×	×	✓	?	?	?	+	–	o	–	+	o	
	open-loop linear matrix modulation (LDC)	✓	×	(✓)	×	✓	×	✓	✓	?	?	?	+	– –	o	o	+	o	
	closed-loop transmit diversity (marks for FDD)	×	×	×	✓	✓	✓	✓	×	?	?	?	o	o	–	o	o	+	
	closed-loop spatial multiplexing (marks for FDD)	×	×	×	✓	✓	✓	(✓)	✓	?	?	?	o	–	– –	o	–	+	
	per antenna rate control (PARC) (marks for FDD)	✓	×	✓	×	✓	×	✓	✓	?	?	?	+	–	–	o	o	+	
	per stream rate control (PSRC) (marks for FDD)	×	×	✓	✓	✓	✓	(✓)	✓	?	?	?	o	–	–	o	o	+	
	long-term MISO beamforming on link (marks for FDD)	×	✓	×	×	✓	✓	×	×	?	?	?	+	o	o	+	+	o	
	long-term MIMO eigenbeamforming (marks for FDD)	×	✓	×	×	✓	✓	×	✓	?	?	?	+	o	o	+	+	o	
	joint design of LT EBF and ST mapping (marks for FDD)	×	✓	(✓)	×	✓	✓	✓	✓	?	?	?	o	o	o	+	+	o	
	MISO/SIMO/MIMO with optimized FEC	?	?	?	?	?	?	?	?	?	?	?	?	– –	?	– –	?	?	
	noncoherent MIMO transmission	✓	×	(✓)	×	×	×	✓	×	×	×	×	o	+	+	–	o	– –	
	mode switching based on channel rank (marks for FDD)	?	?	✓	?	✓	(✓)	(✓)	(✓)	?	?	?	o	?	o	+	+	+	
	(advanced) ST equalization	?	?	?	?	✓	?	?	?	✓	✓	✓	?	–	?	+	+	+	
	(advanced) space frequency equalization	?	?	?	?	✓	?	?	?	✓	✓	✓	?	–	?	+	+	+	
multi-user concepts	grid-of-beams beamforming + scheduling (marks for FDD)	×	(✓)	✓	×	✓	✓	(✓)	✓	?	?	?	o	?	o	+	+	–	
	opportunistic beamforming + scheduling (marks for FDD)	×	(✓)	✓	×	✓	✓	(✓)	✓	?	?	?	o	?	o	+	+	o	
	multi-user long-term beamforming (optimization at BS)	×	✓	×	×	✓	✓	(✓)	✓	?	?	?	o	?	o	+	+	–	
	antenna selection + scheduling (marks for FDD)	×	×	✓	×	✓	×	✓	(✓)	?	?	?	+	?	o	o	+	o	
	antenna hopping + scheduling	×	×	×	×	✓	×	✓	(✓)	?	?	?	+	?	o	o	+	o	
	multi-user linear precoding (marks for TDD)	×	✓	×	✓	✓	✓	✓	✓	✓	✓	✓	–	+	–	+	–	+	
	multi-user non-linear precoding (marks for TDD)	×	×	×	✓	✓	✓	✓	✓	✓	✓	✓	– –	?	–	+	– –	+	

Table 2.1: Qualitative classification and characterisation of MISO/SIMO/MIMO link-level transmission and multi-user MISO/SIMO/MIMO concepts via ✓ (correct), (✓) (may be correct, depends on implementation), × (not correct), ? (depends on other issues, general answer not possible), ++ (large advantage of this scheme), + (advantage of this scheme), o (average behaviour), – (drawback of this scheme), -- (large drawback of this scheme).

2.4 Spatial Modes

For assessment of different multi-antenna *techniques* it is important to define *spatial modes*, which are the combination of enabling requirements and spatial processing gains leveraged. For simplicity, in Figure 2.8 the multitude of enabling requirements is represented by one important parameter, which is the required channel knowledge at the transmitter. We distinguish open-loop techniques, where no channel knowledge is available at the transmitter, from cases where we have CQI or CSI, either as long-term or short-term information. Out of the fundamental spatial processing gains introduced in Section 2.2, Figure 2.8 shows spatial multiplexing (i.e., multiple different data streams to one user), SDMA (multiple data streams to different users), diversity (multiple identical data streams to one user), and beamforming (directive transmission, array gain). For the sake of this figure, the interference suppression gain is included in the beamforming and SDMA component.

Based on the concept of *spatial modes*, we can differentiate between *competing techniques* and *complementary techniques*. Ideally, *competing techniques* are those that allow implementing identical *spatial modes*, whereas *complementary techniques* implement different *spatial modes*. In reality there is a gradual transition between these two ideal cases and almost any two techniques will contain complementary aspects, in particular since almost all techniques differ in their detailed requirements regarding support functions and control information.

Thus, one of the goals of this document is to identify important *spatial modes* that enable a high-performance WINNER radio interface in all major scenarios and use cases. The initial proposal for the WINNER multi-antenna concept investigated in this deliverable will therefore contain a selection out of *competing techniques* and a combination of *complementary techniques* that actually implements the superposition of the requirements and spatial processing gains of all important spatial modes identified.

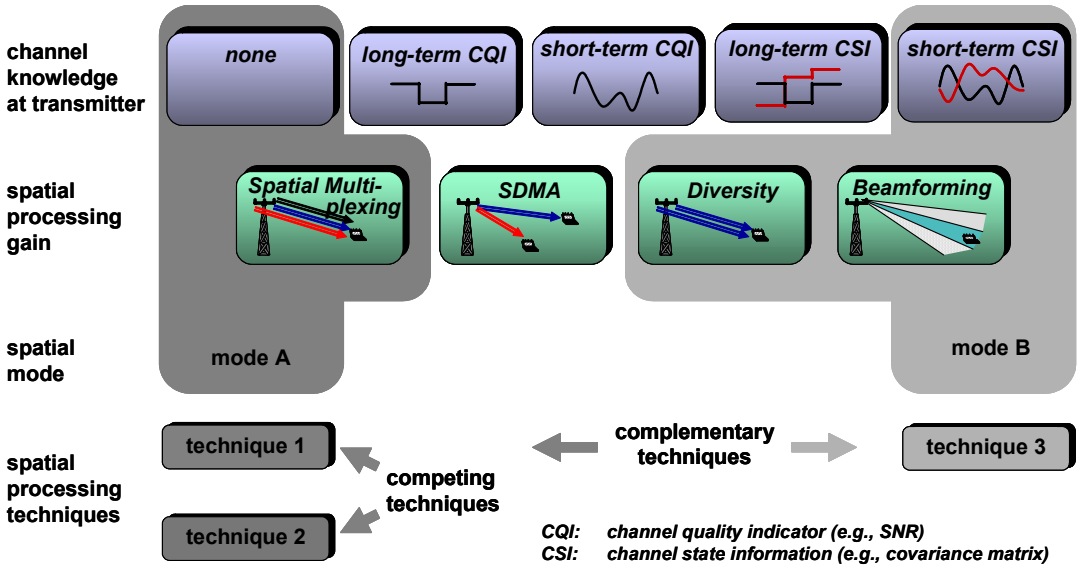


Figure 2.8: Spatial modes and multi-antenna techniques

2.5 Generic Spatial Processing

In Section 2.3, several spatial processing methods and techniques have been identified and also to some extent unified. The identified techniques affect the following functionalities:

- Resource management in general, multiple access and scheduling in particular.
- Transmit Processing.
- Receive Processing.

Spatial processing techniques are thus integrated parts of the above-mentioned functions, and a spatial processing method in general involves interaction and joint optimisation of them all. This optimisation may target for example performance as well as the complexity in different nodes, such as actual process-

ing and the required signalling.

This section provides an overview of the functions affected by spatial processing, and introduces a generic transmitter processing chain. Most of the studies and proposed techniques in Section 2.3 are compatible with this generic transmit chain and will be related to the function blocks described hereafter. For the receiver processing, no generic block diagram is presented. Suitable and intended receiver processing is treated in detail in Task 2.2, and partly in the present document along side the description of the transmission methods.

2.5.1 Scope and Assumptions on the Air Interface

In alignment with [T2.7], it is foreseen that a number of different transport channels are to be supported and each transport channel may be intended for a single user, a group of users, or all users. Further, there is a distinction between scheduled channels and channels that access the medium on a contention basis. Although this report investigates primarily scheduled channels, referred to as scheduled data channels (SDCH) and scheduled control channels (SCCH), with data transmission intended to a single user, i.e., dedicated scheduled channels, techniques for all the other transport channels should also be part of the final WINNER multi-antenna concept.

In general, each scheduled channel consists of several multiplexed data streams (flows) with separate channel coding. However, in order to simplify the description in this section, only one flow per user will be assumed.

The physical channel structure currently assumed divides the available time-frequency resources into chunks, each chunk consisting of a set of (orthogonal) waveforms. The chunks are considered two-dimensional, and at least for the downlink, each chunk consists of a number of subcarriers and a number of consecutive OFDM symbols. For the uplink, and in the general case, other modulating waveforms may also be considered, and in the description below, “frequency” or “subcarrier” should then be replaced by “waveform”. Note that in the present deliverable, no specific assumptions are made regarding the size of the chunks.

When multiple antennas are introduced, the spatial dimension is added to the chunks. The chunks may thus be viewed as three dimensional, and the third dimension will be referred to as a layer. Furthermore, in a simple form of spatial transmit processing, each layer may be encoded, modulated, possibly spread and transmitted from a single physical antenna. In the general case, each layer is though dispersed and transmitted from all physical antennas. Note that this kind of layering can also be realised with a single transmit antenna by using (orthogonal) spreading codes.

2.5.2 Multiple Access and Scheduling

A number of different users are scheduled and concurrently use the available resources. In the most general case, the multiple access can use the time, frequency, code, and space domain. An exemplary scheduling process is illustrated in Figure 2.9. In this example, the code domain (i.e., CDMA) is not used, and the space dimension represents spatial multiplexing. Furthermore, the assignment for one user is contiguous in time, frequency, and space, although especially for non-adaptive systems an interleaved assignment can be envisaged to maximise diversity in the corresponding dimension. It is important to note that the quality of the scheduling process is of crucial importance to the overall system performance. A long-term goal is therefore the joint optimisation of resource allocation in all dimensions. As an initial step towards this aim, this document focuses primarily on the scheduling opportunities in the spatial domain. Channel dependent scheduling/resource allocation can be used to realise multi-user diversity gains in addition to spatial multiplexing gains (spatial division multiplexing and spatial division multiple access).

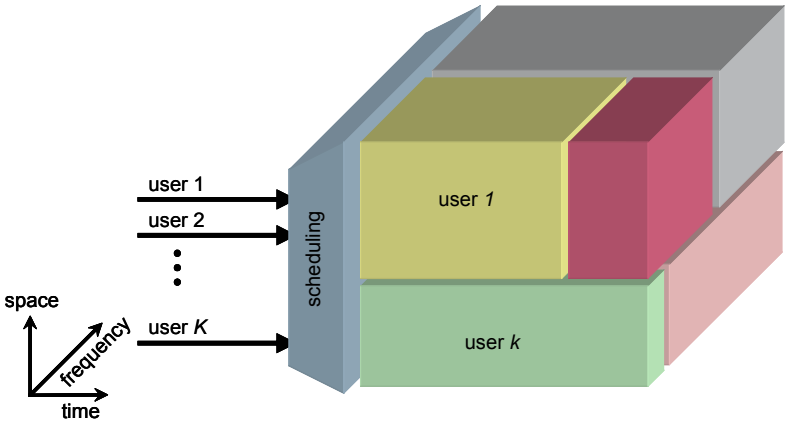


Figure 2.9: Multiple Access and Scheduling

2.5.3 Multi-Antenna Transmission

Figure 2.10 shows a generic multi-antenna transmission chain that serves as a framework to explain the general properties of the spatial transmit processing, as well as for further detailed discussions of the individual techniques in the following chapters. Note that not all building blocks of this generic chain may be used in different techniques, and that the generic transmission chain should be viewed as a functional sketch and not as an implementation specification. Furthermore, it currently appears as if the same antenna transmission chain is applicable both for uplink and downlink, although the waveforms used for uplink may not necessarily be OFDM subcarriers. In multi-hop and relaying scenarios, the transmission chain may be applicable for the individual hops.

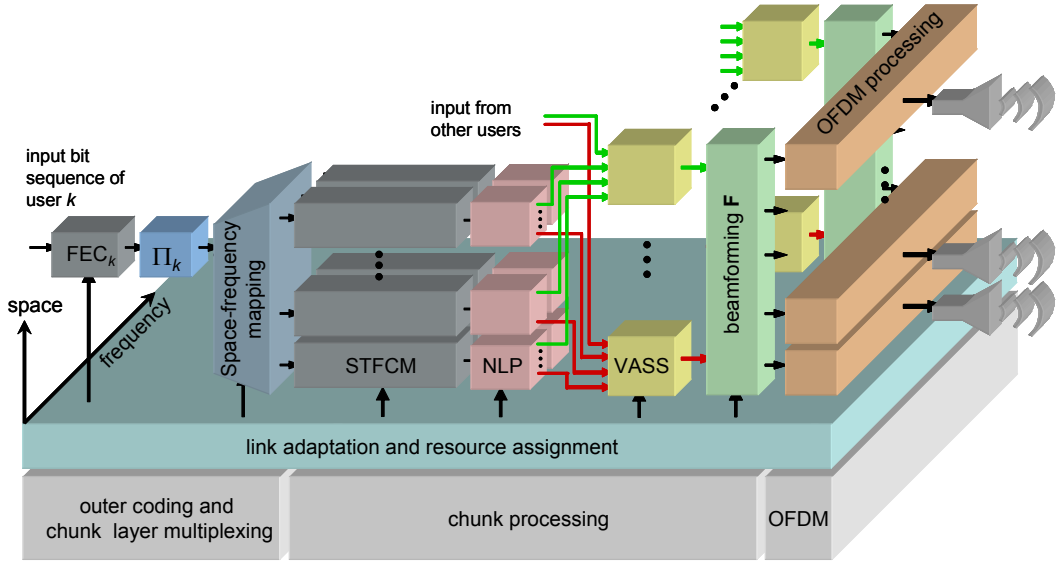


Figure 2.10: Generic Multi-antenna Transmission Technique.

Resource assignment (e.g., scheduling) and *link adaptation*, i.e., any kind of adaptation to the channel properties, have a major impact on the operation of the building blocks. *Link adaptation* might use *mode selection* (selection of FEC *type*, modulation *type*, HARQ *type*, and spatial processing *type*) or simply *rate adaptation* (adaptation of code rate, modulation level and spreading in space, time and frequency). In general, *mode selection* will adapt to widely varying operating conditions, whereas *rate adaptation* can be used for matching the channel conditions on a smaller scale.

In the generic multi-antenna transmission chain depicted in Figure 2.10, the input bit sequence of a particular user k passes forward error coding (FEC_k) and interleaving Π_k. There is also a possibility that

the data stream of a user is first segmented into several parallel subflows, in which case each subflow is processed with an individual channel code and interleaver and then split into substreams. As mentioned above, to simplify the description only a single flow per user is considered here.

Depending on the decisions of the resource management, this stream is mapped onto different substreams that each covers n_{sub} subcarriers, and one spatial layer. It is assumed that this substream processing operates on a frame basis, i.e., on n_{frame} OFDM symbols. Therefore, each substream represents one spatial layer q of a chunk according to the notation in [T2.7]. For each chunk c, the maximum number of spatial layers $Q(c)$ can be different.

In the most general case, each substream is then mapped onto a number of three-dimensional *virtual antenna chunks* that span M_{va} virtual antennas in space, n_{frame} OFDM symbols in time, and n_{sub} subcarriers in frequency. This function will be referred to as space-time-frequency coding and modulation (STFCM). Although there are many possibilities, it is currently believed that many of the previously proposed techniques will use the baseline STFCM given in Figure 2.11. Then each substream q undergoes modulation M(c,q), and power assignment P(c,q). The modulated and power assigned stream is then encoded with a linear dispersion code (LDC) and each resulting parallel LDC encoded stream is then written into the two-dimensional chunk associated with a virtual antenna, either frequency-by-frequency or symbol-by-symbol. Here we note that:

- Our initial investigations will not consider the use of an inner FEC on a chunk layer level (since only few bits may constitute a chunk layer and multiple coding stages are difficult to design).
- LDC, which may be viewed as a straightforward generalisation of multi-code DS-CDMA, can be used to represent not only spreading, but also a large class of so-called vector and matrix modulation schemes, see Section 2.3.2.2.1.

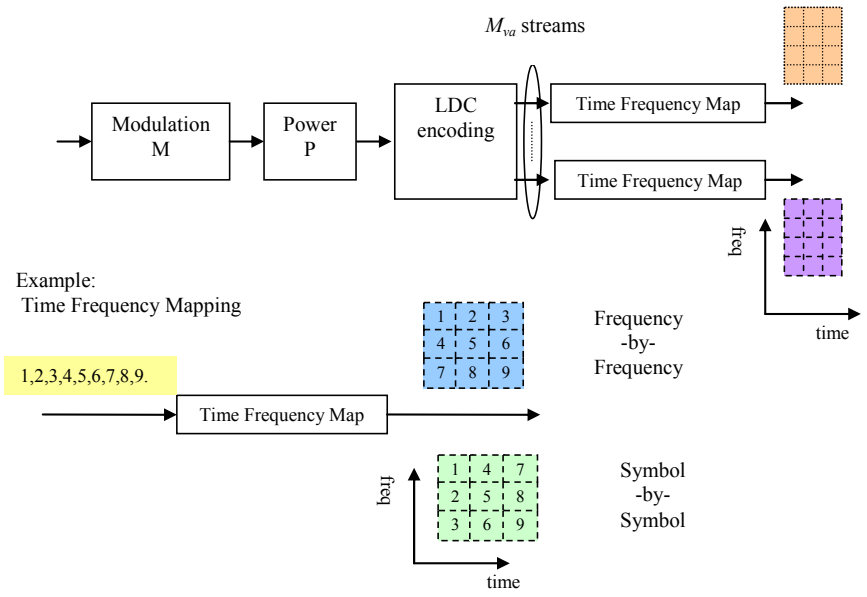


Figure 2.11: Baseline Space-Time-Frequency Coding and Modulation.

Then, referring to Figure 2.10 again, the modulated and coded virtual antenna streams are passed to the block denoted NLP (non-linear precoding) that performs non-linear pre-processing, typically in terms of non-linear feedback and a modulo operation on a per virtual antenna stream basis. Afterwards, the virtual antenna streams are combined in the *virtual antenna stream summation* (VASS) block. The number of virtual antenna streams is also expected to depend on the channel properties, the degree of multi-user optimisation, and the scheduling. Subsequently, beamforming is applied to produce the desired directivity. The beamforming function, which maps the combined and processed *virtual antenna streams* into *antenna streams* typically uses some form of CSI, such as an estimate of the instantaneous channel or

the second order statistics. Beamforming in the present context not only covers techniques such as closed-loop transmit diversity and long-term beamforming, but also random beamforming employed by opportunistic beamforming approaches.

The three dimensional antenna chunks are then concatenated and the OFDM processing is performed per antenna (*N*-point IFFT, cyclic prefix addition, parallel/serial conversion).

It is worth stressing that scheduling, multiplexing, and link adaptation are done jointly with the physical layer receive and transmit processing in terms of the selection of linear dispersion codes and beamforming weights. As an example, beamforming weights and dispersion/spreading codes are typically not chosen independently, but rather optimised jointly when multiple streams are transmitted simultaneously.

The focus and purpose of the present work is not necessarily to optimise a certain block of this generic transmitter, e.g., to design optimum linear dispersion codes and linear precoders/beamforming weights. Rather, the work will motivate the inclusion of various function blocks as well as combinations so as to develop the WINNER concept and leave room for future evolution.

2.5.4 Receiver Processing

As mentioned above, a multi-antenna technique means tight interaction and joint optimisation of receiver processing, transmitter processing and resource management. Thus, various transmission and resource management methods typically have implications on adequate receiver structures. This holds also the other way around in the sense that various receiver methods impose requirements on the transmit processing, and the way to proceed is to consider them jointly.

Suitable receiver methods for spatial processing have been identified in Section 2.3.5, and additional advanced receiver techniques are studied in detail in Task 2.2. Within the framework of this work, advanced receiver processing with multiple receive antennas (e.g., iterative techniques for demodulation, equalisation and detection) is seen as a multi-antenna technique in itself although it most likely has only low requirements on functionality in the transmitter and the resource management. Joint optimisation and design of multi-antenna transmitter and (iterative) receiver structures is a possibility to improve system performance even further.

While many spatial transmit techniques can be received with standard receiver structures, particular spatial processing algorithms (like noncoherent codes and non-linear precoding described later on in this document) require dedicated receiver functionality. For the scope of this document, specific requirements for the receiver architecture will be stated along with the detailed description of the corresponding transmit techniques.

The following is noted at this point:

- The use of non-linear precoding at the transmitter will affect the receiver in the sense that a modulo operation is performed at the receive side as well. This will be described in the next chapter of this report.
- The so-called noncoherent codes currently considered have no structure, and is expected to require a slightly different receiver structure than the conventional coherent transmit structures.
- Iterative techniques for demodulation, equalisation and detection appear to be feasible to approach the performance of joint approaches at a fraction of their complexity. For such approaches, there is room for further improvement by joint design of transmitter and receiver.

2.5.5 Relaying

Although the processing performed at each stage of the multi-hop communication can be mapped onto the generic multi-antenna transmission technique described above, the idea of relaying and multi-hop communication goes far beyond simply cascading individual links. The benefits and gains of multi-antenna relaying techniques over the entire communication path, and the interaction with the specific constraints and requirements of relaying need to be investigated in detail. Although first investigations on multi-antenna techniques for relaying are included in this document, significant future work will be conducted in later phases of the WINNER project.

3. Multi-Antenna Techniques

Chapter 3 provides a more detailed description of the selected spatial processing techniques. Section 3.1 presents techniques that focus on the optimisation of a single link, whereas in Sections 3.2 and 3.3 linear- and non-linear multi-users techniques are considered.

Multi-user MIMO systems provide a high capacity with the benefits of space division multiple access. For the downlink, the channel state information at the base station (BS) or access point (AP) is very important since it allows joint processing of all users' signals, and this improves performance and increases data rates significantly. When channel state information is available at the BS/AP, it can be used to efficiently eliminate multi-user interference (MUI) by beamforming or by using "dirty-paper" codes. The precoding also allows us to perform most of the complex processing at the BS/AP, which simplifies users' terminals. Linear precoding techniques that are covered in Section 3.2 have an advantage in terms of computational complexity. Non-linear techniques treated in Section 3.3 have a higher computational complexity and require some signalling overhead but can provide a better performance than linear techniques. The sensitivity of these proposed techniques to the channel estimation errors at the transmitter will be also analyzed.

Finally, multi-antenna techniques for relaying terminals are described in Section 3.4, before Section 3.5 summarises the main conclusions of this chapter.

3.1 Techniques Focusing on the Optimisation of a Single Link

3.1.1 A Spatial Domain Link Adaptation Technique

3.1.1.1 Overall Description of the Technique

Section 2.2 defined the spatial diversity and SMUX gains of multiple antennas, and Section 2.3.4 described techniques that trade off these generally competing gains. In particular, spatial domain link adaptation is a multi-antenna technique that allows hard switching between multi-antenna methods based on monitoring the short-term characteristic of the radio channel at the receiver. Conceptually, each method corresponds to a single spatial signal-processing algorithm that provides either “pure multiplexing” or “pure diversity” gains. However, it is also possible to adapt between methods that already trade to different extents by themselves the two performance gains.

The properties of the wireless medium given by the channel matrix **H** determine in the end what trade-off between multiplexing and diversity gains can be achieved. For example, spatial multiplexing is not amenable for rank-deficient channel matrices because the medium does not provide the maximum number of spatial channels. Therefore, in scenarios with a strong line-of-sight component or high transmit and/or receive correlation, this technique would fail to achieve the full multiplexing gain and diversity would be advisable. More generally, when a set of multi-antenna methods is available, it is of great interest to be able to decide which method is better suited for the given actual channel conditions. As selection criteria between the two multi-antenna methods, various metrics have proposed.

A first measure is the practical rank (Prank), which is defined by the essentially non-zero eigenvalues of the matrix $\mathbf{H}\mathbf{H}^H$. The underlying idea of the Prank is that singular values of the channel that are of the order of the noise power should not be relied on when separating streams from each other. Thus, the absolute Prank may be defined as

$$Prank = \max\left\{ i \mid \lambda_i\left(\mathbf{H}\mathbf{H}^H\right) > t_{Prank} / SNR \right\}, \tag{3.1}$$

assuming that the eigenvalues are in descending order. For use of the practical rank with concrete signalling schemes, the threshold t_{Prank} may vary depending on the details of the schemes.

In [HP02], R. Heath and A. Paulraj proposed to use the Demmel condition number,

$$\kappa_D = \frac{\lambda_{\min}\left(\mathbf{H}\mathbf{H}^H\right)}{\left\|\mathbf{H}\right\|^2}, \tag{3.2}$$

for quantifying the trade-off between diversity and spatial multiplexing. In the definition of the Demmel condition number, λ_{min} is the smallest eigenvalue of $\mathbf{HH}^H$. $\|\mathbf{H}\|^2$ is the Frobenius norm of the channel matrix, which measures the total power conveyed over all channel paths to the receiver. The performance of "pure spatial multiplexing" significantly depends on the rank of the channel and is strongly determined by the minimum eigenvalue λ_{min}, whereas "pure diversity" is insensitive to the channel rank and only depends on the Maximum Ratio Combined channel power measured by $\|\mathbf{H}\|^2$. Therefore, the Demmel condition number captures the essence of the diversity/multiplexing trade-off and can be used to switch between diversity transmission and spatial multiplexing use.

Another measure that can be used is the relative condition number of the channel matrix. The relative condition numbers are given by

$$\frac{\lambda_n\left(\mathbf{HH}^H\right)}{\lambda_{\max}\left(\mathbf{HH}^H\right)}, \tag{3.3}$$

where we compare the n-th smallest eigenvalue to the largest eigenvalue assuming normalised channel power. Depending on the supported symbol rates, different higher relative condition numbers could be used to decide on the transmission symbol rate in combination with the SNR.

Please note that these two measures, the Demmel condition number and the relative condition number are just characterised by a single number. Similarly, several measures (e.g., more than one relative condition number or all eigenvalues themselves) could simultaneously be used to choose the best-suited transmission scheme.

3.1.1.2 Required Support Functions

To optimally select the best-suited multi-antenna method, the spatial link adaptation algorithm requires full channel state information (CSI) at the receiver, that is, the complete knowledge on the instantaneous realisation of the radio channel. Therefore, spatial link adaptation requires the implementation of means and algorithms for fast and reliable channel estimation at both ends of the link. The transmitter does not require full CSI because the receiver can feedback to the transmitter a single parameter CQI, indicating the instantaneous channel quality and thus the best-suited method for the next frame transmission. Simultaneous switching between methods at the transmitter and receiver has to be controlled by the radio protocols to be implemented in the radio access network.

3.1.1.3 Additional Assumptions and Simplifications for Simulations

The spatial link adaptation technique for switching between different spatial signal-processing methods has been investigated for two multi-antenna arrangements with different numbers of transmit and receive antennas. The first arrangement is characterised by two transmit and two receive antennas and has been studied for a short-range communication environment. The second arrangement considers four transmit and four receive antennas and the investigations focus on the wide-area usage scenario.

A) Spatial Link Adaptation for 2x2 MIMO-OFDM System for Short-Range Usage Scenario

Figure 3.1 illustrates the investigated spatial link adaptation technique for a MIMO-OFDM system with two antennas at the transmitter and receiver. The two available spatial signal-processing methods are space-time coding according to the Alamouti (AM) scheme and spatial multiplexing (SMUX). The OFDM system parameters have been chosen according to the WINNER short-range usage scenario, that is, the available bandwidth is 20/40 MHz at a carrier frequency of 5 GHz, the number of subcarriers is 64/128, etc. The QAM modulation and coding scheme has been selected so that both the AM and SMUX method transmit in average the same number of bits in a symbol period. Alamouti coding is performed on two subsequent OFDM tones; therefore, it is assumed that the frequency response of the channel is constant over any two neighbour subcarriers. The radio channel models a typical large office (IEEE channel model D). For the simulations, it is assumed that the receiver perfectly knows the channel. To obtain estimates of the transmitted symbols, the processing performed at the Alamouti receiver basically orthogonalises the channel irrespective of its realisation. For spatial multiplexing, a low-complexity MMSE receiver has been selected instead of a high-complexity successive cancellation scheme that would outperform the linear receiver at high SNRs. To decide for the Alamouti or spatial multiplexing scheme as the best-suited method for a given channel realisation**,** the receiver computes a metric for the

two candidate spatial signal-processing techniques and feedbacks the corresponding optimal mode as link quality parameter to the transmitter. As selection criterion, the Demmel condition number κ_D is used. It has been shown in [HP02] that spatial multiplexing outperforms the Alamouti method in terms of Euclidean distance if κ_D is above a threshold that is independent of the channel realisation:

$$\kappa_D = \frac{\lambda_{\min}\left(\mathbf{H}\mathbf{H}^H\right)}{\|\mathbf{H}\|^2} \geq \frac{d^2_{\min,SM}}{d^2_{\min,AM}}. \tag{3.4}$$

The threshold is given by the ratio of $d^2_{\min,SM}$ and $d^2_{\min,AM}$, which are the minimum Euclidean distances in the transmit constellations of the modulation scheme used for the two spatial processing methods.

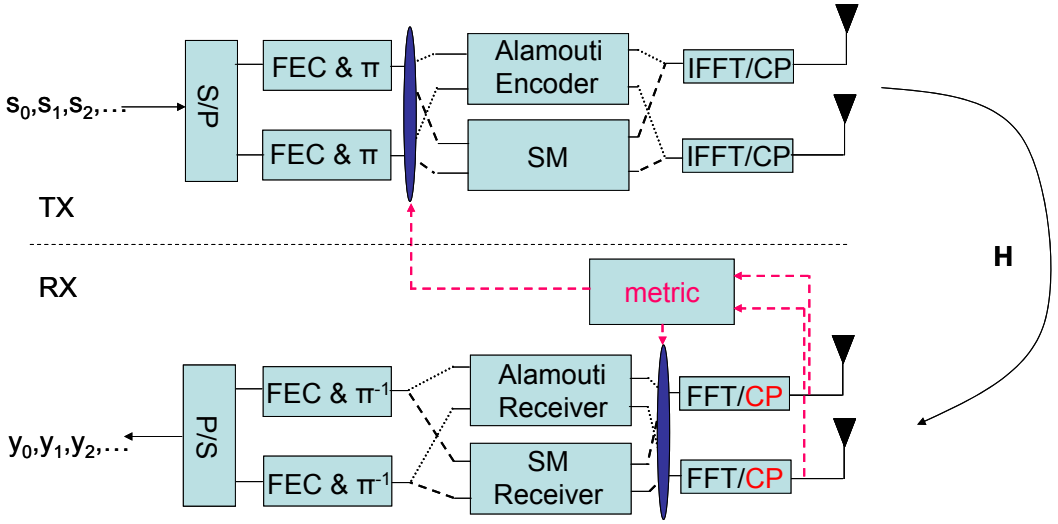


Figure 3.1: An example of a hard-switching scheme for MIMO-OFDM (M_T=2)

B) Spatial Link Adaptation for 4x4 MIMO-OFDM System for Wide-Area Usage Scenario

In case of 4x4 MIMO systems we consider switching between diagonal ABBA, double ABBA and vector modulation (spatial multiplexing). Diagonal ABBA (Diag-ABBA) is a rate 1 matrix modulation and provides diversity degree 4, whereas Double ABBA (DABBA) has rate 2 and provides diversity degree 2. The 4x4 modulation matrix for Diag-ABBA is given as

$$\mathbf{X}_{\text{Diag-ABBA}} = \frac{1}{\sqrt{2}}\begin{bmatrix} \mathbf{X}_A + \mathbf{X}_B & \mathbf{0} \\ \mathbf{0} & -\mathbf{X}_B + \mathbf{X}_A \end{bmatrix}, \tag{3.5}$$

where $\mathbf{X}_A$ and $\mathbf{X}_B$ are 2x2 STTD blocks of the form,

$$\mathbf{X}_A(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{bmatrix} \qquad \mathbf{X}_B(x_3, x_4) = \begin{bmatrix} e^{i\pi/4}x_3 & e^{i\pi/4}x_4 \\ -e^{-i\pi/4}x_4^* & e^{-i\pi/4}x_3^* \end{bmatrix}, \tag{3.6}$$

with x_i denoting the modulation symbols. Here the rows of $\mathbf{X}_{\text{Diag-ABBA}}$ correspond to the transmit antennas and its columns correspond to the neighbouring subcarriers (or time in case of space-time coding).

The 4x4 modulation matrix for DABBA (rows correspond to antennas and columns to subcarriers) is given as

$$\mathbf{X}_{\text{DABBA}} = \frac{1}{\sqrt{2}}\begin{bmatrix} \mathbf{X}_A + \mathbf{X}_B & \mathbf{X}_C + \mathbf{X}_D \\ -\mathbf{X}_D + \mathbf{X}_C & -\mathbf{X}_B + \mathbf{X}_A \end{bmatrix}, \tag{3.7}$$

where $\mathbf{X}_A$ and $\mathbf{X}_B$ are as above, $\mathbf{X}_C$ and $\mathbf{X}_D$ are of the STTD form

$$\mathbf{X}_C(x_5, x_6) = \begin{bmatrix} x_5 & x_6 \\ -x_6^* & x_5^* \end{bmatrix} \qquad \mathbf{X}_D(x_7, x_8) = \begin{bmatrix} e^{i\pi/4}x_7 & e^{i\pi/4}x_8 \\ -e^{-i\pi/4}x_8^* & e^{-i\pi/4}x_7^* \end{bmatrix}, \tag{3.8}$$

and x_i are the modulation symbols. For more information on these matrix modulations, see [HTW03] and references therein.

Traditional adaptive modulation and coding in SISO channels is based on the received SNR, which is the most viable Channel Quality Indicator (CQI) also in MIMO channels. In addition to received SNR, we use the second relative condition number

$$\frac{\lambda_2\left(\mathbf{H}\mathbf{H}^H\right)}{\lambda_{\max}\left(\mathbf{H}\mathbf{H}^H\right)}, \tag{3.9}$$

where $\lambda_2\left(\mathbf{H}\mathbf{H}^H\right)$ denotes the second smallest eigenvalue of $\mathbf{H}\mathbf{H}^H$, and $\lambda_{\max}\left(\mathbf{H}\mathbf{H}^H\right)$ is the largest eigenvalue. This particular CQI has been empirically seen to offer the best correlation between performance of symbol rate 1/2/4 modes and a single parameter describing the eigenvalue distribution of the channel.

If adaptation between different transmission modes is not done on each data subcarrier of an OFDM system, the CQIs have to be averaged over the subcarriers where the spatial transmission mode is not changed. The simulation results present in subsequent chapters are based on adaptation over one OFDM symbol. All results are based on 2D optimisation. For each value of received SNR[1] (1 dB step) and the averaged second condition number (8 quantisation bins), the mode that provides the highest packet throughput is chosen.

3.1.1.4 Conclusions

Spatial link adaptation is a promising multi-antenna technique that implements hard switching between two spatial signal-processing methods based on the short-term characteristic of the radio channel. The adaptation can be performed on the basis of several metrics that directly depend on the particular channel realisation. The benefits of such an approach have been proven to lead to substantial gains in error performance. It can be foreseen that like in traditional link adaptation algorithms, further performance improvements can be achieved when more than two transmission modes are allowed. The extremes modes "pure diversity" and "pure multiplexing" can be complemented by "intermediate" modes that trade to different extents the transmission rate for diversity gain. The disadvantage of the proposed multi-antenna technique is the required fast feedback of the channel-quality indicator from the receiver to the transmitter.

3.1.2 Hybrid Spatial Division Multiplexing/Space-Time Block Coding Schemes

3.1.2.1 Overall Description of the Technique

Open-loop multi-antenna approaches benefit from multiple antennas at the transmitter and receiver for spatial diversity or for SMUX gain. Individually, the methods that leverage these gains suffer from limitations highlighted in Section 2.3.4.

This motivates the search of high rate multi-antenna schemes combining the advantages of Space-Time Block Coding and Spatial Division Multiplexing. A first one has been proposed by Texas Instruments to the 3GPP committee [TI01] to transmit only 2 parallel data streams on 4 antennas, as presented in Figure 3.2. Other schemes have been presented to the IEEE802.11n committee by Realtek Semiconductors [MB04]. The first one is a rate 1 scheme based on the transmission of Alamouti codes on 2 antennas among the N available, with a cycling process to use successively all sets of 2 antennas. The second one applies the classical SMUX approach to N_t antennas among N, with a cycling process to use successively all sets of N_t antennas, but it has to face a limited amount of transmit spatial diversity compared to the use of space-time block coding.

[1] A definition of *SNR* for MIMO-OFDM systems is provided in [D7.2]

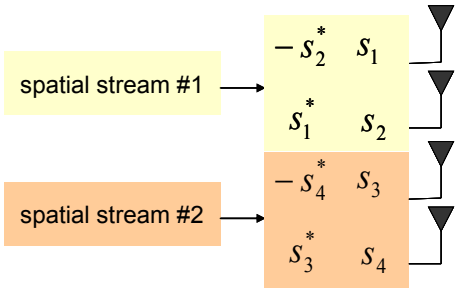


Figure 3.2: Transmission of 2 spatial streams on 4 antennas (SMUX/STBC)

New multi-antenna solutions based on different tradeoffs between data rate and robustness have then to be proposed, and are investigated here. These solutions consist in transmitting N-1 parallel streams on N antennas, one of these streams being encoded by an Alamouti code. The corresponding SMUX/STBC schemes are illustrated in Figure 3.3 and Figure 3.4 for 3 and 4 transmit antennas, respectively.

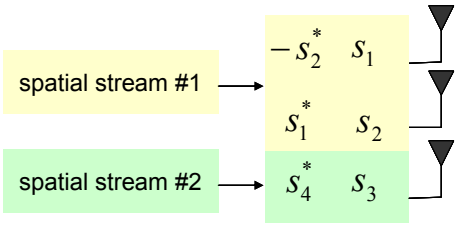


Figure 3.3: Transmission of 2 spatial streams on 3 antennas (SMUX/STBC)

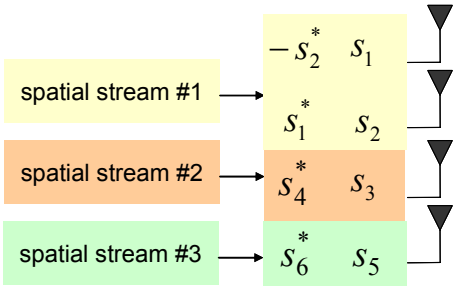


Figure 3.4: Transmission of 3 spatial streams on 4 antennas (SMUX/STBC)

The 3 hybrid schemes relying on a combination of SMUX and STBC are particularly well suited to achieve high data rates in asymmetric antenna configurations, where the number of transmit antennas is superior to the number of receive antennas. These configurations can occur for instance in the downlink between an access point and a station.

To illustrate the interest of these approaches, we present here a few physical layer (PHY) simulation results to compare their performance in upgraded asymmetric configurations to the performance of classical SMUX in square configurations. The simulation parameters used for this performance assessment are: 20 MHz bandwidth, packet size: 1000 bytes, IEEE802.11a convolutional code with coding rates 2/3, 3/4 and 5/6, 16-QAM and 64-QAM constellations, OFDM modulation based on 64-point IFFT (48 data subcarriers), channel IEEE TGn D NLOS (typical office environment), MMSE MIMO detection, perfect channel estimation.

Three comparison scenarios are investigated: 1) comparison of 2x2 and 3x2 antenna configurations, 2) comparison of 2x2 and 4x2 antenna configurations, and 3) comparison of 3x3 and 4x3 antenna configurations. The corresponding PER vs. SNR curves are plotted in Figure 3.5, Figure 3.6, and Figure 3.7 respectively. For each comparison scenario, we observe that the hybrid scheme combined with 1 or 2

additional transmit antennas provides a significant gain compared to SMUX in a square configuration. The gains at PER=10^{-2} vary from 2.7 dB to 5.1 dB for scenario 1, from 4.7 dB to 8.1 dB for scenario 2, and from 2.2 dB to 4.8 dB for the scenario 3.

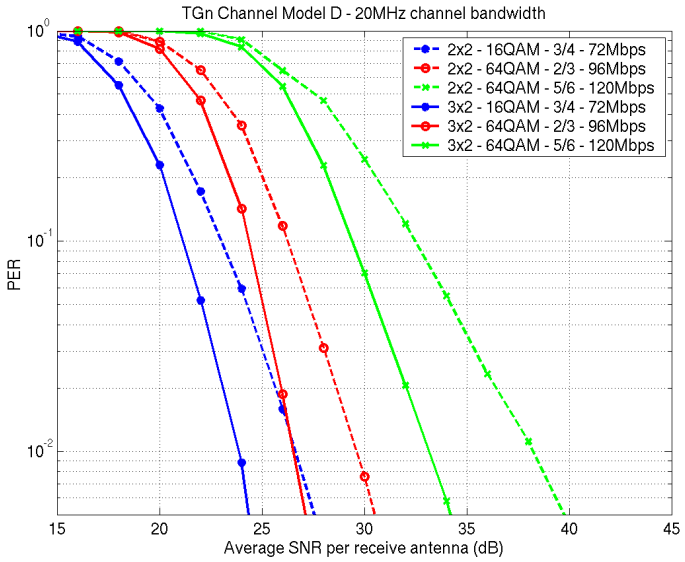


Figure 3.5: Performance comparison of 2x2 SMUX and 3x2 SMUX/STBC, both transmitting 2 parallel streams

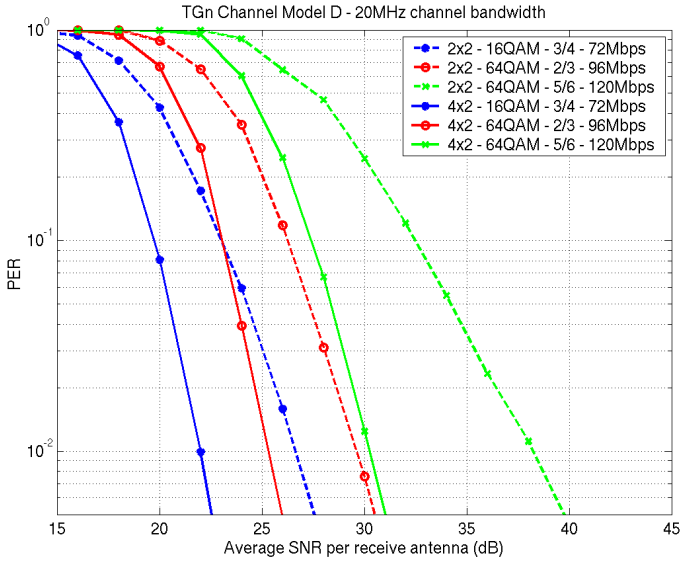


Figure 3.6: Performance comparison of 2x2 SMUX and 4x2 SMUX/STBC, both transmitting 2 parallel streams

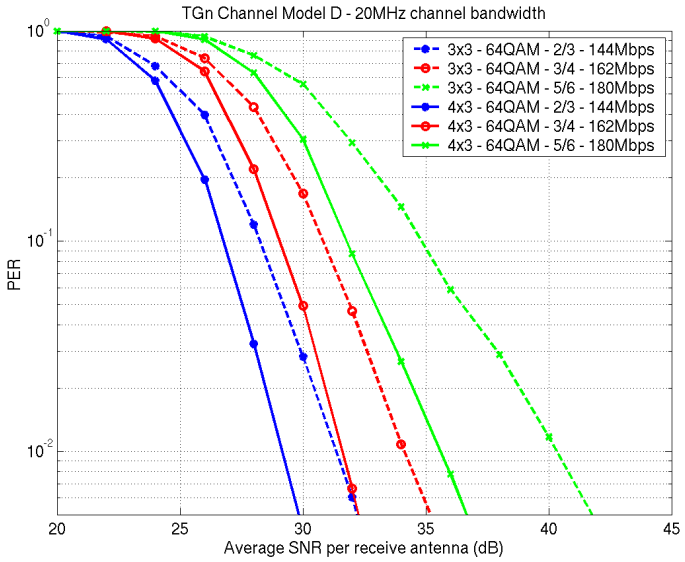


Figure 3.7: Performance comparison of 3x3 SMUX and 4x3 SMUX/STBC, both transmitting 2 parallel streams

3.1.2.2 Additional Assumptions and Simplifications for Simulations

We propose here to define some short-range PHY modes for 2, 3 or 4 transmit antennas. For each transmit configuration, we select two multi-antenna transmit schemes: the first one is dedicated to the support of robust low to medium data rate modes, and the second one is chosen to achieve high data rates. These transmit schemes are presented in Table 3.1 for all antenna configurations.

Table 3.1: Parameters of the multi-antenna transmit schemes

Number of transmit antennas	Multi-antenna scheme	Number of spatial streams	Number of symbols per transmit block
2	STBC	1	2
2	SMUX	2	2
3	SMUX/STBC	2	4
3	SMUX	3	3
4	SMUX/STBC	2	4
4	SMUX/STBC	3	6

The PHY modes defined in this section are based on a set of simulation parameters that has been identified for the short-range scenario: 20 MHz bandwidth, 64-point IFFT, 48 data subcarriers, and 0.8 µs cyclic prefix length. They are presented in Table 3.2 for 2 transmit antennas, and in Table 3.3 for 3 or 4 transmit antennas.

Table 3.2: Definition of the transmission modes for 2 Tx

Data rate (Mbits/s)	Number of spatial streams (NS)	Modulation	Coding rate (R)	Number of data subcarriers (NSD)
6Mbps	1	BPSK	1/2	48
12Mbps	1	QPSK	1/2	48
18Mbps	1	QPSK	3/4	48
24Mbps	1	16-QAM	1/2	48
36Mbps	1	16-QAM	3/4	48
48Mbps	1	64-QAM	2/3	48
60Mbps	1	64-QAM	5/6	48
72Mbps	2	16-QAM	3/4	48
96Mbps	2	64-QAM	2/3	48
108Mbps	2	64-QAM	3/4	48
120Mbps	2	64-QAM	5/6	48

Table 3.3: Definition of the transmission modes for 3 or 4 Tx

Data rate (Mbits/s)	Number of spatial streams (NS)	Modulation	Coding rate (R)	Number of data subcarriers (NSD)
12Mbps	2	BPSK	1/2	48
24Mbps	2	QPSK	1/2	48
36Mbps	2	QPSK	3/4	48
48Mbps	2	16-QAM	1/2	48
72Mbps	2	16-QAM	3/4	48
96Mbps	2	64-QAM	2/3	48
120Mbps	2	64-QAM	5/6	48
144Mbps	3	64-QAM	2/3	48
162Mbps	3	64-QAM	3/4	48
180Mbps	3	64-QAM	5/6	48

3.1.3 Linear Precoders for STBC by Exploiting Long-Term Channel Knowledge

As explained in Section 2.3.2.2.3, linear precoding can reduce the performance degradation due to transmit antenna correlation by forcing transmission on the nonzero eigenmodes of the transmit antenna correlation matrix [SP02]. The main advantage of this precoder is that it does not have to track fast fading but only the slowly varying antenna correlations. The latter can be fed back to the transmitter using a low-rate feedback link or can be based on uplink channel estimation that exploits channel reciprocity.

3.1.3.1 Overall Description of the Technique

We consider an M_T-transmit and M_R-receive antennas where the transmitter has some knowledge of the channel second order statistics. This information is available through the feedback bits sent by the mobile station, or can be based on uplink channel estimation. As depicted in Figure 2.2, the space-time block

encoder **X** maps the input data sequence x=(x_1,x_2,..x_Q) into a M_T×Q matrix **X** of codewords that are split onto M_T parallel sequences. These codewords are then transformed by an M_T×M_T linear transformation **F** in order to adapt the code to the available channel information. The resulting sequences, encompassed in an M_T×Q matrix **C**=**FX**, are sent on the M_T transmit antennas over Q adjacent OFDM subcarriers. The channel is assumed to be constant over the Q adjacent subcarriers.

Transmitted data are recovered at the receiver by means of a Maximum Likelihood (ML) receiver.

Each entry h_{ji} of the M_R×M_T channel matrix **H** represents the channel response between transmit antenna *i* and receive antenna *j* for the considered subcarrier. The received signal (at the mobile) is assumed to be a linear combination of several paths reflected from local scatterers, which result in uncorrelated fading across the receive antennas and therefore uncorrelated rows of matrix **H.** However, limited scattering at the base station (BS) can result in antenna correlation and therefore correlated columns of matrix **H**.

According to the correlation-based model the channel **H** can be written as follows:

$$\mathbf{H} = \mathbf{H}_{\mathbf{W}} \mathbf{R}_{\mathbf{T}}^{\mathbf{1/2}} , \tag{3.10}$$

where $\mathbf{H_W}$ is an M_R×M_T i.i.d complex matrix and $\mathbf{R_T}$ is the M_T×M_T transmit antenna correlation matrix.

The received signal **Y** is corrupted by additive white Gaussian noise denoted by the M_T×Q matrix **W** with covariance matrix $\sigma^2 \mathbf{I}_{\mathbf{M_R}}$:

$$\mathbf{Y} = \mathbf{HC} + \mathbf{W} = \mathbf{HFX} + \mathbf{W} . \tag{3.11}$$

3.1.3.2 Performance Criterion and Solution

In the following analysis, perfect knowledge of the channel correlation matrix ($\mathbf{R}_T$), is assumed at the transmitter.

The linear transformation **F** is determined so as to minimise a given criterion, such as *an upper bound on the pair-wise error probability* (PEP) of the codeword as considered in [SP02]. The PEP is defined as the error probability of choosing in favour of the codeword $\mathbf{X}^l$ instead of the actually transmitted codeword $\mathbf{X}^k$. The codeword $\mathbf{X}^l$ is given by:

$$\mathbf{X}^{l} = \arg\min_{\mathbf{X} \in Code} \left\| \mathbf{Y} - \mathrm{E}_s^{1/2} \mathbf{HF X} \right\|_F^2 , \tag{3.12}$$

with E_s the symbol energy.

Let $\mathbf{E}(k,l) = \mathbf{X}^k - \mathbf{X}^l$ be the code error matrix. We define Ω as the set of non-zero code error matrices $\Omega = \{k,l : \mathbf{E}(k,l) \neq 0\}$.

The *average* PEP denoted by $\overline{PEP}$ was upper-bounded in [TSC98] by:

$$\overline{PEP}(\mathbf{E}) \leq \left[\det\left(\mathbf{I} + \frac{E_s}{4\sigma^2} \mathbf{R}_T^{1/2} \mathbf{FEE}^H \mathbf{F}^H \mathbf{R}_T^{1/2H} \right) \right]^{-M_R} , \tag{3.13}$$

where E_s is the symbol energy.

The design of the precoder should minimise the average pair-wise error probability of the worst error code matrix: $\min_{\mathrm{L}} \max_{\mathrm{E}} \overline{PEP}(\mathbf{E})$.

In [SP02] the worst performance was approximated by $\max_{\mathrm{E}} \overline{PEP}(\mathbf{E}) \approx \overline{PEP}(\mathbf{E}_{\min})$, where $\mathbf{E}_{\min}$ is the code error matrix with the minimum determinant.

This approximation works well if the hermitian of $\mathbf{E}_{\min}$ ($\mathbf{E}_{\min}$ $\mathbf{E}_{\min}{}^{\mathrm{H}}$) is unique. If ($\mathbf{E}_{\min}$ $\mathbf{E}_{\min}{}^{\mathrm{H}}$) is non-unique, for example $\mathbf{E}_1$ and $\mathbf{E}_2$ have the same minimum determinant but ($\mathbf{E}_1$ $\mathbf{E}_1{}^{\mathrm{H}}$) and ($\mathbf{E}_2$ $\mathbf{E}_2{}^{\mathrm{H}}$) are different, minimising the averaged PEP of one of them doesn't ensure the performance of the other, and it can even degrades $\max_{\mathrm{E}} \overline{PEP}(\mathbf{E})$ rather then improving it.

In order to avoid this we introduce $\mathbf{\Sigma} = \sup(\Omega^-)$, the matrix supremum of Ω^-, where $\Omega^- = \{\mathbf{M} / \mathbf{M} \leq \mathbf{EE}^H, \forall \mathbf{E} \in \Omega\}$, and the inequality is in the matrix sense.

If $\mathbf{\Sigma}$ is non-zero, using that ($\mathbf{M}_1 \geq \mathbf{M}_2 \geq \mathbf{0} \Rightarrow \det(\mathbf{M}_1) \geq \det(\mathbf{M}_2)$ we show that

$$\min_{\mathbf{E}} \overline{PEP}(\mathbf{E}) \le \left[\det\left(\mathbf{I} + \frac{E_s}{4\sigma^2}\mathbf{R}_T^{1/2}\mathbf{F}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^H\mathbf{R}_T^{1/2H}\right)\right]^{-M_R}. \quad (3.14)$$

This reduces the design of the precoder to

$$\mathbf{F} = \arg\max_{\mathbf{F}} \det\left(\mathbf{I} + \frac{E_s}{4\sigma^2}\mathbf{R}_T^{1/2}\mathbf{F}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^H\mathbf{R}_T^{1/2H}\right). \quad (3.15)$$
$$s.t: Trace(\mathbf{F}\mathbf{F}^H) = P_0$$

Upon defining the singular value decomposition (SVD) of $\mathbf{R}_T^{1/2}$ and $\mathbf{E}\mathbf{E}^H$: $\mathbf{R}_T^{1/2} = \mathbf{U}_T\boldsymbol{\Lambda}_T^{1/2}\mathbf{V}_T{}^H$ and $\boldsymbol{\Sigma} = \mathbf{V}_s\boldsymbol{\Lambda}_s\mathbf{V}_s{}^H$, the solution of the optimisation problem [SP02] is waterfilling:

$$\mathbf{F} = \mathbf{V}_\mathrm{T}\boldsymbol{\Phi}_\mathbf{f}\mathbf{V}_s{}^H$$
$$\boldsymbol{\Phi}_\mathbf{f}^2 = \left(\gamma\mathbf{I} - \left(\frac{E_\mathrm{s}}{4\sigma^2}\right)^{-1}\boldsymbol{\Lambda}_T^{-1}\boldsymbol{\Lambda}_s^{-1}\right)_+, \quad (3.16)$$

where $\gamma > 0$ is a constant computed from the trace constraint and $(\cdot)_+$ stands for max(·,0). Without loss of generality it is assumed that the constraint on the transmitted power described in Equation (3.15), is such that the trace of $\mathbf{F}\mathbf{F}^H$ is equal to 1.

3.1.3.3 Linear Precoding for Orthogonal and Quasi-Orthogonal Space-Time Coded Systems with M_T Transmit Antennas

In many applications long-term CSI is available at the transmitter, e.g., via a low rate feedback channel. This knowledge can be used to improve the performance of space-time coding schemes by applying precoding techniques.

When orthogonal space-time codes ([Ala98] , [TJC99]) are considered, code error matrices are of the form $\mathbf{E}\mathbf{E}^H = d^2\mathbf{I}$, where $d^2 = \sum_{k=1}^{Q} |x_k - \tilde{x}_k|^2$, for $\tilde{x} = (\tilde{x}_1,...,\tilde{x}_Q)$ decided instead of the transmitted $x = (x_1,...,x_Q)$.

For this case $\boldsymbol{\Sigma} = d_{\min}^2.\mathbf{I}_4$, where $d_{\min}^2$ is the minimum distance of the constellation of x_1 (assuming all the component of x have the same constellation). The solution of optimisation problem is then

$$\mathbf{F} = \mathbf{V}_\mathbf{r}\boldsymbol{\Phi}_\mathbf{f}$$
$$\boldsymbol{\Phi}_\mathbf{f}^2 = \left(\gamma\mathbf{I} - \left(\frac{E_\mathrm{s}d_{\min}^2}{4\sigma^2}\right)^{-1}\boldsymbol{\Lambda}_\mathbf{r}^{-1}\right)_+. \quad (3.17)$$

On the other hand, quasi-orthogonal space-time block codes ([Jaf01], [SPC03]) were derived for 4 transmit antennas from the permutation of two Alamouti codes, with codeword:

$$\mathbf{Z} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ -x_2^* & x_1^* & -x_4^* & x_3^* \\ x_3 & x_4 & x_1 & x_2 \\ -x_4^* & x_3^* & -x_2^* & x_1^* \end{bmatrix} = \begin{bmatrix} A & B \\ B & A \end{bmatrix}. \quad (3.18)$$

This Quasi-orthogonal Space-Time Block Code, called also ABBA, provides full transmission rate to the detriment of orthogonality.

Improved quasi-orthogonal code have been recently proposed [PF01], achieving full diversity through constellation rotation (of x_3 and x_4) by an optimal angle selected to maximise the minimum distance code error matrix (Equation (3.13)). In this case:

$$\mathbf{E}\mathbf{E}^H = \begin{bmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ b & 0 & a & 0 \\ 0 & b & 0 & a \end{bmatrix}, \tag{3.19}$$

where $a = \sum_{i=1}^{4} \left| x_i - \tilde{x}_i \right|^2$ and $b = 2\,\mathrm{Re}\left\{ \left(x_1 - \tilde{x}_1\right)\left(x_3^* - \tilde{x}_3^*\right) + \left(x_2 - \tilde{x}_2\right)\left(x_4^* - \tilde{x}_4^*\right) \right\}$.

For all $\mathbf{E}$, the hermitian matrix $\mathbf{E}\mathbf{E}^H$ has the following eigenvector decomposition $\mathbf{E}\mathbf{E}^H = \mathbf{V}_e \mathbf{\Lambda}_e \mathbf{V}_e{}^H$, where $\mathbf{\Lambda_e} = diag\left(a+b, a+b, a-b, a-b\right)$ and $\mathbf{V_e}$ is square unitary Walsh Hadamard matrix of size 4.

The study of this case shows that $\mathbf{\Sigma} = d_{\min}^2 . \mathbf{I}_4$, where $d_{\min}^2$ is now the minimum distance of the constellation ($x_1 + x_3$), sum of the non-rotated and rotated constellation [SPC03].

The solution for the precoder is again

$$\mathbf{F} = \mathbf{V_r}\mathbf{\Phi_f}$$
$$\mathbf{\Phi_f^2} = \left(\gamma \mathbf{I} - \left(\frac{E_s d_{\min}^2}{4\sigma^2} \right)^{-1} \mathbf{\Lambda_r^{-1}} \right)_+ . \tag{3.20}$$

Depending on the antenna correlation, the linear precoder allows to distribute, in an intelligent fashion, the power over the different eigenmodes of the transmit covariance. However two extreme cases can be observed:

i) When the antenna correlation is zero, the matrix $\mathbf{R}_T$ is multiple of the identity and the precoder is the identity matrix without any action on the space-time block coding,

ii) When the antenna correlation is one, only one eigenvalue of matrix $\mathbf{R}_T$ is non-zero. **F** is then of rank one and the precoder is equivalent to a beamformer.

3.1.3.4 Conclusions

The linear precoding technique allows the adaptation of existing space-time coding schemes to different states of long-term channel knowledge. This knowledge is in the form of the transmit covariance and can be obtained via low-rate feedback link or by uplink channel estimation that exploits channel reciprocity. Furthermore, long-term channel knowledge contains only information on the antenna array response and ignores fast fading. Therefore, the feedback can even be reduced, as the transmit covariance is the same for all subcarriers. For orthogonal and quasi-orthogonal space-time codes, the linear precoding allows a soft adaptation of the link between two extreme cases: pure space-time block coding and beamforming.

3.1.4 Limited CSI Spatial Multiplexing; Per Antenna Rate Control and Per Stream Rate Control

3.1.4.1 Overall Description of the Technique

The technique described here employs multiple antennas to achieve spatial multiplexing gains by transmitting different data streams via each available antenna. Restrictions upon the available signalling bandwidth favour the use of schemes requiring only limited channel state information to be fed back to the transmitter. In this technique only simple indications of channel quality are signalled to the transmitter. These enable the transmitter to select a suitable modulation and coding combination, and additionally determine transmit power levels, antenna selection and scheduling decisions. This represents a variant of Per-Antenna Rate Control (PARC) that has been moulded to the WINNER system assumptions.

For background information, PARC has been proposed within 3GPP [Luc02], as has a variant named Selective-PARC (S-PARC) [Eri04], which permits transmit power savings by turning off some antennas. As described in [Luc02], the receiver applies successive interference cancellation, using MMSE detection of each antenna-stream in order of SINR. The technique has been shown to be capable of approaching the theoretical capacity limit [CLH01], [VG97], which means that PARC in principle can utilise the available diversity and multiplexing gain offered by the channel given that appropriate rate adaptation is used.

For the MMSE-SIC receiver, it is currently believed that it is preferable to do the cancellation after channel decoding. By appropriately tuning the rate, a desired target error rate can be achieved, which in turn means that the error propagation in the SIC can be controlled. For this purpose, it may be advantageous to do channel encoding over all available chunks. This procedure is integrated in the general transmit scheme outlined in the previous chapter by segmentation of a user's data flow into several subflows as illustrated in Figure 3.8. Each subflow is encoded with a separate channel code, segmented into substreams which are then mapped onto particular spatial layers in particular chunks (chunk layers). The number of layers in each chunk is typically equal to the number of transmit antennas, but may be less if antenna selection is employed to conserve transmit power at antennas experiencing deep fades and to realise some transmit directivity.

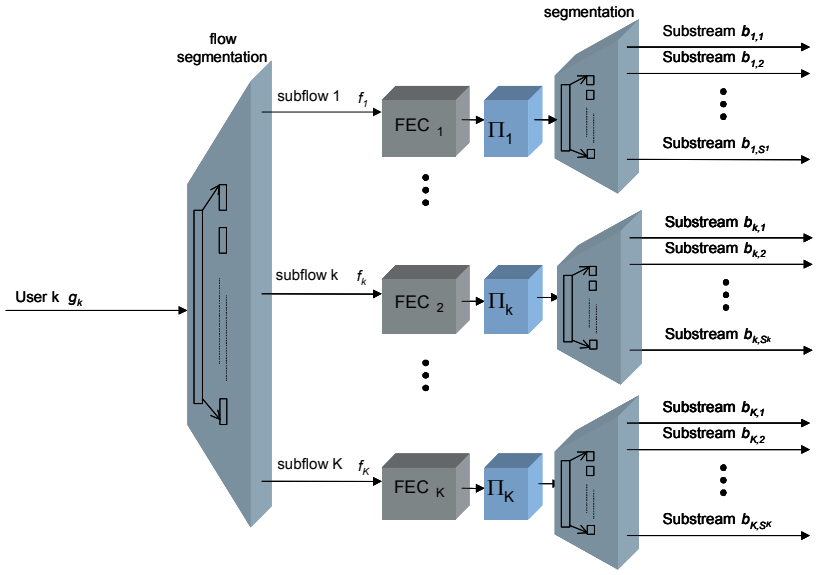


Figure 3.8: Segmentation of a user's data stream into several subflows

Space-time-frequency coding and modulation (STFCM) is then applied to each chunk layer as according to Figure 2.11. As a baseline, no linear dispersion code, no non-linear precoding and no beamforming is applied and this means that each chunk layer corresponds to a virtual antenna, which in turn corresponds to a physical antenna. In essence each channel encoded flow is mapped directly to one of the physical transmit antennas typically using the same modulation and coding for all subcarriers. The CQI provided by return signalling is used to determine appropriate modulation and power level for each subflow as well as the code rate used by the channel codes. Further, CQI captures the receiver's capabilities, and PARC may thus work with both purely linear receivers, such as an MMSE detector, as well as more advanced receivers such as an MMSE concatenated with a SIC (MMSE-SIC).

The goal of the technique described here is to optimise throughput (for a given transmit power) via spatial multiplexing, rate control, and scheduling. The technique facilitates channel dependent scheduling by employing higher order modulation and high coding rates when users' propagation conditions permit.

Although the technique could also be employed for SDMA by transmitting different antenna streams to different users simultaneously, this is not covered within this deliverable, but may be studied later. The main difference in this case is that with PARC, a number of different (sub) flows are transmitted from the antennas directly instead of from beams, and that several flows are allocated to a single user.

3.1.4.2 Required Support Functions

In general, the receiver must make estimates of the SINR for each chunk of each antenna stream in order to determine the appropriate modulation and coding choices back at the transmitter. The SINR estimates should capture the capabilities of the receiver and may be condensed into one CQI indicating preferred modulation and coding rate for each subflow, i.e., typically each transmit antenna. To further improve performance it is foreseen that modulation and power levels may be adjusted on a chunk basis in order to

improve the performance at the cost of higher signalling overhead.

Additionally, the receiver needs accurate estimates of the channel response between each antenna pair for all subcarriers (or chunks) in order to estimate the spatially multiplexed streams. This requires pilot transmissions for each chunk from each transmit antenna.

Forward signalling is required to instruct the receiver regarding power allocations to different antenna chunks, and/or of unused transmit antennas (if antenna selection is employed) as well as code rate.

Finally, it is currently believed that SIC after channel decoding is the preferred demodulation method, and for this purpose the use of several rate controlled channel codes is perhaps the most demanding requirement. This can potentially move the complexity from the receiver detector to the link adaptation and its associated signalling overhead.

3.1.4.3 Applicability and Practical Relevance

Limited CSI Spatial Multiplexing is applicable to high mobility devices (i.e., wide-area scenario) where signalling bandwidth constraints limit the return channel to just CQI measurements. Further, highly mobile devices are often necessarily small and portable, and hence lower terminal classes must be supported. The technique may also be applied in short-range scenarios, e.g., for devices typically operating in wide-area, but temporarily in a more stationary short-range environment.

The PARC technique has been studied within 3GPP and therefore the basic approach may be considered ‘mature’, but there are new aspects in adapting it to the WINNER system specification, e.g., the use of OFDM, and applying ‘per-chunk’ adaptation in addition to per-antenna control.

3.1.4.4 Scalability and Relation to Other Concepts

As previously described, the Limited CSI Spatial Multiplexing technique fits within the generic concept described in Section 2. In essence, each data stream is multiplexed into several subflows, which are separately channel encoded, similar to the case when several different users are served. Thus, on each sub flow it appears straightforward to apply some space-time coding (via LDC encoding, Figure 2.11) as well as a ‘beamforming’ mapping (Figure 2.10) to transmit the virtual antenna chunks via multiple actual antennas for enhanced directivity or diversity. This is referred to as per stream rate control (PSRC). Further, in the absence of CQI for the individual antenna streams at the transmitter, the technique would employ the same transmission rate on each stream and hence reduce to a BLAST-like scheme.

As mentioned above, PARC may be viewed as an alternative to linear dispersion codes when it comes to exploiting the so-called open-loop multiplexing and diversity gains offered by the channel. A major difference is probably the demodulators employed as well as the signalling overhead and use of several channel codes. Whereas PARC originally uses rate control together with SIC after channel decoding, demodulation of linear dispersion codes may require (iterative) non-linear (approximate) maximum likelihood demodulators. On the other hand, linear receiver architecture may also be appropriate depending on the operating point.

3.1.4.5 Summary

The proposed concept achieves throughput gains (enhanced spectral efficiency) due to spatial multiplexing and it is scalable towards other multi-antenna techniques. At the transmitter only CQI measurements are required. If detailed CSI is available (e.g., low mobility TDD device), the scheme does not make full use of it unless adapted towards other techniques better suited to such a scenario. It should also be stressed that the base line PARC concept can be generalised to Per Stream Rate Control (PSRC), where instead of transmitting each code block directly from an antenna, a linear dispersion code and a beamforming component may additionally be applied.

3.2 Linear Techniques Using Multi-User Optimisation

This section provides a deeper insight in the optimisation of multi-user mobile communication systems. The choice of the multi-antenna/multi-user signal processing algorithm depends on various factors and constraints of the system design, such as the scenario or the allowable computational complexity. From the computational complexity point of view the linear techniques are superior to the non-linear techniques

described in Section 3.3, whereas the non-linear techniques usually show a slightly better performance, e.g., in terms of throughput. Another important factor to choose the algorithm is the available amount of CSI at the transmitter. This strongly depends on the system design, e.g., user mobility or duplexing mode. In case of good CSI at the transmitter, beamforming schemes as described, for example, in Section 3.2.1 are used, whereas for limited or no CSI at the transmitter, techniques based on ST codes or higher order statistics as, for example, shown in Section 3.2.7 are applied.

3.2.1 Overall Description and Comparison of Beamforming Techniques

In most propagation conditions, energy reaches the receiving terminal only via a subsection of the solid angle in space. All energy transmitted in other directions is lost for the receiver and even creates harmful interference to other terminals. The concept of beamforming therefore is the idea to concentrate the transmitted energy in relevant areas or equivalently to receive energy from preferred directions. The associated gain is called beamforming or array gain and results in an increase of the SINR of the corresponding link or equivalently in a shift of the $BER(E_b/N_0)$ curve. Right from these initial considerations it is evident that beamforming is a means to efficiently improve the interference conditions by actively controlling or avoiding interference. Additionally, the delay and Doppler spread decrease with increasing directivity of the transmission and the corresponding frequency and time variances are reduced.

Due to the selectivity in space, beamforming and space division multiple access (SDMA) are closely interrelated and therefore considered jointly in this document. In particular, any kind of beamforming can be thought of as an extension (in form of an overlay or replacement technique) of the well-known sectorisation principle in cellular networks – enhanced by flexibility and adaptivity of the beams.

Beamforming requires high correlation of the signals or equivalently relatively low angular spread and few scatterers. Typical scenarios include outdoor transmitters or receivers located above rooftop in typical urban macro cell or in rural environments. Moreover close antenna element spacing will increase the correlation of the signals and are therefore preferred for beamforming.

Different beamforming techniques can be classified according to a number of criteria, including:

- ***Adaptivity***
  - In the *resource assignment algorithm*, e.g.,
    - Scheduling algorithm, beam-to-user mapping,
    - Number of beams per user,
    - Granularity of resource assignment in time and frequency,
  - In the *beamforming algorithm*, e.g.,
    - Beam direction,
    - Interference reduction (null steering, etc.),
    - Beam shape (tapering),
  - In the *physical transmission mode*, e.g.,
    - Transmit and/or receive beamforming,
    - Code rate,
    - Modulation level,
    - Substream power,
    - Presence of additional spatial processing components, like, SDMA, or spatial multiplexing, space or polarisation diversity.
- ***Available channel information***
  - *Type of information*,
    - Explicit feedback, e.g.,
      - CQI, like SINR,
      - Beam selection information,
      - Beamforming vector,
      - Covariance matrix,
    - Based on measurements at the transmitter, e.g.,
      - Covariance matrix (in case of FDD, a proper frequency transformation is necessary).
  - *Granularity of information*
    - Time,

- Static, long-term, short-term,
- Frequency,
  - Per subcarrier, per chunk, per resource allocation,
- Space,
  - All spatial substreams, *n* best streams.

These classification criteria will be used to distinguish the investigated beamforming techniques. Additionally, different beamforming techniques can have different primary *objectives*, e.g., link improvement (e.g., coverage/range extension, user throughput increase) or solution of the multi-user problem (increase of spectral efficiency, cell throughput, interference control/avoidance). Depending on the degree of adaptivity, beamforming can either be implemented at RF or at the baseband. Furthermore, each beamforming technique can be deployed and parameterised in many ways, e.g., with respect to the antenna configuration, and the number of beams per cell or site.

An overview of parameters and enablers of beamforming techniques is shown in Figure 3.9. In the most general case, the beamforming and beam assignment includes an SDMA component and is different for each frequency/time chunk. Also, the number of beams per user and chunk $Q(c)$ is not restricted to one, and therefore spatial multiplexing and space (and/or polarisation) diversity components can be included. Apart from information extracted from the past, the beamforming may rely on information either obtained from measurements of the return channel at the transmitter (with frequency transformations if necessary), or from explicit feedback. Table 3.4 provides a comparison of the beamforming techniques investigated in this document with respect to these classification criteria. In the area of fixed and adaptive beams, three different variants are proposed by different partners. These variants have, to a large extent, identical degrees of adaptivity and requirements for channel state information.

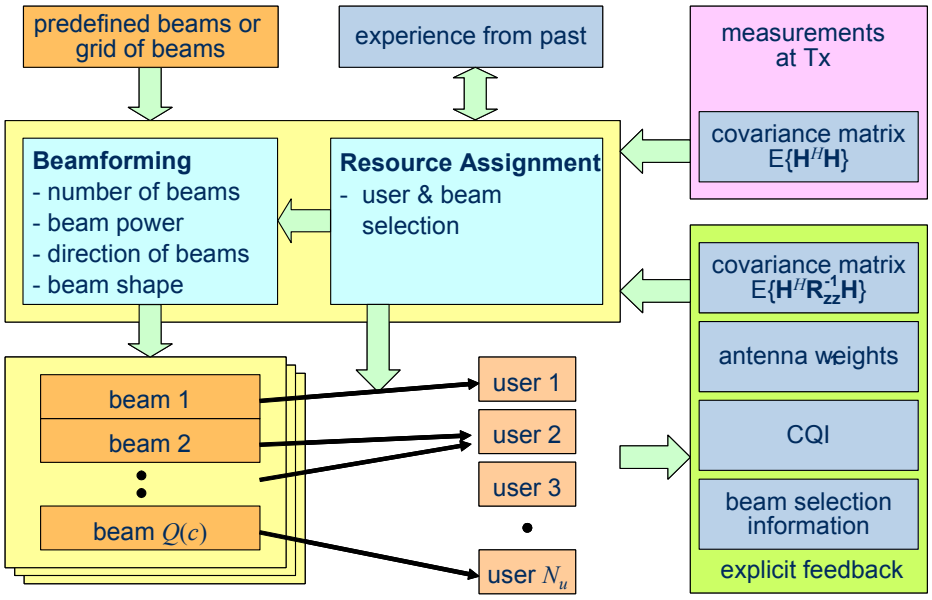


Figure 3.9: Basic enablers and parameter of beamforming techniques

General advantages of beamforming techniques include the support of low-cost single-antenna terminals the possibility of range and coverage extension, and allowing the usage of simple mobile receivers. Beamforming in general is well understood and has proven its capability and robustness in many deployments already. It is a technique that can be used in single-carrier and multi-carrier systems and any kind of multiple access scheme. Due to the high directivity, the interference conditions become very location dependent. In case of adaptivity, special precautions are required to ensure predictability and consistency of channel measurements and corresponding resource assignment and beamforming. Furthermore, if digital beamforming is performed (i.e., baseband beamforming), many techniques require phase coherency throughout the whole processing chain until the radiated signal from the antennas. Therefore calibration is required to remove, e.g., the impacts of varying signal path length, temperature drift and coupling.

Table 3.4: Comparison of beamforming techniques

Tx characteristics	Grids of fixed beams + SDMA	Opportunistic beams + SDMA	Adaptive beams + SDMA SMMSE
number of beams	fixed or adaptive; short/long-term adaptation based on, e.g., instantaneous/average traffic load		
beam direction and shape	fixed; typically pointing at different directions	random or semi-random	Fully adaptive based on short-term CSI $\mathbf{H}$, $\mathbf{H}^H\mathbf{H}$ or long-term CSI $E\{\mathbf{H}^H\mathbf{H}\}$, also trying to avoid/minimise interference to other users
beam power	fixed or adaptive based on short-term or average SINR		adaptive based on CSI, see above
user → beam assignment	adaptive based on short- or long-term CQI		fixed or adaptive (selecting groups of users to be served simultaneously on same resources)
number of beams per user	typically 1 in wide-area scenario		typically 1 in wide-area scenario but adaptive
constraints	requires spatial separability of beams to be efficient	requires large number of users to be efficient	sensitive to CSI precision

Within this document several beamforming techniques and combinations of beamforming with further spatial processing techniques are investigated. Figure 3.10 shows a block diagram that concretises the generic block diagram of Chapter 2 (see Figure 2.10) to an extent that it still accommodates all beamforming techniques proposed and investigated in T2.5. In particular, it can be seen that a combination of beamforming with precoding techniques will not be considered initially[2]. Therefore, the virtual antenna stream combiner is a simple adder, and the non-linear precoding (NLP) blocks are omitted.

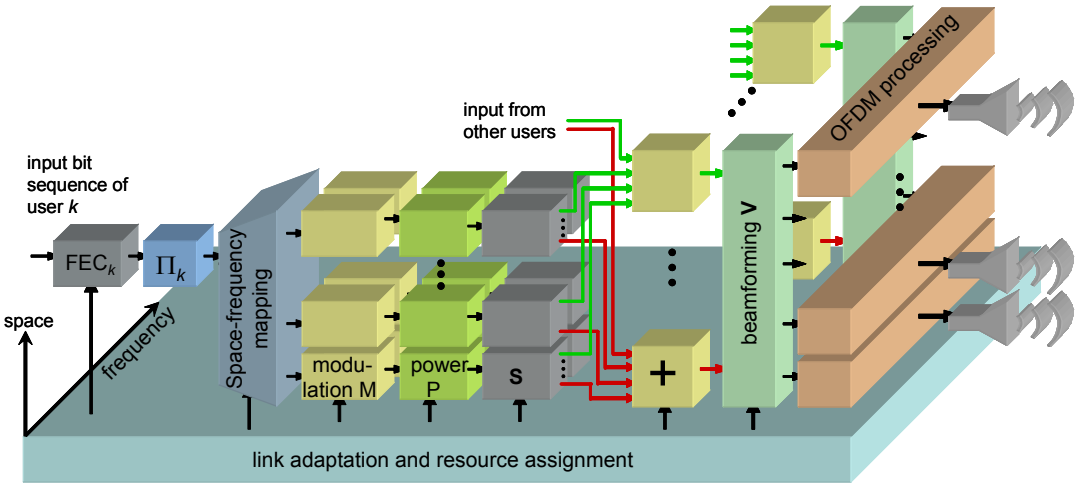


Figure 3.10: Generic block diagram of considered beamforming techniques

[2] Linear precoding techniques can actually be understood as multi-user beamforming themselves.

According to Figure 3.10, the beamforming techniques can be described with the following signal model for each user and chunk:

$$\underset{M_R \times T}{\underbrace{\mathbf{Y}}} = \underset{M_R \times M_T}{\underbrace{\mathbf{H}}} \; \underset{M_T \times Q}{\underbrace{\mathbf{V}}} \; \underset{Q \times Q}{\underbrace{\mathbf{P}}} \; \underset{Q \times T}{\underbrace{\mathbf{S}}} + \underset{M_R \times T}{\underbrace{\mathbf{Z}}} \;, \tag{3.21}$$

where M_T and M_R are the numbers of transmit and receive antennas, respectively, $Q(c)$ is the number of spatial layers of chunk c, T the number of symbol periods over which space-time coding is applied. **H** denotes the channel matrix. **V** is the generalised beamforming matrix containing the beamforming weights. Note that **V** can also be interpreted as a linear precoding matrix. For techniques that operate in the antenna domain **V** is the identity matrix. **P** is a diagonal matrix containing the square roots of the power allocation to the different spatial substreams. **S** is the linear space-time mapping of the N_c coded symbols onto $q \le Q(c)$ spatial substreams and T symbol periods. For matrix modulation (or LDC) in general $T > 1$ and $q > 1$ holds, further investigations will focus on schemes for which $1 < T \le n_{frame}$, i.e., the space-time coding will be confined to one OFDM frame. Note that for mathematical clarity in the model it is assumed that the space-time mapping matrix **S** spans the dimension $Q(c)$ in space. If a particular user only performs encoding over a subset of spatial streams $q < Q(c)$, the appropriate entries in **S** are filled with zeros. The restriction of $T = 1$, $Q(c) > 1$ implements the class of the so-called vector modulation schemes. $T = 1$, $Q(c) = 1$ would result in the well-known scalar modulation without spatial processing. Beamforming techniques without space-time mapping can be achieved by a proper configuration of **S**. For the sake of clear presentation, Figure 3.10 does not show the individual subcarriers and therefore the time frequency mapping function (cf. Figure 2.11) has been omitted in Figure 3.10 but is of course an integral part of the processing and performed immediately after space-time mapping **S**.

3.2.2 Beamforming and SDMA Based on a Grid of Fixed Beams

3.2.2.1 Overall Description of the Technique

Fixed beams can be seen as an extension of the cellular concept. Each transmitting user selects his beamforming vector **v** from a predefined static set of antenna weights, so-called beams. The number of spatial layers Q is here reduced to one, the lack of space-time coding reduces T to one. Therefore, the receive vector of each transmitting user and subcarrier has the following form,

$$\underset{M_R \times 1}{\underbrace{\mathbf{y}}} = \underset{M_R \times M_T}{\underbrace{\mathbf{H}}} \; \underset{M_T \times 1}{\underbrace{\mathbf{v}}} \; \underset{1 \times 1}{\underbrace{p}} \; \underset{1 \times 1}{\underbrace{s}} + \underset{M_R \times 1}{\underbrace{\mathbf{z}}} \;, \tag{3.22}$$

where $\mathbf{v} \in \{\mathbf{v}_1 \quad \mathbf{v}_2 \quad \ldots \quad \mathbf{v}_N\}$, and N denotes the number of available beams.

In the simplest case, the selection of beams is done for all subcarriers and thus requires a minimal feedback rate. Figure 3.11 below shows a possible design example of these beams. Chebychev tapering is used to design the shape of the beams. Here a sidelobe suppression of 21 dB is selected. The beam spacing is non-linear in order to have equal beam crossing levels.

To increase the throughput in this proposed technique a spatial reuse of beams is done based on the beam selection information of each user (or more advanced: CQI). As the overlap of two neighbouring beams is rather large they cannot transmit simultaneously reusing the same time-frequency chunks for different users. With the limited angular spread of a wide-area scenario an SDMA-reuse of resources of non-neighboured beams is possible.

Figure 3.12 shows an exemplary space-frequency spectrum snapshot of the urban macro scenario. The y-axis gives the derivation in degrees from the centre of the example beam, the x-axis shows subcarrier number 1 to 1024 for 20 MHz bandwidth and the colour temperature gives the Rx power depending on beam direction and subcarrier number.

A simple heuristic rule can be given: In non-neighboured beams the time-frequency resources can be reused; neighboured beams must be separated via the multiple access scheme, e.g., in time for SDMA/TDMA. This gives the space-time resource allocation pattern of Figure 3.13 for 8 beams with 8 Tx antennas.

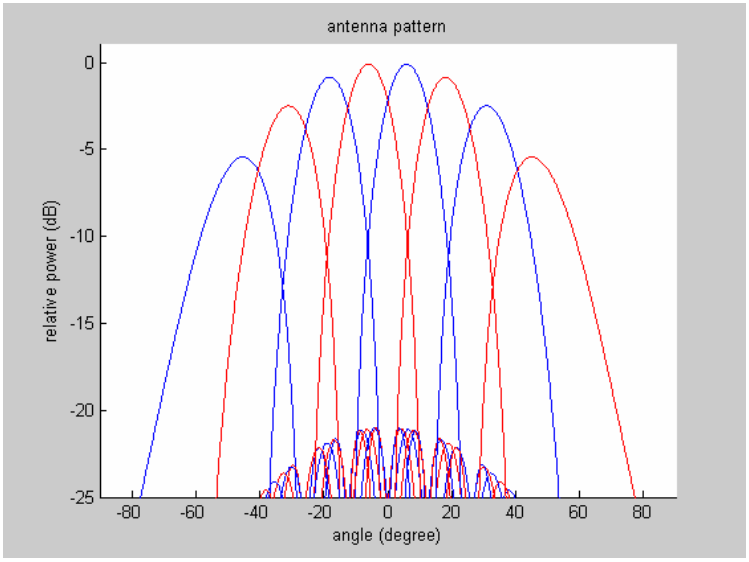


Figure 3.11: A grid of 8 beams for a 120° sector of an 8-element ULA with 70° HPBW

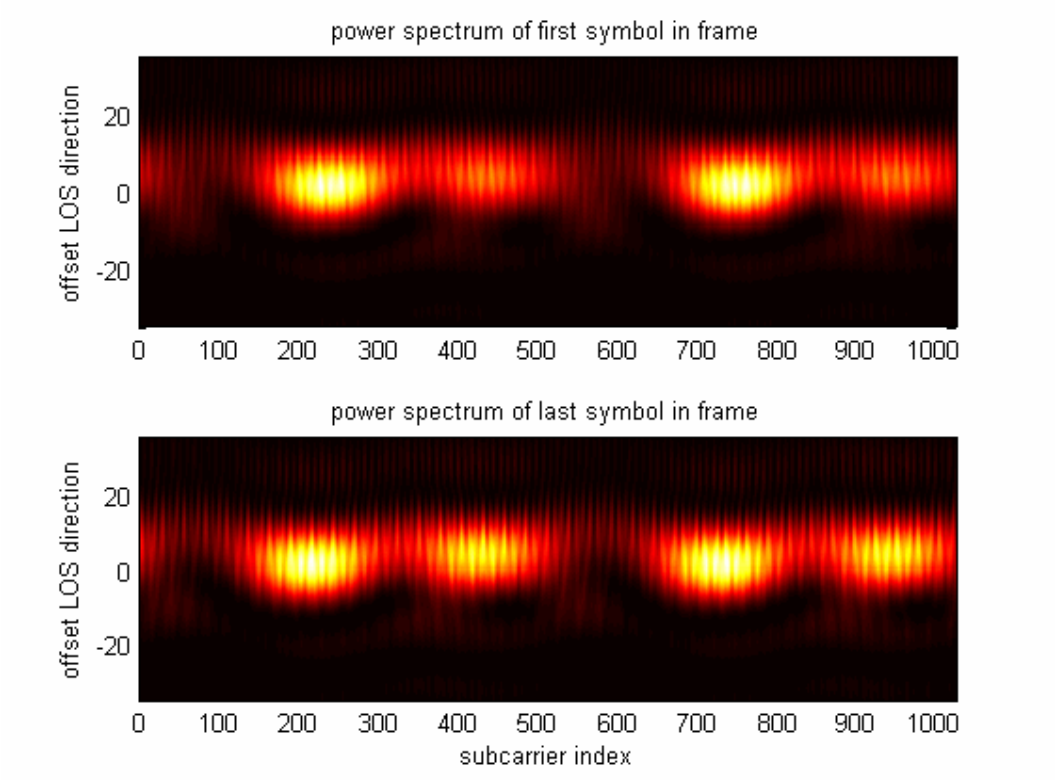


Figure 3.12: 3GPP SCM space-frequency snapshot of urban macro scenario with 8° angular spread

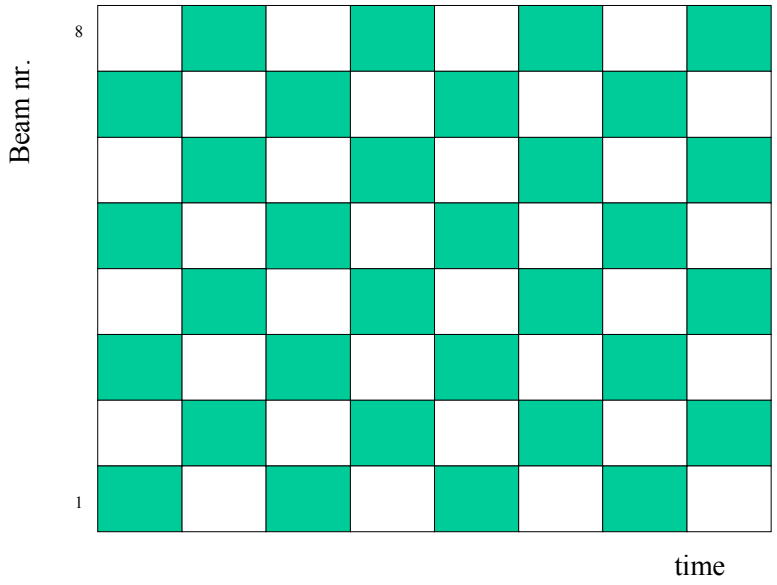


Figure 3.13: Beam/Time Resource Allocation Pattern for SDMA/TDMA with 8 antennas and 8 beams

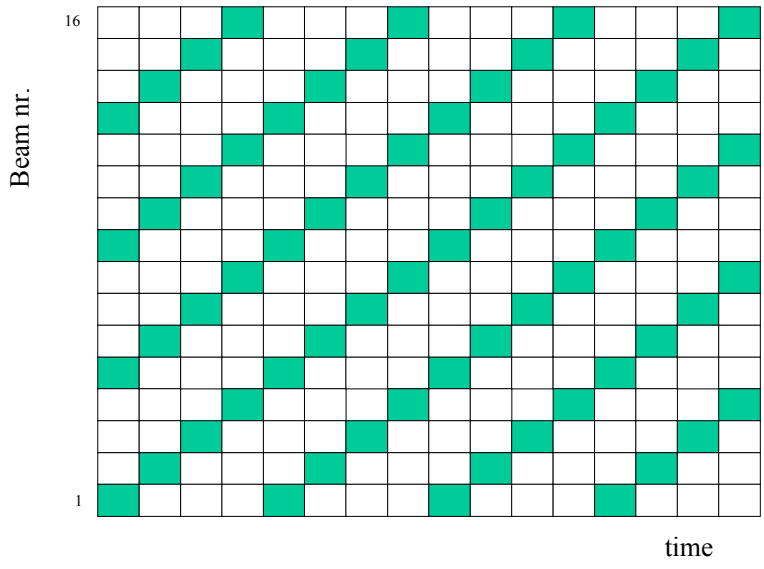


Figure 3.14: Beam/Time Resource Allocation Pattern for SDMA/TDMA with 8 antennas and 16 beams

A further possible refinement for 8 antennas would be to use 16 beams with the same shape as the 8 beams above but with increased overlap. For simultaneous transmission again 4 of them can be selected for reusing the same resources (e.g., transmitting at the same time slot). Now the granularity of spatial coverage is increased, hereby increasing the chances of maximum transmission power for a certain user direction.

SDMA on top of fixed beams with the above resource allocation schemes has thus the potential to increase the throughput of one sector by factor $M_T/2$ when there are enough spatially distributed users available.

With the help of coding and interleaving it is possible to use all subcarriers for one user, resulting in minimal feedback signalling and maximum system simplicity. But obvious from Figure 3.12 is the fact that it is rather suboptimal, because some regions of subcarriers of one user do have bad channel conditions. Thus extra benefits can be obtained from fast scheduling on chunk level.

Theoretically fixed beams could also be created on RF level. For WINNER to keep the amount of required flexibility (e.g., switching of different spatial processing modes) it is suggested to apply the fixed beam weights in the baseband.

3.2.2.2 Required Support Functions

No support function, but an important prerequisite is the need of a certain spatial distribution of users to enable SDMA. User positions spread in space are preferable to clustered users. Large numbers of users per cell give of course increasing chances to schedule spatially separated users in parallel.

This proposed technique needs antenna array calibration, as most beamforming techniques do. Even when using the same oscillator frequency for the power amplifiers, their PLLs cause different phase shifts. Additionally, differences in cable length, temperature and mutual coupling will affect the phases of the array elements. Phase offsets have to be compensated to keep up proper directivity.

Another mandatory support function is the beam selection information. This has to be derived in the base station on a long-term time scale. It can either be based on feedback from the mobiles (for example, signalling the average Rx power of each received beam, or just the strongest beam, or the average Rx power of the two strongest beams). Alternatively also measurements from the uplink long-term CSI could be used to select the beams for each user.

An important support function is the pilot structure. Figure 3.15 shows a possible solution, an example for 8 fixed beams.

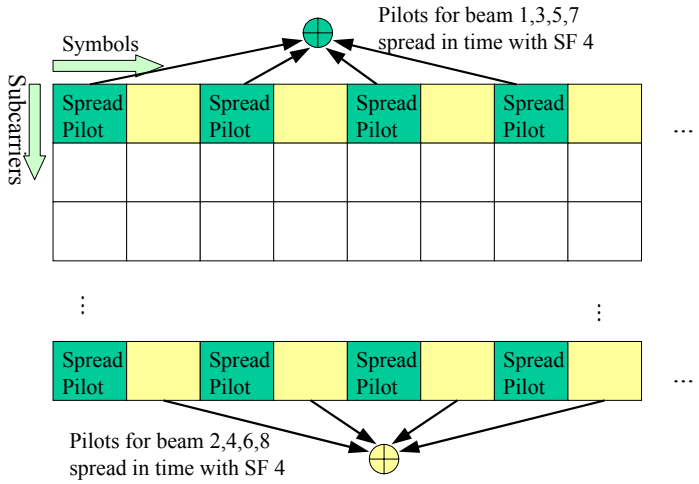


Figure 3.15: Pilot structure example for fixed beams with SDMA/TDMA using pilots spread in time

Four pilot symbols are put on each second OFDM symbol, where transmission for one half of all beams occurs. The channel in time direction typically does not have large variations in this interval, thus spreading pilot symbols in time direction gives a good basis for orthogonality. Each beam that uses this grid (either all even numbered beams or all odd numbered beams) now gets its own orthogonal spreading code chosen from the Walsh-Hadamard codes. An additional benefit besides separation of beams is that the SNR of the despreaded pilots is increased by the spreading gain of 6 dB, which improves the reliability of the channel estimation. This row of 8 successive spread pilots can now be repeated in time and frequency as needed and, e.g., also be shifted diagonally.

Optionally, a CQI short-term feedback per beam and (super-)chunk would allow fast scheduling but

would also create more control overhead. The CQI per beam could be derived from the received pilot power per beam of the above pilot pattern.

3.2.3 Beamforming and SDMA Based on Adaptive Beams

3.2.3.1 Overall Description of the Technique

The achievable SINR of SDMA is limited by the amount of mutual cross-talk. Adaptive beams give more degrees of freedom than fixed beams for reducing interbeam interference and thus have the potential of a larger throughput by allowing, e.g., more users for simultaneous transmission, a larger code rate or higher modulation order.

A joint optimisation of beamforming weights is a very promising approach. Possible optimisation criteria are SIR balancing algorithms, power minimising algorithms or greedy algorithms based, e.g., on minimum induced interference.

Another large potential for SDMA with adaptive beams is the possibility of combining fast spatial scheduling with the design of antenna weights, which can be considered as cross-layer optimisation.

Figure 3.16 shows the block diagram of the principle approach for the proposed technique.

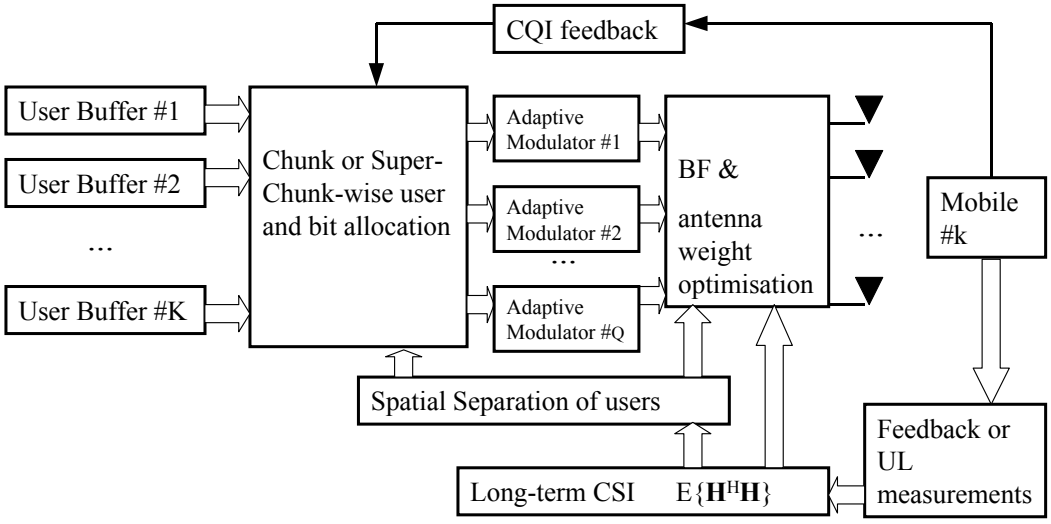


Figure 3.16: SDMA with Adaptive Beamforming and Scheduling

Long-term CSI via feedback or UL measurements is needed at the transmitter and will be used for deciding on the spatial separation between users. A pre-selection of users intended for simultaneous SDMA transmission is based on their spatial separability and the current short-term channel quality, which has to be fed back from the mobile. (Additional scheduler constraints like fairness can be taken into account). Depending on the CQI, the modulation order can be adapted for each transmitting user. (The code rate is fixed for the whole frame due to FEC). Then the weights can be optimised, depending on the chosen criterion, again using long-term CSI.

3.2.3.2 Required Support Functions

As for fixed beams an important prerequisite is the need of a certain spatial distribution of users to enable SDMA. User positions spread in space are preferable to clustered users. Large numbers of users per cell give of course increasing chances to schedule spatially separated ones of them in parallel.

This proposed technique needs antenna array calibration, as most beamforming techniques do. Even when using the same oscillator frequency for the power amplifiers their PLLs cause different phase shifts. Additionally differences in cable length, temperature, and mutual coupling will effect the phases of the array elements. Phase offsets have to be compensated to keep up proper directivity.

Long-term CSI information $\mathrm{E}\left\{\mathbf{H}^{\mathrm{H}}\mathbf{H}\right\}$ is required at the transmitter. It can be obtained via feedback or via uplink covariance matrices with the help of a proper frequency transformation.

3.2.4 Opportunistic Clustered OFDM with Partial Feedback

3.2.4.1 Overall Description of the Technique

Multi-user diversity systems rely on the time-varying fading of the users and scheduling of users that are in favourable instantaneous fading conditions. For time-dispersive channels, OFDM is a suitable technique to exploit multi-user diversity even further. In the downlink, it is feasible to schedule different users on different frequency sub-carriers (OFDMA). This enables the scheduling of users also on their frequency fading peaks, in addition to the temporal fading peaks.

The multi-user diversity scheduling requires short-term CSI from all active users. The feedback of the complete short-term CSI from all users before each scheduling decision can create a very high feedback overhead. One way to reduce the feedback overhead is to use the correlation between adjacent sub-carriers and to remove feedback information that will not likely be used by the scheduler:

- Group adjacent sub-carriers into clusters and feedback only the supportable rate of the weakest sub-carrier within the cluster.
- Feed back only the supportable rate of the strongest clusters.

The cluster-size and the number of rates fed back per user are important design parameters. If the cluster-size is small, the users feed back information about only a few clusters and there are few users in the system, there is a high risk that only a small part of the available downlink frequency band will be allocated. This can be avoided by introducing an adaptive feedback rate. The larger the number of users in the system, the smaller the required feedback from each user is. This can be motivated if there is a relatively high correlation between the number of active users in the downlink and the load on the uplink.

To obtain both multi-user diversity gain and some fairness in terms of delay, users must be fading at fast enough rates. Slowly moving users may experience too long time between fading peaks. Applying opportunistic beamforming can solve this problem. By using several transmit antennas at the base-station; random or semi-random beams can be formed for each scheduling block and changed between blocks. This will change the fading of each user between each scheduling block, enabling fairer scheduling without compromising multi-user diversity gain. A suitable opportunistic beamforming structure would be to use the same beamforming vector across all sub-carriers within a cluster, but independent beamforming vectors in different clusters. This will decrease the sub-carrier correlation between the clusters of a particular user, and will increase the frequency fading, especially for users experiencing relatively flat channels.

The original opportunistic beamforming idea uses no CSI to form the beams [VTL02]. For opportunistic OFDM, where several beams are formed simultaneously (one for each cluster), the idea of using the fed back short-term CSI for beamforming purposes is more attractive. From the reduced short-term CSI feedback, the base station knows which beams were the best for each user. If the temporal fading rate of a user is moderate, this knowledge can be used to form the beam during the next scheduling block. Several strategies are possible. If cell throughput is in focus, the beams that resulted in highest fed back supportable rates from the users could be kept for the next scheduling block. On the other hand, if fairness is in focus, the strongest beams of the most unfairly treated users can be kept. This “semi-random” beamforming based on some partial CSI is still an open issue.

Opportunistic beamforming also gives an effect of interference suppression for free, called opportunistic nulling. Just as for the desired channel, the sum of the interfering channels will be stronger in some clusters and weaker in others. Since the users base their supportable rates for each cluster on SINR measurements, clusters with high interference level will not likely be fed back by the user. This, however, requires synchronisation between base-stations so that all base-stations change beamforming weights and transmit orthogonal training signals simultaneously.

For multi-user diversity systems, the scheduling algorithm is important. The proportional fair scheduler is a well-known approach to combine multi-user diversity gains with some degree of fairness. One problem with the standard proportional fair scheduler is that it treats each user identically, even though users may

have significantly different quality-of-service (QoS) requirements. The proportional fair scheduler can be extended to incorporate both bit rate and delay requirements by introducing the scheduling criterion

$$user_q = \arg\max_k \frac{C_{q,k}}{\left(\frac{A_k}{T_k R_k} + \varepsilon\right)^{\kappa}}, \tag{3.23}$$

where $user_q$ is the scheduled user at cluster q, $C_{q,k}$ is the fed back supportable rate of cluster q and user k, A_k is the achieved throughput in the user-specific historical time window T_k, R_k is the target bit rate of user k, ε is a regularisation term and κ is a fairness parameter. This scheduler differs from the proportional fair in that it lets users have different target bit rates, R_k, and delay requirements, T_k. Furthermore, it is possible to tune the fairness level of the scheduler by adjusting the parameter κ. A very high κ will give a scheduler with Round Robin behaviour, whereas a κ close to zero will give the maximum throughput scheduler.

Semi-analytical simulations have been done to test the performance of the proposed system. Simulation results and assumptions can be found in [SWC04], [SWO04].

A short summary of the proposed scheme:

- A multi-user diversity scheme based on OFDM,
- Highly reduced feedback from each user,
- Opportunistic beamforming with different beams in different frequency bands (clusters) to increase temporal and frequency fading (only one beam however),
- Modified proportional fair scheduling with user-specific bit-rate and delay requirements,
- Semi-random beamforming by utilising the fed back CSI to the scheduler.

3.2.4.2 Required Support Functions

The terminals regularly have to estimate the SINR on all sub-carriers based on a short training signal. Fast CSI feedback from users to base-stations is required. In the forward link, a user identifier has to be inserted for each cluster and scheduling block to notify the users of the scheduling decision.

3.2.4.3 Additional Assumptions and Simplifications for Simulations

In order to make intercell interference predictable for the users, the base stations have to be synchronised, i.e., they change beamformers, transmit training and data at the same time. For the multi-user diversity scheme to work well, the coherence time of the channel has to be larger than the length of a scheduling block, i.e., the time between two consecutive scheduling decisions.

3.2.5 Successive Minimum Mean-Square-Error Precoding (SMMSE)

In recent years, there has been a considerable interest in wireless multiple-input, multiple output (MIMO) communications systems because of their promising improvement in terms of performance and bandwidth efficiency. An important research topic is the study of multi-user (MU) MIMO systems. Such systems have the potential to combine the high capacity achievable with MIMO processing with the benefits of space division multiple access. In the downlink scenario, a base station (BS) or an access point (AP) is equipped with multiple antennas and it is simultaneously transmitting to a group of users. Each of these users is also equipped with multiple antennas. Motivated by the need for cheap mobiles with low power consumption, we focus on systems where the complex signal processing is performed at the BS/AP. The BS/AP will use the channel state information (CSI) available at the transmitter to allow these users to share the same channel and mitigate or ideally completely eliminate multi-user interference (MUI). In a TDD system, using the reciprocity principle it is possible to use the estimated uplink channel for downlink transmission [PNG03]. This information can be used to perform joint precoding at the BS/AP of the users' signals. When the channel is varying too rapidly, then precoding based on the transmit correlation matrix $\mathbf{R}_t = E\{\mathbf{H}^H\mathbf{H}\}$, can also be used to reduce or eliminate multi-user interference. MU precoding based on long-term CSI will be discussed at the end of the section.

All precoding techniques can be classified according to the amount of the MUI that is suppressed at the

transmitter (as zero or non-zero MUI techniques) and linearity (as linear and non-linear techniques) as in Table 3.5. Linear precoding techniques do not require the overhead of sending the demodulation information to the mobile and are computationally less expensive than non-linear techniques. However, non-linear techniques can provide much higher capacity. For example, compared to BD, which is a linear precoding technique, SO THP can provide extra 2 bps/Hz at low SNRs and more than 8 bps/Hz extra at high SNRs.

MMSE precoding was first proposed in [JKG+02]. It can improve the system performance by introducing a certain amount of interference especially for users equipped with a single antenna. However, it suffers from a performance loss when it attempts to mitigate the interference between two closely spaced antennas if the user terminal is equipped with more than one receive antenna.

Table 3.5: Classification of different MU MIMO precoding techniques according to the level of MUI and linearity.

Zero MUI precoding techniques	Non-zero MUI precoding techniques
Channel inversion [GP94], [QTM02]	MMSE [JKG+02]
Block diagonalisation [SSH04]	SMMSE [SH04]
SO THP [SH05], [SHF04]	MMSE THP [JBU04]

Linear MIMO precoding techniques	Non-linear MIMO precoding techniques
Channel inversion [GP94], [QTM02]	MMSE THP [JBU04]
Block diagonalisation [SSH04]	SO THP [SH05], [SHF04]
MMSE [JKG+02]	BD MMSE THP [SHF04]
SMMSE [SH04]	

We consider a MU MIMO downlink channel where M_T transmit antennas are located at the base station, and M_{R_i} receive antennas are located at the *i*-th mobile station (MS), $i = 1,2,\ldots,K$. There are K users (or MSs) in the system. The total number of receive antennas is

$$M_R = \sum_{i=1}^{K} M_{R_i} . \tag{3.24}$$

We will use the notation $\{M_{R_i},\ldots,M_{R_K}\}\times M_T$ to describe the antenna configuration of the system. Let us consider one subcarrier, i.e., a flat fading channel. The MIMO channel to user *i* is denoted as $\mathbf{H}_i \in \mathrm{C}^{M_{R_i}\times M_T}$. Moreover, the combined channel matrix is given by

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1^T & \mathbf{H}_2^T & \cdots & \mathbf{H}_K^T \end{bmatrix}^T . \tag{3.25}$$

The precoding matrix for each subcarrier is equal to $\mathbf{F} = \mathbf{VP}$, $\mathbf{D}$ is the demodulation matrix, and the received vector is equal to

$$\underset{Q\times 1}{\underline{\mathbf{Y}}} = \underset{Q\times M_R}{\underline{\mathbf{D}}} \left(\underset{M_R\times M_T}{\underline{\mathbf{H}}} \; \underset{M_T\times Q}{\underline{\mathbf{V}}} \; \underset{Q\times Q}{\underline{\mathbf{P}}} \; \underset{Q\times 1}{\underline{\mathbf{S}}} + \underset{M_R\times 1}{\underline{\mathbf{Z}}} \right). \tag{3.26}$$

If we denote the *i*-th user's received vector, demodulation matrix, precoding matrix, transmitted vector and noise vector as $\mathbf{Y}_i \in \mathrm{C}^{Q_i\times 1}$, $\mathbf{D}_i \in \mathrm{C}^{Q_i\times M_{R_i}}$, $\mathbf{F}_i \in \mathrm{C}^{M_T\times Q_i}$, $\mathbf{S}_i \in \mathrm{C}^{Q_i\times 1}$, and $\mathbf{Z}_i \in \mathrm{C}^{Q_i\times 1}$, respectively, then we can write equation (3.26) as

$$\begin{bmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_K \end{bmatrix} = \begin{bmatrix} \mathbf{D}_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{D}_K \end{bmatrix} \cdot \left(\begin{bmatrix} \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_K \end{bmatrix} \cdot \begin{bmatrix} \mathbf{F}_1 & \cdots & \mathbf{F}_K \end{bmatrix} \cdot \begin{bmatrix} \mathbf{S}_1 \\ \vdots \\ \mathbf{S}_K \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_1 \\ \vdots \\ \mathbf{Z}_K \end{bmatrix} \right), \tag{3.27}$$

where Q_i is the number of spatial streams transmitted to the *i*-th user. Hence, the received matrix of the *k*-th user can be written as

$$\mathbf{Y}_k = \mathbf{D}_k \left(\mathbf{H}_k \sum_{i=1}^{K} \mathbf{F}_i \mathbf{S}_i + \mathbf{Z}_k \right). \tag{3.28}$$

A) Transmit precoding for short-range scenario

In case of we assume that short-term channel knowledge is acquired at the base station via the measurements on the uplink channel. The linear MMSE transmit filter is defined as

$$\mathbf{F} = \beta \left(\mathbf{H}^H \mathbf{H} + \alpha \mathbf{I}_{M_T \times M_T} \right)^{-1} \mathbf{H}^H, \tag{3.29}$$

where

$$\beta = \sqrt{\frac{P_T}{\mathrm{tr}\left(\mathbf{F}\mathbf{s}\mathbf{s}^H \mathbf{F}^H \right)}} \quad \text{and} \quad \alpha = \left(\frac{P_T}{M_R \sigma_n^2} \right)^{-1}. \tag{3.30}$$

P_T denotes the available transmit power, $\mathbf{s}$ is a data vector and σ_n^2 denotes the variance of the zero mean circularly symmetric complex Gaussian (ZMCSCG) noise.

The MMSE

$$\mathrm{MSE}_{opt} = \alpha \cdot \mathrm{tr} \left\{ \left(\mathbf{H}\mathbf{H}^H + \alpha \mathbf{I}_{M_R \times M_R} \right)^{-1} \right\}, \tag{3.31}$$

results if the precoder is chosen according to equation (3.53). In order to reduce the performance loss due to the cancellation of interference between two closely spaced antennas at the same terminal, the precoding matrix is generated by successively calculating the columns of the precoding matrix $\mathbf{F}$ for each of the receive antennas separately as explained in the sequel. This technique is called successive MMSE (SMMSE) [SH04].

The columns in the precoding matrix $\mathbf{F}$, each corresponding to one receive antenna, are calculated successively. For the *i*-th user, $i = 1,2,\ldots,K$, and *j*-th receive antenna $j = 1,2,...,M_{R_i}$ we define the matrix

$$\overline{\mathbf{H}}_i^{(j)} = \begin{bmatrix} \mathbf{h}_{i,j}^T \\ \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_{i-1} \\ \mathbf{H}_{i+1} \\ \vdots \\ \mathbf{H}_K \end{bmatrix}, \tag{3.32}$$

where $\mathbf{h}_{i,j}^T$ is the *j*-th row of the *i*-th user's channel matrix $\mathbf{H}_i$. The corresponding column of the precoding matrix $\overline{\mathbf{F}}$ is equal to the first column of the following matrix:

$$\mathbf{F}_{i,j} = \beta \left(\overline{\mathbf{H}}_i^{(j)H} \overline{\mathbf{H}}_i^{(j)} + \alpha \mathbf{I}_{M_T \times M_T} \right)^{-1} \overline{\mathbf{H}}_i^{(j)H}. \tag{3.33}$$

After calculating the beamforming vectors for all receive antennas in this fashion, the equivalent combined channel matrix of all users is equal to $\mathbf{H}\overline{\mathbf{F}} \in \mathrm{C}^{M_R \times M_R}$ after the precoding

$$\begin{bmatrix} \mathbf{H}_1\overline{\mathbf{F}}_1 & \cdots & \mathbf{H}_1\overline{\mathbf{F}}_K \\ \vdots & \ddots & \vdots \\ \mathbf{H}_K\overline{\mathbf{F}}_1 & \cdots & \mathbf{H}_K\overline{\mathbf{F}}_K \end{bmatrix}. \tag{3.34}$$

For high SNR ratios, this matrix is also block diagonal. We can now apply any other previously defined SU MIMO technique on the *i*-th user's equivalent channel matrix $\mathbf{H}_i\overline{\mathbf{F}}_i$. After the precoding using the matrix $\overline{\mathbf{F}}_i \in \mathrm{C}^{M_T \times M_{R_i}}$, we first perform the singular value decomposition (SVD) of $\mathbf{H}_i\overline{\mathbf{F}}_i$

$$\mathbf{H}_i\overline{\mathbf{F}}_i = \mathbf{U}_i\mathbf{\Sigma}_i\mathbf{V}_i^H. \tag{3.35}$$

If the number of spatial substreams transmitted to the *i*-th user is $Q_i \le M_{R_i}$, the resulting precoding matrix of the *i*-th user is equal to the product of the matrix $\overline{\mathbf{F}}_i$ and the Q_i right column vectors of $\mathbf{V}_i$ denoted by $\mathbf{V}_i^{(1)} \in C^{M_{R_i} \times Q_i}$

$$\mathbf{F}_i = \overline{\mathbf{F}}_i\mathbf{V}_i^{(1)}. \tag{3.36}$$

For example, there are several possibilities to determine the number of spatial streams. If we want to maximise the capacity of the system we use water-filling (WF) on the eigenmodes of all users $\mathbf{\Sigma} = diag(\mathbf{\Sigma}_1,\ldots,\mathbf{\Sigma}_K)$, or if we want to extract the maximum diversity and array gain, we transmit only on the dominant eigenmode of each user. Dominant eigenmode transmission (DET) can provide maximum SNR at the receiver and minimum BER performance. By using this algorithm the system performance is efficiently improved by introducing MUI and by eliminating inter-stream interference. Figure 3.17 describes the design of the precoding matrix using SMMSE.

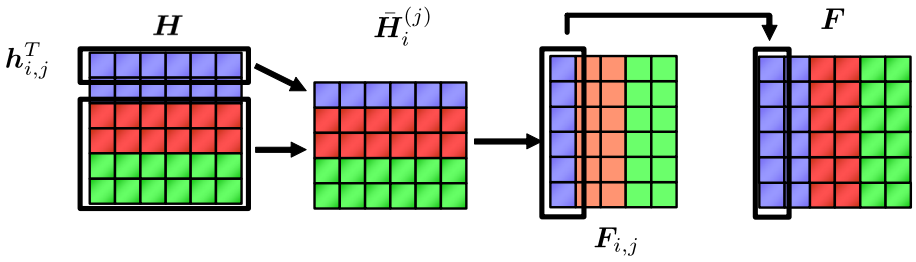


Figure 3.17: Successive MMSE – Block diagram

B) Transmit Precoding for Wide-Area Scenario

If we assume that the channel varies too rapidly to be tractable, the information regarding the relative geometry of the propagation paths is captured by a non-white correlation matrix [SH05a].

If we define the singular value decomposition (SVD) of the *i*-th user's correlation matrix $\mathbf{R}_{t_i}$ as

$$\mathbf{R}_{t_i} = \mathbf{Q}_i\mathbf{\Lambda}_i\mathbf{Q}_i^H, \tag{3.37}$$

then the matrix $\mathbf{R}_{t_i}^{1/2}$ can be written as $\mathbf{R}_{t_i}^{1/2} = \mathbf{Q}_i\mathbf{\Lambda}_i^{1/2}\mathbf{Q}_i^H$. Next, we introduce the matrix

$$\hat{\mathbf{H}}_i = \mathbf{\Lambda}_i^{1/2}\mathbf{Q}_i^H. \tag{3.38}$$

The same algorithm used in case when we have short-term channel knowledge at the transmitter is also applied, but instead of $\mathbf{H}$ we use $\hat{\mathbf{H}} = \left[\hat{\mathbf{H}}_1^T \quad \hat{\mathbf{H}}_2^T \quad \cdots \quad \hat{\mathbf{H}}_K^T\right]^T$.

3.2.5.1 Required Support Functions

This technique requires full channel knowledge at the mobile terminals. At the base station/access point either short-term channel knowledge or long-term CSI can be used to perform the precoding. The pilots transmitted in each chunk can be used to estimate the channel at the BS/AP. This and previous measure-

ments can then be used to predict the channel on the downlink. The information about the noise level at the MS is fed back to the base station/access point.

3.2.5.2 Additional Assumptions and Simplifications for Simulations

In the simulations we will assume a perfect decorrelation of the users in space. This assumption is valid if the number of users is large enough and if a proper spatial scheduling taking into account time and space dimension is applied as described in [FGH05], [GH04]. We assume that the channel is varying slowly enough so that the reciprocity principle can be applied.

3.2.5.3 Summary

Channel knowledge at the base station/access point allows us to perform joint precoding of all users' signals that in turn can result in a significant performance improvement. In a TDD system the channel measurements on the uplink channel can be used to acquire the necessary channel knowledge. In a short-range scenario we can directly use the short-term channel knowledge for this purpose. However, if the channel is changing too fast to keep track of the channel mean, second order statistics can be also used for precoding. Linear precoding techniques do not require any signalling overhead for the purpose of demodulation at the mobile station and they are in general less computationally expensive than the non-linear precoding techniques. MMSE precoding is very attractive with users equipped with only one antenna. However, when the users are also equipped with multiple antennas it causes a performance loss due to the inter-stream cancellation between two closely spaced antennas. SMMSE addresses this problem and handles each receive antenna separately. SMMSE has the advantage over other precoding techniques that the total number of receive antennas at the users' terminals can be greater than the number of antennas at the base station. This technique is especially useful at low SNRs and the possibility of having more antennas at the mobile stations allows further improvements of the performance.

3.2.6 Block Diagonalisation (BD)

Block diagonalisation (BD) is also a linear precoding technique that allows spatial multiplexing between different users (SDMA) and the transmission of multiple data streams per user. In contrast to SMMSE, it does not allow any multi-user interference (MUI). Two versions of BD are presented here. The first one assumes short-term channel state information (CSI) at the transmitter (transmit precoding for short-range scenarios). In the wide-area case, however, we assume that only second order statistics of the CSI, i.e., long-term spatial covariance matrices, are available at the transmitter.

A) Transmit precoding for short-range scenarios

Block diagonalisation was first proposed in [SH02] and further investigated in [SSH04] and [SPS+04]. It is restricted to channels where the number of transmit antennas M_T is greater or equal to the total number of receive antennas in the network M_R.

Let us define the precoder matrices as

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_1 & \mathbf{F}_2 & \cdots & \mathbf{F}_K \end{bmatrix} \in \mathrm{C}^{M_T \times Q}, \tag{3.39}$$

where $\mathbf{F}_i \in \mathrm{C}^{M_T \times r_i}$ is the *i*-th user's precoder matrix. Moreover, $Q \le M_R$ is the total number of transmitted spatial substream sequences, while $q_i \le M_{R_i}$ is the number of data stream sequences transmitted to the *i*-th user. We can find the optimal precoding matrix $\mathbf{F}$ such that all MUI is zero by choosing a precoding matrix $\mathbf{F}_i$ that lies in the null space of the other users' channel matrices. Thereby, a MU MIMO downlink channel is decomposed into multiple parallel independent SU MIMO channels [SH04], [FGH05].

If we define $\widetilde{\mathbf{H}}_i$ as

$$\widetilde{\mathbf{H}}_i = \begin{bmatrix} \mathbf{H}_1^T & \cdots & \mathbf{H}_{i-1}^T & \mathbf{H}_{i+1}^T & \cdots & \mathbf{H}_K^T \end{bmatrix}, \tag{3.40}$$

then the zero MUI constraint forces $\mathbf{F}_i$ to lie in the null space of $\widetilde{\mathbf{H}}_i$. From the singular value decomposition (SVD) of $\widetilde{\mathbf{H}}_i$ whose rank is $\widetilde{L}_i$,

$$\widetilde{\mathbf{H}}_i = \widetilde{\mathbf{U}}_i \widetilde{\mathbf{\Sigma}}_i \begin{bmatrix} \widetilde{\mathbf{V}}_i^{(1)} & \widetilde{\mathbf{V}}_i^{(0)} \end{bmatrix}^H, \tag{3.41}$$

we choose the last right $M_T - \widetilde{L}_i$ singular vectors $\widetilde{\mathbf{V}}_i^{(0)} \in \mathrm{C}^{M_T \times (M_T - \widetilde{L}_i)}$, which form an orthogonal basis for the null space of $\widetilde{\mathbf{H}}_i$. The equivalent channel of user *i* after eliminating the MUI is identified as $\mathbf{H}_i \widetilde{\mathbf{V}}_i^{(0)}$, whose dimension is $M_{R_i} \times \left(M_T - \widetilde{L}_i\right)$ and is equivalent to a system with $M_T - \widetilde{L}_i$ transmit antennas and M_{R_i} receive antennas. Each of these equivalent SU MIMO channels has the same properties as a conventional SU MIMO channel. Define the SVD,

$$\mathbf{H}_i \widetilde{\mathbf{V}}_i^{(0)} = \mathbf{U}_i \mathbf{\Sigma}_i \left[\mathbf{V}_i^{(1)} \quad \mathbf{V}_i^{(0)}\right]^H, \tag{3.42}$$

and let the rank of the i-th user's equivalent channel matrix be L_i. The product of the first L_i singular vectors $\mathbf{V}_i^{(1)}$ and $\widetilde{\mathbf{V}}_i^{(0)}$ produces an orthogonal basis of dimension L_i and represents the transmission vectors that maximise the information rate for user *i* subject to the zero MUI constraint.

B) Transmit precoding for wide-area scenarios

As in the case of SMMSE we assume that if the channel is varying too rapidly to track its mean, the base station/access point has perfect channel knowledge about the transmit correlation matrices of each user.

If we define the singular value decomposition (SVD) of the *i*-th user's transmit correlation matrix $\mathbf{R}_{t_i}$ as

$$\mathbf{R}_{t_i} = \mathbf{Q}_i \mathbf{\Lambda}_i \mathbf{Q}_i^H, \tag{3.43}$$

then the matrix $\mathbf{R}_{t_i}^{1/2}$ can be written as $\mathbf{R}_{t_i}^{1/2} = \mathbf{Q}_i \mathbf{\Lambda}_i^{1/2} \mathbf{Q}_i^H$. Next, we introduce the matrix

$$\hat{\mathbf{H}}_i = \mathbf{\Lambda}_i^{1/2} \mathbf{Q}_i^H. \tag{3.44}$$

The same algorithm is applied as in case when we have short-term channel knowledge at the transmitter but instead of $\mathbf{H}$ we use $\hat{\mathbf{H}} = \left[\hat{\mathbf{H}}_1^T \quad \hat{\mathbf{H}}_2^T \quad \cdots \quad \hat{\mathbf{H}}_K^T\right]^T$.

3.2.6.1 Required Support Functions

This technique requires full channel knowledge at the mobile terminals. At the base station/access point either short-term channel knowledge or long-term CSI can be used to perform the precoding.

3.2.7 Precoding with Long-Term/Second Order Statistics CSI

In general, a better understanding of multi-antenna systems in different transmission scenarios is necessary in order to obtain performance and capacity improvements. Here, we investigate a multi-user MIMO downlink system where the channels are modelled by Ricean fading. We assume that the receivers have perfect channel state information while the transmitter has only knowledge of the means of the channel realisations. Such a scenario can occur if we consider a TDD system in which the base station estimates the channel during the uplink transmission and uses this estimation for downlink transmission. Let's consider two scenarios that can be modelled by Ricean fading.

I: Consider the scenario in which the flat-fading MIMO channel has zero-mean, i.e., Rayleigh fading, and in which the transmitter has imperfect CSI. The imperfect CSI is due to channel estimation errors during the channel estimation phase and due to the time-varying channel. If we assume that the channel estimation is performed during uplink transmission using an MMSE estimator, the estimated channel matrix is the sum of the correct channel matrix plus the estimation error. During downlink transmission the channel estimation is used. The estimation error itself is an i.i.d. complex Gaussian distributed matrix with variance proportional to the pilot signal length. As a result, the transmitter 'sees' an i.i.d. complex Gaussian distributed fading matrix with mean that equals the estimated channel matrix. This scenario leads directly to a Ricean fading channel matrix with knowledge about the mean at the transmitter.

II: If there is a direct LOS connection between the transmitter and the receiver, the flat-fading MIMO channel can be modelled as i.i.d. complex Gaussian distributed with nonzero mean that equals the LOS component. The receiver has perfect CSI. The transmitter gets CSI via a feedback link from the receiver. In order to have high spectral efficiency the bandwidth of the feedback link is kept small. As a result, the receiver collects and averages over the channel realisations and feeds back the average. Even if the channel has zero-mean, the receiver would feed back an average channel realisation. Therefore, the transmitter knows only the mean of the fading process.

Then Scenario I and Scenario II have the same channel model and the transmitter have the channel mean. The real channel varies (Gaussian distributed) around the estimation (channel mean). Even this type of imperfect channel state information (CSI) can help the transmitter adapting to the fading situation in order to increase the spectral efficiency and performance. We assume further that the receivers (the mobile terminals) perform optimal linear processing in terms of multi-user MMSE receiver. Therefore, our performance metric is the average normalised MSE at the receiver. We will derive the optimal precoding strategy and the achievable minimum average normalised MSE. Using the mean channel information, we will derive optimal power allocation and beamforming strategy in terms of minimising the MSE and analyse the impact of the channel estimation error or Ricean factor K on the performance on the MSE performance.

In Figure 3.18, the transmission line for one user k including the interference induced by the next user k+1 is illustrated in an overview style. The main parts, i.e., the transmitter structure (abbreviated by Q), the channel H, and the MMSE receiver are identified.

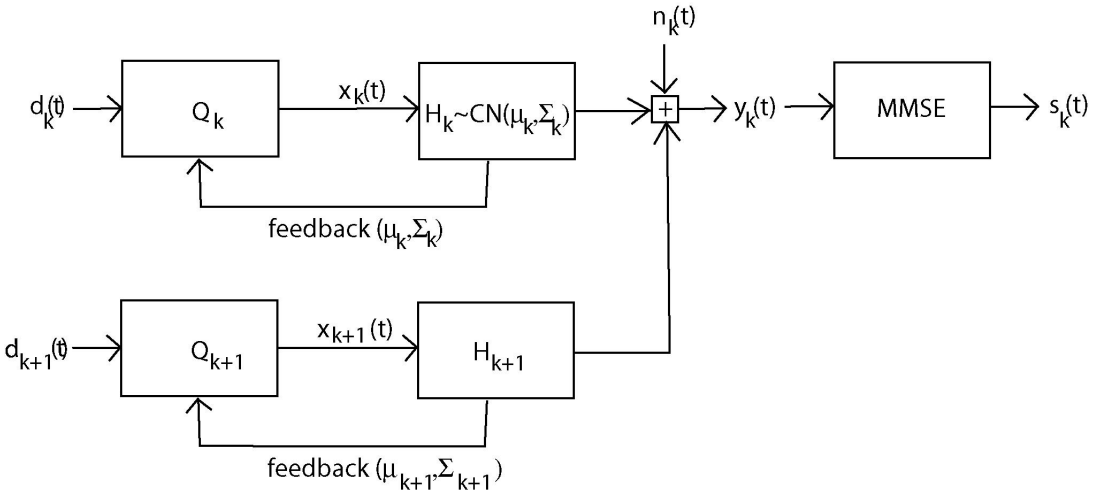


Figure 3.18: Transmission model multi-user MIMO with flat fading Ricean channel and partial CSI at the transmitter, perfect CSI at the receiver and linear multi-user MMSE receiver.

The information data $\mathbf{d}(t)$ at time point t is processed at the transmitter. The resulting vector $\mathbf{x}(t)$ is transmitted over the available antennas. The transmit covariance matrix consists of the data intended for the considered user and all data for other users in the same cell. The sum of all these transmit signals passes through the flat fading MIMO channel **H,** which is assumed to be distributed according to a complex Gaussian distribution with mean matrix $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. At the receiver the additive white Gaussian noise $\mathbf{n}(k)$ and the multi-user interference is added. Next, the received signal $\mathbf{y}(k)$ is then processed by linear multi-user MMSE filtering, which is the optimal linear filter. Finally, at the output of the MMSE receiver an estimate $\mathbf{s}(k)$ of the transmitted vector $\mathbf{d}(k)$ is obtained.

Finally, we look at the receiver side and describe the MMSE receiver. The optimal linear receiver estimates the transmitted signals by taking the received covariance matrix and the noise covariance matrix into account. The estimate of the transmitted signal is given by

$$s_k(k) = \rho \mathbf{Q}_k \mathbf{H}_k{}^H \left(\mathbf{I} + \rho \mathbf{H}_k \mathbf{Q}_k \mathbf{H}_k{}^H + \rho \sum_{\substack{l=1 \\ l \neq k}}^{K} \mathbf{H}_k \mathbf{Q}_l \mathbf{H}_k{}^H \right)^{-1} \mathbf{y}(k) , \tag{3.45}$$

for time point k. The transmit covariance matrix of the user k is given by $\mathbf{Q}_k$. The SNR is ρ. The resulting performance metric is obviously the mean square error (MSE) itself. Since we deal with a vector transmission system, the scalar notion of the MSE has to be extended to the vector case. In order to have fair comparisons, the metric is the normalised by the power of the transmit signals $\mathbf{x}(k)$. Finally, the metric is averaged over the channel realisations $\mathbf{H}_k$ and we obtain the average normalised MSE_k of user k as a

metric for this system [VAT99], [HB03], [JB03]

$$MSE_k = M_T - E\left\{ tr\left(\rho \mathbf{H}_k \mathbf{Q}_k \mathbf{H}_k{}^H \left[\mathbf{I} + \rho \mathbf{H}_k \mathbf{Q}_k \mathbf{H}_k{}^H \rho \sum_{\substack{l=1 \\ l \neq k}}^{K} \mathbf{H}_k \mathbf{Q}_l \mathbf{H}_k{}^H \right]^{-1} \right) \right\}, \tag{3.46}$$

which is a function of the transmit covariance matrix $\mathbf{Q}_k$, which in turn depends on the transmit strategy. The MSE in the formula above is a convex function with respect to the transmit covariance matrix $\mathbf{Q}_k$. The MSE is a one-shot criterion that measures the quality of a transmission and it is closely related to the SINR and the BER.

The optimal transmit strategy of the users is characterised in the following. Each user has an independent Gaussian codebook. The data stream is then serial parallel multiplexed and power is allocated to the parallel data streams. Finally, each data stream of each user is multiplied by a beamforming vector. The Beamforming vectors correspond to the eigenvectors of the transmit covariance matrix and the power allocation corresponds to the eigenvalues of the transmit covariance matrix. In order to find the optimal transmit covariance matrix, the expression for the MSE of user k in (3.46) is minimised with respect to $\mathbf{Q}_k$ [JSB+04]. Then the individual MSE is minimised. In addition to this, the average sum MSE can be jointly minimised with respect to set of transmit covariance matrices.

3.2.8 Variable Collision Multiple Access

3.2.8.1 Overall description of technique

Variable Collision Multiple Access (VCMA) is an uplink multiple access method for MIMO frequency-selective channels with CSIR and no CSIT that allows the system designer to gradually vary the amount of OFDM tone collision among users by assigning potentially overlapping subsets of OFDM tones to different users. The benefits of this flexibility are described below in terms of spectral efficiency (in bps/Hz). In addition, VCMA is a suitable framework to compare different multiple access schemes.

A short technical description of VCMA is as follows. Let K and N denote the total number of users and OFDM tones, respectively. VCMA assigns tone m ($m = 1, \ldots, N$) to a subset of users denoted by U_m, where $U_m \subseteq \{1, \ldots, K\}$. For instance, the case where $|U_m| = K$ for all m corresponds to a full-collision multiple access scheme where all OFDM tones are assigned to all users, and thus models CDMA (in the sense that every user occupies the entire frequency spectrum). On the other hand, the case where $|U_m| \leq 1$ for all m corresponds to a no-collision multiple access scheme where each OFDM tone is assigned to at most one user, and thus models FDMA (in the sense that users are allocated non-overlapping frequency bands). Therefore, VCMA is a unifying framework for multiple access based on a simple parameterisation by the partitioning subsets $\{U_m\}$.

In the following sections, we describe the channel and signalling models, discuss the benefits of full collision and conditions under which the amount of collision can be reduced, and finally quantify the loss of suboptimal (but simpler) techniques in terms of spectral efficiency. In particular, we emphasise that for rich scattering and a small number of receive antennas, very little collision in frequency is needed to realise a significant fraction of the available sum capacity. Minimising the amount of user collision in frequency is desirable in practice, as this minimises the receiver complexity incurred by having to separate the interfering signals.

3.2.8.1.1 Channel and Signalling Models

We assume a frequency-selective L-tap fading MIMO K-users multiple access channel. Each user employs OFDM with N tones, such that the cyclic prefix L_{cp} is greater than L (we ignore the loss of spectral efficiency due to this cyclic prefix). It follows from this assumption that each user's frequency-selective fading MIMO channel decomposes into a set of parallel frequency-flat fading MIMO channel. In addition, we assume fast fading, i.e., the channel fading coefficients change independently from one transmission to the next, and uncorrelated channels among users and tones. Consequently, the received signal vector for the m-th tone is

$$\mathbf{y}_m = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{x}_{k,m} + \mathbf{n}_m, \quad \text{where} \quad m = 1,\ldots,N \quad \text{and} \quad \mathbf{R}_{n_m,n_m} = \sigma_n^2 \mathbf{I}_{M_R}. \tag{3.47}$$

We have used the fact that the $M_R \times M_T$ channel matrices for user k are identically distributed for all tones [VB04]. Thus, $\mathbf{H}_k$ does not dependent on the tone index m. Moreover, we assume spatially uncorrelated fading at the transmit antennas, and we model the spatial fading correlation at the receive antenna array by decomposing the k-th user channel as

$$\begin{aligned} \mathbf{H}_k &= \mathbf{R}_k^{1/2} \mathbf{H}_{k,w}, \text{where} \quad \mathbf{R}_k \text{ is the receive correlation matrix} \\ &\text{and} \quad \mathbf{H}_{k,w} \text{ has i.i.d. complex normal entries (mean = 0 and variance = 1)} \end{aligned} \tag{3.48}$$

The M_T vector $\mathbf{x}_{k,m}$ is the data vector transmitted by user k on the m-th OFDM tone. It has power P_k which we write as a fraction of the total power $P = P_1 + \ldots P_K$, that is, $P_k = d_k P$. The M_R vector $\mathbf{n}_m$ represents the additive white Gaussian noise on tone m, and is such that the noise at different tones is independent.

3.2.8.1.2 Multiple Access with Variable Amount of Collision

VCMA is in fact a parametric family of multiple access schemes obtained by assigning each OFDM tone $m = 1,\ldots, N$ to a subset of users denoted by U_m. A tone assignment is thus parameterised by the subsets $\{U_1,\ldots,U_N\}$. For instance, a fully collision-based multiple access scheme where all tones are assigned to each user (i.e., $|U_m| = K$) is referred to as CDMA. We emphasise that the terminology CDMA is used here solely to indicate that all users occupy the entire bandwidth, and that the effect of redundancy-introducing spreading are not considered. FDMA is obtained for $|U_m| \leq 1$.

The total power P_k of user k is split among the N OFDM tones, and we denote by $b_{k,m}$ the fraction of power allocated to the m-th tone, with the convention that $b_{k,m} = 0$ if user k is not assigned this tone (that is if $k \notin U_m$).

3.2.8.1.2.1 Ergodic Sum Capacity

In this section, we present results on the ergodic sum capacity of VCMA for simplicity, but the results extend to the entire capacity region [VB04].

Assuming joint decoding at the receiver, the ergodic sum capacity for a variable amount of collision is given by

$$C_{S-VCMA} = E\left\{ \frac{1}{N} \sum_{m=0}^{N} \log_2 \det\left(\mathbf{I}_{M_R} + \frac{P}{\sigma_n^2 M_T} \sum_{k=1}^{K} b_{k,m} d_k \mathbf{H}_k \mathbf{H}_k^{\mathsf{H}} \right) \right\}. \tag{3.49}$$

$C_{S\text{-}VCMA}$ is upper bounded by the fully collision-based scheme, that is, when every user allocates an equal fraction of power to every OFDM tone; in other words, when $b_{k,m} = 1/N$ for all $k = 1,\ldots, K$ and $m = 1,\ldots, N$. Consequently, $|U_1| = \ldots = |U_N| = N$, and hence

$$C_{S-VCMA} \leq C_{S-CDMA} = E\left\{ \log_2 \det\left(\mathbf{I}_{M_R} + \frac{P}{\sigma_n^2 M_T N} \sum_{k=1}^{K} d_k \mathbf{H}_k \mathbf{H}_k^{\mathsf{H}} \right) \right\}. \tag{3.50}$$

For any other tone allocation $\{U_1,\ldots,U_N\}$, we denote by N_k the number of tones that user k is allocated to, and assume that it uniformly splits its power among these N_k tones:

$$b_{k,m} = \begin{cases} \dfrac{1}{N_k}, & k \in U_m \\ 0, & k \notin U_m \end{cases}. \tag{3.51}$$

This assumption of uniform power allocation across allocated tones is for simplicity, and is not necessarily optimal (in terms of sum capacity). With these definitions, the sum capacity is

$$C_{S-VCMA}\{U_1,\ldots,U_N\} = E\left\{ \frac{1}{N} \sum_{m=0}^{N} \log_2 \det\left(\mathbf{I}_{M_R} + \frac{P}{\sigma_n^2 M_T} \sum_{k=1}^{K} \frac{1}{N_k} d_k \mathbf{H}_k \mathbf{H}_k^{\mathsf{H}} \right) \right\}. \tag{3.52}$$

3.2.8.1.2.2 Multi-User Spatial Multiplexing Gain

Analogous to the spatial multiplexing gain for a point-to-point link [VB04], we define the *multi-user* spatial multiplexing gain for a given tone allocation $\{U_1,\ldots,U_N\}$ as

$$m\left(\{U_1,\ldots,U_N\}\right) = \lim_{\frac{P}{\sigma^2}\to\infty} \frac{C_{S-VCMA}\{U_1,\ldots,U_N\}}{\log_2\left(P/\sigma^2\right)}. \tag{3.53}$$

From the previous section, it is clear that the multi-user spatial multiplexing gain achieved for the full-collision scheme, denoted by m_{CDMA}, is an upper bound for any other scheme.

It is shown in [VB04] that $m_{CDMA} = \mathrm{E}[\mathrm{rank}(\mathbf{H})]$, where $\mathbf{H} = [\mathbf{H}_1,\ldots,\mathbf{H}_K]$ and that the circularly symmetric complex Gaussian assumption on the fading channel matrices $\mathbf{H}_k$ yields that rank($\mathbf{H}$) is constant with probability one (w.p. 1). Hence, $m_{CDMA} = \mathrm{rank}(\mathbf{H})$ w.p. 1. A general expression for m_{CDMA} in terms of $\mathbf{R}_k$, M_T, M_R, and K cannot be given. For $M_R \le M_T$, however, it can be shown that

$$m_{CDMA} = \mathrm{rank}\left(\sum_{k=1}^{K}\mathbf{R}_k\right), \tag{3.54}$$

and for $M_R > M_T$ that the following bounds hold

$$\max_k\left[\min\left(\mathrm{rank}(\mathbf{R}_k), M_T\right)\right] \le m_{CDMA} \le \min\left[\mathrm{rank}\left(\sum_{k=1}^{K}\mathbf{R}_k\right), KM_T\right]. \tag{3.55}$$

Both the exact expression and the lower and upper bounds show that the multi-user multiplexing gain can be significantly higher than the single-user multiplexing gain that would be obtained if only one user were present. This is due to the fact the in the multi-user case, all the users contribute to the multiplexing gain (when sum capacity is the measure of interest). More specifically, the presence of multiple users increases the effective number of transmit antennas from M_T to KM_T. On the receiver side, the limiting factor for multiplexing gain is rank($\Sigma_k\mathbf{R}_k$) rather than rank($\mathbf{R}_k$) in the single-user case. Since in practice the number of users K is generally large, rank($\Sigma_k\mathbf{R}_k$) typically determines m_{CDMA}.

In order to obtain a high-rank sum-correlation matrix $\Sigma_k\mathbf{R}_k$, we either need the receive antenna spacing to be large so that the individual correlation matrices $\mathbf{R}_k$ are high-rank, or alternatively the $\mathbf{R}_k$ have to span different subspaces. In practice, the latter requirement tends to be satisfied if the individual users are well separated in space, which is typically the case in a cellular system. We finally note that for large M_R and rich scattering/large user separation, m_{CDMA} can be K-times higher than the single-user multiplexing gain that would be obtained if only one of the users were present.

As mentioned earlier, minimising the amount of collision is desirable in practice. Next, we quantify the impact of suboptimum multiple access schemes, that is, of user collision in frequency (or lack thereof), on the multi-user multiplexing gain. For a general tone assignment $\{U_1,\ldots,U_N\}$ we have

$$m\left(\{U_1,\ldots,U_N\}\right) = \frac{1}{N}\sum_{m=1}^{N}\mathrm{rank}\left(U_m\right), \quad \text{where} \quad \mathrm{rank}\left(U_m\right) = \mathrm{rank}\left(\sum_{k\in U_m}\mathbf{H}_k\mathbf{H}_k^{\mathsf{H}}\right). \tag{3.56}$$

Since rank(U_m) ≤ rank(U) where $U = \{1, \ldots, K\}$, it follows that $m(\{U_1,\ldots,U_N\}) \le m_{CDMA}$ (as we expect), thereby showing that the multi-user multiplexing gain for any tone assignment (and hence any amount of collision) is upper-bounded by the multi-user multiplexing gain obtained for CDMA (full collision). The above equation, however, shows that one does not have to enforce full collision in frequency to achieve m_{CDMA} : it suffices to choose tone assignments that result in rank(U_m) = rank(U) for $m = 1,\ldots, N$. The extent to which this is possible depends on the channel's spatial fading statistics and the number of transmit and receive antennas. It is shown in [VB04] that collision in frequency between users i and j does not affect the multi-user multiplexing gain if and only if rank($\mathbf{R}_i$) = rank($\mathbf{R}_j$) ≤ M_T and span($\mathbf{R}_i$) = rank($\mathbf{R}_j$), which corresponds to the case where the channel does not provide any spatial separation between users i and j. This observation suggests a simple strategy for optimum (in the sense of multiplexing gain) tone assignment when rank($\mathbf{R}_k$) ≤ M_T for $k = 1, \ldots, K$. On a given tone, only users with different span($\mathbf{R}_k$) should collide. Intuitively, this ensures that the spatial degrees of freedom offered by the multiple access

channel (and responsible for the multiuse multiplexing gain) are indeed exploited. Recall that in point-to-point MIMO links, in order to realise spatial multiplexing gain, it is crucial that the signals transmitted from the individual antennas are co-channel (or equivalently collide in signal space). Strategies for optimum tone assignment in the general case are discussed in [VB04].

Finally, we consider two extreme cases to further illustrate the role of user collision in frequency. For $\text{span}(\mathbf{R}_1) = \ldots = \text{span}(\mathbf{R}_K)$ and $\text{rank}(\mathbf{R}_1) \le M_T$, the exact expression for m_{CDMA} above (for $M_R \le M_T$) and the lower and upper bounds above (for $M_R > M_T$) all meet and yield $m_{CDMA} = \text{rank}(\mathbf{R}_1)$. For this case, it can be shown that $m_{FDMA} = m_{CDMA}$, that is, orthogonal multiple access achieves the full multi-user multiplexing gain. On the other hand, when $\mathbf{R}_i\mathbf{R}_j = 0$ for any $i \neq j$ (the extreme case of perfect spatial separation between all the users and large number of receive antennas), $m_{FDMA} = m_{CDMA}/K$.

To summarise and conclude, we emphasise that in practice, for good spatial separation between the users and for large M_R, collision in frequency (signal space) is critical to achieve a high multi-user multiplexing gain. On the other hand, for poor spatial separation and/or a small M_R, little or no collision is needed to achieve m_{CDMA}. It should be noted that in the latter case, m_{CDMA} will be smaller than in the former case.

3.2.8.1.2.3 The Two-User Case

The two-user case lends itself to an analytical study of how much collision is actually needed in different scenarios to achieve or approach the optimal (sum capacity) performance of the full-collision scheme. Results for a two-user scenario are discussed in [VB04] for both cases of joint and single-user decoding. These results are briefly summarised below.

Joint decoding:

- In the low-SNR regime, the amount of collision has no impact on the capacity region, nor does the amount of spatial fading correlation.
- In the high-SNR regime, the individual user rates are halved in the no-collision case compared to the full-collision case. In terms of sum capacity, collision is required either in the spatial dimension or in frequency.

Single-user decoding:

- In the low-SNR regime, the amount of collision has no impact on the capacity region.
- In the high-SNR regime, it is optimum for good spatial separation to have the users fully collide in frequency. When the separation is poor, the capacity region is maximised for orthogonal accessing.

3.2.8.1.2.4 Asymptotic Analysis

An asymptotic analysis in the number of users is another powerful framework that lends itself to analytical comparison of FDMA and CDMA. These results can be viewed as an extension to the MIMO case of the single-antenna results of [KH95].

For simplicity, we assume $\mathbf{R}_1 = \ldots = \mathbf{R}_K = \mathbf{I}_{M_R}$ and equal power users, i.e., $P_1 = \ldots = P_K = P$. Furthermore, we set $N = \eta K$, with η an integer, and we consider the large SNR P/σ^2 case. With these assumptions, the asymptotic (as K goes to infinity) difference in sum capacity is given by

$$C_{S-CDMA} - C_{S-FDMA} \approx \left(M_R - \min\left(M_T, M_R\right)\right)\log_2\left(P/K\right) + \min\left(M_T, M_R\right)\log_2\left(M_T\right) - \frac{1}{\ln 2}\left(\sum_{j=1}^{\min\left(M_T,M_R\right)\min\left(M_T,M_R\right)-j}\sum_{p=1}^{}\frac{1}{p} - \gamma\min\left(M_T, M_R\right)\right), \quad (3.57)$$

where $\gamma \approx 0.5772$ denotes Euler's constant.

Figure 3.19 shows this difference for $M_R = 2$, 4 and $P/K = 20$ dB as a function of M_T. In the regime $M_T < M_R$, we observe that the asymptotic sum capacity performance benefits significantly from collision in frequency. As M_T increases, the performance gap between CDMA and FDMA closes. For $M_R \le M_T$, the difference no longer depends on P, which implies that FDMA achieves the same (asymptotic in K) multi-user multiplexing gain as CDMA. This is due to the fact that for $M_R \le M_T$, the multi-user multiplexing gain is "bottle-necked" by M_R, and collision of the transmit signals across the M_T antennas of an individ-

ual user is sufficient to achieve full multi-user multiplexing gain. For fixed $M_T < M_R$, the performance gap increases with M_R, which can be explained as follows: increasing M_R opens up more spatial dimensions and hence collision in frequency becomes mandatory to achieve full multi-user multiplexing gain.

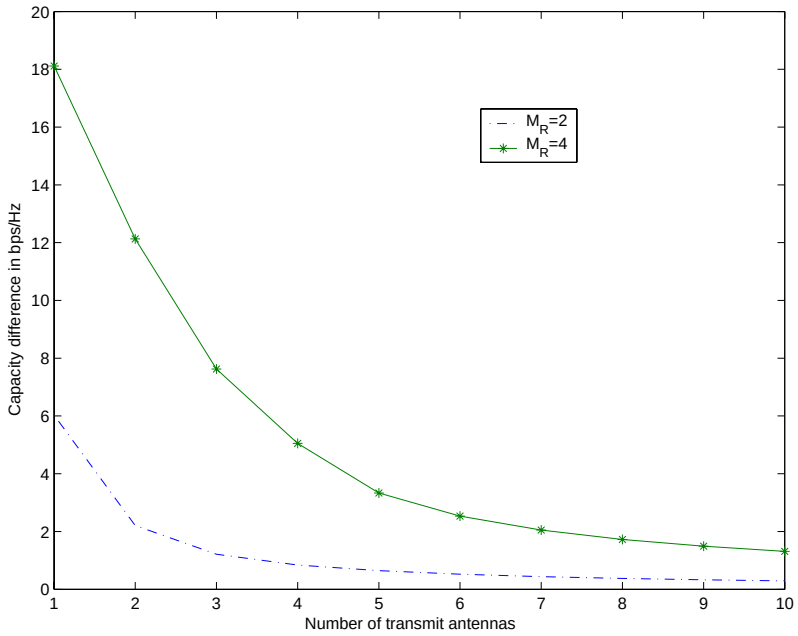


Figure 3.19: Asymptotic sum capacity difference between CDMA and FDMA for different values of M_R

3.2.8.2 Summary and Conclusion

VCMA is a family of MIMO multiple access schemes for frequency-selective fading channels that allows to gradually vary the amount of user collision in frequency (signal space) by assigning different subsets of the available OFDM tones to different users. Therefore, VCMA provides a unifying framework to compare multiple access schemes ranging from FDMA (no-collision) to CDMA-like (full-collision). The performance is assessed in terms of spectral efficiency by computing the sum capacity (it actually extends to the entire capacity region), the multi-user spatial multiplexing gain, and by considering the two-user and asymptotic scenarios.

The VCMA framework highlights significant differences for different SNR regimes: at low SNR, the capacity region is little affected by the amount of collision, whereas at high SNR it is governed by the spatial separation between users. With regard to multi-user multiplexing gain, it is typically limited by the richness of scattering at the receiver or the number of receive antennas. Depending on the propagation conditions and the number of transmit and receive antennas, the multi-user multiplexing gain can be significantly higher that the multiplexing gain that would be obtained if only one of the users were present. For good spatial separation between users and large M_R, frequency-collision is crucial to achieve high multiplexing gain. On the other hand, for poor spatial separation and/or small M_R, little or no collision is needed.

Finally, we mention that it is possible to extend VCMA to include time-domain collisions, and thus cover the entire range of multiple access schemes. Such a unifying framework for the forward link would also be a useful tool.

3.3 Non-Linear Techniques Using Multi-User Optimisation

3.3.1 Successive Optimisation THP (SO THP)

By applying BD on the combined channel matrix of all users the MU MIMO channel can be transformed into a set of parallel single-user MIMO channels. However, there is a capacity loss due to the cancellation of overlapping subspaces of different users. In [SH02], the authors propose a successive precoding algo-

rithm in order to define a simplified solution of the power control problem. By allowing a certain amount of interference, this algorithm reduces the capacity loss due to the subspace cancellation.

First, we have to assume or determine a certain optimum ordering of the users. Using SO, the modulation matrix for each user is designed in such a way that it lies only in the null space of the channel matrices of previous users. As a consequence, only they will generate the interference to this user. Let us define the previous *i*-1 users' combined channel matrix as

$$\hat{\mathbf{H}}_i = \left[\mathbf{H}_1^T \quad \mathbf{H}_2^T \quad \cdots \quad \mathbf{H}_{i-1}^T\right], \tag{3.58}$$

and its corresponding SVD as

$$\hat{\mathbf{H}}_i = \hat{\mathbf{U}}_i \hat{\mathbf{\Sigma}}_i \left[\hat{\mathbf{V}}_i^{(1)} \quad \hat{\mathbf{V}}_i^{0}\right]^H. \tag{3.59}$$

If the rank of $\hat{\mathbf{H}}_i$ is equal to $\hat{L}_i$, then $\hat{\mathbf{V}}_i^{(0)}$ contains $M_T - \hat{L}_i$ right singular vectors. As in the BD solution, we force the modulation matrix $\mathbf{F}_i$ to lie in the null space of $\hat{\mathbf{H}}_i$ by setting $\mathbf{F}_i = \hat{\mathbf{V}}_i^{(0)}\mathbf{F}_i'$ for some choice of $\mathbf{F}_i'$. Thereby, the *i*-th user does not see any interference from any subsequent user $(i+1,\ldots,K)$.

The combination of SO and THP improves the use of the available subspace of different users and eliminates any residual MUI. The resulting equivalent channel matrix is also block diagonal, which facilitates the definition of an ordering algorithm of the users [SH05], [SHF04].

The combination of SO and THP (SO THP) is performed by successively calculating first BD, then the reordering of users, and in the end precoding with THP. Here, instead of examining all $K!$ ordering possibilities that minimise the total capacity loss in the system, we make the heuristic simplification to minimise the capacity loss of each user separately.

In short, we first calculate the maximum capacity that an individual user can achieve. Then, we identify the user with the smallest difference between its maximum capacity and its BD capacity and generate its precoding matrix such that it lies in the null space of the remaining users' channel matrices. Afterwards, we form the new combined channel matrix without this user's channel matrix. We repeat these steps until the combined channel matrix is empty. The order of the users in which they are precoded using THP is the reverse of the order in which their precoding matrices are generated.

With the reordering of the users in the reverse order of precoding we achieve that the equivalent combined channel matrix after precoding and demodulation is lower block diagonal with the singular values on the main diagonal. The lower triangular feedback matrix, used in THP precoding, is generated from this equivalent combined channel matrix after the elements in each row are divided by the elements on the main diagonal, i.e., the corresponding singular values.

In Figure 3.20 we show the block diagram of the SO THP system. The individual user's channel matrices and demodulation matrices are grouped in matrices $\mathbf{H}$ and $\mathbf{D}$. The feedback matrix $\mathbf{B}$, generated in the last step of the SO THP algorithm is now used to precode the users' data streams starting with the data stream of the first user whose precoding matrix $\mathbf{F}_1$ had been generated last.

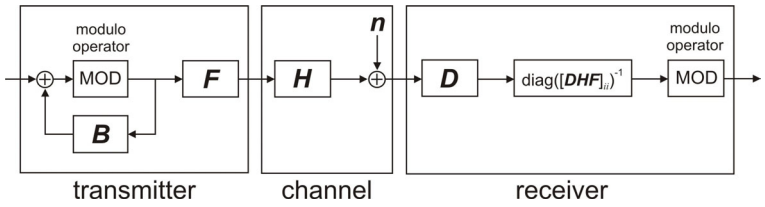


Figure 3.20: Block diagram of the SO THP system.

By using THP at the transmit side we significantly increase the transmit power. Therefore we introduce the modulo operator at the transmitter and the receiver in order to reduce the constellation size into certain boundaries. Before applying the modulo operator at the receiver we divide each spatial stream by the cor-

responding singular value so that the constellation boundaries at the receiver are the same as at the transmitter. The ordering algorithm described here forces the modulation matrix of the user that in the current step has minimum capacity loss to lie in the null space of the remaining users' channel matrices.

for $i = 1 : K$
  $\boldsymbol{H}_i = \boldsymbol{U}_i \boldsymbol{\Sigma}_i \left[\begin{array}{cc} \boldsymbol{V}_i^{(1)} & \boldsymbol{V}_i^{(0)} \end{array} \right]^H$;
  $\boldsymbol{F}_{\max,i} = \boldsymbol{V}_i^{(1)}$;
  $C_{\max,i} = \log_2 \det \left(\boldsymbol{I} + \boldsymbol{R}_{n,i}^{-1} \boldsymbol{H}_i \boldsymbol{F}_{\max,i} \boldsymbol{F}_{\max,i}^H \boldsymbol{H}_i^H \right)$;
end;
$\mathrm{S} = \{1, \ldots, K\}$;
$\boldsymbol{G} = \boldsymbol{H}$;
for $i = K : 1$
  $\left[\begin{array}{cccccc} \boldsymbol{P}_1, & \ldots & \boldsymbol{P}_i, & \boldsymbol{U}_1, & \ldots & \boldsymbol{U}_i \end{array} \right] = \mathrm{BD}(\boldsymbol{G})$;
  for $k = 1 : i$
    $C_k = \log_2 \det \left(\boldsymbol{I} + \boldsymbol{R}_{n,k}^{-1} \boldsymbol{H}_k \boldsymbol{P}_k \boldsymbol{P}_k^H \boldsymbol{H}_k^H \right)$;
  end;
  $k_i = \arg\min_{k \in S} \left(C_{\max,k} - C_k \right)$;
  $\boldsymbol{F}_i = \boldsymbol{P}_{k_i}$;
  $\boldsymbol{D}_i = \boldsymbol{U}_{k_i}^H$;
  $\mathrm{S} = \mathrm{S} \backslash \{k_i\}$;
  $\boldsymbol{G} = \left[\begin{array}{cccccc} \boldsymbol{H}_1^T & \ldots & \boldsymbol{H}_{k_i-1}^T & \boldsymbol{H}_{k_i+1}^T & \ldots & \boldsymbol{H}_K^T \end{array} \right]^T$;
end;
$\boldsymbol{F} = \left[\begin{array}{ccc} \boldsymbol{F}_1 & \ldots & \boldsymbol{F}_K \end{array} \right]$;
$\boldsymbol{D} = \left[\begin{array}{ccc} \boldsymbol{D}_1 & & \\ & \ddots & \\ & & \boldsymbol{D}_K \end{array} \right]$;
$\boldsymbol{B} = \text{lower triangular} \left(\boldsymbol{DHF} \cdot \mathrm{diag} \left(\left[\boldsymbol{DHF} \right]_{ii}^{-1} \right) \right)$;

Figure 3.21: Summary of the SO THP algorithm.

3.3.1.1 Evaluation

We evaluate the performance of the following multi-user precoding techniques: BD, SMMSE, SO THP and MMSE THP. The channel $\mathbf{H}$ is assumed to be spatially white and flat fading. We will use the notation $\{M_{R_i}, \ldots, M_{R_K}\} \times M_T$ to describe the antenna configuration of the system, where M_{R_i} is the number of receive antennas at the i-th mobile terminal and M_T is the number of receive antennas at the base station.

In Figure 3.22 we compare the performance of BD, SMMSE, and SO THP when there is perfect CSI available at the transmitter. We assume that there are two users in the system receiving either one data stream using dominant eigenmode transmission (DET) or multiple data streams (MDS). With BD and SO THP we use water-filling (WF) and with SMMSE no power loading (NoPL). The number of antennas at the base station is set to 4. SMMSE has a clear advantage over BD at low SNRs while BD can provide higher capacity with MDS at high SNRs. However, BD is limited to cases when the total number of receive antennas at the MSs is less or equal to the total number of transmit antennas. From this figure we can see that when the users are equipped with 4 antennas each, SMMSE improves the capacity at some SNRs by more than 4 bits/sec/Hz. The receive signal-to-noise ratio is defined as $\mathrm{SNR}_r = P_T / \sigma_n^2$, where σ_n^2 is the variance of the white additive Gaussian noise.

The uncoded BER performance of SO THP, MMSE THP, SMMSE and BD precoding techniques are shown in Figure 3.23. We compare the performance of these techniques in case when there are three users in the system each equipped with two antennas. The number of antennas at the base station is 6. BD is a linear precoding technique that has a zero MUI interference constraint. By introducing MUI, SMMSE provides both, a higher diversity and array gain than BD. SO THP does not have the same diversity gain as BD that can be explained with the influence of the modulo operator used in THP. The transmit signal-to-noise ratio is defined as $\mathrm{SNR}_t = P_T / \left(M_R \sigma_n^2 \right)$.

In Figure 3.23 we show the performance of SMMSE when there is only long-term CSI available at the transmitter. SMMSE offers the advantage of having no dimensionality restrictions and it provides a higher capacity using more receive antennas at the MSs. Also, in this case we can see that the optimum power loading strategy would be to transmit only on the dominant eigenmode, which is in agreement with the results previously published [PNG03].

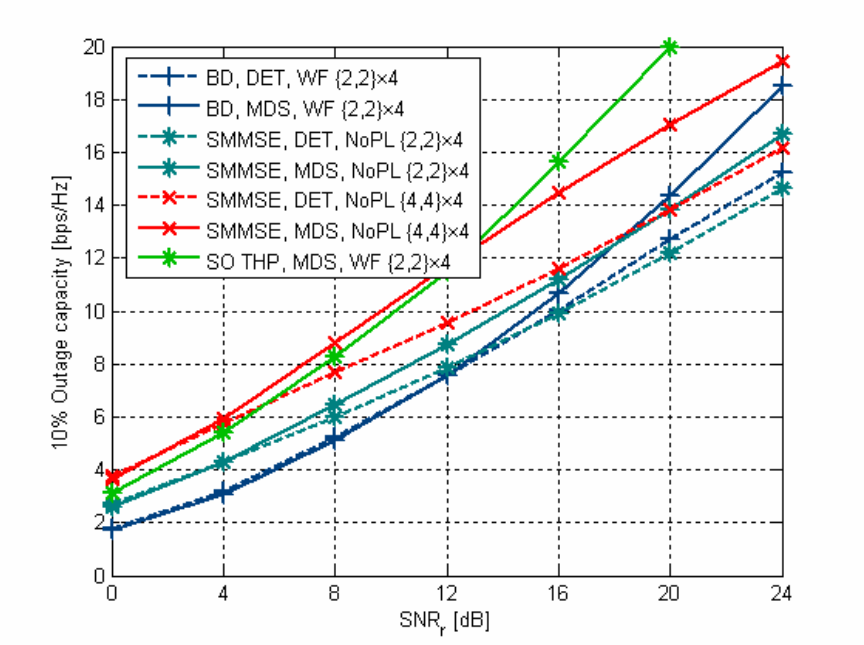


Figure 3.22: 10% outage capacity as a function of the receive SNR_r.

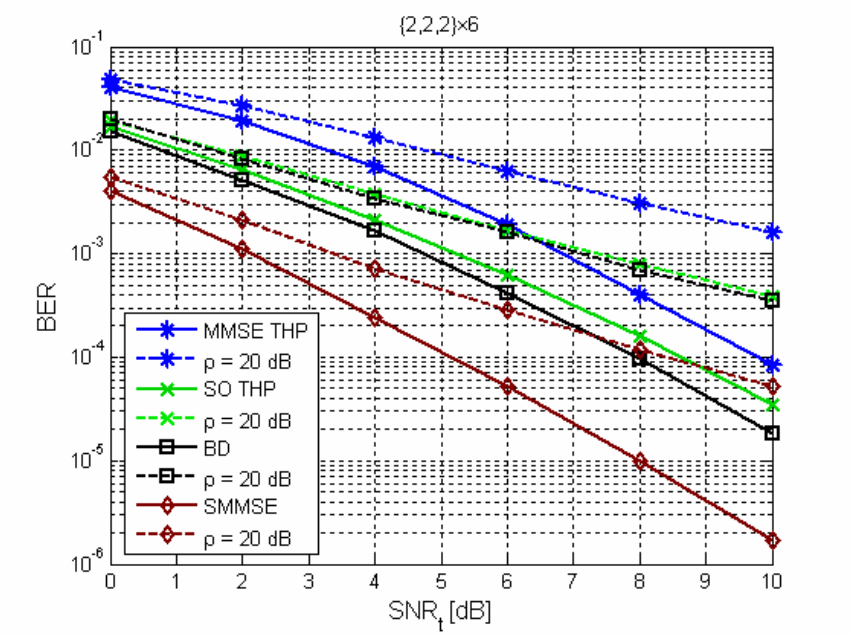


Figure 3.23: BER performance comparison of MMSE THP in configuration $\{1,1,1,1,1,1\}\times 6$ and BD, SMMSE, and SO THP in configuration $\{2,2,2\}\times 6$

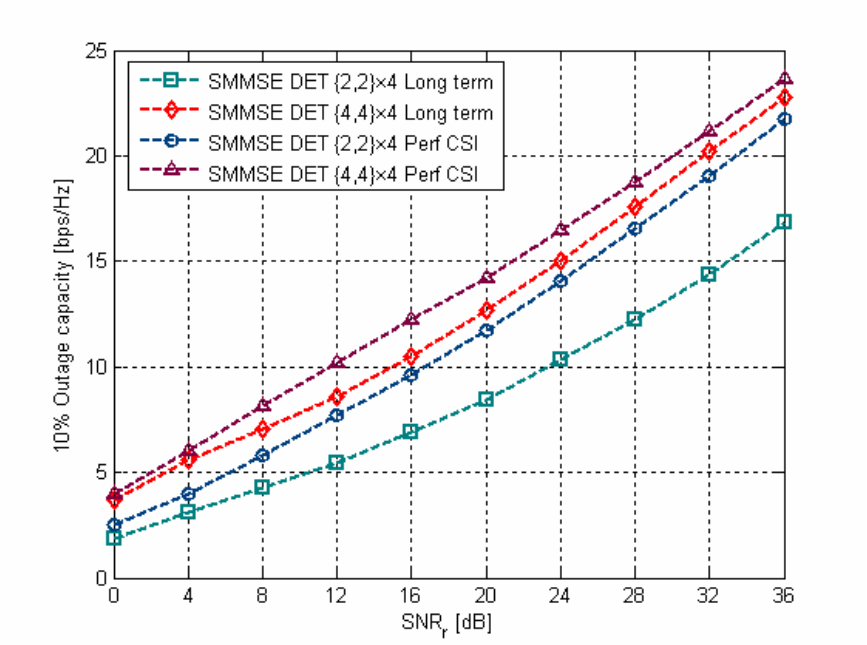


Figure 3.24: 10 % outage capacity with long-term and perfect CSI at the transmitter as a function of the receive SNR_r DET – Dominant Eigenmode Transmission

3.3.1.2 Summary

Depending on the set of constraints, like the size of the overhead or the amount of the MUI allowed, different techniques can be optimal. Linear techniques are computationally less expensive and generally require minimum signalling overhead. On the other hand, the non-linear techniques can provide better performance. Block diagonalisation (BD) is a linear pre-coding technique for the downlink of MU MIMO systems. It decomposes a MU MIMO downlink channel into multiple parallel independent single-user MIMO downlink channels. BD is attractive if the users are equipped with more than one antenna. However, the zero MUI constraint can lead to a significant capacity loss when users' subspaces significantly overlap. By combining a linear pre-coding technique called successive optimisation and a non-linear technique called Tomlinson-Harashima precoding we are able to completely eliminate MUI in the system when there is perfect CSI available at the transmitter. The equivalent channel matrix is block diagonal after precoding. This technique is especially attractive in cases when the users and the base station/access point are equipped with multiple antennas. In these cases SO THP provides higher capacity than BD or MMSE. By transmitting only on the dominant eigenmodes of each user, SO THP can provide a better BER performance than MMSE THP for low SNR ratios. SMMSE precoding reduces the performance loss due to the zero MUI constraint and the cancellation of the interference between the antennas located at the same terminal. SMMSE outperforms BD, another linear precoding technique that also performs well when the users are equipped with multiple antennas but with zero MUI. Moreover, in a system, where all users are equipped with multiple antennas, SMMSE outperforms both SO THP and MMSE THP that are non-linear precoding techniques. SMMSE has relatively low computational complexity and reduces the overhead needed for demodulation. For high SNR ratios it also results in a block diagonal combined network channel matrix and it can be combined with any other previously proposed precoding technique. Another big advantage of SMMSE is that the users can be equipped with more antennas, i.e., the total number of receive antennas in the downlink can be greater than the number of transmit antennas. Both SMMSE and BD can be used also with long-term CSI. The main advantage of SMMSE is again no dimensionality limitation.

3.3.2 Joint THP with Diversity Techniques

In the literature and in the preceding section, THP has been shown to be a powerful means to eliminate the multi-user interference. With multi-antenna systems, diversity techniques have been shown to be an effective mean to combat channel fading. Orthogonal space-time block codes (OSTBC) have been demonstrated to be a powerful diversity technique. In particular, the Alamouti code [[Ala98]] offers very

simple encoding and decoding. Conventional OSTBC code does not require channel state information (CSI) at transmitter side. When CSI is available at transmitter side, we show that nonlinear Tomlinson-Harashima precoding can work jointly with OSTBC. Dominant eigenmode transmission (DET) is another diversity technique that can achieve better diversity gain than OSTBC [CM04] when CSI is known both to transmitter and receivers. In this work, we will demonstrate that DET and OSTBC can be combined with THP to improve the system performance in terms of BER. Furthermore, we show that THP+DET exhibits better performance than OSTB+THP.

THP+DET corresponds to SO THP described in the previous section, where it was investigated for different design criteria. For maximising the information rate, SO THP should be used in conjunction with water-filling over all modes. To maximise the SNR at the receiver, SO THP should be used in conjunction with DET, which is called THP+DET in this section. Also, some other power loading schemes already proposed in the literature could be used in conjunction with SO THP, such as MMSE or minimum BER design approaches [SSB+02], [SSP01].

Although the OSTBC is designed for no channel information at transmitter, it is a kind of diversity technique that takes advantage of block triangularisation of the channel and serves therefore perfectly as a basic example. Of course, the simulations show that the dominant eigenmode transmission (DET) works better than OSTBC when they both are combined with the THP scheme. However, with respect to the complexity of the two schemes, the THP+DET is more complex than THP+OSTBC. In case of DET a singular value decomposition is needed at both the receiver and the transmitter whereas only linear combining is needed in the case of OSTBC. And in this sense THP+OSTBC is another optional technique that can be used for taking use of the diversity of the block triangular channel.

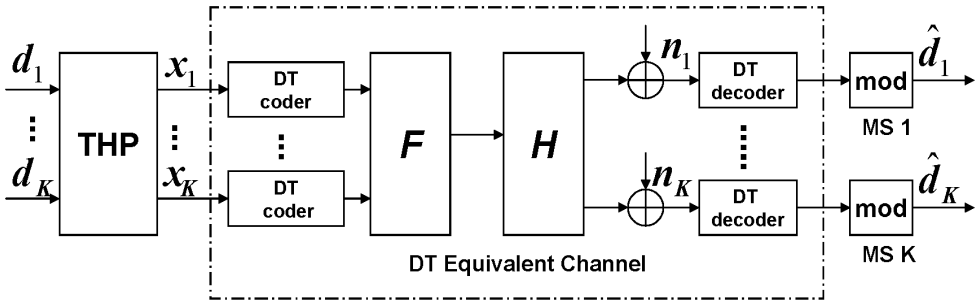


Figure 3.25: Diagram of THP joint with Diversity Techniques (DT)

The proposed diversity techniques (DT) joint with THP is plotted in Figure 3.25. In Figure 3.25 the forward filter $\boldsymbol{F}$ is intended for forming the block triangular channel so that THP can be applied. Taking the advantage of the block triangular channel structure, we combine THP with diversity techniques, which greatly improve the system performance. With our proposed scheme, we assume that the system is equipped with N_t transmit antennas, and there are K users communicating over the frequency selective channel (over N_c subcarriers) with it or flat channel simultaneously. The kth user is equipped with N_k (k=1,…,K) antennas and in total there are N antennas at the receivers. We further assume that N_t is greater or equal to N. The proposed strategy can be applied to frequency flat channels or frequency selective channels. The block diagram of our proposed scheme is for one subcarrier of an OFDM system or for a complete single carrier system.

Let us see how THP+DT works. Let d_k denote the $N_k \times 1$ transmit data symbol vector intended for user k and the stacked vector of K users is $\boldsymbol{d} = [d_1,...,d_K]^T$. Taking use of the known CSI, the THP coder first precodes the $\boldsymbol{d}$ to $\boldsymbol{x}$, and then $\boldsymbol{x}$ is taken by the DT coder and the space-time coded vector $\boldsymbol{X}$ is formed. $\boldsymbol{X}$ is further processed by a forward matrix $\boldsymbol{F}$ before it is launched into the MIMO channel through the N_t antennas at BS. The forward matrix $\boldsymbol{F}$ is intended for the THP scheme and it is designed such that $\hat{\boldsymbol{H}}=\boldsymbol{HF}$ is of lower block triangular form. The idea behind generating the lower block triangular channel is that the first user gets no interference from others, and the second user is interfered only by the first user and the third user has interference only from the first and the second users and so on. Since the CSI is well known by the BS, the interference for a certain user can be pre-subtracted before each data block is transmitted for the users respectively. The unitary property of the forward matrix $\boldsymbol{F}$ guarantees the transmit power follows the constraint. The design of $\boldsymbol{F}$ will be explained in the following section. The channel

between BS and user k is indicated by $\boldsymbol{H}_k$. Stacking the K users' channels into $\boldsymbol{H}$, the whole transmission process can be expressed by $\boldsymbol{y} = \boldsymbol{HFX} + \boldsymbol{n}$, where additive white Gaussian noise (AWGN) $\boldsymbol{n}$ follows with complex normal distribution $\mathcal{CN}(0,\sigma^2\mathbf{I}_K)$ and $\mathbf{I}_K$ is the identity matrix. Throughout the following discussion, we assume the BS has obtained the CSI either by recursive channel estimation in time-division-duplex (TDD) or feedback in frequency-division-duplex (FDD). The mobile users acquire CSI through an appropriate signalling channel (e.g., pilots). We further assume that the channel is quasi-static over a time of one data block transmission.

3.3.2.1 Forward Filter Design

To take the advantage of the Tomlinson-Harashima precoding of eliminating multi-user access interference, our first aim is to find a unitary forward matrix $\boldsymbol{F}$ that is multiplied by a channel matrix $\boldsymbol{H}$ to obtain a lower block triangular matrix. Let $\boldsymbol{H}_k$ denote the channel between user k and BS. The stacked channel can be expressed as $\boldsymbol{H} = [\boldsymbol{H}_1^T, \ldots, \boldsymbol{H}_K^T]^T$. When we ignore the K-th user channel, the remainder of $\boldsymbol{H}$ denotes as

$$\boldsymbol{H}^{(K)} = [\boldsymbol{H}_1^T, \boldsymbol{H}_2^T, \ldots, \boldsymbol{H}_{K-1}^T]^T. \tag{3.60}$$

Applying the singular value decomposition (SVD) to equation (3.60), we get

$$\boldsymbol{H}^{(K)} = \boldsymbol{U}_K [\boldsymbol{\Lambda}_K, \mathbf{0}] \boldsymbol{V}_K^H. \tag{3.61}$$

Note that $\boldsymbol{V}_K$ is an $N \times N$ unitary matrix, $\boldsymbol{\Lambda}_K$ is a $(N-N_K)\times(N-N_K)$ matrix. Multiply $\boldsymbol{V}_K$ with $\boldsymbol{H}$ in order to obtain $\hat{\boldsymbol{H}}^K$. Note that except for the last N_K rows, the 1 to $N-N_K$ rows' last N_K columns of $\hat{\boldsymbol{H}}^K$ are zeros. Then choose the elements of the first $N-N_K-N_{K-1}$ rows of $\hat{\boldsymbol{H}}^K$ as the new matrix $\boldsymbol{H}^{(K-1)}$. Applying SVD to $\boldsymbol{H}^{(K-1)}$, we get

$$\boldsymbol{H}^{(K-1)} = \boldsymbol{U}_{K-1} [\boldsymbol{\Lambda}_{K-1}, \mathbf{0}] \boldsymbol{V}_{K-1}^H. \tag{3.62}$$

Notice $\boldsymbol{V}_{K-1}$ is still an $N \times N$ unitary matrix, $\boldsymbol{U}_{K-1}$and $\boldsymbol{\Lambda}_{K-1}$ are $(N-N_K-N_{K-1})\times(N-N_K-N_{K-1})$ matrices. Multiply $\boldsymbol{V}_{K-1}$ with $\hat{\boldsymbol{H}}^K$, we get a new $\hat{\boldsymbol{H}}^{K-1}$ and repeat the steps as we have performed from $\hat{\boldsymbol{H}}^K$ to $\hat{\boldsymbol{H}}^{K-1}$ above. At each step, one user channel matrix size is reduced and SVD is applied to the reduced $\boldsymbol{H}^{(k)}$ $k=2,\ldots,K$ until to the first user, finally we can get $\boldsymbol{V}_{2,\ldots,}\boldsymbol{V}_{K,}$. Since $\boldsymbol{V}_2$ to $\boldsymbol{V}_K$ are unitary matrices, it is easy to prove that $\boldsymbol{F} = \boldsymbol{\Pi}_{i=2}^{K} \boldsymbol{V}_i$ is a unitary matrix. Summarily the forward matrix $\boldsymbol{F}$ can be obtained by the following algorithm:

$\boldsymbol{F} = \boldsymbol{I}$; $\boldsymbol{H}_t = \boldsymbol{H}$;

for k = K step:-1 to 2

S = sum row size of $\boldsymbol{H}_1$ to $\boldsymbol{H}_{k-1}$; /* channel size k to K is ignored */

[$\boldsymbol{U}$ $\boldsymbol{D}$ $\boldsymbol{V}$] =svd ($\boldsymbol{H}_t$ (1:S,:));

$\boldsymbol{H}_t = \boldsymbol{H}_t * \boldsymbol{V}$;

$\boldsymbol{F} = \boldsymbol{F} * \boldsymbol{V}$;

end

The algorithm follows the Matlab notation and the diagonal $\boldsymbol{D} = [\boldsymbol{\Lambda}, \mathbf{0}]$ is assumed. With the above decomposition steps, we have assumed that the fixed ordering - the data for the first user is precoded first, and the data for the K-th user is precoded last. Actually we can permute the stacked channel $\boldsymbol{H}$ regarding to $\boldsymbol{H}_k$, $k=1,\ldots,K$ to obtain $\boldsymbol{F}$ such that the users' data is precoded in arbitrary user order. In fact the precoding order of users affects the overall performance.

3.3.2.2 OSTBC Equivalent Channel Formulation

The main idea behind combining OSTBC with nonlinear THP is to treat space-time coder, channel and space-time decoder as an equivalent channel, which we call “OSTBC equivalent channel”. Shortly we will show that this equivalent channel is of a lower block triangular matrix form and it is the dashed box shown in Figure 3.25. We will use the Alamouti two branches transmit diversity scheme to derive the OSTBC equivalent channel model; however this derivation can be extended to any OSTBC. To derive the equivalent channel, a system with four antennas at BS and two users each having two antennas will be

presented as an example. The data $\boldsymbol{x}= [\boldsymbol{x}_1^{\mathrm{T}}, \boldsymbol{x}_2^{\mathrm{T}}]^{\mathrm{T}}$. from THP arriving at the OSTBC coder will be coded to be transmitted in two time slots, and we can stack the two time slots of transmitted data into one matrix as

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x}_{11} & -\boldsymbol{x}_{12}^{*} \\ \boldsymbol{x}_{12} & \boldsymbol{x}_{11}^{*} \\ \boldsymbol{x}_{21} & \boldsymbol{x}_{22}^{*} \\ \boldsymbol{x}_{22} & \boldsymbol{x}_{11}^{*} \end{bmatrix}. \tag{3.63}$$

After the coded data $\boldsymbol{X}$ is multiplied by the forward matrix $\boldsymbol{F}$, it is launched into the channel. We know that the result of $\boldsymbol{F}$ multiplied by $\boldsymbol{H}$ is a lower block-triangle matrix, in other words, the coded data $\boldsymbol{X}$ passes an equivalent lower block triangular channel, which is of the form

$$\boldsymbol{H} = \begin{bmatrix} \hat{h}_{11} & \hat{h}_{12} & 0 & 0 \\ \hat{h}_{21} & \hat{h}_{22} & 0 & 0 \\ \hat{h}_{31} & \hat{h}_{32} & \hat{h}_{33} & \hat{h}_{34} \\ \hat{h}_{41} & \hat{h}_{42} & \hat{h}_{43} & \hat{h}_{44} \end{bmatrix}. \tag{3.64}$$

Because of the lower block triangular channel structure, it is easy to see that the first user has no interference from the second user and the second user has interference from the first user. Decoding the first user data is just a regular Alamouti decoding process. Let us focus on the second user. For simplicity, the noise element is ignored. Following notation used in [Ala98] , the second user receives

$$\begin{array}{ll} y_{20} = \hat{h}_{31}x_{11} + \hat{h}_{32}x_{12} + \hat{h}_{33}x_{21} + \hat{h}_{34}x_2 & y_{22} = \hat{h}_{41}x_{11} + \hat{h}_{42}x_{12} + \hat{h}_{43}x_{21} + \hat{h}_{44}x_2 \\ y_{21} = -\hat{h}_{31}x_{12}^{*} + \hat{h}_{32}x_{11}^{*} - \hat{h}_{33}x_{22}^{*} - \hat{h}_{34}x_{21}^{*} & y_{21} = -\hat{h}_{41}x_{12}^{*} + \hat{h}_{42}x_{11}^{*} - \hat{h}_{43}x_{22}^{*} - \hat{h}_{44}x_{21}^{*} \end{array}. \tag{3.65}$$

After the Alamouti decoding process, we obtain

$$\begin{array}{l} s_{20} = \hat{h}_{33}^{*}y_{20} + \hat{h}_{34}y_{21}^{*} + \hat{h}_{43}^{*}y_{22} + \hat{h}_{44}y_{23}^{*} = \lambda_{31}x_{11} + \lambda_{32}x_{12} + \lambda_2 x_{21} \\ s_{21} = \hat{h}_{34}^{*}y_{20} - \hat{h}_{33}y_{21}^{*} + \hat{h}_{44}^{*}y_{22} - \hat{h}_{43}y_{23}^{*} = \lambda_{41}x_{11} + \lambda_{42}x_{12} + \lambda_2 x_{22} \end{array}. \tag{3.66}$$

The λ elements in (3.66) are calculated with the elements in (3.64). Recall user 1 and the matrix form of (3.66) can be rewritten as

$$\begin{bmatrix} s_{10} \\ s_{11} \\ s_{20} \\ s_{21} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ \lambda_{31} & \lambda_{32} & \lambda_2 & 0 \\ \lambda_{41} & \lambda_{42} & 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \\ s_{21} \\ s_{22} \end{bmatrix}. \tag{3.67}$$

Λ_{kk} (k = 1,2) are diagonal matrices having equal λ_k entries along the diagonal, Λ_{12} is a non-diagonal matrix and $\boldsymbol{0}$ is with only zeros entries. The preceding scheme can be extended to arbitrary OSTBC. Denote $\boldsymbol{\Lambda}$ as equivalent channel between the transmitter THP block and receivers' modulo blocks shown in Figure 3.25 with K users' case, and $\boldsymbol{\Lambda}$ can be expressed as

$$\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\Lambda}_{11} & 0 & 0 & 0 \\ \boldsymbol{\Lambda}_{21} & \boldsymbol{\Lambda}_{22} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Lambda}_{K1} & \boldsymbol{\Lambda}_{K2} & \cdots & \boldsymbol{\Lambda}_{KK} \end{bmatrix}, \tag{3.68}$$

where $\boldsymbol{\Lambda}_{\mathrm{kk}}$ (k=1,...,K) are diagonal matrices with equal entries λ_{k} along the diagonals. The $\boldsymbol{\Lambda}_{\mathrm{kk}}$ can be looked at as effective channel for user k and $\boldsymbol{\Lambda}_{\mathrm{ki}}$ (k ≠ i) can be seen as the interference from user 1 to k-1. Equivalently, a lower block triangular channel exists between transmitter's THP block and the receivers' modulo blocks in Figure 3.25. We call this equivalent channel *OSTBC equivalent channel.*

3.3.2.3 DET Equivalent Channel Formulation

As mentioned above, the forward matrix is intended for forming the channel to a block triangular channel. Now let us see how the diversity gain of this block triangular structure can be combined with DET. Here

still a two users' case is given as an example, with the block triangular channel

$$\boldsymbol{HF} = \begin{bmatrix} \hat{h}_{11} & \hat{h}_{12} & 0 & 0 \\ \hat{h}_{21} & \hat{h}_{22} & 0 & 0 \\ \hat{h}_{31} & \hat{h}_{32} & \hat{h}_{33} & \hat{h}_{34} \\ \hat{h}_{41} & \hat{h}_{42} & \hat{h}_{43} & \hat{h}_{44} \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{H}}_{11} & \boldsymbol{0} \\ \hat{\boldsymbol{H}}_{21} & \hat{\boldsymbol{H}}_{22} \end{bmatrix}, \tag{3.69}$$

$\hat{\boldsymbol{H}}_{11}$ and $\hat{\boldsymbol{H}}_{22}$ can be looked as effective channels for user 1 and user 2 respectively, and $\hat{\boldsymbol{H}}_{21}$ is taken by user 2 as the interference from user 1. Using the effective channels, we can apply the DET to user 1 and user 2 respectively by using singular value decomposition.

First we find out the right vectors $\boldsymbol{v_{m1}}$ and $\boldsymbol{v_{m2}}$ and the left vectors $\boldsymbol{u_{m1}}$ and $\boldsymbol{u_{m2}}$ of $\hat{\boldsymbol{H}}_{11}$ and $\hat{\boldsymbol{H}}_{22}$, which correspond to the maximum singular values of $\hat{\boldsymbol{H}}_{11}$ and $\hat{\boldsymbol{H}}_{22}$ respectively. Multiply $\boldsymbol{v_{m1}}$ and $\boldsymbol{v_{m2}}$ with the transmission scalar data x_1 and x_2 to form the transmit vectors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. Stack the $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ to the transmit vector $\boldsymbol{x}$, and after further processed by vector $\boldsymbol{F}$, $\boldsymbol{Fx}$ is launched to the channel. At the receiver, $\boldsymbol{u_{m1}}$ and $\boldsymbol{u_{m2}}$ are applied to sum weight the received vectors $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$. Then the equivalent transmission can be expressed as

$$\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{H}}_{11} & \boldsymbol{0} \\ \hat{\boldsymbol{H}}_{21} & \hat{\boldsymbol{H}}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix} \Rightarrow \begin{bmatrix} \boldsymbol{u}_{m1}^H \boldsymbol{y}_1 \\ \boldsymbol{u}_{m2}^H \boldsymbol{y}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{u}_{m1}^H \hat{\boldsymbol{H}}_{11} \boldsymbol{v}_{m1}^H & \boldsymbol{0} \\ \boldsymbol{u}_{m2}^H \hat{\boldsymbol{H}}_{21} \boldsymbol{v}_{m2}^H & \boldsymbol{u}_{m2}^H \hat{\boldsymbol{H}}_{22} \boldsymbol{v}_{m2}^H \hat{\boldsymbol{H}}_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{3.70}$$

Here, the noise is ignored for simplicity. After some manipulation, equation (3.70) can be further simplified as

$$\begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} \xi_{m1} & 0 \\ \xi_{21} & \xi_{m2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{3.71}$$

With the triangular channel structure, the THP scheme can be applied. Equation (3.71) is derived with only two users each having two antennas, however it can be generalised to the K users case and each user can have an arbitrary number of antennas. In the K users case the equivalent transmission can be expressed as

$$\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_K \end{bmatrix} = \begin{bmatrix} \xi_{m1} & 0 & 0 & 0 \\ \xi_{21} & \xi_{m2} & 0 & 0 \\ \vdots & \cdots & \ddots & 0 \\ \xi_{K1} & \xi_{K2} & \cdots & \xi_{mK} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_K \end{bmatrix}. \tag{3.72}$$

3.3.2.4 THP with Regard to the DT Equivalent Channel

First, let us take the THP+OSTBC scheme into account. When precoding is not applied, after the OSTBC decoding the k user receives $\boldsymbol{\Lambda}_{\mathrm{kk}}\boldsymbol{d}_{\mathrm{k}}$ and interference $\sum_{i=1}^{k-1}\boldsymbol{\Lambda}_{ki}\boldsymbol{d}_i$ and noise. Since the transmitter knows the channel and the data intended for the users, the interference is pre-subtracted before the symbols being launched to the DT (OSTBC) coder $\boldsymbol{x}_k = \boldsymbol{d}_k - \sum_{i=1}^{k-1}\boldsymbol{\Lambda}_{kk}^{-1}\boldsymbol{\Lambda}_{ki}\boldsymbol{x}_i$ with $\boldsymbol{x}_1=\boldsymbol{d}_1$. Because the pre-subtraction operation increases the transmit power significantly, modulo-A operation is applied to the x_k, where A is defined as $A=\sqrt{M}$ and M is the size of M-ary QAM constellation. Note that modulo operation is applied to the real and imaginary parts of each entry of vector x_k respectively and the same operation will apply to the received data at the receiver side. With the modulo operation, the transmit vectors become

$$\boldsymbol{x}_k = \left(\boldsymbol{d}_k - \sum_{i=1}^{k-1}\boldsymbol{\Lambda}_{kk}^{-1}\boldsymbol{\Lambda}_{ki}\boldsymbol{x}_i\right) \bmod A \triangleq \boldsymbol{d}_k - \sum_{i=1}^{k-1}\boldsymbol{\Lambda}_{kk}^{-1}\boldsymbol{\Lambda}_{ki}\boldsymbol{x}_i + \boldsymbol{A}_k, \tag{3.73}$$

with $\boldsymbol{A}_{\mathrm{k}}$ a vector of the same size as $\boldsymbol{x}_{\mathrm{k}}$ and its real and imaginary entries are integer times of A. $\boldsymbol{A}_{\mathrm{k}}$ can be considered as power constraint reduction parameter, which varies according to the data block and channel. At the receiver side, the data received by user k become

$$\boldsymbol{s}_k = \boldsymbol{\Lambda}_{kk}\boldsymbol{d}_k - \boldsymbol{\Lambda}_{kk}\sum_{i=1}^{k-1}\boldsymbol{\Lambda}_{kk}^{-1}\boldsymbol{\Lambda}_{ki}\boldsymbol{x}_i + \sum_{i=1}^{k-1}\boldsymbol{\Lambda}_{ki}\boldsymbol{x}_i + \boldsymbol{\Lambda}_{kk}\boldsymbol{A}_k. \tag{3.74}$$

After the modulo-A operation on $\Lambda_{kk}^{-1}\, \boldsymbol{x}_{k,}$ the user k recovers the desired data $\boldsymbol{d}_k$.

Second, with DET equivalent transmission in (3.72), the precoding process is simple and the precoded data can be expressed as

$$x_k = \left(d_k - \sum_{i=1}^{k-1} \xi_{mk}^{-1} \xi_{ki} x_i\right) \bmod A, \text{ with } x_1 = d_1, \tag{3.75}$$

At receiver, after sum weighting on the receiver vector, the same modulo operation will recover the intended data.

3.3.2.5 Simulation Evaluation

Simulation results are provided in this section to evaluate the performance of our proposed downlink transmission scheme in terms of BER. OFDM systems with QAM modulation are simulated. The 16 taps MIMO channels with exponentially decreasing power profile used for simulation are of i.i.d. zero mean complex Gaussian random entries with unit variance. Perfect CSI is assumed available both BS and MSs. The bit error rate (BER) in the figures is the average over all users. The SNR is the average power of the precoded transmission symbols to noise ratio.

Figure 3.26 shows the comparison of a THP system with 6 single-antenna users and a THP+DT system with 3 users having each two antennas. For fair comparison, the transmit rates of the two systems are equal, a 4-QAM modulation scheme is used for the first system and 16-QAM modulation for THP+DT based system. It is shown in Figure 3.26 that with the same transmit rate, the multi-antenna system employing our proposed scheme performs much better than the one-antenna multi-user system using THP scheme only. When we look at the two systems as one system adopting different schemes for transmission, we say the THP system employing diversity techniques has much better performance than the THP system with "spatial multiplexing" under the same transmit rate condition. Figure 3.26 also shows the diversity gain of the two diversity schemes are the same. However, THP+DET shows better performance than THP+OSTBC.

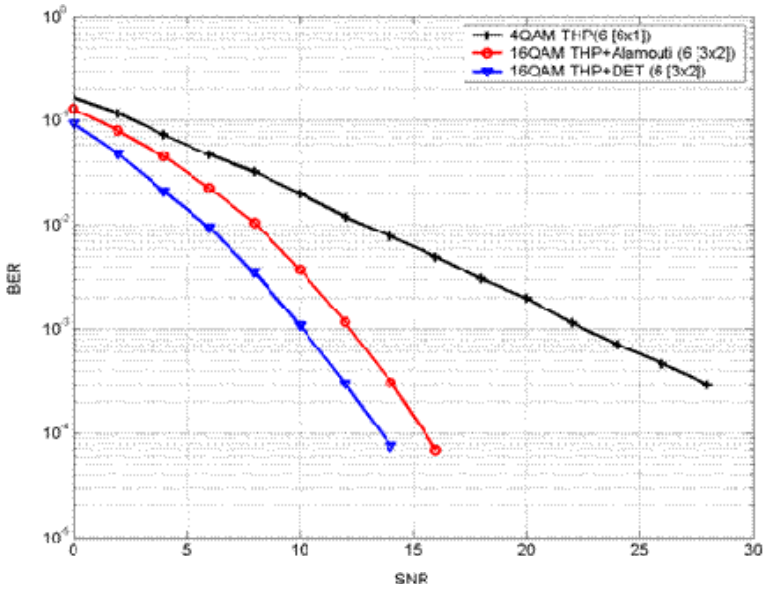


Figure 3.26: Comparison of one-antenna per receiver THP scheme and two-antennas per receiver THP+DT scheme

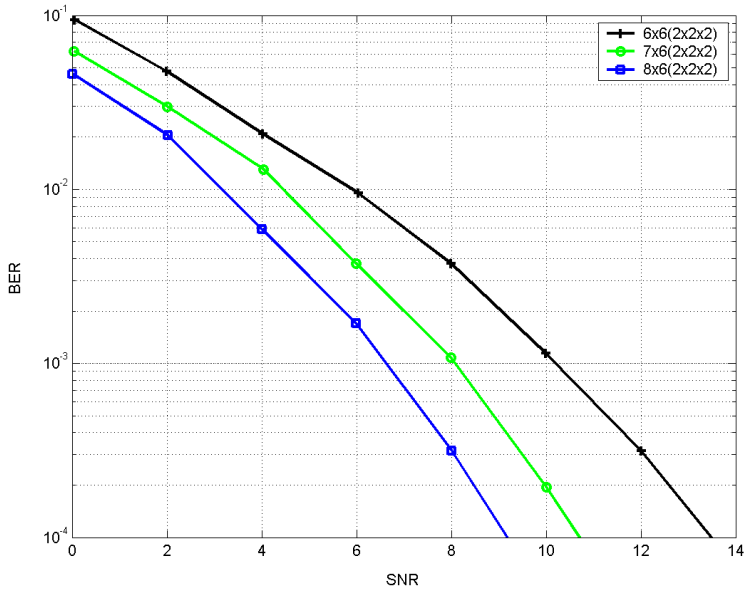


Figure 3.27: The impact of different sizes of antennas at receivers on the THP+DET

The simulation in Figure 3.27 shows the impact of different number of antennas at the base station on the performance of THP+DET system. Three users, each having two antennas and the base station possessing 6, 7, and 8 antennas systems are simulated respectively and 16-QAM modulation is used. By the simulation results, we see in addition to the diversity gain achieved by the diversity technique employed, the additional diversity gain can be achieved by installing more antenna at transmitter. The more transmit antennas are at transmitter side, the more diversity gain can be achieved.

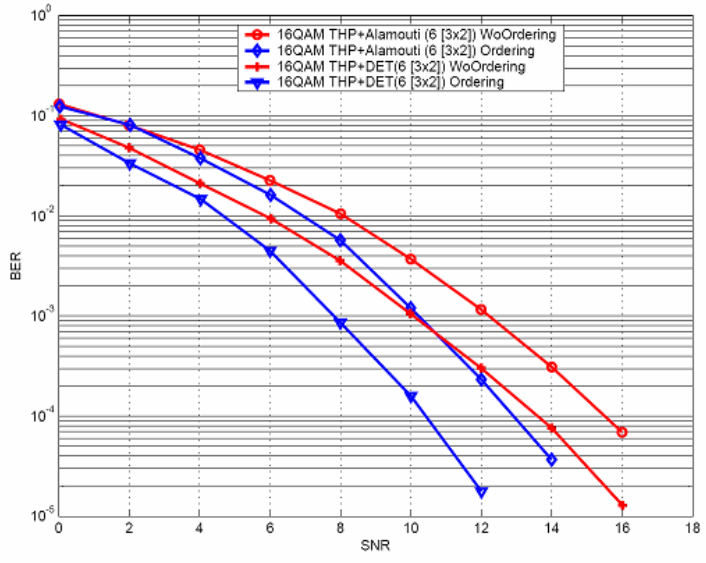


Figure 3.28: Comparison of ordering and without ordering schemes

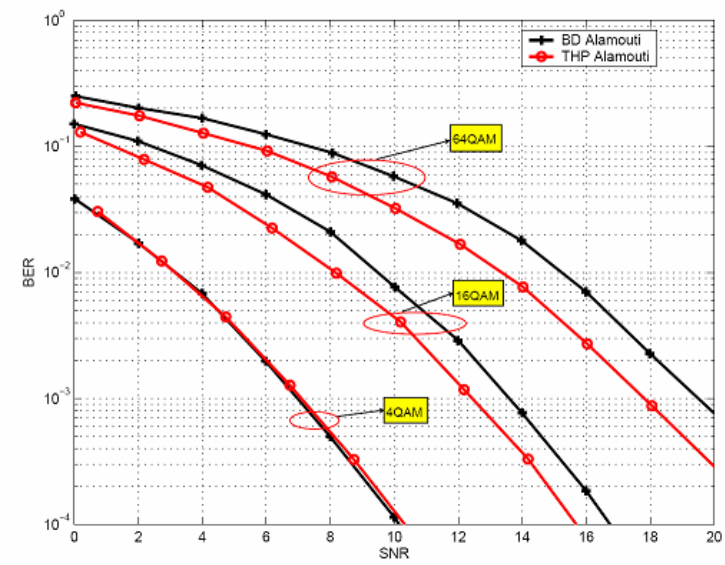


Figure 3.29: Performance comparison of the Alamouti THP and the scheme proposed in [CM04] with Alamouti space-time coding

The simulation result of an optimal ordering scheme is shown in Figure 3.28. The optimal ordering is the ordering that maximises the least element along the diagonal elements of the equivalent channel over all possible $K!$ permutations of K users' channels. About 2 dB gain at BER of 10^{-3} can be observed compared without ordering. And it can be proved that it is optimal ordering at high SNR. In low SNR region, the maximising the sum of the diagonal elements of the equivalent channel is optimal.

The simulation result of Figure 3.29 compares the THP Alamouti scheme with the block diagonal (BD) scheme proposed in [CM04]. For fair comparison, we simulate the block diagonal (BD) scheme with Alamouti space-time code. For a 4-QAM signal constellation, the performance of the both schemes is almost the same. However with the 16-QAM and 64-QAM modulation, the performance of the THP Alamouti scheme is better.

3.3.2.6 Summary

A new scheme that joins THP with DT for downlink of a MIMO multi-user communication system has been proposed. It decomposes the downlink channel to unitary matrix and lower block triangular matrix, which is joint with DT to form an equivalent DT block triangular channel and triangular channel such that THP can be employed. Under the same transmit rate condition, the conducted simulation evaluation shows that multi-user MIMO systems employing THP joint with “diversity technique” scheme has much better performance than the multi-user MIMO system adopting “spatial multiplexing” with THP only. The optimal ordering precoding with regards to maximising the least values among the DT equivalent channel matrix is simulated and analyzed. Compared with the block diagonal scheme in combination with Alamouti proposed in [CM04], the combination of THP and DT shows better performance.

3.3.3 Signal Design for Noncoherent Multi-User MIMO-OFDM

3.3.3.1 Overall Description of the Technique

3.3.3.1.1 Introduction

The assumption of perfect CSI at the receiver in many MIMO data communication schemes is based on a low mobility scenario where the channel stays constant sufficiently long such that the receiver has the opportunity to obtain precise CSI with small or negligible overhead. However, for scenarios with very high vehicle speeds, the accurate acquisition of CSI at the receiver requires frequent retransmission of training sequences prior to data transmission, which results in a significant loss in data throughput. This

loss can be substantial for a MIMO broadband system where a large number of channel coefficients has to be estimated. Therefore, noncoherent design approaches avoiding the need for CSI have recently attracted some interest in the research community. A code design for a single-link was proposed in [GS03] for single-carrier and in [BB04] for multi-carrier systems when the unknown MIMO channel is frequency-selective and in [HMR+00] and [MBV02] and when it is frequency-flat. Signal design for a multi-user communication scenario over unknown flat fading MIMO channels was considered in [BV01]. Several results concerning the available amount of diversity and requirements on efficient signalling exploiting this diversity are presented in the mentioned references. We propose to consider now the generalised problem of noncoherent multi-user communication for multi-antenna systems over unknown frequency-selective channels using OFDM and investigate the requirements on signal design for data communication systems operating in the uplink. The single-link data model for noncoherent MIMO-OFDM from [BB04] is equivalent to a data model for the transmission of pseudocodewords (with implicit structure) in a noncoherent space-time system. We exploit this analogy to make use of results by Brehler and Varanasi [BV01] obtained for noncoherent uplink multi-user space-time communication to design constellations used for space-frequency modulation. With reference to Figure 2.11, we assume that each user employs a user-specific STFCM block with the simplifying assumption that only coding in space and frequency is considered. Assuming that these space-frequency coding and modulation blocks, also referred to as codebooks, allow the exploitation of full space and frequency diversity at the receiver, we obtain a joint performance criterion on the set of codebooks based on an approximate expression (in the high SNR case) of the pairwise detection error probability at the base station. Similar to the approach presented in [GS03] we apply this criterion as a cost function for numerical optimisation of the coding gain provided by the codebooks when full space and frequency diversity gain is exploited. Numerical simulations in an example setup illustrate the potential performance gain due to multi-user signal design.

3.3.3.1.2 System Model

An illustration of the uplink system model is presented in Figure 3.30.

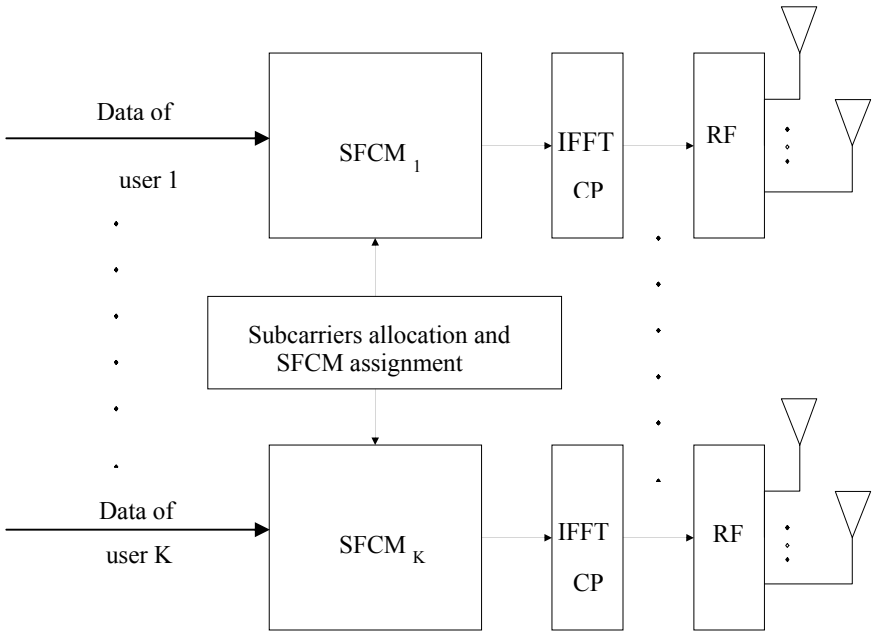


Figure 3.30: Uplink with K users applying user-specific space-frequency coding and modulation

The data of each user is mapped to complex symbols on each available carrier and antenna using a user-specific codebook (illustrated in Figure 3.30 with the SFCM blocks). The mathematical formulation of this system model is an extension of the single-link model used in [BB04] to multiple users. Starting out from the single link model, the received signal after the transmission of a space-frequency signal matrix $\mathbf{C}$ can be written in equivalent form as

$$\mathbf{Y} = \mathbf{E}(\mathbf{C})\underline{\mathbf{H}} + \mathbf{W} \ , \tag{3.76}$$

where we have arranged the channel taps into

$$\underline{\mathbf{H}} \coloneqq \left[\mathbf{H}_0^T \quad \ldots \quad \mathbf{H}_{L-1}^T\right]^T , \tag{3.77}$$

and absorbed the power constant into the $N \times LM_T$ pseudocodeword matrix

$$\mathbf{E}(\mathbf{C}) \coloneqq [\mathbf{C} \quad \mathbf{DC} \quad \ldots \quad \mathbf{D}^{L-1}\mathbf{C}] . \tag{3.78}$$

The noise $\mathbf{W}$ is modelled as a $N \times M_R$ matrix whose elements are independent zero-mean complex Gaussian random variables of variance σ^2.

When K synchronised users, each equipped with M_T antennas, transmit simultaneously to a receiving base station, the received signal becomes

$$\mathbf{Y} = \sum_{k=1}^{K} \mathbf{E}(\mathbf{C}_k)\underline{\mathbf{H}}_k + \mathbf{W} = \overline{\mathbf{E}}(\mathbf{C}_k,...,\mathbf{C}_k)\,\overline{\mathbf{H}} + \mathbf{W} \ , \tag{3.79}$$

where we use the definitions

$$\overline{\mathbf{E}}(\mathbf{C}_1,\ldots,\mathbf{C}_K) \coloneqq \left[\mathbf{E}(\mathbf{C}_1) \quad \ldots \quad \mathbf{E}(\mathbf{C}_K)\right] \text{ and } \overline{\mathbf{H}} \coloneqq \left[\underline{\mathbf{H}}_1^T \quad \ldots \quad \underline{\mathbf{H}}_K^T\right]^T . \tag{3.80}$$

We emphasise that the value of $\overline{\mathbf{H}}$ is known to neither the transmitter nor the receiver. In contrast, the length of the cyclic prefix and therefore the maximum length of the impulse response are known. Moreover, we note here that we assume perfect synchronisation among the signals arriving at the base stations. This is not easy to achieve in practice, but we keep this standard assumption as a first approximation to obtain an insight into the benefit of noncoherent multi-user processing.

The task of the receiver is to determine the transmitted signals $\mathbf{C}_1,\ldots,\mathbf{C}_K$ based on the observation of the received signal $\mathbf{Y}$. Among the many receiver structures that have been proposed in literature, we will focus on the so-called generalised likelihood ratio test (GLRT) receiver. This choice is mainly motivated by the fact that no statistical knowledge about the channel is required. The receiver is therefore robust with respect to changes or uncertainties in statistical properties of the channel or the received powers from the different users. Moreover, the GLRT receiver leads to performance criteria that are relatively simple to evaluate. The GLRT receiver performs a joint channel estimation and data detection in

$$(\hat{\mathbf{C}}_1,\ldots,\hat{\mathbf{C}}_K,\overline{\mathbf{H}}) = \arg \min_{\mathbf{C}_1,\ldots,\mathbf{C}_K,\mathsf{H}} \left\| \mathbf{Y} - \overline{\mathbf{E}}(\mathbf{C}_1,\ldots\mathbf{C}_K)\overline{\mathbf{H}} \right\| , \tag{3.81}$$

which can be concentrated to a minimisation over $\mathsf{E} = \mathsf{E}(\mathbf{C}_1,\ldots,\mathbf{C}_K)$ by replacing the channel with its ML estimate $\hat{\mathsf{H}} = (\mathsf{E}^{\mathsf{H}}\mathsf{E})^{-1}\mathsf{E}^{\mathsf{H}}\mathbf{Y}$ given E. The resulting receiver is

$$(\hat{\mathbf{C}}_1,\ldots,\hat{\mathbf{C}}_K) = \arg \min_{\overline{\mathbf{E}}} \left\| \mathbf{Y} - \overline{\mathbf{E}}(\overline{\mathbf{E}}^H\overline{\mathbf{E}})^{-1}\overline{\mathbf{E}}^H\mathbf{Y} \right\| = \arg \min_{\overline{\mathbf{E}}} \left\| \mathbf{P}_{\overline{\mathbf{E}}}^{\perp}\mathbf{Y} \right\| , \tag{3.82}$$

where we denote with $\mathbf{P}_{\overline{\mathbf{E}}}^{\perp}$ the projection matrix on the complement of the column space of $\overline{\mathbf{E}}$ and where the minimisation is carried out over any $\overline{\mathbf{E}}$ that could possibly have been transmitted. In order to transmit data at a rate R_k in bits per carrier, user k chooses the matrix $\mathbf{C}_k$ from a codebook with $2^{R_k N}$ codewords. Each codeword in that particular codebook can thus be identified using an index i_k with $0 \le i_k < 2^{R_k N}$. In the following, we will denote the codeword with index i_k of user k with $\mathbf{C}_{k,i_k}$ and the corresponding pseudocodewords $\mathbf{E}_{k,i_k} \coloneqq \mathbf{E}(\mathbf{C}_{k,i_k})$. The main goal of the present contribution is to determine a criterion for the design of the users' codebooks. We will present an approach for this design in the next section.

3.3.3.1.3 Signal Design

The aim of the design of each user's codebook is the transmission of data with low probability of error, i.e., the data detected at the receiver is different from the transmitted signal for $e \ge 1$ users. Optimising the error probability directly appears to be an intractable problem. Instead, we consider the minimisation of an approximate bound on the error probability that is valid for high SNR. Apart from an average power

constraint, we do not impose any structure such as a linear dependence of the transmitted codewords on data symbols chosen from a preselected alphabet (as with LDC codes in Figure 2.11). In particular, a linear dependence on the data symbols will lead to an ineffective codebook design because the signal spaces spanned by two codewords multiplied with the unknown channel may coincide. Thus, these codewords cannot be distinguished at the receiver.

We start the derivation of the design criterion by considering the event that e users are detected erroneously and $K-e$ are detected correctly. Define $p_1,\ldots,p_e$ as the indices of the erroneously detected users and $p_{e+1},\ldots,p_K$ as the indices of the correctly detected users. Then arrange the pseudocodewords of the correctly detected users into

$$\widetilde{\mathbf{E}}_q := [\mathbf{E}_{p_{e+1},t_{p_{e+1}}},\ldots,\mathbf{E}_{p_K,t_K}], \tag{3.83}$$

and the transmitted but erroneously received pseudocodewords into

$$\widetilde{\mathbf{E}}_t := [\mathbf{E}_{p_1,t_{p_1}},\ldots,\mathbf{E}_{p_e,t_{p_e}}] \tag{3.84}$$

$$\widetilde{\mathbf{E}}_r := [\mathbf{E}_{p_1,r_{p_1}},\ldots,\mathbf{E}_{p_e,r_{p_e}}], \tag{3.85}$$

where t_k and r_k are the indices of the transmitted and detected codewords of user k, respectively. We are now interested in the *pairwise error probability* (PEP), i.e., the probability that the receiver computes a value in favour of the codewords contained in $\overline{\mathbf{E}}_r = [\widetilde{\mathbf{E}}_q \quad \widetilde{\mathbf{E}}_r]$ when codewords in $\overline{\mathbf{E}}_t = [\widetilde{\mathbf{E}}_q \quad \widetilde{\mathbf{E}}_t]$ were transmitted. This pairwise error probability will obviously depend on the value of the channel coefficients. Since we are interested in a code design that is robust with respect to variations in the channel, we can average the PEP over an assumed distribution. Assuming i.i.d. zero-mean Gaussian elements of $\mathbf{H}$, we apply results by Brehler and Varanasi [BV01] developed for multi-user space-time communication in our context. For high SNR, i.e., small σ^2, the PEP can then be approximated by

$$\Pr(\widetilde{\mathbf{E}}_t \to \widetilde{\mathbf{E}}_r) \approx \sigma^{2eLM_TM_R} \frac{B}{\det(\widetilde{\mathbf{E}}_t^H \mathbf{P}_{\overline{\mathbf{E}}_r}^{\perp} \widetilde{\mathbf{E}}_t)}, \tag{3.86}$$

where B is a constant independent of the design of the signalling matrices. This expression holds under the condition that $\widetilde{\mathbf{E}}_t^H \mathbf{P}_{\overline{\mathbf{E}}_r}^{\perp} \widetilde{\mathbf{E}}_t$ is full rank. Considering all possible error events, the minimum of the slope eLM_TM_R of the pairwise error probability is an indication of the available diversity gain. It is therefore apparent that a diversity gain LM_TM_R can be achieved, which is the maximum diversity available for a each user. Moreover, it is clear that for high SNR, the pairwise error probability for events with $e>1$ will be negligible when compared to single-error events. We will therefore only consider single-error events in our optimisation of the codebooks. A cost function for this optimisation can now be defined as a sum over all possible single-error events over all possible transmitted codewords as

$$C = \sum_{i_1,\ldots,i_K} \sum_{u=1}^{K} \sum_{j_u \neq i_u} \frac{1}{\det(\mathbf{E}_{u,i_u}^H \mathbf{P}_{\overline{\mathbf{E}}_r}^{\perp} \mathbf{E}_{u,i_u})}, \tag{3.87}$$

where $\overline{\mathbf{E}}_r$ contains all pseudocodewords $\mathbf{E}_{k,i_k}$ for $k \neq u$ as well as $\mathbf{E}_{u,j_u}$. This cost function resembles a scaled union bound in the sense that approximations for the PEP of all error events are summed up neglecting all error events with more than one erroneously detected user.

The cost function can now be optimised using gradient search with a constraint on the transmitted power (average power transmitted per user is N). This optimisation is similar to the search in [GS03] for a single link. Note that the optimisation is in general a formidable task but can be managed for small problem sizes. Performance evaluations are presented in the following section.

3.3.3.1.4 Evaluation

The purpose of this section is the investigation of potential multi-user gains in a noncoherent MIMO OFDM system. Thus we are primarily interested in how far a noncoherent MIMO OFDM system with dedicated multi-user processing, i.e., where multiple users communicate on the same subcarriers (they "collide") performs better than a system avoiding collision of users by assigning orthogonal frequency

bands.

In order to assess the benefit of the multi-user design, we therefore compare our scheme with three benchmark schemes designed for single-link noncoherent MIMO OFDM, where all users employ the same codebook on orthogonal carriers:

1. We design a code using the methodology outlined in Section 3.3.3.1.3 but restrict ourselves to the single-user case. All users then use the same codebook, however on orthogonal carriers.
2. Noncoherent space-frequency codes as designed in [BB04] are applied to design unitary codebook matrices.
3. We design a code based on space-frequency differential unitary transmission as outlined in [GS04] but use the GLRT receiver detecting several symbols jointly. Note that the necessary initialisation for the differential encoding requires an overhead that has to be compensated by an appropriate rate of the chosen constellation.

The channel is assumed to match with the assumptions of the code design, i.e., the elements of H are i.i.d. zero-mean Gaussian random variables with variance $1/L$. We assume equal power levels from all users and define the $\mathrm{SNR} = 1/\sigma^2$. The performance of a multi-user code with $R = 0.5$ bits per subcarrier and user for $K = 2$ users each employing $M_T = 2$ transmit antennas over a channel with $L = 2$ taps with $N = 16$ is presented in Figure 3.31.

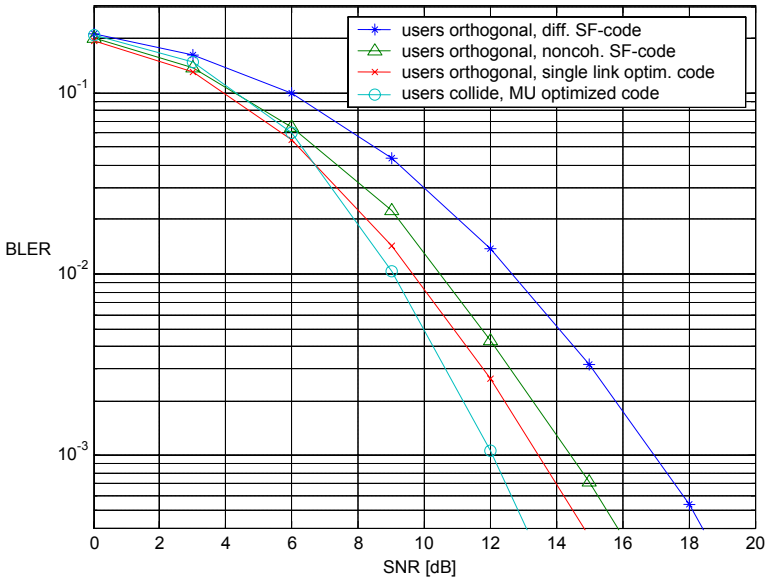


Figure 3.31: Example with two users equipped with 2 transmitter antennas using N = 16 subcarriers. The channel has L = 2 i.i.d. taps

The benchmark schemes have the same parameters and support the same bit-rate per user. The performances do not differ among the users within a given system. All proposed systems allow the exploitation of space and frequency diversity. The proposed design method leads to superior code performance in a multi-user setting. For example, the system using the proposed multi-user signal design method outperforms a system based on optimised single-link design by more than 1 dB for a BER of 10^{-3}. The gain becomes even more significant in comparison to the other benchmark systems using noncoherent space-frequency coding and space-frequency differential transmission.

3.3.3.2 Required Support Functions

The technique investigates the benefit of signalling design when CSI is available at neither transmitter nor receiver. To isolate this effect, we exclude other factors that might influence the performance. In particular, we assume synchronisation of users in the uplink and general complex-valued signalling alphabets without peak-to-average power constraints. Future research will be required in order to investigate how to

handle the much more complicated case when these effects are taken into account jointly.

3.3.3.3 Scalability

We expect to obtain significant gain compared to single-user design in scenarios according to WINNER assumptions. However, to our present knowledge, we do not know how to scale the proposed design technique to scenarios with a large number of users transmitting at high spectral efficiencies over large bandwidths without significant increase of design and receiver complexity. Finding low complexity approximations of our design technique is in the focus of our current research.

3.3.3.4 Summary

A signal design methodology for noncoherent multi-user MIMO-OFDM systems was presented. The designed signals allow the users to occupy the same bandwidth, provide space and frequency diversity gain and outperform several other schemes where users are assigned orthogonal frequencies. We are therefore confident that dedicated multi-user signal design can be beneficial for an improvement in error performance.

However, to our present knowledge, it is not clear how to exploit the available gain in typical WINNER scenarios if design and receiver complexity are constrained. It is therefore not possible to assess the presented method in full detail given the time plan of the WINNER project.

3.4 Relaying in Multi-antenna Networks

Future communications systems are expected to fulfil stringent requirements in terms of data rates, coverage and efficiency. Relaying techniques will contribute to the achievement of these goals. In general, relays are used to extend the network connectivity, to use more efficiently the resources and obtain cost-efficient network architectures. By reducing transmission distance, the aggregated path loss is smaller than for direct communication, and consequently, the transmission power can be lowered or higher data rates can be targeted. However, relaying also incurs a number of drawbacks including for instance higher delays, a radio transceiver complexity increase and more traffic on the network.

There are various types of relaying, each of which results in different degrees of complexity and performance. A low-complexity repeater relays the signal after amplification, whereas more complexity is required for a relay that decodes and re-encodes the packet before forwarding it. Additionally, a more sophisticated processing can be applied (e.g., distributed selection diversity, distributed space-time coding…) achieving contrasted performance.

The assessment of relaying techniques should not be dissociated from the multi-antenna techniques that will be used in the WINNER system concept. Studying and understanding the impact of relaying in a MIMO network is required if these techniques are to be successfully integrated in the concept.

3.4.1 Amplify-and-Forward and Decode-and-Forward for Multi-Antenna Stations

3.4.1.1 Overall Description of the Technique

We investigate by means of simulations the relaying techniques known as amplify-and-forward (AF) and decode-and-forward (DF) in a network where the stations have multiple antennas. Figure 3.32 illustrates the 2-slot scheduling structure assumed for a relaying operation. The underlying multiple-access scheme is TDMA, with the time slots being assigned by a central controller or access point (AP). Thus the AP always assigns 2 consecutive time slots when a packet scheduled for transmission has to be relayed. The AP transmits a packet on the first time slot to the destination or mobile station (MS). In addition, a relay station (RS) also receives the packet transmitted by the AP. During the second time slot, the RS retransmits the packet to the destination (MS) using one of the aforementioned methods (i.e., either by amplifying the received packet or by completely decoding and re-encoding it).

In the considered network, all stations can be equipped with multiple antennas, enabling therefore different spatial signalling mode on different links. Different strategies achieving different diversity-multiplexing tradeoffs may be used. In what follows, the processing performed by the relay station is described in terms of the effective channel experienced by the transmitted signal. For example, in case Alamouti space-time coding is utilised, the effective channel is the well-known scalar channel with gain equal to the

Frobenius norm of the channel matrix. For spatial multiplexing, the effective channel is the channel matrix itself.

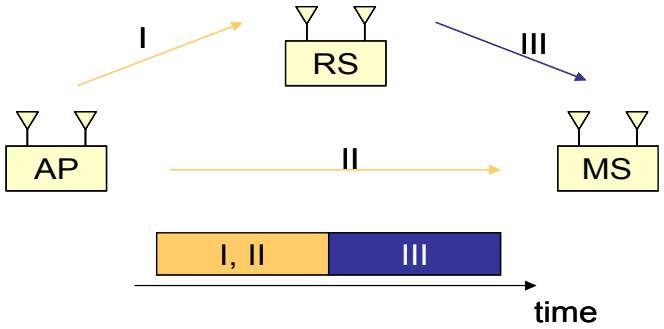


Figure 3.32: Relaying Scheduling Structure

Due to the fact that the same information is conveyed over two different time slots, the relaying protocol takes a hit of ½ on its achievable information rate. However, it has been shown that under some conditions, relaying is in fact able to recover from this loss (see for example [NBK04]). This theoretical result will be investigated for typical propagation conditions in a short-range scenario later in this document.

Let us denote by the subscripts I, II and III the link on which a signal is received. Therefore, the relayed communication illustrated in Figure 3.32 is completely described by the following set of equations:

$$\mathbf{y}_k = \sqrt{\frac{E_k}{M_T}}\mathbf{H}_k\mathbf{s}_k + \mathbf{n}_k, \quad k \in \{\mathrm{I}, \mathrm{II}, \mathrm{III}\}, \tag{3.88}$$

where k denotes the link, E_k the average signal energy received on link k and the matrix $\mathbf{H}_k$ is of dimension $M_{R,k}$ x $M_{T,k}$. The transmitted vector $\mathbf{s}_k$ has components drawn from a finite constellation. Channel knowledge is assumed to be perfect at the receiver, but the transmitter has none. Therefore, the transmitted power is evenly split among antennas. The differences between the two relaying methods under consideration are described below:

A) Amplify-and-Forward

The signal received at the RS is amplified by a factor β that ensures a normalised average energy of the signal retransmitted by the relay. Therefore, we have:

$\mathbf{s}_{\mathrm{III}} = \beta\ \mathbf{y}_{\mathrm{I}}$, where

$$\beta = \sqrt{\frac{M_{T,\mathrm{III}}}{\mathrm{E}_{\mathrm{I}}\left\|\mathbf{H}_{\mathrm{I}}\right\|_{\mathrm{F}}^2 + N_0 M_{R,\mathrm{I}}}}. \tag{3.89}$$

The signal received on the concatenation of links I and III is thus given by:

$$\mathbf{y}_{\mathrm{III}} = \sqrt{\frac{E_{\mathrm{I}}E_{\mathrm{III}}}{M_{T,\mathrm{I}}M_{T,\mathrm{III}}}}\mathbf{H}_{\mathrm{I}}\mathbf{H}_{\mathrm{III}}\beta\ \mathbf{s}_{\mathrm{I}} + \sqrt{\frac{E_{\mathrm{III}}}{M_{T,\mathrm{III}}}}\mathbf{H}_{\mathrm{III}}\beta\ \mathbf{n}_{\mathrm{I}} + \mathbf{n}_{\mathrm{III}}, \tag{3.90}$$

where, conditionally on $\mathbf{H}_k$, the effective noise is normal with covariance matrix

$$\mathbf{C}_{NN} = N_0\left(\mathbf{I} + \frac{\mathrm{E}_{\mathrm{III}}}{\mathrm{E}_{\mathrm{I}}\left\|\mathbf{H}_{\mathrm{I}}\right\|_{\mathrm{F}}^2 + N_0 M_{R,\mathrm{I}}}\mathbf{H}_{\mathrm{III}}\mathbf{H}_{\mathrm{III}}^{\mathrm{H}}\right). \tag{3.91}$$

B) Decode-and-Forward

The signal received at the RS and at the MS during the first time slot using DF is identical to that received

using AF. In contrast to the AF case, the relay does not forward any packet when it is not able to decode it correctly. This observation can be stated as follows:

$$\mathbf{y}_{\mathrm{III}} = \sqrt{\frac{E_{\mathrm{III}}}{M_{T,\mathrm{III}}}}\mathbf{H}_{\mathrm{III}}\mathbf{s}_{\mathrm{III}} + \mathbf{n}_{\mathrm{III}}, \text{where } \mathbf{s}_{\mathrm{III}} = \begin{cases} \mathbf{s}_{\mathrm{I}}, & \text{if decoded successfully at RS} \\ \mathbf{0}, & \text{if } \mathbf{s}_{\mathrm{I}} \text{ not decoded success. at RS} \end{cases}. \tag{3.92}$$

The MS does not require any information on link I, apart from knowing whether the signal has been retransmitted. It already appears that with DF the performance on the concatenated link will be directly impacted by the quality of link I, i.e., the relay will only be useful if link I is good enough so as to allow a successful decoding at the relay.

3.4.1.2 Required Support Functions

Although channel state information is not required at the transmitter side, all the receivers must have information on the channel in order to decode the received signal correctly. For the concatenation of links I and III and in case of performing AF relaying, we note that the MS requires channel state information of link I in addition to that related to link III. DF does not suffer from this additional requirement and can correctly decode the signal received on link III independently of the channel I. However, we still note that the relay should be able to notify any absence of signal on III due to a decoding error.

3.4.1.3 Additional Assumptions and Simplifications for Simulations

The required channel state information described in the previous paragraph is assumed to be perfectly known. Moreover, since we are considering an underlying OFDM system, we assume that the cyclic prefix perfectly copes with the inter-symbol interference. Finally, all the stations in the network are perfectly synchronised with a common clock reference given by the AP.

3.5 Review of Multi-Antenna Techniques

To round off the chapter on multi-antenna techniques, Table 3.6 provides an overview on existing schemes and their characteristics. Note that well known techniques, such as, STBC, STTC, BLAST or beamforming approaches for single links are not described in this deliverable as they have been already studied in literature extensively. For completeness also this well known schemes have been included in Table 3.6.

Table 3.6 Different multi-antenna techniques and their implementation in the generic block diagram that is depicted in Figure 3.10.

		FEC	SF mapping / multiplexing	modulation / bit loading	power loading	LDC	VASS	NLFB	BF	CSI at Tx
focus on link-level processing	**space-time block codes (STBC)**	✓[3]	×	single stream, adaptive	single stream, adaptive	✓ $R \leq 1$	×	×	×	×
	space-time trellis codes (STTC or STTrC)	✓	×	trellis coded modulation			×	×	×	×
	space-time Turbo codes (STTuC)	✓	×	space-time Turbo coding			×	×	×	×
	open-loop vector modulation (BLAST)	✓	×	single stream, adaptive	single stream, adaptive	✓ $R > 1$	×	×	×	×
	open-loop linear matrix modulation (LDC)	✓	×	single stream, adaptive	single stream, adaptive	✓	×	×	×	×

[3] Applicable "✓", not applicable "×"

		FEC	SF mapping / multiplexing	modulation / bit loading	power loading	LDC	VASS	NLFB	BF	CSI at Tx
	closed-loop transmit diversity (FDD)	✓	×	single stream, adaptive	single stream, adaptive	×	×	×	✓ single stream	short-term CSI
	closed-loop spatial multiplexing (FDD) with outer coding	✓	✓	stream-specific, adaptive		×	×	×	✓	short-term CSI
	closed-loop spatial multiplexing (FDD) with stream-specific coding	≥1	✓	stream-specific, adaptive		×	×	×	✓	short-term CSI
	per stream rate control (PSRC) (FDD)	≥1	✓	stream-specific, adaptive		(✓)	×	×	(✓)	short-term CQI (CSI)
	long-term MISO beamforming on link (FDD)	✓	×	single stream, adaptive		×	×	×	✓	Long-term CSI
	long-term MIMO eigenbeamforming (FDD) using LDC	✓	×	single stream, adaptive	single stream, adaptive	✓	×	×	✓	long-term CSI
	long-term MIMO eigenbeamforming (FDD)	≥1	✓	stream-specific, adaptive	stream-specific, adaptive	×	×	×	✓	long-term CSI
	MISO/SIMO/MIMO with optimised FEC	✓	×	optimised FEC			×	×	✓	short-term
	noncoherent MIMO transmission	✓	×	single stream, fixed		✓ nonco-herent	×	×	×	×
	mode switching based on channel rank (FDD)	✓	×	single stream, adaptive	single stream, adaptive	✓ different codes used	×	×	×	short-term CQI
multi-user concepts	**grid-of-beams beamforming + scheduling (FDD)**	✓	✓	single stream, adaptive	single stream, adaptive	×	✓	×	✓ fixed	short-term CQI
	opportunistic beamforming + scheduling (FDD), single-stream	✓	×	single stream, adaptive	single stream, adaptive	×	✓	×	✓ random	short-term CQI
	opportunistic beamforming + scheduling (FDD)	✓	✓	stream-specific, adaptive	stream-specific, adaptive	×	✓	×	✓	long-term CSI
	antenna selection + scheduling (FDD)	✓	✓	stream-specific, adaptive		×	✓	×	✓	short-term CQI
	antenna hopping + scheduling	✓	✓	stream-specific, adaptive		×	✓	×	×	×
	multi-user linear precoding (TDD)	✓	✓	stream-specific, adaptive		(✓)	✓	×	✓ jointly opt.	short-term CSI
	multi-user non-linear precoding (TDD)	✓	✓	stream-specific, adaptive		(✓)	✓	✓	✓ jointly opt.	short-term CSI

4. Simulation Methodology

Performance gains achieved on a single communication link (*link level*) do not necessarily translate into the same (relative) gains at *system level*, where multiple base stations communicate with multiple users. In order to perform a realistic evaluation of the enhancements achieved by advanced multi-antenna and coding schemes, system-level simulations are needed. In addition to modelling *self interference* (e.g., due to remaining ISI) and *intra-cell interference* (i.e., multiple access interference within the coverage region of a BS/AP), such system-level simulations would also model the interaction between multiple cells in the form of *inter-cell interference* (i.e., interference from other BS/APs and/or mobile terminals associated to other BS/APs). Interference may also arise from cross-channel interference due to unwanted transmissions from the same or other systems, or from imperfect receiver rejection of unwanted signals.

One of the major difficulties in system-level simulations is the complexity involved in characterising the performance of the radio links between all terminals and base stations. Link-level simulation of all such links is clearly prohibitive. As a result, the performance of the radio links has traditionally been evaluated in terms of the frame error rate (FER) for a defined frame size, as a function of signal-to-interference plus noise ratio (SINR), *averaged* over all channel realisations of one specific channel model. *FER versus average SINR performanc*e has therefore been used as the interface between the link- and system-level simulators. This may be adequate as long as every transmitted packet encounters similar channel statistics, which implies very large packet sizes/coding blocks with respect to the channel coherence time. But in case of data-centric radio networks, this condition is generally not fulfilled. The *specific* channel realisation encountered may result in performance that is significantly different from the one predicted from the *average* FER versus SINR performance. This will have a significant impact on crucial system-level mechanisms such as *packet scheduling*. Useful modelling of the performance of *fast scheduling* (in time, frequency, and/or space domain), *fast link adaptation* in form of fast power control and adaptive modulation and coding (AMC), or other advanced schemes such as HARQ is impossible with only *averaged/large-scale parameters* (e.g., path loss, shadowing) modelled on system level. Therefore, system-level simulations must include (maybe simplified) modelling of *small-scale/fast fading effects*.

Use of multiple antennas at the transmitter and/or at the receiver of a communication link will affect the *properties of the interference* experienced in the system as well as the ability to counteract the interference. This holds for schemes such as beamforming, closed-loop/open-loop transmit diversity, and closed-loop/open-loop spatial multiplexing. Therefore, to evaluate the system performance with schemes employing multiple Tx and/or Rx antennas, *interference modelling* is a topic that deserves attention. Interference modelling is of interest also in the case with conventional transmit and receive schemes. Just as the average FER versus average SINR performance may depend on the quality variations during a code block caused by channel properties such as delay spread and Doppler spread, we also expect that performance changes with interference variations caused by resource management.

To be able to generate comparable system-level simulation results within the WINNER project, a harmonised “standard” on how to model link level behaviour (i.e.,FER) in such simulations has been proposed.

In this chapter, first the scope of the link and system-level simulations is presented. Then, the problem of interfacing the link-to-system level results is addressed and different link level performance (quality) models are reviewed for this purpose. The selected quality model for the investigations performed in WINNER is described and some validation results are presented. Finally, complexity issues are discussed.

4.1 Link-to-System-Level Interface

4.1.1 Link Level Abstraction

An overview about the processing blocks involved in the system-level simulations and their dependencies is presented in Figure 4.1. It is assumed that the system to be simulated is a cellular system employing a *central controller/scheduler at each base station* (or access point, respectively). The green blocks represent *system specific behaviour*, which is generally a function of the yellow blocks, i.e., the *radio channels*, *environment characteristics*, as well as the *traffic characteristics*. (Note that although marked yellow, the interference block can also depend on the particular system concept.) From the BER/FER values to be determined different *figures of merit* can be derived such as, e.g., aggregate cell throughput

or delay distributions for different traffic queues.

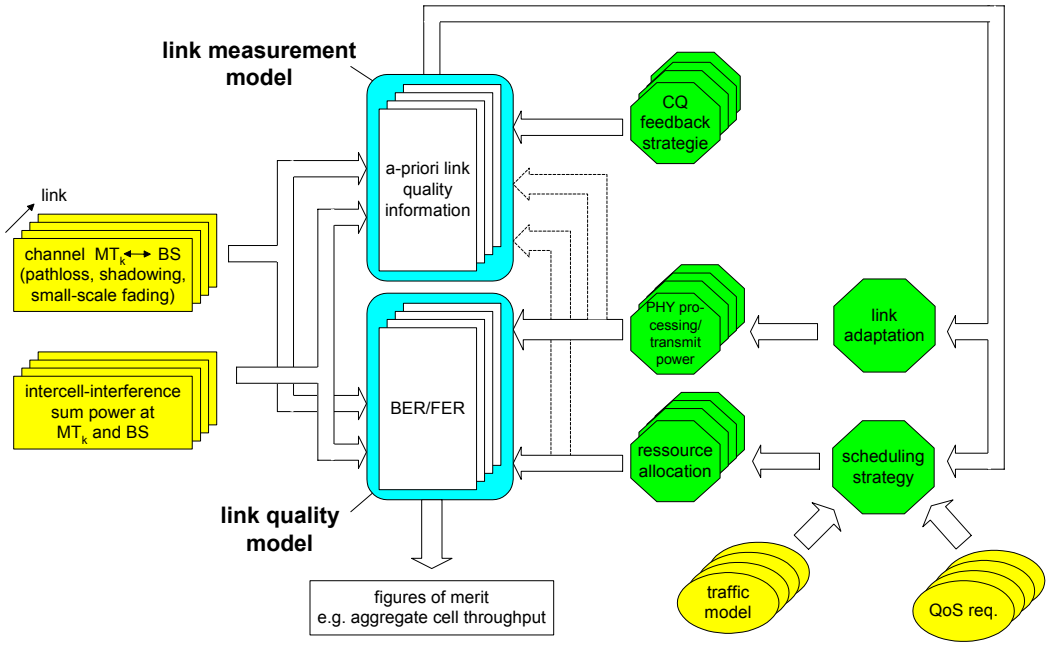


Figure 4.1: Example block diagram of system-level simulations

It can be observed in Figure 4.1 that two link models are required for system simulations. Besides estimating the link quality (e.g., BER, PER/FER/BLER) for each link after the resource allocation and link processing have been fixed, some sort of system dependent channel quality measures must be provided beforehand, based on which scheduling and link adaptation schemes can be performed. In the following, these models shall be referred to as:

- (Link) quality models
- (Link) measurement models

Measurement models reflect base station and terminal measurements, such as SINR estimates used for channel dependent scheduling and link adaptation. In fact, the measurement results not only depend on the channel and intercell interference, but also on the system settings during the measurement phase as for example Tx power, beamforming weights (indicated by the dashed arrows). Note that the measurement models are needed in the system-level simulator to provide appropriate estimates of the channel quality (depending on a certain channel quality feedback strategy) that will be exploited by RRM, scheduling, and/or link adaptation. Hence, measurement models are strongly related to the system concept.

Quality models, on the other hand, provide an estimate of the link performance (e.g., PER) when *decisions* on the RRM, scheduling, and link adaptation are *already known*. Thus, quality models are less complicated, since, for instance, channel quality mismatch due to possibly unknown scheduling decisions is not an issue. Quality models generate bit or block error probabilities and are utilised for *quality estimation at run-time*, which in turn can generate, for example, re-transmissions and affect also slow link adaptation, such as outer loop power control. Quality models should be able to handle HARQ and could also be required for modelling of decoding of physical layer signalling, such as TPC commands and feedback of channel state information for multi-antenna transmit schemes. A quality model may be viewed as a conditional probability; the probability that the transmitted code word is decoded erroneously given the channel, Tx power, beamforming weights, and interference plus noise during the interleaving period. Note that all the above-mentioned quantities may vary during the interleaving period.

Measurement models and quality models are very closely related with respect to their objectives: both allow performance prediction, either in the system itself to enable adaptive transmission, or in the simulation to evaluate the chosen transmission characteristics. However, they may differ considerably with respect to their output. The measurement model must reflect the system specific link quality estimation procedure and consequently each system requires its own model. Unlike the measurement model, the

same quality model can be used in all system simulations. Since it is not possible to provide a uniform measurement model for all potentially upcoming system proposals, this WINNER approach for system-level modelling focuses on quality models only.

4.1.2 Quality models

The purpose of the (*link) quality model* is to describe a procedure how to *accurately estimate the link performance* (e.g., BER, PER, goodput) of each user for known block length, selected PHY transmission mode, scheduling decision, resource allocation, propagation conditions, and intercell interference, without really performing the complete link level signal processing. It is desirable that the modelling captures transceiver characteristics, including, e.g., radio-frequency front-end functions and receiver algorithms. Furthermore, in order to evaluate advanced techniques, such as, e.g., interference cancellation and multi-antenna techniques at the system level, the model should capture the essential features in a fair way, by preserving the relative merits. The gain of different techniques can then be compared both with a reference system and with each other.

Considering an *adaptive multi-user MIMO system that* exploits *short-term* (instantaneous) channel (and maybe even interference) knowledge of all active users as a worst case scenario (i.e., scenario with the highest degree of adaptivity), it is required to *generate fast (=small-scale) fading channel realisations on system level* as indicated in Figure 4.1. It seems that without this detailed channel knowledge, fast link adaptation and scheduling and thus the system behaviour cannot be modelled accurately.

Figure 4.2 illustrates the task to be performed by the link quality model as well as a potential realisation covering most of the known models.

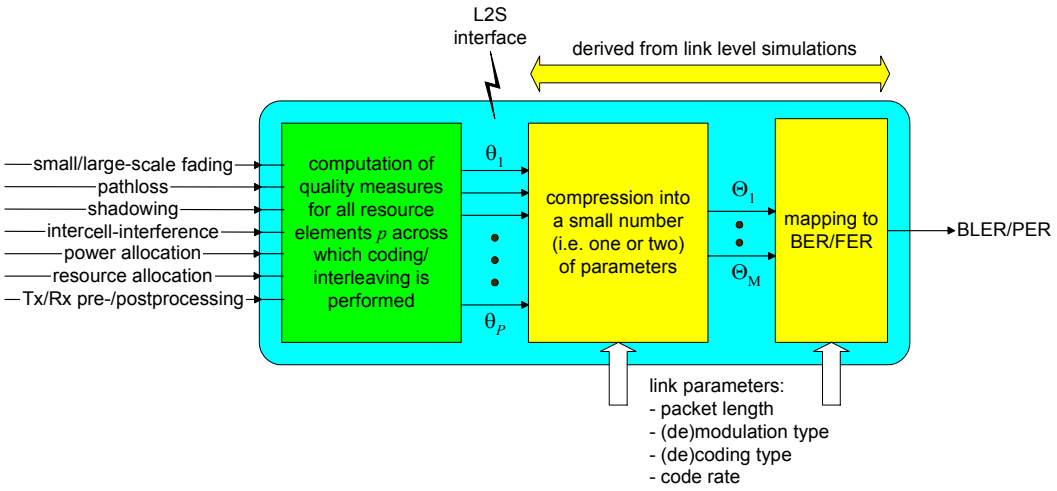


Figure 4.2: Potential realisation of a link quality model

Once the scheduler has determined what resources are allocated to which users and at what power level, the effective (MIMO) channel, including path loss, shadowing, small- (and large) scale fading as well as inter-cell interference, can be computed. From these values, a set of *quality measures* θ_p is derived *for all resource elements* p=1,..., P, covered by a single codeword taking the actual spatial, temporal and spectral Tx/Rx pre- and postprocessing into account. A possible approach is, for instance, to compute the SINR after equalisation at the receiver.

Since the number of resource elements can be very large (e.g., MIMO-OFDMA) and subsequent multi-dimensional mapping to PER is not feasible, some compression must be applied before mapping the compressed values Θ_1, Θ_2... (typically one or two scalars) to a bit or packet error rate. For example, the computation of the average SINR or average SINR plus variance reflects such compression. Separate link-level simulations are *only* used to create appropriate look-up tables or analytical equations in advance, in order to map the set of quality indicators $\{\Theta_1,\Theta_2...\}$ to the required BER or FER. The mapping function depends on the applied coding, interleaving and modulation scheme as well as the packet length. Therefore, if no parameterised analytical expression can be provided, a set of mapping

functions must be determined covering all possible test cases.

Modelling of HARQ is not considered in this document, but may be studied later.

In the following sections, *examples* of different quality models and corresponding compression functions, proposed in the literature, are presented.

4.1.2.1 Channel Capacity

In [3GPP2] and [FITD331]it was proposed to use the *channel capacity* as a quality measure. In general, *different capacity equations* apply for different MIMO transmission schemes (e.g., open-loop vs. closed-loop, ergodic capacity for coding/interleaving across fading realisations or outage capacity for quasi static channel within code word). Note also that the capacity approach does *not* necessarily directly account for any specific *modulation and coding format* and its efficiency. However, any transceiver imperfections may be incorporated into the subsequent PER mapping.

Let us assume a multi-carrier transmission system with K subcarriers that converts the frequency selective fading channel into a set of K frequency flat fading channels. Let $\mathbf{F}_k$ stand for the linear transformation applied at the transmitter to allocate power over the different channel modes (e.g. spatial modes) of the k-th OFDM subcarrier. Let also $\mathbf{H}_k$ stand for the MIMO channel response and $\mathbf{R}_k$ for the correlation matrix of the Gaussian coloured noise $\boldsymbol{\nu}_k$, respectively, so the received signal at the k-th subcarrier is written as:

$$\mathbf{y}_k = \mathbf{H}_k \cdot \mathbf{F}_k \mathbf{c}_k + \boldsymbol{\nu}_k \,. \tag{4.1}$$

Assuming a data vector $\mathbf{c}_k$ with complex Gaussian distributed elements and

$$E\{\boldsymbol{\nu}_k \boldsymbol{\nu}_k^{\mathrm{H}}\} = \mathbf{R}_k \,, \tag{4.2}$$

then the channel capacity in this scenario is obtained as an extension of the MIMO channel capacity given in [FITD331] to multi-carrier transmission:

$$C = \frac{1}{K} \sum_{k=1}^{K} \log\left[\det(\mathbf{I} + \mathbf{F}_k^{\mathrm{H}} \mathbf{H}_k^{\mathrm{H}} \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{F}_k)\right]. \tag{4.3}$$

The suitability of the proposed channel capacity metric has been investigated in [FITD331] for a number of scenarios depending on the validity of the Gaussian assumption for the interference. It was shown that under the assumption of additive white Gaussian noise and unstructured interference the metric provides a good interface. However, in the presence of structured interference and for small number of interferers the metric cannot provide a suitable interface due to considerable deviation from the Gaussian assumption that is made for noise plus interference. For large number (>2) of interferers the Gaussian assumption becomes more and more accurate and the channel capacity metric can provide a suitable interface.

Note that the channel capacity considered here inherently provides already a *sum across spatial modes* (e.g., Tx antennas for V-BLAST or spatial layers for general matrix modulation), rendering the assessment of different receiver options impossible. Furthermore, a *capacity metric* will most likely work well for systems that exploit a large amount of the capacity (e.g., for open-loop capacity and adaptive matrix modulation or closed-loop MIMO capacity and SVD-based MIMO), but might be less appropriate, if an adaptation of the transmission mode is not applied.

The mutual information measure given in the previous equation assumes infinite coding horizons and Gaussian waveforms. Typically, finite modulation alphabets such as QPSK and 16-QAM are used. To include the effects of the modulation alphabet, the constrained capacity can be calculated as for example in [CTB96].

4.1.2.2 Cut-off Rate

Regarding the cut-off rate derivation, a similar signal model as in equation 4.1 is considered. In contrast to the previous section, the transmitter produces a coded sequence $\boldsymbol{c}$ consisting of N symbols $\mathbf{c}=(\mathbf{c}_1,..,\mathbf{c}_N)$, where each symbol $\mathbf{c}_k$ is a vector of Q symbols in a constellation of size M (i.e., $\mathbf{c}_k \in M^Q$). In the general case, before transmitting the data through the M_T antennas, each symbol is pre-filtered at the transmitter, using matrix $\mathbf{F}_k$ of size ($M_T \times Q$) that allocates power over the different channel modes [FITD331]:

$$\boldsymbol{x}_k = \boldsymbol{F}_k \boldsymbol{c}_k \,. \tag{4.4}$$

When no pre-filtering is employed $\mathbf{F}_k$ is the identity matrix.

Let $\mathbf{H}_k$, $\mathbf{G}_k$, stand for the memoryless channel response of the desired user and the interference during the transmission of the k-th data symbol respectively. (In fact, as the channel is considered invariant during one codeword transmission and each symbol is transmitted through a different OFDM subcarrier, $\mathbf{H}_k$, $\mathbf{G}_k$ can only take K different values). Let also $\mathbf{i}_k$, $\mathbf{w}_k$ stand for the interfering signal and receiver noise degrading the transmission of symbol $\mathbf{c}_k$. Then, the received signal can be expressed as:

$$\boldsymbol{y}_k=\boldsymbol{H}_k\boldsymbol{F}_k\boldsymbol{c}_k+\boldsymbol{G}_k\boldsymbol{i}_k+\boldsymbol{w}_k\,, \tag{4.5}$$

Consider now a receiver that does not take into account the structure of the interference but it rather ignores it or models it as a coloured noise whose covariance can be measured at the receiver input. The simplified signal model is in this case:

$$\boldsymbol{y}_k=\boldsymbol{H}_k\,\boldsymbol{F}_k\boldsymbol{c}_k+\boldsymbol{v}_k\,, \tag{4.6}$$

where $\mathbf{v}_k$ is a vector of zero-mean coloured noise that is assumed to be Gaussian (default assumption).

This signal model includes as particular cases the receivers that model the interference as a Gaussian noise with the appropriate spatial correlation:

$$\mathbf{C}_{\mathbf{v}_k}=E\left\{\mathbf{v}_k\mathbf{v}_k^H\right\}=\mathbf{C}_{\mathbf{w}_k}+\sigma_i^2\mathbf{G}_k\mathbf{G}_k^H\,. \tag{4.7}$$

The codeword error probability is minimised by selecting the most likely codeword, i.e., the candidate that maximises the probability:

$$\mathrm{p}(\boldsymbol{y}\mid\hat{\boldsymbol{c}})=\mathrm{p}(\boldsymbol{y}_1,\ldots,\boldsymbol{y}_N\mid\hat{\boldsymbol{c}}_1,\ldots,\hat{\boldsymbol{c}}_N)\,, \tag{4.8}$$

or equivalently the candidate codeword that minimises the decoding metric:

$$\Delta(\hat{\mathbf{c}})=\sum_k\left(\mathbf{y}_k-\mathbf{H}_k\mathbf{F}_k\hat{\mathbf{c}}_k\right)^H\mathbf{C}_{\mathbf{v}_k}^{-1}\left(\mathbf{y}_k-\mathbf{H}_k\mathbf{F}_k\hat{\mathbf{c}}_k\right). \tag{4.9}$$

The pairwise error probability (PEP) of deciding for codeword $\hat{\mathbf{c}}$ when $\mathbf{c}$ was transmitted can be written as [FITD331]:

$$\begin{aligned}P(\mathbf{c}\to\hat{\mathbf{c}})&=P(\Delta(\mathbf{c})>\Delta(\hat{\mathbf{c}}))=\\&=P\Bigg(\sum_k\left(\mathbf{G}_k\mathbf{i}_k+\mathbf{w}_k\right)^H\mathbf{C}_{\mathbf{v}_k}^{-1}\mathbf{H}_k\mathbf{F}_k\left(\mathbf{c}_k-\hat{\mathbf{c}}_k\right)+\\&\quad\sum_k\left(\mathbf{H}_k\mathbf{F}_k\left(\mathbf{c}_k-\hat{\mathbf{c}}_k\right)\right)^H\mathbf{C}_{\mathbf{v}_k}^{-1}\left(\mathbf{G}_k\mathbf{i}_k+\mathbf{w}_k\right)+\\&\quad\sum_k\left(\mathbf{H}_k\mathbf{F}_k\left(\mathbf{c}_k-\hat{\mathbf{c}}_k\right)\right)^H\mathbf{C}_{\mathbf{v}_k}^{-1}\mathbf{H}_k\mathbf{F}_k\left(\mathbf{c}_k-\hat{\mathbf{c}}_k\right)\ <0\Bigg)\end{aligned} \tag{4.10}$$

A bound on the mean value of the PEP can be expressed in terms of the *channel cut-off rate*:

$$E_{\mathbf{c},\hat{\mathbf{c}}}\left\{P(\mathbf{c}\to\hat{\mathbf{c}})\right\}\le 2^{-N\cdot R_o}\,. \tag{4.11}$$

It can be shown [FITD331] that in the interference-free case, an exact expression for the channel cut-off rate can be computed:

$$R_{o\,\text{No interf}}=-\tfrac{1}{K}\sum_{k=1}^{K}\log_2\left(\tfrac{1}{Q^{2M}}\sum_{\mathbf{c}_k}\sum_{\hat{\mathbf{c}}_k}\exp\left(-\tfrac{1}{4}\left(\mathbf{H}_k\mathbf{F}_k\left(\mathbf{c}_k-\hat{\mathbf{c}}_k\right)\right)^H\mathbf{C}_w^{-1}\mathbf{H}_k\mathbf{F}_k\left(\mathbf{c}_k-\hat{\mathbf{c}}_k\right)\right)\right), \tag{4.12}$$

which in the case of white noise $\boldsymbol{C}_\mathrm{w}=\sigma_\mathrm{w}^2\mathbf{I}$ can be expressed as

$$R_{o\,\text{No interf white}}=-\tfrac{1}{K}\sum_{k=1}^{K}\log_2\left(\tfrac{1}{Q^{2M}}\sum_{\mathbf{c}_k}\sum_{\hat{\mathbf{c}}_k}\exp\left(-\tfrac{1}{4\sigma_w^2}\left\|\mathbf{H}_k\mathbf{F}_k\left(\mathbf{c}_k-\hat{\mathbf{c}}_k\right)\right\|^2\right)\right). \tag{4.13}$$

In the interference-limited scenario the above expression can only be used to provide a *lower bound on the channel cut-off rate.* We introduce a constant λ, with $\lambda>0$, which determines how tight the bound is. Due to the need of using the same value of λ for all codeword pairs in order to derive a simple expression

and assuming

$$m_{p_k} = (\boldsymbol{G_k i_k})^H \boldsymbol{C}_{v_k}^{-1} \boldsymbol{H_k F_k}(\boldsymbol{c_k} - \hat{\boldsymbol{c}}_{\boldsymbol{k}}) + (\boldsymbol{H_k F_k}(\boldsymbol{c_k} - \hat{\boldsymbol{c}}_{\boldsymbol{k}}))^H \boldsymbol{C}_{v_k}^{-1} \boldsymbol{G_k i_k} + (\boldsymbol{H_k F_k}(\boldsymbol{c_k} - \hat{\boldsymbol{c}}_{\boldsymbol{k}}))^H \boldsymbol{C}_{v_k}^{-1} \boldsymbol{H_k F_k}(\boldsymbol{c_k} - \hat{\boldsymbol{c}}_{\boldsymbol{k}})$$
$$\sigma_{p_k}^2 = 2(\boldsymbol{H_k F_k}(\boldsymbol{c_k} - \hat{\boldsymbol{c}}_{\boldsymbol{k}}))^H \boldsymbol{C}_{v_k}^{-1} \boldsymbol{C}_{w_k} \boldsymbol{C}_{v_k}^{-1} \boldsymbol{H_k F_k}(\boldsymbol{c_k} - \hat{\boldsymbol{c}}_{\boldsymbol{k}})$$

a lower bound for the channel cut-off rate can be written as follows:

$$R_o \geq \min_{\lambda} \; -\frac{1}{K}\sum_{k=1}^{K} \log_2\left(\frac{1}{Q^{2M}} \sum_{\mathbf{c}_k} \sum_{\hat{\mathbf{c}}_k} B_\lambda(\mathbf{c}_k, \hat{\mathbf{c}}_k)\right) = \min_{\lambda} \; -\frac{1}{K}\sum_{k=1}^{K} \log_2\left(\frac{1}{I}\frac{1}{Q^{2M}} \sum_{\mathbf{c}_k} \sum_{\hat{\mathbf{c}}_k} \sum_{\mathbf{i}_k} \exp\left(-\lambda m_{p_k} + \frac{\lambda^2}{2}\sigma_{p_k}^2\right)\right). \quad (4.14)$$

4.1.2.3 Signal-to-Interference-plus-Noise Ratio (SINR)

Let us consider a transceiver structure employing bit-interleaved coded modulation (BICM), as depicted in Figure 4.3.

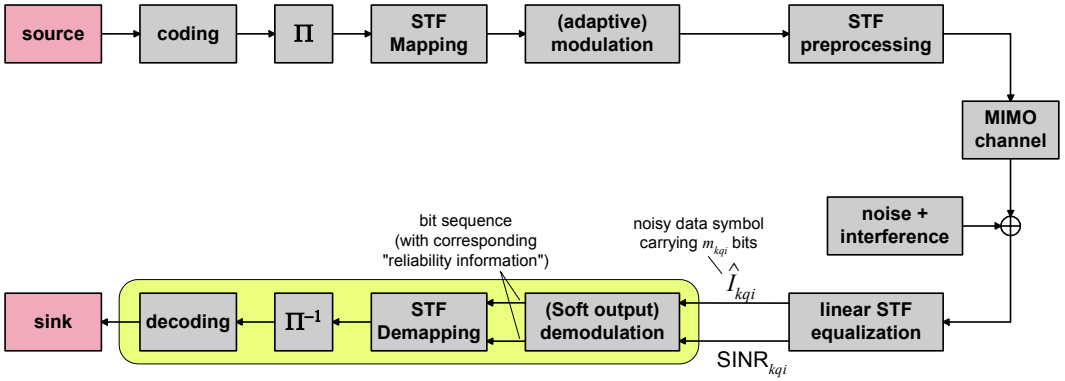


Figure 4.3: Transceiver block diagram employing BICM

The use of SINR after receiver space-time-frequency postprocessing as L2S interface metric is based on the assumption that these values provide sufficient information for predicting the demodulator/decoder behaviour in terms of PER (without really performing demodulation and decoding) and moreover, account accurately enough for the Tx/Rx specific pre- and postprocessing capabilities. However, analytical expressions are difficult (if not impossible) to obtain if space-time-frequency equalisation and decoding are performed *jointly* or *iteratively*. Hence, MIMO *Turbo equalisation* or other advanced schemes on system level may be difficult to assess. It should be stressed that the SINR values (similar to the capacity values in Section 4.1.2.1) only need to be calculated for the time/frequency elements *i*, *k* and the spatial layers *q* (in case of MIMO vector/matrix modulation), *across which the encoded bits are transmitted*.

Predicting the PER based on the computed SINR values requires an appropriate model for demodulator and decoder. Whenever the modulation is identical for all resource elements (e.g., all subcarriers, spatial layers, OFDM symbols, and/or HARQ transmissions) we can directly utilise the respective SINR values as input for the decoder model. However, in the case of adaptive modulation within a code word, the bit quality varies not only as a function of the SINR, but also as a function of the applied modulation scheme. Different solutions can be considered in this case:

- Weighting of SINR values, i.e., $SINR_{kqi}' = w_{\text{mod}} \cdot SINR_{kqi}$, depending on modulation format/level
- Use of mutual information

In any case, the SINR characterises the *soft bit quality* either directly (*non-adaptive modulation*) or indirectly via an *additional mapping function* (*adaptive modulation*).

4.1.3 Data Compression and Mapping to Packet Error Rate

Since different bits can be transmitted on different *spatial layers*, in different *OFDM symbols*, and on different *OFDM subcarriers* (i.e., coding/interleaving across space, time, and/or frequency) they may have different quality. Thus, we have a *multi-state channel*. The task of the decoder model is to map the

set of soft bit quality values (such as SINRs) to a block error probability. As emphasised in the previous section, the *total number of resource elements* in *frequency* (i.e., subcarriers or blocks of subcarriers), *time* (i.e., subsequent OFDM symbols or HARQ retransmissions), and *space* (i.e., layers for MIMO vector or general matrix modulation) is typically by far too large for direct use of the associated quality measures $\{\theta_1...\theta_P\}$ in the mapping to PER. To reduce look-up tables or analytical equations for the PER to only a few independent parameters, the set of quality measures needs to be compressed to a much smaller set $\{\Theta_1,\Theta_2...\}$ of typically only 1 or 2 scalar indicators. Since *compression* and *mapping to PER/BLER* (denoted by CPM in the following paragraphs) are inherently linked and cannot be designed independently, both are treated together.

The presentation of the compression and mapping methodology will be focused on the *SINR or respective bit reliabilities*, since this is the selected quality model in WINNER. However, similar ideas may be adapted for the other quality models presented in the previous sections. Note that the mapping does not only depend on the *average* SINR value, but on the entire *statistics* (or at least the variance) and the *correlation* (for non-ideal interleaving). There are *two variants of CPM*, which are either *independent of the packet length* or *packet length specific*.

Although rather simple from the conceptual point of view, the *complexity* of the CPM approach might be critical for an OFDM system with a large number of subcarriers. The computation of SINRs after demodulation for a large number of subcarriers, taking into account the current channel realisations and the linear receivers applied, appears challenging. *Decimation in time and frequency* corresponding to the coherence time and the coherence bandwidth is a natural way to reduce complexity. Once we have clearly defined *suitable channel indicators*, *system simulations for certain system concepts* may be considerably simplified by *modelling the statistics and correlations* instead of the complete set of channel matrices. However, initially some "full-blown" simulations are needed in order to verify further simplifications.

4.1.3.1 One-Dimensional Data Compression and Mapping

For the *multi-state channel*, the decoder model has to map the *set of soft bit quality values* $\{\theta_1,...,\theta_P\}$, such as SINRs, to a block error probability. The basic idea is to find a compression function that maps the sequence of varying SINRs to a single SINR value that is strongly correlated with the actual BLER.

The concept of an effective E_b/N_0 has been introduced in [NR98] for a CDMA system. Regarding *effective SINR*, a generalised approach,

$$\mathrm{SINR}_{\mathrm{eff}} = I^{-1}\left(\frac{1}{P}\sum_{p=1}^{P} I\left(\mathrm{SINR}_p\right)\right), \tag{4.15}$$

was derived in [3GPP2E], where the function $I(x)$ is referred to as an "information measure" and $I^{-1}(x)$ is its inverse. Several different "information measures" were considered like mutual information, AWGN channel capacity, cut-off rate, as well as linear, logarithmic and exponential mapping. Sigmoidal functions such as mutual information and exponential mapping revealed superior performance.

In [3GPPE], the exponential averaging method was extended by a scaling factor β yielding an information measure:

$$I(x) := \exp\left(-\frac{x}{\beta}\right). \tag{4.16}$$

This gives the following expression for effective SINR:

$$\mathrm{SINR}_{\mathrm{eff}} = -\beta \ln\left[\frac{1}{P}\sum_{p=1}^{P}\exp\left(-\frac{\mathrm{SINR}_p}{\beta}\right)\right]. \tag{4.17}$$

The scaling factor β allows to adjust the compression function in a way that the mismatch between the actual BLER and the predicted BLER (in conjunction with the mapping) is minimised in certain sense. As BLER mapping function, it is proposed to use the AWGN curves. Generally, for any set of physical layer parameters different β values and mapping functions apply (Figure 4.4).

Validation results for the so-called ***exponential effective SINR*** **mapping (EESM)** may be found in

[3GPPE2][3GPPN][3GPPN2], both for a single link case, but also for the case with coloured intercell interference. The accuracy of the method was shown to be sufficient for system-level evaluations.

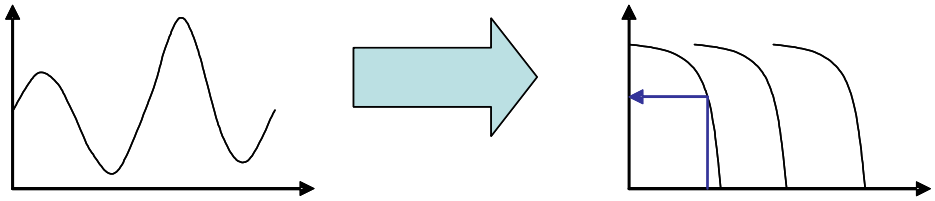


Figure 4.4: Exponential effective SINR mapping illustration.

A channel capacity based approach is obtained using a logarithmic information measure:

$$I(x) := \log_2\left(1+\frac{x}{\beta}\right). \tag{4.18}$$

This gives the following expression for effective SINR:

$$\mathrm{SINR}_{\mathrm{eff}} = \beta \cdot \left(2^{\frac{1}{P}\sum_{p=1}^{P}\log_2\left(1+\frac{\mathrm{SINR}_p}{\beta}\right)} - 1 \right). \tag{4.19}$$

Similarly it was proposed in [IEEE80211] to use an effective SINR according to

$$\mathrm{SINR}_{\mathrm{eff}} = 10^{-\beta\,\mathrm{var}(\log_{10}\mathrm{SINR}_p)} \cdot 10^{\frac{1}{P}\sum_{p=1}^{P}\log_{10}(\mathrm{SINR}_p)}. \tag{4.20}$$

Again in both cases, β is a parameter to minimise the gap between predicted and measured BLER. Note that the last compression function was presented as part of the IEEE 802.11n standardisation assuming a MIMO-OFDM transmission with individually encoded and modulated data streams at each spatial layer q (i.e., no coding across spatial substreams). For any stream q the effective SINR is determined and mapped to a BER, again via the AWGN simulation results. The BLER is then derived from BER values using the block length as parameter. In the following, equation (4.18) shall be referred to as ***capacity effective SINR mapping*** **(CESM)** and equation (4.20) as ***logarithmic effective SINR*** **mapping (LESM)**, respectively.

The use of mutual information as "information measure" has been proposed and investigated in [3GPP2E]. This approach appears attractive in the sense that the *modulation alphabet* is at least conceptually accounted for and that such a measure may therefore be *applicable also when the bits of a code word are mapped onto symbols from different modulation alphabets*. Candidate information measures to be considered include the coded modulation capacity as well as the bit-interleaved coded modulation capacity. The latter can be written as [CTB96]:

$$I_{m_p}(x) = m_p - E_Y\left\{\frac{1}{2^{m_p}}\sum_{i=1}^{m_p}\sum_{b=0}^{1}\sum_{z\in X_b^i}\log\frac{\sum_{\hat{x}\in X}\exp(-|Y-\sqrt{x}(\hat{x}-z)|^2)}{\sum_{\tilde{x}\in X_b^i}\exp(-|Y-\sqrt{x}(\tilde{x}-z)|^2)}\right\}, \tag{4.21}$$

where m_p is the bits per symbol, X is the set of 2^{m_p} symbols, X_b^i is the set of symbols for which bit i equals b. Further, Y is a zero mean unit variance complex Gaussian variable. Note that the use of the BICM capacity expression depends on the demodulator used. If an approximate demodulator is used to calculate log likelihood ratios, another capacity expression would apply.

Based on the generalised approach in equation (4.15), the **mutual information effective SINR mapping** (termed **MIESM** in the following) in conjunction with an additional scaling parameter β can be described by

$$\mathrm{SINR}_{\mathrm{eff}} = \beta \cdot I_{m_{\mathrm{ref}}}^{-1}\left(\frac{1}{P_u}\sum_{p=1}^{P_u} I_{m_p}\left(\frac{\mathrm{SINR}_p}{\beta}\right)\right), \tag{4.22}$$

where I_{m_p} is the mutual information function of the applied modulation alphabet of size 2^{m_p} at the pth data symbol. Note that averaging should only be performed across those resource elements that are actually occupied. If in case of OFDM, for example, some subcarriers are not used due to bad fading conditions, then the corresponding resource elements must not be counted, yielding $P_u < P$. The value m_{ref} can be set to the average number of transmitted bits per resource element. Obviously, the corresponding function I_{mref} only exists for integer values of m_{ref}. However, in case of high correlation between the computed average mutual information and the PER, the choice of I^{-1} mainly influences the shape of the subsequent mapping function and not (or only to a small extent) the PER prediction quality.

In summary, a possible decoder model could first determine an *effective SINR* and then apply this effective SINR to determine the BLER/PER from a *one-dimensional look-up table* determined either from a *single AWGN simulation* or derived directly from the recorded PER data (as function of the effective SINR). It should be kept in mind that such a model in contrast to the approach in [NR98] neglects the structure of the interleaver.

Rough models to capture the impact of inter-symbol interference, channel estimation errors, frequency offsets and other imperfections, as well as particular receiver structures may be included in the demodulator model that generates the SINR. It should be emphasised that for a mismatched demodulator, SINR does not necessarily alone dictate the channel decoder performance. In this case, both the signal level and the interference level may affect the performance.

4.1.3.2 Two-Dimensional Data Compression and Mapping

In [LRZ03][SBO04] it is proposed, to calculate the *SINR mean and (normalised) SINR variance indicator*

$$\mathrm{SINR}_{\mathrm{avg}} = \frac{1}{P}\sum_{p=1}^{P}\mathrm{SINR}_p \qquad I_{\mathrm{var},\beta} = \frac{1}{P}\sum_{p=1}^{P}\left(\sqrt{\frac{\mathrm{SINR}_p}{\mathrm{SINR}_{\mathrm{avg}}}} - E\left(\sqrt{\frac{\mathrm{SINR}_p}{\mathrm{SINR}_{\mathrm{avg}}}}\right)\right)^{\beta}, \tag{4.23}$$

and use these values as input parameters of a two-dimensional mapping function to determine the BLER, i.e.,

$$\mathrm{BLER} = f_M(\mathrm{SINR}_{\mathrm{avg}}, I_{\mathrm{var},\beta}). \tag{4.24}$$

The parameter β is again included for optimisation. In case β=2 the variance indicator equals the common variance definition.

4.1.4 Validation Results

It is believed that the same quality model should be applicable both for simulations with SISO systems as well as with various MIMO configurations. This since SISO is a special case of MIMO and since it allows a fair comparison of the benefits of advanced antenna systems as compared to SISO systems.

Therefore, the SINR based approach in Section 4.1.2.3 was taken, and we note that the channel capacity approach in Section 4.1.2.1 is similar to CESM in the SISO case. Further, from the results of [3GPP2E], we expect that the approach of cut-off rate approach of Section 4.1.2.2 to perform similar to EESM and MIESM in the SISO case.

A simulation campaign was performed, primarily with two different link level simulators, in an effort to identify which SINR based quality model to use. In order to limit the simulation time, the majority of the evaluations were done for the SISO case with the same modulation format for all transmitted symbols, but some results with adaptive modulation and MIMO spatial multiplexing with a linear receiver were also done verifying that the quality model may be used also for such cases. In the case with MIMO, the SINR calculation has to capture the characteristics of the liner space-time-frequency equalisers so that a post

receiver processing SINR is calculated. The conclusion of the simulation campaign was that MIESM is preferred over the other SINR based approaches.

In here, only a fraction of the results are presented, and only for MIESM, EESM and CESM. In all three cases, the models were trained, that is the value of the scaling parameter β was tuned, as described later in Section 4.1.6.2. Further, the validation is such that for each simulated BLER point, the difference in average SINR between the link level simulation and the simulation of the trained quality model is taken as an error, see further Section 4.1.6.3. Only points for which the BLER is within an assumed operating region between 0.3% and 30% are used. All available channel realisations are used in the validation, despite the fact that they have also been used to train the model. In addition to some scatter plots, the average of the magnitude of the error as well as the 95 percentile of the magnitude of the error are considered.

4.1.4.1 Simulation Assumptions

OFDM modulation with a carrier separation of 50 kHz and 1664 sub-carriers were considered. The simulations were done in the frequency domain, and the cyclic prefix was assumed to cover the time dispersion of the channel. 3GPP channel codes were used together with random bit interleaving. Note that a new random interleaver was generated for each coding block. BPSK, QPSK and 16QAM modulation with Gray mapping were assumed and the modulated symbols were mapped to all or a subset of the carriers. The modulation and coding schemes (MCS) considered are given in Table 4.1.

Table 4.1: Modulation and coding schemes

MCS	Channel coding	#uncoded bits	#coded bits	Modulation	#carriers
1	Conv. ½	824	1664	BPSK	1664
2	Conv. ½	824	1664	16QAM	416
3	Turbo 1/3	1104	3324	QPSK	1662
4	Turbo 1/3	1104	3324	16QAM	831
5	Punc. Conv ¾	2488	3328	QPSK	1664
6	Punc. Conv. ¾	2488	3328	16QAM	832

The demodulator performed log map demodulation to generate soft values fed to the channel decoders. The turbo decoder used log max decoding with eight iterations and a scale factor of 0.75 was applied to the extrinsic information

In addition to the AWGN channel, the following power delay profiles were used to generate channel realisations:

- Rayleigh fading independent from carrier to carrier
- 3GPP Spatial Channel Model
    - Suburban Macro with 6 paths
    - Urban Macro with 60 paths

For each MCS, between 22 and 40 channel realisations from these power delay profiles were generated and kept constant, i.e., the speed was set to zero. For each channel realisation, the noise power was then varied. For each pair of channel realisation and noise power at most 10000 (50000 in AWGN case) or 200 erroneous blocks were simulated. For each block a new random interleaver was generated in addition to a noise realisation. The AWGN simulation results are the only results used to determine the mapping from effective SNR to BLER. Since this mapping should cover the relevant operating region with some margin and it will be extrapolated as described in Section 4.1.6.2, a larger number of blocks are simulated so that lower BLERs with reasonable confidence interval can be determined.

4.1.4.2 Simulation Results

First, the parameters of the model, i.e., the look-up table and the scaling parameter, were trained separately for each MCS. MIESM validation results for MCS 4 and MCS 6 are plotted in Figure 4.5 and Figure 4.6 respectively. Each figure contains the following plots:

- The BLER as a function of the average SINR (top left)

- The BLER as a function of the effective SINR with tuned scaling parameter (top right)
- A scatter-plot of the difference in average SINR between the trained model and the link level simulation for simulated BLERs in the range [0.003,0.3] (bottom left)
- A histogram of the difference in average SINR differences (bottom right)

Further, in the two top plots the following line-styles for the different channel models:

- Crosses (x) joined with lines are AWGN simulations,
- Circles (o) joined with lines are fading independent from carrier to carrier,
- Cyan dots are 3GPP SCM Suburban Macro with 6 paths
- Red dots are 3GPP SCM Typical Urban with 60 paths

As can be seen from the top left plots in the Figures, the performance in terms of BLER as a function of average SNR depends on the channel realization. There is a significant difference between the AWGN results and the other channel realizations. However, as can be seen from the top right plots of the Figures, the performance in terms of BLER, as a function of effective SNR, is similar for all channel realizations indicating that the mapping from effective SNR to BLER is suitable. To quantify the model error, the trained quality model is executed for the same channel realizations. The difference in average SNR between the link level simulation and the simulation of the quality model is then taken as an error for each simulated BLER point, see Section 4.1.6.3. The scatter plots (bottom left) show how the error correlates with the BLER whereas the histograms (bottom right) show the distribution of the errors for all BLERs in the region considered, between 0.3% and 30%. As can be seen, the error in required average SNR predicted by the trained quality model is rather small, less than 0.2 dB for all simulated points in the operating region.

The results for all the MCS are summarised below in Table 4.2 not only for MIESM but also for EESM and CESM. It is noted that the accuracy of MIESM is as good as or better than EESM, especially for 16QAM. This is perhaps not so surprising since the derivation of the EESM is not directly applicable to higher order modulation such as 16QAM. However, the BICM mutual information expression appears to take the multi-state nature of the higher order modulation into account in a suitable way. It can further be seen that the CESM has lower accuracy, especially for MCS 5 and 6. It should also be mentioned that the cost function used to train the scaling parameter β was a decreasing function of β so that no unique minima were found when training CESM for MCS 5 and 6.

Table 4.2: Summary of results with modulation and code specific model parameters.

	MIESM			**EESM**			**CESM**		
MCS	Tuned β	Average absolute error (dB)	95 perc. absolute error (dB)	Tuned β	Average absolute error (dB)	95 perc. absolute error (dB)	Tuned β	Average absolute error (dB)	95 perc. absolute error (dB)
1	1.235	0.04	0.09	0.911	0.04	0.12	0.094	0.26	0.65
2	1.250	0.04	0.12	6.150	0.19	0.53	0.437	0.19	0.60
3	1.061	0.03	0.09	1.464	0.04	0.10	0.588	0.03	0.07
4	1.097	0.01	0.02	3.578	0.04	0.11	1.020	0.02	0.05
5	1.13	0.06	0.16	1.844	0.10	0.27	0.001	2.28	3.51
6	1.109	0.07	0.19	8.223	0.23	0.55	0.001	1.90	2.75

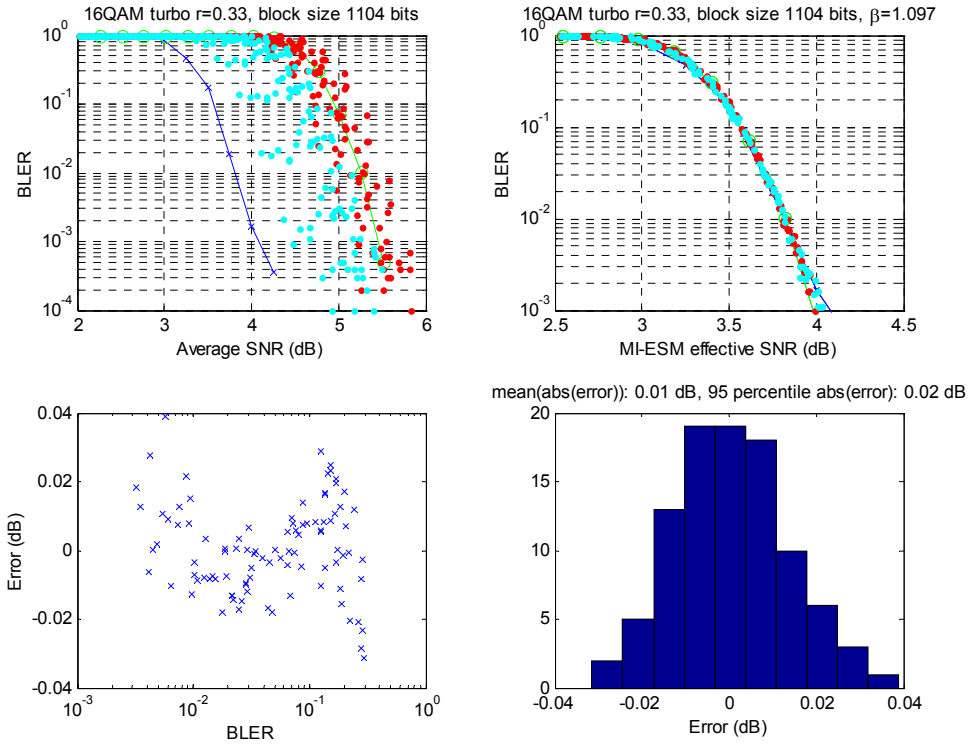


Figure 4.5: Simulation results for MIESM and MCS4, rate 1/3 turbo code and 16QAM

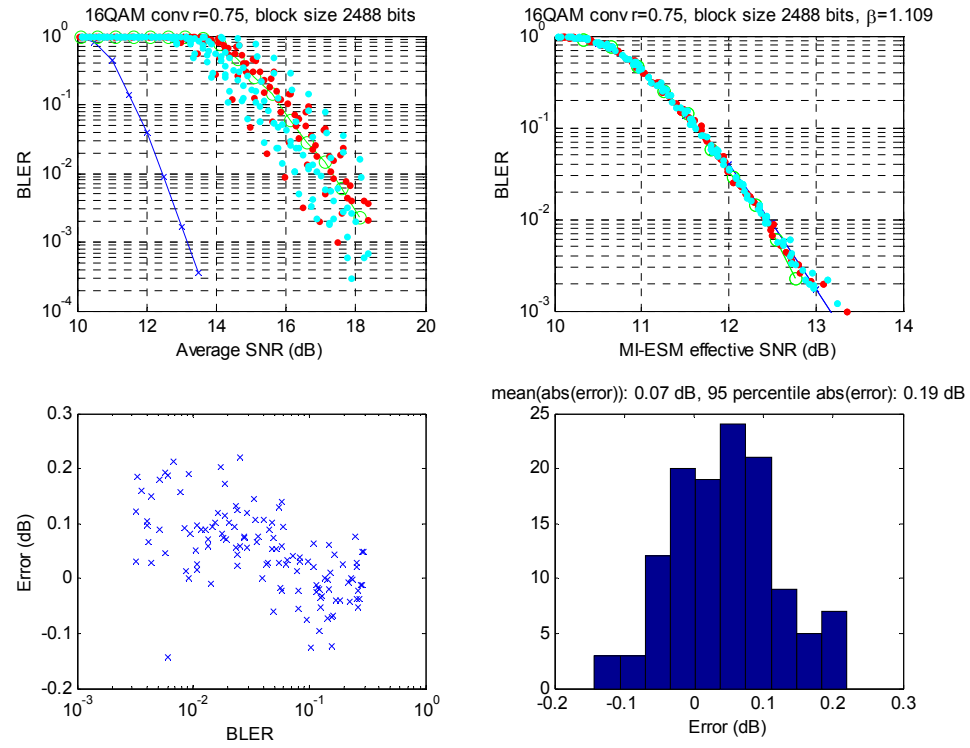


Figure 4.6: Simulation results for MIESM and MCS 6, rate 3/4 punctured convolutional code and 16QAM

As can be seen from Table 4.2, the scaling parameters for MCSs with the same coding scheme are indeed similar, and this suggests using a modulation independent coding scheme parameter β. This is of interest when it comes to modelling adaptive modulation schemes and has been validated in the simulation campaign. The applicability of this is illustrated by training and validating the model for pairs of MCSs, i.e., one set of parameters for MCS 1 and 2, another for MCS 3 and 4 as well as yet another set for MCS 5 and 6. An example for MCS 1 and 2 using BPSK as reference modulation is shown Figure 4.7. The results for all three pairs of MCSs are also summarised in Table 4.3.

Table 4.3: Summary of results with modulation independent parameters. MCS 1+2 use BPSK modulation as reference modulation whereas MCS 3+4 and MCS 5+6 use QPSK as reference

	MIESM		
MCS	Tuned β	Average absolute error (dB)	95 perc. absolute error (dB)
1+2	1.233	0.04	0.10
3+4	1.057	0.04	0.10
5+6	1.133	0.06	0.15

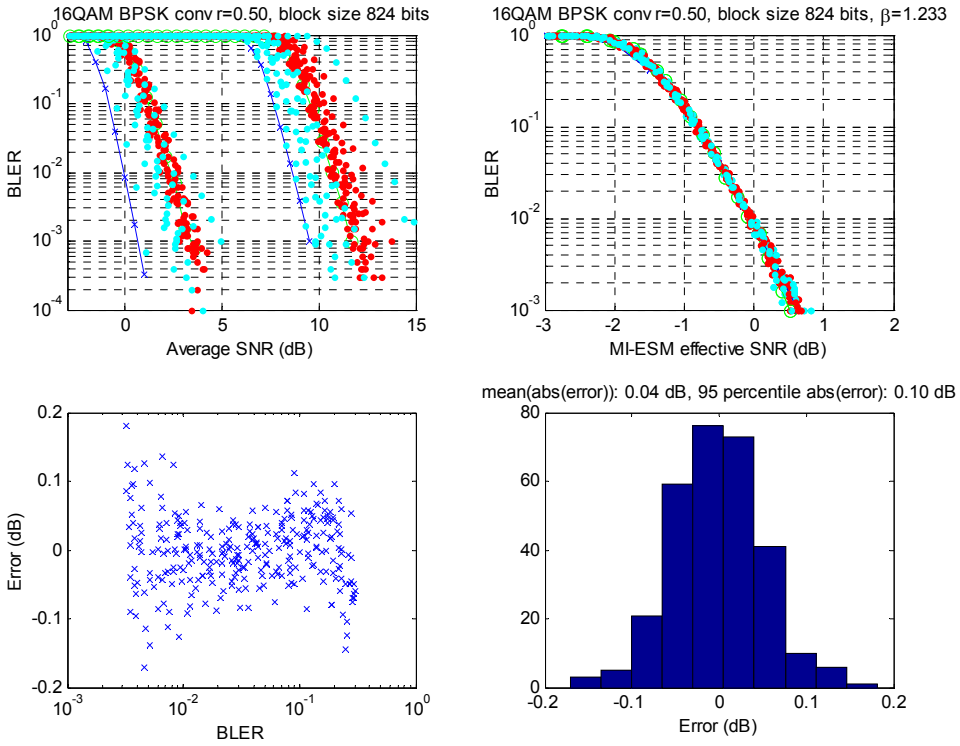


Figure 4.7: Simulation results for MIESM and MCS 1+2, rate 1/2 convolutional code and BPSK and 16QAM

4.1.5 Complexity issues

In the general case, the proposed quality model structure may seem overwhelming; quality measures have to be calculated for a possibly very large number of resource elements and space-time-frequency equalisers and imperfections such as channel estimation errors are in essence simulated on system level.

In this section, possible complexity reduction approaches are discussed.

4.1.5.1 Decimation of the Number of SINR Samples

Besides the SINR calculation itself, the involved complexity is obviously determined by the number of different SINRs on the resource elements required to calculate the effective SINR. This number has to be determined as a trade-off between complexity and reliability.

For the example of SISO OFDM, only a reduced number of sub-carriers may be considered to determine the effective SINR. The frequency offset for the *p*-th sub-carrier involved in the effective SINR calculation is then given by

$$f_p = \frac{N_C}{N_{bin}} \Delta f \left(p - \frac{N_{bin}+1}{2} \right), \tag{4.25}$$

where

Δf is the sub-carrier separation.

N_C is the number of useful sub-carriers of the OFDM system.

N_{bin} is the number of frequency bins used to determine the effective SINR.

The number of frequency bins required to obtain a reliable effective SINR depends on the coherence bandwidth of the channel, which can be approximated by

$$B_{coh} = \frac{1}{\tau_{\max}}, \tag{4.26}$$

where τ_{max} is the maximum relative channel delay. As a rule of thumb, the frequency spacing between two SINR values should not exceed half the coherence bandwidth. Note that this condition may introduce a channel dependence in the interface, which is clearly not desirable. To avoid this channel dependence a worst case design w.r.t. the expected channel realisations should be implemented.

Note that this reasoning is only based on the variation of the channel transfer function itself. It does not take into account variations of the interference powers and a modification of the transfer function due to MIMO processing. In any case, it has to be ensured that the reduced set of SINR samples provides a sufficiently detailed image of the varying SINR within a coded block.

4.1.5.2 Avoid Re-Conversion to SINR

In EESM and MIESM, the SINRs are averaged in the exponential or in the mutual information domain. A simplification can here be obtained by using BLER look-up tables directly in the respective domain. Then the re-conversion using the natural logarithm in EESM or the inverse mutual information function in MIESM can be avoided.

Note that this simplification also holds for the training of the model. If the error minimisation is based on the BLER, the other axis is of minor importance and can be either an effective SINR or the exponential or mutual information domain.

4.1.5.3 Average Value Interface

As opposed to an actual value interface as proposed in the precious sections of this chapter, it is suggested to use the so-called *average value interface* (which is of much lower complexity) if it is adequate for evaluations of certain system concepts. This should be the case where system-level analysis aims at longer-term (macro) system functionalities and therefore link level shorter-term (micro) fluctuations can be assumed averaged out. The actual value interface may be used to *train* this average value interface given that the quality model is adequate.

4.1.6 The Selected WINNER Quality Model - Overview

4.1.6.1 L2S Interface Methodology Outline

The selected WINNER quality model provides a method to predict code word error rates (also denoted as block error rate, BLER, or packet error rate, PER) in system-level simulations without actually performing (de)modulation, (de)interleaving and (de)coding explicitly, but by utilizing pre-computed curves,

which 'map' a certain system realisation – through an effective SINR value – to link quality (BLER/PER).

The basic procedure to be conducted for any code word of interest is depicted in Figure 4.2 and comprises three steps:

- Step 1: Computation of the SINR values for all data symbols (or a subset with sufficient statistics) of the considered code word.
- Step 2: Compression of the set of SINR values obtained in step 1, to a single effective SINR value using averaging in the "mutual information domain":

$$\mathrm{SINR}_{\mathrm{eff}} = \beta \cdot I_{m_{\mathrm{ref}}}^{-1}\left(\frac{1}{P_u}\sum_{p=1}^{P_u} I_{m_p}\left(\frac{SINR_p}{\beta}\right)\right), \tag{4.27}$$

with

$$I_{m_p}(x) = m_p - E_Y\left\{\frac{1}{2^{m_p}}\sum_{i=1}^{m_p}\sum_{b=0}^{1}\sum_{z\in X_b^i}\log\frac{\sum_{\hat{x}\in X}\exp(-|Y-\sqrt{x}(\hat{x}-z)|^2)}{\sum_{\tilde{x}\in X_b^i}\exp(-|Y-\sqrt{x}(\tilde{x}-z)|^2)}\right\}, \tag{4.28}$$

and the definitions

parameter	Definition
P_u	number of transmitted/used data symbols per codeword
$SINR_p$	SINR for symbol p
m_p	bits per symbol p
m_{ref}	average number of transmitted bits per data symbols, i.e., $m_{ref} = \frac{1}{P_u}\sum_{p=1}^{P_u} m_p$
X	Set of 2^{m_p} data symbols
X_b^i	Set of symbol for which bit i equals b
Y	Zero mean unit variance complex Gaussian variable
β	Free parameter; to be optimised

Note that *the use of the BICM capacity expression depends on the demodulator used. If for example an approximate demodulator is used to calculate log likelihood ratios, another capacity expression would apply.*

- Step 3: Mapping of the effective SINR value obtained in step 2 to the corresponding BLER value using either a look-up table or an approximate analytical expression if available.

  Note that the mapping function depends on the packet length as well as the coding scheme. Generally, for any combination different mapping functions have to be provided. It is also possible to use different mappings for different modulations, but the results obtained so far indicate that modulation independent mappings may be used at a relatively small cost in terms of accuracy. The recommendation is though to always validate the trained model as described in Section 4.1.6.3 to ensure that modulation independent mappings can be used.

4.1.6.2 Training the Model

For each coding scheme, including block length and code rate, the L2S interface model has two parameters, which have to be determined in advance based on link level simulations:

- The mapping from SINR to BLEP obtained from an AWGN simulation with the reference modulation
- The scaling parameter β obtained by training the model as described below

Link level simulations with an AWGN channel and the reference modulation are used to determine the mapping from SINR to BLEP. Simulations are performed to cover the relevant operating region (e.g., 3% target BLER) with some margins, say a decade above (30%) and below (0.3%). Linear extrapolation and cubic interpolation are then used (in the logarithm domain) to extend the mapping for SINRs outside the range of the simulated SINRs and to interpolate for SINRs between the simulated points. An alternative implementation, which has been found suitable for codes such as convolutional codes, will use a polynomial approximation (again in the logarithmic domain) to cover a suitable range of SINRs in addition to extrapolation outside this range.

To train the scaling parameter β, L channel realisations are used to create multi-state channels with SNR variations during the encoded block. For each channel realisation, link simulations are performed for several different noise powers, where for each code block simulated a new random channel interleaver is generated. In this way the average block error rate is determined for different signal to noise ratios for each channel realisation. If the coding scheme is to be used in combination with different modulation alphabets, simulations are also done with different modulation.

An operating region of a decade above and below a target BLER of 3% is considered, and therefore only simulated points in the region 0.3%-30% are kept. In this way, for each channel realisation $l=1,..,L$, N_l simulated BLER values are used to train the adjustment factor β. More specifically, the adjustment factor is chosen as

$$\hat{\beta} = arg \min_{\beta} \sum_{l=1}^{L} \sum_{n=1}^{N_l} \left(\mathrm{BLEP}_{pred,l,n}(\beta) - \mathrm{BLER}_{sim,l,n} \right)^2 , \tag{4.29}$$

where $\mathrm{BLER}_{pred,l,n}(\beta)$ is the predicted BLEP for a candidate value of the adjustment factor β and $\mathrm{BLER}_{sim,l,n}$ is the corresponding simulated BLER.

In case that a parameterised model for the SINR to BLEP mapping is used, such as a polynomial approximation as mentioned above, the parameters of the mapping may be trained jointly with the scaling parameter β. The current base line approach is though to determine the mapping from SINR to BLEP from the AWGN simulation alone.

4.1.6.3 Validating the Trained Model

To assess the accuracy of a trained model, link level simulation results for a number of channel realisations are used. For each channel realisation, simulations with varying noise power are performed to determine the BLER as a function of the linear average SINR. For P carriers, this linear average SINR is defined as

$$\overline{SINR} = \frac{1}{P} \sum_{p=1}^{P} SINR_p , \tag{4.30}$$

where $SINR_p$ is the SINR of carrier p. Further, for each channel realisation, the trained quality model can also be executed with varying noise power, so that the BLEP as a function of the average SINR may be determined. Points for which the BLER is within the "operating region" between 0.3% and 30% are considered and all available channel realisations are used in the validation.

As "error" for each simulated BLER point, we take the difference in average SINR between the link level simulation and the simulation of the trained quality model, for which the BLEP coincides with the BLER. It is then important to verify that the error is reasonable for the target BLER operating region.

4.1.7 Conclusions and Open Issues

In this section the methodology and implementation guidelines for interfacing link and system-level performance within the WINNER project were presented.

First, the objectives of the link and system-level simulations within WINNER, the purpose of quality models as performance metrics and a number of candidate quality models were discussed. For the actual interface implementation a number of compression functions and mapping to packet error rate performance options were described. Validation results on the proposed quality model and mapping functions

have been provided in order to demonstrate the training and validation of the interface. Finally complexity issues were discussed.

A brief set of instructions on how to use the selected quality model and mapping was given. The computation of the selected quality model, SINR, along with the mutual information expression, to be used as the compression function, was described. Guidelines on the training of the model and validation of its accuracy were provided. Note that the actual proposed expression of the compression function based on the mutual information depends on the demodulator and should be modified accordingly when a different demodulator is employed.

Based on these guidelines, the partners in WINNER are expected to use a unified approach as baseline for calibration of and simulations with their system-level simulators. However, for the actual implementation, partners may introduce different mapping functions to suit their transceiver architecture, or simplification and approximations appropriate for the scope of their analysis.

The current model does is not outlined for HARQ type II with chase combining or incremental redundancy and estimation errors, such as channel estimation errors, are not considered.

We also expect that the validation measure and method is refined as the understanding of system performance dictating characteristics is developing. Ideally, the error measure should give an indication of the accuracy of the system-level simulation results.

5. Simulation Assumptions

This chapter summarises the simulation parameters and assumptions that have been applied when generating the simulations results in Chapter 7. However, it should be mentioned that an absolute comparability among simulation results from different WINNER project partners could not be accomplished mainly due to lack of time for calibration of the different simulation chains. As a workaround, the simulation results for the different multi-antenna methods are collected relative to the single antenna case, and thus a relative comparison is achieved by the use of common parameters.

The system parameter assumptions presented in this chapter are to a large extent aligned with the draft system concept proposal, assumptions in WP2 [T2.7]. Basically, two different system parameter settings are considered denoted as *wide-area* and *short-range* according to their application scenarios. Hence the chapter is structured according to this classification.

Obviously, these scenarios may be further subdivided into "subscenarios" according to the detailed environment characteristics, propagation conditions, user density and mobility distribution, traffic type, etc. The attempt was made to keep the number of different subscenarios reasonable small to shorten simulation time. Moreover, some simplifications have been applied mainly with respect to the user mobility model and traffic characteristics.

Note that an extensive discussion of the different parameters options is omitted here for the sake of brevity. More details regarding this issue can be found in [T2.5]. In the following sections of this chapter the main simulation assumptions are summarised in table format. Some of the parameter values therein are not defined as decisions on the WINNER system concepts are not yet finalised.

The assumptions presented in this chapter will be used for the simulation studies presented in chapter 7. Possible deviations from this set of assumptions or additional assumptions will be described in detail in the corresponding section of Chapter 7.

5.1 Wide-Area scenario

5.1.1 Carrier Frequency System Parameters

Parameter	DL	UL
Centre frequency	5.384 GHz	4.616 GHz
Bandwidth/ channelisation	20 MHz	20 MHz
Duplex scheme	(T+F)DD	
Average transmit power	20 W per site	200 mW
Transmission technique	OFDM Δf=19531.25 Hz $\rightarrow T_N$=51.2 µs f_c=kΔf, k∈[-416:416], k≠0 T_G=5 µs $\rightarrow T=T_N+T_G$ =56.2 µs	not specified
FEC	convolutional $(561,753)_{oct}$	
code rates	r=1	r=2/3
corresponding puncturing pattern	1 1	1 0 1 1
interleaver	Random	
modulation alphabet	BPSK, QPSK, 16-QAM, 64-QAM	
chunk size	n_{frame}= 6 n_{sub}= 8	
pilot/information carrying units per chunk	12/36 resource elements ⇒ 25% overhead for parameter estimation	

<table>
<tr><th>Parameter</th><th colspan="2">DL</th><th>UL</th></tr>
<tr><td>MCS and corresponding packet lengths</td><td colspan="3">
<table>
<tr><th>MCS</th><th>modulation</th><th>code rate</th><th>data bits</th><th>coded bits (incl. tail bits)</th><th>number chunks</th></tr>
<tr><td>1</td><td>BPSK</td><td>½</td><td>460</td><td>936</td><td>26</td></tr>
<tr><td>2</td><td>QPSK</td><td>½</td><td>928</td><td>1872</td><td>26</td></tr>
<tr><td>3</td><td>16-QAM</td><td>½</td><td>1864</td><td>3744</td><td>26</td></tr>
<tr><td>4</td><td>64-QAM</td><td>2/3</td><td>3736</td><td>5616</td><td>26</td></tr>
</table>
</td></tr>
<tr><td>multiple access</td><td colspan="2">– non-adaptive TDMA+OFDMA
– adaptive/non-adaptive SDMA according to spatial processing technique under investigation</td><td>not specified</td></tr>
<tr><td>Number of sectors</td><td colspan="2">3</td><td>1</td></tr>
<tr><td>Sector antenna pattern</td><td colspan="2">$A(\theta) = -\min\left[12\left(\frac{\theta}{\theta_{3dB}}\right)^2, 20\right]$ [dB]
θ3dB = 70o</td><td>omni directional</td></tr>
<tr><td>Antenna configuration per sector</td><td colspan="3">ULA with vertical/cross polarised antennas, see below</td></tr>
<tr><td></td><td colspan="2">AW0: |
AW3: | | | | (0.5λ)
AW6: X X (4λ)
AW7: | | | | | | | | (0.5λ)</td><td>TW0: |
TW1: | | (0.5λ)
TW2: | | | | (0.5λ)</td></tr>
<tr><td>To be considered in link level simulations</td><td colspan="3">a. AW0 - TW0 (reference)
b. AW3 - TW0
c. AW6 - TW2
d. AW6 - TW1
e. AW7 - TW0 (beamforming/SDMA)</td></tr>
<tr><td>To be considered in system level simulations</td><td colspan="3">1) AW0 - TW0 (reference)
2) AW6 - TW0 50%, TW1 30%, TW2 20% (diversity/SMUX)
3) AW3 - TW0 50%, TW1 30%, TW2 20% (beamforming)
4) AW7 - TW0 50%, TW1 30%, TW2 20% (beamforming/SDMA)</td></tr>
<tr><td></td><td colspan="3"></td></tr>
<tr><td>channel estimation</td><td colspan="3">perfect</td></tr>
<tr><td>synchronisation</td><td colspan="3">perfect</td></tr>
<tr><td>Baseline detector</td><td colspan="3">MMSE equalisation (if required) with ML based soft output demodulation
(ML criterion is applied to equaliser output signal if present)</td></tr>
<tr><td></td><td colspan="3"></td></tr>
<tr><td>Adaptivity</td><td colspan="3">– adaptive MCS selection per superchunk (=26 chunks)
– optional: adaptive MCS selection per chunk
– optional: adaptive resource assignment
– adaptivity of spatial processing according to method under investigation</td></tr>
<tr><td>CSI quality</td><td colspan="3">– perfect
– optional: Gaussian error model</td></tr>
<tr><td>(H)ARQ</td><td colspan="3">no</td></tr>
<tr><td>Power control</td><td colspan="2">no</td><td>perfect</td></tr>
</table>

5.1.2 Scenario Parameter

Parameter	C.2 (typical urban)	
channel model	WINNER WP5 Urban Macro (NLOS only)	
Path loss	PL(d)=28.3·log10(d)+61.5	
Shadow fading	Log-normal with standard deviation σ_{SH}=5.7dB	
Cell range	500m· 2km	
Cellular structure	hexagonal grid	
User distribution	uniform in two dimensions	
Number of users	1, 2, 4, 8, 16, 32	
User mobility model	50%: 0 km/h 50%: 70 km/h	80%: 0km/h 20%: 250km/h
Traffic model	full queue assumption	
intercell-interference level	not specified	not specified

5.2 Short-Range Scenario

5.1.1 Carrier Frequency System Parameters (general WINNER assumptions)

Parameter	DL	UL
Centre frequency	5.0 GHz	5.0 GHz
Bandwidth/ channelisation	100 MHz	100 MHz
Duplex scheme	TDD	
Average transmit power	1 W per site	200 mW
Transmission technique	OFDM Δf=48828.125 Hz $\rightarrow T_N$=20.48 µs f_c=kΔf, k∈[-832:832], k≠0 T_G=0.8 µs $\rightarrow T=T_N+T_G$ =21.28 µs Optional: Δf=195312.5 Hz $\rightarrow T_N$=5.12 µs f_c=kΔf, k∈[-208:208], k≠0 T_G=0.8 µs $\rightarrow T=T_N+T_G$ =5.92 µs	not specified
FEC	convolutional $(561,753)_{oct}$	
code rates	r=1	r=2/3
corresponding puncturing pattern	1 1	1 0 1 1
interleaver	Random	
modulation alphabet	BPSK, QPSK, 16-QAM, 64-QAM	
chunk size	n_{frame}=15 n_{sub}= 16	
pilot/information carrying units per chunk	60/180 resource elements ⇒ 25% overhead for parameter estimation	
MCS	MCS 1: BPSK, ½ MCS 2: QPSK, ½ MCS 3: 16-QAM, ½ MCS 4: 64-QAM, 2/3	
multiple access	– non-adaptive TDMA+OFDMA – adaptive/non-adaptive SDMA according to spatial processing technique under investigation	not specified

Parameter	DL	UL
Number of sectors	3	1
Sector antenna pattern	$A(\theta) = -\min\left[12\left(\frac{\theta}{\theta_{3dB}}\right)^2, 20\right]$ [dB] $\theta_{3dB} = 70^o$	omni directional
Antenna configuration per sector	ULA with vertical/cross polarised antennas, see below AS0: | AS5: | | | | (0.5λ)	TS0: | TS1: | | (0.5λ) TS3: | | | | (0.5λ)
To be considered in link level simulations	1) AS0 - TS0 (reference) 2) AS5 – TS3	
To be considered in system level simulations	1) AS0 - TS0 (reference) 2) AS5 - TS0 20%, TS1 40%, TS3 40%	
channel estimation	perfect	
synchronisation	perfect	
Baseline detector	MMSE equalisation (if required) with ML based soft output demodulation (ML criterion is applied to equaliser output signal if present)	
Adaptivity	− adaptive MCS selection per superchunk (=? chunks) − optional: adaptive MCS selection per chunk − optional: adaptive resource assignment − adaptivity of spatial processing according to method under investigation	
CSI quality	− perfect − optional: Gaussian error model	
(H)ARQ	No	

5.2.2 Carrier Frequency System parameters (to be considered in this deliverable)

Parameter	**Value**	**Unit/Notes**
Centre frequency	5	GHz
Number of subcarriers in OFDM	64	Equals the length of FFT
FFT BW	20	MHz
Signal BW	15	MHz
Number of data subcarriers	48	[-24:24] , subcarrier 0 not used
Sub-carrier spacing	312.5	KHz
OFDM symbol length(excluding cyclic prefix)	3.2	μs
Cyclic prefix length	0.8	μs
Physical Frame size	15 x (96+18)	KHz x μs (TDD guard interval 18μs)
Frame size in symbols	48 x 24	Frequency x time

5.2.3 Scenario Parameters

Parameter	Typical office
channel model	IEEE 802.11n C-LOS, D-NLOS
Path loss	Included in the model
Shadow fading	Log-normal with standard deviation σ_{SH}=3,5dB
Cell range	30m
Cellular structure	Isolated cell
User distribution	uniform in two dimensions
Number of users	1, 2, 4, 8, 16, 32
User mobility model	0 Km/h
Traffic model	full queue assumption
intercell-interference	Not specified

6. Assessment Criteria and Performance Measures

This chapter defines the assessment criteria and performance measures as they are used for this document. Since many system parameters are not yet defined and no common WINNER simulator is available, apart from absolute figures of merit many result discussions will be based on *relative* figures of merit F_r, which are based on the quotient of the performance of the investigated multi-antenna technique F_x to the corresponding performance of the calibration scenario F_{cal}:

$$F_r = \frac{F_x}{F_{cal}},$$

where F is a placeholder for the assessment criteria defined in the sequel. A number of assessment criteria are defined as CDFs/PDFs. In the presentation of results of the test configurations specific measurement points will be shown in addition to the plot of the distribution to facilitate comparison. Unless otherwise stated these measurement points are the 95th, 50th and 10th percentiles of the CDF, inline with [D7.2] .

The main focus of simulations within T2.5 is on the multi-antenna aspects, which requires detailed simulation of the spatial processing algorithms. Therefore multiple users, their local distribution and antenna parameters need to be simulated including slow and fast link adaptation, as well as slow and fast resource allocation. This means that the simulators need to resolve the single resource elements in space (beams, antennas), frequency (subcarrier, group of subcarrier), and time (code block, OFDM symbol). Due to the involved implementation efforts, simulation complexity, and runtime an evolutionary approach for the simulation assumptions has been adopted (see Chapter 4). In particular, all simulations for D2.7 will adopt a simplified traffic model based on the full queue assumption without detailed modelling of packets. Most investigations will be based on snapshots without short-term evolution of the individual and uncorrelated drops of large-scale parameters (like user location, channel realisation, etc), which reduces interpretability and requires adaptation of many assessment criteria. Apart from link-level results, mostly multi-link single-site simulations are performed, which still suffer considerable limitations since intercell interference is not accounted for properly. However, in this document also first system-level simulations based on several sites will be compared to single-site simulations to investigate the importance of accurate intercell interference modelling.

The full queue traffic model can be related to any other traffic model considering the following impacts

- The load offered from higher layers at each time instant is the summation of the maximum data rate supported on each user connection over all existing terminals. When comparing results to simulations with other traffic models, the statistics of the offered load should be comparable.
- For a given number of users per cell, the full queue traffic model offers maximum multi-user diversity. A fair comparison in terms of multi-user diversity to other traffic models should be based on an equivalent average number of user that have concurrently ongoing packet calls (i.e., active users that have data in the transmit buffer). Therefore the number of users shown in the results of this document must conceptually be seen as the number of concurrently active users for a given traffic model. However, a fair comparison between techniques, in terms of load and multi-user diversity is maintained and the full queue assumption does not create any imbalance or bias of the relative performance results that will be included in this document
- Conceptually the full buffer model is most closely related to traffic like ftp, while packet data traffic with irregular patterns of short and burst packet calls and reading times has significantly different characteristics.
- In general there is a mismatch in size between the encoded and modulated higher layer packets and the available channel resources, which is either accounted for by fine-tuning of the code rate (puncturing/repetition) and/or by padding. The effect of these granularity issues is not accounted for in simulations without detailed modelling of packets. However, it is anticipated that for any reasonable system design, these impacts are minor
- Packet delay cannot be investigated directly, however, at least for high-data rate traffic, there is normally a correlation between user throughput and user packet delay.

The use of snapshot simulations and a full queue traffic model without detailed modelling of packets requires adaptation and redefinition of some figures of merit compared to the definition in [D7.2]. The

exact definition of the assessment criteria in T2.5 is given in the following. Once corresponding traffic modelling approach is available and the simulator capabilities allow for it, the assessment criteria used in T2.5 will be aligned with [D7.2].

6.1 User Throughput

The user throughput T_u is defined as the ratio of correctly received information bits on one link to one user to the total simulation time (i.e., the time that elapses in the real system) for this link. Statistics are collected from all links within the evaluation area (e.g., centre cell or original cells in case of simulator using wrap-around technique). The cumulative distribution function (CDF) of these user throughput values is provided and comparisons are based on the corresponding percentiles.

For pure snapshot simulations each link exists only the minimum time step of the simulator (e.g., one OFDM symbol). Therefore any user that is not scheduled instantaneously will have zero user throughput and any TDMA component of the scheduling, i.e., any averaging of user throughput over time, will not be considered. Consequently reliable user throughput values can only be obtained by averaging over an ensemble of instantaneous realisations with identical characteristics, such as distance, as detailed in the following section.

6.2 Coverage and Fairness

Coverage and fairness with respect to user location are investigated using the scatter plot of user throughput versus normalised distance. The distance is normalised to the cell radius and distance bin are equally spaced every 1% of the distance. For each distance bin the average user throughput is calculated and plotted.

In a similar way, user throughput values are often accumulated and averaged for user realisations having similar path loss and shadowing. Since fairness is to a large extend introduced by the scheduling algorithm, which are out of the scope of this document (mostly a simple Round Robin scheduling will be used), this presentation is not adopted in this document.

6.3 Cell Throughput

The cell throughput T_c is defined as the aggregate number of correctly received information bits within one cell per simulation time step. Samples are taken at each time step from all cells within the evaluation area (e.g., centre cell or original cells in case of simulator using wrap-around technique). From these samples the CDF of cell throughput is calculated. Cells are defined as the parts of a site with a fixed allocation of antenna resources. A site itself is the location of the base station hardware.

6.4 Spectral Efficiency

Given the limitations in the simulator capabilities, spectral efficiency values cannot be obtained with high accuracy. In the following we first present a possibility to obtain spectral efficiency along the methodology in [D7.2], [UMTS30.03] under the constraint of full queue simulations consisting of pure snapshots of uncorrelated drops of long-term parameters. In particular the definition of the satisfied user criterion becomes difficult in this case. In the last part we provide a pragmatic approach to obtain spectral efficiency values based on the CDF of user throughput obtained from simulations including short-term evolution.

6.4.1 Basic Spectral Efficiency Calculation

The spectral efficiency is defined similar to [D7.2], [UMTS30.03]. A single service scenario is assumed in the simulations. Therefore the average normalised throughput can be expressed as

$$v_{site} = \frac{\sum_{k=1}^{N_c} \overline{T}_c(k)}{N_s B},$$

where the average cell throughput for all cells k is summed up and divided by the number of sites N_s and

the bandwidth B. The average normalised throughput is equivalent to the term "system load" in [D7.2], [UMTS30.03]. Note that in case of sectorisation, the number of cells N_c is a multiple of N_s. The site is defined as physical co-location of hardware serving the set of antennas. Users may be connected to a site either directly or through relay stations. Both types should be taken into account when calculating the spectral efficiency of a site.

The site spectral efficiency $v_{site,\ 98\%}$ is the average normalised throughput for which 98% of the users are satisfied according to the satisfied user criterion below.

The area spectral efficiency $v_{area,\ 98\%}$ is obtained by multiplying the site spectral efficiency by the number of sites N_s per covered area A:

$$v_{area,98\%} = v_{site,98\%} \frac{N_s}{A}.$$

For single-cell or single-site simulations an area spectral efficiency cannot be predicted with sufficient accuracy due to the simplified intercell interference modelling. In this case only site spectral efficiency is used. The spectrum efficiency is given separately for uplink and downlink.

6.4.2 Satisfied User Criterion

The satisfied user criterion is normally based on a threshold of the active session throughput (or packet call throughput), i.e., based on an average user throughput. In snapshot simulations each user exists only for one time instant and therefore no averaging over time can be performed. Any TDMA component of scheduling will not be considered and any user that is not scheduled instantaneously is necessarily a dissatisfied user. Therefore a straightforward application of the common satisfied user criterion as in [D7.2], [UMTS30.03] is not applicable.

In order to overcome this shortcoming of snapshot simulations, we will resort to the use of average user throughput with respect to distance. Assuming a homogenous user density, 98% of the users will be located within a sector having the radius of $\sqrt{0.98}$ times the nominal cell radius. Assuming a monotonically decreasing function of average user throughput versus normalised distance, we can define a satisfied user criterion as follows:

We assume that 98% of users are satisfied if the average user throughput versus normalised distance is above 500 kbps for all normalised distances less or equal than $\sqrt{0.98}$. Note that the average user throughput of 500 kbps, corresponding to the minimum sustained bit rate in WINNER deliverable D7.1 is also used as threshold for internet or ftp traffic models in [D7.2] .

6.4.3 Spectral Efficiency Calculation for Simulations Using Short-Term Evolution

As an pragmatic approach and in particular as a basis for relative comparisons of different spatial processing techniques or variants, spectral efficiency can be based on the CDF of average user throughput for simulations including short-term evolution. Here the site spectral efficiency is obtained based on the load for which 90% of users have an average throughput greater or equal than 500 kbps. The corresponding average cell throughput is normalised by the bandwidth and further to the number of cells per site or to the area, depending on whether site spectral efficiency or area spectral efficiency is presented.

6.5 Control Overhead

The evaluation of the associated control overhead of each technique in the forward and return link is an important part of the assessment process. While first throughput values are given without considering such overhead, a dedicated discussion will investigate the impact of overhead on the overall trade-off of performance versus complexity/overhead.

At this stage, the overhead due to forward and return link control channels cannot be included directly in the figures of merit, since the control channel structure is unknown. Therefore for each technique the required additional control information is estimated separately for the forward and return link according to the following procedure. For each control information (e.g., spatial mode selector, number of spatial streams used, loading information, etc.) a rough estimate of the number of required bits in the forward and return control channel is given in Table 6.1 and Table 6.2. For each technique under consideration the

required control information is gathered. For each control information the required update rate regarding the frequency, time, space, and user dimension is determined and will be detailed in the simulation result chapters. Note that we focus on control information that is related to spatial processing, or that differs for other reasons amongst the spatial processing techniques considered. Apart from different requirements on control information that originate from the spatial processing technique directly, major differences between the control overhead of different spatial processing techniques will be encountered based on the amount of adaptivity, e.g., whether one or several FEC are used per data flow, and whether, e.g., power and/or bit loading is used. With respect to the link level retransmission overhead we assume no differences between the different techniques and therefore exclude it from our discussion.

Table 6.1: Preliminary estimation of Return Control Channel Overhead

Control Information	Parameter	Estimated Information Bits	Comment
beam selection information	ld(n_{va})	[1, 2, 3] bits	allows to select one virtual antenna stream out of n_{va}
CQI	N_{CQI}	5 bit	allows to report 32 AMC levels, e.g., per spatial substream in adaptive systems,
CSI	$\boldsymbol{N_{CSI}}$	8 bit	transmission of one complex channel coefficient

Table 6.2: Preliminary estimation of Forward Control Channel Overhead

Control Information	Parameter	Estimated Information Bits	Comment
Modulation M(*c,q*)	N_M	2 bits	assuming BPSK, QPSK, 16-QAM, 64-QAM (e.g., per chunk or spatial substream in adaptive systems, per transmission in non-adaptive systems),
Code Rate of FEC_o	N_{FECo}	6 bits	per transmission; identical for all techniques
Code Rate of FEC_i(*c,q*)	$\boldsymbol{N_{FECi}}$	6 bits	if FEC_i is used; per chunk or spatial substream in adaptive systems, per transmission in non-adaptive systems,
Power P(*c,q*)	N_p	6 bits	technique dependent, i.e., per chunk in adaptive systems using power loading, per transmission in non-adaptive systems,
STFC information	ld(N_c)	[1, 2, 3]	e.g., LDC code selection if N_c different codes or in general spatial processing modes are employed
terminal ID		0 bits	dedicated control channel or implicit signalling assumed
resource map	N_{RM}		chunk and spatial layer IDs dedicated to one user in FDMA/SDMA systems; in pure SDMA systems this reduces to a beam/virtual antenna stream map; temporal update per scheduling interval
pilots	N_p		per spatial substream of each chunk if additional pilots are required compared to a chunk in SISO processing

Reasonable update rates in frequency domain include per chunk, per user, per cell. In the time domain per frame (= chunk duration), and per scheduling interval are some natural update intervals. In the spatial dimension signalling might be required, e.g., per layer of a chunk, per chunk, per user, or per cell. Further it needs to be distinguished, whether signalling is required only to/from scheduled users, or from all active users. Thus it will be possible to calculate an average overhead rate, based on the average number information bits per time required in uplink and downlink.

While for most of the control overhead it is at least feasible to obtain a first estimate of the information bits required, the additional requirements for pilots is for further study, see also discussion in Chapter 7. Since the control channel structure is not known at this time, it is only feasible to obtain a rough estimate of the control channel overhead after coding. To this aim, the calculated estimated control information bits per cell and per second are multiplied by $Z = \{1, 2, 4, 8\}$ to obtain optimistic, realistic and pessimistic coded control channel overhead estimates. For all these four estimates the effective user throughput is calculated by subtracting the corresponding coded control channel overhead from the throughput obtained in the simulations.

In this initial stage, also the return channel overhead will be simply compiled and compared per technique, since no dedicated modelling of the return channels will be pursued.

Although this procedure is very preliminary and only allows coarse estimations, it will give us indications,

- Which techniques provide higher efficiency including all required signalling,
- About the trade-off between adaptivity versus control overhead,
- Which techniques are favourable, provided an efficient signalling can be implemented,
- Which control information consumes the major part of the signalling bandwidth and therefore is in the focus for optimisation.

As a very important additional benefit, a detailed listing of all control signalling involved for each technique will help in identifying techniques, which nicely complement each other in the scenarios where they provide high performance but result in low overall overhead in terms of control signalling (since they are based on a similar set of required control information).

6.6 Robustness

Channel estimation errors at the receiver cause degradation in the decoding performance of the current data. Additionally they impact performance of closed-loop multi-antenna techniques since the erroneous measurement is used as basis for CSI feedback for spatial processing. On the feedback path additional errors are introduced by quantisation, signalling errors (i.e., erroneous decoding of the CSI feedback), as well as by delay.

Robustness needs in particular to be investigated for spatial processing techniques using short-term CSI knowledge at the transmitter, like most precoding techniques. To obtain first insight into this topic, a simple Gaussian error model is used to generate the erroneous channel estimates and erroneous CSI feedback. The relative performance loss compared to the ideal case is used to characterise robustness with respect to channel estimation and CSI feedback errors.

6.7 Flexibility and Scalability

The flexibility and scalability of multi-antenna techniques is considered regarding the supported scenarios, the supported user terminal classes, and supported user density and velocity distributions. Due to the importance of these criteria they are discussed in the derivation of the WINNER multi-antenna concept proposal in Chapter 8.

Flexibility and scalability is assessed in terms of the ability of the multi-antenna technique to be easily adapted and perform well in different usage conditions, and scenarios, like short-range, wide-area, relaying and peer-to-peer. Of particular interest is also whether the technique is scalable with respect to the amount of CSI required at the transmitter. The required adaptation and supporting functions to implement

a particular technique in all scenarios is a major criterion that allows to estimate overhead and complexity involved in providing such flexibility.

For the scope of this document, user terminal classes are based on the number of antennas at the mobile terminal. For single-antenna terminals an additional low-end terminal class can be assumed, which only provides basic processing capabilities. If possible, simulations are performed with a mix of user terminal classes as in real scenarios. While quantitative investigations for different terminal classes will be in the focus at a later stage of the project, a first qualitative comparison of different techniques is given in this document.

In order to investigate the impact of different numbers of users per cell and different user speeds, simulations are performed for varying numbers of users and mobility profiles as defined in Chapter 5. Additionally particular remarks on preferable spatial user distributions and user velocities are given along with the technique descriptions.

6.8 System and Terminal Complexity

Clearly spatial processing has impact on system and terminal complexity, which are important criteria for system design. For example, directivity will change applicability of and requirements for pilots. A further enabler for many advanced spatial processing techniques is short-term channel knowledge at the transmitter. This requirement might have critical impact on system complexity, since either feasible ways to exploit channel reciprocity and the corresponding impact on system design need to be elaborated or significant return link control information is required. Additionally the spatial processing techniques differ in baseband and RF complexity and in split of the total complexity into the transmitter and receiver part. A discussion of these complexity aspects can be found in Chapter 8.

7. Simulation Results

7.1 Introduction

The objective in this chapter is to discuss performance assessment issues and present simulation results for the wide-area, the short-range scenario and relaying in order to evaluate the potential gains from the deployment of the multi-antenna concept described in Chapter 2 and the associated complexity and overhead signalling. The evaluation is done in the framework of the WINNER system concept and assumptions.

The performance of the multi-antenna techniques described in Chapter 3 is evaluated in terms of the assessment criteria presented in the previous chapter, namely cell and user throughput, coverage and spectral efficiency and control signalling overhead.

The following three clusters of multi-antenna techniques are investigated:

- Beamforming-oriented techniques, employing SDMA with a grid of fixed beams;
- Spatial multiplexing with limited CSI (PARC);
- Spatial link adaptation across different matrix modulation schemes and application of linear precoding to exploit CSI.

Comparisons of performance gains will be performed with respect to the assessment criteria considered for a number of scenarios and antenna configurations, in order to understand the impact of critical parameters on a certain algorithm. As different simulators (of different partners) have been used for the evaluations of different techniques and calibration of different simulators was not performed, comparisons among different algorithms will only be performed in a quantitative manner in order to evaluate the relative merits.

7.2 Performance Results for the Wide-Area Scenario

7.2.1 Beamforming-Oriented Techniques with a Grid of Fixed Beams

7.2.1.1 Objective

Using fixed beams at the base stations of a cellular system is a relatively simple but yet efficient spatial processing technique, which is in the meantime already employed in third generation mobile networks. Beamforming exploits the relatively small angular spread of the wide-area channel (assuming BS antennas above rooftop), which makes it possible to focus a large portion of the transmitted energy to a restricted angular zone. The obvious advantages are enhanced SNR or coverage with reduced intercell-interference. Moreover, multiple users/terminals can be served simultaneously with controllable interference by forming grids of fixed beams. Clearly, the amount of interference depends on the applied beam pattern as well as the number of beams and the beam distance.

Whereas adaptive beamforming aims at directing one or more beams to preselected user(s), the fixed beam approach works to some extent reversely. Here, the beam or grid of beams (GoB) – to be applied on a certain time-frequency unit, i.e., chunk – is selected first and then the users are assigned to those beams typically based on some fed back channel quality indicator, e.g., SINR. Thus, the "system intelligence" is shifted from spatial processing to resource assignment. The advantage of such a scheme compared to adaptive spatial processing techniques is a higher robustness since misadjustment of the antenna weights due to defective channel state information is excluded.

A fair assessment of the GoB approach requires multi-cell system-level simulations including some adaptive resource assignment functionalities. In this framework, results are presented in Section 7.2.3.4 where the GoB approach is compared to other spatial processing techniques. In this section, however, only single sector multi-link/system-level simulation results are given constrained by the individual simulator capabilities. Therefore, this section focuses on the impact of different parameter settings and environment characteristics on the performance of the GoB approach.

7.2.1.2 Assumptions

Any deviating/additional simulation assumptions with respect to Chapter 5 are listed in Table 7.1 below.

Table 7.1: Applied simulation assumptions deviating from or additional to the assumption specified in Chapter 5 (no input indicates compliance)

Parameter	Multi-link level simulator (**MLS**)	System-level simulator (**SLL**)
Path loss	Not included, all users "see" same noise variance σ^2 (located at same distance to BS)	
Channel model	Variable (Suburban Macro, Urban Macro 8° AS, urban micro)	Suburban Macro
Mobility model	variable but identical velocity of all users	
Transmit power	Normalised to noise power	20/3 W
Frame size	24+1 OFDM symbols (1 dummy control symbol)	24 OFDM symbols
Chunk size	832 subcarriers x 12 OFDM symbols	416 subcarriers x 3 OFDM symbols
Block size	8 blocks per chunk 16 QAM: 2488 bits + 8 (trellis term.) bits 4 QAM : 1240 bits + 8 (trellis term.) bits	A virtual MAC frame length of 3736 data bits is considered. Each MAC frame is mapped to 1, 2, 4 or 8 chunks (depending on the MCS)
MCS	Fixed	adaptive
#of antennas at terminal	Either 100% single omni or 100% 2Rx omni using MRC	
Beam tapering	Chebychev, 21 dB	Variable
Grid shape	Two grids with 4 beams each according to Chapter 3. First grid (with odd beam numbers) is used in odd time slots and second grid (with even beam numbers) is used in even time slots	Variable. The grids are altered every MAC frame (in Round Robin manner)
Resource Assignment	Round Robin (**RR**): - assignment of beam-id's per user on a long-term basis (fixed per drop) - cyclic user activity inside each beam, regardless of channel conditions Best effort (**BE**): based on CQI for a whole frame, allocation based on a greedy algorithm, meaning: - each user feeds back one CQI value per frame for each beam - user with best overall CQI is allocated first on its best beam - take remaining set of unallocated users and beams and continue allocating second best user etc. - allocate until all beams or users are used up or certain CQI threshold is underrun	Round Robin (**RR**): - users are served in a round robin fashion but only if their throughput exceeds a threshold of 0.1bit per resource element Proportional fair scheduling (**PFS**): operates close to (but not identical) "best effort" (**BE**) scheduling due to the small number of MAC frames simulated and the resulting short throughput history.
Throughput (goodput) calculation	Cell Throughput for each SNR: $TP_{cell} = \sum_u (1 - BLER_u) \cdot R_u$, with R_u as the rate of transmitted information bits per second of user *u*	The chunk error rate $BLER_c$ is determined according to the L2S interface approach specified in Chapter 4. The corresponding MAC frame BLER (and the resultant throughput) is derived from $\mathrm{BLER} = 1 - \prod_c (1 - \mathrm{BLER}_c)$
Number of snapshots	30 - 120	200

7.2.1.3 Results

The MLS simulation results for different channel characteristics, i.e., suburban macro, urban macro and urban micro, are shown in Figure 7.1, Figure 7.2 and Figure 7.3, respectively. The dashed black line indicates the maximum possible throughput of 113.3 Mbit/s when 4 beams transmit at the same time without any errors. Different curves are shown for best effort (BE) and Round Robin (RR) resource assignment and varying numbers of users.

In the urban micro scenario, a fixed 4-QAM modulation was used since the throughput with 16-QAM is hardly larger due to inter-beam interference caused by the large angular spread of the urban micro scenario.

The normalised throughput gain of SDMA GOB vs. non-SDMA GOB for an SNR of 15dB is summarised in Table 7.2,

Table 7.3 and Table 7.4 The throughput can be increased up to nearly 4 times with SDMA compared to non-SDMA fixed beams with adaptive scheduling. In case of a suburban macro environment (with inherently small angular spread) the non-adaptive (RR) scheduler reaches factor 3.3 throughput increase. As expected the results show that beamforming for propagation channels with large angular spread is not as promising as for small angular spread. Especially the non-adaptive (RR) resource assignment shows a strong degradation when all beams are used (with larger numbers of users) due to increased inter-beam interference.

Note that the throughput for the urban micro scenario is no monotonous function of the number of users any more for Round Robin in these simulation results. The large angular spread causes here an increasing BLER floor due to mutual cross-talk between beams which is above 40% for 4 QAM when all beams transmit simultaneously.

Compared to the simulator capabilities this effect can be to some extend avoided in a real world system. Better codes (using, e.g., turbo coding instead of convolutional codes) tolerate a larger raw BER floor. Adaptive modulation and coding (which is not implemented in the MLS) also avoid these high BLER floors by adapting to the (at least) long-term channel quality.

Another important parameter especially for large angular spreads is the impact of beam tapering, which was kept constant for the MLS simulations. Here an adaptation of tapering also could improve the performance.

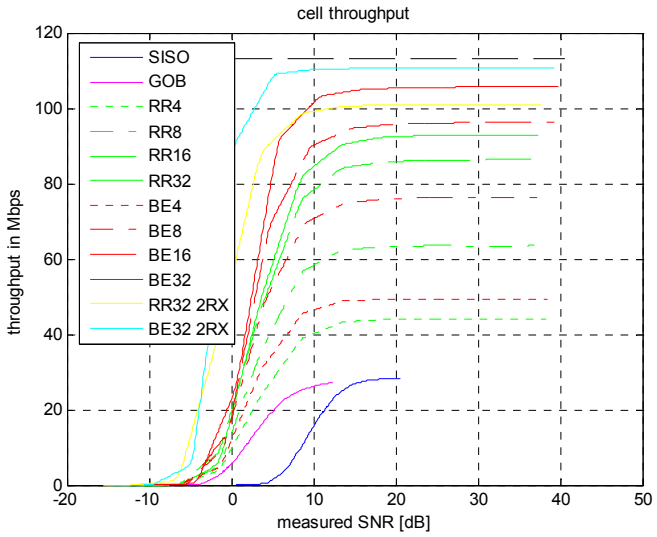


Figure 7.1: Suburban Macro, Speed 70 km/h, 16-QAM fixed modulation

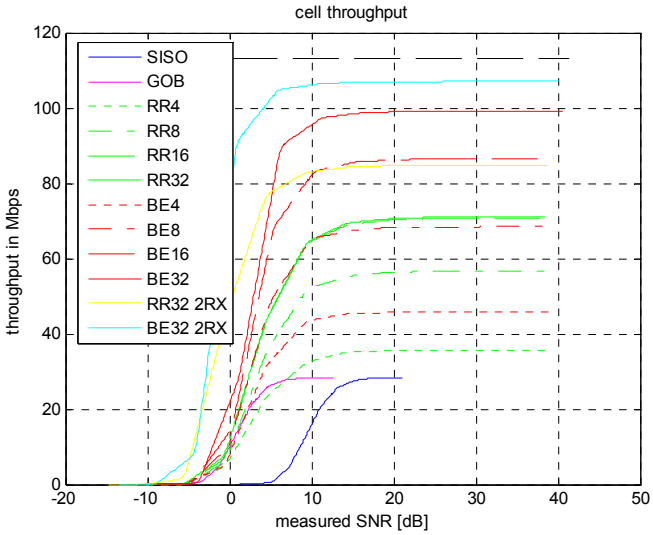


Figure 7.2: Urban Macro 8°, Speed 70 km/h, 16-QAM fixed modulation

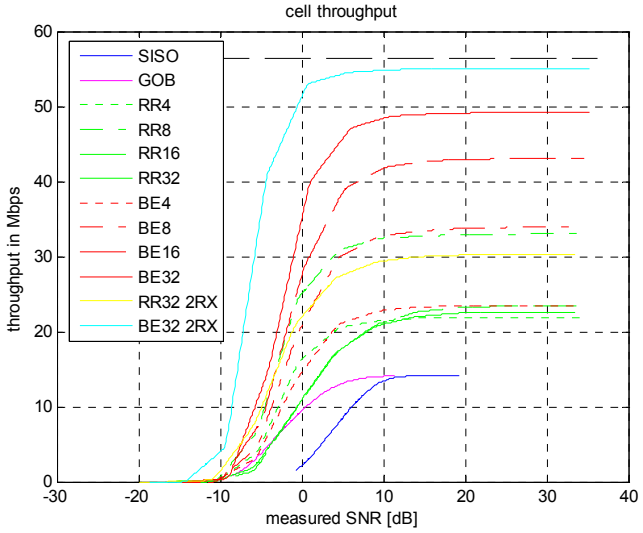


Figure 7.3: Urban Micro, speed 50km/h, 4-QAM fixed modulation

Table 7.2: Normalised throughput of SDMA GoB vs. non-SDMA GoB, suburban macro

Number of users	Rx antennas	Normalised Throughput at SNR 15dB		
		GOB	SDMA+GOB Round Robin Per Beam	SDMA+GOB Adaptive Scheduling
4	1	1	1.60	1.79
8	1	1	2.30	2.75
16	1	1	3.11	3.48
32	1	1	3.34	3.85
32	2	1	3.70	4.06

Table 7.3: Normalised throughput of SDMA GoB vs. non-SDMA GoB, urban macro

Number of users	Rx antennas	Normalised Throughput at SNR 15dB		
		GOB	SDMA+GOB Round Robin Per Beam	SDMA+GOB Adaptive Scheduling
4	1	1	1.24	1.60
8	1	1	1.95	2.38
16	1	1	2.44	3.01
32	1	1	2.45	3.47
32	2	1	2.97	3.77

Table 7.4: Normalised throughput of SDMA GoB vs. non-SDMA GoB, urban micro

Number of users	Rx antennas	Normalised Throughput at SNR 15dB		
		GOB	SDMA+GOB Round Robin Per Beam	SDMA+GOB Adaptive Scheduling
4	1	1	1.5	1.65
8	1	1	2.32	2.37
16	1	1	1.61	3.03
32	1	1	1.56	3.47
32	2	1	2.12	3.90

The system level simulation results are depicted in Figure 7.4. Focusing on a propagation scenario with low angular spread, i.e. suburban macro, and an isolated cell environment, it can be observed that the average cell throughput scales with the number of beams (at least up to the considered maximum of 5 beams). A significant throughput gain is obtained by adaptive resource assignment (PFS) provided that the number of users is sufficiently high. For example, in case of four beams per grid and 32 users, this gain amounts to a factor of ≈1.5 (96.4 Mbps compared to 64.4 Mbps in Figure 7.4). The measured performance differences between adaptive and non-adaptive resource assignment are clearly larger in system level simulations than in multi-link level simulations due to the additional exploitation of "path loss diversity".

The user throughput vs. distance curves depicted on the bottom of Figure 7.4 (for an exemplary case of 32 users) indicate the increasing unfairness when turning from non-adaptive to adaptive resource assignment. However, for more than one beam adaptive resource assignment outperforms non-adaptive assignment for any distance. It is an important aspect of a future system design to balance cell throughput vs. user needs not only with respect to throughput but also with respect to delay requirements.

The simulation results also reveal the trade-off between multiplexing gain and array gain since doubling the number of beams does not lead to a cell throughput gain by the same factor. Adaptive resource assignment is able to counteract the array gain loss to a certain extent albeit at the expense of fairness. A necessary requirement is that the signal-to-intracell-interference (SIR) level caused by adjacent beams is small compared to the required SNR level for the "highest" PHY mode not excluding its application a priori. Moreover, the service area for the highest PHY mode must remain large enough to contain a sufficient number of potential users to be served.

The SIR level depends not only on the number of beams but also on the applied beam pattern which can be controlled by antenna amplitude tapering (subject to the constraints given by the number of antennas). The impact of different tapering variants, i.e., $\cos^n$ with n=0,1,2 and Chebychev is depicted in Figure 7.5 considering 2 and 4 beams per grid. For the two beam case, the different beam patterns under investigation are illustrated in Figure 7.6.

Note that in case of $\cos^n$ tapering the beam distance within each grid is set so that no interference is caused in the main direction of the other beams. This limits the degree of freedom with respect to the number of beams per grid in a 120^o sector. Since Chebychev tapering is characterised by an equiripple sidelobe bounded by "any" parameterisable value, grids of beams can be designed more flexibly in that case.

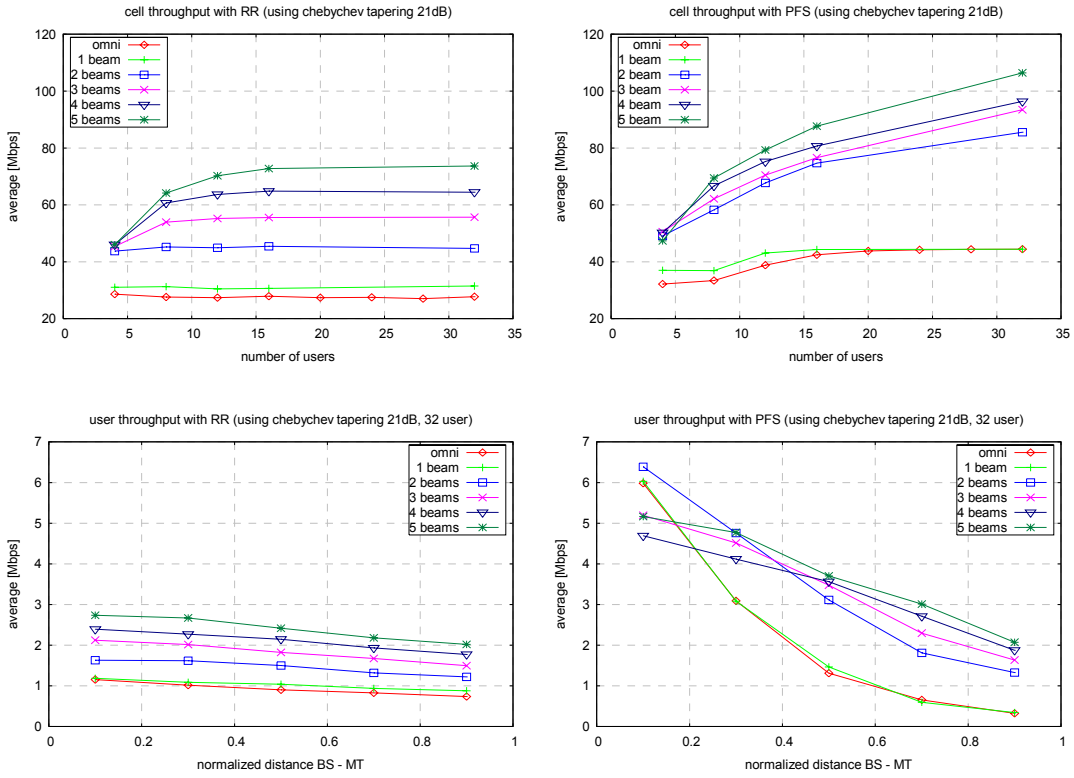


Figure 7.4: average cell throughput (top graphs) and user throughput (bottom graphs) for different number of beams and resource assignment strategies (left: round robin, right: proportional fair)

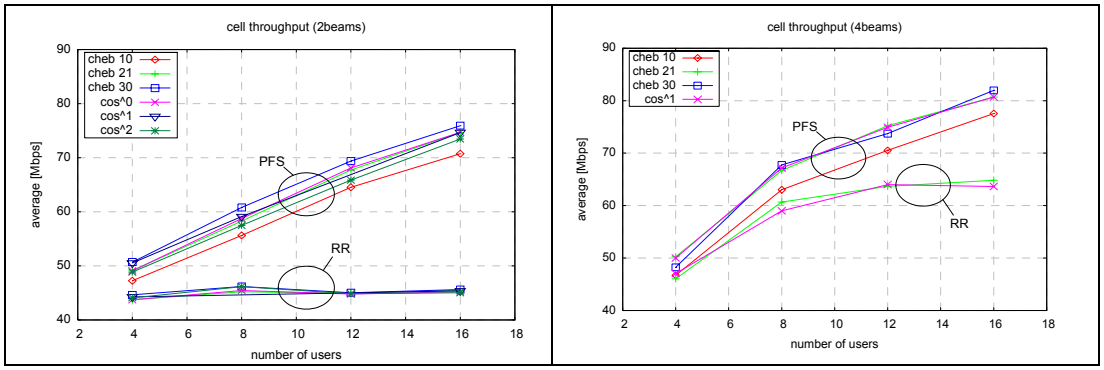


Figure 7.5: average cell throughput with different beam amplitude tapering in case of 2 beams (left graph) and 4 beams (right graph)

In fact, the “optimum” tapering is not only a function of the number of beams per grid but depends also - probably even more - on the intercell-interference scenario. Thus, the results presented here can only be regarded as a snapshot without the potential of being generalised. From Figure 7.5 follows that tapering has only a measurable impact on cell throughput if adaptive resource assignment is applied. In case of non-adaptive processing the average SINR across all users is dominated by bad transmission conditions (i.e., being located at the edge of the beam or far away from the BS) and not by the tapering itself. But the type of tapering has an impact on the SINR statistics, which can be exploited in conjunction with adaptive scheduling. In the specific scenario under investigation, the differences between the tapering variants are relatively small. If only 2 beams are formed, Chebychev tapering with high sidelobe suppression performs best. The variant “cheb 10”, i.e. Chebychev tapering with 10dB sidelobe suppression, causes too much interference and should be avoided in any case.

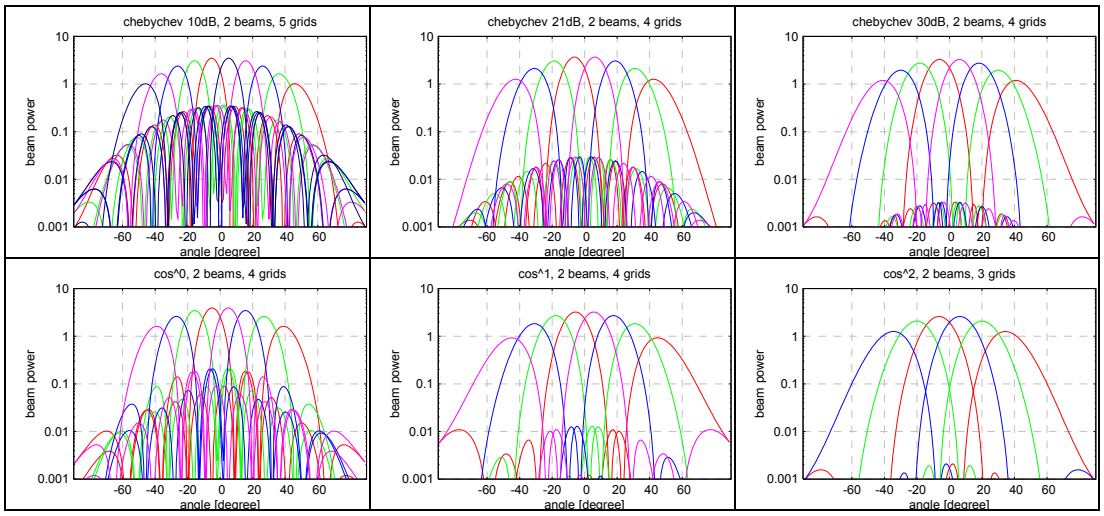


Figure 7.6: Applied GoBs in case of cos^n and Chebychev tapering

7.2.1.4 Conclusions

The results show that SDMA with a grid of beams is an interesting candidate for the wide-area scenario for increasing the cell throughput. The prerequisite is of course a certain number of available users for simultaneous transmission, which are spatially separated. Adaptive scheduling (based on CQI) can exploit extra multi-user diversity at the cost of increased feedback rate. Even with the coarse granularity used, adaptive scheduling pays off, especially for increase of angular spread.

7.2.2 Spatial Link Adaptation

7.2.2.1 Adaptation between Matrix Modulations

Adaptation between different matrix modulations is described in Chapter 3 mostly for the case when adaptation is performed on each data subcarrier of an OFDM system. However, adaptation on each subcarrier might not be possible due to the limits in feedback signalling. The objective here is to study adaptation between different spatial transmission formats when adaptation is performed only once during a transmission of a frame and hence averaging of the second relative condition number (See Section 3.1) has to be applied.

OFDM parameters used in the simulations correspond to those described in Chapter 5. The applied linear matrix modulations are diagonal ABBA (rate 1), double ABBA (rate 2) and vector modulation (rate 4). In case of diagonal and double ABBA, coding dimension is frequency, i.e., coding is extended over 4 neighbouring subcarriers. The channel code concatenated with inner matrix modulations is a rate ½ convolutional code. The interleaver is random and the frame length is 6656 bits, with 3320 information bits per frame. In the receiver, ML soft output demodulation is performed. The results are based on 2D optimisation. For each value of received SNR (1 dB step) and the averaged second condition number (8 quantisation bins), the mode that provides the highest packet throughput is chosen. Channel models used are WINNER SCM urban micro and macro with the antenna configuration AW3-TW2. It is assumed that the channel stays constant during transmission of one frame.

Figure 7.7 and Figure 7.8 show the throughput results for the studied spatial transmission methods and the adaptive scheme. It can be seen that adaptation between different transmission modes performs well and hence the two chosen CQIs can be considered to be a reasonable option for link-adaptation in the WINNER system (naturally the received SNR is the most viable CQI also in the MIMO case). When comparing the results between the two channel models, it can be noticed that the performance of diagonal ABBA does not depend on the channel model used whereas for double ABBA and vector modulation the performance degrades when moving from urban micro environment to the urban macro environment. This is naturally due to the decreased angular spread of the channel.

Finally, it should be noted that the considered CQIs can also be used to form a link-to-system interface in the case of advanced multi-antenna methods such as the matrix modulations considered here. This possibility should be considered especially in the cases where usage of a "SINR per stream-based" link-to-system interface (see Chapter 4) becomes computationally too demanding.

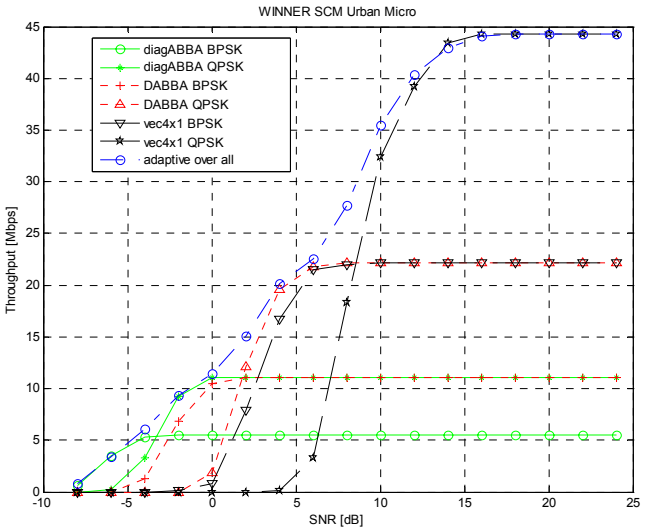


Figure 7.7: Throughput for WINNER SCM Urban Micro

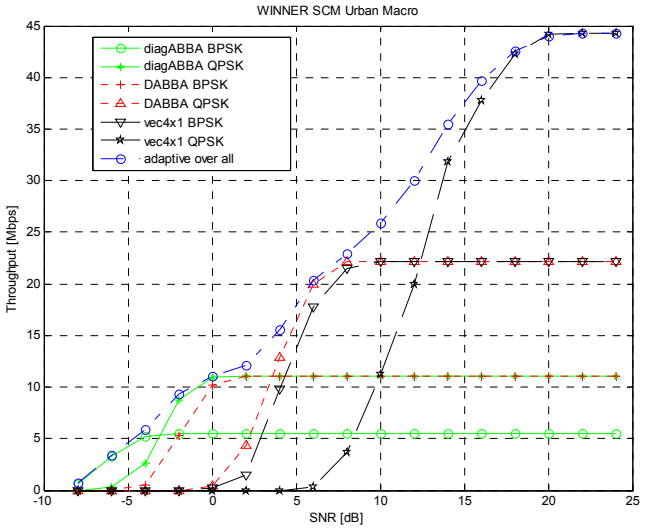


Figure 7.8: Throughput for WINNER SCM Urban Macro

7.2.2.2 Linear Precoding to Account for Antenna Correlation

Linear Precoding is proposed in order to exploit the knowledge of antenna correlation at the transmitter. This long-term CSI can be obtained based on feedback signalling or channel reciprocity.

Using Linear Precoding allows re-configurability of space-time coding techniques to different cases of channel correlation. The use of Linear Precoding with ABBA matrix modulation is studied in this section and simulations are conducted for different channel and modulation scenarios.

Simulation assumptions are aligned to those described in Chapter 5. Channel generation is performed using the WINNER SCM channel model. Scenarios investigated are the Urban Micro and Macro. The channel is either 4 x 1 (antenna array configuration AW0-TW0) or 4 x 4 (antenna array configuration AW6-TW2). Constellations are either BPSK or QPSK and the code rate is ½. Performance is depicted in terms of throughput versus SNR. We compare the three following cases 'Linear Precoding over ABBA', 'no precoding over ABBA' and 'beamforming' (transmitting over the strongest eigenmode of the transmit covariance matrix).

For ABBA the coding dimension is frequency, i.e., the ABBA block is transmitted over 4 neighbouring subcarriers. At the receiver the ABBA demodulation is done via ML hard decision.

Figure 7.9-Figure 7.12 show performances for different environments and antenna settings. In Figure 7.9 the environment is Urban Macro and the channel has a very high correlation (ρ= 0.97 for the closest couple of antennas). In such conditions the beamforming technique is almost optimal; however, the Linear Precoding still provides some gain of 1.2 dB in SNR for 2 Mbps throughput in the BPSK case. The use of ABBA with no precoding (no channel knowledge) gives very poor performances w. r. t. the two previous techniques that exploit the knowledge of antenna correlation at the transmitter.

In Figure 7.10 the environment is Urban Micro and the correlation is now lower (ρ=0.75 for the closest couple of antennas). This improves the performance of the Linear Precoding over beamforming. Especially for high throughput (low error probability and better quality of service), the gain is 3dB in SNR for 10.4 Mbps throughput in the QPSK case. The gain w.r.t. the case of ABBA with no precoding is comparable for high throughput and more important for low throughput.

For Figure 7.11 and Figure 7.12 the system is 4x4 and the antenna spacing has a greater impact. Environments are respectively Urban Macro and Urban Micro. Overall the increase of the number of transmit antennas provides an array gain and allows to improve the performance of all techniques.

Furthermore, the increase of the antenna spacing (lower correlation: resp. ρ=0.64 and ρ= 0.42 for the closest couple of antennas) is more favourable for Linear Precoding (and beamforming), especially in the Urban Micro environment where the correlation is the lowest. The gain of Linear Precoding over no linear precoding is significant, especially for low to middle values of the throughput.

We conclude that Linear Precoding over STBC results in significant performance enhancements in the presence of channel correlation. When compared with beamforming based on the same CSI, Linear Precoding allows substantial gains, in particular for low correlation (increased antenna spacing and/or rich environment like Urban Micro) and for high throughput. The latter corresponds also to lower error probability and better quality of service.

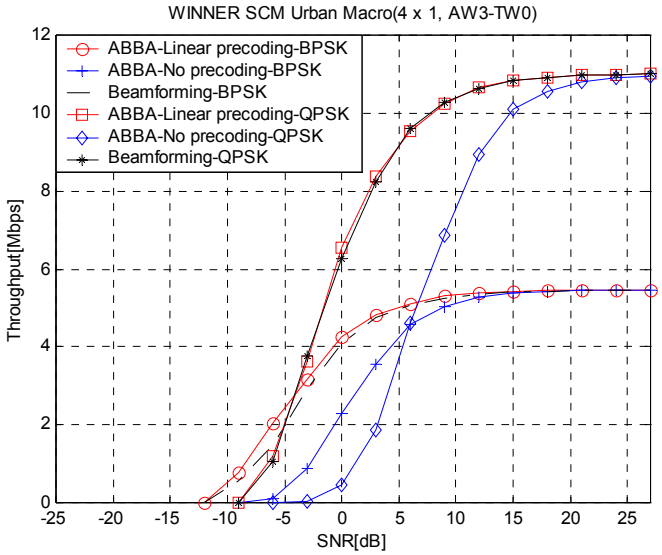


Figure 7.9: Throughput for 4x1 system and Urban Macro environment

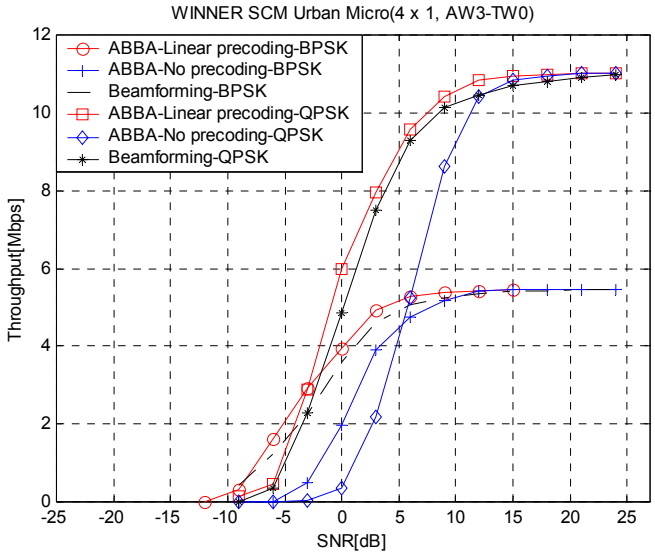


Figure 7.10: Throughput for 4x1 system and Urban Micro environment

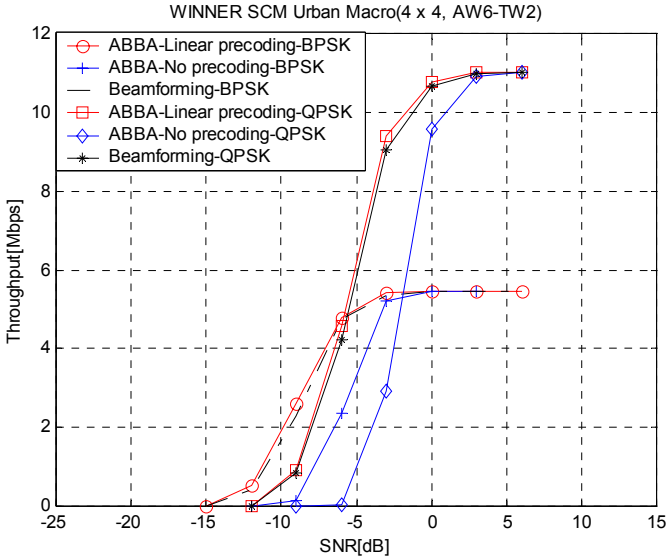


Figure 7.11: Throughput for 4x4 system and Urban Macro environment

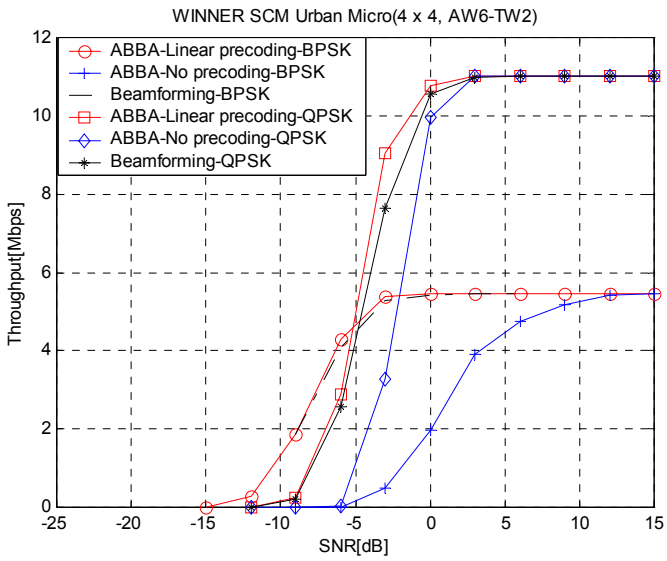


Figure 7.12: Throughput for 4x4 system and Urban Micro environment

7.2.2.3 Conclusions

From the simulation results for spatial link adaptation and linear precoding it can be observed that the benefit obtained from the knowledge of the transmit covariance is significant, especially in the case of high antenna correlation. This gain is independent of the number of receive antennas. Furthermore, the knowledge of the transmit covariance matrix is available in most cases, since for both of the TDD and FDD modes the reciprocity of the transmit covariance holds. Therefore, linear precoding should naturally be used to exploit this knowledge.

Moreover, for a number of receive antennas greater than one, higher multiplexing gain can be achieved using link adaptation. Simulations show that it is possible to find two CQIs per coded frame and base the link adaptation on these indicators.

It is also possible to combine both approaches by using linear precoding as an additional feature of the spatial link adaptation scheme. For such a combination, most of the gain is expected to occur in the low to middle SNR region.

7.2.3 A Preliminary Comparison of Some Techniques

In the present section, a preliminary comparison of two different techniques, beamforming based on a set of fixed beams, and per antenna rate control (PARC) is performed. Both techniques are similar in the sense that a number of parallel separately channel encoded spatial streams are transmitted. With beamforming, an SDMA component may be added so that the multiple streams are transmitted in different beams to different users. For PARC, on the other hand, instead of transmitting streams to different users, all streams are transmitted to the same user, and the streams are in the present comparison transmitted in the element space directly from the antennas.

Three different wide-area deployments are considered,

- A single omni directional site, i.e., a single cell
- A single three sector site, i.e., three cells
- Nineteen three sector sites, i.e., 57 cells

An extended version of the 3GPP SCM urban macro channel model with 15 degrees angle spread was used and the cell range was 500m. Most of the assumptions are consistent with the assumptions made in previous chapters, and for convenience and clarity, they are repeated in addition to some additional assumptions made.

In the case of a single site, users are uniformly distributed in space on a disc with radius equal to the cell range. In the multi-site deployment, users are uniformly distributed in space over the coverage area of the

system and the site-to-site distance equals to 3/2 times the cell range.

The total transmit power per sector cell was 20W, independent of the number of transmit antennas. At the terminals a noise-power spectral density of -167 dBm/Hz, corresponding to a noise figure of 7dB, was assumed. Omni directional antenna elements with 10 dBi gain were assumed in the single cell case, whereas sector antennas with 70 degrees half power beam width and gain 14dBi were assumed in the case with multiple cells. Cell selection was based on the path loss, and the path loss calculation includes the impact of angle spread at the AP, shadow fading and the element patterns. In the case with beamforming, cell selection was done based on the beam with the lowest path loss of each cell. Further, the set of beams was generated with an FFT in addition to 21 dB Chebychev tapering. This means that for a four element ULA (AW3), four beams are created, and for an eight element ULA (AW7), eight beams are used. For the studies of PARC, a combination of space and polarisation diversity is considered; two cross-polarised elements separated four wavelengths (AW6).

As mentioned in the previous chapter, all users were assumed to have full buffers, and time-domain round robin scheduling was used. The scheduling strategy was slightly modified for the case with beamforming with an SDMA component. In this case, the set of beams was divided into a number of subset of beams, referred to as grids. With four antennas and four beams, two grids with two non-adjacent beams were created for possible transmission of two parallel streams. With eight antennas and eight beams, both two-stream transmission and four-stream transmission were considered. For two-stream transmission, four grids with two beams each were used, and for four-stream transmission, two grids with four beams each were created. When two or four grids are used, each grid is active half or a quarter of the time respectively. For each beam in the active grid, round robin scheduling of the users connected to it was then done. Since only cases with up to 32 users were studied, there were not always users connected to all active beams, and the available power is then equally split between active beams actually serving a user. It should be noted that this kind of resource assignment is probably far from optimum, and that the basic idea is to get an initial assessment of the potential for SDMA.

Coding and interleaving was done over super-chunks, covering 26 adjacent chunks. The 25% overhead for parameter estimation was for simplicity kept the same for all techniques and was modelled as randomly interleaved with the data. An error-free channel-quality measurement is assumed to be available with no delay so that the MCS can be chosen to maximise throughput under the constraint that the estimated BLEP is below 10%. For multi-stream transmission with PARC to a single user, it is assumed that all streams are retransmitted if there is an error in one or more of the stream. In this case, the BLEP is constrained to be below 10% divided by the number of transmitted streams. Finally, the transmit power allocated to a user and a certain stream is uniformly distributed over all the 832 carriers, and any additional control signalling overhead on top of the 25% overhead for parameter estimation is neglected.

Since the focus of the study is on the spatial processing, complexity was reduced and only a quarter of all carriers were actually simulated. To estimate the performance of an OFDM/TDMA system, the results are scaled with a factor of four. Simulations, not presented here, indicate that this may be a fair extrapolation method for non channel-dependent scheduling. A potential drawback is that the benefit of the possible frequency diversity is not visible in the results.

Block errors were generated according to the previously defined link-to-system interface, and the SINR calculations capture the post-receiver processing SINR considering the channels and transmit weights used by all active transmit antennas in the system. An ideal MMSE receiver with weights determined on a per carrier basis, taking the spatial colour of the intra-stream and inter-cell interference into account, was assumed. For the case with beamforming and an SDMA component, this means that terminals with multiple antennas also suppress interference from other beams transmitting to other users. For the case with PARC transmission, an ideal SIC after channel decoding was used, and if one stream was erroneous, it was assumed that all streams were in error. Note that imperfections, such as estimation errors and time dispersion greater than the cyclic prefix were neglected. The same holds for channel variations in time during a chunk.

For the simulations with a single site, 50 snapshots were simulated whereas only 10 snapshots were considered for the multi-cell deployment. Each snapshot has a duration of 12x24x6 OFDM symbols, which corresponds to roughly 100ms. The results presented here include the average aggregated throughput per cell as well as percentiles of the user throughput averaged in time.

7.2.3.1 Downlink Beamforming Using a Set of Fixed Beams

First, beamforming using a grid of fixed beams is considered. Users connected to a cell are then scheduled in a round-robin fashion, and the best beam is used for downlink transmission to the served user. At the terminal side, 50% of the users have single antenna (TW0), 30% have two antennas (TW1) and the remaining 20% have four antennas (TW2). For comparison, the case with single antenna transmission is included, both for the case that all terminals have a single antenna as well as for the case that there is a mix of terminals with different number of antennas.

The results are plotted in Figure 7.13, both for a single three-sector site as well as for the multi-site deployment. As expected, there is a significant gain from the use of multiple terminal antennas, about 50% in cell throughput in the multi-cell deployment relative to the case with only single antenna terminals. Downlink beamforming further improves performance, and it may be noted that the performance improvement in terms of cell throughput from doubling the number of transmit antennas is significantly larger in the multi-cell case as compared to the single cell case. This may be partly due to the limitation of available MCSs in addition to the reduction of intercell interference obtained by transmitting the energy over a narrower angular sector.

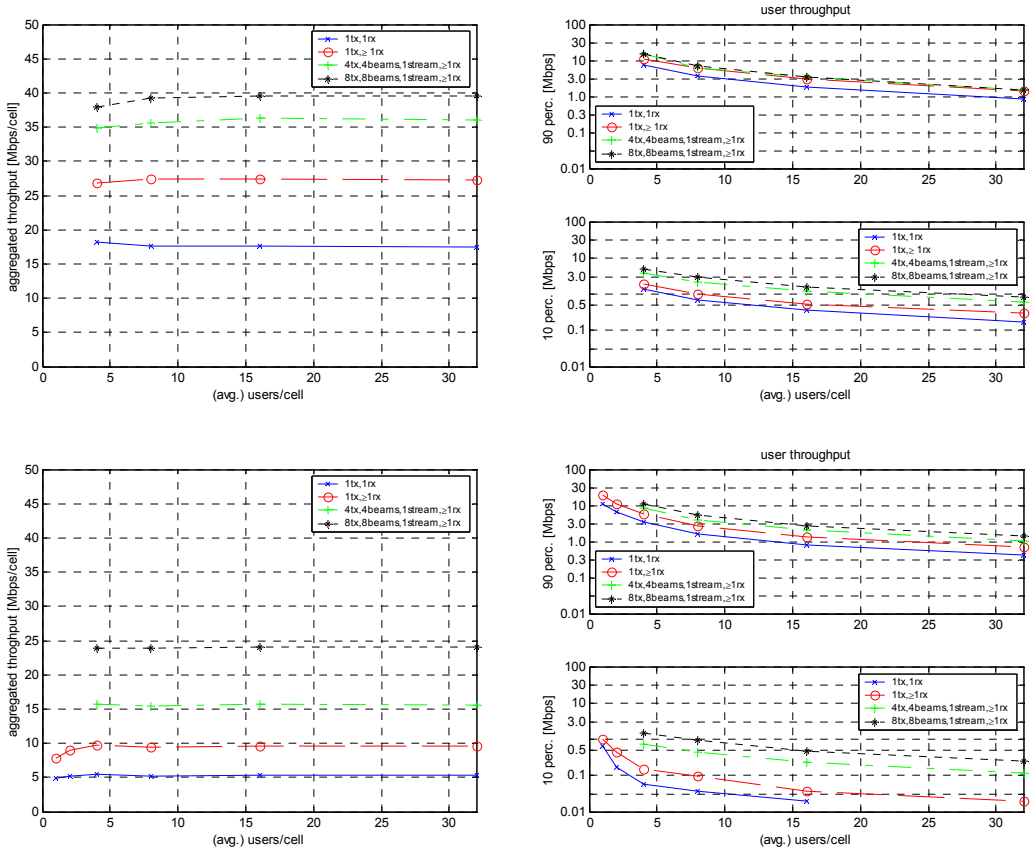


Figure 7.13: Downlink beamforming using a set of fixed beams. Estimates of cell throughput (left) and 10 and 90 percentiles of user throughput (right) for OFDM/TDMA and a single three-sector site (top) and a multi-site deployment (bottom)

7.2.3.2 Beamforming Using a Set of Fixed Beams with an SDMA Component

Next, an additional SDMA component was considered together with beamforming based on a set of fixed beams. Recall that the resource assignment is slightly different in this case, the beams are partitioned into grids, each grid is active a fraction of the time, and round-robin scheduling is performed among the users connected to each beam in the active grid.

The results are shown in Figure 7.14. In the legend of the figure, the maximum number of streams refers

to the number of beams in each grid. Performance of an SDMA component in terms of cell throughput improves with increasing number of users. This since the probability of several spatially separated users increases. In both the single site and the multi-site case, serving up to four users in parallel may increase performance in terms of cell throughput for a large number of users with around 160-180% as compared to single antenna transmission. It may also be seen that the relative gain in transmitting two streams is slightly larger in the multi-site case as compared to the single site case, indicating as expected that there is a trade-off between re-use within the same site and re-use between different sites which is not seen from single cell or single site simulations. It should be noted here that the terminals with multiple antennas also may suppress inter-beam interference and that a very simple scheduling strategy is employed.

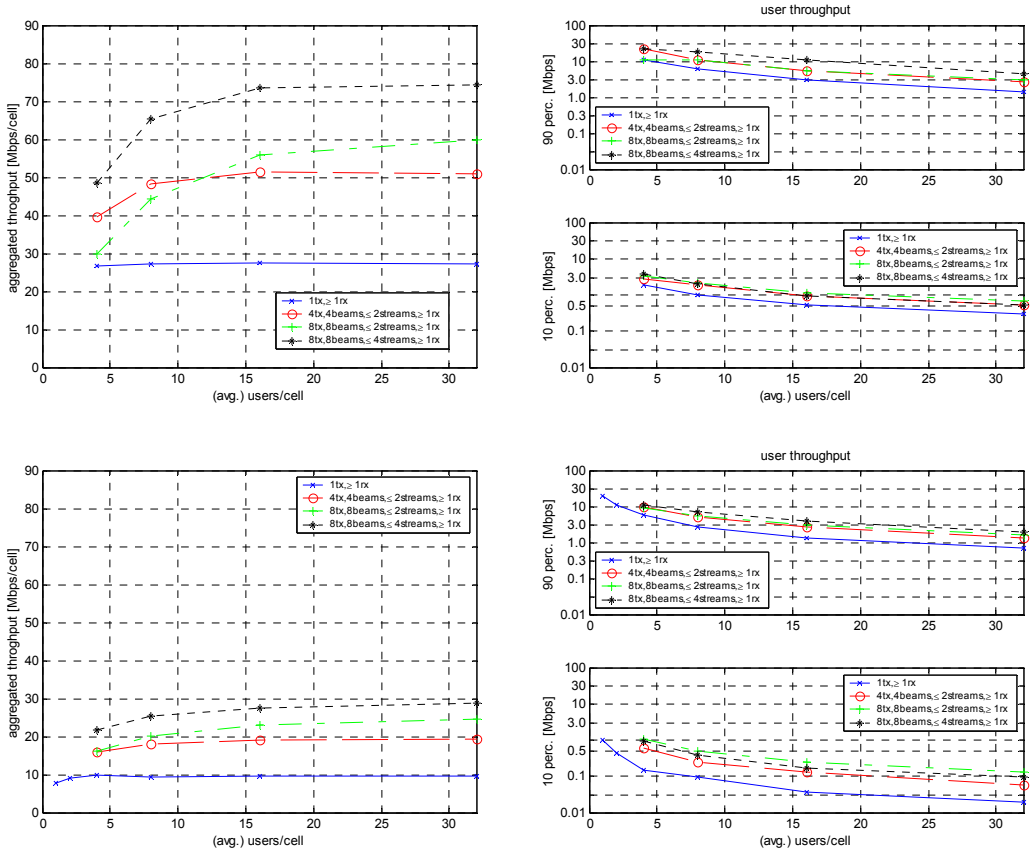


Figure 7.14: Downlink beamforming using a set of fixed beams with an SDMA component. Estimates of cell throughput (left) and 10 and 90 percentiles of user throughput (right) for OFDM/TDMA/SDMA and a single three-sector site (top) and a multi-site deployment (bottom)

7.2.3.3 Per Antenna Rate Control (PARC)

For the first set of results with PARC, convolutional coding was considered for the case that all terminals have four antennas separated half a wavelength (TW2). As mentioned above, the terminal receivers employed ideal MMSE combining and SIC of multiplexed streams after channel decoding. The number of streams transmitted from the base stations was varied in a crude, static way, by only using a subset of the antennas. The results do therefore not illustrate any benefits of antenna (subset) selection based on short-term feedback. If one stream is transmitted, then a single slanted antenna element is used. If two streams are transmitted, then a dual polarised antenna is used, and when four streams are transmitted, two dual polarised elements separated four wavelengths are used, i.e., the configuration referred to as AW6 in the chapter on assumptions. It should be kept in mind that a physical realisation of an omni-directional dual polarised antenna is not obvious, and that this case is mainly included to study a scenario with no inter-cell interference. The available transmit power is split equally between the number of antennas so that a total transmit power of 20 W per cell is used independent on the number of streams transmitted (given that a user is served in the cell). Multiple terminal antennas can be used for a combination of

spatial multiplexing and intercell interference suppression, and by varying the number of antennas and streams used in each cell by the AP, it is then possible to roughly assess possible tradeoffs. The results are shown in Figure 7.15: As can be seen from the single omni-site case, impressive performance gains in terms of user peak rates and cell throughput may be obtained in the absence of intercell interference with spatial multiplexing as compared to single stream transmission, which is limited by the highest rate of the considered MCSs.

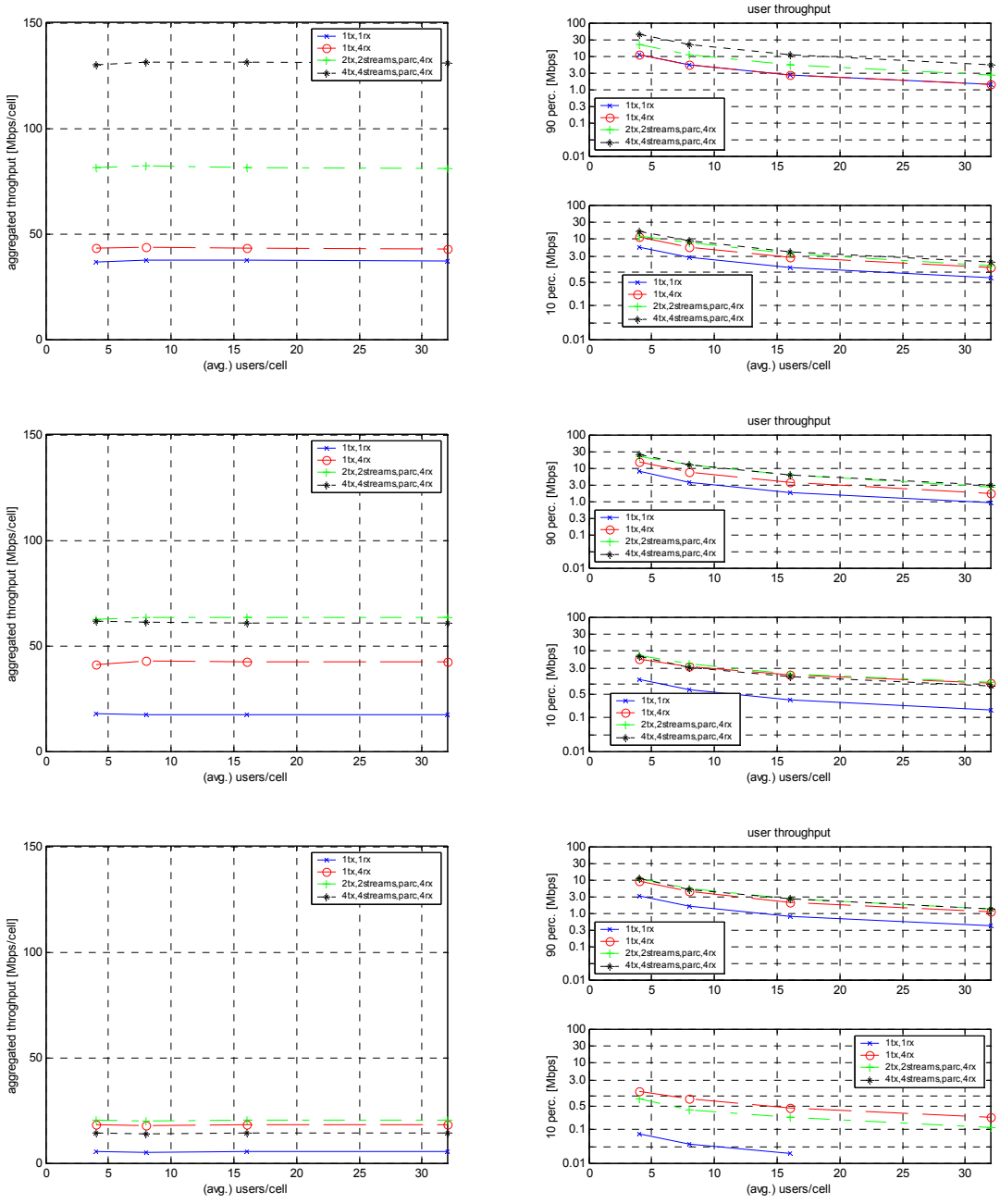


Figure 7.15: Multiple terminal antenna and spatial multiplexing with PARC. Estimates of cell throughput (left) and 10 and 90 percentiles of user throughput (right) for OFDM/TDMA and a single omni-site (top), a single three-sector site (middle) and a multi-site deployment (bottom)

Another observation seen by comparing the results for the three-sector site with the results for the omni-site is that the inter-sector interference has a considerable impact on the relative performance of different transmission strategies for the system under consideration. In the multi-site deployment it may further be seen that the gain of transmitting two streams instead of one is only about 15% and that this comes at the

cost of a relative performance degradation for the users with the lowest throughputs (see the 10 percentile). In a multi-cell deployment a limited amount of spatial multiplexing may make sense, depending on the load, but this may introduce unfairness in the sense that a larger spread in user throughput may be seen. Multi-stream transmission, as expected, may thus come at the cost of reduced spatial interference suppression possibilities for the terminal.

In the considered case, all terminals had four receive antennas. In a possible future deployment scenario, a mix of terminals with different number of receive antennas is expected, and the reference scenario with 50% single antenna terminals (TW0), 30% two antenna terminals and 20% three antenna terminals (TW2) was considered. Performance of spatial multiplexing with PARC and SIC after channel decoding was investigated and as expected found poor. An explanation is that the lowest code rate per stream for the MCSs considered is too high in the absence of HARQ type two with chase combining or incremental redundancy. Two solutions to this problem include adapting the number of transmitted streams, and then possibly also benefit from the gains from antenna subset selection, and to extend the set of MCSs by low rate channel coding or spreading. In here, the set of MCSs was extended, and used a rate 1/3 turbo code together with rate matching and BPSK, QPSK, 16-QAM and 64-QAM modulation to generate nine MCSs with 60 to 4203 data bits per superchunk.

The results are shown in Figure 7.16, and indicate that PARC with more than one stream may improve certain performance measures also without adapting the number of streams in scenarios with little interference and with 50% single antenna terminals. In the multi-cell case, the situation is a bit more complex, due to the trade-off between receiver intercell interference rejection and the use of multiple transmit antennas. Multi-cell results are not presented in here, as the objective of the example is to illustrate the need for low rate MCSs when PARC is used and that PARC may have certain flexibility when it comes to supporting terminals with different number of receive antennas.

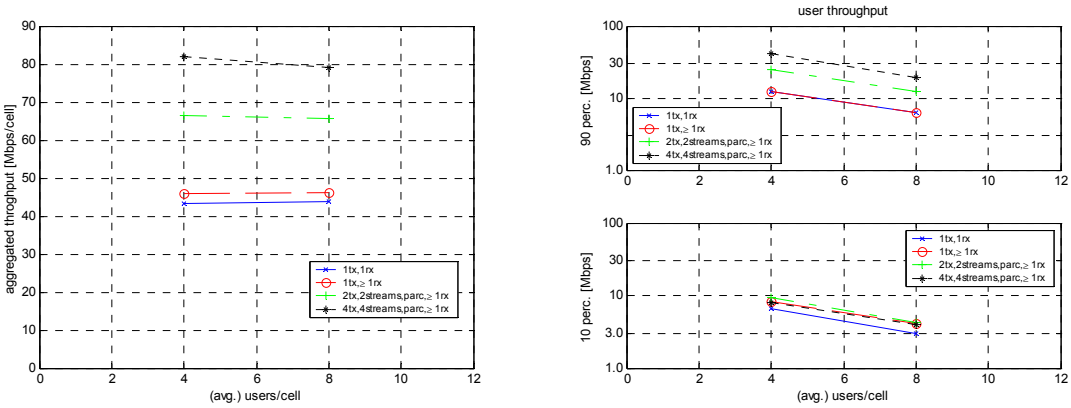


Figure 7.16: Spatial multiplexing with PARC, a mix of terminals and extended set of MCSs. Estimates of cell throughput (left) and 10 and 90 percentiles of user throughput (right) for OFDM/TDMA and a single omnisite

7.2.3.4 A Comparison in Terms of Site Spectral Efficiency and some Remarks

A comparison is next presented. The spectral efficiency in terms of bps/Hz/site is estimated for the different basic techniques considered. More specifically, the spectral efficiency is taken as the cell throughput multiplied with the number of sectors per site at the load, in terms of average number of users per cell, for which the 10 percentile of the extrapolated average user throughput, averaged over 100ms, is at least 0.5 Mbps. Only a finite number of loads were simulated and simple linear interpolation and extrapolation of the results in Figure 7.13, Figure 7.14 and Figure 7.15: were done to be able to estimate the site spectral efficiency.

The results are presented in Table 7.5 assuming a total bandwidth of 20 MHz. It should be noted that the results do not include any control signalling overhead in addition to the 25% overhead for parameter estimation. Further, the short snapshot-like simulations and considered definition of spectral efficiency as

well as the scenario means that absolute numbers should be used with care. However, some relevant relative comparisons can hopefully still be made to guide future work.

Table 7.5: Preliminary estimates of spectral efficiencies in bps/Hz/site with unknown confidence intervals. '-' means that a low enough number of users were not simulated, and '≈' means that the estimated average user throughput was higher than 0.5 Mbps, and that the performance of 32 users per cell (on average) was taken.

	Transmitter	Receiver	Deployment		
			Omni site	Three sector site	Multi-site
Single antenna	1 antenna	1 antenna	1.9	2.6	0.7
	1 antenna	1 (50%), 2 (30%), and 4 (20%) antennas, MMSE combining		4.1	1.3
	1 antenna	4 antennas, MMSE combining	≈2.1	≈6.3	2.7
Downlink BF with fixed beams	4 antennas, downlink BF with 4 fixed beams, one user served	1 (50%), 2 (30%), and 4 (20%) antennas, MMSE combining		≈5.4	2.3
	8 antennas, downlink BF with 8 fixed beams, one user served	1 (50%), 2 (30%), and 4 (20%) antennas, MMSE combining		≈5.9	3.6
Downlink BF with fixed beams and SDMA component	4 antennas, downlink BF with 4 fixed beams, up to 2 users served	1 (50%), 2 (30%), and 4 (20%) antennas, MMSE combining		≈7.6	2.5
	8 antennas, downlink BF with 8 fixed beams, up to 2 users served	1 (50%), 2 (30%), and 4 (20%) antennas, MMSE combining		≈9.0	3.6
	8 antennas, downlink with 8 fixed beams, up to 4 users served	1 (50%), 2 (30%), and 4 (20%) antennas, MMSE combining		≈11.1	3.7
PARC	Dual polarised antenna, 2 streams with PARC	4 antennas, MMSE combining and SIC	4.0	≈9.5	3.1
	Two dual polarised antennas, 4 streams with PARC	4 antennas, MMSE combining and SIC	6.5	≈9.1	-

From the simulations results in this section, Section 7.2.3, the following observations can be made

- The spectral efficiency and performance in terms of cell throughput observed in the single cell and single site cases may be very different as compared to the corresponding measures in the multi-site case. As expected, interference modelling is of great importance.
- Multiple terminal antennas, in here used not only for diversity, but also for interference suppression may as expected offer significant gains, and a relevant benchmark for MIMO schemes is therefore believed to be receiver diversity with a possible component of interference rejection.
- By comparing the results obtained downlink beamforming with and without an SDMA component as well as the results for single stream transmission with multi-stream transmission with PARC it may be seen that the large performance gains observed in single cell or site become very modest in the uniformly loaded multi-site case. This indicates again that modelling of intercell interference in a multi-site deployment is important, or example when considering the benefits of SDMA.
- Both SDMA and PARC foster spatial multiplexing gains, and as is well known, the benefit from this depends on the operating point in terms of SINR. However, the results here indicate that beamforming with no SDMA component and terminal receive diversity with interference suppression is a strong candidate for the heavily loaded uniform wide-area scenario considered. On the other hand, if all terminals have four receive antennas; PARC with half the number of

antennas at the AP has similar spectral efficiency with possibly lower deployment cost and antenna design requirements.

- The single cell and single site results may perhaps be used to extrapolate the performance for a sparsely loaded scenario or a hot spot. For such a case beamforming with an SDMA component and spatial multiplexing with, e.g., PARC may be used to achieve significant spectral efficiencies. However, as was seen, SDMA requires that there are enough users to serve, and therefore spatial multiplexing with PARC may then be a strong candidate to serve a small number of heavy users. An adaptive combination, such as PARC armed with a beamforming/directivity component, referred to as per stream rate control previously, is then an interesting candidate.
- Spatial intercell interference management is of great interest and importance. This was illustrated in the study of PARC with a varying number of streams in combination with terminal interference suppression.
- It should be kept in mind that the considered techniques have different complexity, for example in terms of the number of antennas at the AP and the terminals.

Note that that the results are obtained for a certain scenario with a lot of specific assumptions, e.g., in terms of traffic modelling, deployment, physical layer parameters such as MCSs and very simple scheduling algorithms. During the course of the project, we expect that assumptions, models and criteria will be refined.

7.2.4 Control Overhead for Wide-Area Scenario Techniques

The overhead signalling considerations, as explained in Section 6.5, will be based on the separate evaluation of the required additional control information to be conveyed in forward and return link control channels, without taking into account a specific structure for the control channel. For each control information (e.g., spatial mode selector, number of spatial streams used, loading information, etc.) an rough estimate of the number of required bits in the forward and return control channel will be provided. Furthermore the required update rate regarding the frequency, time, space, and user dimension for the control signalling information is determined.

7.2.4.1 Link Adaptation Overhead

In principle, link adaptation requires to signal the applied modulation and coding format in the forward control channel to enable decoding in the receiver. This applies to all spatial processing techniques. The overhead for the individual techniques however differs, depending whether a single modulation and coding format is applied to all spatial streams, whether adaptive modulation and coding is used per stream, or whether power and bit loading is applied. Table 7.6 shows the basic overhead estimations for these three cases when using a dedicated control channel (estimations for common control channels, or deviating assumptions for particular techniques are provided in the sections of individual techniques). In cases, where the same modulation and coding format is used for all spatial streams $M_t = 1$, otherwise it is typically below four. However, it should be noted that it is an overall multiplicative factor. Also the update rates in frequency is an overall factor and in general much larger than M_t for a fully adaptive system in frequency. Therefore it can be concluded that adaptivity in frequency can be significantly more costly than in space domain and therefore must provide accordingly more gain, when a final comparison and decision is made, which dimension should be used for adaptivity and which should bring in averaging and diversity components. It is understood that the update rates in frequency and time are linked to the coherence bandwidth and coherence time. Since both have major impact on the overhead it is recommended that the WINNER system concept provides means to adapt the granularity of link adaptation in frequency and time domain in a user-specific fashion. As a distinguishing criterion for different spatial processing techniques we note, that adaptive modulation and coding per stream multiplies the net overhead, unless adaptivity in other domains is reduced. For techniques based on bit and power loading (i.e., adaptation of modulation and power), N_{FECo} is replaced by N_p. However, at this stage, we expect approximately the same amount of information bits required for both, so that these two types of link adaptation will not differ in terms of forward control channel overhead. To quantify this we assume a temporal update rate of n_{ch_t} =150, i.e., link adaptation every 50 ms. In this case the forward control channel overhead is $M_t \cdot n_f \cdot$ 158.2 kbps.

Table 7.6: Estimation of Forward Control Channel Overhead for adaptive modulation and coding

Control information		Modulation and coding format
Estimated information bits per message		$(N_M + N_{FECo})$
Required update rate	**in time**	$(n_{ch\_t} \cdot 337.2\ \mu s)^{-1}$
	in frequency	n_f
	in space	M_t
	in user domain	1
total control information rate per user		$n_f \cdot (n_{ch\_t} \cdot 337.2\ \mu s)^{-1} \cdot (N_M + N_{FECo}) \cdot M_t$ kbps

7.2.4.2 SDMA with Grid of Beams

The overhead estimation for the return control channel in case of SDMA with GoB is provided in Table 7.7. The given figures are based on the assumption that - for any chunk in the frequency domain - a terminal feeds back a single CQI value for the "best" beam only (SDMA/FDMA is included in the overhead calculation). Thus, the amount of feedback information at a certain time instant comprises the sum of beam selection information plus corresponding CQI multiplied by the number of chunks per channel. The update rate depends to a large extent on the basic transmission mode, i.e., short-term adaptive, long-term adaptive or non-adaptive. It is believed that in the wide-area scenario at least long-term adaptivity can be supported. Note that depending on the terminal velocity, different adaptation modes and even different update rates may be used for different terminals. In Table 7.7 the update rate is parameterised by the update period given by the product of chunk duration n_{sub}·56.2 µs and number of chunks between updates n_{ch_t}.

Using the exemplary parameters from Table 6.1 as well as n_{ch_t} = 8, n_{sub} =3, n_{va} = 4 and N_c/n_{frame}=2 (as assumed in Chapter 5) yields an overhead of approximately 10.4 kbit/s/user to support short-term adaptivity. For the update rate of n_{ch_t}= 150, which is used for comparison, it results in 0.55 kbps. Since the average user throughput is in the order of several Mbit/s the amount of overhead seems to be acceptable. In case of long-term adaptivity the overhead becomes even negligible.

Optionally a more general kind of feedback instead of single CQI for the best beam would be one CQI per beam. This generates more overhead but also gives additional scheduling information. Thus the short- or long-term CQI information yields $N_{CQI} \cdot n_{va}$ instead of N_{CQI} but the beam ID can be dropped. Deviating from Table 7.7 the total information rate per user then becomes $17.8 \cdot N_c \cdot N_{CQI} \cdot n_{va} / (n_{frame} \cdot n_{sub} \cdot n_{ch\_t})$ kbit/s. For the short-term example with n_{sub} =3 the resulting return control rate becomes 29.7 kbps.

Table 7.7: Estimation of return control channel overhead for SDMA+GoB and single CQI for best beam

Control information		Beam ID	Short-term or long-term CQI
Estimated information bits per message		$\mathrm{ld}(n_{va})$	N_{CQI}
Required update rate	**in time**	$(n_{ch\_t} \cdot n_{sub} \cdot 56.2\ \mu s)^{-1}$	
	in frequency	N_c/n_{frame}	
	in space	1	
	in user domain	1	
total control information rate per user		$17.8 \cdot N_c\ [\mathrm{ld}(n_{va}) + N_{CQI}] / (n_{frame} \cdot n_{sub} \cdot n_{ch\_t})$ kbit/s	

The overhead of the forward control channel is generally larger especially if link adaptation *and* resource assignment is done adaptively on a short-term basis. Assuming a common control channel per beam, the amount of overhead is summarised in Table 7.8. Note that no additional TDMA component per scheduling interval is taken into account here. In case of a pure SDMA/FDMA scheme per scheduling interval the number of served users N_{Us} equals $n_{va}N_c/n_{frame}$ and the required control information rate per user is

simply 17.8 $[\mathrm{ld}(N_M)+\mathrm{ld}(N_{FECo})+\mathrm{ld}(N_{MT})]/(n_{sub}\cdot n_{ch\_t})$ kbit/s. Using again the figures from Table 6.1 and Table 6.2 yields around 11.1 kbit/s/user with $\mathrm{ld}(N_{MT})=7$, and for $n_{ch\_t}= 150$ we obtain 0.59 kbps.

Table 7.8: Estimation of forward control channel overhead for SDMA+GoB

Control information		Modulation and Coding information	terminal ID
Estimated information bits per message		$N_M + N_{FECo}$	$\mathrm{ld}(N_{MT})$
Required update rate	**in time**	$(n_{ch\_t}\cdot n_{sub}\cdot 56.2\ \mu s)\text{-}1$	
	in frequency	N_c/n_{frame}	
	in space	n_{va}	
	in user domain	$1/N_{Us}$	
total control information rate per user		$17.8\cdot N_c\, n_{va}\, [N_M+N_{FECo}+\mathrm{ld}(N_{MT})]/(N_{Us}\cdot n_{frame}\cdot n_{sub}\cdot n_{ch\_t})$ kbit/s	

7.2.4.3 Spatial Domain Link Adaptation

Assuming an update rate of n_{ch_t} chunks in the time direction, no adaptation in frequency domain, and $2^{N_{CQI}}$ different modulation and coding combinations, the estimated forward and return channel overheads are given in Table 7.9. For example, with $n_{ch\_t} = 4$ and $N_{CQI} = 35$, total control information rate per user is 3.7 kbps. Using $n_{ch\_t}= 150$ for comparison, the update rate is 0.10 kbps. For the forward control overhead N_{CQI} is replaced by $(N_M + N_{FECo})$.

Table 7.9: Preliminary estimation of Return and Forward Control Channel Overhead for Spatial Domain Link Adaptation

Control information		Modulation and Coding information
Estimated information bits per message		N_{CQI}
Required update rate	**in time**	$(n_{ch\_t}\cdot 337.2\ \mu s)^{-1}$
	in frequency	1
	in space	1
	in user domain	1
total control information rate per user		$(n_{ch\_t}\cdot 337.2\ \mu s)^{-1}\cdot N_{CQI}$ bps

7.2.4.4 Linear Precoding

Linear Precoding assumes the knowledge of the transmit covariance at the BS. The transmit covariance is a M_T x M_T hermitian matrix (M_T: number of antennas at the BS), that has $M_T{}^2$ real coefficients, which corresponds to $M_T{}^2/2$ complex coefficients.

The transmit covariance depends only on slowly changing large-scale parameters (AoA, AoD and PDP). These parameters remain pretty constant over distances of the order of several meters. For the update rate, we assume an update distance of 1 m and maximum user velocity in an urban environment: 70km/h (worst case scenario). This corresponds to an update time period of 50ms, i.e., n_{ch_t} =150 chunk durations. For this temporal update rate and using $M_T = 4$, the overhead according to Table 7.10 results in 3.2 kbps.

It is worth emphasising that 70km/h is the worst-case scenario in an urban environment, for low velocity (static user) the control information rate per user is far less than reported in the table. Furthermore, the CSI can also be provided by channel reciprocity. However, we still need lower rate control information in order to correct the RF impairments.

Simulations were conducted for fixed modulation. However, the system will be adaptive over 4 choices of combined constellation-code rate. 2 bits are needed to carry this information in the forward link.

Furthermore, spatial link adaptation can be used using matrix modulation (ABBA, Double-ABBA, and BLAST). This requires further 2 bits to carry matrix modulation information. The total needed bits are then 4. Additionally, N_{FECo} bits are required for signalling the parameters of the channel code. For the

parameters discussed above this yields 0.51 kbps. Note that compared to Table 7.6 ($n_f = M_T = 1$) this is an increase of around 25% to $M_t \cdot n_f \cdot$ 197.7 kbps.

Table 7.10: Preliminary estimation of Return Control Channel Overhead for Linear Precoding

Control information		**CSI**
Estimated information bits per message		$M_T^2/2 \cdot N_{CSI}$
Required update rate	**in time**	$(n_{ch\_t} \cdot 337.2\ \mu s)^{-1}$
	in frequency	1
	in space	1
	in user domain	1
total control information rate per user		$1/2 \cdot M_T^2 \cdot N_{CSI}\ (n_{ch\_t} \cdot 337.2\ \mu s)^{-1}$ kbps

Table 7.11: Preliminary estimation of Return Control Channel Overhead for Linear Precoding

Control information		**Modulation information**
Estimated information bits per message		$(4+N_{FECo})$
Required update rate	**in time**	$(n_{ch\_t} \cdot 337.2\ \mu s)^{-1}$
	in frequency	1
	in space	1
	in user domain	1
total control information rate per user		$(4+N_{FECo})\ (n_{ch\_t} \cdot 337.2\ \mu s)^{-1}$ kbps

7.2.4.5 Limited CSI Spatial Multiplexing

Limited CSI Spatial Multiplexing (for brevity, referred to as 'PARC' here forth) requires only a CQI measurement for each transmit antenna stream to be signalled on the return channel. This is shown in the table below. The calculation corresponds to a single user being scheduled on all subcarriers in all OFDM symbols of each scheduling time-frequency resource unit, of which there are assumed to be n_f in the frequency direction, and each consists of n_{ch_t} chunks in the time direction.

The simplest (but least flexible) implementation would involve scheduling the same user for all available subcarriers (i.e., $n_f = 1$), which, assuming 4 transmit antennas, 2 bits of CQI and an update rate of once per chunk duration ($n_{ch\_t} = 1$), leads to 23.7 kbps overhead (and to 0.16 kbps for $n_{ch\_t} = 150$). Lower overhead can be achieved if a reduced update rate (larger n_{ch_t}) is acceptable.

The approach with the maximum scope for channel dependent scheduling (but consequently high overhead) would have the flexibility to schedule different users on each chunk in the frequency direction, i.e., $n_f = 104$ (for 832 subcarriers and a chunk size of $n_{sub} = 8$). This leads to an (unacceptable!) overhead of 2.6Mbps for $n_{ch\_t} = 1$ (and to 17.3 kbps for $n_{ch\_t} = 150$), which shows clearly the necessary trade between increasing throughput via smaller scheduling units and the required consequent increase in signalling capacity.

Table 7.12: Preliminary estimation of Return Control Channel Overhead for PARC

Control information		**CQI**
Estimated information bits per message		N_{CQI}
Required update rate	**in time**	$(n_{ch\_t} \cdot 337.2\ \mu s)^{-1}$
	in frequency	n_f
	in space	M_t
	in user domain	1
control information rate per user		$n_f \cdot (n_{ch\_t} \cdot 337.2\ \mu s)^{-1} \cdot N_{CQI} \cdot M_t$ kbps
total control information rate per user		$n_f \cdot (n_{ch\_t} \cdot 337.2\ \mu s)^{-1} \cdot N_{CQI} \cdot M_t$ kbps

Here M_t is the number of transmit antennas, n_{ch_t} is the number of chunks in the time direction for which users are scheduled, and n_f is the number of scheduling resource units employed in the frequency direction.

In short, the return control channel overhead is expected to increase linearly with the number of transmitted antenna streams as compared to transmission with a single antenna. It should be kept in mind that the CQI measurements are not expected to be independent, and one could therefore consider joint encoding of the CQIs of different antennas in effort to trade signalling requirements for lower flexibility. Also, it remains to be understood how to quantise the measurements. Should a fine quantisation be done in the frequency domain, or should one use a much more coarse granularity in the frequency domain when multiple antennas are used and so on.

The forward control channel must provide signalling of the modulation and coding format, and may employ more bits for this than the return link CQI (e.g., 8 bits instead of 5). In all other respects though, the overhead is identical to the return link and the calculation follows Table 7.6 for $n_f > 1$ and $M_T > 1$.

There is additional forward control channel overhead (not quantified here) relating to:

1. Resource map – signalling the IDs of which time, frequency, and antenna stream resources are allocated to which user
2. Pilots – pilot transmissions are required from each transmit antenna (preferably orthogonally) in order to permit the receiver to form channel estimates

Again, to allow for maximum flexibility, it is expected that the forward control overhead increases linearly with the number of transmit antennas as compared to the case with single antenna transmission.

It is also expected that by joint encoding of control information for different transmit antennas the overhead can be significantly reduced at the cost of flexibility and scalability.

7.2.5 Conclusions

In this section simulation results for the wide-area scenario have been presented. It was shown that beamforming has high potential for scenarios with small angular spread, since it provides a good performance versus complexity trade-off. For a given number of users, cell throughput scales nearly linearly with the number of beams up to a certain maximum, which is given by the achievable beam width in combination with the required sidelobe suppression. When using a fixed grid of beams, it is important to carefully design the grid and the beam pattern (e.g., sidelobe suppression), since performance is sensitive to the inter-beam interference. Inter-beam interference must be reduced to a level that allows use of the highest PHY mode even with an additional margin for intercell interference. Means to reduce this sensitivity to inter-beam interference are strong channel coding and link adaptation. When beamforming is combined with adaptive resource assignment based on short-term CQI, considerable performance enhancements can be obtained due to the exploitation of multi-user diversity, especially for high load and increasing angular spread in the propagation channel.

For the space-time-frequency coding and modulation, linear dispersion codes (matrix modulation) are considered as baseline for further investigations. Regarding their applicability to different channel conditions diagonal ABBA is more robust than DABBA or vector modulation. Linear precoding to account for antenna correlation adds significant gain to pure matrix modulation (STBC) at a very low overhead signalling cost, especially in the case of highly correlated channels, as for example in the urban macro scenario. The linear precoding can be applied to any matrix modulation mode and will be considered as an additional feature to spatial link adaptation when long-term CSI is available at the transmitter. Spatial mode selection with adaptivity to long-term CSI can provide substantial gains within each scenario, and considering the fact that the WINNER multi-antenna concept must adaptively support many different scenarios it will be an essential feature of the proposed design.

The forward control channel overhead will in most cases be dominated by the information required due to link adaptation. In particular adaptivity in frequency and time domain is considered costly. Even if adaptivity per spatial stream is used, the corresponding multiplicative factor is much lower than for the frequency domain. These considerations are important to perform a performance versus overhead trade-off evaluation in order to decide when to select adaptivity as opposed to introducing averaging and diversity components. It is therefore recommended that the WINNER system concept provides means to adapt

the granularity of link adaptation in frequency and time domain in a user-specific fashion to the coherence bandwidth and coherence time.

Comparing the return link overhead for identical temporal update rates, low overhead is encountered for spatial domain link adaptation schemes based simply on the selection of different codes. The initial estimations show that SDMA based on a grid of beams has approximately 5 times higher return link overhead. For identical update rates, linear precoding has 30 times higher overhead than space-time code selection. However, for a fair comparison, the required update rates for long-term spatial covariance matrices must be put in relation to the update rates for CQI required in different scenarios. For the PARC technique, the return link overhead can be comparable to the simple selection of codes up to an overhead of 10 times that of the linear precoding technique, depending on the degree of adaptivity. These considerations clearly indicate, that return link overhead due to spatial processing spans a large range and needs future investigations and optimisation.

The multi-cell system evaluations indicate that downlink beamforming based on a set of fixed beams and terminals with multiple receive antennas and interference suppression capabilities are suitable for the highly loaded wide-area scenario considered. For low-load scenario, single site results indicate that PARC may be used to achieve high cell throughput and user peak rates. Another important observation is that single-cell and single-site evaluations may overestimate the benefits of SDMA and spatial multiplexing.

One important focus of future research will be the assessment of the optimum amount of adaptivity in terms of the associated 'penalty' of overhead signalling. For techniques exploiting long-term CSI, as for example linear precoding for matrix modulations, the resulting gains are large at very low expense in terms of overhead signalling. At the example of limited CSI spatial multiplexing (PARC, PSRC), it was shown that the control overhead grows linearly with the number of transmitted streams, or transmit antennas, and for the case that the transmission is adaptive also in the time and frequency domains, this may result in an unacceptably high overhead. After a certain break-even point further adaptivity will merely increase complexity and overhead without providing corresponding performance gain.

7.3 Performance Results for the Short-Range Scenario

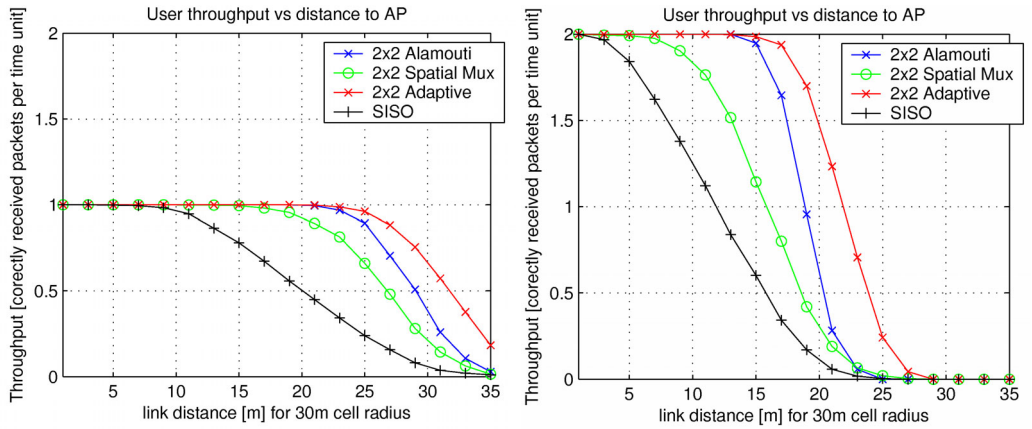
7.3.1 Spatial Link Adaptation

In this section, the performance of the spatial link adaptation technique for the short-range usage scenario is assessed. We present simulation results for a MIMO-OFDM system with two antennas at the transmitter and receiver. The system supports switching between two spatial signal-processing methods, namely space-time coding according to the Alamouti (AM) scheme and spatial multiplexing (SMUX). To decide for the AM or SMUX scheme as the best-suited method for a given channel realisation**,** the receiver computes for each subcarrier of the OFDM signal a metric for the two candidate spatial signal-processing techniques and feed-backs the corresponding optimal mode as link quality parameter to the transmitter. The simulations have been performed for a signalling rate of 1 and 2. In case of signalling rate 1, the symbols to be Alamouti encoded are QPSK modulated while BPSK modulated symbols are used for spatial multiplexing; in case of signalling rate 2, 16-QAM or 4-QAM symbols are fed to the two spatial processing units, respectively. Further details on the assumed parameters for the MIMO-OFDM system and the selection criteria for link adaptation are given in Section 5.2. In the simulations, we assume a typical large office environment that can be modelled by the IEEE 802.11n channel model D with enabled/disabled line-of-sight (LOS) component.

Figure 7.17 and Figure 7.18 show the throughput performance of a point-to-point link in a short-range communication system using the described MIMO-OFDM transmission techniques for the IEEE 802.11n channel model D, where the (LOS) component of the channel response is disabled in the latter figure. The throughput is measured as function of the distance between the mobile station (MS) and the access point (AP). The AP nominally covers a cell area with a radius of up to 30 meters, and MSs at the cell border receive signals from the AP with a SNR of 20 dB.

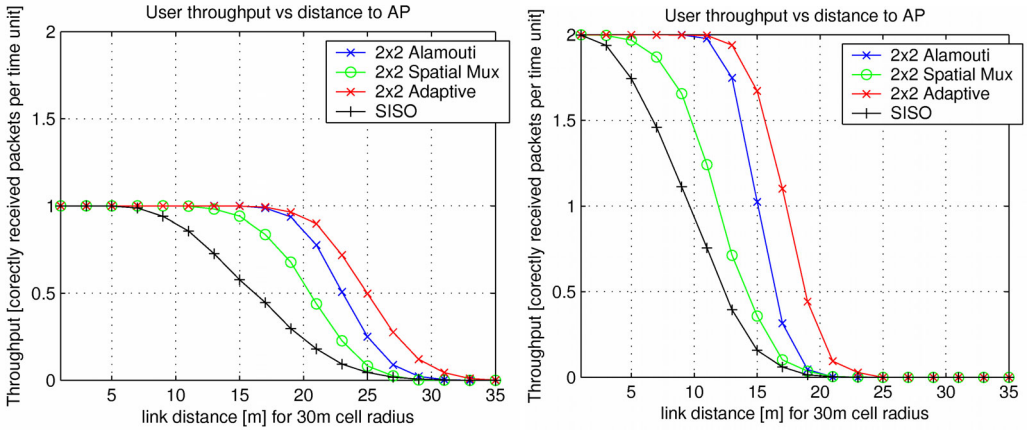
The transmit packet queue of the mobile station is always fully loaded with packets of fixed size and the throughput is measured by counting the number of correctly received packets per time unit. The graphs show the throughput performance of the MIMO-OFDM point-to-point link with two antennas at the transmitter and receiver for three different spatial processing techniques, namely the 2x2 Alamouti

scheme, 2x2 spatial multiplexing, and adaptively switching between the two schemes. For comparison purposes, we also plotted the performance of a corresponding transmission system, where the MS and AP are equipped with a single transmit and receive antenna (SISO). The simulations have been performed for a system with signalling rate 1 and 2.



Signalling Rate = 1 Signalling Rate = 2

Figure 7.17: Throughput performance of a point-to-point link with and without spatial-link adaptation for IEEE channel model D with LOS

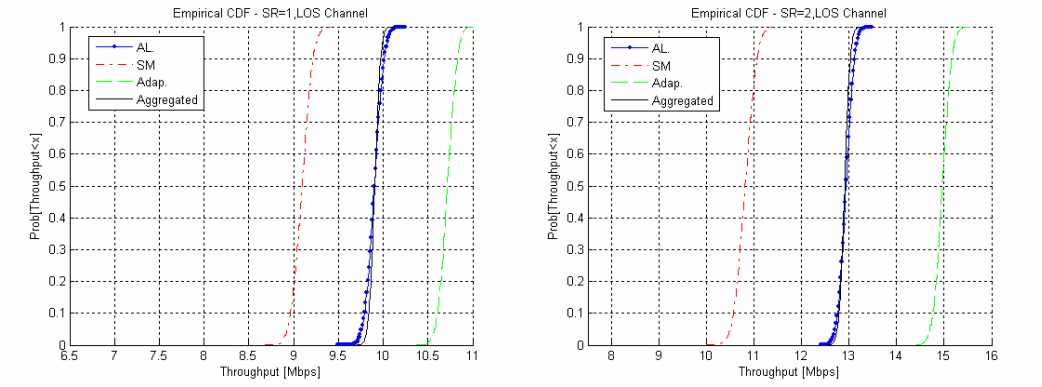


Signalling Rate = 1 Signalling Rate = 2

Figure 7.18: Throughput performance of a point-to-point link with and without spatial-link adaptation for IEEE channel model D without LOS

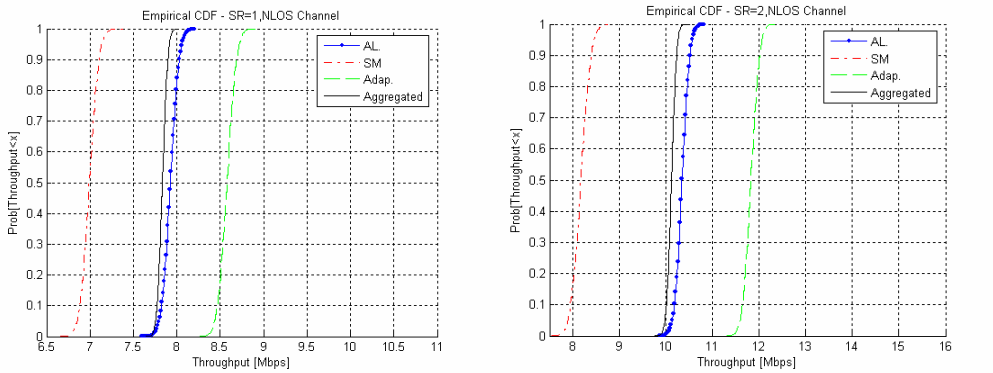
The simulation results indicate substantial performance gains in terms of coverage and throughput in case of using one of the three MIMO-OFDM transmission techniques. For the two investigated radio channels, the Alamouti scheme outperforms spatial multiplexing, and these performance gains can be further increased by applying spatial link adaptation. At low distances, the transmission schemes with signalling rate 2 provide the highest throughput. However, since higher data rates require a larger SNR for low-error transmissions and the mean SNR decreases with distance, a high throughput cannot be achieved close to the cell border. To get cell coverage at the expense of a slightly decreased throughput, MIMO-OFDM transmission techniques with signalling rate 1 - possibly in conjunction with spatial link adaptation - should be employed instead. The simulation results obtained for the channel model with and without LOS components indicate that the presence of a line-of-sight component leads to a higher throughput for all distances and investigated transmission techniques.

Figure 7.19 and Figure 7.20 show the throughput performance of a mobile user. As spatial-processing techniques, we have applied the 2x2 Alamouti scheme, 2x2 spatial multiplexing, and the spatial link adaptation technique for switching between the two multi-antenna methods. Moreover, the cell throughput is shown that has been obtained by aggregating the throughput over the users in the network. The radio channel is modelled by the IEEE 802.11n channel model D with and without LOS. The cell radius r is equal to 30 meters, and it is assumed that a mobile located at the cell border receives signals from the AP with a SNR of 10 dB. In our simulations, the mobile is placed at randomly selected locations at a distance of d<r to the AP. Since the result depends on the realisation of the channel, we have plotted the cumulative distribution function (CDF) of the throughput in the two figures.



Signalling Rate = 1 Signalling Rate = 2

Figure 7.19: User and aggregated throughput performance with 2x2 Alamouti (AL), 2x2 spatial multiplexing, and with spatial-link Adaptation for IEEE channel model D with LOS



Signalling Rate = 1 Signalling Rate = 2

Figure 7.20: User and aggregated throughput performance with 2x2 Alamouti (AL), 2x2 spatial multiplexing, and with spatial-link adaptation for IEEE channel model D without LOS

The simulation results indicate that the 2x2 Alamouti scheme provides significantly better throughput results than 2x2 spatial multiplexing. For a signalling rate of 1, the gains are in the order of 10%, while at signalling rate of 2 gains of 20% have been measured. Applying spatial-link adaptation further enhances the throughput performance by 10 to 15%. The presence of a LOS component in the radio channel leads to higher throughput values as expected.

7.3.2 Successive Minimum Mean-Square-Error Precoding (SMMSE)

In this section, we present the results of system-level simulations of the system employing SMMSE, previously described in Chapter 3. We investigate the performance of this system in a typical office environment using the IEEE802.11n D N-LOS channel model. We assume that perfect CSI is available at the

transmitter. The transmitter and the receiver are equipped with 4 antennas each. The goal was to investigate the spatial multiplexing gain that this technique can provide to a single user as a special case since this technique was originally designed for the multi-user case. In Figure 7.21 the user throughput vs. distance is shown. The cumulative distribution function (CDF) of the user throughput in a cell with the radius of 30 m is shown in Figure 7.22.

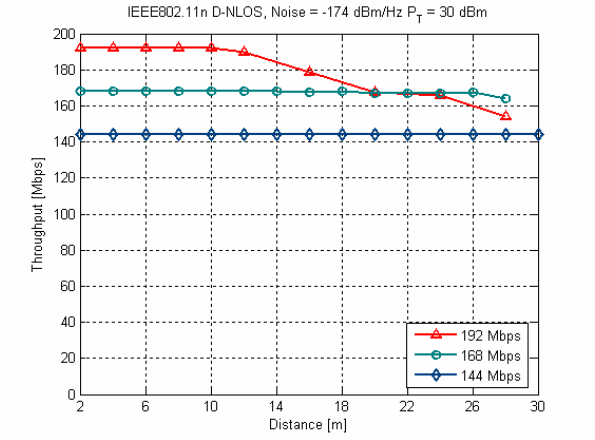


Figure 7.21: User throughput as a function of distance. D – NLOS channel model

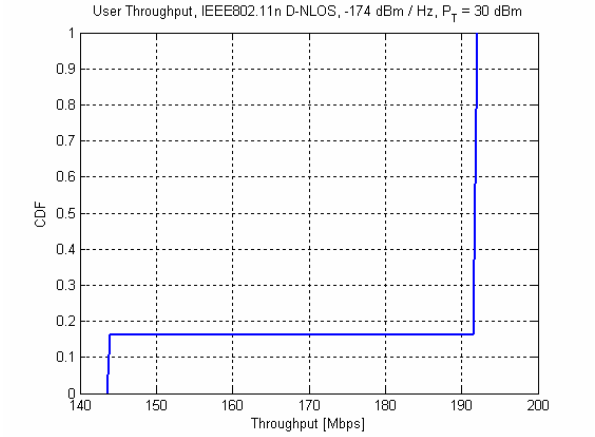


Figure 7.22: CDF of user throughput

There are three curves represented in Figure 7.21. The curve for 192 Mbps is obtained by multiplexing 4 spatial streams modulated using 64-QAM 2/3 and the curve for 168 Mbps by multiplexing 3 spatial streams modulated using 64-QAM 2/3 plus one spatial stream modulated using 16-QAM 1/2. Finally, the third curve for 144 Mbps is obtained by multiplexing 3 spatial streams modulated using 64-QAM 2/3. The data rates and the coverage can be explained with the idealistic assumption of the availability of the perfect CSI at the transmitter. In the single-user case, SMMSE results in transmission on the eigenmodes of the channel with MMSE detection at the receiver.

SMMSE is designed to optimise the performance of the downlink of a multi-user MIMO system. However, when the number of users is equal to one, it can also provide a very high SMUX gain as shown in this section. This depends on the CSI available at the transmitter and the influence of channel estimation errors. The impact of non-perfect CSI on the performance of the system will be investigated in further studies.

7.3.3 Joint THP with Diversity Techniques

In this section, we assess the performance of joint THP with Diversity Techniques using the IEEE802.11n channel model D, in LOS and NLOS configurations. This scheme is associated to rate ½ turbo-coding,

and each slot consists of 2 OFDM symbols.

Figure 7.23 and Figure 7.24 show the average throughput of a multiple-link in a short-range communication system using the described joint THP +Diversity Techniques (DT) for the IEEE 802.11n channel model D with line-of-sight (LOS) component. The individual channels of the users are independently distributed from each other. The throughput is measured as a function of the SNR. The base station or AP has 6 antennas and communicates with 3 users each having 2 antennas. The mobile station is always fully loaded with packets of fixed size and the throughput is measured by counting the number of correctly received frames per time unit. In Figure 7.23, the average throughput of 3 users as a function of the SNR is depicted. Furthermore, in Figure 7.24 the average throughput as a function of the number of active users is depicted.

Figure 7.23 shows that the average user throughput decreases as the number of the active users increase. However, the aggregating throughput increases when more users are activated (Figure 7.24)

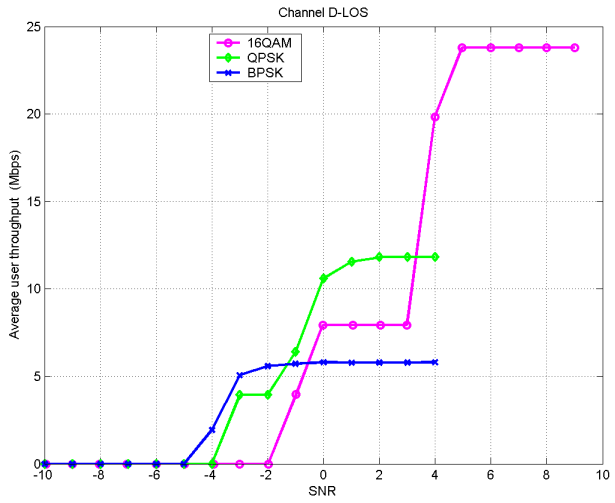


Figure 7.23: Average throughput with multiple links for the IEEE Channel Model D with LOS, where 6 antennas are deployed at the BS or AP and the three terminal users each having 2 antennas

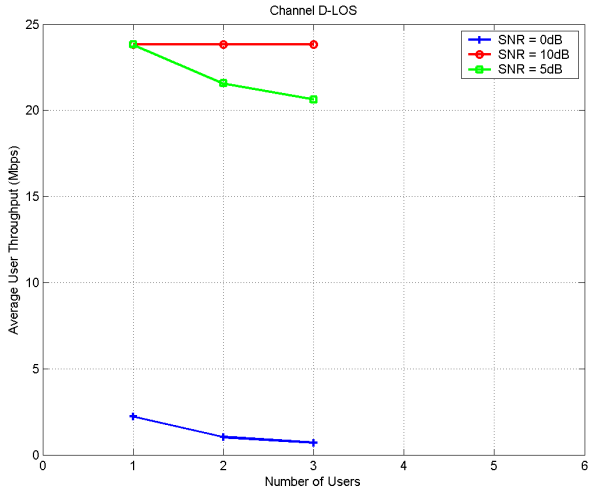


Figure 7.24: Average throughput for the IEEE Channel Model D with LOS, where a maximum of three terminal users each having 2 antennas are activated with multiple links

Figure 7.25 and Figure 7.26 give the simulation results with the IEEE802.11n downlink channel, where the BS has 4 antennas and 2 users having two antennas each. The two users are assumed to be always fully loaded by communicating with the base station. For comparison purposes, the simulation results of THP+Dominant Eigenmode Transmission (DET), THP+Alamouti and SISO, where BS and MS are equipped with single transmit and receive antenna respectively, are plotted in one graph. The average throughput over the channel model D with LOS component and without LOS component is plotted in Figure 7.25 and Figure 7.26, respectively. The simulations have been performed for a system with rate 2 and 4, i.e., 4-QAM and 16-QAM modulations are used.

Figure 7.25 and Figure 7.26 are plotted to compare the two diversity techniques OSTBC, DET with a SISO system. Since the BER performance of THP+DET is better than THP+OSTBC, the user throughput of THP+DET scheme is shown to be superior to the THP+OSTBC. Compared with the SISO case, the throughput gain from THP+DET is tremendous. Notice that with the THP+DT scheme the first precoded user will always get the best effective channel, the last precoded user gets the worst effective channel, and the channels of the other users are between. Until a certain SNR threshold, the first user can get the maximum throughput and other users still have no throughput. When the SNR is greater than a certain threshold, other users can also get some throughput. As the SNR increases, all users can get maximum throughput at end.

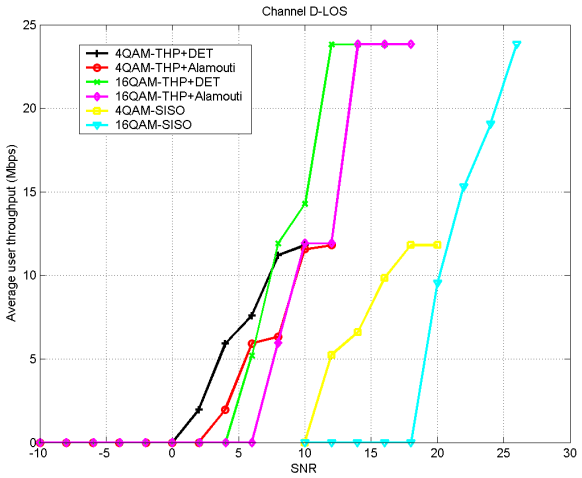


Figure 7.25: Throughput performance of a point to multipoint system, where the BS has 4 antennas and each of the two mobile users has 2 antennas with the IEEE Channel Model D with LOS

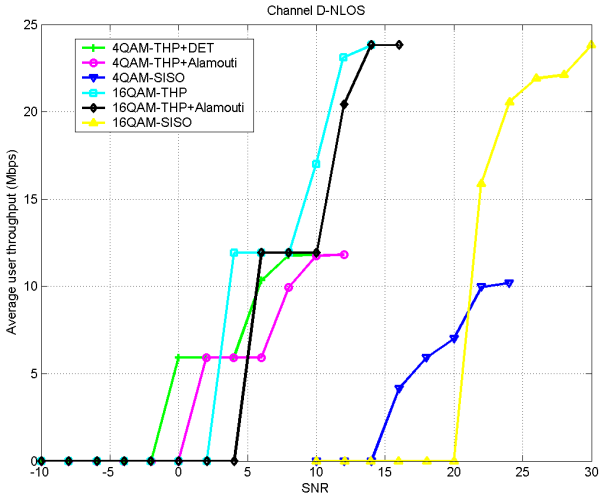


Figure 7.26: Throughput Performance with the IEEE Channel Model D without LOS

7.3.3.1 Robustness

The robustness of the technique considered here is investigated by the impact of the channel estimation errors on the BER performance and throughput. The channel used for simulations is the IEEE80211n D-NLOS. Complex Gaussian noises, which have -20 dB, -26 dB, -30 dB and -34 dB variance respectively, are modelled as channel estimation errors added to the channels. The channel estimation errors are at the transmitter side, which may occur due to an erroneous feedback channel or the CSI used at the transmitter may be outdated. The CSI at receiver side is assumed to be perfect. Furthermore, turbo coding is included.

In Figure 7.27, the BER performance curves of the THP+Alamouti scheme are depicted. For comparison purposes, the SISO case with perfect channel and estimation channel with -30 dB estimation errors are also depicted in the same figure (red and black curves). The magenta curves are the results without any coding, and suffer from a lack of precision on the BER computation; the blue curves are the results with Turbo Coding for THP+Alamouti.

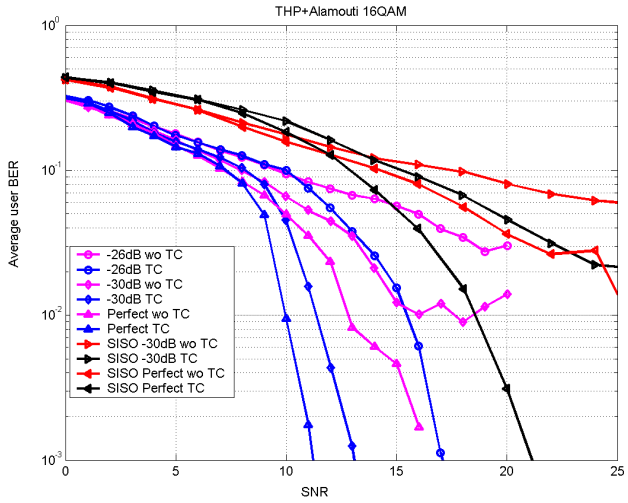


Figure 7.27: The impact of channel estimation errors on the BER performance

Actually, the channel estimation error power compared to the power of the channel itself is small. However, from the Figure 7.27 we see that the performance degradation due to the channel estimation error is quite significant, and that the THP+Alamouti scheme is more robust than the SISO case. We also observe that Turbo Coding can greatly reduce the BER performance degradation due to channel estimation errors.

Figure 7.28 gives the throughput performance when different levels of channel estimation errors are added to the real channel. The throughput is the average over the two users. On Figure 7.28, we observe that at low SNR, the throughput degradation due to channel estimation errors is less than the degradation at high SNR case, i.e., the channel estimation errors have great impact on the throughput performance in the high SNR region.

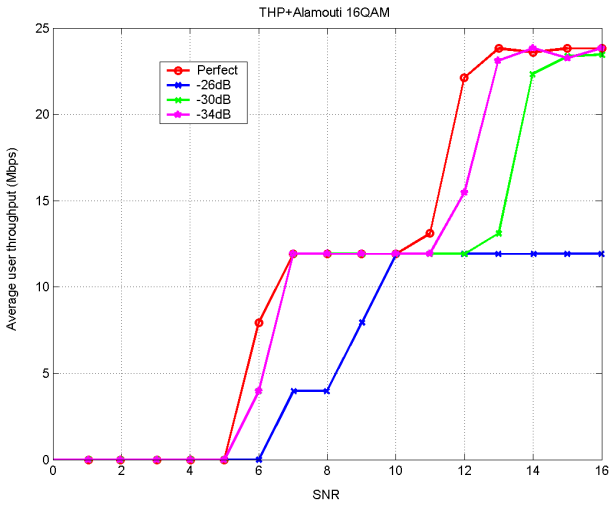


Figure 7.28: The impact of channel estimation errors on the throughput performance

7.3.4 Space-frequency coded MIMO OFDM systems

In this section, computer simulation results for our proposed space-frequency coded OFDM system [D2.3] in the short-range environment are illustrated. Figure 7.29 and Figure 7.30 illustrate the cell throughput cumulative distribution function (CDF) for the 4-QAM modulated OFDM systems. Different channel coding schemes have been adopted, more specifically, 3GPP turbo code with coding rate ½ based on $(13, 15)_{oct}$ and space-time turbo coded modulation (STTuCM) [Tuj03] with max-Log-MAP decoding algorithm. The fading channel considered in these two figures is IEEE 802.11n C NLOS. Number of subcarriers is 128. Perfect channel synchronisation and channel estimation are assumed. The receiver is a symbol-level minimum mean-squared error (MMSE) filtering linear detector, which is described in more detail in the Deliverable 2.3 [D2.3].

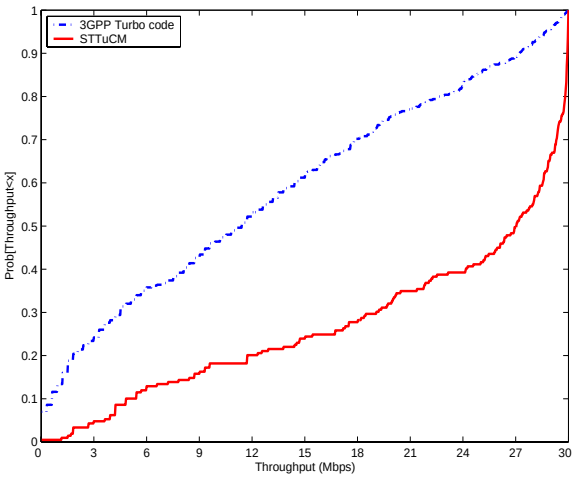


Figure 7.29: Throughput CDF performance of OFDM system with 3GPP turbo code and STTuCM and SNR = 4 dB

Average power control is assumed. From the results shown in these two figures, obviously the STTuCM provides much better performance than the 3GPP turbo coding. It is also concluded that SNR = 4 dB is not enough to get satisfied performance for both coding schemes. At SNR = 6 dB, the OFDM system with STTuCM can achieve satisfying performance with very high user throughput.

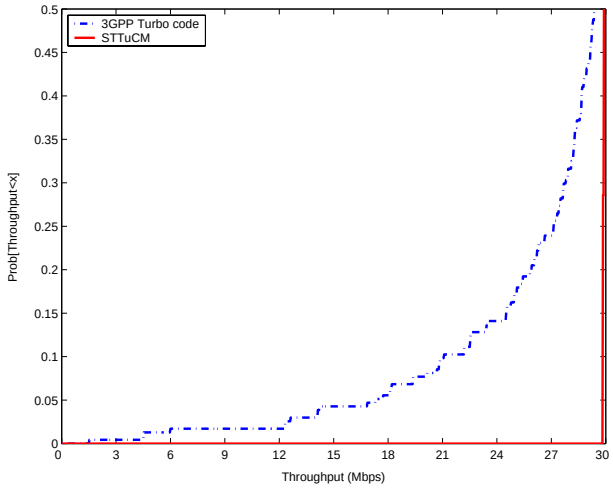


Figure 7.30: Throughput CDF performance of OFDM system with 3GPP turbo code and STTuCM and SNR = 6 dB

After a closer observation of the curves above, if 95% is the satisfying threshold, the throughput provided by turbo coded system is about 20 Mbps, however, STTuCM coded system can offer nearly the maximum throughput for the system, i.e., 30 Mbps. This is because when turbo code is used in multi-antenna systems, the transmitted signal from different antennas will interfere with each other and the adopted receiver is not powerful enough to completely remove this interference. Only spatial diversity gain provided by the receiver can be obtained. On the contrary, for STTuCM coded system, since the transmitted signals from these two antennas are correlated due to the coding structure and their correlation is used by the decoder, in addition to no interference, both spatial diversity gain and coding gain can be achieved. So the advantages offered by STTuCM can be seen clearly.

7.3.5 Control Overhead

7.3.5.1 Spatial Link Adaptation

The control overhead required for spatial domain link adaptation depends on both, the number of subcarriers in the OFDM system and the number of transmission modes that can be supported. No channel state information is required at the transmitter: based on the estimated channel, the receiver can feedback the appropriate mode on each subcarrier. In order to ensure a reasonable performance, this information needs to be updated as the channel changes. Therefore, multiple updates have to be done during a channel coherence time.

In Table 7.13, we assume that the adaptation is performed every 10 frames, or equivalently, every 240 OFDM symbols. As only two transmission modes are envisaged, namely spatial multiplexing and Alamouti, only 1 bit is required to convey the mode information. The control information is the same for forward and return channels.

Table 7.13: Preliminary Estimation of Return and Forward Control Channels Overhead for Spatial Link Adaptation

Control information		**e.g., CQI**
Estimated information bits per message		$1 \cdot 48$
required update rate	**in time**	$(10 \cdot 96\ \mu s)^{-1}$
	in frequency	1
	in space	1
	in user domain	1
control information rate per user		50 kbps
total control information rate per user		**50 kbps**

The control overhead could be substantially reduced if the update rate was decreased. However, attention has to be paid to the performance penalty incurred by transmission modes that are not optimally suited to the actual channel. Additionally, control overhead could be decreased by exploiting the correlation among subcarriers and grouping them into subsets using the same mode.

7.3.5.2 Successive Minimum Mean-Square-Error Precoding

In case of SMMSE we assume that the reciprocity assumption is valid and that this requires a calibration of the uplink measurements approximately once a day. The mobile terminal feeds the estimated complex channel coefficient for each transmit-receive antenna pair back to the base station/access point and for this we assume that it requires 8 bits. The number of antennas at the base station/access point and the number of antennas at the *i*-th mobile terminal are denoted by M_T and M_{R_i}, respectively.

Next, we assume that the base station provides the information about the number of spatial streams used and the MCS used on each spatial stream to the mobile terminal before each transmission. Here we assume that 4 spatial streams and 16 MCSs can be used. The number of spatial streams of the *i*-th user is denoted by Q_i. The parameters are kept fixed during of one frame of duration 337.2 μs.

Table 7.14: Preliminary estimation of Return Control Channel Overhead for SMMSE

Control information		**CSI**
Estimated information bits per message		8
required update rate	**in time**	Once a day
	in frequency	48
	in space	$M_{R_i} \times M_T$
	in user domain	1
control information rate per user		$M_{R_i} \times M_T$ 0.005 bps
Total control information rate per user		$M_{R_i} \times M_T$ 0.005 bps

Table 7.15: Preliminary estimation of Forward Control Channel Overhead for SMMSE

Control information		**Modulation information**	**#spatial streams Q_i**
Estimated information bits per message		4	2
required update rate	**in time**	337.2 μs^{-1}	337.2 μs^{-1}
	in frequency	4	4
	in space	Q_i	1
	in user domain	1	1
control information rate per user		Q_i 23.7 kbps	17.8 kbps
total control information rate per user			Q_i 23.7 + 17.8 **kbps**

7.3.5.3 Joint THP with Diversity Techniques

The THP+DT scheme is used for the downlink transmission, therefore only the forward control channel overhead is considered. The return control channel overhead can be neglected. The overhead calculation is according to the system settings. Assume that there are N_t antennas deployed at the BS and N_r users each having N_a antennas, the overhead transmission is calculated in the following tables.

In the non-adaptive mode, the modulation for one link is fixed. One transmission of this overhead information holds for all the frames. Therefore the overhead rate is very low and can be ignored.

Table 7.16: Preliminary estimation of CSI for mobile users (Forward Control Overhead per user)

Control information	**e.g., CSI**
Estimated information bits per message	$48\cdot(N_a\cdot 8+5)$ bits
required update rate in time	24·4 μs
control information rate per user	$(4\cdot N_a+2.5)$Mbps
total control information rate per user	$N_r\cdot(4\cdot N_a+2.5)$Mbps

Table 7.17: Preliminary estimation of Forward Control Channel Overhead with non-adaptive transmission

Control Information	**Estimated Information Bits**	**Comment**
Modulation M(c,q)	2 bits	per transmission
Code Rate of FEC_o	6 bits	per transmission;
Power P(c,q)	6 bits	Power control per subcarrier
resource map		No chunk and spatial layer IDs needed, since the user interference is completely cancelled with THP+Alamouti.
pilots		No pilots needed

In the adaptive mode, the channel is assumed to be stable over the transmission of one frame.

Table 7.18: Preliminary estimation of Forward Control Channel Overhead with adaptive transmission

Control Information	**Estimated Information Bits**
Modulation M(c,q)	48·2 bits
Code Rate of FEC_i(c,q)	6 bits
Power P(c,q)	48·6 bits
Total bits per frame	198bits
Total time need	24·4μs
Total rate per user	2.06 Mpbs

7.3.6 Conclusions

Section 7.3 presents a first evaluation of the performance of the following techniques:

- Spatial Multiplexing Adaptation
- Successive Minimum Mean-Square-Error Precoding
- Joint THP with Diversity Techniques
- Space-Frequency Coded OFDM System

in a short-range scenario. The performance of these techniques is mainly evaluated in terms of coverage, user and cell throughput.

The performance of a 2x2 system using either the Alamouti scheme, the spatial multiplexing approach, or the spatial link adaptation technique for switching between these two methods is compared to the performance of a single antenna system. The simulation results show that a significant gain is provided by the spatial link adaptation technique compared to Alamouti coding or spatial multiplexing.

The Successive Minimum Mean-Square-Error Precoding approach is investigated for a single user, in order to evaluate its performance in a simplified context before addressing the multi-user context. A very high spatial multiplexing gain is confirmed for all modes that have been simulated.

The performance of joint THP with diversity techniques is presented in a multi-user scenario. The throughput gain that is provided by a multi-antenna system using this approach compared to a single antenna system is tremendous. Besides the robustness of joint THP with diversity techniques to channel estimation errors at the transmitter, due to an erroneous feedback channel for instance, is evaluated. The system is shown to still provide significant gain compared to a single antenna system, and to be more robust.

These three techniques rely on some level of channel quality information or channel state information at the transmitter. Consequently a preliminary study of the control overhead in the forward and return link has been completed. The information that is provided for each technique gives some inputs on its complexity and can be correlated to the level of performance provided by this technique. For instance spatial link adaptation provides a significant gain with low overhead, whereas very important gains can be demonstrated for joint THP with diversity techniques requiring a very high amount of overhead transmission.

The last approach investigated in the short-range scenario consists in a space-frequency coded OFDM system. Two systems, based respectively on the 3GPP turbo code and space-time turbo-coded modulation are compared. A significant increase in throughput is provided by the last solution compared to the turbo code based one, thanks to the coding structure across the antennas of space-time turbo-coded modulation.

7.4 Performance Results for Relaying

In this section, we assess the performance of Decode-and-Forward relaying as introduced in Section 3.4 for the short-range scenario.

7.4.1 MIMO Relaying

In this Section, we assess the performance of some well-known multi-antenna methods in a relaying network. We consider an OFDM transmission scheme, which has been identified as a potential candidate for the WINNER concept in the short-range usage scenario [D7.2]. Different MIMO transmission modes are taken into account: in a first step we consider diversity coding using the Alamouti scheme with diversity orders 2 and 4 (i.e., 2x1 and 2x2 systems respectively), 2x2 spatial multiplexing, and as a reference system, the single-antenna (SISO) mode. Subsequently, we study systems with a higher number of antennas by considering some selected schemes out of the ones proposed in Section 3.1.2.

We study here decode-and-forward relaying as described in Section 3.4.1. Our TDMA scheme reserves two time slots for each link between source and destination. During the first time slot, the source can transmit its packet to the relay and/or the destination. Obviously, there will be network setups were the relative position of the three stations will make it possible that both, the relay as well as the destination get this packet. Other setups will also see only one of those stations receiving successfully the transmitted packet, and finally, the packet might not be received at all when the source is placed excessively far. During the second TDMA time slot, the relay will retransmit the packet received during the first time slot, and the source will remain silent (see description in Section 3.4.1.1 for more details). From an information-theoretic point of view, this scenario is not optimal: the best performance is achieved when the source transmits again to the destination in the second time slot [NBK04]. In our setting, we declare that a packet has been successfully received when the destination has been able to decode it either during the first or the second time slot. Note that packets successfully decoded by the destination during the second time slot must have been necessarily decoded at the relay during the first one.

We also assume that the packet processing time by the relay is much smaller than the time slot duration such that it can be neglected. The radio channels have been modelled using the IEEE 802.11n channel model D for small antenna configurations and channel model E for larger ones [D5.1]. The first one covers the typical office environment whereas the second assumes a large office environment. For each of those channel models, we have fixed the cell size to a maximum of 30 meters.

To assess the performance gains of MIMO relaying in terms of cell coverage, we simulated the basic TDMA-based relaying network consisting of a single access point, mobile station, and relay station. The

distance between the AP and the MS has been set to a fixed value of 30 meters, while the relay has been located at various positions within the square of 30x30 meters between the AP and MS. We applied the relaying scheduling strategy as discussed by means of Figure 3.32. Since the AP assigns two consecutive time slots for the transfer of each packet from the AP to the MS, the throughput normalised to the packet rate can never exceed 0.5. In all simulations, the signal-to-noise ratio has been set so that the MS experiences a normalised throughput performance of 10% in the absence of a relay.

Figure 7.31 shows the normalised throughput as a function of the relay position. The position is normalised so that the access point is located at the coordinates (0,0) and the mobile station at (1,0). We have obtained throughput performance results in case of utilizing in all radio transceivers either the Alamouti space-time coding scheme with two transmit and one (or two) receive antennas, the spatial multiplexing scheme with two Tx/Rx antennas, or the basic underlying OFDM transmission scheme with a single Tx/Rx antenna. The blue reference plane in Figure 7.31 represents the throughput if no relaying is used and shows the user's throughput if the communication would solely rely on the direct link between the AP and MS. Adding a relay at any location in the considered square improves the throughput performance and cell coverage. The biggest improvements are obtained for all signalling modes if the relay is located in the dark read area between AP and MS. Note that for diversity modes the size of this area is bigger than for SISO or spatial multiplexing modes. More generally, using relays in conjunction with the 2x2 or even the less complex 2x1 Alamouti scheme results in a higher throughput that can be guaranteed over an extended area compared to SISO relaying. Spatial multiplexing combined with relaying also leads to improvements over SISO relaying in terms of cell coverage.

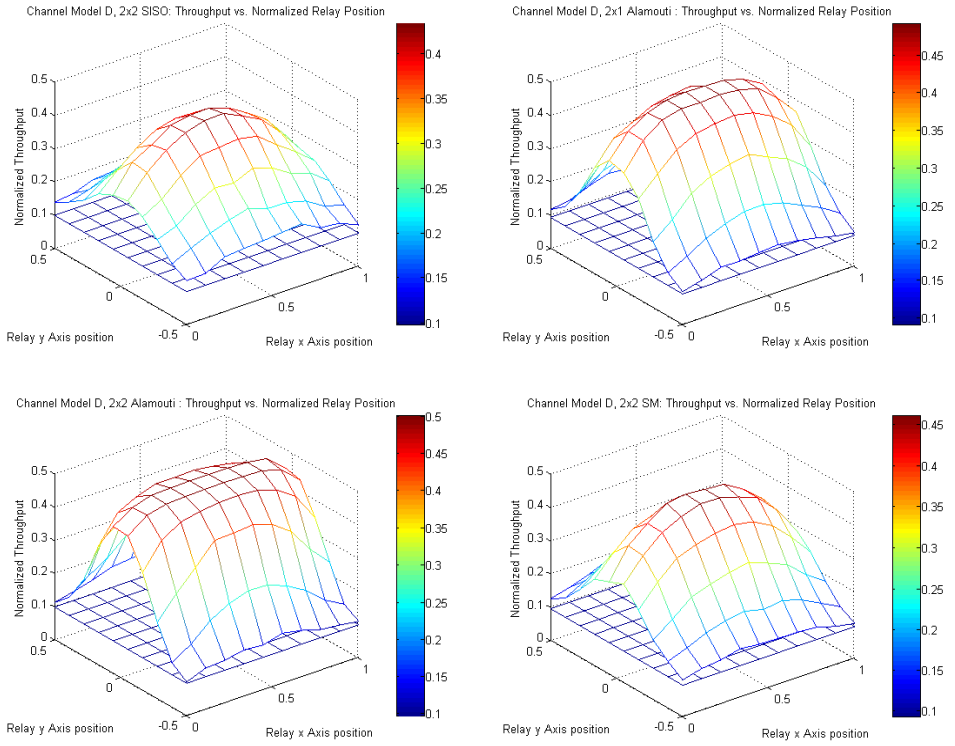


Figure 7.31: Normalised throughput as a function of relay position using DF combined with SISO, 2x1 Alamouti, 2x2 Alamouti, and 2x2 Spatial-Multiplexing

We turn our attention now to the individual user throughput. Since we consider a perfectly synchronised TDMA network, there is no intra-cell interference. The inter-cell interference is not part of our current single-cell assessment. We study first a network with 4 stations, each of those using one of the transmission schemes considered above. We set the relay at 10 meters from the AP, and set a rather low SNR of 10 dB. Figure 7.32 shows the empirical throughput for each of the considered users. The maximum throughput achievable with the considered modes is of 12 Mbps. However, in order to take into account the throughput hit due to the use of two subsequent time slots when relaying and allow a fair comparison,

the relaying modes are in fact transmitting at 24 Mbps.

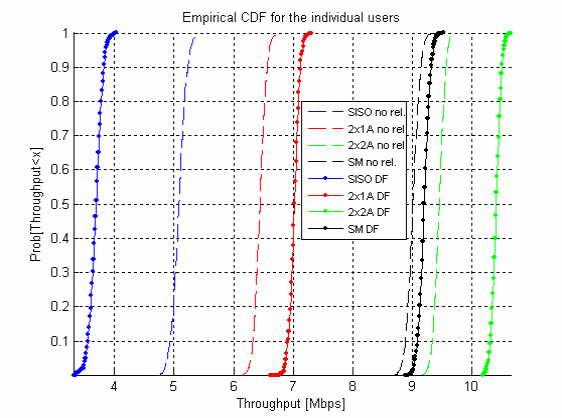


Figure 7.32: User Throughput of SISO, 2x1 Alamouti, 2x2 Alamouti, and 2x2 Spatial-Multiplexing with and without DF Relaying

We note that relaying improves the distribution of the users throughput distribution function for all the cases but one: the throughput degrades for the SISO mode. This can be explained as follows: the relatively low SNR regime at which we are working prevents the SISO packets to be systematically decoded at the relay. Therefore, relaying does not help at all in the transmission process and the system takes the hit due to the usage of two time slots without really getting the gains expected from relaying.

Furthermore, we note that the improvements seem to be more important as the diversity in the system increases: the second-order diversity systems gains in throughput but less than the fourth-order diversity system. Spatial multiplexing is using a considerable part of the antenna diversity to separate the transmitted streams at the receiver, which translates in a smaller relaying gain. However, it can be expected that the gains over the non-relaying scenario will be much more important as the SNR increases.

Figure 7.33 does the same experiment as above but with higher antenna configurations. In fact, the four considered transmission modes are chosen out of the modes proposed in Section 3.1.2 and based on hybrid SMUX/STBC schemes. The relay node is positioned at a distance of 10 meters from the access-point, and the SNR is set to 20 dB. Again, we conclude that the modes that have more diversity show more important gains. This is particularly noticeable in the case of 3 transmitted streams, where the relaying gain for the considered SNR is marginal. When comparing the 4x4 modes with 2 streams, it is interesting to note that for reasonably high SNR the relaying gains seem to be more important for higher rates.

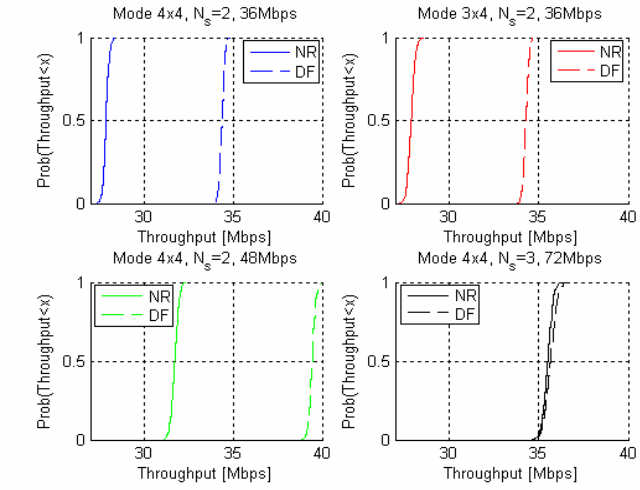


Figure 7.33: User Throughput with and without DF Relaying for different modes based on hybrid SMUX/STBC techniques

We now consider the cumulative distribution function of the aggregated throughput for the modes considered above. For a multiple-access interference-free system such as the one assumed here, the average throughput over the four users is representative of a larger network.

Figure 7.34 shows the aggregated throughput for the two. Although direct comparison between both is not possible due to a different SNR, we can still conclude that higher gains are visible when higher MIMO configurations are used.

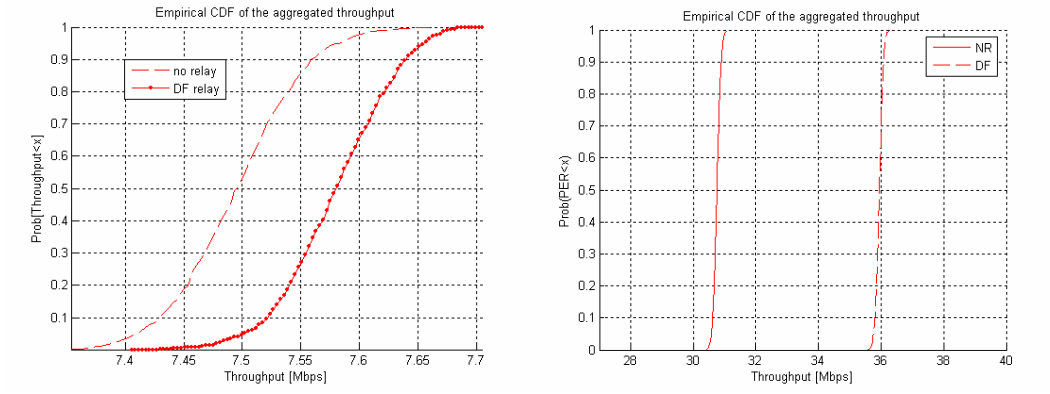


Figure 7.34: Aggregated Throughput with and without DF Relaying for transmission schemes with up to 2x2 (left) and 4x4 (right) antenna configurations

7.4.2 Control Overhead

Relaying does not incur additional control overhead at the physical link. However, considerable overhead is required at upper layers to control and coordinate a network comprising relays. These topics are out of the scope of this deliverable.

7.4.3 Conclusions

The performance of DF relaying has been assessed by simulations in a short-range scenario. The impact of multi-antenna transmission techniques on the relaying performance has been illustrated in terms of coverage, user throughput and aggregated throughput. In terms of coverage, the spatial diversity introduced by multiple antennas extends considerably the area in which a relay can be placed without throughput degradation. Concerning the user and aggregated throughputs, multi-antenna systems not only recovered the loss incurred by using two temporal time-slots for relaying, but they were also shown to outperform the direct link performance. The gain depends on the distances source-relay and relay-destination, and for fixed distance and power, higher diversity modes achieve higher gains.

7.5 Conclusions and Overall Comparison of Techniques

In this chapter, performance results have been presented for the wide-area, the short-range scenario and relaying in order to evaluate the potential gains from the deployment of the multi-antenna concept and the associated complexity and overhead signalling in the framework of the WINNER system concept and assumptions. The performance of the multi-antenna techniques was evaluated in terms of the selected assessment criteria, namely cell and user throughput, coverage and spectral efficiency and control signalling overhead.

From the simulation results in the wide-area scenario it was observed that directivity obtained by beamforming has high potential for scenarios with small angular spread, since it provides a good performance versus complexity trade-off. For a given number of users, cell throughput scales nearly linearly with the number of beams up to a certain maximum, which is given by the achievable beam width in combination with the required sidelobe suppression. When beamforming is combined with adaptive resource assignment based on short-term CQI, considerable performance enhancements can be obtained due to the exploitation of multi-user diversity, especially for high load and increasing angular spread in the propagation channel.

For the space-time-frequency coding and modulation, linear dispersion codes (matrix modulation) are considered as baseline for further investigations. Linear precoding to account for antenna correlation adds significant gain to pure matrix modulation (STBC) at a very low overhead signalling cost, especially in the case of highly correlated channels, as for example in the urban macro scenario. The linear precoding can be applied to any matrix modulation mode and will be considered as an additional feature to spatial link adaptation when long-term CSI is available at the transmitter. Spatial mode selection with adaptivity to long-term CSI can provide substantial gains within each scenario, and considering the fact that the WINNER multi-antenna concept must adaptively support many different scenarios it will be an essential feature of the proposed design.

The multi-cell system evaluations indicate that downlink beamforming based on a set of fixed beams and terminals with multiple receive antennas and interference suppression capabilities are suitable for the highly loaded wide-area scenario considered. Another interesting observation is that the relative benefit of an additional SDMA component may be relatively small, at least when no channel dependent scheduling is used. For low-load scenario, single-site results indicate that PARC may be used to achieve high cell throughput and user peak rates. Further, in the high load multi-cell case, it appears to be beneficial not to use all spatial degrees of freedom for spatial multiplexing and instead use some of them for terminal interference suppression.

The multi-user multi-antenna signal design problem has been investigated with focus in the short-range scenario both from a communication theory and a simulations perspective. Multi-user precoding techniques are restricted to linear precoding (e.g., SMMSE or BD based on long-term statistics). Techniques like BD provide performance improvements primarily for multi-antenna mobile terminals. SMMSE is especially attractive since it does not suffer from the restrictions considering the number of antennas in the system, provides good performance for various terminal classes, and can exploit either long-term or short-term CSI at the transmitter to perform the precoding. For the short-range scenario precoding techniques for further consideration include also non-linear techniques that relay only on the short-term CSI at the transmitter due to their sensitivity to the channel estimation errors. SO THP, SMMSE THP and SMMSE are of interest in realistic scenarios with a mixture of terminal classes.

Relaying and multi-hop communication is a concept with high potential. It can provide additional performance enhancements when spatial processing is applied on the corresponding links. Spatial multiplexing combined with relaying also leads to improvements over SISO relaying in terms of cell coverage.

One important focus of future research will be the investigation of the optimum amount of adaptivity in terms of the associated overhead signalling requirements. For techniques exploiting long-term CSI, as for example linear precoding for matrix modulations, the resulting gains are large at very low expense in terms of overhead signalling. At the example of limited CSI spatial multiplexing (PARC, PSRC), it is shown that the control overhead estimations have a dynamic range of the order of a hundred, depending on the degree of link adaptivity and flexibility in channel dependent scheduling. This impressive range clearly shows that after a certain breakpoint further adaptivity will merely increase complexity and overhead without providing corresponding performance gain.

Further investigations are required once a baseline WINNER channel code is available, since coding reduces the impairments due to channel estimation errors significantly. In particular for the techniques that rely on short-term CSI at the transmitter further studies regarding robustness and the feasibility of exploiting reciprocity in TDD systems are required.

8. System and Terminal Complexity

In this Chapter we address the impacts of spatial processing on system and terminal complexity. Directivity brings restrictions and additional requirements to pilot design and use and therefore has impact on the overall system design (Section 8.1). Since several spatial processing techniques require short-term channel knowledge at the transmitter, it is important to investigate the feasibility of exploiting channel reciprocity and to understand the complexity and overhead of an alternative solution based on return link feedback (Section 8.2). First investigations and conclusions regarding baseband and RF complexity, as well as the split of complexity between the network and the user terminal are presented in Section 8.3 and 8.4.

8.1 The Relation of Pilots and Spatial Processing

Control information is required to enable certain physical layer support functions, e.g. connection setup, mobility support (handover), power control, synchronisation, CQI measurements and most importantly the channel estimation. Pilots, i.e. symbols with a priori known content on particular resources, are generally used for implementing such support functions. In order to realise an efficient system, synergy effects should be exploited by reusing the same pilots for different physical layer support functions. Spatial processing, however, limits the potential reuse of pilots and brings along additional requirements for pilot design.

Employing multiple antennas at the base stations yields directivity not only with respect to the transmitted information but also with respect to the pilots. This directivity is generally intended and controlled by the transmit antenna weights **f**. If user-specific beams are formed it is beneficial to transmit data and pilots with the same antenna weights $\mathbf{f}_U$ since the user U is only interested in his effective channel $\mathbf{H}_U \cdot \mathbf{f}_U$. However, these so-called dedicated pilots prevent other users V from estimating their effective channel characteristics $\mathbf{H}_U \cdot \mathbf{f}_V$ at the same time. In case of beamforming with a grid of fixed beams $\mathbf{f}_B$, pilots are not user-specific but beam-specific and thus may be exploited from all users within a beam. In 3G systems this is referred to as "Secondary Common Pilot Channels". Pilots of fixed beams which use the same time-frequency resources have to be orthogonalised via time, frequency or code (see below).

Certain spatial processing techniques need to estimate the complete channel matrix $\mathbf{H}_U$ at the receiver. Therefore, the pilots have to be separated per transmit antenna. Orthogonality can be obtained via separation in frequency or time. Another option is to obtain orthogonality by code via spreading with the advantage of spreading gain and the disadvantage of multi-code interference and loss of temporal resolution (if spreading in time direction is applied). It is desirable to reuse these "pilot patterns per antenna" in a common channel to create pilot overhead only once for all users. Such a common channel is required to broadcast information to all users as, e.g., cell identifier or other management information. A disadvantage of the "pilots per antenna" is that they cannot benefit from array gain and thus need a larger Tx power to obtain a satisfactory channel estimation quality. The needs for common and dedicated pilots that have been identified above are summarized in Table 8.1.

Table 8.2: Overview of the different mentioned types of pilots for spatial processing

Pilot Type	Pilot Directivity	Intended Users	Purpose
Dedicated Pilot	Effective channel $\mathbf{H}_U \cdot \mathbf{f}_U$ of user U	1	Effective pilot caused by e.g. beamforming
Common Pilot per Cell	For a whole cell (omni) or sector (sector pattern)	All	Mobility
Common Pilot per Beam	One fixed beam covering a part of a sector, thus effective channel $\mathbf{H}_U \cdot \mathbf{f}_B$ of beam B, separated per beam	All using fixed beams	Effective pilot of fixed beams: "Secondary Common Pilot Channel"
Common Pilot per Antenna	Omni, separated per antenna	All which need more than effective channel	Pilot to obtain unweighted channel matrix **H**

The quality of the channel estimation depends on the power and time-frequency resolution of the pilot pattern. The required frequency resolution depends on the maximum length of the channel impulse response, whereas the required time resolution depends on the maximum Doppler shift of the mobile. These questions of OFDM pilot design are addressed in [D2.1] and will be covered by Task 2.2. The needs for common and dedicated pilots within the WINNER multiple antenna concept have to be further assessed, based on the selected techniques and in a joint effort with T2.2. Obviously, the physical channel structure of the WINNER air interface has to support these necessary pilot functions to enable all relevant spatial processing techniques. Additionally, the robustness of each spatial processing technique against channel estimation error has to be taken into account. A technique which is less robust requires more pilot overhead which has to be subtracted from the total Tx power budget. Consequently, this pilot overhead must be traded off with the associated gains of such techniques. As simulation results of D2.7 are based on perfect CSI and the kind of multiple access and physical channel multiplexing is still under study, these questions can only be investigated in more detail later on.

8.2 Obtaining Short-Term Channel Knowledge at the Transmitter

Multi-antenna algorithms can be split in two categories, depending whether CSI is supposed to be available at the transmitter or not: when CSI is assumed (resp. not assumed), multi-antenna techniques are referred to as closed-loop (resp. open-loop) techniques. Closed-loop techniques are known to provide substantial performance benefit in wireless communication systems compared to open-loop techniques [Tel95]. However, these techniques require obtaining the knowledge of the MIMO channel prior to data transmission.

In the general case, a feedback link can be established and used by the receiver to forward the estimated channel to the transmitter; the main drawback of this solution is the additional overhead required and its throughput loss induced. For TDD systems in which uplink and downlink channels share the same frequency, the pure propagation channel can be assumed to be reciprocal from Babinet's principle [Bal82]. Hence in principle the transmitter can acquire CSI by estimating the propagation channel over the signal it receives, and no feedback link is required. However, the effective MIMO channel is a combination of the MIMO propagation channel and the RF front-end contributions. The RF front-ends cannot be considered to be exactly the same on the UL and the DL, since i) the components used are not necessarily the same and ii) their characteristics are guaranteed only within a production spread even if they correspond to the same product device. As a consequence, the imperfections of the RF front-ends might be inconsistent with the reciprocity assumption.

Moreover, in TDD systems, UL and DL are separated in time, implying that between the MIMO channel estimation on the UL and the use of these estimates on the DL, the channel experienced might have changed, as for instance in a mobility scenario. Due to the delay between UL and DL, the CSI at the transmitter might be incorrect.

In the following, we investigate the channel reciprocity assumption in the context of the TDD short-range scenario. We first explain how the imperfections of the RF front-ends might modify the UL and DL MIMO channels. Then, based on observed characteristics of actual components, a model for these imperfections is introduced, that will be used to assess their actual impact on the performance of closed-loop methods requiring either perfect or partial CSI. Finally, simulation results are presented, and the impact of the delay between the MIMO channel estimation on the UL and the use of these estimates on the DL is also discussed in a mobility scenario.

8.2.1 RF Front-end Imperfections and Model

Fundamentally, the reciprocity of the propagation channel is known as Babinet's principle [Bal82] and is a basic tenet of antenna theory. However, this holds only for the propagation channel which does not take into account the RF front-end components. It is the imperfections of the RF front-ends, which also contribute to the effective channel experienced that might put in question the exploitation of the reciprocity principle.

These imperfections can be considered as being due to amplification (or loss) or due to filtering of the transmitted or received signal. In general, the difference between these two effects is that amplifiers

generally have constant gain over the bandwidth of interest, while a filtering effect might have different amplitude and phase effects upon each of the subcarriers of the OFDM-based system considered herein. These effects can be summarised in Figure 8.1 for a TDD system having a transmitter (access point) with two transmit antennas, and a receiver (mobile terminal) with one receive antenna. In this figure, h_1 (resp. h_2) is the propagation channel (reciprocal) between the first (resp. second) antenna of the transmitter and the receive antenna.

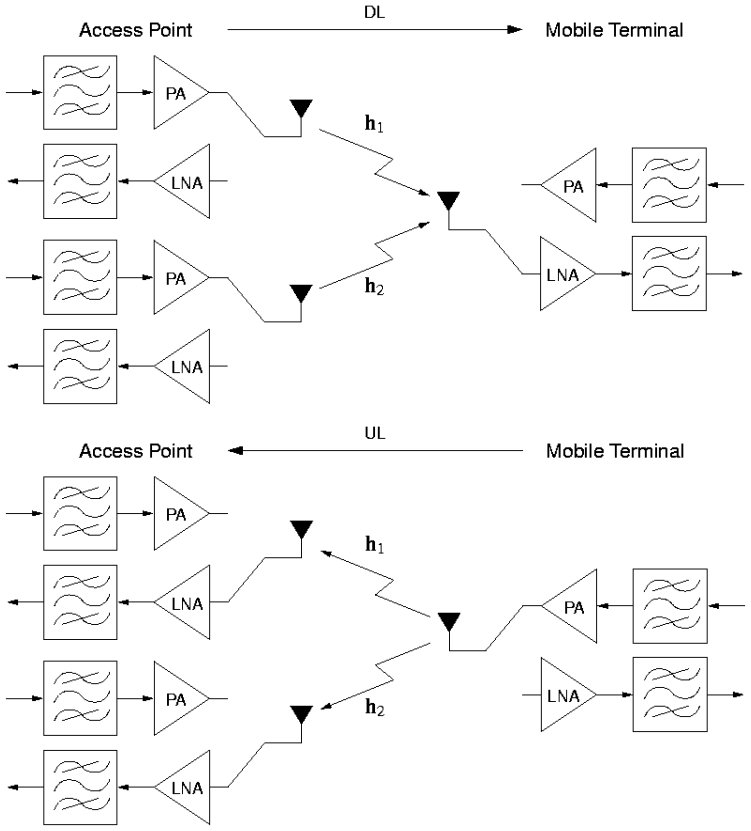


Figure 8.1: Contributions of RF front-ends

It is the relative difference in the UL/DL blocks that is important when considering the channel reciprocity assumption. For instance, if the UL filter differs from the DL filter, then so will the channels. In fact, even if both filters correspond to the same product device, they will never have exactly the same characteristics due to the production spread of the device used. As a consequence, the conclusion that can be drawn from this remark and Figure 8.1 is that, in practice, the channel reciprocity assumption (including RF front-ends) does not hold since UL and DL channels are not identical.

However, the impact of these differences on the performance of closed-loop methods needs to be investigated. The major advantage of CSI at the transmitter based on the channel reciprocity assumption, compared to introducing a feedback link, is that it is a low complexity solution. Actually, the choice of assuming the channel to be reciprocal should be based on a performance/complexity trade-off: implementing a feedback link should be chosen only if the performance gain is significant.

We introduce a model for the RF front-end imperfections, derived from Figure 8.1. This model is based on observed characteristics of actual components, and can be considered as representative of current RF technology. The goal is to take into account these imperfections in order to quantify their impact on the performance of closed-loop methods based on channel reciprocity. The model is depicted in Figure 8.2 for a system with two transmit and one receive antenna, but it can be generalised to any multi-antenna system.

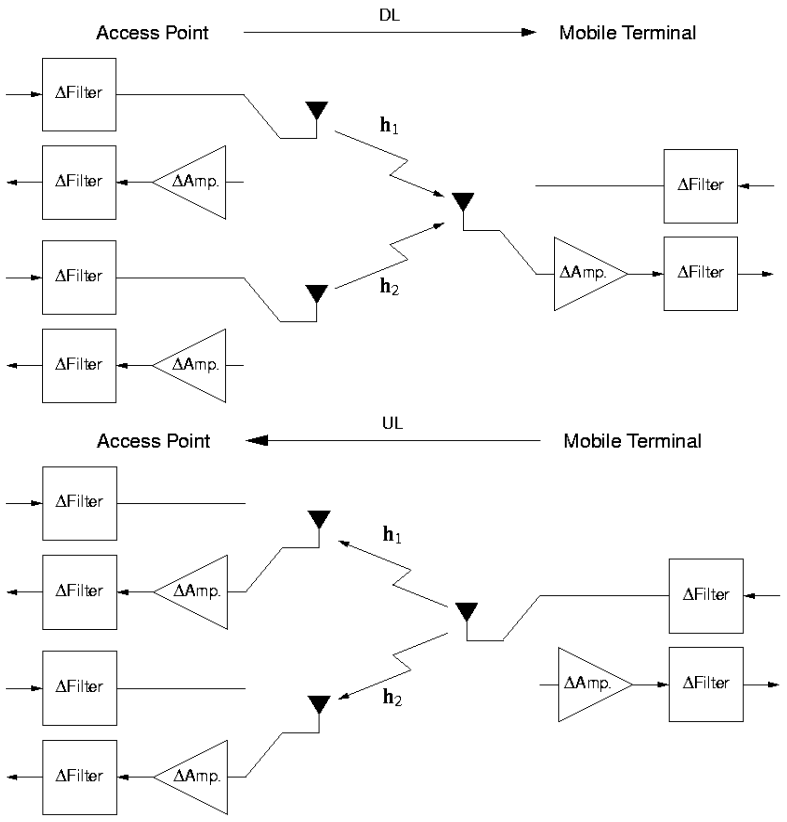


Figure 8.2: RF imperfections model used

Each component in Figure 8.2 represents a deviation from a nominal device, in terms of gain and phase. The deviations are independent from one component to another. Two different types of deviations are considered, corresponding to two different devices:

- ΔAmp. represents amplifiers' effects, as well as the deviation due to the automatic gain control module (i.e., uncertainty on measured power). The gain deviation in dB is modelled as a normally distributed random variable with zero mean and a standard deviation equal to 1.5dB. The phase deviation will be set to various deterministic arbitrary values, and is assumed to be the same for all subcarriers (not very high in practice).
- ΔFilter represents filters' effects. The gain deviation in dB depends on the subcarrier index and is modelled for each subcarrier as a uniformly distributed random variable over the interval ±1dB, representing the ripple of the filter and its production spread. The phase deviation will be set to various deterministic arbitrary values, and is assumed to be the same for all subcarriers (we consider that the group delays of the filters are constant in-band); unlike amplifiers, the phase deviation can be rather important for filters.

Of course this model could be further refined, but it is a good basis for performance assessment of closed-loop methods using the reciprocity assumption.

8.2.2 Simulation Results for RF Imperfections in Closed-loop Multi-antenna Solutions

Two closed-loop methods will be simulated with the model for the RF front end imperfections introduced above: i) the TxAA scheme, which assumes perfect CSI at the transmitter and ii) the Transmit Selection Diversity scheme, which requires only partial CSI at the transmitter.

The transmit adaptive array (TxAA) technique makes use of multiple transmit antennas to maximise the signal to noise ratio at the receiver. For that purpose a transmit weight factor chosen to maximise the received signal power is applied to the transmit signal. As it requires perfect CSI at the transmitter, this scheme is sensitive to RF front-end imperfections.

The Transmit Selection Diversity (TSD) scheme, which requires only partial CSI knowledge at the trans-

mitter, aims at selecting the best channel for data transmission. The interest for studying TSD lies in its simplicity: only partial CSI is required at the transmitter side. The sub-optimality of TSD is given by the fact that we transmit over the best channel, instead of transmitting over the strongest eigenvalue of the MIMO channel matrix, as for TxAA. When applying the TSD scheme to OFDM-based multi-antenna systems, we can consider two different approaches, i) selecting the best transmit antenna for all subcarriers or ii) selecting the best transmit antenna for each subcarrier. We will focus on the second approach since it is more interesting from a performance point of view. In addition to its simplicity, TSD can tolerate reasonable channel estimation errors, since it only needs to estimate the best channel. Therefore, the TSD scheme might be robust to RF front-end imperfections.

We now present simulations results obtained with the model for the RF front-end imperfections, for both the TxAA and TSD closed-loop schemes. We consider an IEEE802.11a-based multi-antenna system with 4 transmit and 1 receive antennas. The performance of the system is evaluated in terms of PER, for 16QAM constellation symbols and a code rate of 1/2. The packet length is set to 48 bytes.

The propagation channel, which does not include RF front-end imperfections and thus is reciprocal, is based on the TGn D power delay profile proposed for typical office environment with NLOS conditions. However, these channels are assumed to be spatially uncorrelated and then correspond to a simplified TGn D channel model. In the simulation results presented here, MIMO channels (either on the UL or the DL) are assumed to be perfectly estimated (i.e., not corrupted by noise). However, it should be noted that channel estimation might contribute to the non-reciprocity of the channel since the UL and DL MIMO channel estimates will have different noise contributions. The simulation results obtained for the single antenna IEEE802.11a system and for an open-loop multi-antenna system using the ABBA STBC [TBH00] are also given for comparison purposes; the latter corresponds to a non-orthogonal STBC that fully exploits the transmit spatial diversity. For each closed-loop algorithm, the performance with perfect feedback is also given.

8.2.2.1 Impact of RF Front-end Imperfections

Figure 8.3 shows the PER simulation results obtained with the model introduced above. The phase deviation between UL and DL in the model is set to various deterministic values, and it is assumed to be the same for the RF front-ends of the transmitter. This implies that the phase deviation between UL and DL is the same regardless the transmit antenna we consider. This is not a realistic assumption but is made here to limit the number of variables. However it corresponds to the model obtained after applying the self calibrating scheme described in [NCTH82].

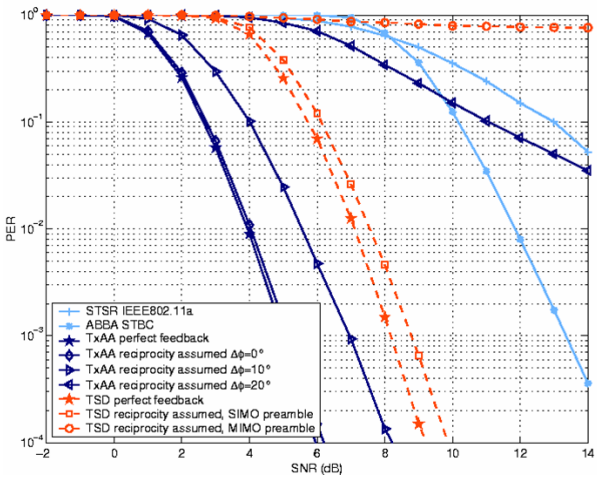


Figure 8.3: Impact of RF front-end imperfections

TxAA simulation results: When channel reciprocity is assumed, and the phase deviation is null, then we can see that the level of performance is close to the perfect feedback case since a loss of approximately 0.1dB is introduced at a PER of 10^{-2}. This means that the TxAA scheme is robust to gain imperfections introduced by the RF front ends when assuming channel reciprocity. However, when the phase deviation

increases, the performance decreases severely since a loss of 2dB is experienced when the phase deviation is set to 10 degrees, while a PER of 10^{-2} might never be reached when it is set to 20 degrees. The conclusion is that if the channel is assumed to be reciprocal, a calibration scheme is necessary in order to compensate for the phase deviation (this is true with one receive antenna; when considering more than one receive antenna both phase and gain would need to be calibrated). This calibration would also result in some additional overhead, but less than for the feedback link since the parameters that need to be calibrated depend only on the RF front-end components and are therefore slowly varying, while feedback is required every time the channel changes.

TSD simulation results: When assuming the reciprocity of the channel, the UL MIMO channel is used for selecting the best antenna for each subcarrier. Two different signalling strategies are considered for DL MIMO channel estimation: i) preambles transmitted using the same TSD scheme, i.e., over used subcarriers only, referred to as Single Input Multiple Output (SIMO) preamble and ii) preambles enabling the estimation of all channels of the system, referred to as MIMO preamble. In the first case, the channel is estimated only on used subcarriers, while in the second case, the receiver estimates the whole DL MIMO channel, from which are detected the subcarriers that were used at the transmitter side. It is clear from Figure 8.3 that the second strategy is not the good one: an error floor is introduced and a PER of 10^{-2} is never reached. This is due to the errors made when detecting for each subcarrier the antenna used at the transmitter side, which occur when two channels have roughly the same power. For the SIMO preamble case, the performance obtained when assuming the channel to be reciprocal is very close to the perfect feedback case (~0.4dB at a PER of 10^{-2}). The practical conclusion here is that, when channel reciprocity is assumed, a mechanism needs to be introduced to ensure that no errors are made when detecting which antenna was used for each subcarrier: this is obtained when using a SIMO preamble. This is no longer necessary when a feedback link is implemented.

8.2.2.2 Impact of Mobility

When channel reciprocity is assumed for closed-loop methods, a delay can exist between the UL on which the CSI is obtained, and the DL on which the CSI is used (this delay will be referred to as CSI delay). Note, that in general the CSI delay will be even greater if CSI feedback is used in stead of exploitation of reciprocity. In a mobility scenario, this implies that the CSI obtained on the UL might become imperfect or even incorrect when used on the DL. It is the impact of both CSI-delay and mobility on the performance of closed-loop techniques that is studied here. Figure 8.4 shows the PER simulation results obtained with mobility conditions (3m/s and Jakes model Doppler spectrum) and various CSI-delays. The model for RF front-end impairments is used with no phase deviation. TSD further uses a SIMO preamble.

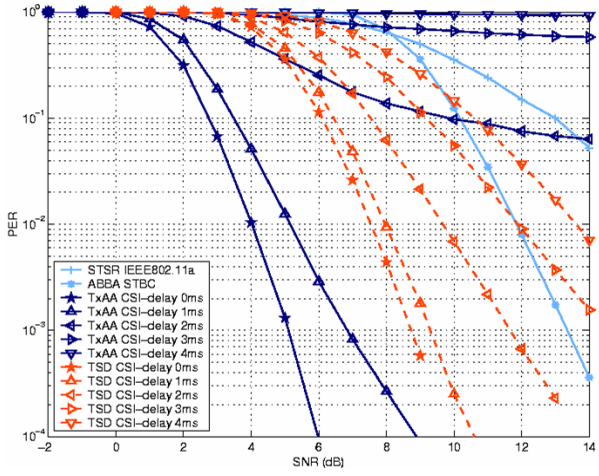


Figure 8.4: Impact of CSI-delay with 3m/s mobility

TxAA simulation results: When the CSI delay increases, the performance of the TxAA scheme rapidly decreases and an error floor is experienced for a CSI delay of 2ms: a PER of 10^{-2} is never reached. However, for a CSI delay of 1ms, the performance loss introduced compared to the instantaneous CSI case is equal to ~1dB for a PER of 10^{-2}, implying that gain over the TSD scheme remains important

(~3dB). Please note that when the CSI delay is equal to 0ms (UL and DL MIMO channels estimated at the same time), the results obtained are similar to those in Figure 8.3, and the MIMO channel is almost static during the packet duration; hence the degradation is due only to the CSI delay.

TSD simulation results: Compared to the TxAA scheme, the performance of the TSD scheme decreases slowly as the CSI-delay increases. For a CSI-delay of 1ms, the loss introduced compared to the instantaneous case is equal to ~0.4dB for a PER of 10^{-2}. Additionally, even for large CSI-delays, the performance is always higher or similar to the STSR IEEE802.11a case, which means that the TSD scheme is robust to CSI-delays. However, when compared to the open-loop system using the ABBA STBC, the performance of the TSD scheme remains higher only for CSI-delays lower than 3ms (for a PER of 10^{-2}), implying that the ABBA STBC should be used instead of the TSD scheme for very large CSI-delays. The advantages of the TSD scheme over the ABBA STBC are i) a large performance gain when no mobility is considered (~4.5dB in Figure 8.3) and ii) low-complexity decoding scheme since it is backward compatible with IEEE802.11a receivers.

8.2.3 Further Aspects

Apart from the impact of RF imperfections there are further basic considerations and constraints that might limit the feasibility to exploit reciprocity:

- If different basic transmission schemes are used in UL and DL (e.g., multiple access and resource allocation, bandwidth, Tx and Rx antenna configurations, spatial processing) it might not be straight forward to extract the required information for the forward link based on return channel measurements at Tx anyhow.
- Return channel measurements might not be possible at required times for non-scheduled users or users which have only activity in the forward link (e.g., downlink transmission only). In this case a permanent return link control channel (including pilots) would be required simply to allow the measurements at the transmitter and this must be taken into account as overhead induced by the exploitation of reciprocity
- The actual interference condition at the receiver can never be obtained by return link measurements, i.e., exploitation of reciprocity, therefore optimum MIMO processing is not possible and the receiver need to account for the remaining interference, e.g., in case of precoding. The other alternative: using return link measurements and feedback of interference values is not a promising option, since the feedback of the interference situation would consume almost the same signalling bandwidth than feedback of the complete CSI.
- To trade off between exploitation of reciprocity and feedback investigations are required regarding the required quantisation (number of bits) and the control information update rate in time, and frequency.

8.2.4 Conclusion

We have investigated the channel reciprocity assumption in the context of the TDD short-range scenario, using the model of the RF impairments that has been introduced. For the TxAA scheme, it has been shown that a phase calibration is required (when considering one receive antenna); with no phase deviation, the performance obtained is very close to the one with a perfect feedback link. For the TSD scheme, it is reasonable to assume the channel to be reciprocal; however a mechanism needs to be introduced in order to ensure that no errors are made when detecting which antenna was used at the transmitter side for each subcarrier. The impact of the CSI-delay with mobility conditions has also been investigated with short packets. While the performance of the TxAA scheme collapses for CSI-delays higher than 1ms, the TSD scheme maintains a good level of performance (never worse than the single antenna case) with decoding complexity similar to single antenna systems.

This study of the exploitation of the reciprocity property in a TDD system gives us some information on how spatial processing techniques relying on short-term CSI at the Tx can be impacted by the RF impairments. These techniques are promising for the short-range scenario, and in this context, calibration techniques should be investigated to evaluate their impact on the performance of the system. In particular an analysis of the overhead due to a calibration process needs to be conducted. Further aspects to consider is the impact of basic differences in the transmission scheme in uplink and downlink, the impact of inhomogeneous user traffic in up- and downlink, and the fact that the actual interference condition at the

receiver can only be obtained by feedback and not by exploitation of reciprocity.

8.3 Baseband Complexity

The results presented in this section constitute only a first rough estimation of the baseband complexity of different transmit and receive strategies and is based on the relative energy and cycle count required for one run of the algorithm. The analysis is only meant to give a first overview of the methodology and issues related to complexity assessments, and complexity performance tradeoffs (see also [METD5.1], [IME03]). It is by no means complete and further in depth studies should be done in upcoming deliverables such as D2.9.

8.3.1 Methodology

Since the general methodology for complexity assessment is already described in[D2.3], only the outline of the argumentation will be presented in this report. For details, please refer to[D2.3]. Different cost functions have to be considered when assessing the baseband complexity of an algorithm, when implemented in silicon, namely the chip area (cost) as well as the energy and number of cycles required for execution of a single operation (hence, power consumption and delay). These costs are often summarised as the well-known "ATE product" (area requirements-time needed for an operation-energy consumption), which is the target of the optimisation in the implementation process. The chip can usually be designed to meet the specific requirements, e.g., the achievable throughput, at the expense of higher chip area and/or power consumption. The focus in our assessment is on power consumption and delay only. Memory requirements should be assessed in future versions of this work.

It is reasonable to classify the operations required for a run of a transceiver algorithm according to their complexity as follows (for further detail please refer to[D2.3]: Simple arithmetic operations (add, abs, shift, max, etc), multiplications, divisions, square roots, and non-linear functions in general. The cost related to these functions, in terms of energy and cycle count are shown in Table 8.3 [JDL+04][D2.3].

Table 8.3: Relative energy and processing time costs

	Simple Operation	Multiplication	Division	Square root	Non-linear function
Relative energy cost factor	1	10	40	50	60
Relative cycle count	1	1	4	5	6

The figures given in the above table are the basis for the complexity assessments given in the subsequent parts of this section. The cost in terms of energy and cycle count will be equal whenever we require only simple arithmetic operations for the execution of the algorithm. Processors can usually be designed to have the same cycle delay for addition and multiplication, at the expense of large area and energy consumption. Divisions and other non-linear functions are implemented via iterative interpolations, explaining the higher cost for these operations. Energy consumption is mostly of interest for the user terminal, since power consumption is usually not a limiting factor at the access point.

The complexity of any algorithm can be decomposed into two parts: preprocessing and precoding/detection, where the former has to be performed only when the channel changes while the latter has to be performed for each transmitted/received (vector) symbol. This enables assessing the trade-off between algorithms that require large preprocessing and require only few operations per transmitted bit (more suited for slow fading regime) and algorithms that require almost no preprocessing but have a high complexity per transmitted bit (more suited for fast fading regime). It appears to be reasonable to assess the complexity of all algorithms relative to the simplest possible precoding/detection strategy, e.g., linear precoding/detection.

8.3.2 Basic Signal Processing Building Blocks

In the following, it is assumed that all algorithms work on complex values (where possible), the number of operations is then appropriately calculated in terms of operations on real numbers (e.g., one complex ADD equals two real ADDs, one complex MUL is equivalent to four real MULs and two real ADDs,

etc.). Similar results are obtained when replacing the complex linear equation system representing the MIMO channel with a real equation system of double dimensions. Table 8.4 summarises the complexity of some basic signal processing operations that are recurring in many of the considered algorithms and can hence be seen as building blocks of these. Here, M_i and N_i correspond to the number of rows and columns of matrix i, respectively. The modified QR decomposition normalises the diagonal entries in R to one. Q is therefore no longer a unitary matrix, since its basis vectors are no longer scaled to length 1, but this approach enables omitting the square root operation during the matrix decomposition and also enables to omit one multiplication per layer for the nulling and cancelling step in the successive interference cancellation/THP precoding.

Table 8.4: Complexity of basic signal processing operations

	Matrix/ vector sum	Modified QR decomposition (square matrix)	Matrix-matrix/matrix vector multiplication	Nulling and Cancelling
Real ADD	$2MN$	$4M^3 - 2M^2$	$2M_1N_2(2N_1 - 1)$	$2(M^2 - M)$
Real MUL	-	$4M^3 + M^2 - M$	$4M_1N_1N_2$	$2(M^2 - M)$
Real DIV	-	M	-	
Total cycle count	$2MN$	$8M^3 - M^2 + 3M$	$2M_1N_2(4N_1 - 1)$	$4(M^2 - M)$
Total energy consumption	$2MN$	$44M^3 + 8M^2 + 30M$	$2M_1N_2(22N_1 - 1)$	$22(M^2 - M)$
4x4 System				
Total cycle count	32	508	120	48
Total energy consumption	32	3064	696	264

The overhead for calculation of an inverse matrix from an available QR decomposition of this matrix is roughly 30-50%, for a 4x4 system. A sorted QR decomposition can be calculated at negligible overhead, compared to a conventional QR decomposition. Extension of the original channel matrix by an identity matrix weighted by the noise variance is one method to achieve MMSE detection. The complexity of decomposing the extended channel matrix is correspondingly higher. Calculating the soft output from the channel detector usually requires only the subtraction of the Euclidean distances of the two constellation points having a 0 or 1 at the considered bit position and being closest to the received signal vector (maxMAP) – for Gray mapping the effort is hence around two real operations per received bit.

8.3.3 Basic Classes of Transmit/Receive Algorithms

Since a corresponding precoding technique exists for almost all receive strategies, the computational complexity required for signal processing at the transmitter and receiver is quite comparable, for the considered class of algorithms. The complexity assessment will therefore only distinguish between different classes of processing strategies, not where they are actually performed (Table 8.5).

Many detection algorithms require only a matrix-vector multiplication plus soft output calculation for detection. Lattice aided algorithms require an additional vector-matrix multiplication to transform the signal back into the original vector space (from the lattice reduced space). Meaningful figures for the complexity of sphere detectors/precoders are hard to establish, since they depend on a large number of assumptions (enumeration strategy, channel preprocessing, required number of candidates, etc.).

Taking the figures in the tables together allows for making the following statements, for a 4x4 MIMO system:

- Preprocessing for linear algorithms based on inverting the channel matrix is roughly 50% more complex than for SIC (QR/QL decomposition) based receiver structures, and about 6 times as complex as the actual detection.

- Detection for SIC (QR/QL decomposition) based receiver structures is 50% more complex than linear detection. The ratio of preprocessing to detection complexity is about 4:1.
- Signal processing for MIMO systems usually requires a large number of multiplications that are hard to avoid. Relative energy consumption is therefore roughly 6 times the cycle count, under the taken assumptions on the complexity of different operations.
- Detection/precoding complexity for lattice aided systems is roughly double that of the corresponding standard algorithm (linear, or SIC/THP based).
- Preprocessing complexity for lattice aided systems and detection complexity for sphere detectors are both random events and therefore hard to characterise. The *average* number of column exchanges in the LLL algorithm can be kept quite low, if appropriate preprocessing is used (namely, a MMSE-SQRD) [WBK04]. Similarly, the size of the search tree for sphere detectors based on a fixed search radius can be reduced substantially when using MMSE preprocessing and layer ordering [ZRF04]. Assessing the complexity of such approaches in more detail should be the focus of upcoming deliverables.

Table 8.5: Complexity of different processing strategies

	Linear precoding/ detection	**THP/Successive Interference Cancellation**	**Lattice reduction**	**Sphere Precoding/Detection**
Preprocessing				
Matrix inverse	1			
(Sorted) QR/QL decomposition		1	1	1
LLL algorithm			1	
Detection				
Vector-matrix product	1	1	2	1
Nulling and cancelling		1	Optional	Search Tree
Slicing/ LLR calculation	1	1	2	Number of candidates

Dividing the number of operations required for linear detection of a single vector symbol by the number of information bits per received vector symbol and adding 2 times the inverse of the code rate for soft output detection hence gives a rough estimate of the detection effort *per received information bit*. It is easily seen that, e.g., for a 4x4 system running rate 1/2 coded 64-QAM transmission, the total cycle count for detection, per information bit, would be only 10+4 = 14 operations. Compared to the effort for decoding, which ranges from several hundreds to several thousands of operations per received bit, this effort is negligible. Similar statements hold for the detection/precoding complexity of THP/SIC receivers, and lattice aided schemes. For sphere detection, the detection complexity becomes a significant issue when aiming for very high performance (hence requiring a large number of candidates), since detection complexity scales with the number of branches/leaves in the search tree.

The complexity required for preprocessing should be of minor interests in environments with low mobility, when the burst size is very long (e.g., in excess of 10 OFDM symbols). However, things look different in all cases where preprocessing has to be done very often, i.e., the channel is rapidly changing in frequency and/or time, burst transmission with short frame lengths is the target, or a user receiving data is allocated chunks that are widely separated in frequency and/or time. In these cases, preprocessing makes up a significant portion of the total complexity.

8.3.4 Receiver Performance-Complexity Trade-Offs

The main (if not only) justification for an increase in receiver complexity is higher performance, e.g., in terms of lower SNR required to achieve the same block error rate. In the following, the relative merits of different receiver algorithms as well as their disadvantages will be shortly discussed. While the statements made hold for a large number of environments, the relative merits of different techniques are highly

dependent on a number of input parameters, such as channel conditions (available diversity, antenna correlation), employed forward error correction technique, and modulation alphabet. Several well-known techniques have already been described in [D2.1], so we will not in detail describe them here. Readers interested in the foundations and detailed descriptions of the discussed algorithms are referred to these documents and the references cited therein. A more in-depth study of the performance of different receiver architectures should be part of future work.

The use of some sort of forward error correction scheme is mandatory for achieving reliable transmission. In order for the decoder to work effectively, it needs to be provided with "soft input" from the detector: information on how reliably a specific bit has been detected. The L-value or "soft output" from the detector is a common measure and denotes the probability of a certain bit being zero or one, conditioned on the channel output. Calculating correct soft outputs is a critical point for many of the receiver architectures considered in literature and the above complexity assessment. An incorrect calculation of the detector soft output can in some scenarios offset any benefit from reduced SNR requirements or diversity gain achieved by the considered technique in uncoded transmission. This is especially true whenever a large amount of diversity is available within one codeword: the operating region is then at such low SNR that even in uncoded transmission, benefits from using the specific technique are quite low (remember that under hard output detection, the decoder essentially faces a binary symmetric channel, for which the Shannon bound for rate 1/2 coding is at a BER of roughly 15%, and advanced FEC schemes have their waterfall region around 12% [RSU01]).

In view of this fact, and assuming that capacity-approaching codes such as PCCC and LDPCC are envisioned for the use in WINNER [D2.3], we make the following statements:

- **Linear detection:** From an uncoded perspective, this receiver architecture is highly inefficient, since it leads to strong noise enhancement whenever the channel matrix is near-singular, thus spoiling detection performance. On the other hand, the multiplication with the weighting matrix (channel inverse) completely cancels out the interference between the data transmitted from different antennas, i.e., the system is fully orthogonalised (at the expense of noise enhancement, of course). Calculating correct soft output is hence very easy, since the noise enhancement can be easily calculated via the row norms of the weighting matrix. Linear detection obtains remarkably good performance in uncorrelated channels with a large amount of diversity in one codeword.
- **Successive Interference Cancellation** (V-BLAST [WFG+98] and SQRD [WRB+02] based algorithms): One main disadvantage of any scheme based on such nulling and cancelling approaches is its susceptibility to detection errors made during the detection of first layers. There exist several approaches to mitigate these problems: appropriate layer ordering [WRB+02], per-antenna forward error correction (as in the traditional V-BLAST [WFG+98]) and taking into account error propagation in the calculation of the soft outputs [CCC00]. However, all of these approaches have their weaknesses: layer ordering achieves good performance when considering uncoded transmission and small modulation alphabets. Gains are substantially lower for higher modulation alphabets and lower SNR (the operating regime when coding over a large amount of fades, as explained above). Another problem is related to the combination of layer ordering and per-antenna FEC: since in a frequency selective OFDM system, the optimal layer ordering may vary from subcarrier to subcarrier, either an average layer ordering has to be assumed, or SIC detection has to be done with interference uncancelled on some subcarriers. The resulting performance is, however, not satisfactory. Introducing detector-decoder iterations alleviates the problem, at substantially increased complexity. The approach proposed in [CCC00] solves the soft output calculation problem for SIC based receivers very efficiently, yet at quite high complexity, since it requires a soft mapping based on the L-values from precedingly detected streams as well as the calculation of the power of the residual noise due to decision errors.
- **Lattice Reduction (LR) Aided Detection** [WLF03]: The benefit from using such techniques is mainly the realization of the full diversity order (equal to that of the ML detector). However, gains are relatively low, whenever the system operates at a quite low SNR. Another main problem is that quantization to the signal lattice is executed in a transformed vector space, where decision boundaries required for calculating detector soft output are unavailable. As a result, performance of lattice-aided detectors is quite low in environments where FEC is done over a

large amount of frequency and/or time diversity. Overcoming this limitation constitutes a main challenge for LR-based receiver architectures.

- **Sphere Detection** [HtB03][DGC03][VH03]: It has been shown in a number of publications that sphere detectors are capable of finding the ML solution to the MIMO detection problem at quite low complexity. Using appropriate preprocessing techniques for schemes based on a fixed search radius [ZRF04] and/or enumeration strategies (namely, Schnorr-Euchner [VH03][DGC03]) results in an *average* detection complexity for finding only the ML point that is significantly below 10 times that of SIC detection, in a large regime of interest. However, in order to produce good soft output, the number of candidates required is usually very high [HtB03] and since the complexity of sphere detection essentially scales with the number of required candidates, the complexity of such a detector scheme is still to be considered prohibitive when compared with simple linear equalization. Approaches that combine sphere detection with ideas from sequential decoding [BHW03] appear to be a promising path for further investigations.

Another approach to achieve very good performance in MIMO detection is using Turbo equalization (i.e., iteratively exchange extrinsic information between detector and decoder). It should, however, be kept in mind that the channel decoder still contributes substantially to the total baseband complexity and that using such iterative equalization techniques thus substantially increases complexity. It is concluded that assessing the performance-complexity trade-off in MIMO detection is a challenging topic, due to the multitude of receiver algorithms invented in recent years. Future work should analyze in detail, which of these techniques achieves the best detection performance, given a certain maximum allowable complexity, the specific environment conditions and the constraints set by the WINNER system design.

8.4 RF Complexity

For D2.7 the HPA, LNA, duplex filter, synthesiser and channel selection filter requirements are initially considered to be similar for all spatial processing techniques investigated. It is assumed that the dominant difference in RF complexity arises from the number of antennas, the number of separate RF chains, and whether the signal paths of the different RF chains need to be calibrated for phase coherency. One beneficial side effect from using multiple antennas is that the total transmitted power is distributed over several RF chains, relaxing the requirements in terms of peak transmitted power for the high power amplifiers, thus enabling the use of cheaper HPA. As “Dirty RF” effects, i.e., RF non-linearities (such as phase noise, I/Q imbalance, etc.), can be expected to worsen as systems are designed for higher carrier frequencies, lower supply voltages and higher SNR regions, such effects should be taken into account when designing multi-antenna systems.

8.5 Conclusion

In this Chapter, we have investigated the impact of spatial processing on system and terminal complexity. Spatial processing relies on control information for different purposes, and in particular to get some inputs on the channel state information. This information can be provided by specific pilots: an overview of the type of pilots required for different spatial processing techniques is thus presented, but can also be obtained by exploiting the reciprocity of the propagation channel, for instance in a short-range scenario. In this context, we propose a model for the RF imperfections that is used to assess their impact on different closed-loop techniques. From the simulation results it is mainly observed that a phase calibration is necessary to get the expected level of performance when the phase knowledge is required.

Finally a first analysis of the baseband complexity of different transmit and receive techniques is given. This analysis compares the levels of complexity they require, and in particular the performance-complexity trade-offs of various types of receiver. For the baseline WINNER multi-antenna concept proposal the baseband complexity (assessed in terms of chip area, number of cycles, and energy consumption) involved in detection is negligible compared to other building blocks, like the decoding algorithm. However, for space-time turbo coded modulation, where initial results have also been presented, or sphere decoding complexity should be investigated in more detail. The required preprocessing power, however, depends very much on coherence time and bandwidth. In the extreme case, e.g. of a singular value decomposition per subcarrier and per OFDM symbol it is a critical part of the overall complexity.

9. Proposals for the WINNER Multi-Antenna Concept

The work in this document has been guided by several sources of input, in particular the already existing results and documents of the WINNER project (e.g. [D2.1], [D2.4], [D2.3], [D7.1], [D7.2]). Out of the vast number of aspects to be considered the major requirements are briefly listed below.

The WINNER overall system requirements clearly state that a single ubiquitous radio access system concept shall be developed that is able to adapt to a comprehensive range of mobile communication scenarios with scalability in complexity. Under the constraint of low deployment efforts and cost, it aims at complete coverage (including rural areas with low population density), while at the same time provide significant performance enhancement compared to legacy systems and their evolutions [D7.1]. Spatial processing has the potential to contribute significantly to the following specific requirements that follow from the overall goal:

- Improved spectral efficiency and increased user peak data rate,
- Increased range or coverage in a cost-efficient manner,
- Enhanced interference management,
- Adaptivity to scenario and channel conditions,
- Concurrent support of adaptive and non-adaptive transmissions,
- Support of different terminal types (including single-antenna terminals),
- Reduced terminal power consumption to increase talk and stand-by time,
- Management and reduction of human exposure to electromagnetic fields.

Based on these guidelines and the investigation detailed before, a first proposal for the WINNER multi-antenna concept is developed in the sequel. First the individual techniques are compared with respect to the gains they leverage and their suitability for different scenarios. Then we distinguish incompatible and competing techniques from those that can be generalised into *umbrella techniques*. The provision of generalised techniques is considered to be of major importance to ensure that the WINNER multi-antenna concept can adapt to widely varying scenarios and support future WINNER system enhancements. In each generalisation step we consider the implications on CSI requirements, forward control overhead, transmitter and receiver complexity, as well as on complexity of radio resource management, interference management, and mobility support. This facilitates estimation of which generalisations are favourable, and also provides important guidance for required investigations in future.

Preferred spatial modes are identified for different scenarios, which highlights the need to support flexible combinations of all spatial processing gains and the need to support techniques with varying amount of channel knowledge at the transmitter. Therefore, the identified *umbrella techniques* are then prioritised according to their scalability regarding the CSI requirements and supported mobile terminal classes. The *WINNER multi-antenna concept* proposal is subsequently formed out of these *umbrella techniques*. Finally, a generic transmitter block diagram that allows implementation of the *WINNER multi-antenna concept* proposals is presented and described.

9.1 Generalisation of Techniques and Impact on Required Enablers

Based on review work and investigations detailed in this document, the multi-antenna techniques are sorted according to the spatial processing gains they leverage, like spatial diversity (DIV), spatial multiplexing/SDMA (SMUX), and beamforming (BF), as depicted in Figure 9.1. The available channel knowledge at the transmitter strongly depends on the terminal velocities, which will have different distributions in different scenarios. The spatial channel properties are also scenario-dependent. Therefore the most promising spatial processing techniques will be different in different scenarios and the vertical placement in Figure 9.1 maps the different techniques to scenarios. This is an indication rather than an exact mapping, since the suitability of techniques for different scenarios is a gradual process, just like the actual conditions relevant for spatial processing may vary significantly in each scenario. Arrows in Figure 9.1 indicate a generalisation process or a combination of different techniques. Boxes represent termination points of generalisation processes or major intermediate steps, called *umbrella techniques*.

To ensure a versatile and future-proof WINNER multi-antenna concept we strive for a high degree of flexibility and generalisation. Figure 9.1 identifies such major generalisation steps, which all need to be

assessed further by trading off the additional performance gain with the increased complexity and requirements regarding support functions and enablers. If a reasonable degree of generalisation has been achieved on each *generalisation tree* the remaining techniques form either *complementary techniques* (if they map to different scenarios, have different requirements regarding channel knowledge at transmitter, and/or foster different spatial processing gains) or *competing techniques* (if they are cover similar scenarios, have similar system requirements, and foster identical spatial processing gains), as explained in Section 2.4.

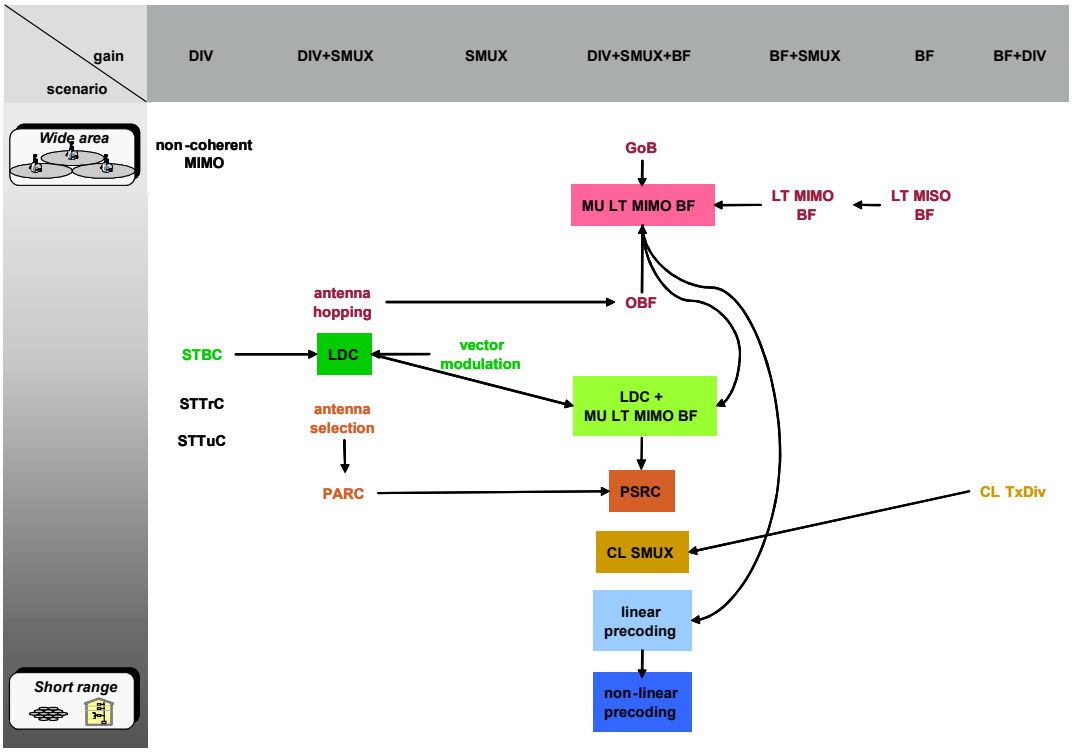


Figure 9.1: Generalisation and mapping of techniques to scenarios

In the following, we discuss each generalisation process in Figure 9.1 and list the corresponding impacts on CSI requirements (including all forms of CSI, like short-term, long-term, and CQI), forward control information, Tx and Rx complexity (including measurements), as well as on complexity of mobility support, interference and radio resource management in Table 9.1. In this table ✓ indicates no significant increase in complexity, whereas ↑ denotes complexity or overhead increase. It is assumed that long-term CSI can be obtained from return link measurements. Therefore techniques based on long-term CSI have no increase with respect to open-loop techniques in terms of CSI feedback but in Tx complexity due to the measurements.

As shown in Figure 9.1, noncoherent spatial processing is an alternative for situations where no CSI can be obtained even at the receiver. Open-loop space-time coding techniques, like space-time block codes (STBC), space-time trellis codes (STTrC), and space-time turbo codes (STTuC) are competing techniques without CSI knowledge at the transmitter, and therefore well-suited in situations where either such channel knowledge cannot be obtained (as a fallback mode) or where the application of closed-loop schemes is not recommended (e.g., very short packet calls, or common channels that require high robustness). The major distinguishing factor between them is receiver complexity (versus performance), their suitability and robustness for code rate adaptation, and their flexibility to be enhanced by additional spatial processing gains, like a SMUX component. STBC have lowest receiver complexity. STBC in this document refers to orthogonal space-time block coding. Linear matrix modulation or LDC can be seen as a generalisation of STBC to include also coding in frequency domain (SFBC, space frequency block codes) and non-orthogonal codes, which include vector modulation. This generalisation step basically increases Rx complexity due to LDC decoding (see Table 9.1).

Competing techniques that include DIV and SMUX are LDC, per antenna rate control (PARC), as well as antenna selection and hopping techniques. The primary discriminating factors between them are the signalling requirements, flexibility and receiver complexity, which need to be considered when detailed performance comparisons are done. All of them can be further generalised by including an additional directivity (beamforming) component. For example, antenna hopping is a special case of opportunistic beamforming (OBF, see Section 3.2.4), the additional complexity due to directivity is minor. OBF itself can be expanded to multi-user long-term MIMO beamforming (MU LT MIMO BF), just like the beamforming techniques based on a grid of fixed beams (GoB, see Section 3.2.2). The performance and complexity increase of these generalisation steps largely depend on the exact implementation of the multi-user long-term MIMO beamforming (in particular the multi-user and scheduling aspect) and are for further study. Also it should be noted that a simple flavour of beamforming, like OBF, in general requires a more complex resource allocation algorithm, than, e.g., MU LT MIMO BF does.

Single-user long-term MISO beamforming techniques (LT MISO BF, i.e., beamforming and SDMA based on adaptive beams, see Section 3.2.3) can be generalised to single-user long-term MIMO beamforming (LT MIMO BF) by allowing, e.g., an additional SMUX component (e.g., by transmitting on more than one long-term eigenbeam). This increases slightly the forward control overhead as well as Tx and Rx complexity. Furthermore allowing parallel data streams per user adds degrees of freedom and thus complexity to the RRM. For the migration of LT MIMO BF to MU LT MIMO BF the same holds as in the corresponding step from OBF to MU LT MIMO BF: the complexity vs. performance trade-off largely depends on the exact implementation of this technique and requires further investigations. For example the multi-user aspect can be simply realised by a selection of users (scheduling) and subsequent single-user LT MIMO BF, but also more sophisticated joint optimisation of the RRM and spatial filtering are conceivable. The block named LDC + MU LT MIMO BF is a combination of both techniques including spatial transmission modes based on directivity and LDC.

A migration path from MU LT MIMO BF to linear precoding exists if short-term CSI is available, further generalisation then includes also non-linear precoding techniques. Due to the exploitation of short-term CSI these are considered major steps in complexity and overhead, as well as a major impact on the required enabling functionalities. Therefore each technique forms a major intermediate step and in particular the linear and non-linear precoding techniques have partly *competing* character. Further studies are required here.

Antenna selection can be seen as a variant of per antenna rate control (PARC, i.e., limited CSI spatial multiplexing, see Section 3.2.5) with coarse on/off rate control. PARC of course has higher Tx and Rx complexity. PARC may in turn be generalised to per stream rate control (PSRC), which means that instead of transmitting each stream directly from a physical antenna, it is spread with a LDC and a beamforming component so that it is transmitted from all physical antennas. Alternatively, PSRC may be seen as a generalisation of LDC + MU LT MIMO BF. As compared to LDC + MU LT MIMO BF, this means that the data flow intended for a certain receiver may be split into several sub flows that are separately channel encoded. This enables the use of successive interference cancellation after channel decoding, which potentially may trade signalling overhead for a reduction of receiver complexity. Naturally, compared to PARC, PSRC has higher Tx and Rx complexity, whereas it at this point is not obvious that PSRC has higher Rx complexity than an advanced non-linear demodulator for an LDC. Finally, it may be noted that PSRC in the present generalisation mainly affects the channel coding and associated functions such as link adaptation. PSRC collapses to LDC+MU MIMO LT BF for the case that a user data flow is directly mapped to a single sub flow, which in turn is encoded with a single code. One may also consider the use of short-term CSI in combination with PSRC. The step from closed-loop transmit diversity (CL TxDiv) to closed-loop spatial multiplexing (CL SMUX) adds a spatial multiplexing component, which increases forward and return link control overhead as well as Tx and Rx complexity. As in the MIMO BF case, multiple streams per user add to the complexity of the combinatorial problem to solve in the RRM.

Table 9.1: Impact of Generalisation Steps

Generalisation	Benefit	overhead		complexity				
		CSI feedback	forward control	Tx	Rx	mobility support	interference management	RRM
STBC → LDC	including SMUX	✓	✓	✓	↑	✓	✓	✓
vector modulation → LDC	including DIV	✓	✓	✓	✓	✓	✓	✓
antenna hopping → OBF	including BF	✓	✓	↑	✓	✓	✓	✓
OBF/GoB → MU LT MIMO BF	smoother adaptation,	✓	✓	↑	✓	(↑)	↑	↑
LT MISO BF → LT MIMO BF	including SMUX	✓	↑	↑	↑	✓	✓	↑
LT MIMO BF → MU LT MIMO BF	higher spectral efficiency	✓	✓	↑	✓	(↑)	↑	↑
LDC → LDC + MU LT MIMO BF	including BF	✓	↑	↑	↑	↑	↑	↑
MU LT MIMO BF → LDC + MU LT MIMO BF	higher flexibility	✓	↑	↑	↑	✓	✓	✓
MU LT BF → linear precoding	higher spectral efficiency	↑	✓	↑	↑	↑	↑	↑
linear precoding → non-linear precoding	higher spectral efficiency	✓	↑	↑	↑	✓	✓	✓
antenna/beam selection → PARC/PSRC	including SMUX	✓	✓	↑	↑	✓	↑	✓
PARC → PSRC	including BF and LDC	✓	✓	↑	↑	✓	✓	✓
LDC+ MU LT MIMO BF→ PSRC	including per stream rate control	✓	↑	↑	(↓)	✓	✓	✓
CL TxDiv → CL SMUX	including SMUX	↑	↑	↑	↑	✓	✓	↑

When interpreting the previous figure and table it must be kept in mind that the complexity and overhead increase may be very different for individual generalisation steps and the actual amount depends on the specific implementation details of the corresponding techniques. Furthermore, a complexity increase in the mobile terminal is considered much more critical than in the access point and it is the goal to concentrate as much complexity as possible in the network elements. For final assessment further specific and detailed studies of the performance vs. complexity trade-off are required to determine the appropriate degree of generalisation required in the WINNER system.

9.2 Flexibility, Scalability and Spatial Modes

In the following the identified *umbrella techniques* are further compared and evaluated regarding the following criteria:

- Capability to foster different types of spatial processing gains (i.e., capability to adapt to a large range of usage scenarios and channel properties),
- Low overhead and complexity of the required enablers,
- Scalability with respect to the CSI requirements at the transmitter,
- Support of a large range of mobile terminal capabilities.

In general, not all spatial processing gains can be exploited to a large extent simultaneously, either due to limitations inherent in the scenario, the terminal capabilities, or due to the specific propagation conditions. Open-loop spatial diversity is important as a fallback mode, when no CSI can be obtained at the transmitter. Furthermore even low-cost mobile terminals can support simple orthogonal space-time coding. However, at this stage of the WINNER project it is not clear how many and to what extent other sources of diversity will be available (e.g., frequency, time, polarisation) and the need of additional spatial diversity is hard to estimate. Therefore an important area of future research is to understand which dimensions are favourably exploited by means of link adaptation, and which should use averaging techniques, like diversity. A major driver for this decision will be the feasibility and required overhead to implement adaptive techniques in the corresponding dimension.

While spatial multiplexing is an enabler for high peak user data rates and represents the capacity-achieving strategy for sparsely loaded sites (few users active), beamforming provides high spectral efficiency by exploiting SDMA and multi-user diversity in case of higher number of users (whose signals naturally exhibit higher decorrelation due to spatial separation than the different streams to one terminal in case of spatial multiplexing) [GVK02][Gen03]. Furthermore beamforming is a means to increase range or coverage and therefore provide cost-effective deployments in rural, sparsely populated areas. Also in the uplink a directivity component allows to save the scarce transmit power of mobile terminals. It is also a promising technique for relay links and supports active management of interference and human exposure to electromagnetic fields.

In Figure 9.2, preferred spatial modes are shown for different scenarios that are based on different types of links. For each scenario the spatial mode is shown as the combination of required channel knowledge at the transmitter and the major spatial processing gains to be exploited (cf. Section 2.4). It is understood that even within one scenario the propagation conditions, the system load, and the performance objectives might be largely different and consequently many different spatial modes might be most suited in particular cases. Nevertheless, we identified one preferred mode for a typical set of conditions in the following scenarios: wide-area, short-range, (multihop or traditional) relay links, peer-to-peer, and in an open-loop fallback mode based solely on spatial diversity.

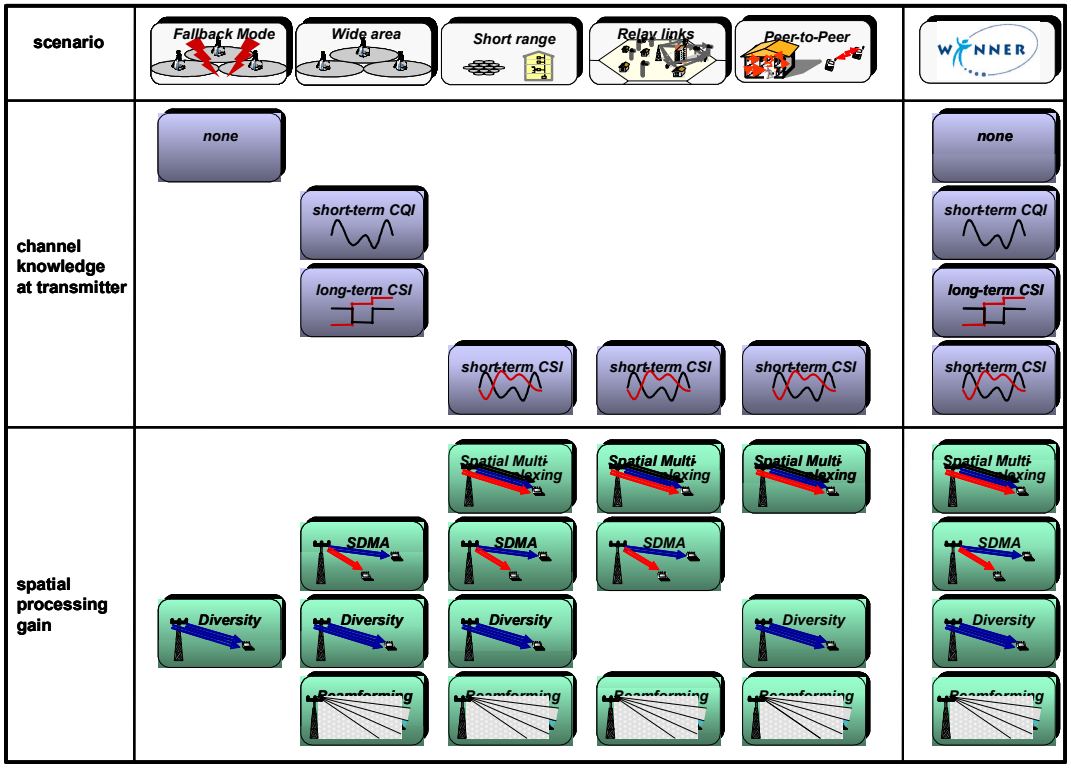


Figure 9.2: Preferred spatial modes for different scenarios and radio interface modes

For wide-area scenarios only long-term CSI seems to be reasonable for spatial processing in the majority of cases, most favourably combined with short-term CQI for link adaptation. In all other scenarios (expect the fallback mode) we consider short-term CSI, either due to reduced mobility or even fixed point-to-(multi)point connections. In these cases we assume that short-term CQI for link adaptation can be obtained by proper processing of the short-term CSI information and therefore it is not stated as additional requirement.

Apart from the fallback mode, beamforming (including interference suppression and precoding) is considered as integral part in all scenarios. A certain amount of spatial diversity is also beneficial in all scenarios, except for the fixed relay links, where due to the quasi-deterministic fixed link we should use the spatial degrees of freedom preferably for increasing spectral efficiency by means of spatial multiplexing and SDMA. As stated above a flexible combination of SDMA and spatial multiplexing is required, since the optimal spatial mode depends on system load. SDMA is an important enabler for high spectral efficiency and therefore included in wide-area, short-range, and relay links (for the point-to-point peer-to-peer link it is not important of course). Spatial multiplexing is another means to increase system performance and in particular user peak data rates. Therefore it should be enabled whenever propagation conditions allow it. Apart from the typically sparse scattering channels that might dominate in rural or wide-area scenarios, a spatial multiplexing component adds significant performance gains in all other scenarios. For the relay link scenario this is in particular true for multihop relay links, while the situation might be different for traditional (above roof-top) relay links. However, even for traditional relay links spatial multiplexing based on different polarisations can be considered.

The last column in Figure 9.2 contains the requirements for the entire WINNER multi-antenna concept and is formed as a superposition of the spatial modes of the individual scenarios. It is evident that the WINNER multi-antenna concept must work with varying degree of available channel knowledge at the transmitter and must be able to foster flexible combinations of spatial multiplexing, SDMA, spatial diversity, beamforming, and means for enhanced interference management. Further aspects need to be considered, e.g., the WINNER multi-antenna concept must support and enable various business opportunities. This means, e.g., that a wide range of different terminal capabilities from low-end single-antenna to high-end multi-antenna must be supported and high flexibility for future evolution of spatial processing must be ensured.

Therefore the major goal is to provide means to support and foster all spatial processing gains by finding the most flexible combination of techniques that requires minimum possible overhead (signalling, measurements, support functions, enablers, etc.) and complexity. Since in different scenarios (different user speeds, potentially different duplexing schemes, etc.) different amount of CSI will be available at the transmitter, techniques that can be adapted to different degrees of CSI while at the same time providing competitive performance will be privileged. In a similar way, scalability and applicability of one technique to different user terminal capabilities is a desired feature since in operational systems a mixture of different terminal classes will be encountered.

Table 9.2 presents an overview of the adaptivity and scalability of the *umbrella techniques*, sorted by their degree of suitability to the wide-area (WA) (upper rows) and to the short-range (SR) scenario (lower rows). Spatial processing based solely on LDC only make sense in scenarios, where no directivity is required, and therefore is of limited flexibility. Out of those umbrella techniques, which include SMUX, DIV, and BF components, CL SMUX has highest requirements in terms of CSI knowledge at the transmitter and does only support terminal classes with multiple antennas. Due to the involved processing complexity, the precoding techniques are not suited for low end terminals.

MU LT MIMO BF is a very flexible and scalable technique, which is however predominantly suited for wide-area coverage. In its combination with LDC (LDC + MU LT MIMO BF), and also by including precoding techniques based on short-term CSI (LDC + MU MIMO BF) it forms very promising umbrella techniques, since it provides support for the whole range of terminals from low-end single-antenna devices up to high-end multi-antenna user equipment. On top of these very promising umbrella techniques, PSRC can be implemented if each data flow to a certain user is multiplexed onto several independently rate controlled and channel coded subflows. Such combinations can be adapted to all kind of CSI knowledge at the transmitter, and are equally-suited for wide-area and short-range scenarios. Note that in different scenarios and in different propagation conditions not all components of LDC + MU

MIMO BF+PSRC might be activated, therefore our proposal is seen as a generalisation of the idea of switching between different spatial modes or techniques (or combinations thereof).

Table 9.2: Adaptivity and scalability of umbrella techniques

Umbrella Technique	**Scenarios**	**Gains**	**CSI**			**terminal classes**		
			long-term CSI	CQI	short-term CSI	low end	1 antenna	> 1 antenna
MU LT MIMO BF	**WA**	SMUX+DIV+ BF	✓	✓		✓	✓	✓
LDC	**WA**, SR	SMUX+DIV		✓		✓	✓	✓
LDC + MU LT MIMO BF	**WA**, SR	SMUX+DIV+ BF	✓	✓		✓	✓	✓
PSRC	WA, SR	SMUX+DIV+ BF	✓	✓		✓	✓	✓
LDC + MU MIMO BF	WA, SR	SMUX+DIV+ BF	✓	✓	✓	✓	✓	✓
CL SMUX	WA, **SR**	SMUX+DIV+ BF			✓			✓
linear precoding	WA, **SR**	SMUX+DIV+ BF	✓	✓	✓		✓	✓
non-linear precoding	**SR**	SMUX+DIV+ BF			✓		✓	✓

The baseline WINNER multi-antenna concept used for further investigation will therefore be a *multi-user spatial domain link adaptation concept* consisting of the following basic components:

- Linear dispersion codes,
- Directive transmission (beamforming),
- Multi-user precoding, and
- Per stream rate control.

To allow such a spatial mode switching return link control information is required to report measurements or information that allows the transmitter to select the best mode. Also in the forward link control information on the actually chosen spatial mode is required to allow detection and decoding of the data. The details of the actual set of spatial processing modes that need to be realised to achieve high performance in all major scenarios and propagation conditions is for further study. Also there are several further investigations required concerning the following aspects of such a combination, such as:

- Relevant parameter that trigger switching between spatial modes,
- Concept and implementation aspects to enable a smooth transition between spatial modes,
- Integration and implementation aspects of the spatial domain link adaptation with more conventional link adaptation such a rate control in combination with identified HARQ concepts,
- Trade-off between performance and receiver complexity for high-rate non-orthogonal matrix modulation compared to simpler space-time codes,
- Design of LDC under the constraint of virtual antenna streams with different average quality, which results from the combination with eigenbeamforming or precoding techniques,
- Optimum algorithms and strategies for solving the multi-user problem with different amount of CSI and scalable degree of complexity,
- Increasing directivity due to beamforming reduces the channel variances in time and space and therefore reduces the achievable diversity gain, i.e., a joint optimisation of beamforming and diversity gain needs to be performed,
- Degree of adaptivity (ranging from on/off, selection to different degrees of link adaptation) vs. performance trade-off, in particular for long-term vs. short-term spatial link adaptation,
- Challenges and opportunities in interference management and RRM due to directivity and SDMA,
- Complexity vs. performance and robustness for the migration from long-term single-user BF to the different precoding techniques,

- Feasibility of exploiting reciprocity in TDD systems,
- Spatial processing concept for the uplink and special scenarios (metropolitan multi-hop, relay, peer-to-peer).

Important conclusions at this point, however, are:

- The WINNER multi-antenna concept shall provide all enablers to support the umbrella technique LDC + MU MIMO BF+PSRC,
- Alternative concepts are much less scalable and flexible and will only be relevant for special scenarios or applications,
- Precoding techniques that will be further investigated for the wide-area scenario are restricted to linear precoding (e.g., SMMSE or BD based on long-term statistics) and provide performance improvements primarily for multi-antenna techniques,
- For the short-range scenario precoding techniques for further consideration include non-linear techniques and may also be attractive for single-antenna terminals, in particular SO THP, SMMSE THP and SMMSE in more realistic scenarios with a mixture of terminal classes,
- The robustness of precoding techniques and their performance under realistic operational conditions (in particular viable ways to obtain short-term CSI at the transmitter) require further studies.

9.3 Proposal for a Generic Multi-antenna Transmission Chain

Chapter 2 has already introduced a generic block diagram, which encompasses most of the techniques currently studied. Based on this we further develop an initial proposal for a spatial processing chain that is capable of supporting all multi-antenna techniques recommended for further study in the previous section. In particular it contains further details for the generic transmitter in Figure 2.10 based on the investigation in this document.

The proposed block structure has been developed according to the following main requirements:

- Concurrent support of adaptive and non-adaptive transmissions modes (i.e., transmissions that use channel-based link adaptation techniques and those that rely on diversity and spreading of resource elements),
- Support bit-loading per spatial layer of one chunk, i.e., varying number of coded bits,
- Support of SDMA on chunk-level, i.e., different users may transmit on the *virtual antenna streams* of one chunk.

These requirements have lead to a two-step processing of the data streams of the concurrently scheduled users based on *segmentation* into substreams (that allows varying number of coded bits per chunk layer) and subsequent *chunk mapping*, which is flexible enough to support adaptive and non-adaptive transmission modes with time-varying relative percentage of resource allocation in each mode. A detailed description of these blocks can be found in the following section. This processing structure will constitute the basis for the majority of further investigations in T2.5.

9.3.1 Channel Coding and Chunk Layer Multiplexing

Let us assume that the scheduler has L users for transmission. As a first step, the corresponding bit stream of user l, g_l, may first be multiplexed or segmented into K_l subflows or segments as depicted in Figure 9.3. Note, in particular also no multiplexing might occur, i.e., joint forward error coding of an entire user flow is contained in this proposal as a special case. In this way, a total of

$$K = \sum_{l=1}^{L} K_l \tag{9.1}$$

bit streams or subflows are obtained. Each subflow k then passes forward error coding (FEC_k) and interleaving Π_k[4]. Depending on the decisions of the resource management the bit stream of each subflow

[4] For the sake of this presentation we assume any means of rate matching to be included in the FEC, i.e., the encoding, segmentation, and chunk mapping will ensure that the correct number of coded bits is mapped to the physical channel resources (including any effect of link adaptation and spreading).

is split up in different segments, called *substreams* (Figure 9.4).

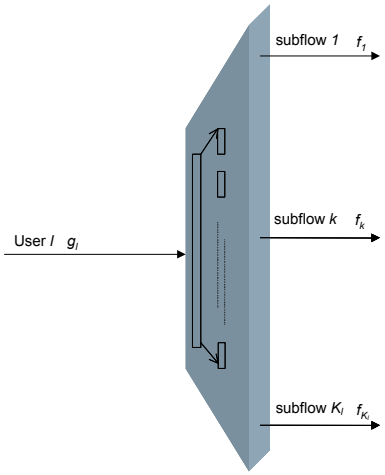


Figure 9.3: Segmentation or multiplexing of a user's data stream/flow into several subflows.

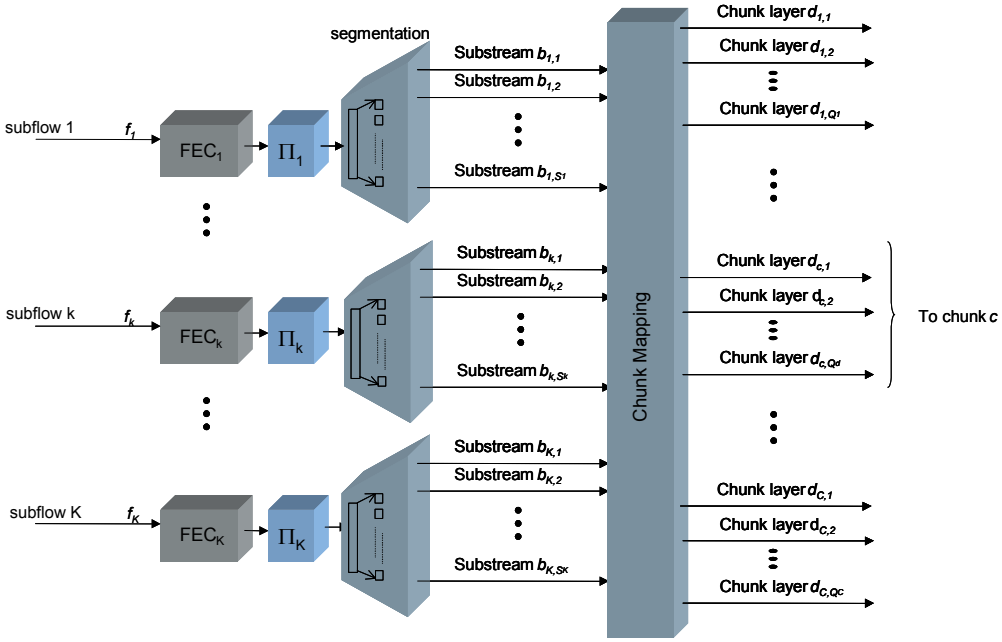


Figure 9.4: Channel coding, segmentation and mapping onto chunk layers

Each substream is then mapped onto a layer in a specific chunk depending on decisions of the resource management. This allows accommodating adaptive and non-adaptive transmissions in a flexible way. Substreams of different users can be multiplexed on arbitrary chunk layers (including SDMA per chunk, as well as FDMA and TDMA). The substreams of one user (sub-)flow might either occupy neighbouring resources in frequency and space (e.g., if link adaptation is employed) or be spread over frequency and space to exploit diversity.

A chunk covers n_{sub} subcarriers and, since it is assumed that the resource allocation operates on a frame basis, n_{frame} consecutive OFDM symbols. The (maximum) number of layers in a chunk c, denoted Q_c, can be different for different chunks. Let $b_{k,s}$, $k=1..K$, $s=1,..S_k$, denote substream k obtained by segmenting the channel encoded and interleaved stream of bits of user k. This stream is mapped to a layer of a chunk with two *chunk mapping functions*, $\gamma(k,s)$, and $\sigma(k,s)$, as

$$d_{c,q} = b_{\gamma(k,s),\sigma(k,s)}, \tag{9.2}$$

where $d_{c,q}$ is the stream of bits to be transmitted on layer q in chunk c, and the mapping functions are determined by the resource management functionality. Function $\gamma(k,s)$ selects the user (sub-)flow, whereas $\sigma(k,s)$ selects the substreams to be mapped to the corresponding chunk layer.

9.3.2 Chunk Processing

After the flows have been segmented and multiplexed into chunk layers, each data is spread in space, time and frequency as illustrated in Figure 9.5.

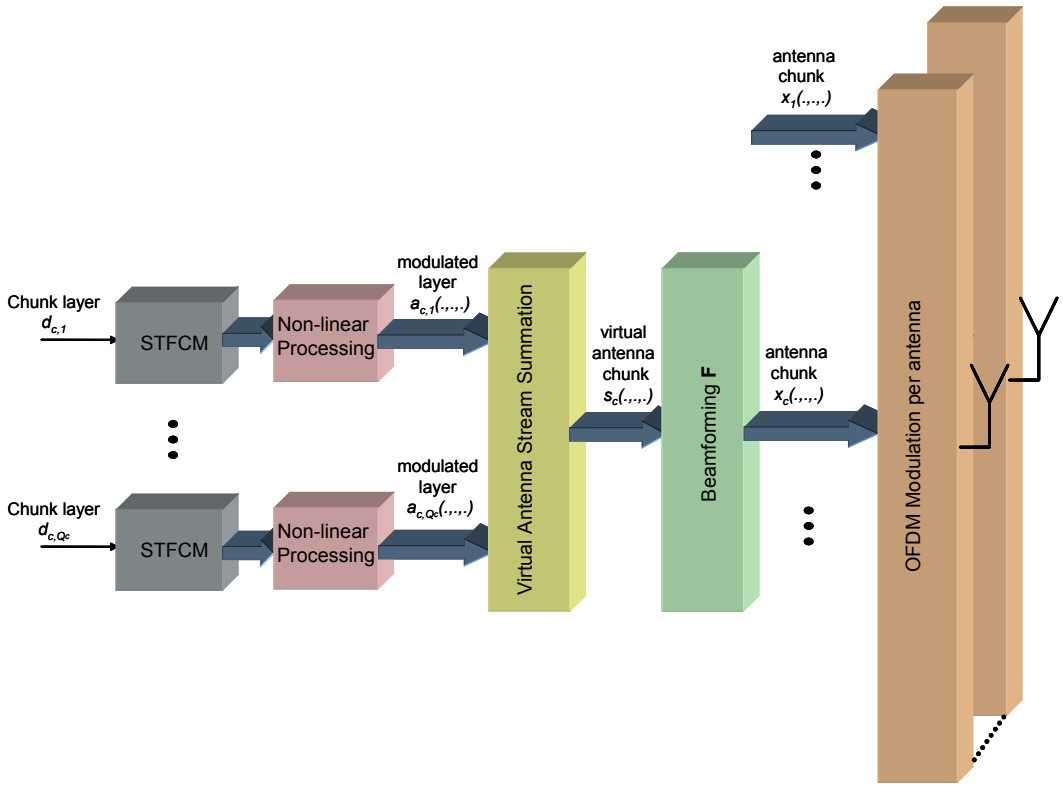


Figure 9.5: Spreading in space, time and frequency of layers allocated to a chunk

In the most general case, each chunk layer is then mapped onto a three-dimensional entity (*modulated layer*) that spans M_{va} virtual antennas in space, n_{frame} OFDM symbols in time, and n_{sub} subcarriers in frequency by means of space-time-frequency coding and modulation (STFCM). The STFCM block has already been detailed in Chapter 2 but is reproduced for simplicity in Figure 9.6. Each chunk layer $d(c,q)$ is subject to power and bit loading, i.e. modulation $M(c,q)$, and power assignment $P(c,q)$. The encoded, modulated and power assigned stream is then encoded with a linear dispersion code (LDC) and each resulting parallel LDC encoded stream is then written into the two-dimensional chunk associated with a virtual antenna, either frequency-by-frequency or symbol-by-symbol. LDC, which may be viewed as a straightforward generalisation of multi-code DS-CDMA, can be used to represent not only spreading, but also a large class of so-called vector and matrix modulation schemes. The generic transmitter chain actually does not contain FEC at chunk-layer level, since the number of bits of a chunk layer may be rather small, and since a combination of nested FEC and space-time-frequency coded modulation is not considered optimum and controllable.

Note here that the number of virtual antennas, M_{va}, can be arbitrary in relation to the number of layers in the chunk, Q_c, and the number of physical antennas, denoted M_T. For example, two chunk layers ($Q_c = 2$) might each undergo space-time-coding that maps these layers onto four virtual antenna streams ($M_{va} = 4$, e.g., like the well-known Alamouti-code would do). These four virtual antenna streams could be transmitted via beams generated by an 8-element antenna array, i.e., $M_T = 8$. The number M_{va} is chosen

depending on the transmit method. Also, the LDC of a particular chunk layer might encode over less than M_{va} streams. For simplicity of mathematical presentation, we will assume however zero-padding of the unused streams and represent any output of the STFCM block by M_{va} streams. Subsequently, non-linear processing is applied in the case of non-linear precoding techniques, typically in terms of feedback and a modulo operation per virtual antenna stream.

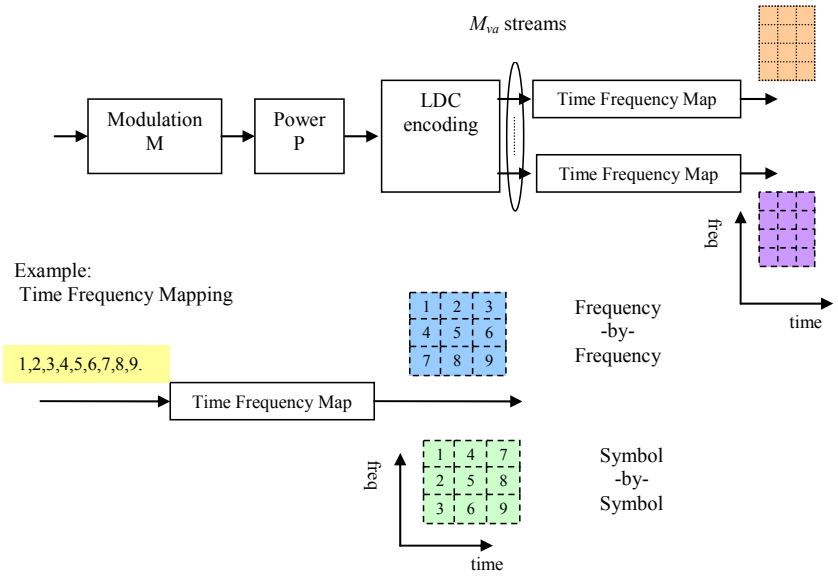


Figure 9.6: Baseline Space-Time-Frequency Coding and Modulation

The modulated layers of chunk c, denoted $a_{c,q}(.,.,.)$, are then combined with a summation in the block denoted Virtual Antenna Stream Summation (VASS) unit in Figure 9.5 (cf. also Figure 2.10). For simplicity and homogeneity, an appropriate choice of M_{va}, and zero-padding of the chunks generated by the STFCM are assumed so that the combination can be represented as a summation. The combined signal in chunk c, denoted $s_c(.,.,.)$, is then

$$s_c(n_{va}, n_t, n_f) = \sum_{q=1}^{Q_c} a_{c,q}\left(n_{va}, n_t, n_f\right), \tag{9.3}$$

where $n_{va}=1,..,M_{va}$, $n_t=1,..,n_{frame}$,$n_f=1,..,n_{sub}$. The number of virtual antenna streams is also expected to depend on the channel properties, the degree of multi-user optimisation, and the scheduling. Subsequently, beamforming is applied to produce the desired directivity. The beamforming block in Figure 9.5 will map the chunk $s_c(.,.,.)$ of size M_{va} *x* n_{frame} *x* n_{sub} onto a three-dimensional *antenna chunk* $x_c(.,.,.)$ of size M_T *x* n_{frame} *x* n_{sub}. Linear mappings are considered, so that one obtains:

$$x_c(n_a, n_t, n_f) = \sum_{m_{va}=1}^{M_{va}} \sum_{m_t=1}^{n_{frame}} \sum_{m_f=1}^{n_{sub}} f_c(n_a, n_t, n_f, m_{va}, m_t, m_f) s_c(m_{va}, m_t, m_f), \tag{9.4}$$

where $x_c(n_a,n_t,n_f)$ is the chip/symbol to be transmitted from transmit antenna $n_a=1,..M_T$, in OFDM symbol $n_t=1,..,n_{frame}$ in the frame on subcarrier $n_f=1,..n_{sub}$ in the set of subcarriers associated with the chunk c. Further $f_c(n_a,n_t,n_f, m_{va},m_t,m_f)$ denotes the beamforming coefficient for chunk c, i.e. the contribution of $s_c(m_{va},m_t,m_f)$ to $x_c(n_a,n_t,n_f)$. Note that in general the beamforming can be different for each subcarrier within one chunk. The beamforming function, which maps the combined and processed virtual antenna streams, typically uses some form of channel state information, such as an estimate of the instantaneous channel or the second order statistics. Beamforming in the present context not only covers techniques such as closed-loop transmit diversity and long-term beamforming, but also antenna or beam selection and hopping as well as random beamforming employed by opportunistic beamforming approaches.

10. Further Evolution of Investigations in WINNER Phase I

Although an initial multi-antenna concept has been presented earlier, there remains much to be studied in order to refine this further, not only in terms of multi-antenna techniques but also the necessary support functions. This chapter summarises the main topics on which attention will be focussed during the remainder of WINNER Phase I. A number of these future studies will require close cooperation with other task-teams within the overall project, particularly with respect to multiple access, MAC/RRM issues, and further development of the link-to-system interface. The description of future work has been split into two categories:

- Future investigations
- Evolution of assessment capability

These are described in subsequent sections. First a discussion follows of the 'critical parameters' on which future work will be focussed.

10.1 Critical Parameters

In the study of multi-antenna techniques for the WINNER system concept there are a number of topics likely to have major impact upon the conclusions that are drawn. These 'critical parameters' therefore deserve particularly thorough assessment and must be modelled with sufficient accuracy. The following topics are regarded as critical parameters and require detailed study in the on-going investigations:

- Propagation conditions – including: antenna correlation; delay spread; Doppler; spatial correlation between users
- Interference – including: link level modelling of inter-cell interference; impact of scheduling, traffic models, and cellular deployment upon interference; simplifying assumptions for interference modelling; interference management techniques
- Traffic modelling – including: flows of data to/from users, user mobility, and their distribution in space
- CSI reliability at the transmitter – including: the effects and modelling of CSI errors; required CSI signalling bandwidth
- CSI reliability at the receiver – including the impact of channel estimation errors upon both link level performance, and upon the link-to-system interface
- Hardware imperfections – including: synchronisation; fixed point implementation
- Resource and mobility management – including: scheduling; power control; adaptivity
- Receiver complexity
- Control overhead – including: signalling requirements; pilot transmissions

The simulations described earlier in this document include state of the art modelling of the propagation conditions, but there remains much scope for refinement in terms of traffic modelling, interference, and CSI reliability. Initial assessments of terminal complexity and control overhead have been provided in this deliverable and these can be improved as the system concept evolves. Additionally a full assessment of the effects of hardware imperfections is necessary. It can be seen in the sections that follow that much of the future work relates to the critical parameters listed above.

10.2 Future Investigations

Future investigations will be aimed at refining the multi-antenna concept previously described. Further, it is also crucial to ensure its integration into higher layers and the overall WINNER system concept as a whole. With these aims in mind, the key areas for further study are:

- Identify a baseline set of spatial modes and techniques used to implement them, investigate possibilities for a smooth transition between these modes, and identify a set of parameters that will allow the system to choose and adapt the spatial modes employed in a cell.

- Consider multiple access in time, frequency and space jointly, including assessment of the benefits of multi-user diversity (e.g., for overcoming capacity losses due to channel correlations), and its impact upon system set up.
- Investigate the impact of realistic channel estimation methods upon performance and detail the required pilots, support functions and other enablers required for effective deployment of the multi-antenna concept.
- Explore further the tradeoffs between performance enhancement from CSI availability at the transmitter and the increased signalling bandwidth required. Additionally, evaluate the potential gain of techniques using short-term CSI at the transmitter, and the feasibility of obtaining the short-term CSI (e.g., exploitation of reciprocity versus return channel signalling).
- Evaluate the impact of other 'practical implementation' issues, for example: assessment of the support required in the system concept for symbol and frequency synchronisation errors; appropriate design of CSI/CQI measurements; the impact of estimation and measurement errors in conjunction with hybrid ARQ; and handover between cells/beams.
- In addition to these refinements of the system concept it will also be beneficial to continue study of new and promising techniques such as superposition coding and "dirty-paper-like" coding.

10.3 Evolution of Assessment Capability

To enable the above investigations some advancement is required in the current simulation capability and methodology. More specifically:

- Inclusion of hybrid ARQ to study the consequences with respect to measurement errors and to enable delay related performance assessments
- More detailed interference and interference-control modelling, and more studies of multi-cell operation
- Realistic traffic modelling and scheduling algorithms
- Advancement of the link-to-system interface to include modelling of non-linear/iterative receivers and the interrelation of channel and space-time-frequency coding.

11. Summary and Conclusions

In order to assess advanced beamforming and MIMO technologies, *methodologies* for simulation and assessment have been developed and *simulative investigation*, as well as *conceptual work* has been conducted.

For a comprehensive assessment of multi-antenna techniques, it is mandatory to consider the performance at system-level, since many effects of spatial processing, like multi-user precoding, the impact of spatially-coloured interference, and the benefits of interference management techniques are not tractable on a link level. Therefore considerable efforts have been dedicated to the development and verification of simulation *methodologies*, in particular to the link-to-system level interface definition. Major challenges included the support of multi-antenna techniques in a multi-state OFDM system like WINNER, while keeping the complexity of the interface within feasible limits. The chosen link-to-system interface (see Chapter 4) allows predicting BLER in system-level simulations without performing modulation, interleaving and coding explicitly. It is based on mapping of the large number of different *SINR* values of the multi-state channel within one codeword to *one* effective *SINR* based on a specific averaging function in the mutual information domain. This one-dimensional compression function reduces complexity considerably, while at the same time maintaining high accuracy, as has been shown by extensive validation. This link-to-system level interface has been adopted as WINNER standard for WP2 and serves as a baseline for other WPs, which adapt it to their particular needs.

In order to ensure a certain degree of comparability across different simulation platforms, common simulation assumptions and parameters have been established for the scope of this document (Chapter 5) and will be adapted throughout the remainder of this project. Further methodological activities included the development of criteria for assessment of multi-antenna techniques. Due to constraints in the available traffic models and the simulation capabilities the assessment criteria proposed by WP7 have been adapted for single-site, snapshot system-level simulations without dedicated modelling of packets and data traffic (Chapter 6). Additionally a method to obtain at first estimate of control overhead in the absence of a defined control channel structure and other relevant system parameters has been established.

Based on this methodology the partners have started to develop or adapt simulators for investigation of individual multi-antenna techniques. Although at this time still constraints apply in the simulator capabilities, first *simulation results* and comparison of techniques are provided in Chapter 7. Evaluation of system and terminal complexity due to spatial processing are presented in Chapter 8.In contrast to most other publications on MIMO technologies, we include the impact of control overhead in our assessment and in the discussion of the figures of merit, like cell throughput, user throughput, coverage, robustness and spectral efficiency. Results are provided predominantly for the downlink of the wide-area and short-range scenario. Also initial results for MIMO relaying are provided.

Although the results obtained until now are preliminary and suffer from considerable constraints as detailed above, several conclusions have been drawn and are reflected in the current proposal for the WINNER multi-antenna concept. The most general and important findings are shortly stated in the sequel.

Initial comparisons between single-site and network system-level simulations indicate that a realistic modelling of inter-cell interference has dramatic impact on cell and user throughput, especially for techniques that include spatial multiplexing. In particular for low load (i.e., few users) single-site simulations tend to overestimate the performance considerably. Therefore the upgrade of the simulators to full network simulations is of highest priority for future work.

Directivity obtained by beamforming has high potential for scenarios with small angular spread, since it provides a good performance versus complexity trade-off. For a given number of users, cell throughput scales nearly linearly with the number of beams up to a certain maximum, which is given by the achievable beam width in combination with the required sidelobe suppression. When using a fixed grid of beam, it is important to carefully design the grid and the beam pattern (e.g., sidelobe suppression), since performance is sensitive to the inter-beam interference. Inter-beam interference must be reduced to a level that allows use of the highest PHY mode even with an additional margin for intercell interference. Means to reduce this sensitivity to inter-beam interference are strong channel coding and link adaptation. If

beamforming is combined with adaptive resource assignment based on short-term CQI, considerable performance enhancements can be obtained due to the exploitation of multi-user diversity, especially for high load and increasing angular spread in the propagation channel. It is interesting to note that although adaptive resource allocation reduces fairness, it has the potential to increase the absolute user throughput throughout the cell compared to, e.g., a simple Round Robin scheduling.

For the space-time-frequency coding and modulation, linear dispersion codes (matrix modulation) are considered as baseline for further investigations. Regarding their applicability to different channel conditions diagonal ABBA is more robust than DABBA or vector modulation. Linear precoding adds significant gain to pure matrix modulation (STBC). Spatial mode selection has proven substantial gain within each scenario, and considering the fact that the WINNER multi-antenna concept must support many different scenarios it is a must. Promising input parameter for the spatial mode selection are received SNR and the second condition number. Conceptually, these two parameters could also serve as a simplified link-to-system-level interface for investigations that require low computational complexity in the physical layer modelling.

One important focus of future research will be the quest for the optimum amount of adaptivity. At the example of limited CSI spatial multiplexing (PARC, PSRC), it is shown that the control overhead estimations range from around 30 kbps to 3.2 Mbps per user depending on the degree of link adaptivity and flexibility in channel dependent scheduling. This impressive range (a factor of one hundred!) clearly shows that after a certain breakpoint further adaptivity will merely increase complexity and overhead without providing corresponding performance gain. However, another associated trade-off is that higher efforts in link adaptation combined with simpler space-time-frequency codes might enable to concentrate the complexity at the network side compared to a reduced-adaptivity technique combined with spatial processing that involves higher decoding complexity.

Relaying and multi-hop communication is a concept with high potential. It gains additional performance if spatial processing is applied on the corresponding links. Even using the simple 2x1 Alamouti scheme higher throughput can be guaranteed over an extended area compared to SISO relaying. Spatial multiplexing combined with relaying also leads to improvements over SISO relaying in terms of cell coverage.

The multi-user multi-antenna signal design problem has been tackled both from a communication theory and a simulative perspective and the following conclusions are drawn: Multi-user precoding techniques that will be further investigated for the wide-area scenario are restricted to linear precoding (e.g., SMMSE or BD based on long-term statistics). Techniques like BD provide performance improvements primarily for multi-antenna mobile terminals. SMMSE is especially attractive since it does not suffer from restrictions considering the number of antennas in the system, provides good performance for various terminal classes, and can exploit either long-term or short-term CSI at the transmitter to perform the precoding. For the short-range scenario precoding techniques for further consideration include also non-linear techniques that rely only on the short-term CSI at the transmitter. SO THP (non-linear), SMMSE THP (non-linear) and SMMSE (linear) are of interest in realistic scenarios with a mixture of terminal classes. However, it is noted, that in particular for the techniques that rely on short-term CSI at the transmitter further studies regarding robustness and the feasibility of exploiting reciprocity in TDD systems are required. Further investigations are also required once a baseline WINNER channel code is available, since coding reduces the impairments due to channel estimation errors significantly.

For the baseline WINNER multi-antenna concept proposal the baseband complexity (assessed in terms of number of cycles, and energy consumption) involved in spatial processing for signal detection or precoding can be considered to be negligible compared to other building blocks, like the decoding algorithm, when only linear or SIC/THP based approaches are considered. However, more advanced schemes like sphere detection have a substantial complexity. Moreover, a substantial amount of processing is required for some of the preprocessing algorithms (matrix inversions and decompositions, etc.) and will be significant even in comparison to the decoding effort. Since such preprocessing is only required whenever changes in the channel conditions occur, the related effort highly dependent on the channel coherence time as well as, e.g., the burst length. Once detailed information on the physical channel allocation for the WINNER system is available, a more in depth study of this topic should be done. For space-time turbo coded modulation, where initial results have also been presented, complexity should be investigated in more detail. While in MIMO detection, substantial performance gains can be obtained by using detector-

decoder iterations, it should be kept in mind that using such iterative techniques requires a repeated execution of the decoding algorithm, which contributes substantially to the total baseband system complexity. The performance gains obtained should hence be weighted against the increase in complexity.

Also in our *conceptual activities*, considerable work has been dedicated to advance from pure spatial processing algorithms to techniques that include all overhead and supporting functions required to implement these algorithms in a radio access network. First a generic framework that allows a uniform description and mathematical model for most of the proposed techniques has been developed (Chapter 2). Details of individual techniques and their implementation in the generic block diagram of Chapter 2 are provided in Chapter 3. Based on this framework and the simulation results an initial proposal for the WINNER multi-antenna concept (Chapter 9) has been developed. The following major guidelines for creating this concept proposal were adopted:

- All spatial processing gains, i.e., spatial diversity, spatial multiplexing, beamforming and interference management by spatial processing must be supported to ensure high performance in all major scenarios, usage conditions (number of users, user distribution, user speeds, channel properties), and for all terminal classes.
- Techniques with scalability to different amount of CSI information at the transmitter are preferred.
- Complexity should be concentrated in the network and minimised in the mobile terminals.
- The processing must support concurrent transmission to different users using different spatial processing techniques and different degrees of adaptivity.
- The possibility for future evolution of spatial processing must be ensured.

To ensure a versatile and future-proof WINNER multi-antenna concept a high degree of flexibility and generalisation is required. Therefore generalisation paths of individual techniques have been identified and the associated impact on system design has been identified. Due to the superior characteristics with respect to scalability and flexibility, a combination of linear dispersion codes (LDC), beamforming and multi-user precoding is proposed as *baseline umbrella technique*, which allows implementation of various different spatial modes.

Based on this *baseline umbrella technique* a transmitter structure has been developed, which includes multiplexing, channel coding, chunk layer multiplexing, and spatial processing on chunk layer basis. The chunk layer multiplexing allows flexible bit stream distribution and implementation of adaptive transmission or non-adaptive transmission that relies on diversity and interleaving. In the spatial chunk layer processing, space-time-frequency coding and modulation is adopted based on power and bit loading, and LDCs. LDCs are considered as a general spatial processing option that allows a flexible trade-off between spatial diversity and multiplexing and includes several classes of space-time codes, like orthogonal STBC, vector modulation, etc. A beamforming block allows implementation of directional transmission and linear multi-user precoding. Non-linear precoding will also be further considered for short-range applications and therefore a corresponding block is included in the proposed transmission block diagram.

Finally, critical constraints and open issues have been identified, and will serve as an input for future work in T2.5. Having the simulation methodology and baseline simulators now available, the simulation accuracy and capability will be constantly further improved. Major next steps include a detailed investigation of the optimum degree of adaptivity, i.e., the trade-off between overhead, complexity and performance. Additionally the robustness of different spatial processing technique with respect to imperfections encountered in an operational system (e.g., channel estimation errors, quantisation, and delay) will be researched. Of major importance is also the development of a joint framework and optimisation of scheduling and radio resource management in frequency, time, and spatial domains in cooperation with other tasks of WP2.

12. References

[3GPP] 3GPP TR25.996 V6.1.0 (2003-09) “Spatial channel model for multiple input multiple output (MIMO) simulations” Release 6.

[3GPP2] 3GPP TSG RAN WG1, “Throughput Simulations for MIMO and Transmit Diversity Enhancements to HSDPA”, TSGR1-001388.

[3GPP2E] Ericsson, “Effective SNR mapping for modelling frame error rates in multiple-state channels,” 3GPP2-C30-20030429-010

[3GPPE] Ericsson, “System-level evaluation of OFDM – further considerations,” 3GPP TSG RAN WG1 #35 R1-031303

[3GPPE2] Ericsson, “System-level evaluation of OFDM – further considerations,” 3GPP TSG RAN WG1 #35 R1-031303

[3GPPN] Nortel Networks, “Effective SIR Computation for OFDM System-Level Simulations,” 3GPP TSG RAN WG1 #35 R1-031370

[3GPPN2] Nortel Networks, "OFDM Exponential Effective SIR Mapping Validation, EESM Simulation Results", 3GPP R1-04-0089, Jan. 2004.

[Ala98] S. M. Alamouti, “A simple transmit diversity technique for wireless communications,” IEEE Journal on Selected Areas in Comm., vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[AVZ02] Agrell E., Vardy A., Zeger K., “Closet point search in lattices,” IEEE Trans. Inform. Theory, vol. 48, pp.2201-2213, Aug. 2002

[Bal82] C. A. Balanis, Antenna Theory: Analysis and Design, Harper and Row, New York, 1982.

[BB03] Brunel L., Boutros J., “Lattice decoding for joint detection in direct-sequence CDMA systems,” IEEE Trans. Inform. Theory, vol. 49, pp.1030-1037, Apr. 2003

[BB04] Borgmann M., Bölcskei H., “Noncoherent space-frequency coded MIMO-OFDM,” IEEE J. on Select. Areas Comm., submitted in April 2004, to appear

[BB04b] Borgmann M., Bölcskei H., “MIMO communication in the wideband regime,” 2004, in preparation

[BBP03] Boelcskei H., Borgamnn M., Paulraj A.J., “Space-frequency coded MIMO-OFDM with variable multiplexing-diversity tradeoff,” in Proc. IEEE Int. Conf. Communications, Anchorage, pp. 2837-2841, May 2003

[BHN00] Brunner C., Hammerschmidt J., Nossek J., “Downlink eigenbeam-forming in WCDMA,” presented at the Eur. Wireless 2000, Dresden, Germany, September 2000.

[BHW03] Bäro S., Hagenauer J., Witzke M., “Iterative Detection of MIMO Transmission using a List-Sequential (LISS) Detector,” IEEE International Conference on Communications 4 (2003), 2653–2657.

[Bin90] Bingham J.A.C., “Multicarrier modulation for data transmission: an idea whose time has come," IEEE Commun. Mag., vol.28, no.5, pp.5 - 14, May 1990

[BJ02] Boche H., Jorswieck E., “Sum Capacity Optimisation of the MIMO Gaussian MAC,” Proceedings of WPMC 2002

[BJ03] Boche H., Jorswieck E., “Multi user MIMO systems, worst case noise and transmitter cooperation,” Proc. ISSPIT 2003, invited paper

[BMM94] Benedetto S., Mondin M., Montorsi G., “Performance evaluation of trellis-coded modulation schemes,” Proc. of the IEEE, Vol. 82, No. 6, pp. 833-855, June 1994.

[BN04] Böelcskei H., Nabar R.U., “Realizing MIMO gains without user cooperation in large single-antenna adhoc wireless networks,” in Proc. IEEE Intl. Symposium on Inform. Theory, Chicago, IL, July 2004

[BP00] Bölcskei H., Paulraj A.J., “Performance of space-time codes in the presence of spatial fading correlation,” Proceedings of the 34th Asilomar Conference on Signals, Systems and Computers, pp. 687-693, Oct. 2000.

[BP00a] H. Bölcskei and A. J. Paulraj, “Space-frequency coded broadband OFDM systems,” in Proc. IEEE WCNC-2000, Chicago, IL, Sept. 2000, pp. 1-6

[BPS98] Biglieri E., Proakis J.G., Shamai S., “Fading channels: information-theoretic and communications aspects,” IEEE Trans. Inform. Theory, vol. 44, no. 6, pp. 2619-2692, Oct. 1998

[BV01] M. Brehler and M. K. Varanasi, “Noncoherent multi-user space-time communications,” in Proc. 2001 Conference on Information Sciences and Systems, The John Hopkins University, March 21—23, 2001

[CCB95] Chow P.S., Cioffi J.M., and Bingham J.A.C., “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, Nov. 1995

[CCC00] Won-Joon Choi; Kok-Wui Cheong; Cioffi, J.M.; “Iterative soft interference cancellation for multiple antenna systems” Wireless Communications and Networking Conference, 2000. WCNC. 2000 IEEE , Volume: 1 , 23-28 Sept. 2000 Pages:304 - 309 vol.1

[CH00] Cui D., Haimovich A.M., “Design and performance of turbo space-time coded modulation,” in Proc. IEEE GLOBECOM’00, Vol. 3, pp. 1627-1631, San Francisco, USA, Nov./Dec. 2000

[CLH01] Chung S.T., Lozano A., Huang H.C., “Approaching Eigenmode BLAST Channel Capacity Using V-BLAST with Rate and Power Feedback,” in Proc. IEEE Vehicular Technology Conf. (VTC Fall 2001), vol. 2, pp. 915-919, Atlantic City, NJ, Oct. 2001.

[CM04] Lai-U. Choi and Ross D. Murch, “A transmit preprocessing technique for multi-user MIMO systems using a decomposition approach,” IEEE Trans. on Wireless Communications, vol. 3, no. 1, pp. 20–24, 2004.

[CTB96] Caire G., Taricco G., and Biglieri E., “Capacity of bit-interleaved channels,” Electronic Letters, vol. 32, no. 12, pp 1060-1061,June 1996.

[CTL04] Codreanu M., Tujkovic D. and Latva-aho M., “Adaptive MIMO-OFDM with Low Signalling Overhead for Unbalanced Antenna Systems”, In Proc. PIMRC2004, Barcelona, Spain, 5-8 Sept. 2004

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

[D2.4] WINNER D2.4 v1.0, “Assessment of Adaptive Transmission Technologies,” due in February 2005

[D2.3] WINNER D2.3, Assessment of radio-link technologies, Feb. 2005.

[D5.1] WINNER D5.1 v1.0 “A set of channel and propagation models for early link and system level simulations”, April 2004

[D7.1] WINNER D7.1 v1.0 “System Requirements” July 2004.

[D7.2] WINNER D7.2 v1.0 “System Assessment Criteria Specification,” July 2004.

[DCB00] Damen O., Chkeif A., Belfiore J.C., “Lattice-code decoder for space-time codes,” IEEE Comm. Letters, pp. 161-163, May 2000

[DGC03] O. M. Damen, H. E. Gamal, and G. Caire, “On the Complexity of ML Detection and the Search for the Closest Lattice Point,” IEEE Transactions on Information Theory, vol. 59, no. 10, pp. 2400–2414, Oct. 2003.

[DKF+] Dinis R., Kalbasi R., Falconer D., Banihashemi A., “Iterative Layered Space-Time Receivers for Single-Carrier Transmission over Severe Time-Dispersive Channels,” accepted for publication in IEEE Comm. Letters Mar. 2004.

[Eri04] Ericsson, “Selective Per Antenna Rate Control (S-PARC),” 3GPP TSG RAN WG1, R1-04-0307, 2004

[FCV+01] Firmanto W., Chen Z., Vucetic B., Yuan J., “Design of space-time turbo trellis coded modulation for fading channels,” in Proc. GLOBECOM’01, Vol. 2, pp. 1093-1097, Dec. 2001

[FGH05] M. Fuchs, G. Del Galdo and M. Haardt, “A Novel Tree-based Scheduling Algorithm for the Downlink of Multi-user MIMO Systems with ZF Beamforming,” In Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP), Philadelphia, PA, March 2005.

[FH96] Fischer R.F.H. and Huber J.B., “A new loading algorithm for discrete multitone transmission," Proc. IEEE Global Telecommun. Conf., London, UK, pp.724 - 728, Nov.18 – 22, 1996

[FITD331] IST-2000-30116 FITNESS Project, “MTMR Baseband Transceivers Needs for Intra-system and Inter-system (UMTS/WLAN) Reconfigurability, D3.3.1.

[FITD5.3] IST-FITNESS D5.3: “Final FITNESS Project Report – MTMR Technologies for UMTS and WLAN transparent operation: Performance evaluation and recommendations”

[FLOD14] IST-2001-32125 FLOWS Deliverable D14, “Selected MIMO Techniques and their Performance,” January 2004

[GB04] Gärtner M.E., Bölcskei H., “Ergodic capacity and outage properties of CDMA in multiple-access fading channels,” in Proc. International Symposium on Information Theory and its Applications (ISITA), Parma, Italy, Oct. 2004

[GC98] Goldsmith A.J., Chua S.G., “Adaptive Coded Modulation for Fading Channels,” IEEE Transactions on Communications, Vol. 46, No. 5, May 1998

[Gen03] Geng N., “Multi-User MIMO Cell Throughput for Uplink and Downlink,” in Proc. of ECWT (European Conference on Wireless Technologies), Munich, Oct. 2003

[GFB+96] Guey J.C., Fitz M., Bell M., Kuo W.Y., “Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels,” in Proc. IEEE VTC’96, pp. 136-140, 1996

[GH04] Del Galdo and M. Haardt, “Comparison of zero-forcing methods for downlink spatial multiplexing in realistic multi-user MIMO channels,” in Proc. 59th IEEE Vehicular Technology Conf. (VTC 2004 Spring), Milan, Italy, May 2004.

[GJJ+03] Goldsmith A., Jafar S.A., Jindal N., Vishwanath S., “Capacity Limits of MIMO Channels,” IEEE Journal On Selected Areas In Communications, Vol. 21, No. 5, June 2003

[GK00] Gupta P., Kumar P.R., “The capacity of wireless networks,” IEEE Trans. Inform. Theory, vol. 46, no. 2, pp. 388-404, March 2000

[GMF00] Geng J., Mitra U., Fitz M., “Optimal space-time block codes for CDMA Systems,” in MILCOM’2000, Vol. 1, pp. 387-391, April 2000

[GP94] D. Gerlach and A. Paulraj, “Adaptive transmitting antenna arrays with feedback,” IEEE Signal Processing Letters, vol. 1, no. 10, pp. 150–152, October 1994.

[GS00] Ganesan G., Stoica P., “Space-Time Diversity using Orthogonal and Amicable Orthogonal Designs,” in IEEE International Conference on Acoustics, Speech and Signal Processing, volume 5, pages 2561-2564, Istanbul, Turkey, June 2000

[GS03] J. Giese and M. Skoglund, “Space-time constellation design for unknown frequency-selective channels,” in Proc. Int. Conf. on Communications, Anchorage, Alaska, USA, May 2003

[GS04] J. Giese and M. Skoglund, “Performance of Differential Unitary Space-Frequency Modulation”, in Proc. International Symposium on Information Theory and its Applications, Parma, Italy, Oct. 2004

[GT02] Grossglauser M., Tse D.N.C., “Mobility increase the capacity of adhoc wireless networks,” IEEE/ACM Trans. Networking, vol. 10, no. 4, pp. 477-486, Aug. 2002

[GV02] Gastpar M., Vetterli M., “On the capacity of wireless networks; the relay case,” in Proc. IEEE INFOCOM, vol. 3, New York, NY, June 2002, pp. 1577-1586

[GVK02] N. Geng, I. Viering, M. Kiessling, “Multi-User MIMO-OFDM Cell Throughput under Real-World Propagation Conditions,” Proc. IEEE Vehicular Technology Conf VTC Fall 2002, Vancouver, Canada, Sept. 2002

[HB03] T. Haustein and H. Boche, “On optimal power allocation and bit-loading strategies for the MIMO transmitter with channel knowlegde”, Proc. of IEEE ICASSP 2003, vol. IV, pp. 405-409, April 2003.

[HBP01] Heath R., Bölcskei H., Paulraj A., “Space-time signalling and frame theory” in Proc. ICASSP’2002, Vol. 2, pp. 1194-1199, 2001

[HE00] Hammons A., El Gamal H., "On the theory of space-time codes for PSK modulation," IEEE Trans. Information Theory, Vol. 46, No. 2, pp. 524-542, March 2000

[HH01] Hassibi B., Hochwald B.M., "High-rate codes that are linear in space and time," Bell Labs internal report, April 2001

[HM00] Hochwald B.M., Marzetta T.L., "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," IEEE. Trans. on Inf. Theory, pp. 543-564, March 2000

[HMP99] Hochwald B.M., Marzetta T.L., Papadias C.B., "A Novel Space-Time Spreading Scheme for Wireless CDMA Systems," in Thirty Seventh Annual Allerton Conference on Communication, Control and Computing, Monticello, Illinois, USA, September 1999

[HMR+00] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens and R. Urbanke, "Systematic Design of Unitary Space-Time Constellations," IEEE Trans. On Info. Theory, vol. 46, no. 6, pp. 1962-1973, Sept. 2000

[HP02] R.W.Jr. Heath and A.. Paulraj, "Switching between multiplexing and diversity based on constellation distance," in Proc. of Allerton Conf. on Comm. Control and Comp., Oct. 2002

[HS00] Hochwald B., Sweldens W., "Differential Unitary Space-time Modulation," IEEE Trans. on Comm., vol. 48, pp. 2041-2052, Dec. 2000

[HT03] Hochwald B.M., Ten Brink S., "Achieving near-capacity on a multiple-antenna channel," IEEE Trans. on Comm., pp. 389-399, March 2003

[HtB03] B. M. Hochwald and S. ten Brink, "Achieving Near-Capacity on a Multiple-Antenna Channel," IEEE Transactions on Communications, vol. 51, no. 3, pp. 389–399, Mar. 2003.

[HTW03] Hottinen A., Tirkkonen O., Wichmann R., Multi-Antenna Transceiver Techniques for 3G and Beyond, John Wiley & Sons, Chichester, 2003

[Hug00] Hughes B.L., "Differential Space-Time Modulation," IEEE Trans. on Inf. Theory, vol. 46, pp. 2567-2578, Nov. 2000

[IEEE80211] IEEE 802.11-04/0269: "PHY abstraction based on PER prediction"; March 2004

[IME03] IST-2000-30148 I-METRA, "Final Report", October 2003

[IMY+01] Ionescu D.M., Mukkavilli K.K., Yan Z., Lilleberg J., "Improved 8- and 16-State Space-Time codes for 4PSK with Two Transmit Antennas," IEEE Commun. Letters, vol. 5, pp. 301-303, July 2001

[Ion03] Ionescu D.M., "On Space-Time Code Design," IEEE Trans. Wireless Comm., vol. 2, pp. 20-28, Jan. 2003

[Jaf01] Jafarkhani H., "A Quasi-Orthogonal Space-Time Block Code," IEEE Trans. Comm., pp. 1-4, Jan. 2001

[JB02] Jorswieck E., Boche H., "On the optimality-range of beamforming for MIMO systems with covariance feedback," in Proc. of 2002 Conf. on Inf. Science and Systems, Princeton University, March 2002

[JB03] E. A. Jorswieck and H. Boche, "Transmission Strategies for the MIMO MAC with MMSE receiver: Average MSE Optimisation and Achievable Individual MSE Region", IEEE Trans. on Signal Processing, vol. 51, no. 11, pp. 2872-2881, Nov. 2003.

[JBS04] Jorswieck E., Boche H., Schubert M., Analysis of Multi-user MIMO Systems with MMSE Receiver Based on Worst Case Noise, Proc. ITG Workshop on Smart Antennas, 2004

[JBU04] Joham M., Brehmer J., Utschick W., "MMSE approaches to multiuser spatio-temporal Tomlinson-Harashima precoding," In Proc. ITG SCC'04, pp 387-394, January 2004.

[JDL+04] J.-P. Javaudin, C. Dubuc, D Lacroux, and M Earnshaw, "An OFDM Evolution for the UMTS High Speed Downlink Packet Access," IEEE VTC-2004 Fall, Los Angeles, California, USA, September 2004

[JKG+02] M. Joham, K. Kusume, M. H. Gzara, W. Utschick, and J. A. Nossek, "Transmit Wiener Filter for the Downlink of TDD DS-CDMA Systems", in Proc. ISSSTA 2002, Sep. 2002.

[JO04] J.Jaldén, B. Ottersten, “On the Complexity of Sphere Decoding in Digital Communications,” IEEE Transactions on Signal Processing, To Appear

[JRK+99] Jalloul L., Rohani K., Kuchi K. and Chen J., “Performance analysis of CDMA transmit diversity methods,” Proc. IEEE VTC'99 Fall, Vol. 3, pp. 1326-1330, Sept. 1999

[JS03] Jafarkhani H., Seshadri N., “Super-orthogonal space-time trellis codes,” IEEE Trans. Information Theory, Vol. 49, No. 4, pp. 937-5650, Apr. 2003

[JSB+04] E. Jorswieck, A. Sezgin, H. Boche and E. Costa, "Optimal transmit strategies in MIMO Ricean Channels with MMSE Receiver," in Proc. 60th IEEE Vehicular Technology Conf. (VTC 2004 Fall), Los Angeles, CA, Sept. 2004.

[JSO02] Jöngren G., Skoglund M., Ottersten B., “Combining beamforming and orthogonal space-time block coding,” IEEE Trans. on Inf. Theory, pp. 611-627, March 2002

[KH95] R. Knopp and P.A. Humblet, “Multiple-accessing over frequency-selective fading channels,” Proc. IEEE PIMRC-95, vol. 3, Apr. 1995, pp. 1326-1330.

[LF99] Liu Y., Fitz M.P., “Space-Time Turbo Codes,” in Proc. Allerton 1999

[LR99] Liberti J.C., Rappaport T.S., Smart Antennas for Wireless Communications, Prentice Hall, 1999

[LRZ03] Lampe M., Rohling H., Zirwas W., “PER Prediction for PHY Mode Selection in OFDM Communication Systems,” Proc. of Globecom 2003, Nov. 2003

[LS01] Lampe L.H.J., Schober R., “Differential modulation diversity for OFDM,” in Proceedings of 6. International OFDM-Workshop , pp. 19.1-19.6, Hamburg, Germany, September 2001

[Luc02] Lucent, TSG-R1-01-0879, “Increasing MIMO Throughput with Per-Antenna Rate Control,” 3GPP TSG RAN WG1 Document, available through ftp://ftp.3gpp.org/, 2002

[MB04] “MIMO Mode Table for 802.11n”, R. Mahadevappa and S. ten Brink, Realtek Semiconductors, document IEEE 802.11-04/553r0, May 2004

[MBV02] M. L. McCloud, M. Brehler and M. K. Varanasi, “Signal design and convolutional coding for noncoherent space-time communication on the Rayleigh fading channel,” IEEE Trans. On Info. Theory, May 2002

[METD3.2] IST-1999-11729 METRA, Deliverable D3.2, “Review and Selection of Relevant Algorithms,” June 2000

[METD5.1] IST-1999-11729 METRA, Deliverable D5.1, “Cost versus Benefit Trade-Offs of Multi-Element Transmission/Reception”, August 2001

[MZ04] Molisch A.F., Zhang X., “FFT-based hybrid Antenna Selection Schemes for Spatially Correlated MIMO Channels,” IEEE Com. Let., Vol. 8, No. 1, pp. 36-38, January 2004

[Nar99] Narayanan K.R., “Turbo Decoding of Concatenated Space-Time Codes,” in Proc. Allerton Conf. 1999

[NBK04] R. U. Nabar, H. Boelcskei, and F. W. Kneubuehler, “Fading Relay Channels: Performance Limits and Space-Time Signal Design”, IEEE Journal on Selected Areas in Communications, June 2004, to appear.

[NCTH82] K. Nishimori, K. Cho, Y. Takatori, and T. Hori, Automatic Calibration Method Using Transmitting Signals of an Adaptive Array for TDD systems, IEEE Trans. On Vehicular Technologies, 50(6):1636-1640, November 1982.

[NR98] Nanda S., Rege K.M., “Frame Error Rates for Convolutional Codes on Fading Channels and the Concept of Effective Eb/N0,” IEEE Trans. on Veh. Techn., vol. 47, no. 4, November 1998, pp. 1245—1250

[ODS02] Onggosanusi E.N., Dabak A.G., Schmidl T.A., “High Rate Space-Time Block Coded Scheme: Performance and Improvement in Correlated Fading Channels,” Wireless Communications and Networking Conf. 2002, vol. 1, pp. 194-199, March 2002

[PF01] Papadias C.B., Foschini G.J., “A Space-Time Coding Approach for Systems Employing Four Transmit Antennas,” in Proc. of the ICASSP Conference, Salt Lake City, Utah, USA, June 2001.

[PNG03] A. Paulraj, R. Nabar and D.Gore, Introduction to Space-Time Wireless Communications, Cambridge University Press, 2003.

[PP97] Paulraj A.J., Papadias C.B., "Space-Time Processing for Wireless Communications," IEEE Signal Proc. Magazine, pp. 49-83, Nov. 1997

[QTM02] W. Qiu, H. Tröger, and M. Meurer, "Joint transmission (JT) in multi-user MIMO transmission systems," EURO-COST, COST 273 TD(02)008,2002.

[RM94] Rupf M., Massey J.L., "Optimum sequence multisets for synchronous code-division multiple-access channels," IEEE Trans. Inform. Theory, vol. 40, no. 4, pp. 1261-1266, July 1994

[RSU01] T. Richardson, M. Shokrollahi, and R. Urbanke, "Design of Capacity-Approaching irregular Low-Density Parity-Check Codes," IEEE Transaction on Information Theory, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[RUY02] Rose C., Ulukus S., Yates R.D., Wireless Systems and Interference Avoidance, IEEE Trans. On Wireless Communications, 1(3):415-428, July 2002

[RW00] Reynolds D., Wang X., "Low-complexity turbo equalization for diversity channels," Signal Processing, Elsevier Science Publishers, vol. 81, no. 5, pp. 989–995, May 2000

[SAG01] Siemens, "Description of the eigenbeamformer concept (update) and performance evaluation," R1-02-0203, 3GPP TSG-RAN WG 1 document, 2001

[Sam03] Samsung, TSG-R1-03-0877, "Overview MIMO Contributions: Throughput Viewpoint," 3GPP TSG RAN WG1 Document, available through ftp://ftp.3gpp.org/, 2003

[SAT03] IST-1999-10322 SATURN, "Final Report", April 2003

[SBO04] Svedman P., Bengtsson M., Ottersten B., "Table Based Performance Evaluation for HIPERLAN/2 Systems – A Multi Parameter Design," Report Royal Institute of Technology (KTH), Stockholm, Jan. 2004

[SD99] Stefanov A., Duman T., "Turbo Coded Modulation for Wireless Communications with Antenna Diversity," in Proc. VTC'99, Vol. 3, pp. 1565-1569, Sept. 1999

[SFG+00] Shiu D., Foschini J.G., Gans M., Kahn J.M., "Fading correlation and its effect on the capacity of multi-element antenna systems," IEEE Transactions on Communication, Vol. 48, pp. 502-513, March 2000.

[SG01] Su H., Geraniotis E., "Space-Time Turbo Codes with Full Antenna Diversity," IEEE Trans. on Comm., vol. 49, pp. 47-57, Jan. 2001

[SH02] Q. Spencer and M. Haardt, "Capacity and downlink transmission algorithms for a multi-user MIMO channel", in Proc. 36th Asilomar Conf. on Signals, Systems, and Computers, Pacific Grove, CA, IEEE Computer Society Press, Nov. 2002.

[SH04] V. Stankovic and M Haardt, "Multi-user MIMO Downlink Precoding for Users with Multiple Antennas", in Proc. of the 12-th Meeting of the Wireless World Research Forum (WWRF), Toronto, ON, Canada, Nov. 2004

[SH05] V. Stankovic and M. Haardt, "Successive optimization Tomlinson-Harashima precoding (SO THP) for multi-user MIMO systems," in Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP), Philadelphia, PA, Mar. 2005.

[SH05a] V. Stankovic and M. Haardt, "Multi-user MIMO downlink beamforming over correlated MIMO channels," in Proc.International ITG/IEEE Workshop on Smart Antennas, Duisburg, Germany, April. 2005.

[SHF04] V. Stankovic, M. Haardt, and M. Fuchs, "Combination of block diagonalization and THP transmit filtering for downlink beamforming in multi-user MIMO systems," in Proc. European Conf. on Wireless Technology (ECWT 2004), Amsterdam, The Netherlands, Oct. 2004.

[SL02] Schober R., Lampe L.H.J., "Noncoherent receivers for differential space-time modulation," IEEE Trans. on Communications, vol. 50, pp. 768-77, May 2002

[SP02] Sampath H., Paulraj A., "Linear precoding for space-time coded systems with known fading correlations," IEEE Comm. Letters, pp. 239-241, June 2002

[SPC03] Sharma N., Papadias C.B., “Improved Quasi-Orthogonal Codes Through Constellation Rotation,” IEEE Trans. Comm. Vol. 51, No.3, pp. 332-335, March 2003.

[SPP02] Sandhu S., Pandit K., Paulraj A., “On non-linear unitary space-time codes,” in Proc. IEEE ISIT, p. 416, June 2002

[SPS+04] Q. H. Spencer, C. B. Peel, A. L. Swindlehurst, and M. Haardt, “An introduction to the multi-user MIMO downlink,” IEEE Communications Magazine, pp. 60-67, Oct. 2004, special issue on MIMO Systems.

[SSB+02] A. Scaglione, P. Stoica, S. Barbarossa, G. Giannakis, and H. Sampath, "Optimal Designs for Space-Time Linear Precoders and Decoders," IEEE Trans. Sig. Proc., 50(5), 1051-1064, May 2002.

[SSH04] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, “Zero-forcing methods for downlink spatial multiplexing in multi-user MIMO channels,” IEEE Trans. Signal Processing, vol. 52, pp. 461-471, Feb. 2004.

[SSP01] H. Sampath, P. Stoica, and A. Paulraj, "Generalized Linear Precoder and Decoder Design for MIMO Channels Using the Weighted MMSE Criterion," IEEE Trans. Comm., 49(12), 2198-2206, Dec. 2001.

[SWC04] Patrick Svedman, Sarah Kate Wilson, Leonard J. Cimini, Jr., and Björn Ottersten, “Opportunistic Beamforming and Scheduling for OFDM Systems,” submitted to IEEE Trans. on Communicaitons, Oct. 2004

[SWO04] Patrick Svedman, Sarah Kate Wilson, and Björn Ottersten, “A QoS-Aware Proportional Fair Scheduler for Opportunistic OFDM,” in Proc. IEEE VTC-2004 Fall, Los Angeles, CA, Sept. 2004

[SWX] Song A., Wang G., Xia X.G., “Some super-orthogonal space-time trellis codes based on non-PSK MTCM,” to appear in IEEE Trans. Wireless Commun, March, 2005.

[T2.5] WINNER T2.5 Document, “Basic Simulation Parameters”, v0.7.

[T2.7] WINNER T2.7 Document, “WINNER Air Interface: Concept Proposal #1,” Issue 01, 17.11.04

[TBH00] Tirkkonen O., Boariu A., Hottinen A., “Minimal Non-Orthogonality Rate 1 Space-Time Block Code for 3+ Tx,” in Proc. Of the International Symposium on Spread Spectrum Techniques & Applications, pp. 429-432, September 2000

[TC04] Tolli A. and Codreanu M., “Compensation of Interference Non-Reciprocity in Adaptive TDD MIMO-OFDM Systems” Proc. of PIMRC2004, Barcelona, Spain, 5-8 Sept. 2004

[Tel95] E. Telatar, Capacity of Multi-antenna Gaussian Channels, AT&T Bell Labs, 1995.

[TH00] O. Tirkkonen, A. Hottinen, “Complex Space-Time Block Codes for Four Tx Antennas”, in Proc. GLOBECOM conf., vol. 2, pp. 1005-1009, 2000

[TH01] Tirkkonen O., Hottinen A., “Improved MIMO performance with non-orthogonal space-time block codes,” in Proc. of Globecom, SanAntonio, vol. 2, pp. 1122-1126, Nov. 2001

[TI01] Texas Instruments. “Double-STTD scheme for HSDPA systems with four transmit antennas: Link Level Simulation Results”. TSG-R WG1 document, TSGR1#20(01)0458, 21st -24th May, 2001, Busan, Korea

[TJC99] V. Tarokh, H. Jafarkhani, A.R. Calderbank, “Space-Time Block Codes from Orthogonal Designs”, IEEE Trans. on Inf. Theory, vol. 45, n°5, pp. 1456-1467, July 1999

[TJL02] Tujkovic D., Juntti M., Latva-aho M., “Space-time turbo coded modulation: Design and applications,” EURASIP Journal on Applied Signal Processing, Special issue on Space-time coding and its applications, Part I, vol. 2002, no. 3, March 2002

[TNS+99] Tarokh V., Seshadri N., Calderbank A.R., “Space-Time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction,” IEEE Trans. Inf. Theory, vol. 44, pp. 744-765, March 1998

[TSC98] Tarokh V., Seshadri N., Calderbank A.R., “Space-Time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction,” IEEE Trans. Inf. Theory, vol. 44, pp. 744-765, March 1998

[Tuj00] Tujkovic D., “Recursive space-time trellis codes for turbo coded modulation,” in Proc. IEEE GLOBECOM’00, Vol. 2, pp. 1010-1015, San Francisco, USA, Nov./Dec. 2000

[Tuj03] D. Tujkovic, Space-time turbo coded modulation for wireless communication systems, Ph.D. thesis, Ouluensis. University of Oulu, Finland, 2003.

[UMTS30.03] TR 101 112 V 3.2.0 (1998), Universal Mobile Telecommunications System (UMTS); Selection procedures for the choice of radio transmission technologies of the UMTS (UMTS 30.03 version 3.2.0)

[VAT99] P. Viswanath, V. Anantharam, D. N. C. Tse, “Optimal sequences, power control and user capacity of synchronous CDMA systems with linear MMSE multi-user receivers”, IEEE Trans. on Information Theory, vol. 45, no. 6, pp. 1968-1983, Sept. 1999.

[VB04] S. Visuri and H. Bölcskei, “MIMO-OFDM multiple access with variable amount of collision,” IEEE Trans. Inform. Theory, 2004, submitted.

[VB99] Viterbo E., Boutros J., “A universal lattice code decoder for fading channels,” IEEE Trans. on Inf. Theory, pp. 1639-1642, July 1999

[Ver02] S. Verdú, “Spectral efficiency in the wideband regime,” IEEE Trans. Inform. Theory, vol. 48, no. 6, pp. 1319-1343, June 2002

[VG97] Varanasi M.K., Guess T., “Optimum decision feedback multi-user equalisation with successive decoding achieves the total capacity of the Gaussian multiple-access channel,” in Proc. Asilomar Conf. on Signals, Systems and Computers, pp. 1405-1409, Monterey, CA, Nov. 1997.

[VH03] H. Vikalo and B. Hassibi, “The Expected Complexity of Sphere Decoding, Part I: Theory, Part II: Applications,” IEEE Transactions on Signal Processing, submitted for publication, 2003.

[VTL02] P. Viswanath, D. N. C. Tse, and R. Laroia, “Opportunistic Beamforming Using Dumb Antennas,” IEEE Trans. on Information, Theory, June 2002

[VTL04] M. Vehkapera, D. Tujkovic, Z. Li and M. Juntti, “Receiver design for spatially layered downlink MC-CDMA system,” to appear in IEEE Trans. Veh. Technol., 2004.

[WBK04] D. Wübben, R. Boehnke, V. Kuehn, K.D. Kammeyer, “MMSE Based Lattice-Reduction for Near-ML Detection of MIMO Systems”, ITG Workshop on Smart Antennas, Munich, Germany, 18.03.2004

[WFG+98] Wolniansky P.W., Foschini G.J., Golden G.D., Valenzuela R.A., “V-BLAST: An Architecture for Realizing Very High Data Rates Over the Rich-Scattering Wireless Channel,” Proc. ISSSE 98, Pisa, Sept. 1998

[WLF03] C. Windpassinger, L. Lampe, R. F. H. Fischer “From Lattice-Reduction-Aided Detection Towards Maximum-Likelihood Detection in MIMO Systems¨ In Proceedings of the International Conference on Wireless and Optical Communications (WOC 2003), Banff, Canada, July 2003

[WOS+03] W. Wang, T. Ottosson, M. Sternad, A. Ahlén and A. Svensson, “Impact of multi-user diversity and channel variability on adaptive OFDM”, IEEE VTC 2003-Fall, Orlando, Fla., Oct. 2003

[WOSAS03] W. Wang, T. Ottosson, M. Sternad, A. Ahlén and A. Svensson, “Impact of multi-user diversity and channel variability on adaptive OFMD”, IEEE VTC 2003-Fall, Orlando, Fla., Oct. 2003.

[WP99] Wang X., Poor H.V., “Iterative (turbo) soft interference cancellation and decoding for coded CDMA,” IEEE Trans Commun., vol. 47, no. 7, pp. 1046–1061, July 1999

[WRB+02] D. Wuebben, J. Rinas, R. Boehnke, V. Kuehn, and K. Kammeyer, “Efficient Algorithm for Detecting Layered Space-Time Codes,” in 4th International ITG Conference on Source and Channel Coding, Berlin, Germany, Jan. 2002.

[YC01] Yu W. and Cioffi J.M., “On constant power water filling," Proc. IEEE Int. Conf. Commun., pp.1665 - 1669, June11 – 14, 2001

[YRB+04] Yu W., Rhee W., Boyd S., Cioffi J., Iterative Water-filling for Gaussian Vector Multiple Access Channels, IEEE Trans. on Information Theory, Jan. 2004

[ZRF04] E. Zimmermann, W. Rave and G. Fettweis, "On the Complexity of Sphere Decoding", Proceedings of the 7th International Symposium on Wireless Personal Multimedia Communications (WPMC04), Abano Terme, Italy, September 2004

[ZT03] L. Zheng and D.N.C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," IEEE Trans. Inform. Theory, vol. 51, no. 11, pp. 2784-2795, Oct. 2003.

[ZVR+03] Zhuang X., Vook F.W., Rouquette-Léveil S., Gosse K., "Transmit Diversity and Spatial Multiplexing in Four Transmit Antenna OFDM," IEEE ICC2003, vol. 2, pp. 1343-1347, April 2003