Real Time Signal Processing for Multi-Antenna Systems and experimental Verification on a reconfigurable Hardware Test-bed

von Diplom-Physiker
Thomas Haustein
aus Berlin

von der Fakultät IV - Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften
- Dr.-Ing. -

genehmigte Dissertation

Promotionsausschuss

Vorsitzender: Prof. Dr. Ing. Adam Wolisz
Berichter: Prof. Dr. Ing. Dr. rer. nat. Holger Boche
Berichter: Prof. Dr. Ing. Johannes Huber

Tag der wissenschaftlichen Aussprache: 23. Januar 2006

Berlin, 2006
D 83

Zusammenfassung

Mehrantennensysteme werden in zukünftigen Mobilfunksystemen der dritten und vierten Generation eingesetzt werden, um die Leistungsfähigkeit der Systeme, ihre spektrale Effizienz sowie die Zuverlässigkeit der Funkverbindung zu verbessern.

Aus der Informationstheorie ist bekannt, dass die Kanalkapazität von Mehrantennensystemen linear mit der Anzahl der verwendeten Sende- und Empfangsantennen ansteigen kann. Wichtige Kenngrößen zur Systemcharakterisierung sind die Kanalkapazität und der mittlere quadratische Fehler (MSE). Beide Kenngrößen hängen von den Eigenschaften des Mehrantennen-Funkkanals ab sowie vom Umfang und der Art der Kanalinformation am Sender und Empfänger.

In der vorliegenden Arbeit werden optimale Übertragungsverfahren für Einnutzer- und Mehrnutzer-Mehrantennen-Systeme hergeleitet, jeweils abhängig von der vorhandenen Kanalkenntnis. Hierzu wird der Summen-MSE als Optimierungskriterium herangezogen. Die sich daraus ergebenden Resultate weichen erwartungsgemäß von denen für die ergodische Kapazität ab. Für das MIMO-Einnutzerszenario mit perfekter Kanalkenntnis am Empfänger und keiner oder ebenfalls perfekter Kanalkenntnis am Sender wird die optimale Übertragungsstrategie hergeleitet. Des weiteren leiten wir die optimale Sendestrategie für den Mehrnutzer-Zugriffskanal mit jeweils einer Antenne pro Mobilterminal und mehreren Antennen an der Basisstation her. Wir untersuchen wie das funktionale Verhalten der Summenrate durch die SNR-Gap-Approximation beeinträchtigt wird.

Um einen Übergang von der etablierten Theorie zu Implementierungen auf echten Systemen zu ermöglichen, werden einige grundsätzliche und durch die Praxis motivierte Betrachtungen für Mehrantennensysteme vorgenommen. Hier betrachten wir insbesondere die Anzahl und Art der verwendeten Antennen auf jeder Seite der Übertragungsstrecke. Wir untersuchen die Auswirkung von Sichtverbindungen im Funkkanal auf systemrelevante Parameter, wie den Rang des Übertragungskanals, die Kapazität und die erreichbare Bitfehlerrate.

Der MIMO Broadcast-Kanal als duales Äquivalent des MIMO-Vielfachzugriffskanals wird für eine Mehrantennenbasisstation und verteilte Mobilterminals mit jeweils einer Antenne betrachtet. Im Fall von ausreichender Kanalkenntnis an der Basisstation können Vorcodiertechniken eingesetzt werden. Wir betrachten verschiedene lineare und nicht-lineare Vorcodierverfahren und vergleichen sie bezüglich der jeweils notwendigen Sendeleistung.

In der Praxis kann die Leistungsfähigkeit eines Übertragungssystems degradieren, durch z.B. Kanalschätzfehler, eine begrenzte Dynamik der Sendeleistungsverstärker oder durch Gleichkanalinterferenzen. Wir untersuchen die Auswirkungen verschiedenen Degradierungsursachen und entwickeln Lösungsvorschläge zu ihrer Beseitigung bzw. Begrenzung.

Die Verwendung kanalangepasster Übertragungsverfahren ist eine Schlüsseltechnik zur Erreichung eines optimalen Datendurchsatzes. Hierzu wird die Datenübertragung an die aktuelle Güte des Funkkanals angepasst, so dass z.B. eine Übertragung in schlechten Kanälen vermieden werden

kann. Wir erweitern das dazu verwendete adaptive Bitladeverfahren um Mehrnutzer-Scheduling-Algorithmen. In einem schichtübergreifenden Ansatz werden sowohl die aktuellen Kanalzustände aller Nutzer als auch Übertragungsgüteanforderungen der einzelnen Nutzer berücksichtigt. Dieser Ansatz ermöglicht einen hohen Zelldurchsatz und eine hohe Systemstabilität bezüglich der Datenwarteschlangen für jeden Nutzer. Eine Echtzeitimplementierung im MIMO-Testbett zeigt, dass eine Umsetzung bereits auf heutiger Hardware möglich ist.

Eine wesentliche Voraussetzung für die Umsetzung der entwickelten Übertragungstechniken ist eine sorgfältige Analyse und Optimierung der notwendigen MIMO-Algorithmen. Wir geben eine Anforderungsanalyse für das Algorithmendesign und diskutieren einige ausgewählte MIMO-Basisbandalgorithmen im Detail besonders bezüglich einer Abbildung auf digitale Signalprozessoren (DSPs).

Der letzte Teil der Arbeit ist den experimentellen Resultaten gewidmet, wo sich Theorie und Praxis treffen. Wir führen in die wesentlichen Übertragungsmodi des MIMO-Testbetts und die dazugehörigen relevanten Systemparameter ein und berichten dann über die Experimente zu Antennendiversität und zum räumlichen Multiplexing. Wir vergleichen verschiedene Übertragngsverfahren bezüglich der gemessenen Bitfehlerrate und des erreichten Datendurchsatzes mit adaptivem Bitladen unter Verwendung von linearen und nicht-linearen Detektionsverfahren. Des weiteren berichten wir über eine erste echtzeitfähige Implementierung von adaptiver Kanalinversion.

Abstract

Multiple-input multiple-output (MIMO) systems will be applied in wireless communications in order to increase the system performance, spectral efficiency, and link reliability. Theoretically, the channel capacity of MIMO systems can grow linearly with the number of transmit and receive antennas. An important performance metric beneath capacity is the normalized mean square error (MSE) under the assumption of optimal linear reception. Both performance measures depend on the properties of the MIMO channel as well as on the available channel state information (CSI) at the transmitter.

In this thesis, we derive optimum transmission strategies of single- and multiuser MIMO systems with respect to the different types of CSI at the transmit and receive side. The optimization is taken under the assumption of the MSE as the objective function. The results differ therefore from those known for ergodic capacity optimizations. We start with a derivation of the optimum transmission strategy for the single user MIMO scenario with perfect channel knowledge at the receiver and no or full channel knowledge at the transmitter. Furthermore, we derive the optimum transmission strategy for a multiple access channel (MAC) with only one antenna per user and several antennas at the base station. We look very close on how the SNR gap approximation, often used for bit-loading approaches, affects the behavior of the sum rate functional which has to be maximized.

To bridge from the well studied multi-antenna theory towards implementations on real-world systems some basic and practical considerations will be made for wireless MIMO systems. Here, the emphasis is put on antenna configurations with respect to the number and kind of antennas at each side of the link. Furthermore we analyze the effect of a line-of-sight on system relevant parameters as the rank of the transmission channel, capacity and achievable bit-error-rates.

The MIMO broadcast channel as the duality equivalent of the MIMO MAC is investigated for a multi-antenna base station and several distributed users. When CSI is available at the BS appropriate pre-coding techniques can be applied. We will look into SVD-MIMO transmission and linear and non-linear pre-coding techniques. Further emphasis is put on a comparison towards the necessary transmit power needed for transmit pre-coding.

In reality the systems can suffer from performance degradation caused by e.g. channel estimation errors, a limited transmitter dynamics or co-channel interference. The impact of each degradation factor will be evaluated and strategies to combat or limit the undesired effects will be proposed.

To achieve optimum system performance adaptive transmission is an important issue. The aim is to adapt the data transmission to the actual channel realization, thus avoiding transmission over bad channels. We will extend channel aware bit-loading with discrete modulation alphabets to multi-user scheduling policies. In a cross-layer approach the optimization considers the instantaneous channel state and quality of service parameters e.g. the data queues of all users as well. This approach allows high cell throughput and stable data queues at the MTs or the BS, both

which may have limited buffer size or user applications may have stringent delay requirements. A real-time implementation on the MIMO test-bed shows the feasibility already on today's hardware.

A prerequisite for a successful implementation is a thorough analysis of the MIMO algorithms necessary to realize the before discussed transmission strategies. We list requirements for real-time algorithms and a few basic algorithms for MIMO base-band signal processing are discussed in detail with respect towards an implementation in a DSP.

The last part of this thesis is dedicated to the experiments where theory and practice will meet. We start with an overview of the real-time MIMO test-bed and the various configurations for the experiments are introduced. We then present experimental results on antenna diversity and spatial multiplexing. We will compare the BER performance and the measured throughput with channel adaptive bit-loading using linear and non-linear detection schemes. Furthermore, we show measurement results of a first implementation of real-time adaptive channel inversion.

Acknowledgements

During my PhD studies, I had the opportunity to meet many interesting and extraordinary individuals. First of all, I thank my advising Professor Holger Boche for his support throughout the years of joint work. He always made it possible to share and discuss ideas and problems. He was a good teacher in information theory and related mathematical techniques. His enthusiasm, knowledge, energy, and visions have motivated me to a great extent. I thank Professor Johannes Huber for serving as second referee and reader of this dissertation.

I want to thank my close colleagues from the implementation team which all supported me in the challenging task of bringing algorithms into practice with real-time implementations and experiments. My special thanks goes to Andreas Forck who became more than a colleague during that time. I thank Volker Jungnickel, Clemens von Helmolt, Udo Krüger, and Holger Gäbler for the inspiring discussions and intense collaboration.

My gratitude extends to Eduard Jorswieck, Aydin Sezgin, Chan Zhou, Stefan Schiffermüller, Volker Pohl, Anastasios Giovanidis, and Martin Schubert for collaborating and co-authoring papers with me. I would like to thank all colleagues at the Heinrich-Hertz-Institut and the Technical University of Berlin for stimulating discussions and additional support.

My wife Viviane deserves my special thanks for her support and encouragement throughout this time.

Contents

List of Figures

List of Tables

1 Nomenclature and abbreviations

Nomenclature

$\mathbf{H}$	Flat Fading MIMO channel matrix
$\mathbf{H}_k$	Channel matrix of user k
h_{ij}	Entry in the i-th column and the j-th row of the channel matrix $\mathbf{H}$
$\mathbf{h}_i$	Transmit or receive vector of the i-th signal
$\mathbf{n}$	Noise vector
$\mathbf{x}$	Transmitted signal vector
$\mathbf{y}$	Received signal vector
σ_N^2	Noise power
K	Number of mobile users
n_T	Number of transmit antennas
m_R	Number of receive antennas
MSE_k	Achievable normalized MSE of user k
P	Transmit power constraint
p_k	Individual power constraint of user k or direction k
R_k	Achievable transmission rate of user k
$\mathbf{Z}$	Noise covariance matrix
$\mathbf{Q}$	Transmit covariance matrix
$\mathbf{Q}_k$	Transmit covariance matrix of user k

Abbreviations

The following list summarizes the acronyms used in this work.

ACI	Adaptive Channel Inversion
ARQ	Automatic Repeat ReQuest
AWGN	Additive White Gaussian Noise
BC	Broadcast Channel
CDMA	Code Division Multiple Access
CI	Channel Inversion
CSI	Channel State Information
cdf	Cumulative Distribution Function
DL	Down-Link
DLL	Data Link Layer
DSL	Digital Subscriber Line
FDD	Frequency Domain Duplex
FDMA	Frequency Division Multiple Access
FEC	Forward Error Correction
i.i.d.	Identically Independent Distributed
ISI	Inter Symbol Interference
JCBF	Joint Costa Beam Forming
JT	Joint Transmission
LHS	Left Hand Side
LCI	Linear Channel Inversion
LMMSE	Linear Minimum Mean Square Estimation
LOS	Line Of Sight

MAC	Multiple Access Channel
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output
ML	Maximum Likelihood
MLD	Maximum Likelihood Detector
MMSE	Minimum Mean Square Error
MRC	Maximum Ratio Combining
MSE	Average normalized Mean Square Error
NLOS	Non-Line Of Sight
PAM	Pulse Amplitude Modulation
pdf	Probability Density Function
PEF	Power Enhancement Factor
PSK	Phase Shift Keying
OFDM	Orthogonal Frequency Division Multiple Access
QAM	Quadrature Amplitude Modulation
QLD	QL-Decomposition
QoS	Quality of Service
QRD	QR-Decomposition
RHS	Right Hand Side
RKI	Ranked Known Interference
Rx	Receiver
SDMA	Space Division Multiple Access
SIC	Successive Interference Cancellation
SIMO	Single Input Multiple Output
SNR	Average Signal to Noise Ratio
SVD	Singular Value Decomposition
TDD	Time Domain Duplex
TDMA	Time Division Multiple Access
THP	Tomlinson Harashima Pre-coding
Tx	Transmitter
UL	Up-Link
UMTS	Universal Mobile Telecommunications System
UPA	Uniform Power Allocation
VBLAST	Vertical BLAST
WLAN	Wireless Local Area Network
ZF	Zero Forcing

2 Introduction

2.1 Motivation

The widespread use of wireless and mobile communication devices has changed everyday life tremendously during the recent decade. The introduction of cellular networks laid the foundation for mobile communication almost everywhere, anytime and with everyone. A growing use of data communication mainly over the internet e.g. email, news or information of any kind, produces an increasing demand in wireless data traffic as well. Since wireless connections are generally no exclusive point-to-point connections as land lines used e.g. for telephone and DSL, the available frequency spectrum has to be shared with other users and radio systems.

The tremendous expectations towards the growth of mobile communications made the available spectrum valuable and expensive for licensing. Therefore it is a prerequisite for all service providers and radio systems to exploit the limited resource frequency spectrum very efficiently.

A new transmission concept proposed by Foschini [Fos96] using multiple-antennas at each side of the radio link promised a significant increase in spectral efficiency. An information theoretic basic work by Telatar [Tel95] on the capacity in multi-antenna channels opened intensive research activities in the MIMO[1] area worldwide. The new domain to be exploited is the spatial domain, taking into account the separability of data streams transmitted from different antennas due to their specific and independent spatial interference patterns e.g. in multi-path environments.

Therefore a MIMO transmission scheme allows that several radio links can be supported simultaneously at the same time, in the same frequency band by using multiple antennas at both end of the radio transmission link.

2.2 Related Work

The increasing demand for faster and more reliable wireless communication links reopened discussions on how to exploit all given degrees of freedom in wireless transmission which can come basically from time, frequency, space or scenarios with many users to choose from. Since the time and frequency domains are already exploited to a high extend the spatial domain offers an additional degree of freedom. The work of Foschini [Fos96, FG98] inspired discussion about radio transmission systems with multiple antennas at both ends of the link - so called multiple-input multiple-output (MIMO) systems. The achievable capacity in a single cell multi-user scenario [KH95] was well understood and it was also well known that the use of several antennas at one side of the transmission link can increase the system capacity and performance due to transmit or receive diversity [Jak74]. In recent years, it was found that MIMO systems have the ability to reach higher spectral efficiency than systems using antenna arrays only at one side of the link

[1] MIMO: Multiple-Input Multiple-Output

[Tel99]. This so-called spatial multiplexing was studied in [Sal85, Fos96, RC98, CS03] and is based on the fact that under a sum power constraint the capacity can be increased by establishing several parallel links (MIMO) instead of one SISO[2] link. When the transmission with spatial multiplexing is separable, then the sum capacity is given by the sum of the individual capacities which is always bigger than that of a single antenna link. [ZT03] showed that there exists a fundamental trade-off between multiplexing and diversity gain for any multi-antenna system.

In 1998 a first successful experimental demonstration [WFGV98] proved the practical feasibility of spatial multiplexing in narrow-band frequency-flat channels which boosted research effort in the MIMO area.

In a basic work Telatar [TT98] showed an equivalence between Gaussian and fading channels when we go into the broadband wireless channel. A recent work from [Ver02] investigated the spectral efficiency in the broadband regime in detail.

The general impact of antenna correlation on the achievable capacity was investigated in [CTK98, SFGK99] while [GBGP00, SL03] pointed out the existence of so-called keyhole channels where the capacity can degrade to that of a SISO system despite the fact that there is low fading correlation between antennas.

In order to prepare a transfer of the promising predictions from information theory into practical applications many groups [SO00, KMJ$^+$00, MBFK00, CLW$^+$02] started extensive channel measurement campaigns to study the spatial, temporal and frequency nature of transmission channels in realistic scenarios. The measurement results indicated that even in outdoor scenarios with sufficient scattering environment [LCW$^+$01] high channel capacities similar to indoor scenarios [KWV00, PJHvH02, JPN$^+$02] can be found.

Recent works of [PCL03, Win04, E.J04] gave general frameworks for the optimization of transmit and receive processing for multi-antenna systems towards capacity, throughput or bit-error-rate performance. An extension to frequency selective channels was given by [PFL00, FH03, Joh04, FSH04, Hun02, Man05].

Another important issue is the power control for the transmission link to optimize throughput and bit-error performance. This was investigated in [Sch02] in a general frame work using duality between up-link and down-link. [HL02] investigated the performance under a slow feed-back channel constraint. A work of [HD03] discussed optimum power allocation in the context with multi-carrier systems.

For the case of channel state information (CSI) at the transmitter, the link performance can be enhanced by appropriate signal processing at the transmitter before emitting the signal from the antennas. The most simple way is exploiting transmit diversity [Win98] while linear transmit pre-coding proposed by [HPJ$^+$02, JHJvH02, WM03] or in the context of CDMA [BF01, BMWT00, Irm05] needs more complex signal processing at the transmit side. A first realtime implementation of adaptive linear pre-coding was presented by [HFG$^+$04a].

If CSI is available at the Tx and the Rx then eigenmode transmission [BHSN00, HBD00, EB99, HB03] is the optimum strategy. The data streams are coupled into the eigenspaces of the channel and decoupled at the Rx providing full decorrelation due to the orthogonal sub-spaces. An ASIC-

[2]Single-Input Single-Output

Implementation of algorithms for slow flat fading channels was presented in [PTB03] while other groups conducted experiments on reconfigurable hardware platforms ranging from single carrier transmissions at low data rates [SJZO05] and high data rates [HFGS05] to multicarrier MIMO-OFDM in [NWK$^+$05].

The fundamental result from Costa [Cos83] was applied for non-linear pre-coding in multi-antenna systems by e.g. [SB02a]. In fact, results concerning the capacity of (MIMO) broadcast channels [WJ01, VJG03, JG04, VT03, CS03] have shown that Costa pre-coding is an essential ingredient to achieve the capacity of these channels. A simple version which is based on a decision feed-back equalizer (DFE) structure equivalent to non-linear detection schemes like V-BLAST [Fos96] is Tomlinson-Harashima pre-coding [Tom71, HM72] which was originally proposed for inter-symbol-interference channels but can be easily adapted for MIMO pre-coding [FWLH02, GC02, CS03, IHRF03]. So-called lattice-based pre-coding techniques help to assure a limited transmit power [ESZ00, WF03, Fis02, FWLH02, IBW04]

A further important contribution for the overall multi-antenna system performance is given by a proper coding to combat badly faded channels e.g. [Goe99, CG01]. The additional spatial dimension allows for so-called space-time-codes [Sze05] which perform optimum in the low SNR range by basically transmitting replicas of the same information over e.g. different antennas in different time slots. In parallel very efficient and powerful error correcting codes like Turbo-Codes [BGT93], Product Accumulate Codes [LNG04] or Low-Density-Parity-Check-Codes (LDPC) [Gal62] have been developed over the recent years which are now entering the application stage [LTS00, ZPBF04]. Coded transmission which is a research area in itself was not considered throughout the thesis without disregarding the impact of channel and source coding on the final system performance.

In reality practical transmission systems normally don't apply neither Gaussian alphabets nor infinite interleaving as would be required from the capacity point of view. Nevertheless we are interested in how to achieve optimum rate and performance with e.g. discrete modulation alphabets and/or symbol-by-symbol decisions. This problem is generally referred to as *bit-loading* and can be performed in time, space and frequency [Bin90]. Ref [Cam98] gave theoretical sufficient conditions for discrete bit-loading to be optimum in the context of OFDM. [CCB95, FH96, Cam99, ATG99, AC02, SS04, MDHF03, Dar04, Lam04] proposed bit-loading strategies for fixed rate applications. A recent work by [PV02] discussed an analytical optimization of the joint-error-rate with successive interference cancellation at fixed rate by means of power and bit-allocation. In ref [VG97] it was shown that a transmission using a MMSE-SIC receiver combined with adaptive modulation and coding is capacity achieving at high SNR.

A slightly different bit-loading approach is followed in this thesis. The idea exploits the fact that CSI is available to the transmission system and that the loading of information on available resources is performed such that transmission in bad channels is avoided. Exploiting CSI and the detector structure we can predict the achieved signal-to-noise-and-interference-ratio (SINR) in front of the decision unit. Based on symbol by symbol decisions we can now adapt power and bit-allocation such that all data streams have a desired error probability [HB03, HLB03]. The proposed scheme has variable rate but an upper limited and assured BER, which requires error-correcting codes only to contribute SNR gain instead of protection against fading. This allows for codes with high code rates and schemes like automatic repeat request (ARQ) [ZLH42, FS98, AVCM04, ACV+04, LZG04] are supported ideally since the achieved BER can be controlled to the desired working point. Experiments of [HFG+04c, HFG+04a, HFGS05] could show the advantages of channel aware bit-loading in experiments at high data rate. The resulting variable data rate in a single user scenario might appear unusual, but with an increasing number of users, a multi-user scheduling algorithm can control the data streams individually.

In the reality of multi-user scenarios the user scheduling becomes a challenging task when spectral efficiency and quality-of-service (QoS) e.g. average rate or delay are included in the optimization. [NMR03] discussed single antenna systems and [YC03, Bor, TH98] used simplified assumptions on the packet arrival rates. Works by [HBH03, BW02, BW03, BW04a, WB04a] proposed a powerful framework to solve the complex scheduling task very efficiently, such that a real-time implementation [HFG+04b, HFG+05a] on today's hardware could show the gains towards sum rate and individual QoS requirements of a new scheduling policy.

Recently, an increasing number of implementations of MIMO techniques e.g. [HKM+04, KSM+05] and real-time over-the-air transmission experiments were published e.g. [JFH+05a, NWK+05]. While some groups e.g. [SJZO05] focussed on narrow-band single carrier transmission like in this thesis [HFG+04c, HFG+04a, HFG+04b, HFGS05] while other groups exploited the MIMO techniques straight forward for multi-carrier systems which can exploit higher bandwidths [JFH+05a, JHF+04, BCK03, HPB+04, NWK+05]. When using OFDM to handle the frequency selective channel, all algorithms discussed in this thesis can be mapped in a sub-carrier-by-sub-carrier fashion and the computational effort scales linearly with the number of OFDM tones. An analysis of the algorithmic complexity connected with MIMO-OFDM [PC04, **?**, HSJ+05] showed that further complexity reduction is a key issue for an efficient implementation.

2.3 Contribution and Outline of the Thesis

Throughout this thesis we will focus on narrow-band MIMO transmission systems where the channel coefficients between antenna pairs can be represented by constant complex values, the so-called flat-fading case. For completeness and especially in the context of complexity we will also address broadband transmission systems using multi-carrier techniques in frequency selective channels.

The contributions of this thesis are summarized in the following:

Chapter 3 introduces into multi-antenna systems. We consider several up- and down-link scenarios, statistical channel models and we will give a short introduction on capacity and bit-error-rate issues with multiple antenna systems.

In section 3.5 we will make some basic considerations for the implementation and applications of multiple antenna techniques for Wireless LAN systems. Here, we put an emphasis on antenna configuration with respect to the number of antennas at each side of the link and a general guide towards antenna design. Furthermore we discuss the effect of a line-of-sight on system relevant parameters as the rank of the transmission channel, capacity and achievable bit-error-rates.

The first part of chapter 4 is dedicated to a detailed discussion of multi-antenna transmission schemes. The optimization is taken under the assumption of the mean square error (MSE) as the objective function. The results differ therefore from those known for ergodic capacity optimizations. In section 4.1.1 we start with a derivation of the optimum transmission strategy for the single user MIMO scenario with perfect channel knowledge at the receiver and no or full channel knowledge at the transmitter.

In section 4.1.2 we derive the optimum transmission strategy for a multiple access channel (MAC) with only one antenna per user and several antennas at the base station which uses a MMSE receiver and successive interference cancellation (SIC). Furthermore we look closer how the SNR gap approximation, often used for bit-loading approaches, affects the behavior of the sum rate functional which has to be maximized. Extensive numerical simulations will illustrate the results.

In section 4.2 we discuss the broadcast channel from a multi-antenna base station to several distributed users and appropriate pre-coding techniques. We have a closer view into SVD-MIMO transmission and linear and non-linear pre-coding techniques. Further emphasis is put on a comparison towards the necessary transmit power needed for transmit pre-coding.

Section 4.3 discusses sources of performance degradation in MIMO systems, e.g. channel estimation errors, a limited transmitter dynamics and co-channel interference. The impact of each degradation factor is evaluated and strategies to combat or limit the undesired effects are proposed.

The last two sections of chapter 4 are dedicated to the very important topic of practical bit-loading strategies with discrete modulation alphabets adapted to the actual channel realization. We extend the concept of channel aware bit-loading to multi-user scheduling policies in a cross-layer approach and will consider additionally quality of service parameters e.g. the data queues of the individual users.

Chapter 5 is dedicated to real-time algorithms necessary to realize the before discussed transmission strategies in a real-time application. We give a list of requirements for real-time algorithms and a few basic algorithms for MIMO base-band signal processing are discussed in detail with respect towards an implementation in a DSP. The various algorithmic approaches will be illustrated by examples.

Chapter 6 is dedicated to the experiments where theory and practice will meet. We start with an overview of the real-time MIMO test-bed and the various configuration for the experiments are introduced. In section 6.2 we present experimental results on antenna diversity and spatial multiplexing. After a comparison of the BER performance of linear and non-linear detection schemes we show throughput measurement results with channel adaptive bit-loading using linear and non-linear detection schemes and the throughput optimum SVD scheme which exploits joint signal processing at the transmit and receive side. In section 6.3 we show measurement results of a first implementation of real-time adaptive channel inversion as a special case of spatial pre-coding for a system using non-cooperative receive antennas.

Section 6.4 shows measurement results on multi-user scheduling where we will show that already with today's hardware efficient multi-user scheduling schemes can be implemented which offer a high sum throughput due to spatial multiplexing and limited delays at the same time since the queue length of the users is evaluated as a QoS parameter.

Chapter 7 concludes the thesis and gives an outlook on open problems and future research topics. The list of publications and the bibliography finalize this thesis.

3 Multi-Antenna Systems - Capacity and Bit Error Rates

3.1 Channel Models

3.1.1 Multiple Antenna MIMO Channel Model

We consider a transmission system which uses n_T transmit antennas and m_R receive antennas thus forming a general multiple-input multiple-output (MIMO) transmission channel $\mathbf{H}$.

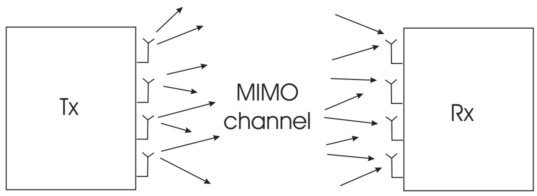

Figure 3.1: A Multiple-Input Multiple-Output (MIMO) channel.

We describe the transmission channel $\mathbf{H}$ in matrix form of dimensions $m_R \times n_T$. For most of this thesis will will assume no frequency selectivity of the transmission channel therefore the channel coefficients h_{ij} reduce to a complex scalar value. Then h_{ij} characterizes the transmission coefficient received at the i-th Rx antenna coming from the j-th Tx antenna. This scalar describes the amplitude and the phase of the transmission from one antenna to another.

Furthermore we apply the commonly accepted flat block fading model.

The **flat fading model** assumes the signal bandwidth of the system to be much shorter than the coherence bandwidth of the radio transmission channel thus reducing the channel coefficients for each antenna pair to a single frequency independent scalar complex value.

The **block fading model** assumes constant fading coefficients over the whole length of a block and i.i.d.[1] coefficients between two following transmission block or channel realizations. In real transmission scenarios two successive channel realizations are at least in part correlated to each other. Nevertheless the block fading model is widely accepted and offers the possibility of comparison with other results and it gives us insight into quantities of interest like e.g. capacity, throughput, outage or BER.

The belonging transmit and receive signals are described as vectors $\mathbf{x}$ and $\mathbf{y}$ of dimensions $n_T \times 1$ and $m_R \times 1$, respectively.

The general MIMO transmission equation has the following form

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{3.1}$$

[1] identically independently distributed

where $\mathbf{n}$ describes the additional noise imposed by the amplifiers at each receive antenna.

The following tableau gives an illustration matrix/vector equation from (3.1)

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{m_R} \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1n_T} \\ h_{21} & h_{22} & & h_{2n_T} \\ \vdots & & \ddots & \vdots \\ h_{m_R 1} & h_{m_R 2} & \cdots & h_{m_R n_T} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n_T} \end{pmatrix} + \begin{pmatrix} n_1 \\ n_2 \\ \vdots \\ n_{m_R} \end{pmatrix}. \tag{3.2}$$

All considerations are based on a uni-lateral transmission view, always keeping in mind that many real communication systems require bi-directional data exchange between the two communication points. Therefore all schematic figures have to be read as a transmission from the left to the right or otherwise it will be indicated by arrows. The term base station (BS), mobile station (MS) or equivalently mobile terminal (MT) is referred to the transmitter (Tx) and the receiver (Rx) or vice versa, depending on the scenario under investigation.

The term ***uplink*** (UL) will be used to describe a transmission from a MS to the BS and the term ***downlink*** (DL) will be used for the opposite transmission direction. This notation pay tribute to the fact that in cellular mobile networks BSs are often placed further up on a hill or on top of a building while the MSs are on ground level in general.

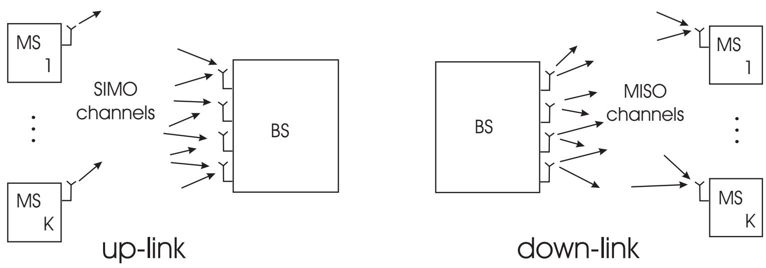

Figure 3.2: Uplink transmission (left) and Downlink transmission (right) in mobile cellular networks when the MSs have only one antenna each.

3.1.2 Single User Channel

A transmission link between two terminals will be referred to as a Single User (SU) channel if only one terminal at each side of the link is involved in the transmission at a time. The terminals can be equipped with only one or several antennas each, which results in SISO[2], SIMO[3], MISO[4] or MIMO[5] channels depending on the number of antennas on each side.

Single User MIMO Channel

We consider a transmitter unit with n_T transmit antennas and a receive unit with m_R receive antennas (Fig. 3.3). We put no limitation on the role of the transmit and receive unit, in gen-

[2]Single-Input Single-Output
[3]Single-Input Multiple-Output
[4]Multiple-Input Single-Output
[5]Multiple-Input Multiple-Output

eral, therefore the notation MT and/or BS is additionally used to identify a specific scenario, if necessary.

It is important to note that in principle both sides of the single user MIMO link are capable of performing joint signal processing. This general ability can then be exploited depending on the amount of **channel state information** (CSI) available at the Tx and/or the Rx. Having said this we can categorize 3 major transmission scenarios.

CSI @ Tx	CSI @ Rx	Transmission Mode
YES	NO	joint Precoding at Tx
NO	**YES**	joint Decoding at Rx
YES	**YES**	joint Precoding at Tx and joint Decoding at Rx

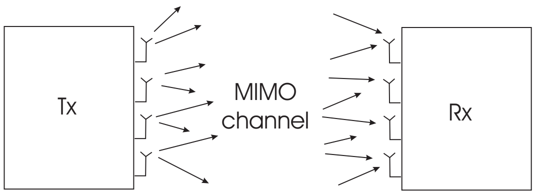

Figure 3.3: Single user MIMO Channel (SU-MIMO).

The performance of the joint pre-coding and joint decoding also depends on the number of transmit and receive antennas. Thus, the minimum antenna configuration provides a single-input single-output (SISO) channel, where pre-coding and decoding reduces to complex scalar multiplications.

3.1.3 Multi-User Channel

If several users / MTs are to be supported by a BS then the separation can be performed in time (TDMA[6]), frequency (FDMA[7]), code (CDMA[8]) and/or spatial domain (SDMA[9]). The first three multi-user techniques are already standard techniques of actual mobile communication systems. FDMA and CDMA are not within the scope of this work and will therefore only slightly touched if necessary for some explanations and to give application examples.

The last multi-user technique (SDMA) can be directly applied due to the use of multiple antennas at the BS and possibly at the MTs (see Fig. 3.5). Throughout this thesis the main emphasis will be on the spatial domain, except when we consider multi-user scheduling techniques. There we use both the spatial and time domain to schedule the transmission from and to all users (sec. 4.5).

Multi-User Multiple Access Channel

In contrast to the more general case discussed above we consider a scenario where several MTs try to transmit to a communication access point over a so called a multiple access channel (MAC). We assume the access point to be a BS therefore in the following we will consider an up-link scenario where the BS has several antennas and the MTs have one or several transmit antennas each.

[6] Time Division Multiple Access
[7] Frequency Division Multiple Access
[8] Code Division Multiple Access
[9] Space Division Multiple Access

Furthermore the BS has full CSI. Since the MTs are decentralized and no joint pre-processing can be performed between the MTs the BS has to perform all the joint signal processing at the receiver side. The only case where the MTs can perform a kind of spatial pre-processing is limited to the case of several transmit antennas at one terminal. The resulting channel is then a MU-MIMO channel which we will not discuss in detail but it should be mentioned for the purpose of completeness of scenarios.

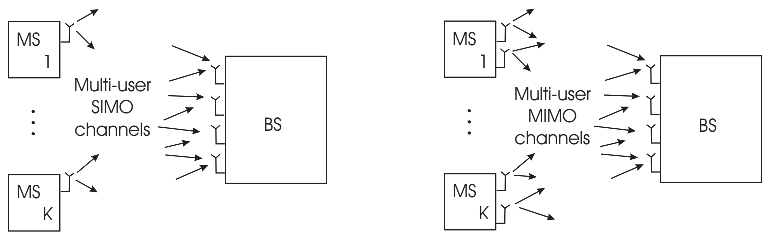

Figure 3.4: Multi-User Multiple-Access Channel left:MU-SIMO MAC, right:MU-MIMO MAC .

Multi-User Broadcast Channel

Corresponding to the up-link MAC scenario the down-link transmission from the BS to several MTs will be referred to as a Multi-User Broadcast Channel (MU-BC). This pays tribute to the fact that in principle all users could receive every message sent by the BS (broadcast). Here, again the MTs can have only one ore several antennas for the reception. Again, the BS has to perform all spatial signal processing since the MTs (one antenna per MT) are not able to perform joint detection and can therefore be low price devices. The necessary CSI to perform spatial pre-processing must be obtained the one or another way e.g. by measuring the channel in the opposite direction and exploiting the channel reciprocity in TDD[10].

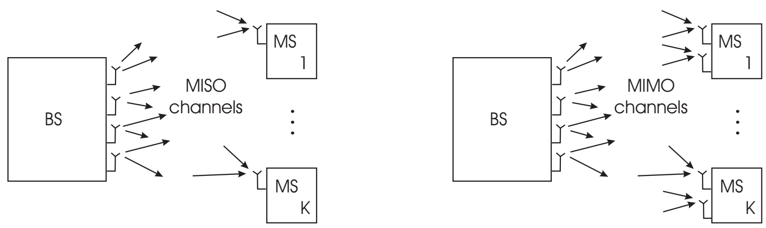

Figure 3.5: Multi-User Broadcast Channel left: MU-MISO BC, right: MU-MIMO BC.

3.1.4 The Statistical Flat-Fading Channel Model

As introduced above the **flat fading model** assumes the symbol length to be much longer than the delay spread of the transmission channel which reduces the channel coefficients to frequency independent scalar complex values. The delay spread describes the maximum width of the temporal distribution of incoming signals belonging all to the same transmitted symbol. The physical reason behind this temporal spread is to be explained by pathes of different length which results in different arrival times after multiplying with the reciprocal of the speed of light in the medium.

[10]Time Division Duplex

In practice the delay spread is calculated only till the last incoming signal peak with reasonable power, all incoming signals after are not considered and will cause inter-symbol-interference (ISI) which does not decrease the system performance if its power is below a acceptable threshold, depending on the transmission system.

The extension is found in the **block fading model** which assumes constant fading coefficients over the whole length of a block and i.i.d. coefficients between two following transmission blocks or channel realizations.

This flat block fading model will be used in the following to evaluate algorithms, predict capacity and / or BER based on numerical simulations. The important statistical component is then always given by the statistical distribution of the channel coefficients, determined by the mean and the variance and whether channel coefficients are i.i.d. or have a distinctive kind of correlation e.g. transmit or receive correlation.

The Rayleigh and Rician Fading Channel

The simplest probabilistic model for the channel filter taps is based on the assumption that there are a large number of statistically independent reflected and scattered paths with random amplitudes in the delay window corresponding to each tap. Note, that with rising measurement bandwidth in channel sounding more taps can be resolved. The phase of i-th path is $2\pi f_c \tau_i$ modulo 2π. Now, $f_c \tau_i = d_i/\lambda$, where d_i is the distance travelled by the i-th path and λ is the carrier wavelength. Since the reflectors and scatterers are far away compared to the carrier wavelength, i.e., $d_i \gg \lambda$, it is reasonable to assume that the phase for each path is uniformly distributed between 0 and 2π and that the phases of different paths are independent. The contribution of each path in the tap gain $h_l(t)$ at time instance t can be modelled as a circular symmetric complex random variable. Each tap $h_l(t)$ is the sum of a large number of such small independent circular symmetric random variables. It follows that the sum of many such independent circular symmetric random variables, according to the Central Limit Theorem can be modelled reasonably as a zero-mean Gaussian random variable. Similarly, because of the uniform phase, $\mathbb{E}\left[h_l(t)\mathrm{e}^{j\phi}\right]$ is Gaussian with the same variance for any fixed ϕ. This assures us that $h_l(t)$ is in fact circular symmetric $\mathcal{CN}(0, \sigma_l^2)$ which means variance $\frac{\sigma_l^2}{2}$ for the real and the imaginary part. It is assumed here that the variance of $h_l(t)$ is a function of the tap l, but independent of time t. With this assumed Gaussian probability density, we know [Xio00, TV04] that the magnitude $r = |h_l(t)|$ of the l-th tap is a Rayleigh distributed with density

$$f(r) = \frac{r}{\sigma_l^2} \exp\left(\frac{-r^2}{2\sigma_l^2}\right), \qquad r \geq 0, \tag{3.3}$$

and the squared magnitude $|h_l(t)|^2$ is χ^2distributed with two degrees of freedom with density

$$f(r^2) = \frac{1}{\sigma_l^2} \exp\left(\frac{-r}{\sigma_l^2}\right), \qquad r \geq 0. \tag{3.4}$$

This model, which is called Rayleigh fading, is quite reasonable for scattering mechanisms where there are many small reflectors, but is adopted primarily for its simplicity in typical cellular situations with a relatively small number of reflectors. The word Rayleigh is almost universally used for this model, but the assumption is that the tap gains are circularly symmetric complex Gaussian random variables.

Since for many scenarios the average path loss can be assumed to be constant, it can be taken out of the channel matrix as a constant scalar factor. A motivation for this can be seen in the separability of any transmission channel. The attenuation which every transmitted signal will suffer over distance or by passing certain media is taken away from the channel coefficient and is treated as a constant scalar which has no implication for the statistical behavior of the system. Hence, the statistical Rayleigh channel will be numerically modelled by channel matrix entries $h(t)$ which are i.i.d. $\mathcal{CN}(0,1)$. This means that between one pair of antennas in average neither attenuation nor amplification is expected.

The statistical behavior of the Rayleigh channel describes a scenario of multi-path transmission without a line-of-sight. The Rayleigh assumption is therefore a suitable model e.g. for an indoor environment.

There is a frequently used alternative model in which a line-of-sight (LOS) path (often called a specular path) is large and has a known magnitude, and that there are also a large number of independent paths. In this case, at least for one value of l the coefficient for the l-th tap, $h_l(t)$, can be modelled as

$$h_l(t) = \sqrt{\frac{K}{K+1}}\sigma_l e^{j\theta} + \sqrt{\frac{1}{K+1}}\mathcal{CN}(0,\sigma_l^2) \tag{3.5}$$

with the first term corresponding to the specular path arriving with uniform phase θ and the second term corresponding to the aggregation of the large number of reflected and scattered paths, independent of θ. The parameter K (so-called K-factor) is the ratio of the energy in the specular path to the energy in the scattered paths; the larger K is, the more deterministic is the channel. The magnitude of such a random variable is said to have a Rician distribution. Its density has quite a complicated form[Xio00] but in many scenarios e.g. with a strong LOS the Rician model reflects the reality more precise than the Rayleigh model.

For numerical simulations the Rician channel will be modelled by channel entries $h(t)$ which are i.i.d. $\mathcal{CN}(K,1)$ with the Rician factor K as a fixed mean. For the composition of a Rician channel which follows (3.5) we construct the Rician channel $\mathbf{H}$ as

$$\mathbf{H} = \sqrt{\frac{K}{K+1}}\mathbf{H}_{LOS} + \sqrt{\frac{1}{K+1}}\mathbf{H}_{NLOS}, \tag{3.6}$$

from a LOS and a non-LOS component. We then obtain

$$K = \frac{\mathbb{E}[\mathrm{trace}(\boldsymbol{H}_{LOS}\boldsymbol{H}_{LOS}^H)]}{\mathbb{E}[\mathrm{trace}(\boldsymbol{H}_{NLOS}\boldsymbol{H}_{NLOS}^H)]} \tag{3.7}$$

for the Rician factor K, using the trace operator

$$\mathrm{trace}(\mathbf{H}\mathbf{H}^H) = \sum_{i=1}^{n_T} \left(\mathbf{H}\mathbf{H}^H\right)_{ii} = \sum_{i=1}^{n_T} |\mathbf{h}_i|^2. \tag{3.8}$$

Rank and Condition Number of the Channel

Another very useful measure is the **rank** of a matrix.

We assume a singular value decomposition of the channel **H**

$$\mathbf{H} = \mathbf{U} \cdot \mathbf{D} \cdot \mathbf{V}^H \tag{3.9}$$

where **U** and **V** are unitary matrices and **D** has only diagonal entries in the upper square sub-matrix. The non-negative real entries on this diagonal are called singular values. The number of singular values which are greater than zero denote **rank**(**H**).

The fraction between the biggest singular value and the smallest non-zero singular value is called the **condition number** of **H** which we denote **cond**(**H**). The condition number gives a measure about the quality ration between the best and the worst sub-channel. This is of importance if an inversion of **H** is needed e.g. for zero forcing detection in a multi-antenna system. A matrix is called singular when some columns or rows are linearly dependent from each other or one column / row can be decomposed as a linear combination of some other columns / rows.

Note, that in practice the number of non-zero SVs is replaced by the number of valid or useful SVs. Here, valid or useful SVs has to be understood under certain side constraints e.g. the dynamic range of a given transmission scheme or the achievable SNR for data transmission. Those factors might limit the number of data streams which can be multiplexed over the transmission channel and thus, can be significant few than non-zero SVs.

3.1.5 Frequency Selectivity of the Transmission Channel

With rising delay spread bandwidth e.g. in outdoor scenarios the channel often becomes much longer than the symbol length [11]. This means that during one symbol length the transmission properties are changing significantly and can no longer be described sufficiently by a single scalar. We then speak of a **frequency selective channel**, where transmission signals of different base-band frequencies face different transmission properties. As a result, the base band signals will then be affected by self-interference or inter-symbol-interference during the transmission which can be mitigated by appropriate signal processing e.g. FIR filters or rake-receiver as used in CDMA[12] systems.

[11] The delay spread of the channel can be approximated by the difference in arrival time of the signal travelling the shortest path, which will be mainly the LOS, and the last relevant signal along the longest path, which e.g. can be a reflection from far away.

[12] Code Division Multiple Access

Another common approach is the slicing of the frequency transmission band into D frequency sub-carriers [13]. Using such multi-carrier or multi-tone transmission schemes e.g. OFDM[14] the channel can be decomposed into D frequency-flat sub-carriers since the overall transmission channel achieves block diagonal structure. Hence, using Fourier Transform (FFT) the received signal at k-th sub-carrier is given by a flat fading MIMO equation

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{n}_k, \tag{3.10}$$

where $\mathbf{x}_k^{N_T \times 1}$ is a zero-mean transmit vector on sub-carrier k, $\mathbf{H}_k^{M_R \times N_T}$ the k-th MIMO channel matrix and $\mathbf{n}_k^{M_R \times 1}$ is circular symmetric zero-mean additive white Gaussian noise with $\mathbb{E}[\mathbf{n}_k \mathbf{n}_k^H] = \sigma_N^2 \mathbf{I}$, where $\cdot^H$ denotes the conjugate transpose and σ_n^2 the noise variance. A cyclic prefix (CP) of length $L_{CP} > L$ assures orthogonality between the sub-carriers and additionally allows a certain timing flexibility at the receiver.

Since we perform the signal processing for each of the sub-carriers separately like in the flat-fading case we can reuse all MIMO algorithms developed for the single carrier flat-fading case. Therefore, all algorithms discussed in this thesis can directly be extended to the frequency selective channel by simply using OFDM techniques.

3.1.6 Equivalent Real-Valued Model

For complex-valued channel models it can turn out to be useful to work with an equivalent real-valued transmission model. Taking the complex-valued model

$$\mathbf{y}_c = \mathbf{H}_c \mathbf{x}_c + \mathbf{n}_c, \tag{3.11}$$

by separating real and imaginary parts we can equivalently write [Tel99]

$$\begin{bmatrix} \Re \mathbf{y}_c \\ \Im \mathbf{y}_c \end{bmatrix} = \begin{bmatrix} \Re \mathbf{H}_c & -\Im \mathbf{H}_c \\ \Im \mathbf{H}_c & \Re \mathbf{H}_c \end{bmatrix} \begin{bmatrix} \Re \mathbf{x}_c \\ \Im \mathbf{x}_c \end{bmatrix} + \begin{bmatrix} \Re \mathbf{n}_c \\ \Im \mathbf{n}_c \end{bmatrix} \tag{3.12}$$

which gives an equivalent $K = 2K_c$-dimensional real model of the form

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{3.13}$$

with the obvious definitions of $\mathbf{y}$, etc.

One reason to use this description is that, if the components of $\mathbf{x}$ are taken from some set of evenly spaced points on the real line, the noiseless received signal $\mathbf{H}\mathbf{x}$ from (3.12) can be interpreted as points in a lattice described by the basis $\mathbf{H}$, and the detection problem can be considered as

[13] Practical systems normally use $D \geq 3L$, where L denotes the order of the frequency selective channel.
[14] Orthogonal Frequency Division Multiplexing

an instance of a *lattice decoding problem* [Win04]. Additionally it was shown by [FW03] that removing the restrictions of complex arithmetics by breaking the coupling between real and imaginary parts can be usefully exploited, e.g. for decision-feedback equalization where the real valued signal precessing allows additional choice for the detection order. Furthermore, real-valued signal processing offers advantages to combat I/Q imbalances in single carrier transmission systems [HFG$^+$04c]. Many low cost systems use direct up- and/or down-conversion (DUC/DDC) techniques to limit sample rates to reasonable values. With imperfections in the I/Q-modulators and demodulators this commonly used technique imposes a severe I/Q imbalance on the end-to-end base-band transmission channel. This signal crosstalk between the real and imaginary part of the signal can be compensated easily by real-valued signal processing.

3.1.7 Signal Constellations

The following gives a description of the signal sets we consider for the components of the complex valued symbols to be transmitted, which we will call $\mathbf{a}_c$, and the corresponding real-valued vector $\mathbf{a}$. Remember that the channel input vector is denoted $\mathbf{x}_c$, and $\mathbf{x}_c = \mathbf{a}_c$ (complex-valued) or $\mathbf{x} = \mathbf{a}$ (real-valued) if no transmit pre-processing is performed, resulting in a one symbol transmission per transmit antenna.

Throughout this thesis we use PAM[15] with square QAM[16] for analytical discussions and numerical simulations and rectangular QAM for the transmission experiments. Following [Hay01] the components of $\mathbf{a}_c = [a_{c,1}, ..., a_{c,n_T}]^T$ are taken from a rectangular grid around zero in the complex plane.

Some signal sets A_c correspond to square M-ary QAM constallations, M=4,16,64,256 (accordingly $R_m = \log_2(M) = 2, 4, 6, 8$ bits are included in one symbol). A real-valued constellation set A can be understood as two sub-sets of a complex-valued constellations projected onto the real and imaginary plane as depicted in Fig. 3.6. Those real-valued signal sets carry then 1,2,3 or 4 bits per symbol, respectively.

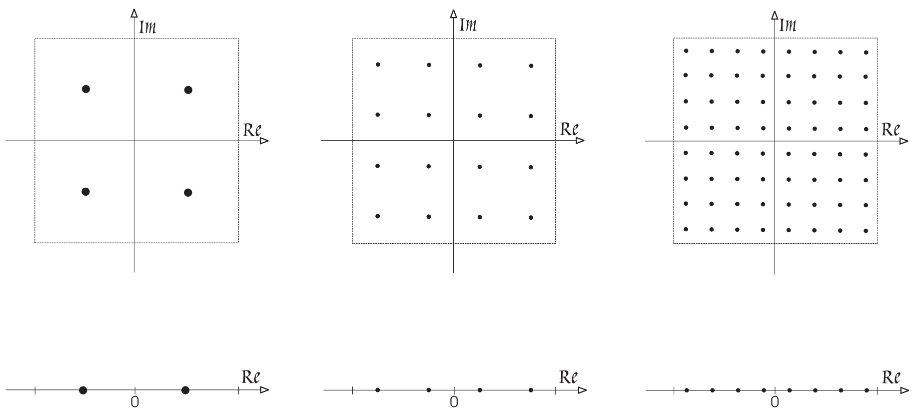

Figure 3.6: Complex-valued signal constellations A_c used for transmission (top) and their projections A onto the real-axis (bottom). Normalization to unit symbol power with uniform probability is assumed. The boundary regions of the constellations are marked by the bounding squares.

[15] Pulse Amplitude Modulation
[16] Quadrature Amplitude Modulation

While it is common to use signal points on the odd-integer grid, $(2a+1)+j(2b+1)$, $a, b \epsilon \mathbb{Z}$, we choose the complex signal points such that the unit average transmit power holds $\mathbb{E}[a_c^H a_c] = 1$ and $\mathbb{E}[a^H a] = 1/2$ for the real-valued case.

For ASK[17] symbols (real-valued) in component a_k, used with uniform probability, the average real-valued constellation point distance d_M has to satisfy

$$\sum_{i=1}^{M/2} ((2i+1)d_M)^2 \cdot \frac{2}{M} = \frac{1}{2} \tag{3.14}$$

with $M = 2, 4, 8, 16$ denoting the number of signal points along the real axis.

For the chosen signal constellations the dominant errors will be caused by symbols distorted to the nearest neighbors of the transmitted symbol. Therefore, Gray labelling will be used as proposed in [J.G00, TAG99] to map the bits to the constellation points such that errors between adjacent symbols only cause 1 bit error. This minimizes the bit error rate for un-coded transmission which will be studied throughout this thesis. Since the nearest neighbors in the complex plane are situated either along the purely real or purely imaginary offset, Gray labelling can be applied independently to the real and imaginary part of the constellation, i.e., the points in A can be Gray labelled regardless of which component of A_c they correspond to.

3.1.8 A mathematical measure of correlation

In order to provide a measure of correlation, we will use the definition used in [E.J04]. This measure will be very useful if e.g. two transmission scenarios are to be compared.

A more detailed introduction following the outline of [E.J04] is given in the appendix of P.2.

We take two arbitrarily chosen transmit correlation matrices[18] $\mathbf{R}_T^{(1)}$ and $\mathbf{R}_T^{(2)}$ with the constraint that $\mathrm{trace}(\mathbf{R}_T^{(1)}) = \mathrm{trace}(\mathbf{R}_T^{(2)}) = n_T$ which is equivalent to

$$\sum_{l=1}^{n_T} \lambda_l^{T,1} = \sum_{l=1}^{n_T} \lambda_l^{T,2}, \tag{3.15}$$

with $\lambda_l^{T,1}$, $1 \leq l \leq n_T$, and $\lambda_l^{T,2}$, $1 \leq l \leq n_T$, are the eigenvalues[19] of the covariance matrix $\mathbf{R}_T^{(1)}$ and $\mathbf{R}_T^{(2)}$, respectively.

This constraint regarding the trace of the correlation matrix $\mathbf{R}_T$ is necessary because the comparison of two transmission scenarios is only valid if the average path loss is equal. Without receive correlation, the trace of the correlation matrix can be written as

$$\mathrm{trace}(\mathbf{R}_T) = \sum_{i=1}^{n_T} \left(\mathbb{E}\left[\mathbf{H}^H\mathbf{H}\right]\right)_{ii} = \sum_{i=1}^{n_T} \mathbb{E}\left[|\mathbf{h}_i|^2\right]. \tag{3.16}$$

[17] Amplitude Shift Keying
[18] transmit correlation matrix $\mathbf{R}_T = \mathbf{H}^H\mathbf{H}$, receive correlation matrix $\mathbf{R}_R = \mathbf{H}\mathbf{H}^H$
[19] $\boldsymbol{\lambda}^T$:vector of eigenvalues of transmit correlation matrix

However, the RHS of (3.16) is the sum of the average path loss from the transmit antenna $i = 1...n_T$. In order to study e.g. the impact of correlation on the achievable capacity separately, the average path loss is kept fixed by applying the trace constraint on the correlation matrices $\mathbf{R}_T^{(1)}$ and $\mathbf{R}_T^{(2)}$.

We will say that a correlation matrix $\mathbf{R}_T^{(1)}$ is more correlated than $\mathbf{R}_T^{(2)}$ with descending ordered eigenvalues $\lambda_1^{T,1} \geq \lambda_2^{T,1} \geq ... \geq \lambda_{n_T}^{T,1} \geq 0$ and $\lambda_1^{T,2} \geq \lambda_2^{T,2} \geq ... \geq \lambda_{n_T}^{T,2} \geq 0$ if

$$\sum_{k=1}^{m} \lambda_k^{T,1} \geq \sum_{k=1}^{m} \lambda_k^{T,2} \quad 1 \leq m \leq n_T - 1. \tag{3.17}$$

The measure of correlation which we will introduce is defined in a natural way: the larger the first m eigenvalues of the correlation matrices are (with the trace constraint in (3.16)), the more correlated is the MIMO channel. As a result, the most uncorrelated MIMO channel has equal eigenvalues, whereas the most correlated MIMO channel has only one non-zero eigenvalue which is given by $\lambda_1 = n_T$.

Before proceeding with our definition of 'more correlated' in terms of the eigenvalue distribution of the channel covariance matrix, we give the necessary definitions we will need in the following.
Definition 1: For two vectors $\mathbf{x}, \mathbf{y} \in R^n$ with descending ordered components $x_1 \geq x_2 \geq ... \geq x_n \geq 0$ and $y_1 \geq y_2 \geq ... \geq y_n \geq 0$ one says that the vector $\mathbf{x}$ majorizes the vector $\mathbf{y}$ and writes

$$\mathbf{x} \succ \mathbf{y} \ \text{ if } \ \sum_{k=1}^{m} x_k \geq \sum_{k=1}^{m} y_k \ \ , m = 1, ..., n-1. \ \text{ and } \ \sum_{k=1}^{n} x_k = \sum_{k=1}^{n} y_k.$$

Using this definition we can compare the correlation of two matrices by comparing there belonging vectors of ordered eigenvalues in the sense of majorization. This means under trace constraint from (3.16) that $\mathbf{R}_T^{(1)}$is more correlated than $\mathbf{R}_T^{(2)}$ if the vector of ordered eigenvalues $\lambda^{T,2}$ is majorized by $\lambda^{T,1}$ ($\lambda^{T,1} \succ \lambda^{T,2}$). The measure of correlation introduced here is limited due to the fact that a comparison of two vectors in the majorization sense is not always possible.

3.1.9 Simulation Setups

The main objective of this thesis will be the evaluation of different transmit and receive configurations, therefore we will concentrate on the single user MIMO channel as discussed in 3.1.2. Depending whether the individual configuration requires joint transmit pre-processing or/and joint receive processing, the results are valid either for the single user MIMO (point-to-point), the multiple access channel (multi-point-to-point) and the broadcast channel (point-to-multi-point).

All simulations will be restricted to direct base-band transmission assuming all RF components to be ideal towards e.g. I/Q-imbalances and non-linearities. Following [Win04] we use the equivalent discrete base-band description [Tre71] where the continuous-time transmit signal is formed using (complex) PAM modulation with a $\sqrt{\text{Nyquist}}$-pulse $g(t)$ of energy E_g.

We will assume for most cases the number of receive antennas to be equal or higher than the number of transmit antennas to profit from the receive diversity in the up-link scenarios, see details in 3.5.1. When down-link scenarios with pre-coding are discussed the number of transmit antennas is assumed to be equal or higher than the number of receive antennas or supported users in a broadcast scenario.

The real $2n_T \times 2m_R$-dimensional channel matrix $\mathbf{H}$ then corresponds to a $n_T \times m_R$ random matrix $\mathbf{H}_c$ of i.i.d. complex Gaussian entries of unit variance and zero mean (Rayleigh channel) or non-zero mean (Rician channel). The normalization of the channel matrix will be mostly such that $\mathbb{E}\left(\text{trace}[\mathbf{H}\mathbf{H}^H]\right) = n_T \cdot m_R$ is satisfied. The real-valued representation was the basis for the implementation in the single carrier system, while the simulations for capacity and BERs use the complex signal representation.

The transmitted power will satisfy a sum power constraint in most cases and capacity and bit error rates are plotted over the average signal-to-noise-ratio per receive antenna (SNR at Rx antenna) under the assumption of equal noise power σ_N^2 from all receive amplifiers. The widely used measure of $\frac{E_b}{N_0}$ which denotes the fraction of the received energy per bit (E_b) and the two-sided power spectral density (N_0) allows a fair comparison of different information transmitting systems but will be of minor importance for this work.

In this work we will use the average SNR per Rx antenna which is a practical measure and can be directly measured during experiments.

3.2 Channel Capacity

A very important measure of a transmission system is given by the mutual information or the channel capacity [Sha48] which gives a measure how many bits per time and frequency bandwidth can be transmitted in a certain channel. Transmission systems with a high spectral efficiency e.g. multi-antenna (MIMO) systems can transmit more bits per time and frequency band than e.g. SISO systems.

In the following we always assume a sum power constraint $\sum_{i=1}^{n_T} p_i = P$ at the transmit side which means that one Tx antenna can use the full sum power for a transmission if the others are not active. The noise power σ_N^2 per Rx antenna is assumed to equal for all antennas.

Starting from the SISO channel capacity [Sha48]

$$C_{SISO} = \log_2 (1 + SNR) \tag{3.18}$$

We extend this to the general frequency selective (FS) MIMO channel

$$C_{MIMO}^{FS} = \frac{1}{B} \int_B \log_2 \det \left[\mathbf{I} + \mathbf{G}(f)\mathbf{\Phi}(f)\mathbf{G}(f)^H \mathbf{K}^{-1}(f)\right] df \tag{3.19}$$

$\mathbf{G}$: channel response matrix, not normalized ($m_R \times n_T$)
$\mathbf{K}$: covariance matrix of impairment ($m_R \times m_R$)
B: bandwidth

$\mathbf{\Phi}$: covariance matrix of transmit signal ($n_T \times n_T$) with $\mathbf{\Phi} = \mathbb{E}[\mathbf{x} \cdot \mathbf{x}^H]$ and the sum transmit power P is held constant.

If we limit ourselves to the flat fading case and require the impairment to be Gaussian, then the integration and its averaging effects disappear

$$C_{MIMO} = \log_2 \det \left[\mathbf{I} + \mathbf{G\Phi G}^H \mathbf{K}^{-1}\right]. \tag{3.20}$$

When entries of $\mathbf{G}$ are independent, the open-loop capacity (CSI is known to the Rx only) is maximized by transmitting Gaussian signals with covariance

$$\mathbf{\Phi} = \frac{P}{n_T}\mathbf{I} \tag{3.21}$$

with P the total radiated sum power.

If entries of $\mathbf{G}$ have same variance (g), we can define a unit-variance normalized channel matrix $\mathbf{H}$ so that

$$C_{MIMO} = \log_2 \det \left[\mathbf{I} + \frac{P}{n_T} g \mathbf{HH}^H \mathbf{K}^{-1}\right]. \tag{3.22}$$

When the impairment consists exclusively of thermal noise $\mathbf{K} = \sigma_N^2 \mathbf{I}$ we find the noise-limited open loop capacity to be

$$C_{MIMO} = \log_2 \det \left[\mathbf{I} + \frac{Pg}{\sigma_N^2} \frac{1}{n_T} \mathbf{HH}^H\right]. \tag{3.23}$$

where $\frac{Pg}{\sigma_N^2}$ denotes the average SNR at any Rx antenna.

The different levels of randomness in the channel can be separated into:

- Large-scale randomness e.g. distance dependence , shadowing, etc. can be absorbed into SNR and regarded as deterministic within a local area.

- Small-scale randomness caused by multi-path will be contained within $\mathbf{H}$.

The asymptotic capacities will then increase with SNR depending on the number of transmit or/and receive antennas assuming the channel to have full rank at least in average and a sum power constraint at the transmitter.

Increasing *SNR* with symmetric antenna numbers $n_T = m_R = n$

$$C = n \log_2 [1 + SNR]. \tag{3.24}$$

Increasing *number of transmitters* , n_T, with m_R constant, $n_T \geq m_R$

$$C = m_R \log_2 [1 + SNR]. \tag{3.25}$$

Increasing *number of receivers* , m_R, with n_T constant, $m_R \geq n_T$

$$C = n_T \log_2 \left[1 + \frac{m_R}{n_T} SNR \right]. \tag{3.26}$$

From (3.24) to (3.26) we see that the asymptotic capacity slope is determined by $\min(n_T, m_R)$ therefore it makes no sense to increase the number of Tx antennas without inceasing the numbers of receive antennas at the same time since the total transmit power is limited, therefore the SNR can not be changed by this means. An increase in receive antennas will improve the SNR, since the effective size of the antenna array is increased and more signal energy is available for signal processing. Then the final capacity slope is determined by the number of Tx antennas.

We can conclude that considering the asymptotic behavior of the MIMO capacity it would be most beneficial towards spatial multiplexing to distribute a certain total number of antennas between two communication points such that the whole antenna configuration becomes ideally symmetric, or in case of an odd number of antennas, place one more antenna at the receiver side to increase the SNR.

Let us now consider a narrow band (flat fading scenario) with multi-path fading (Rayleigh) and the large scale randomness (distance, shadowing, Rician factor etc.) to be held fixed. Furthermore the total power constraint holds and different data streams are transmitted from each Tx antenna.

The capacity formula by [Fos96] describes how much capacity is available in a certain MIMO channel realization.

$$C_{n_T, m_R} = \log_2 \det \left[\mathbf{I}_{mR} + (\rho / n_T) \mathbf{H} \mathbf{H}^H \right] \text{bit/s/Hz}. \tag{3.27}$$

where $\rho = \frac{Pg}{\sigma_N^2}$ is the average SNR at any Rx antenna under the assumption of i.i.d. noise at any Rx antenna with variance σ_N^2 as introduced in (3.23). Then (3.27) can be reformulated as

$$C = \sum_{i=1}^{n_T} \log_2 \left[1 + \frac{\rho}{n_T} \lambda_i^2 \right] \tag{3.28}$$

where λ_i are the singular values of the parallel sub-channels of $\mathbf{H}$ which describe the properties of the parallel spatial sub-channels. λ_i can be obtained by singular value decomposition (see section 5.2.2) or λ_i^2 can be obtained directly by eigenvalue decomposition.

We see that the distribution of the singular values (SVs) / eigenvalues (EVs) is very important for the achievable channel capacity especially in the low to mid SNR region (below 20 dB) where many applications will be working. Therefore it is of great interest to us to know how we can influence the distribution of the singular values by an appropriate system design. If we consider a per channel normalization to unit variance we can evaluate the best and worst case scenario. For a better illustration we plot the capacity curves per channel realization versus the SNR for a fixed number of Tx and Rx antennas ($n_T = m_R = 8$) in Fig. 3.7. The best curve (red, most left) belongs to the channel where all SVs are the same ($\mathbf{H}$ has diagonal form) and the capacity grows with 8 bits/s/Hz per 3 dB SNR increase at high SNR. The worst channel (e.g. dyad or key-hole channel [GAY$^+$02, JPN$^+$02]) (most right curve, black) belongs to a SV-distribution where only one SV is non-zero , therefore the capacity only grows with 1 bit/s/Hz per 3 dB SNR increase even at high SNR. The green dashed curve denotes the ergodic (average) capacity over 10^5 random

Rayleigh distributed channel realizations.

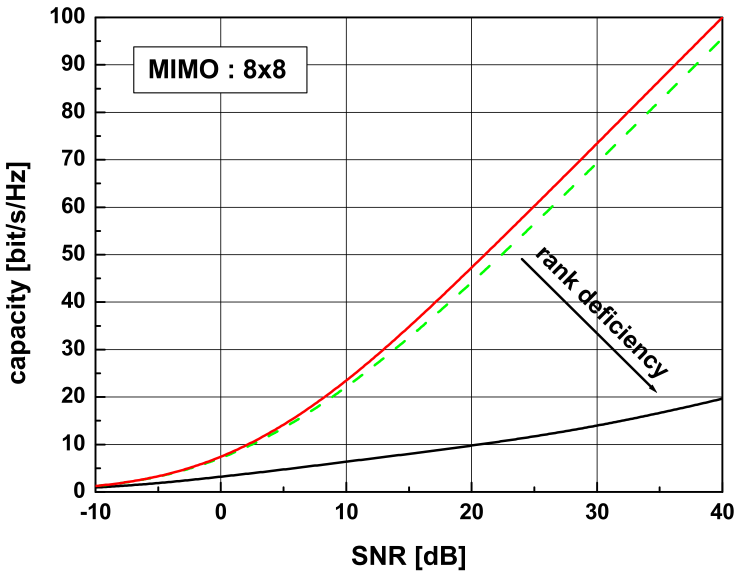

Figure 3.7: Channel capacities for different channel realizations under trace constraint. Red curve left: all SVs are identical, green dashed curve middle: the average capacity over 10^5 Rayleigh channel realizations, black curve right: only one SV is non-zero.

In reality we will have a mixture of many channel realizations which can follow e.g. Rayleigh or Rician statistics, therefore the ergodic (average) or outage capacity[20] are more adequate measures, compare also Fig. 3.13 and Fig. 3.14.

We now consider closed-loop transmission systems where channel state information is available at the transmit side as well. Again, a total transmit power constraint is assumed. Then the optimum strategy which maximizes the capacity is a transmission of each independent data stream into the direction of one of the eigenvectors of the channel matrix $\mathbf{H}$ [Tel99, E.J04, HB03].

Now the capacity formula from (3.27) changes to

$$C_{n_T,m_R} = \log_2 \det \left[\mathbf{I}_{mR} + (\rho/n_T)\mathbf{HQH}^H\right] \text{ bit/s/Hz.} \tag{3.29}$$

where $\mathbf{Q}$ denotes the transmit correlation matrix which can be decomposed into

$$\mathbf{Q} = \mathbf{V} \cdot \mathbf{D}_Q \cdot \mathbf{V}^H \tag{3.30}$$

where V is a unitary beam-forming matrix derived from SVD ($\mathbf{H} = \mathbf{UDV}^H$) which arranges the transmission of the data streams into the eigenmodes of the channel and a weight matrix $\mathbf{D}_Q$ which weighs each data stream towards the percentage received from the total transmit power.

[20] An outage capacity of 1 % means that only 1 % of the channels do not carry the specified capacity, therefore outage will occur with probability 0.01.

We now find [EB99]

$$\begin{aligned} C_{n_T,m_R} &= \log_2 \det\left[\mathbf{I}_{mR} + (\rho/n_T)(\mathbf{UDV}^H)(\mathbf{VD}_Q\mathbf{V}^H)(\mathbf{VD}^H\mathbf{U}^H)\right] & (3.31)\\ &= \sum_{i=1}^{\min(n_T,m_R)} \log_2\left[1+\frac{\rho}{n_T}\cdot|d_{ii}|^2\cdot\lambda_i^2\right] \text{bit/s/Hz.} & (3.32)\end{aligned}$$

where $|d_{ii}|^2 = p_i$ which is the power used for each data stream and λ_i the singular values of the channel.

The optimum power allocation is the well known water-filling solution which also includes the beam-forming case at low SNR, where only one data stream is transmitted with full power from all antennas and the high SNR case where uniform power allocation is optimum.

The water-filling solution can be described by two sets of equations

$$\frac{\partial C(\mathbf{p})}{\partial p_i}_{|_{\mathbf{p}=\mathbf{p}^{opt}}} = const, i = 1...n_T, \text{ and } \sum_{i=1}^{\min(n_T,m_R)} p_i^{opt} = P \quad (3.33)$$

or

$$p_i^{opt} = \left[\mu - \frac{n_T}{\rho}\frac{1}{\lambda_i^2}\right]^+ \text{ with } \mu \text{ such that } \sum_{i=1}^{\min(n_T,m_R)} p_i^{opt} = P \quad (3.34)$$

where $a^+ = max(a,0)$ is the positive part of a.

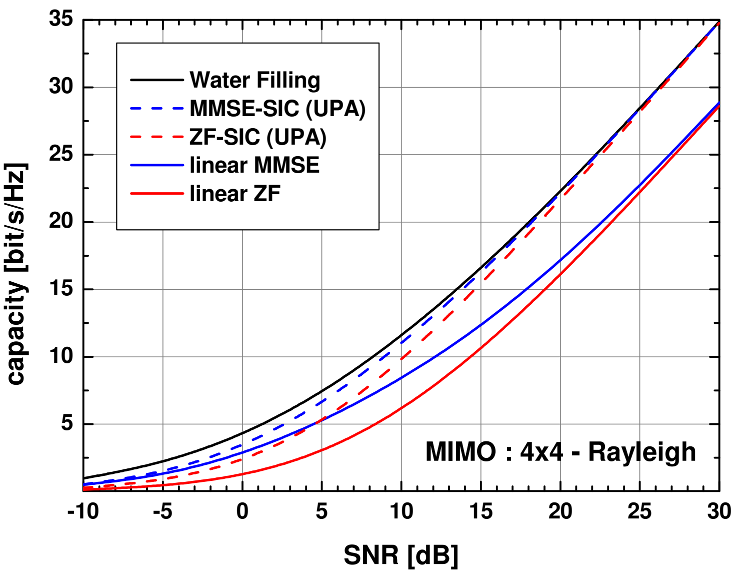

Figure 3.8: Achievable spectral efficiency / capacity with different transmission schemes in a Rayleigh flat fading channel with 4 Tx and 4 Rx antennas.

Fig. 3.8 shows the achievable capacity for different transmission and detection schemes[21]. The optimum transmission strategy with CSI at the Tx and Rx is the water-filling solution from (3.34) represented by the black solid line. Multiplexing one data stream per antenna and applying successive interference cancellation (SIC) at the Rx is known to be capacity achieving in the high SNR region as well [VG97]. Unfortunately this is only valid under the assumption of negligible error propagation during the successive detection process which is not valid in real applications. Therefore we have to expect a rightward shift of the capacity curves with SIC, leaving a gap between the spectral efficiency of the SVD scheme and SIC detection even at high SNR. For the linear detection schemes (solid lines red and blue) we observe a throughput loss which increases with the number of parallel data streams. We also observe that under the perspective of capacity that the achievable capacities with ZF and MMSE converge at high SNR in contrast to BER plots with a fixed rate where MMSE always outperforms ZF at any reasonable SNR value.

3.3 Bit Error Rates

Another way to characterize the performance of a transmission scheme is an evaluation of the fraction of bits, symbols or frames which are detected correctly after transmission of a channel. This can be done using the measures bit-error-rate (BER), symbol-error-rate (SER) or frame-error-rate (FER).

Following the argumentation and notations in [TV04] this gives an example for the transmission with BPSK symbols over one single channel (AWGN or Fading) to show that e.g. the bit-error probability can be calculated for certain statistical channel models. Knowing the channel gain, coherent detection of BPSK can be performed on a symbol by symbol basis.

$$y = hx + n \tag{3.35}$$

The detection of x from y can be done exactly as in the AWGN case, except that the decision is now based on the sign of the real sufficient statistic. If the transmitted symbol is $x = \pm a$, then for a given value of h, the error probability of detecting x is

$$Q\left(\frac{a|h|}{\sqrt{N_0/2}}\right) = Q\left(\sqrt{2|h|^2 SNR}\right) \tag{3.36}$$

where $SNR = a^2/N_0$ is the average received signal-to-noise ratio per symbol time and Q as Q-function. The channel gain was normalized such that $\mathbb{E}[|h|^2] = 1$. We average over the random gain h to find the overall error probability. For Rayleigh fading when $h \sim \mathcal{CN}(0,1)$, direct integration yields

$$p_e = \mathbb{E}[Q\left(\sqrt{2|h|^2 SNR}\right)] = \frac{1}{2}\left(1 - \sqrt{\frac{SNR}{1+SNR}}\right) \tag{3.37}$$

At high SNR, we get the approximation

$$p_e \approx \frac{1}{4SNR} \tag{3.38}$$

[21] UPA: Uniform Power Allocation

which decays inversely proportional to the SNR. This is in contrast with AWGN where the error probability decays exponentially with the SNR, therefore e.g. at an error probability of 10^{-3}, there is a 17 dB difference between the performance on the AWGN channel and coherent detection on the Rayleigh fading channel. This difference in the required SNR to attain the same error probability will be used in the thesis to compare schemes. This corresponds to a horizontal gap the SNR curves of the two schemes at the same error probability.

The main reason why detection in fading channel has poor performance is due to the fact that the channel gain is random and there is a significant probability that the channel is in a "deep fade". At high SNR, we can in fact be more precise about what a "deep fade" means by inspecting (3.36). The quantity $|h|^2 SNR$ is the instantaneous received SNR. Under typical channel conditions, i.e., $|h|^2 SNR \gg 1$, the conditional error probability is very small, since the tail of the Q-function decays very rapidly. In this regime, the separation between the constellation points is much larger than the standard deviation of the Gaussian noise. On the other hand, when $|h|^2 SNR$ is of the order of 1 or less, the separation is of the same order as the standard deviation of the noise and the error probability becomes significant. The probability of this event is

$$\mathbb{P}\{|h|^2 SNR < 1\} = \int_0^{1/SNR} e^{-x} dx = \frac{1}{SNR} + O(\frac{1}{SNR^2}). \tag{3.39}$$

This probability has the same order of magnitude as the error probability itself (3.38). Thus, we can define a "deep fade" via an order-of-magnitude approximation

$$\begin{aligned} \text{deep fading event: } |h|^2 &< \frac{1}{SNR} \\ \mathbb{P}\{\text{deep fading}\} &\approx \frac{1}{SNR}. \end{aligned}$$

[TV04] conclude that high-SNR error events most often occur because the channel is in deep fade and not as a result of the additive noise being large. In contrast, in the AWGN channel the only possible error mechanism is for the additive noise to be large. Thus, the error probability performance over the AWGN channel is much better.

3.4 Diversity versus Multiplexing Trade-off

Following the basic idea of [ZT03] which investigated the fundamental trade-off between multiplexing gain and diversity gain for any multi-antenna system for SNR$\rightarrow \infty$ we concentrate on the definitions and results which are important for this thesis.

Zheng and Tse [ZT03] argue that multiple antenna channels provide spatial diversity, which can be used to improve the reliability of the link. The basic idea is to supply the receiver with multiple independently faded replicas of the same information symbol, so that the probability that all the signal components fade simultaneously is reduced. As an example, they consider uncoded BPSK signals over a single antenna fading channel ($n_T = m_R = 1$). We known [J.G00] that the bit error probability at high SNR (averaged over the fading gain of $\mathbf{H}$ as well as the additive noise) is

$$P_e(\text{SNR}) \approx \frac{1}{4}(\text{SNR})^{-1}. \tag{3.40}$$

In contrast, transmitting the same signal to a receiver equipped with 2 antennas, the bit error probability is

$$P_e(\mathrm{SNR}) \quad \approx \quad \frac{3}{16}(\mathrm{SNR})^{-2}. \tag{3.41}$$

Having the extra receive antenna, the error probability decreases with SNR at a steeper slope than $(\mathrm{SNR})^{-2}$. This phenomenon implies that at high SNR, the error probability is much smaller. Similar results can be obtained if the modulation is changed to other signal constellations. Since the performance gain at high SNR is determined by the SNR exponent of the error probability, this exponent is called the **diversity gain**. It corresponds to the number of independently faded pathes that a symbol passes through; in other words, the number of independent fading coefficients that can be averaged over to detect a symbol. In a general system with n_T transmit and m_R receive antennas, there are in total $m_R \times n_T$ random fading coefficients, therefore the maximal diversity gain provided by the channel is $m_R \cdot n_T$.

Besides providing diversity to improve reliability, multiple antenna channels can also support a higher data rate than single antenna channels. Let us consider an ergodic block fading channel. Then the ergodic capacity (bits/s/Hz) of this channel [Tel95, Fos96] for high SNR grows with

$$C_{MIMO}(\mathrm{SNR}) \quad \approx \quad M \cdot \log_2(\mathrm{SNR}); \text{ with } M = \min(n_T, m_R) \tag{3.42}$$

in contrast to $\log_2(\mathrm{SNR})$ as for single antenna channels. This result suggests that the multiple antenna channel can be viewed as M parallel spatial channels and the number $M = \min(n_T, m_R)$ is the total number of degrees of freedom to communicate. Now one can transmit independent information symbols in parallel through the spatial channels. This idea is also called **spatial multiplexing**.

Since in any realistic scheme we achieve only a fraction of the actual capacity we say that a scheme achieves a spatial multiplexing gain of r if the supported data rate follows

$$R(\mathrm{SNR}) \quad \approx \quad r \log_2(\mathrm{SNR})[bit/s/Hz] \tag{3.43}$$

at high SNR.

[ZT03] then introduces the following definitions which we will also use despite the fact that they are mainly useful for SNR $\to \infty$ and e.g. fixed rate systems might have no multiplexing gain at all at high SNR, according to this definition.

Definition 2: A multi-antenna transmission scheme is said to achieve spatial multiplexing gain r and diversity gain d if the data rate

$$\lim_{SNR \to \infty} \frac{R(SNR)}{\log_2 SNR} = r \tag{3.44}$$

and average error probability

$$\lim_{SNR \to \infty} \frac{\log_2 (P_e(SNR))}{\log_2 SNR} = -d. \tag{3.45}$$

For $T \geq m_R + n_T - 1$ [ZT03] then concludes the following theorem about the optimum trade-off curve as the main result which we will use in this thesis.
Theorem 1: Assume $T \geq m_R + n_T - 1$. Let $M = \min(n_T, m_R)$: The optimal trade-off curve $d^\star(r)$ is given by the piecewise linear function connecting the points $(k; d^\star(k)); k = 0, ..., M$, where

$$d^\star(k) = (n_T - k)(m_R - k) \tag{3.46}$$

and k is the number of multiplexed data streams. As a consequence $d^\star_{max} = m_R \cdot n_T$, and $r^\star_{max} = \min(n_T, m_R)$. The function $d^\star(r)$ is plotted in Fig. 3.9.

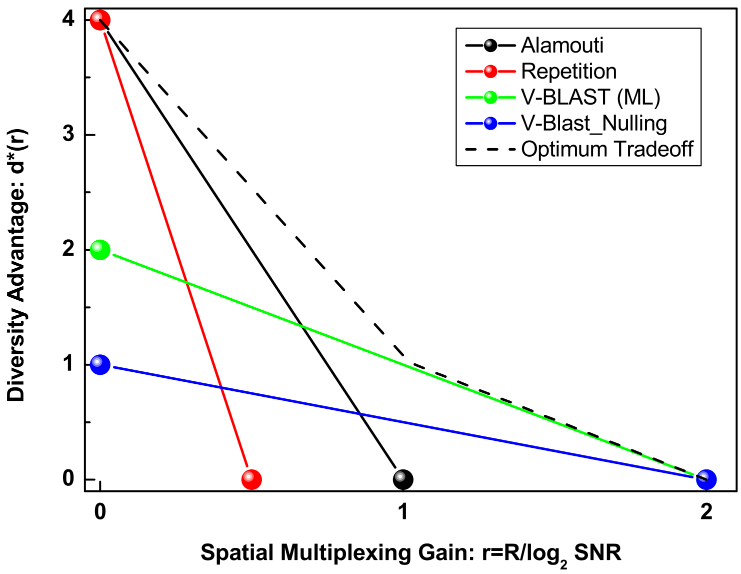

Figure 3.9: The diversity-multiplexing tradeoff of the 2×2 i.i.d. Rayleigh fading MIMO channel along with those of four schemes.

The optimal trade-off curve intersects the r axis at $M = \min(n_T, m_R)$. This means that the maximum achievable spatial multiplexing gain $r^\star_{max} = M$, which is the total number of degrees of freedom provided by the channel, as suggested by the ergodic capacity result in (3.44). The results are to be understood as gain compared to a SISO channel.

The theorem says that at this point, however, no extra positive diversity gain can be achieved. Intuitively, as $r \to M$, the data rate approaches the ergodic capacity and there is no protection against the randomness in the fading channel. On the other hand, the curve intersects the d axis at the maximal diversity gain $d^\star_{max} = n_T \cdot m_R$, corresponding to the total number of random fading coefficients that a scheme can average over. There are known designs that achieve the maximal diversity gain at a fixed data rate [Ala98]. The theorem says that in order to achieve the maximal diversity gain, no positive spatial multiplexing gain can be obtained at the same time. The optimal trade-off curve $d^\star(r)$ bridges the gap between the above two design criteria, by connecting the two extreme points: $(0; d^\star_{max})$ and $(r^\star_{max}; 0)$. This result says that positive diversity gain and spatial multiplexing gain can be achieved simultaneously. However, increasing the diversity advantage comes at a price of decreasing the spatial multiplexing gain, and vice versa. The trade-off curve is thus a more complete concept than the two extreme points corresponding to the maximum diversity gain and maximum multiplexing gain. For example, the ergodic capacity result suggests that by increasing the minimum of the number of transmit and receive antennas, $M = \min(n_T, m_R)$, by one, the channel gains one more degree of freedom, corresponds to $r^\star_{max}$ is increased by 1. The theorem makes here a more informative statement: if we increase both n_T and m_R by 1, the entire trade-off curve is shifted to the right by 1, as shown in Fig. 3.10; i.e., for a given diversity gain requirement d, the supported spatial multiplexing gain is then increased also by 1.

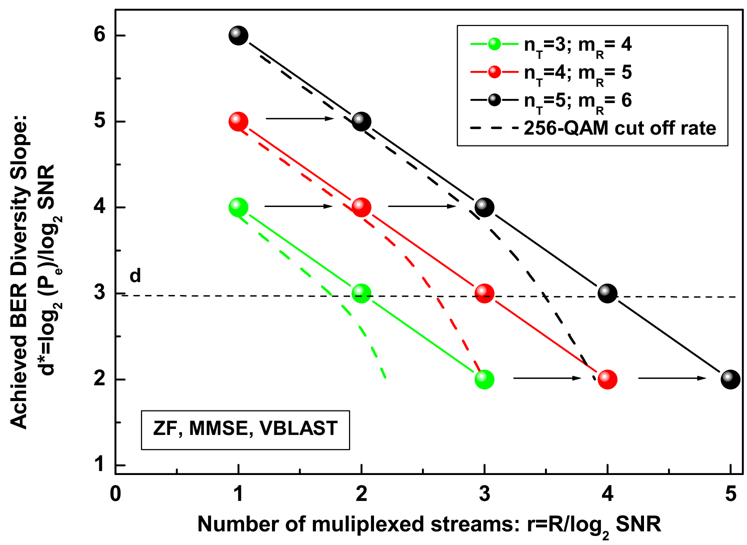

Figure 3.10: Diversity versus Multiplexing Trade-off curve $d^\star(r)$ for $m_R = n_T + 1$ with $n_T = 3$ (green), $n_T = 4$ (red) and $n_T = 5$ (black) assuming Rx detectors like ZF, MMSE or VBLAST. The bullets belong to the transmission schemes with an unlimited number of modulation levels, while the dotted line shows the behavior achieved with a rate saturated at 256-QAM.

For a better understanding of the operational meaning of the trade-off curves, we have a closer look at an example depicted in Fig. 3.10. Here, we assume uncoded transmission with independent data streams from each antenna and all MIMO signal processing is performed at the Rx with either a ZF-, MMSE- or VBLAST-detector. The trade-off curve now has the direct dependance between the achieved data stream multiplexing and the available antenna diversity when using all receive antennas. Furthermore, we assume that from one to n_T data streams are transmitted from the transmit antennas in a BLAST like fashion[22]. Here, we do not consequently follow the notation of [TV04] which defined their diversity advantage and multiplexing gain always compared to a

[22]BLAST transmission mode: one data stream is transmitted from one transmit antenna

SISO system. This means that a diversity order of 1 for the BER equals a diversity advantage of zero against a SISO system. We see that adding one more antenna at both sides of the link still increases the multiplexing gain by one and leaves the diversity gain unchanged. If a limitation of the modulation levels becomes effective which has to be expected for a real transmission system, we will not obtain the maximum multiplexing gain as predicted by the solid lines in Fig. 3.10. This can be explained by the fact that before the sum rate curve reaches its theoretical maximum slope, the cut-off rate limits a further accent and the theoretical expected slope is not reached to the full extend. Transmission experiments conducted in our lab are in good accordance with the dashed lines in Fig. 3.10 which give the diversity - multiplexing curve which is achievable in reality with the above mentioned linear and non-linear detectors (see also Fig. 6.27).

Summarizing the results from this section we always have to keep in mind that in no way we can maximize the number of multiplexed data streams and the antenna diversity gain at the same time. We have to trade between the two extremes in a reasonable way. This basic result is reflected in the following section on considerations on the design parameters of WLAN systems.

3.5 Considerations for the Application in WLAN Systems

In the this section we discuss some basic parameters and their influence on the performance of a MIMO system. This gives a very useful insight into the relevance of certain parameters for a dedicated application and might help the system designer to optimize the overall performance of the whole system.

The discussion will be underlined with figures for a better illustration. Throughout the next three subsections we use the following system model assumptions.

We assume a MIMO transmission in a flat block-fading multi-antenna channel. Within the Rician model for the fading channel, we describe $\mathbf{H}$ as a composition of a line of sight (LOS) component and a non-line of sight (NLOS) component [DF99].

$$\mathbf{H} = \sqrt{\frac{K}{K+1}}\mathbf{H}_{LOS} + \sqrt{\frac{1}{K+1}}\mathbf{H}_{NLOS} \tag{3.47}$$

As an example, the LOS component is modelled with a circular transmitter (Tx) array and a uniform linear receiver (Rx) array which faces broad-side to the Tx (see Fig. 3.11). All antennas have the same polarization and are adjusted perpendicular to one plane. The radiation pattern is assumed to be omni-directional and the path-loss is proportional to $1/d^2$ (d : distance between the Tx and Rx arrays). The NLOS component is modelled as a Rayleigh channel with the matrix entries being complex Gaussian having zero mean and unit variance.

The composite Rician channel will be characterized by the effective Rician factor K which we define as the expectation of the ratio of the LOS-power and the NLOS-power averaged over all

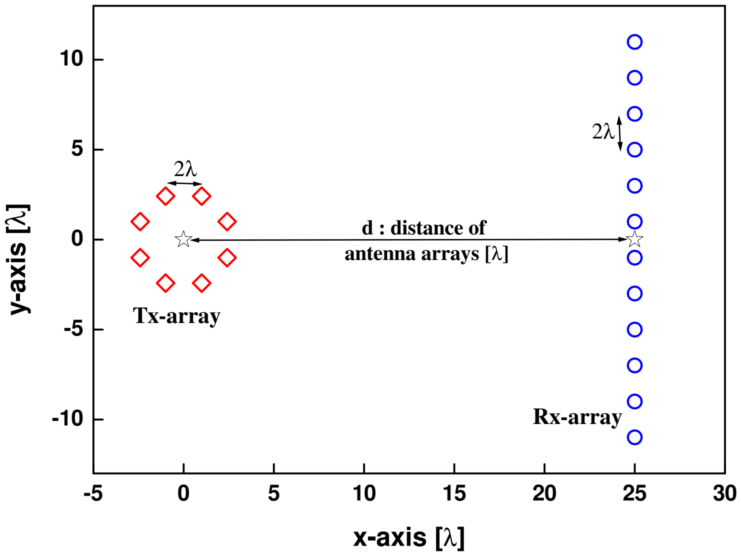

Figure 3.11: Antenna configuration for the LOS model Tx-antenna: circular, Rx - antenna: linear-broad-side.

Tx and Rx antennas similar as in (3.7).

$$K = \frac{\mathbb{E}\left[\sum_{i,j=1}^{n_T,m_R} |h_{ij}^{LOS}|^2\right]}{\mathbb{E}\left[\sum_{i,j=1}^{n_T,m_R} |h_{ij}^{NLOS}|^2\right]} = \frac{\mathbb{E}[\mathrm{trace}(\mathbf{H}_{LOS} \cdot \mathbf{H}_{LOS}^H)]}{\mathbb{E}[\mathrm{trace}(\mathbf{H}_{NLOS} \cdot \mathbf{H}_{NLOS}^H)]}. \tag{3.48}$$

In the following it is assumed that the transmit symbols satisfy a long term unit power constraint

$$\mathbb{E}[\mathbf{x}^H \cdot \mathbf{x}] = 1. \tag{3.49}$$

Two different constraints will be applied on total power transferred in average over the transmission channel. We take the expectation of the transmitted power and apply

1. a total power constraint on $\mathbf{H}$:

$$\mathbb{E}\left[\mathrm{trace}(\mathbf{H} \cdot \mathbf{H}^H)\right] = \mathrm{const}$$

to discuss systems using a Tx power control. This is motivated by the fact that a practical system will try to control the average received power with a feed back to the Tx to limit the necessary dynamic range at the Rx and to reduce the interference to other users as well.

2. a Rayleigh power constraint on $\mathbf{H}_{NLOS}$:

$$\mathbb{E}\left[\mathrm{trace}(\mathbf{H}_{NLOS} \cdot \mathbf{H}_{NLOS}^H)\right] = \mathrm{const}$$

to discuss the effect of sudden blocking or adding of a LOS in a transmission scenario without or with a very slow power control. Practically, blocking the LOS simply reduces the power at the Rx, while the received power from the NLOS component is almost unchanged.

In order to see how the parameters of interest have influence on a MIMO system we do numerical system simulations. We assume Rayleigh and Rician channels and look at

1. the distribution of the singular values of the channel matrix $\mathbf{H}$.
2. the capacity of $\mathbf{H}$.
3. the BERs using fixed rate (QPSK) per antenna, no error protection coding and a linear Detector (Zero-Forcing) at the Rx.

3.5.1 Antenna Diversity

Distribution of the Singular Values:

Assuming a pure Rayleigh channel $\mathbf{H}$ with no LOS we investigate the distribution of the ordered SVs (λ_i) of $\mathbf{H}$ obtained by SVD of 10^5 random channel matrices.

Fig 3.12 shows the computed probability density function (pdf) of the ordered SVs for a MIMO system having 8 Tx antennas and 8 / 12 Rx antennas, respectively. The numerical results are consistent with an analytical formula (3.50) for the ordered eigenvalues $\Lambda_i = \lambda_i^2$ given by [Tel95] and [Ede89]

$$p_{\Lambda,ord}(\Lambda_1, ..., \Lambda_m) = \frac{1}{K_{m,n}} \prod_i e^{-\Lambda_i} \Lambda_i^{n-m} \prod_{i<j} (\Lambda_i - \Lambda_j)^2, \qquad \Lambda_1 \geq ... \geq \Lambda_m \geq 0 \tag{3.50}$$

where $K_{m,n}$ is a normalizing factor.

With 8 Rx and 8 Tx antennas, the distributions are well separated from each other and there is only a slight increase of the distance between the maxima from the smallest to the highest SV. From this unique property of the ordered SV distribution, a Rayleigh-like channel can quickly be identified, when measured MIMO channels are inspected, for instance.

The more additional antennas are used at one side of the link, the more all SVs are shifted to higher values. This is due to the higher signal power received by using either receive diversity or transmit diversity. Additional antennas successfully combat the fading. Note that that the smallest SV has the largest relative shift when compared to the case of 8 Rx antennas. The condition number $\mathrm{cond}(\mathbf{H}) = \lambda_{\max}/\lambda_{\min}$ is therefore significantly reduced.

Capacity and Antenna Diversity:

The average capacity calculated from 10^5 random channels is shown in Fig. 3.13. It is observed that no additional multiplexing gain is achieved if antennas are added only at one side of the link. Following the definition given in (3.44) [ZT03] a transmission scheme achieves a maximum spatial

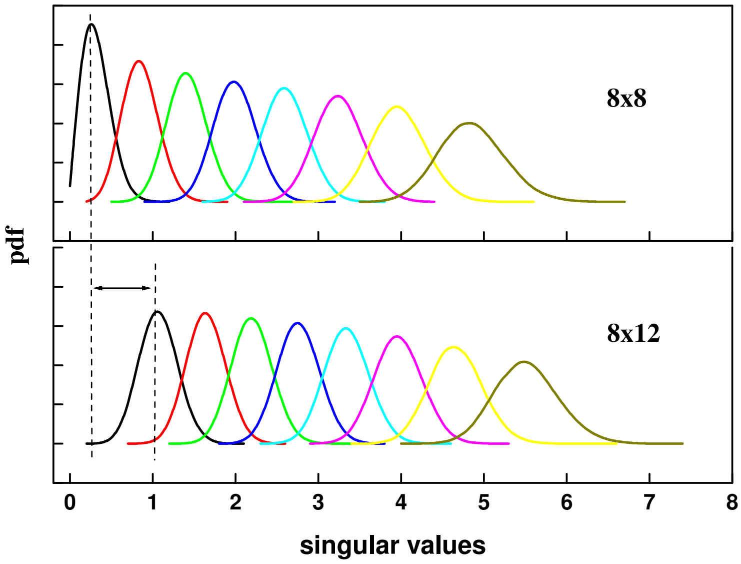

Figure 3.12: Pdf of the ordered singular values for a MIMO transmission matrix with 8x8 and 8x12 antennas.

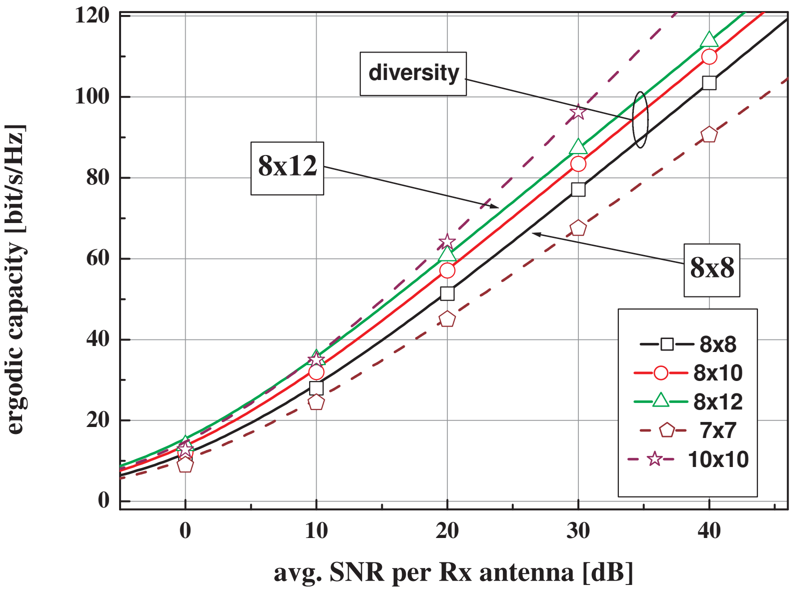

Figure 3.13: Average Capacity for various MIMO configurations in the Rayleigh channel plotted over average SNR at Rx antenna.

multiplexing gain $r = \min(n_T, m_R)$ if $\mathbf{H}$ has full rank. For a sufficiently large SNR we observe a slope of 8 bits/s/Hz every 3 dB for the MIMO systems using 8 Tx antennas in Fig. 3.13. This slope r which is also called effective degree of freedom (edof) [CFV$^+$00] is a very important parameter. A MIMO system has therefore to be adaptive to the edof otherwise much of the effort of parallel transmission will be wasted. In indoor scenarios [KWV00], [JPN$^+$02] almost the Rayleigh capacity was found and even in multi-scattering urban areas like Down-town Manhattan [LFV01a] about

80% of the Rayleigh capacity were substantial when 16 antennas were used at both sides of the link.

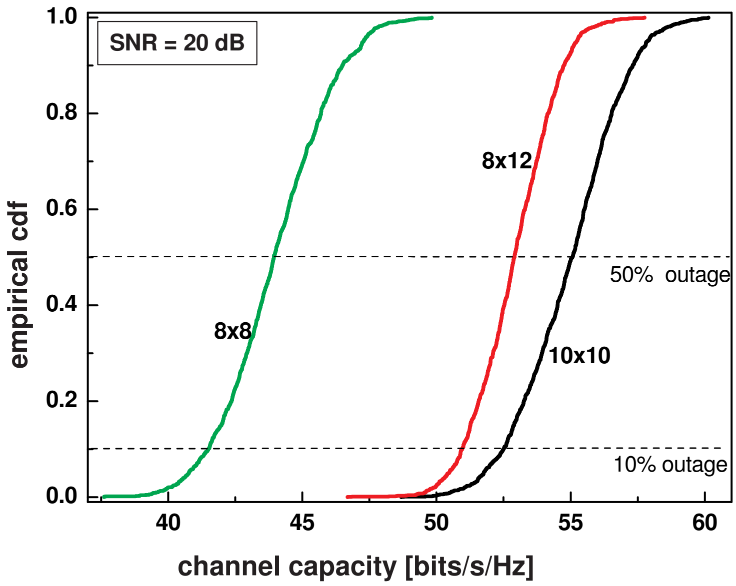

Figure 3.14: Empirical cumulative density function for the capacity for various antenna configurations in the i.i.d. Rayleigh channel (SNR=20 dB).

Fig. 3.14 shows the empirical cumulative density function (ecdf) of the capacity obtained from the simulations. The average capacity which was already depicted in Fig. 3.13 increases when additional antennas are used on both sides. Adding antennas at one side not only increases the average capacity. Interestingly, a steeper accent of the ecdf is always observed which narrows the capacity distribution. For a total number of 20 antennas, the capacity loss would be only 1.5 Bits/s/Hz (~3%) for an 10% outage capacity with SNR=20 dB when a 8x12 MIMO configuration is used instead of a 10x10 configuration and the difference reduces when the outage shall be further reduced.

Bit Error Rates and Antenna Diversity:

We assume equal data rate and modulation on n_T parallel data streams transmitted each from one transmit antenna and no CSI at the Tx. The receiver with $m_R \geq n_T$ receive antennas has perfect channel knowledge and performs ZF to extract the data from the receive signals.

$$\hat{\mathbf{x}} = \mathbf{H}^{\dagger}\mathbf{y} = \mathbf{H}^{\dagger}(\mathbf{H} \cdot \mathbf{x} + \mathbf{n}) = \mathbf{x} + \mathbf{H}^{\dagger}\mathbf{n} \tag{3.51}$$

with $\mathbf{H}^{\dagger}$ denoting the Moore-Penrose pseudo inverse of the channel matrix $\mathbf{H}$.

The resulting BER performance can be obtained analytically for the special case of binary phase shift keying (BPSK) [J.G00], [JJR94], and simulation results in Fig. 3.15 agree very well with

(3.52),

$$p_b = \left(\frac{1}{2}(1-\mu)\right)^l \cdot \sum_{k=0}^{l-1} \begin{pmatrix} l-1+k \\ k \end{pmatrix} \cdot \left(\frac{1}{2}(1+\mu)\right)^k \tag{3.52}$$

where $l = m_R - n_T + 1$ and $\mu = \sqrt{(\frac{\gamma_c}{1+\gamma_c})}$ and $\gamma_c = \frac{\langle SINR \rangle}{n_T}$.

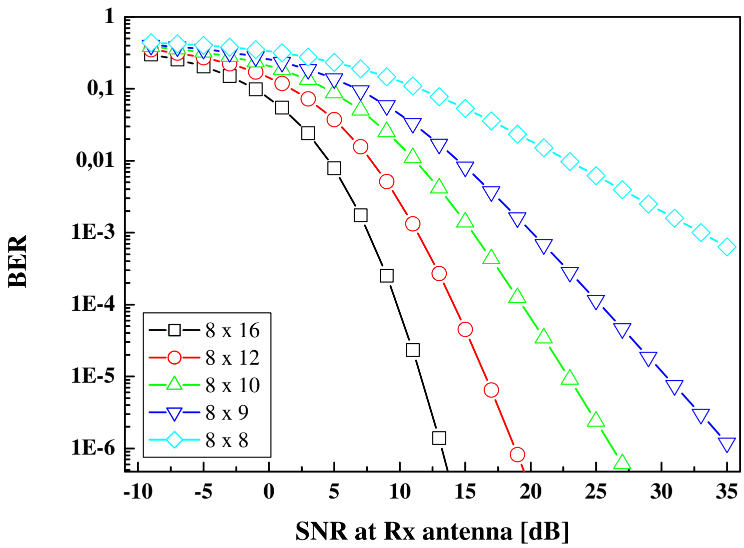

Figure 3.15: Bit Error Rates with antenna diversity plotted over average SNR at Rx antenna.

Fig. 3.15 illustrates the remarkable effect of antenna diversity on the achievable BER. For high SNR, the BER with ZF follows a decay of a $m_R - n_T + 1$-diversity order system. This gain in BER performance is easy to understand if we remember that the unitary matrix $\mathbf{V}$ obtained from SVD (3.9) performs a projection of the data streams on the parallel sub-channels. In general, this projection feeds parts of each data stream into each sub-channel. The sub-channels corresponding to the smaller SVs (see Fig. 3.12) have the worst effective SNR which means they cause most of the bit errors while the part of the data which is feed into the better sub-channels suffers less degradation. Since additional antennas shift the whole bunch of ordered SVs to higher values, especially the smallest SVs are improved which caused most of the bit errors. This explains the drastic enhancement effect on the BER.

3.5.2 Line-of-Sight in Rician Channels

In a general wireless transmission scenario we have to consider the existence of a LOS and therefore something like a Rician channel. The composite channel consisting of a NLOS component and the LOS component may suffer performance degradation when the Rician factor K is high and rank($\mathbf{H}_{LOS}$) is low and the received power is kept constant by a closed loop transmit power control. But for the first part of the simulations we discuss the LOS effect on the SVs considering the LOS as additional power to the NLOS component (no power control at Tx).

Distribution of the Singular Values:

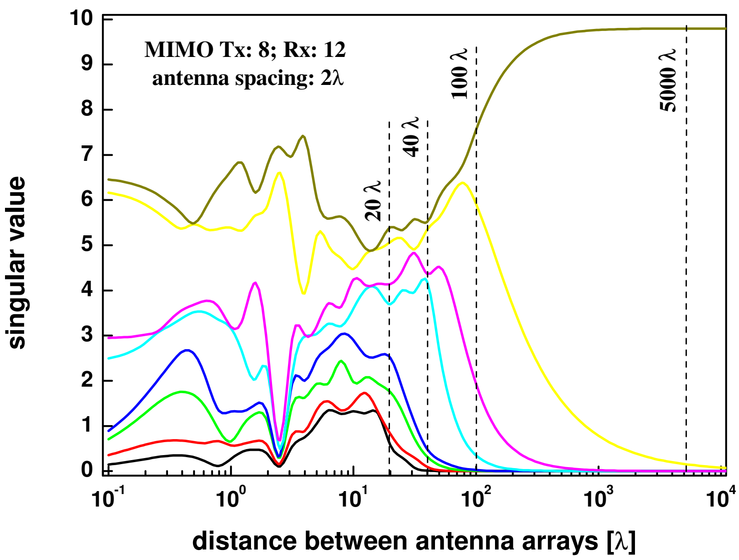

Figure 3.16: Ordered singular values of the LOS channel $\mathbf{H}_{LOS}$ versus the distance of the antenna arrays, given in unit wavelength [λ].

The LOS channel is modelled with a circular Tx array and a linear Rx array which faces broadside to the Tx as depicted in Fig. 3.11. The channel matrix is then computed for a fixed distance (d) between the antenna arrays. Fig. 3.16 shows the distribution of the SVs of the LOS matrix obtained by SVD. At short distances the matrix $\mathbf{H}_{LOS}$ has high or even full rank which means that MIMO may already work on a LOS based connection with special antenna configurations. In [HK03] it was shown analytically and in an experiment that under well known geometrical conditions between the Tx and the Rx like in an office scenario, the MIMO channel can be enhanced by forcing the LOS channel component to achieve full rank. This also includes the most desirable antenna configuration where all SVs are the same and the higher the Rician factor the better. This allows very stable high data rate transmission without a need of sophisticated detection schemes like V-BLAST because already a linear detector e.g. ZF or MMSE achieves a very good BER performance.

In the far field approximation where d is of several thousand carrier wavelengths λ, the matrix $\mathbf{H}_{LOS}$ reduces its effective rank to one which means that the transmission should be reduced to one data stream with beam-forming by adaptation at the Tx side [HFG+04c, HJJ+01a, HFG+04a] since otherwise the transmission channel might be over-loaded with multiple data streams. Assuming a Rician channel we compose $\mathbf{H}_{LOS}$ and $\mathbf{H}_{NLOS}$ and investigate the distribution of the SVs for a fixed distance $d = 10^5\lambda$ and varying Rician factors K in Fig. 3.17 and for a fixed Rician factor $K = 10$ dB and several distances d for the antenna positions in Fig. 3.18.

Fig. 3.17 shows that the additional power from the LOS improves only the highest SVs while the lower SVs remain unchanged. With rising Rician factor K the quality of the best and worst transmission channel is spreading apart meaning that with high K the number of useful parallel transmission channels will decrease under a Tx power control scheme.

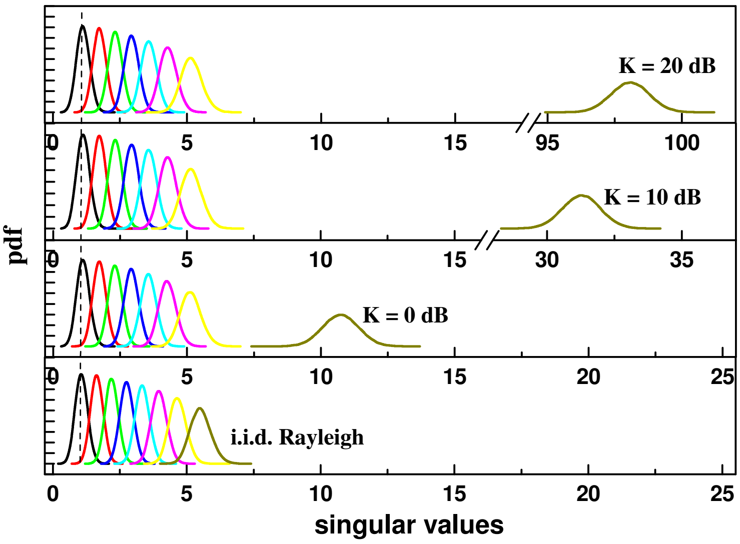

Figure 3.17: Distribution of the ordered singular values of a 8x12 MIMO system for a fixed array distance (far field:$d = 10^5\lambda$) and antenna spacing $\delta = 2\lambda$ and several Rician factors K and no power control at Tx.

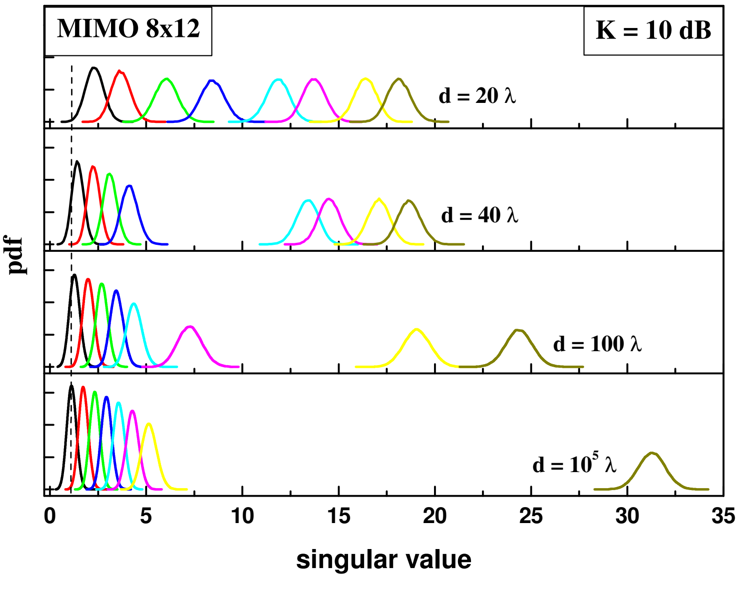

Figure 3.18: Distribution of the singular values for several array distances $d = 20, 40, 100, 10^5\lambda$ (antenna spacing $\delta = 2\lambda$ and Rician factor $K = 10$dB) and no power control at Tx.

Fig. 3.18 shows that the effect of the LOS strongly depends on the actual antenna configuration. In case of a static LOS configuration we see that the LOS improves exactly rank($\mathbf{H}_{LOS}$) singular values. For small distance between Tx and Rx the Rician channel has full rank even for high Rician factors K. This has to be kept in mind for the antenna design towards real applications, either the antennas enhance the multi-path signals or if there is not much multi-path signal available or the Rician factor is high then the antennas should support a LOS channel with maximum rank possible (at least 2 due to polarization multiplexing [AMdC01, SBH+02]) and the data stream multiplexing should be adapted to the actual channel quality [JHPvH03, HFG+04c].

Capacity and Line-of-Sight

In order to discuss the effect on the capacity we plot the average capacity over the average SNR in the Rayleigh channel (see Fig. 3.19). This is chosen to show the effect of the additional LOS without any normalization which would be caused by a Tx power control.

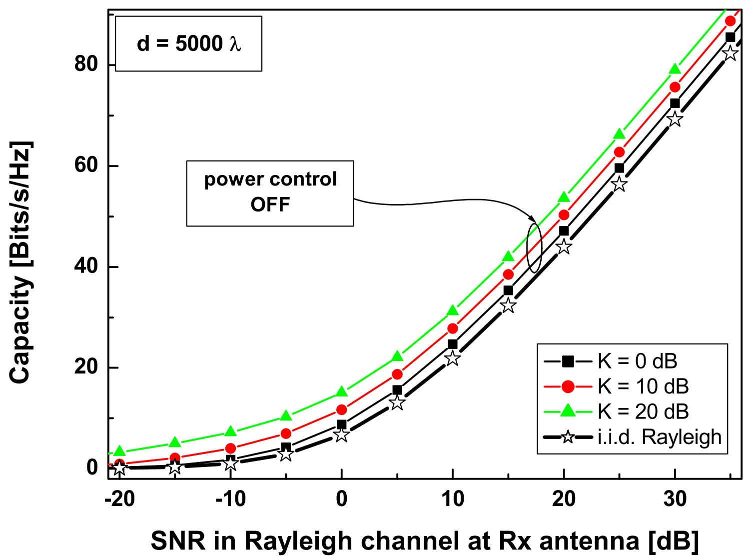

Figure 3.19: Average capacity for a MIMO system with 8 Tx and 8 Rx antennas plotted for various Rician factors K without power control.

When a LOS is suddenly added or blocked (we assume that the slow power control can not follow fast enough), then the LOS simply adds or subtracts power and capacity because some of the SVs are increased/decreased while the characteristic slope for high SNR remains unchanged. This means a constant multiplexing gain.

Fig. 3.20 shows the capacity loss if we assume a slow power control at the Tx like in [DF99]. This power control assumption is justified by the fact that real transmission systems always require a sort of power control to match the received power and the dynamic range of the receive branches and to reduce the interference power into neighboring cells. Therefore the average transmitted power will be automatically reduced if more power is received due to an additional LOS.

In Fig. 3.20 the capacity of a single-input single-output (SISO) channel is compared with a MIMO channel of 8 Tx and 8 Rx antennas for several Rician factors K at SNR=20 dB. The curves represent some specific antenna configurations according to the SV distribution in Fig. 3.18. Under the assumption of a power control we have to conclude that if $\mathbf{H}_{LOS}$ is of reduced rank we expect always a capacity loss for a high Rician factor K. If a static LOS antenna configuration is apparent which supports a high rank of $\mathbf{H}_{LOS}$ then the capacity loss can be limited or even avoided. Note

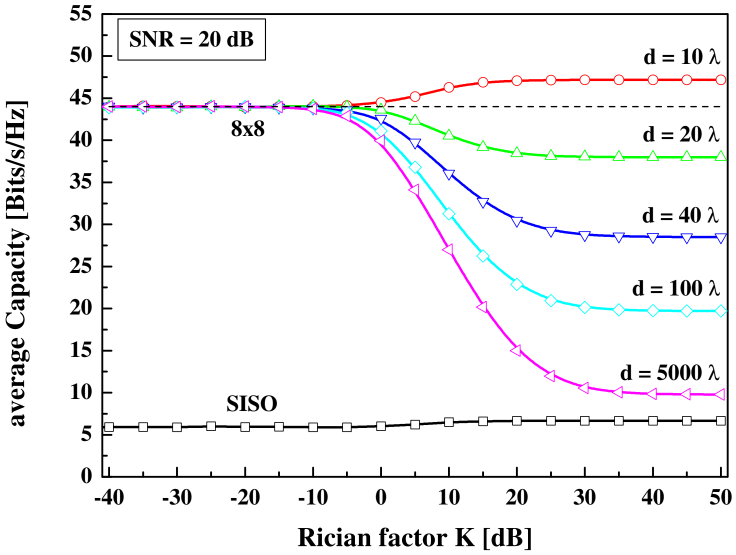

Figure 3.20: Average capacity (8×8) for various antenna array distances (d=10, 20, 40, 100, 5000 λ) versus Rician factor K. The Tx power control holds long term receive power constraint at Rx (constant dynamic range at Rx). SNR at Rx antenna: 20 dB.

if K can be kept low with an appropriate antenna design, therefore the capacity loss might be tolerable.

Bit Error Rate and Line-of-Sight

The BER simulations in Fig. 3.21 and Fig. 3.22 assume perfect CSI and ZF at the Rx. Likewise in the discussion about capacity we start with the assumption of no power control where a LOS simply adds additional power to the Rx.

The curves in Fig. 3.21 show the effect of the LOS on the BER depending on rank($\mathbf{H}_{LOS}$) with a fixed Rician factor $K = 10$ dB and perfect CSI at the Rx (no power control). For comparison we have used the same antenna configurations as in Fig. 3.16. The open squares (□) represent the BER in the Rayleigh channel. The open circles (∘) stand for the worst case scenario with a LOS in the far-field approximation which means that rank($\mathbf{H}_{LOS}$) = 1, so only the highest SV is improved and the small SVs remain unchanged and so does the BER which is determined by the bit errors caused by these channels. With rising rank of $\mathbf{H}_{LOS}$ the smaller SVs are also shifted to higher values thus benefiting from the additional power of the LOS. The left curve is the best case scenario where $\mathbf{H}_{LOS}$ is of full rank causing all SVs to be shifted. The resulting effect does not only shift the BER curve to a lower SNR region but also improves the slope to be more like in a channel with additive white Gaussian noise (AWGN) instead of a Rayleigh fading channel. This is easy to understand because in the latter case already $\mathbf{H}_{LOS}$ enables full multiplexing gain but the LOS does not follow fading statistics which explains the BER performance more similar to a wired data transmission.

If a long term Tx power control is assumed, then the SNR axis in Fig. 3.21 has simply to be

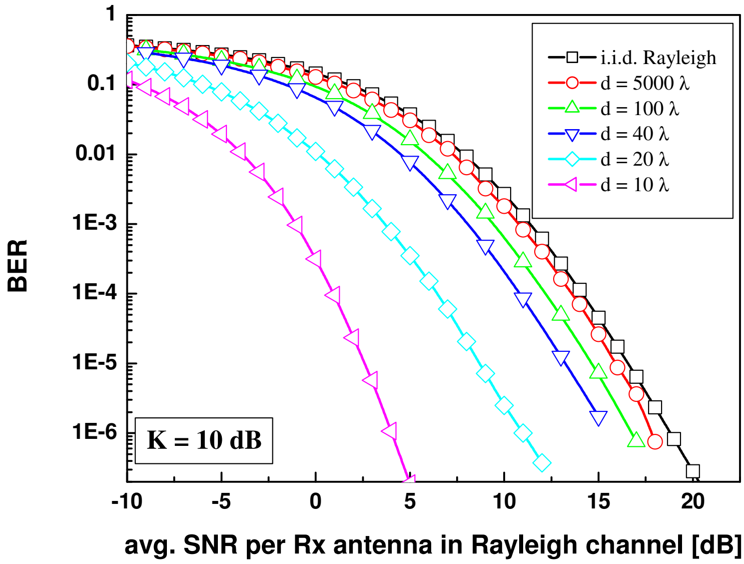

Figure 3.21: BER for a 8x12 MIMO system with $K = 10$dB and various antenna array distances d (no power control at the Tx).

shifted by a factor of $1 + K$ for all BERs. This is done in Fig. 3.22 for four distinctive antenna configurations and variable Rician factors.

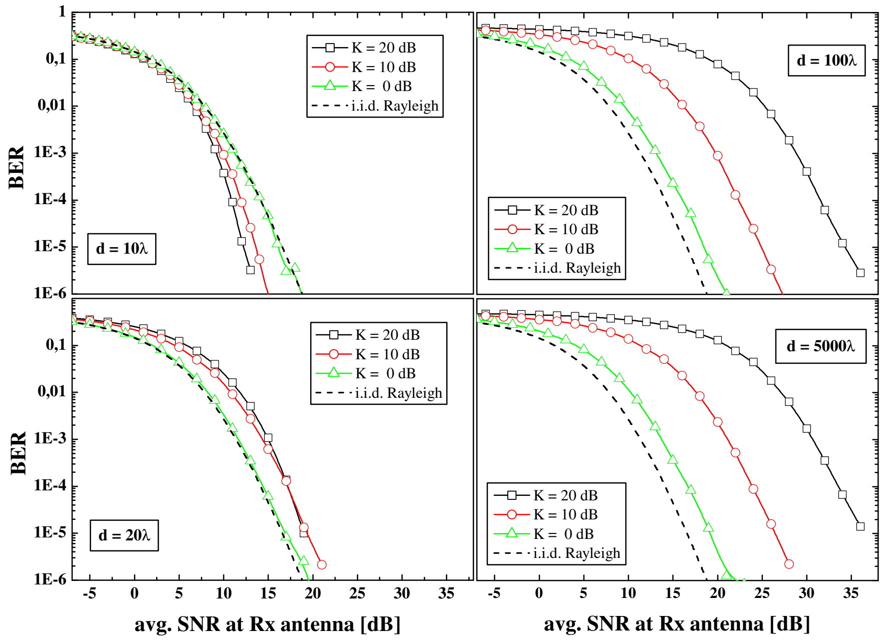

Figure 3.22: BER performance of a 8x12 MIMO system for four distinctive antenna array distances of 10, 20, 100 and 5000 λ and various Rician factors K and BPSK (long term power control at the Tx is applied).

We see that the BER performance always degrades if rank($\mathbf{H}_{LOS}$) $< \min(n_T, m_R)$ and a long term power control is activated. The worst degradation is found when $\mathbf{H}_{LOS}$ is of rank one that means the LOS supports only one sub-channel while the power transmitted over the remaining parallel MIMO sub-channels is reduced dramatically by the long term power control. Note that the BER performance is not much effected by a reduction of the rank of $\mathbf{H}$ if maximum likelihood detection would be used, which seems technically not very realistic at the moment, assuming many parallel data streams and a high modulation level, e.g. 256-QAM.

Summarizing the effects of a LOS on the BER performance and the achievable capacity of a MIMO system we have to state that we have to face BER performance degradation and capacity loss with a LOS, in general. For any practical system long term power control can be expected because of SNR and dynamic range requirements at the Rx. This power control will generally respond to the overall power received at the Rx regardless of its origin (LOS or NLOS). The degradation in BER and capacity strongly depends on the rank($\mathbf{H}_{LOS}$) and the Rician factor K which was also found in [LFV01b]. Recent experiments conducted by [ESA04] and [HK03] showed the crucial performance dependence on rank($\mathbf{H}$). For some specific antenna configurations rank($\mathbf{H}_{LOS}$) can be high or equal to the $M = \min(n_T, m_R)$ but this could only be realized or enforced for quasi-static connections [HK03]. Generally, this can not be taken for granted, therefore we have to expect a lower rank of the LOS channel. This means that the degradation depends mainly on the Rician factor K. For measurements in typical indoor scenarios we used antennas which support the NLOS component and achieved Rician factors between -5dB and +10dB [PJHvH02, JPN+02]. We used $\lambda/4$-antennas (see Fig. 6.8) which were mounted on the three planes of a metal cube of edge size λ= 6 cm. The measurement results indicate that adequate antenna design is very important. Suitable antennas for MIMO WLAN applications should have high spatial diversity achievable by polarization diversity, diverse and broad antenna pattern to support the explicable properties of the rich multi-scattering environment. Similar results could be concluded from narrow band [KMJ+00] and broadband [JPN+02] measurements in rich multi-path environments.

4 Transmission strategies for multiple antenna channels

4.1 Transmission and Detection Strategies

4.1.1 Single User MIMO Transmission Scenario

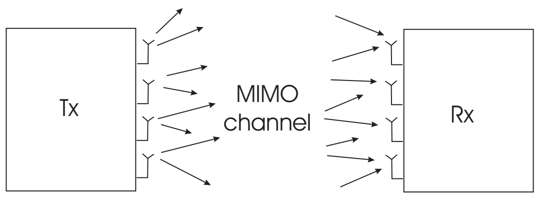

Figure 4.1: Single user MIMO Channel (SU-MIMO).

In the following we assume a single user MIMO system with n_T Tx and m_R Rx antennas with $n_T \leq m_R$ and up to L data streams are transmitted over the MIMO channel $\mathbf{H}$. At the Rx we assume a Minimum Mean Square Error (MMSE) detector with perfect CSI. We assume either no CSI (4.1.1) at the transmitter or perfect CSI at the transmitter and the receiver and we always consider a total sum power constraint at the Tx. The transmission scenario is depicted in Fig. 4.1.

The complex valued MIMO transmission model in matrix form reads

$$\mathbf{y} = \mathbf{H} \cdot \mathbf{x} + \mathbf{n} \tag{4.1}$$

with $\mathbf{y}$ the receive vector of length m_R, $\mathbf{x}$ the transmitted vector of size n_T, $\mathbf{n}$ is the additive Gaussian noise vector of size m_R.

CSI only at the Receiver

If no CSI is available at the Tx then uniform power allocation and one data stream per antenna is optimal: $E[\mathbf{x}\mathbf{x}^H] = \frac{P}{n_T} \cdot \mathbf{I}_{n_T}$ where $E[.]$ means the expectation, $[.]^H$ means Hermitean conjugate, P is the total sum power and $\mathbf{I}_{n_T}$ is the identity matrix of size $n_T \times n_T$.

The data symbol estimate by the linear MMSE receiver is

$$\hat{\mathbf{x}} = \frac{P}{n_T} \mathbf{H}^H \left[\sigma^2 \mathbf{I}_{m_R} + \frac{P}{n_T} \mathbf{H}\mathbf{H}^H \right]^{-1} \mathbf{y} \tag{4.2}$$

with σ^2 the noise variance at the Rx. The covariance matrix $\mathbf{K}_\varepsilon$

$$\begin{aligned}\mathbf{K}_\varepsilon &= \mathbb{E}[(\hat{\mathbf{x}}-\mathbf{x})(\hat{\mathbf{x}}-\mathbf{x})^H] \\ &= \frac{P}{n_T}\mathbf{I}_{n_T} - \frac{P}{n_T}\mathbf{I}_{n_T}\mathbf{H}^H[\sigma^2\mathbf{I}_{m_R} + \frac{P}{n_T}\mathbf{H}\mathbf{H}^H]^{-1}\mathbf{H}\frac{P}{n_T}\mathbf{I}_{n_T}\end{aligned}$$

yields with a normalization

$$\frac{n_T}{P}\mathbf{K}_\varepsilon = \mathbf{I}_{n_T} - \sqrt{\frac{P}{n_T}}I_{n_T}\mathbf{H}^H[\sigma^2\mathbf{I}_{m_R} + \frac{P}{n_T}\mathbf{H}\mathbf{H}^H]^{-1}\mathbf{H}\sqrt{\frac{P}{n_T}}\mathbf{I}_{n_T}. \tag{4.3}$$

trace $(\mathbf{K}_\varepsilon)$ gives the normalized MSE at the Rx.

$$\mathrm{trace}\left(\frac{n_T}{P}\mathbf{K}_\varepsilon\right) = n_T - \mathrm{trace}\left([\sigma^2\mathbf{I}_{m_R} + \frac{P}{n_T}\mathbf{H}\mathbf{H}^H]^{-1}\frac{P}{n_T}\mathbf{H}\mathbf{H}^H\right). \tag{4.4}$$

We consider the singular value decomposition (SVD) of

$$\mathbf{H} \;=\; \mathbf{U}\Lambda_{\mathbf{H}}^{1/2}\mathbf{V}^H \tag{4.5}$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices and $\Lambda^{1/2}$ is a diagonal matrix with the square root of the ordered eigenvalues of $\mathbf{H}\mathbf{H}^H$ on its diagonal. We now decompose $\mathbf{H}\mathbf{H}^H$

$$\begin{aligned}\mathbf{H}\mathbf{H}^H = \mathbf{U}\Lambda_H^{1/2}\mathbf{V}^H\mathbf{V}\Lambda_H^{1/2}\mathbf{U}^H &= \mathbf{U}\Lambda_H\mathbf{U}^H \\ \text{therefore } \sigma^2\mathbf{I}_{m_R} + \frac{P}{n_T}\mathbf{H}\mathbf{H}^H &= \mathbf{U}\left(\sigma^2\mathbf{I}_{m_R} + \frac{P}{n_T}\Lambda_H\right)\mathbf{U}^H.\end{aligned}$$

We define $\mathbf{D} = \sigma^2\mathbf{I}_{m_R} + \frac{P}{n_T}\Lambda_H$ then

$$[\sigma^2\mathbf{I}_{m_R} + \frac{P}{n_T}\mathbf{H}\mathbf{H}^H]^{-1} = \mathbf{U}\mathbf{D}^{-1}\mathbf{U}^H. \tag{4.6}$$

We now apply (4.6) to substitute the last part in (4.4)

$$\begin{aligned}&\mathrm{trace}([\sigma^2\mathbf{I}_{m_R} + \frac{P}{n_T}\mathbf{H}\mathbf{H}^H]^{-1}\frac{P}{n_T}\mathbf{H}\mathbf{H}^H) \\ =\;& \mathrm{trace}\left(\frac{P}{n_T}\mathbf{U}\mathbf{D}^{-1}\mathbf{U}^H\mathbf{U}\Lambda_H\mathbf{U}^H\right) = \mathrm{trace}\left(\frac{P}{n_T}\mathbf{U}\mathbf{D}^{-1}\Lambda_H\mathbf{U}^H\right) \\ =\;& \sum_{l=1}^{m_R}\frac{\frac{P}{n_T}\lambda_H(l)}{\sigma^2+\frac{P}{n_T}\lambda_H(l)} = m_R - \sigma^2\sum_{l=1}^{m_R}\frac{1}{\sigma^2+\frac{P}{n_T}\lambda_H(l)}\end{aligned}$$

with $\lambda_H(i)$ denoting the i-th eigenvalue of $\mathbf{H}$. The normalized MSE of (4.4) (also see Fig. 4.3) now reads

$$\frac{n_T}{P}\text{trace}(\mathbf{K}_\varepsilon) = n_T - m_R + \sigma^2 \sum_{l=1}^{m_R} \frac{1}{\sigma^2 + \frac{P}{n_T}\lambda_H(l)} \tag{4.7}$$

$$= n_T - m_R + \sigma^2 \left(\sum_{l=1}^{n_T} \frac{1}{\sigma^2 + \frac{P}{n_T}\lambda_H(l)} + \sum_{l=n_T+1}^{m_R} \frac{1}{\sigma^2} \right) \tag{4.8}$$

$$= \sum_{l=1}^{n_T} \frac{1}{1 + \frac{P}{n_T\sigma^2}\lambda_H(l)}. \tag{4.9}$$

The RHS in (4.9) is a Schur-convex function which leads to the following theorem.
Theorem 2: For trace($\mathbf{HH}^H$) = constant, rising correlation[1] in $\mathbf{H}$ increases the normalized MSE at the MMSE receiver.
Proof: Let trace($\mathbf{HH}^H$) $= \sum_{l=1}^{N} \lambda_H(l) \doteq 1$. $\mathbf{H}_1$ and $\mathbf{H}_2$ be two channel matrices and $\mathbf{H}_1$ has more correlation than $\mathbf{H}_2$ which we write $\sum_{l=1}^{m} \lambda_{\mathbf{H}_1}(l) \geq \sum_{l=1}^{m} \lambda_{\mathbf{H}_2}(l) \quad m = 1, ..., n_T$. The MSE is of the form MSE= $\sum_{l=1}^{m} f(x)$ with the Schur-convex function $f(x) = \frac{1}{1+x}$. According to theorem C1 from chapter 3 in [MO79] also MSE= $\sum_{l=1}^{m} f(x)$ is Schur-convex. Therefore always holds

$$\sum_{l=1}^{m} \frac{1}{\sigma^2 + \frac{P}{n_T}\lambda_{H_1}(l)} \geq \sum_{l=1}^{m} \frac{1}{\sigma^2 + \frac{P}{n_T}\lambda_{H_2}(l)} \qquad \square \tag{4.10}$$

CSI at the Transmitter and the Receiver

Let L be fixed and perfect CSI is available at the Tx, then the data symbol vector $\boldsymbol{s}\epsilon\mathbb{C}^L$ is preprocessed and then $\mathbf{x}\epsilon\mathbb{C}^{n_T}$ is emitted from the M Tx antennas while $n_T - L$ data streams are switched off. The transmission scheme is depicted in Fig. 4.2 and

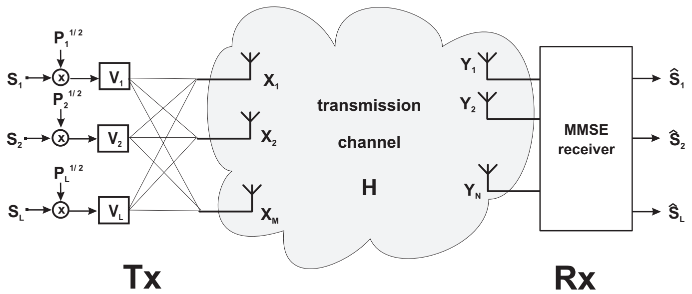

Figure 4.2: MIMO transmission setup with channel knowledge at Tx.

$\mathbf{x} = \mathbf{WDs}$, where $\mathbf{D} = \text{diag}(\sqrt{P_1}, ..., \sqrt{P_L})$ is the power allocation matrix and $\mathbf{W}$ is a unitary beam-forming matrix of size $n_T \times L$. Now (4.1) reads

$$\boldsymbol{y} = \mathbf{Hx} + \mathbf{n} = \mathbf{HWDs} + \mathbf{n} \tag{4.11}$$

We define $\mathbf{s}_p = \mathbf{Ds}$ then $\mathbb{E}[\mathbf{ss}^H] = \mathbf{I}_L$ and $\mathbb{E}[\mathbf{s}_p\mathbf{s}_p^H]$=$\mathbf{D}$ with $\sum_{l=1}^{L} P_l \leq P$. The estimated data

[1]Correlation is used here in the sense of the distribution of the ordered eigenvalues (EW)[H. 02]. Uncorrelated - best case, when all EW are the same, fully correlated - worst case, when there is only one EW bigger then zero.

$\hat{\mathbf{s}}\epsilon\mathbb{C}^L$ at the MMSE receiver is then

$$\hat{\mathbf{s}} = \mathbf{D}\mathbf{W}^H\mathbf{H}^H[\sigma^2\mathbf{I}_{m_R} + \mathbf{H}\mathbf{W}\mathbf{D}\mathbf{W}^H\mathbf{H}^H]^{-1}\mathbf{y}. \tag{4.12}$$

The covariance matrix $\mathbf{K}_\varepsilon$ is

$$\mathbf{K}_\varepsilon = \mathbf{D} - \mathbf{D}\mathbf{W}^H\mathbf{H}^H[\sigma^2\mathbf{I}_{m_R} + \mathbf{H}\mathbf{W}\mathbf{D}\mathbf{W}^H\mathbf{H}^H]^{-1}\mathbf{H}\mathbf{W}\mathbf{D} \tag{4.13}$$

or in normalized form

$$\mathbf{D}^{-\frac{1}{2}}\mathbf{K}_\varepsilon\mathbf{D}^{-\frac{1}{2}} = \mathbf{I}_L - \mathbf{D}^{\frac{1}{2}}\mathbf{W}^H\mathbf{H}^H[\sigma^2\mathbf{I}_{m_R} + \mathbf{H}\mathbf{W}\mathbf{D}\mathbf{W}^H\mathbf{H}^H]^{-1}\mathbf{H}\mathbf{W}\mathbf{D}^{\frac{1}{2}}. \tag{4.14}$$

With $\mathbf{H} = \mathbf{U}\Lambda_H^{1/2}\mathbf{V}^H$, $\mathbf{V} = [V_1, ..., V_{n_T}]$ and $\mathbf{W} = [V_1, ..., V_L]$ the normalized MSE is given by

$$\begin{aligned}
\mathrm{trace}(\mathbf{D}^{-\frac{1}{2}}\mathbf{K}_\varepsilon\mathbf{D}^{-\frac{1}{2}}) &= L - \mathrm{trace}\left([\sigma^2\mathbf{I}_{m_R} + \mathbf{H}\mathbf{W}\mathbf{D}\mathbf{W}^H\mathbf{H}^H]^{-1}\mathbf{H}\mathbf{W}\mathbf{D}\mathbf{W}^H\mathbf{H}^H\right) \\
&= L - \left(m_R - \sigma^2\sum_{l=1}^{m_R}\frac{1}{\sigma^2 + \lambda_H(l)P_l}\right) && (4.15) \\
&= L - m_R + \sigma^2\left(\sum_{l=1}^{L}\frac{1}{\sigma^2 + \lambda_H(l)P_l} + \sum_{l=L+1}^{m_R}\frac{1}{\sigma^2}\right) && (4.16) \\
&= \sum_{l=1}^{L}\frac{1}{1 + \frac{\lambda_H(l)P_l}{\sigma^2}}. && (4.17)
\end{aligned}$$

In order to minimize the sum of the MSE's for all data streams we solve the following minimization problem

$$\min_{\substack{\sum_{l=1}^L P_l \le P \\ P_l \ge 0}} \sum_{l=1}^{L}\frac{1}{1 + \frac{\lambda_H(l)P_l}{\sigma^2}}. \tag{4.18}$$

We find the Lagrange function $\mathcal{L}(\boldsymbol{P}, \mu, \boldsymbol{\omega})$

$$\mathcal{L}(\boldsymbol{P}, \mu, \boldsymbol{\omega}) = \sum_{l=1}^{L}\frac{1}{1 + \frac{\lambda_H(l)P_l}{\sigma^2}} + \mu\left(\sum_{l=1}^{L}P_l - P\right) - \sum_{l=1}^{L}\omega_l P_l \tag{4.19}$$

where μ is the Lagrange multiplier to satisfy $\sum_{l=1}^L P_i \le P$ and $\boldsymbol{\omega}$ guarantees all $P_i \ge 0$. Partial differentiation of (4.19) gives

$$\frac{\partial\mathcal{L}}{\partial P_r} = -\frac{\frac{\lambda_H(r)}{\sigma^2}}{\left(1 + \frac{\lambda_H(r)P_r}{\sigma^2}\right)^2} + \mu - \omega_r = 0. \tag{4.20}$$

With a closer look at (4.18) we see that the sub-channels have different impact on the MSE. We expect a "waterfilling"-like solution, which means that with little sum power, sub-channels corresponding to smaller eigenvalues λ_i are switched off.

$$\text{if } P_l^{opt} = 0, \text{ then } \omega_l \ge 0 \text{ and if } P_l^{opt} > 0, \text{ then } \omega_l = 0 \tag{4.21}$$

There exists a maximum index L for which holds: if $l > L$, then

$P_l^{opt} = 0$ and $\frac{\lambda_H(l)}{\sigma^2} = \mu - \omega_l$. With (4.19) and for $l \leq L$ the optimal solution is then given by

$$P_l^{opt} = \left[\sqrt{\frac{\sigma^2}{\mu \lambda_H(l)}} - \frac{\sigma^2}{\lambda_H(l)} \right]^+ \tag{4.22}$$

where μ satisfies $\sum_{l=1}^{L} P_l^{opt} = P$. This leads to:

Theorem 3: In case of perfect CSI at the Tx and Rx and a MMSE receiver, then the optimal transmit strategy is given by transmitting L data streams with the transmit vector $\mathbf{x}$:

$$\mathbf{x} = \mathbf{W}\mathbf{D}\boldsymbol{s}. \tag{4.23}$$

The unitary beam-forming matrix $\mathbf{W}$ is given by the first L columns of $\mathbf{V}$ obtained from SVD of $\mathbf{H} = \mathbf{V}\boldsymbol{\Lambda}^{1/2}\mathbf{U}^H$ and the power allocation matrix $\mathbf{D} = \text{diag}(\sqrt{P_1}, ..., \sqrt{P_L})$ with P_l in (4.22) from the solution of the minimization problem formulated in (4.18).

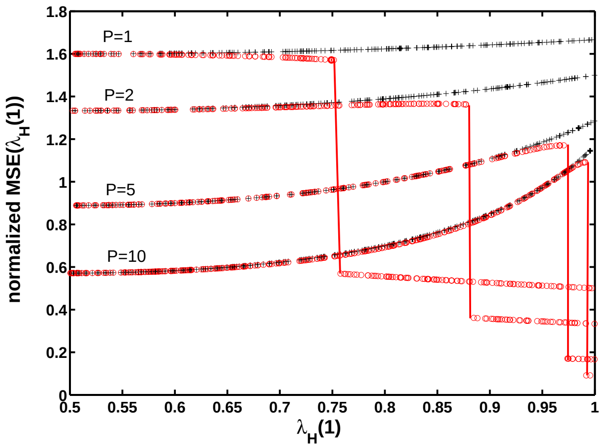

Figure 4.3: MSE as a function of the eigenvalues $\lambda_H(1)$ with $\lambda_H(1) = (1 - \lambda_H(2))$ and $\sum_{i=1}^{2} P_i =$ P $= 1, 2, 5, 10$ and $\sigma^2 = 1.0$

Fig. 4.3 shows the MSE functions for the ordered 2 Eigenvalues example. The upper curves(+) belong to the system with no CSI at the Tx and are Schur-convex. The lower curves ($\circ$) of the 4 sets are the MSE's with optimum power allocation. These functions are not Schur-convex, in general, which we show in the following. On the right hand side from the jump discontinuity only the effective MSE of the remaining data stream is depicted because P_2=0. Let us assume σ and P to be fixed and $\lambda_{\mathbf{H}}(1)$ be the parameter for the MSE like in Fig. 4.3 ($\lambda_{\mathbf{H}}(1) + \lambda_{\mathbf{H}}(2) = 1$). The solution of the minimization task be $L = 2$ for $\lambda_{\mathbf{H}}(1) = \lambda_{\mathbf{H}}(2)$. With rising correlation ($\lambda_{\mathbf{H}}(1) \uparrow$) we find a $\tilde{\lambda}_{\mathbf{H}}(1)$, so that $L = 1$. Now we consider $\lambda_{\mathbf{H}}^{(1)} \geq \lambda_{\mathbf{H}}^{(2)} \geq \tilde{\lambda}_{\mathbf{H}}(1)$ then

$$\text{MSE}^{(1)} = \frac{1}{1 + \frac{\lambda_{\mathbf{H}}^{(1)} P}{\sigma^2}} \leq \frac{1}{1 + \frac{\lambda_{\mathbf{H}}^{(2)} P}{\sigma^2}} = \text{MSE}^{(2)} \leq \text{MSE}(\tilde{\lambda}_{\mathbf{H}}) \tag{4.24}$$

which is Schur-concave for $\lambda_{\mathbf{H}}^{(i)} \geq \tilde{\lambda}_{\mathbf{H}}$. The more general case reads:

$$\sum_{l=1}^{n_T} \lambda_l^{(1)} = \sum_{l=1}^{n_T} \lambda_l^{(2)} \text{ and } \lambda^{(1)} \geq \lambda^{(2)}. \tag{4.25}$$

We substitute P_l in (4.18) with (4.22) and find for the L best channels in use $\mathrm{MSE}(L) = \sum_{l=1}^{L} \frac{1}{\sqrt{\mu\lambda(l)}}$. Since we know that $\sum_{l=1}^{L} \frac{1}{\sqrt{\lambda^{(1)}(l)}} \geq \sum_{l=1}^{L} \frac{1}{\sqrt{\lambda^{(2)}(l)}}$ a comparison of $\mathrm{MSE}^{(1)} = \sum_{l=1}^{L} \frac{1}{\sqrt{\mu^{(1)}\lambda(l)}}$ and $\mathrm{MSE}^{(2)} = \sum_{l=1}^{L} \frac{1}{\sqrt{\mu^{(2)}\lambda(l)}}$ depends on $\mu^{(1)}$ and $\mu^{(2)}$. Therefore it can not be generally stated whether $\mathrm{MSE}^{(1)} \geq \mathrm{MSE}^{(2)}$ or vice versa. This complex behaviour is to be seen in Fig. 4.3.

We now find the critical power when a channel has to be switched off, assuming a fixed correlation and noise. We consider the two eigenvalue example. We assume P to be the sum power, so that $P_2^{opt} > 0$. We choose a $\hat{P}$, with $P \geq \hat{P}$, which holds $P_1^{opt} = P$ and $P_2^{opt} = 0$. We find the function $f(P_1^{opt}, P_2^{opt})$ and parameterize it with $P_1^{opt} = P - \varepsilon$ and $P_2^{opt} = \varepsilon$.

$$\begin{aligned} f(P_1^{opt}, P_2^{opt}) &= \frac{1}{1 + \frac{\lambda_{\mathbf{H}}(1) P_1^{opt}}{\sigma^2}} + \frac{1}{1 + \frac{\lambda_{\mathbf{H}}(2) P_2^{opt}}{\sigma^2}} \\ f(P - \varepsilon, \varepsilon) &= \frac{1}{1 + \frac{\lambda_{\mathbf{H}}(1)(P-\varepsilon)}{\sigma^2}} + \frac{1}{1 + \frac{\lambda_{\mathbf{H}}(2)\varepsilon}{\sigma^2}} \end{aligned}$$

Now we look at the point where the derivative becomes positiv

$$\frac{df(P-\varepsilon, \varepsilon)}{d\varepsilon} \Big|_{\varepsilon=0} \geq 0. \tag{4.26}$$

$$\frac{df}{d\varepsilon} \Big|_{\varepsilon=0} = \frac{\frac{\lambda_{\mathbf{H}}(1)}{\sigma^2}}{\left(1 + \frac{\lambda_{\mathbf{H}}(1)(P-\varepsilon)}{\sigma^2}\right)^2} - \frac{\frac{\lambda_{\mathbf{H}}(2)}{\sigma^2}}{\left(1 + \frac{\lambda_{\mathbf{H}}(2)\varepsilon}{\sigma^2}\right)^2} \geq 0. \tag{4.27}$$

which leads to
Theorem 4: A necessary and sufficient condition for beam-forming to be optimum is given by

$$P \leq \frac{\sigma^2}{\lambda_H(1)} \left(\sqrt{\frac{\lambda_{\mathbf{H}}(1)}{\lambda_{\mathbf{H}}(2)}} - 1 \right) = P^{crit}. \tag{4.28}$$

Below the critical sum power P^{crit} only one channel is active.

4.1.2 Multi-User SIMO MAC scenario

System model

We consider a multi-user SIMO scenario where all K users have one transmit antenna each and the base station (BS) has m_R receive antennas. The overall formed MU-SIMO or MIMO channel is of size $m_R \times K$. We investigate the up-link from the mobiles to the BS with perfect channel state information.

Fig. 4.4 shows the signal model for the multi-user SIMO MAC with a MMSE receiver. K mobile terminals with one antenna each transmit to a BS with m_R antennas.

The transmit signal from user k is given by x_k and the transmission channel to the BS is described by the channel vector $\mathbf{h}_k$. The received signal $\boldsymbol{y}$ at the BS is given by

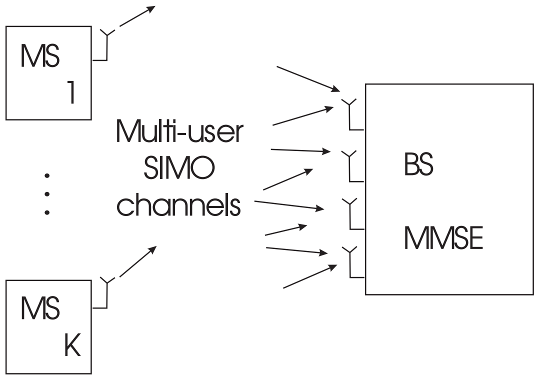

Figure 4.4: MU SIMO MAC with MMSE multi-user receiver

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{h}_k x_k + \mathbf{n}. \tag{4.29}$$

We assume a flat fading channel $\mathbf{h}_k$ for all users in (4.29). Additionally, we assume an i.i.d. additive white Gaussian noise vector $\mathbf{n}$ with noise variance σ_N^2. Equation (4.29) can be rewritten in compact form as

$$\mathbf{y} = \mathbf{H}\hat{\mathbf{x}} + \mathbf{n} \tag{4.30}$$

where $\mathbf{H} = [\mathbf{h}_1, ..., \mathbf{h}_K] \epsilon \mathbb{C}^{m_R \times K}$ is the overall channel matrix and the transmit signals from all users are collected in $\hat{\mathbf{x}} = [x_1, ..., x_K]^T$.

Analytical Description of the SINR region

We follow the considerations of [BS02b] and assume i.i.d. Gaussian noise at each receive antenna. Then the SINR of all links can be controlled by jointly adjusting the transmission powers $p_1...p_K$ and the beamformers $\mathbf{u}_i \epsilon \mathbb{C}^{m_R}, 1 \leq i \leq K$. The uplink SINR of the i-th user is given by

$$\begin{aligned} \mathrm{SINR}_i^{\mathrm{UL}}(\mathbf{u}_i, p_1, ..., p_K) &= \frac{p_i \mid \mathbf{u}_i^H \mathbf{h}_i \mid^2}{\mathbf{u}_i^H \mathbf{Z}_i(\mathbf{p}) \mathbf{u}_i}, \quad 1 \leq k \leq K \\ \text{where } \mathbf{Z}_i(\mathbf{p}) &= \sigma_N^2 \mathbf{I} + \sum_{\substack{k=1 \\ k \neq i}}^{K} p_k \mathbf{h}_k \mathbf{h}_k^H \end{aligned} \tag{4.31}$$

For given transmission powers $p_1...p_K$, the SINR_i from (4.31) are individually maximized by the normalized MMSE solution

$$\mathbf{u}_i^{opt} = \alpha \mathbf{Z}_i(\mathbf{p})^{-1} \mathbf{h}_i, \tag{4.32}$$

where α is chosen such that $\| \mathbf{u}_i^{opt} \| = 1$. Substituting $\mathbf{u}_i^{opt}$ in (4.31) we have

$$\mathrm{SINR}_i^{\mathrm{UL\text{-}MMSE}}(p_1...p_K) = p_i \mathbf{h}_i^H \mathbf{Z}_i(\mathbf{p})^{-1}\mathbf{h}_i. \tag{4.33}$$

A simple two user example given in Fig. 4.6 may illustrate the shape of the SINR region obtained. But the more general problem is to find the optimum power allocation which maximizes the sum capacity for given $\mathbf{h}_1, ..., \mathbf{h}_K$.

Multi User Sum Capacity without Successive Interference Cancellation (noSIC)

It is well known that under the assumption of i.i.d. Gaussian signaling, there is a one-to-one monotonic relationship between user rates R_i and $SINR_i$ values [Lap96], i.e.,

$$\begin{aligned} R_i^{\mathrm{noSIC}} &= \log_2(1 + SINR_i^{\mathrm{MMSE}}(p_1...p_K)) \\ &= \log_2 \mid \mathbf{I} + \mathbf{Z}_i(\mathbf{p})^{-1/2} p_i \mathbf{h}_i \mathbf{h}_i^H \mathbf{Z}_i(\mathbf{p})^{-1/2} \mid \quad (4.34) \\ &= \log_2 \mid p_i \mathbf{h}_i \mathbf{h}_i^H + \mathbf{Z}_i(\mathbf{p}) \mid - \log_2 \mid \mathbf{Z}_i(\mathbf{p}) \mid \text{[bits/symbol]} \quad (4.35) \end{aligned}$$

where | . | denotes det(.).

Hence, the sum rate without SIC at the receiver is given by

$$f^{\mathrm{noSIC}}(\mathbf{p}) = \sum_{i=1}^{K} R_i^{\mathrm{noSIC}} = K \log_2 \mid \sigma_N^2 \mathbf{I} + \sum_{k=1}^{K} p_k \mathbf{h}_k \mathbf{h}_k^H \mid - \sum_{k=1}^{K} \log_2 \mid \mathbf{Z}_k(\mathbf{p}) \mid . \tag{4.36}$$

As stated in [BS02b] the function $f^{noSIC}(\mathbf{p})$ is the sum of a convex and a concave term, therefore the result needs neither to be convex nor to be concave, as can be seen in the given 2-user example depicted in Fig. 4.10. Thus, known strategies for sum capacity maximization [BSJ03, BJ02, VBW98] cannot be applied here.

4.1.3 Multi-User Sum Capacity with SIC

Successive interference cancellation (SIC) is a common technique to reduce known interference step by step. We assume a detection order for the data streams $d_1, ..., d_K$. We start the process by making a decision for data symbol d_1. Since d_1 is known from now on its interference on data symbols from other users can be eliminated prior to making a decision for the next data stream. Effectively every transmission layer sees only the interference from the layers which will be detected later.

Using SIC the functional describing the sum rate becomes concave in the transmission powers $p_1...p_K$. The rate of layer k is

$$R_k^{SIC} = \log_2(1 + p_k \mathbf{h}_k^H [\sigma_N^2 \mathbf{I} + \sum_{l=k+1}^{K} p_l \mathbf{h}_l \mathbf{h}_l^H]^{-1} \mathbf{h}_k). \tag{4.37}$$

For convenience we define $\mathbf{B} = [\sigma_N^2 \mathbf{I} + \sum_{l=k+1}^{K} p_l \mathbf{h}_l \mathbf{h}_l^H]$ and $\sigma_N^2 = 1$.

$$R_k^{SIC} = \log_2 | \mathbf{I} + p_k \mathbf{B}^{-1/2} \mathbf{h}_k \mathbf{h}_k^H \mathbf{B}^{-1/2} | \tag{4.38}$$

$$= \log_2 | \sum_{l=k}^{K} p_l \mathbf{h}_l \mathbf{h}_l^H + \mathbf{I} | - \log_2 | \sum_{l=k+1}^{K} p_l \mathbf{h}_l \mathbf{h}_l^H + \mathbf{I} | \tag{4.39}$$

Then, the sum rate with SIC is given by

$$f^{SIC}(\mathbf{p}) = \sum_{k=1}^{K} R_k^{SIC} = \tag{4.40}$$

$$= \sum_{k=1}^{K} (\log_2 | \sum_{l=k}^{K} p_l \mathbf{h}_l \mathbf{h}_l^H + \mathbf{I} | - \log_2 | \sum_{l=k+1}^{K} p_l \mathbf{h}_l \mathbf{h}_l^H + \mathbf{I} |) \tag{4.41}$$

$$= \log_2 | \mathbf{I} + \sum_{k=1}^{K} p_k \mathbf{h}_k \mathbf{h}_k^H | \,. \tag{4.42}$$

or in a more general form with $\sigma_N^2 \neq 1$

$$f^{SIC}(\mathbf{p}) = \log_2 | \mathbf{I} + \sum_{k=1}^{K} \frac{p_k}{\sigma_N^2} \mathbf{h}_k \mathbf{h}_k^H | \,. \tag{4.43}$$

The optimum power allocation can be found by convex optimization techniques, like the *maxdet* algorithm [VBW98].

We make the observation that the sum rate in (4.43) is independent on the detection order of the data streams (see fig. 4.5). This ***conservation law of the sum capacity*** is of great importance and is very interesting from a providers point of view. If the sum rate is independent on the user detection order, then the provider can run its system at maximum sum rate while supporting the users with different SINR demands by simply changing the detection order accordingly at the BS.

4.1.4 Analysis of the SNR Gap Concept

A common approximation to predict the achievable rate for a transmission system with a certain symbol alphabet and code is the so-called SNR gap approximation. The assumption is that a transmission system can achieve a rate that simply corresponds to the Shannon capacity shifted on the SNR axis. The SNR shift is then called a SNR gap Γ which depends on the modulation and coding scheme used.

This concept was used e.g. for bit-loading algorithms for OFDM [CCB95] and is proposed by [CLH+03] to perform bit-loading with the VBLAST algorithm. The idea is to predict the SINR of each layer by using the gap approximation.

If the SNR gap approximation is used in order to simplify the problem of finding a practical solution with SIC and realistic modulation alphabets and coding, then the desirable and nice

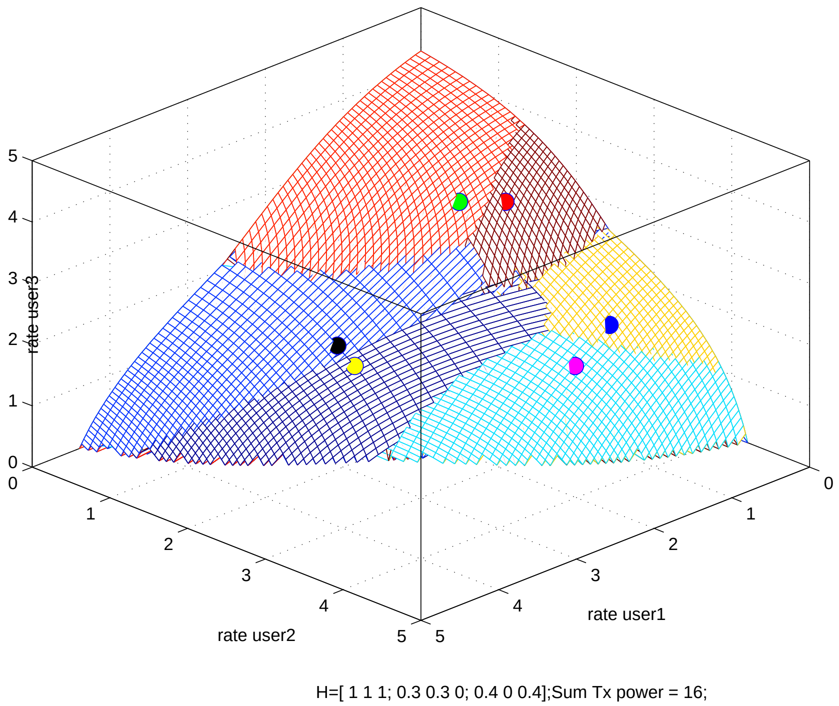

Figure 4.5: Achievable rate region for a 3 user SIMO channel w/o time sharing policies. The colored hyperplanes show the rate region for one fixed SIC order each. The dots indicate where the sum rate for every possible SIC order is maximized. The maximum sum rate is independent on the SIC order.

concave behavior of the sum capacity functional and the independence on the detection order can be lost. To see this, we have to look into the mathematical expression in more detail.

Following the outlines of [CLH+03] we assume the SNR gap $\Gamma > 1$ and a detection order 1...K. The rate for the k-th layer is then

$$
\begin{aligned}
R_k^\Gamma &= \log_2(1 + \frac{1}{\Gamma}\mathrm{SINR}_k) && (4.44)\\
&= \log_2 |\, \mathbf{I} + \sum_{l=k+1}^{K} \frac{p_l}{\sigma^2}\mathbf{h}_l\mathbf{h}_l^H + \frac{1}{\Gamma}\frac{p_k}{\sigma^2}\mathbf{h}_k\mathbf{h}_k^H \,| - \log_2 |\, \mathbf{I} + \sum_{l=k+1}^{K} \frac{p_l}{\sigma^2}\mathbf{h}_l\mathbf{h}_l^H \,| \,. && (4.45)
\end{aligned}
$$

The sum rate is then given by

$$
\begin{aligned}
\sum_{k=1}^{K} R_k^\Gamma &= \sum_{k=1}^{K} \log_2 |\, \mathbf{I} + \sum_{l=k+1}^{K} \frac{p_l}{\sigma^2}\mathbf{h}_l\mathbf{h}_l^H + \frac{1}{\Gamma}\frac{p_k}{\sigma^2}\mathbf{h}_k\mathbf{h}_k^H \,| \\
&\quad - \sum_{k=1}^{K} \log_2 |\, \mathbf{I} + \sum_{l=k+1}^{K} \frac{p_l}{\sigma^2}\mathbf{h}_l\mathbf{h}_l^H \,| \,. \qquad (4.46)
\end{aligned}
$$

Here we make the observation that the interference terms do not cancel pairwise as seen in (4.43) therefore the achieved sum rate may become dependent on the detection order. Furthermore the two sum terms are concave each, therefore the difference of two concave functionals corresponds to a sum of a concave and a convex functional. Thus the functional for the sum rate with gap approximation needs neither to be concave nor to be convex and standard algorithms to find the maximum sum rate can not be applied as discussed in 4.1.2.

As long as the SNR gap $\Gamma \simeq 1$ or the channels are orthogonal (e.g. MIMO-OFDM [GVK02]) the gap approximation does neither affect the concave behavior of the functional nor the sum capacity is depending on the detection order of the data streams. But this can not be assumed in general, therefore the SNR gap approximation might be misleading sometimes.

4.1.5 Two User Examples

Detection without SIC

We consider a 2 user SIMO scenario without SIC and a MMSE receiver. Starting from (4.33) and using the matrix inversion lemma we find

$$\mathrm{SINR}_1 = p_1 \mathbf{h}_1^H [\sigma^2 \mathbf{I} + p_2 \mathbf{h}_2 \mathbf{h}_2^H]^{-1} \mathbf{h}_1 \tag{4.47}$$

$$= \frac{\frac{p_1}{\sigma^2}}{1 + \frac{p_2}{\sigma^2}\|\mathbf{h}_2\|^2} \left[\|\mathbf{h}_1\|^2 + \frac{p_2}{\sigma^2} \left(\|\mathbf{h}_1\|^2 \|\mathbf{h}_2\|^2 - \|\mathbf{h}_1^H \mathbf{h}_2\|^2 \right) \right] \tag{4.48}$$

$$\mathrm{SINR}_2 = \frac{\frac{p_2}{\sigma^2}}{1 + \frac{p_1}{\sigma^2}\|\mathbf{h}_1\|^2} \left[\|\mathbf{h}_2\|^2 + \frac{p_1}{\sigma^2} \left(\|\mathbf{h}_1\|^2 \|\mathbf{h}_2\|^2 - \|\mathbf{h}_1^H \mathbf{h}_2\|^2 \right) \right] \tag{4.49}$$

We see that for a fixed power allocation the $SINR_i$ is maximal if $h_1 \perp h_2$. The resulting capacity

$$C_{1+2} = \log_2(1 + SINR_1) + \log_2(1 + SINR_2)$$

is also maximized.

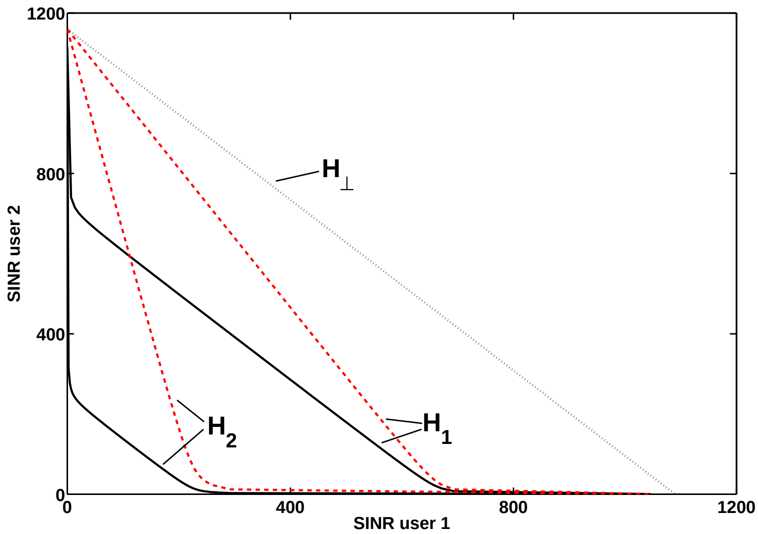

Figure 4.6: SINR region in the 2 user example for fixed sum powers P_{total} and varying correlation in H with SIC (dashed lines) and without SIC (solid lines) ($n_T = 2, m_R = 4$, $\sigma_N^2 = 1$).

For illustration purposes we choose three gain normalized channels with varying correlation [2]. Fig. 4.6 shows the achievable SINR regions. The regions appear to be convex which could not be proven so far. The maximum achievable SINR region is given for $\mathbf{h}_1 \perp \mathbf{h}_2$ and nothing can be gained with SIC.

If correlation between $\mathbf{h}_1$ and $\mathbf{h}_2$ is apparent then the achievable SINR region is reduced (see curves for $\mathbf{H}_1$ and $\mathbf{H}_2$ in Fig. 4.6) and SIC gains according to the correlation. For this example the channels were chosen as

$$\mathbf{H}_\perp = \begin{pmatrix} 0 & 0.4 \\ 0.3 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{H}_1 = \begin{pmatrix} 1 & 0.4 \\ 0.3 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \mathbf{H}_2 = \begin{pmatrix} 0 & 0.4 \\ 0.3 & 0 \\ 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

The SINR regions of another example are depicted in Fig. 4.7, note the dB scale. We discuss this example in the following regarding the theoretically achievable sum rate and what can be achieved with discrete bit-loading in 4.4.2. In Fig. 4.7 we see that the SINR gain for the second data stream is about 10 dB at maximum. Therefrom we would expect a higher achievable rate of about 3 bits/symbol if SIC is applied. The calculation is based on the fact that we expect the rate to rise by 1 bit/s/Hz per 3 dB SINR increase. This gain in data throughput can be approved in Fig. 4.10 and Fig. 4.29. The channel used here assumes 2 users and 4 BS antennas. Without loss of generality it is real-valued for convenience.

$$\begin{aligned} \mathbf{h}_1 &= [+0.5918 \;\; -0.5107 \;\; -0.5102 \;\; -0.1423]^T \\ \mathbf{h}_2 &= [-1.2379 \;\; +0.4931 \;\; +0.5738 \;\; +0.4500]^T \end{aligned} \tag{4.50}$$

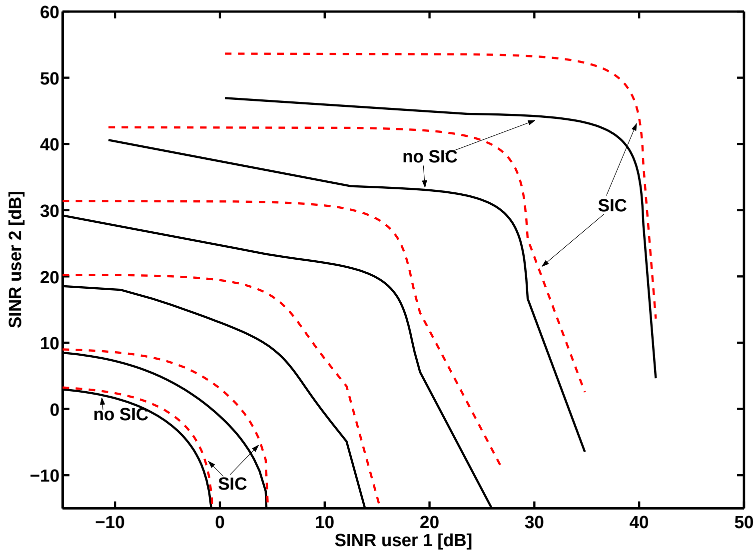

Figure 4.7: SINR region in the 2 user example for different sum powers P_{total} with/without SIC ($n_T = 2, m_R = 4,\ \sigma_N^2 = 1$). Channel from (4.50).

[2] Correlation between two channel vectors is meant here to be the normalized scalar product $\frac{\mathbf{h}_1}{\|\mathbf{h}_1\|} \cdot \frac{\mathbf{h}_2}{\|\mathbf{h}_2\|}$ where zero means uncorrelated and one means full correlation

Detection with SIC and Individual Power Constraints

In the following we will always assume a detection order 2,1 which means the data from user 2 is detected first and then subtracted before user 1 is detected. This makes user 1 virtually interference free. We assume σ_N^2 =1. Furthermore we assume the allocatable power to each user to be $p_{1,2} = [0, P_{max}]$. In direct consequence of maximizing the sum throughput this yields to $p_1 = p_2 = P_{max}$ for $\Gamma = 1$.

SIC and $\Gamma = 1$

The rates of the 2 users with SIC and individual power constraint are

$$\begin{aligned} R_1 &= \log_2(1 + p_1|\mathbf{h}_1\|^2) && (4.51)\\ R_2 &= \log_2(1 + p_2\mathbf{h}_2^H[\mathbf{I} + p_1\mathbf{h}_1\mathbf{h}_1^H]^{-1}\mathbf{h}_2) && (4.52)\\ &= \log_2\left(1 + p_2\left(\|\mathbf{h}_2\|^2 - \frac{p_1\|\mathbf{h}_1^H\mathbf{h}_2\|^2}{1 + p_1\|\mathbf{h}_1\|^2}\right)\right). && (4.53)\end{aligned}$$

The sum rate is then given by

$$\begin{aligned} C^{SIC} &= R_1 + R_2 && (4.54)\\ &= \log_2\overbrace{(1 + p_1\|\mathbf{h}_1\|^2 + p_2\|\mathbf{h}_2\|^2 + p_1p_2(\|\mathbf{h}_1\|^2\|\mathbf{h}_2\|^2 - \|\mathbf{h}_1^H\mathbf{h}_2\|^2))}^{g_1} && (4.55)\\ &= G(p_1, p_2). && (4.56)\end{aligned}$$

With $p_1 \le p_2 = P_{max}$ the functional G is only dependent on p_1.

Since always holds $\|\mathbf{h}_1\|^2\|\mathbf{h}_2\|^2 \ge \|\mathbf{h}_1^H\mathbf{h}_2\|^2$ the functional $G(p_1)$ is monotonic rising in p_1 therefore maximum sum rate is achieved for $p_2 = p_1 = P_{max}$.

SIC and $\Gamma > 1$

The rates of the 2 users with SIC, individual power constraint and Γ-approximation are

$$\begin{aligned} R_1^\Gamma &= \log_2(1 + \frac{p_1}{\Gamma}\|\mathbf{h}_1\|^2) && (4.57)\\ R_2^\Gamma &= \log_2(1 + \frac{p_2}{\Gamma}\mathbf{h}_2^H[\mathbf{I} + p_1\mathbf{h}_1\mathbf{h}_1^H]^{-1}\mathbf{h}_2) && (4.58)\\ &= \log_2\left(1 + \frac{p_2}{\Gamma}\left(\|\mathbf{h}_2\|^2 - \frac{p_1\|\mathbf{h}_1^H\mathbf{h}_2\|^2}{1 + p_1\|\mathbf{h}_1\|^2}\right)\right). && (4.59)\end{aligned}$$

The sum rate is then given by

$$\begin{aligned} C^\Gamma =& R_1^\Gamma + R_2^\Gamma && (4.60)\\ =& \log_2\overbrace{(1 + p_1\|\mathbf{h}_1\|^2 + \frac{p_2}{\Gamma}\|\mathbf{h}_2\|^2 + \frac{p_1p_2}{\Gamma}(\|\mathbf{h}_1\|^2\|\mathbf{h}_2\|^2 - \|\mathbf{h}_1^H\mathbf{h}_2\|^2))}^{g_1} && (4.61)\\ &+ \log_2\underbrace{\left(\frac{1 + \frac{p_1}{\Gamma}\|\mathbf{h}_1\|^2}{1 + p_1\|\mathbf{h}_1\|^2}\right)}_{g_2} = \log_2 g_1(p_1, p_2) + \log_2 g_2(p_1) = G(p_1, p_2). && (4.62)\end{aligned}$$

Because the power of user 2 does not influence the rate of user 1, user 2 will always transmit at maximum transmit power. With $p_1 \leq p_2 = P_{max}$ the functional G is only dependent on p_1. We differentiate $G(p_1)$ in order to find a maximum for the sum rate

$$G'(p_1) = \frac{g_1'(p_1)}{g_1(p_1)} + \frac{g_2'(p_1)}{g_2(p_1)} = 0.$$

$$G'(p_1) = \frac{\|\mathbf{h}_1\|^2 + \frac{(\|\mathbf{h}_1\|^2\|\mathbf{h}_2\|^2 - \|\mathbf{h}_1^H\mathbf{h}_2\|^2)P_{max}}{\Gamma}}{\log(2)(1 + p_1\|\mathbf{h}_1\|^2 + \frac{P_{max}}{\Gamma}\|\mathbf{h}_2\|^2 + \frac{p_1 P_{max}}{\Gamma}(\|\mathbf{h}_1\|^2\|\mathbf{h}_2\|^2 - \|\mathbf{h}_1^H\mathbf{h}_2\|^2)} \tag{4.63}$$

$$+ \frac{\left(1 + p_1\|\mathbf{h}_1\|^2\right)\left(-\frac{\|\mathbf{h}_1\|^2\left(1+\frac{p_1\|\mathbf{h}_1\|^2}{\Gamma}\right)}{(1+p_1\|\mathbf{h}_1\|^2)^2} + \frac{\|\mathbf{h}_1\|^2}{(1+p_1\|\mathbf{h}_1\|^2)\Gamma}\right)}{\left(1 + \frac{p_1\|\mathbf{h}_1\|^2}{\Gamma}\right)\log(2)} \tag{4.64}$$

After some transformations we finish with the following numerator which is set equal to zero.

$$0 = \|\mathbf{h}_1\|^2 P_{max}\left(-\|\mathbf{h}_1^H\mathbf{h}_2\|^2 p_1(2 + \|\mathbf{h}_1\|^2 p_1) + (\|\mathbf{h}_2\| + \|\mathbf{h}_1\|^2\|\mathbf{h}_2\| p_1)^2)\right) \tag{4.65}$$

$$+ \left((\|\mathbf{h}_1\| + \|\mathbf{h}_1\|^3 p_1)^2 - \|\mathbf{h}_1^H\mathbf{h}_2\|^2 P_{max}\right)\Gamma \tag{4.66}$$

The two zeros for $p_1 are$:

$$p_1 = -\frac{1 \pm \frac{\|\mathbf{h}_1^H\mathbf{h}_2\|^2\sqrt{P_{max}}\sqrt{\Gamma - 1}}{\sqrt{-\|\mathbf{h}_1^H\mathbf{h}_2\|^2 P_{max} + \|\mathbf{h}_1\|^2(\|\mathbf{h}_2\|^2 P_{max} + \Gamma)}}}{\|\mathbf{h}_1\|^2}.$$

One solution p_1^+ is always negativ and therefore not a valid transmit power while the other solution p_1^- can become positiv exactly when:

$$\frac{\|\mathbf{h}_1^H\mathbf{h}_2\|^2\sqrt{P_{max}}\sqrt{\Gamma - 1}}{\sqrt{-\|\mathbf{h}_1^H\mathbf{h}_2\|^2 P_{max} + \|\mathbf{h}_1\|^2(\|\mathbf{h}_2\|^2 P_{max} + \Gamma)}} > 1.$$

Fig. 4.10 shows the sum rate for a 2 user example ($n_T = 2, m_R = 4$, $\sigma_N^2 = 1$) and varying total transmit power P_{tot}. We see that the sum rate functional with SIC is always concave while the behavior without SIC depends on the total transmit power. For very large transmit powers ($P_{tot} > 10^5$) the sum rate functional also has a concave behavior for this example.

Detection with SIC and Sum Power Constraints

SIC and $\Gamma = 1$

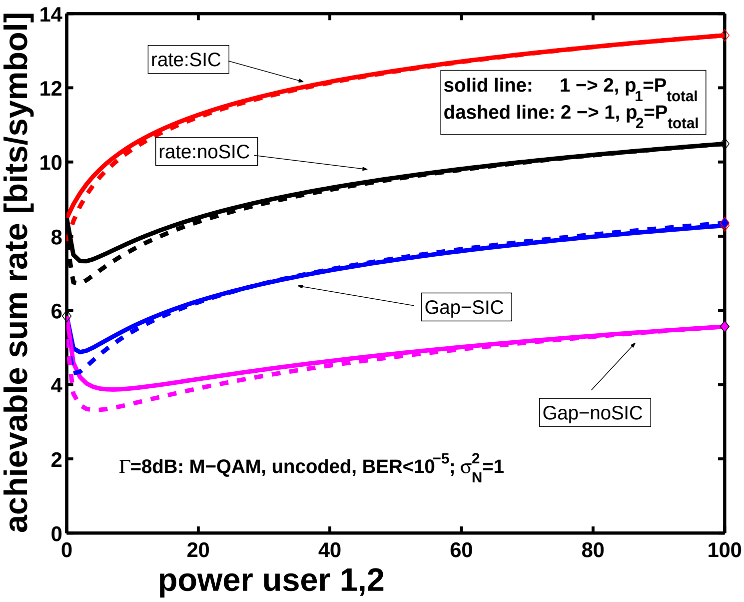

Figure 4.8: Achievable sum rate in the 2 user case for $P_{max} = 100$ as a function of p_1, p_2 under individual power constraint, ($n_T = 2, m_R = 4$, $\sigma_N^2 = 1$) and parallel detection (black and magenta) and SIC (red and blue) is assumed. Channel from (4.50).

The rates of the 2 users with SIC and sum power constraint are

$$R_1^\Gamma = \log_2(1 + p_1|\mathbf{h}_1\|^2) \tag{4.67}$$

$$R_2^\Gamma = \log_2(1 + p_2\mathbf{h}_2^H[\mathbf{I} + p_1\mathbf{h}_1\mathbf{h}_1^H]^{-1}\mathbf{h}_2) \tag{4.68}$$

$$= \log_2\left(1 + p_2\left(\|\mathbf{h}_2\|^2 - \frac{p_1\|\mathbf{h}_1^H\mathbf{h}_2\|^2}{1 + p_1\|\mathbf{h}_1\|^2}\right)\right). \tag{4.69}$$

The sum rate is then given by

$$C^{SIC} = R_1 + R_2 \tag{4.70}$$

$$= \log_2\overbrace{(1 + p_1\|\mathbf{h}_1\|^2 + p_2\|\mathbf{h}_2\|^2 + p_1p_2(\|\mathbf{h}_1\|^2\|\mathbf{h}_2\|^2 - \|\mathbf{h}_1^H\mathbf{h}_2\|^2))}^{g_1} \tag{4.71}$$

$$= \log_2 g_1(p_1, p_2). \tag{4.72}$$

With $p_2 = P_{total} - p_1$ the functional G is only dependent on p_1.

$$G(p_1) = \log_2\overbrace{(1 + p_1\|\mathbf{h}_1\|^2 + (P_{total} - p_1)\|\mathbf{h}_2\|^2 + p_1(P_{total} - p_1)(\|\mathbf{h}_1\|^2\|\mathbf{h}_2\|^2 - \|\mathbf{h}_1^H\mathbf{h}_2\|^2))}^{g_1} \tag{4.73}$$

Next we differentiate $G(p_1)$ in order to find the maximum sum rate.

$$G'(p_1) = \frac{g_1'(p_1)}{g_1(p_1)} = 0.$$

The maximum is found for p_1:

$$p_1 = \frac{1}{2}\left(\frac{(\|\mathbf{h}_1\| - \|\mathbf{h}_2\|)(\|\mathbf{h}_1\| + \|\mathbf{h}_2\|)}{\|\mathbf{h}_1\|^2\|\mathbf{h}_2\|^2 - \|\mathbf{h}_1^H\mathbf{h}_2\|^2} + P_{total}\right) \tag{4.74}$$

SIC and $\Gamma > 1$

The rates of the 2 users with SIC, sum power constraint and Γ-approximation are

$$R_1^\Gamma = \log_2(1 + \frac{p_1}{\Gamma}\|\mathbf{h}_1\|^2) \tag{4.75}$$

$$R_2^\Gamma = \log_2(1 + \frac{p_2}{\Gamma}\mathbf{h}_2^H[\mathbf{I} + p_1\mathbf{h}_1\mathbf{h}_1^H]^{-1}\mathbf{h}_2) \tag{4.76}$$

$$= \log_2\left(1 + \frac{p_2}{\Gamma}\left(\|\mathbf{h}_2\|^2 - \frac{p_1\|\mathbf{h}_1^H\mathbf{h}_2\|^2}{1 + p_1\|\mathbf{h}_1\|^2}\right)\right). \tag{4.77}$$

The sum rate is then given by

$$C^\Gamma = R_1^\Gamma + R_2^\Gamma \tag{4.78}$$

$$= \log_2\overbrace{(1 + p_1\|\mathbf{h}_1\|^2 + \frac{p_2}{\Gamma}\|\mathbf{h}_2\|^2 + \frac{p_1 p_2}{\Gamma}(\|\mathbf{h}_1\|^2\|\mathbf{h}_2\|^2 - \|\mathbf{h}_1^H\mathbf{h}_2\|^2))}^{g_1} \tag{4.79}$$

$$+ \log_2\underbrace{\left(\frac{1 + \frac{p_1}{\Gamma}\|\mathbf{h}_1\|^2}{1 + p_1\|\mathbf{h}_1\|^2}\right)}_{g_2} = \log_2 g_1(p_1, p_2) + \log_2 g_2(p_1) = G(p_1, p_2). \tag{4.80}$$

With $p_2 = P_{total} - p_1$ the functional G is only dependent on p_1. We differentiate $G(p_1)$ in order to find the maximum sum rate

$$G'(p_1) = \frac{g_1'(p_1)}{g_1(p_1)} + \frac{g_2'(p_1)}{g_2(p_1)} = 0.$$

The lengthy differentiation finishes with a polynomial equation which is of third order in p_1. Therefrom we know that all zeros have to be real valued or conjugated complex and that their number is either 1 or 3. In the latter case we find two local maxima or minima.

To illustrate the two user behavior we choose a fixed channel and vary the power allocated to each user under a sum power constraint. Fig. 4.9 shows the sum rate functional plotted over the transmit power of user 1. For convenience we set $\sigma_N^2 = 1$, the total transmit power $P_{tot} = 1000$ and Γ=8 dB which corresponds to uncoded M-QAM and a BER ¡ 10^{-5}. Due to the entanglement of the two channel vectors SIC can gain up to 3 bits/s/Hz for this example. The upper curve is the concave sum rate functional with SIC with a square to mark the maximum achievable rate. The curve below which describes the achievable sum rate without SIC is neither convex nor concave. The maximum sum rate is achieved when all power is given to user 2.

Below we see the resulting sum rate calculated with the SINR gap approximation. We clearly see that the curves are neither concave nor convex and several local maxima appear. Furthermore the sum rate is clearly depending on the detection order at the receiver (dotted and dashed curve). The sum rate curve without SIC remains neither convex nor concave.

Fig. 4.10 shows the sum rate functional again plotted versus the transmit power of user 1 under sum power constraint. We varied the total transmit power while keeping the channel fixed. The

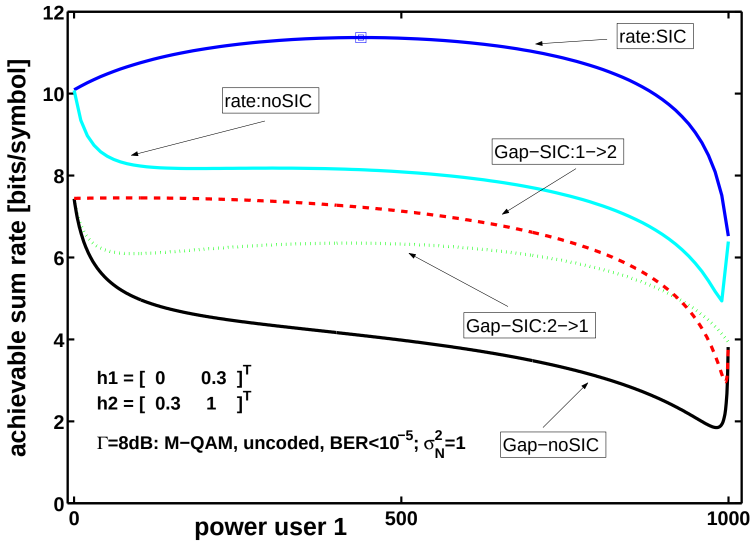

Figure 4.9: Achievable sum rate in the 2 user example for $P_{total} = 1000$ and sum power constraint with/without SIC (upper curves), $n_T = m_R = 2$. The sum rate is depicted as a function of p_1 if the SNR gap approximation is applied which is a central part of the SRPQ-algorithm proposed by Chung et. al

black solid lines represent the sum rate functional without SIC and up to $P_{tot} = 140$ the optimum power allocation is given by only supporting user 2 at full power. The red and dotted lines represent the functional when SIC is used. Here we clearly see again the concave behavior and already from $P_{tot} = 10$ both users are supported to reach the maximum sum rate.

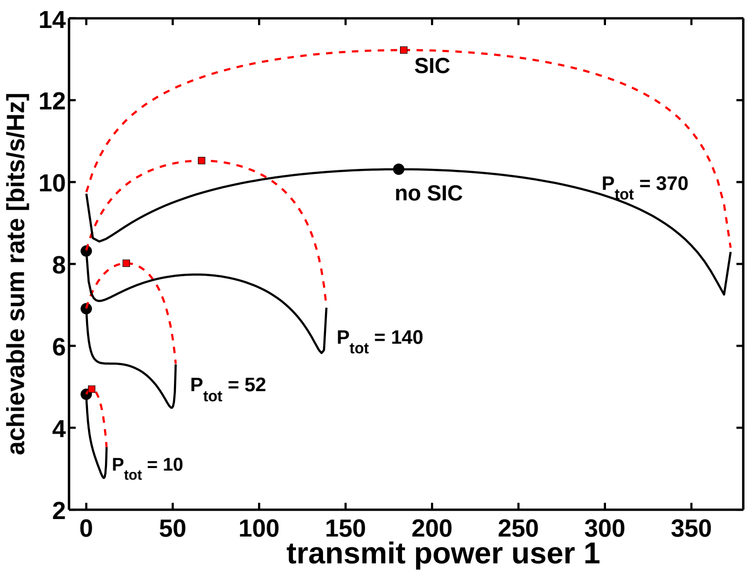

Figure 4.10: Achievable sum rate in the 2 user case as a function of p_1 under sum power constraint, ($n_T = 2, m_R = 4$, $\sigma_N^2 = 1$) and parallel detection (solid line) and SIC (dashed line) is assumed. Circles / squares indicate maximum rate without / with SIC. Channel from (4.50).

4.2 Pre-coding Strategies

4.2.1 The idea of Pre-coding

Pre-coding strategies are means or methods applied at the transmitter to facilitate detection at the receiver. In contrast to the detection methods, now channel state information (CSI) is required at the transmitter side, and not necessarily at the receiving side. This is naturally the case in symmetric time-division-duplex (TDD) systems, where each receiver in turn also acts as a transmitter. Provided the time between up-link and down-link is short relative to the fading coherence time [J.G00], the estimates from the receiving direction can be used for the subsequent transmission. In other settings e.g. quasi-static channels the CSI can be made available to the transmitter via some backward channel.

We formalize this into a pre-coding operation $\tilde{\mathbf{F}}$ applied on the data symbols $\mathbf{d}$, thus sending $\mathbf{x}$ from the transmitter

$$\mathbf{x} = \tilde{\mathbf{F}}\mathbf{d}, \tag{4.81}$$

The belonging detection operation at the receiver is

$$\begin{aligned} \mathbf{d}' &= \mathbf{F}_{Rx}\mathbf{y} && (4.82) \\ &= \mathbf{F}_{Rx}(\mathbf{H}\mathbf{x} + \mathbf{n}) && (4.83) \\ &= \mathbf{d} + \mathbf{F}_{Rx}\mathbf{n} && (4.84) \end{aligned}$$

Note, that pre-coding can not be applied in the multiple access channel at all, since the users are not able to cooperate. However, for the dual setting of the broadcast channel, some pre-coding methods that we will discuss are able to improve the detection performance of the users, even though the users/MTs can not cooperate on the detection task. Note that for the those pre-coding methods, the matrix $\mathbf{F}_{Rx}$ from (4.84) has diagonal form.

4.2.2 SVD-MIMO Transmission and Waterfilling

If CSI is available at the transmitter and receiver, we can apply the rather straightforward concept of SVD-MIMO transmission, which is in fact information-theoretically optimum towards capacity as discussed in sec. 3.2.

The SVD based transmission scheme requires joint processing both on the transmit and the receive side of the communication system, which means that it can only be implemented in centralized systems, such as the single user MIMO system described in sec. 4.1.1.

The necessary pre-coding and decoding matrices are obtained from the singular-value-decomposition

(SVD) [PTVF92, GL96] of $\mathbf{H}$, the matrices $\mathbf{U}$, $\mathbf{D}$, and $\mathbf{V}$ are obtained like in (5.9) as

$$\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^H \tag{4.85}$$

with $\mathbf{U}$ and $\mathbf{V}$ unitary matrices and $\mathbf{D} = \mathrm{diag}(\lambda_1, ..., \lambda_k)$ a diagonal matrix with the singular values of $\mathbf{H}$.

Using the transmit signal $\mathbf{x} = \mathbf{V}\mathbf{d}$, the appropriate receiver processing is

$$\begin{aligned} \mathbf{d}' &= \mathbf{D}^{-1}\mathbf{U}^H\mathbf{y} && (4.86) \\ &= \mathbf{D}^{-1}\mathbf{U}^H(\mathbf{U}\mathbf{D}\mathbf{V}^H\mathbf{x} + \mathbf{n}) && (4.87) \\ &= \mathbf{d} + \mathbf{D}^{-1}\mathbf{U}^H\mathbf{n} = \mathbf{d} + \mathbf{n}' && (4.88) \end{aligned}$$

where $\mathbb{E}[\mathbf{n}'\mathbf{n}'^H] = \sigma_N^2\mathbb{E}[\mathrm{diag}(\lambda_1^{-2}, ..., \lambda_k^{-2})]$ gives the average SNR for the parallel AWGN channels under the assumption of i.i.d. noise at all Rx antennas.

The multiplication with the unitary matrices do neither enhance the transmit power nor enhance the noise power at the Rx, hence the separation into parallel channels is performed in an ideal manner. In fact, the information theoretic capacity of the MIMO channel is equal to the sum of the capacities of the parallel sub channels obtained from the SVD transmission scheme, cf., e.g., [Tel99]. It turns out that the noise powers resulting are distributed over a wide range and consequently some form of adaptive transmission is necessary to achieve a reasonable average (uncoded) BER. In [Ven02] it was observed that without any power and rate adaptation the performance of this scheme approaches that of linear equalization for high SNR. The author argues that here the sum of the noise powers of the overall transmission channel is the same for linear pre-coding (ZF) and the SVD scheme, since $\mathrm{trace}(\mathbf{D}^{-2}) = \mathrm{trace}(\mathbf{H}^{-1}{\mathbf{H}^{-1}}^H)$.

Since the parallel channels are independent, mere power-adaptation, i.e., multiplying the transmit vector by $\mathbf{D}^{-1}$, does not significantly change the error rate performance, since it the transmit power enhancement is limited due to the limited dynamic range of the Tx amplifiers or the limited resolution of the DAC, therefore this strategy is not advised for an application.

The right way to deal with these parallel sub-channels is to adapt the transmission rates according to the SNR of each sub-channel. If this is done properly, asymptotically the (constellation-constrained) capacity of the MIMO channel can be fully utilized. An additional power allocation at the Tx is possible if sub-channels are switched off in the low SNR regime. Now, the saved power of one or several streams can be redistributed to the remaining data streams to increase throughput. Ideally, this is done in a water-filling like fashion as long as the per antenna transmit power limit is not violated. For the uncoded transmission, this adaptive bit and power allocation has the effect of equalizing the bit error rates for all bits transmitted while achieving the maximum sum throughput. The measurement results of an implementation in the real-time test-bed are found in Fig. 6.20.

[Win04] showed that with both the transmitter side adaptation to the channel matrix and the spatial loading (Off or QPSK), the resulting bit error rate is superior to that of maximum-likelihood detection for the Rx detection scenario with fixed rate per antenna transmission, because of the ideal separation of the sub-channels. The BER resulting without loading approaches that of linear ZF detection.

For MIMO systems where joint processing and (rate/power) adaptation is possible, the SVD based scheme is the optimum approach.

4.2.3 Linear Pre-coding

The transmitter-side equivalent to the linear Rx equalization discussed in section 3.5.1, and the equivalent approach is linear pre-equalization. In the ZF or MMSE case this corresponds to a multiplication with a pre-coding matrix $\tilde{\mathbf{F}}$ and transmitting $\mathbf{x} = \tilde{\mathbf{F}}\mathbf{d}$

with

$$\tilde{\mathbf{F}} = \frac{1}{\alpha}\mathbf{H}^{\dagger} \tag{4.89}$$

which leads to a received signal of

$$\mathbf{y} = \mathbf{d} + \alpha\mathbf{n} \tag{4.90}$$

where the parameter α is used to keep the transmitted power constrained.

A sum power constraint leads to

$$\alpha = \sqrt{\text{trace}(\mathbf{H}^{\dagger}\mathbf{H}^{\dagger^H})}, \tag{4.91}$$

while the more realistic constraint of limited transmit power per antenna leads to

$$\alpha = \sqrt{\max(\tilde{\mathbf{f}}^i\tilde{\mathbf{f}}^{i^H})}, \tag{4.92}$$

where $\tilde{\mathbf{f}}^i$ is the i-th row of the pre-coding matrix $\tilde{\mathbf{F}}$. The transmit constraint from (4.92) was implemented for the experiments on Adaptive Channel Inversion described in sec. 6.3.

In the following no such transmit power will be assumed in order to study the general performance and to evaluate the necessary dynamic range at the transmit side which is very important towards a reduction of transmitted radiation [Min04], which helps reducing the inner-cell and inter-cell interference.

If transmit power limitations of any kind are taken into account then this transforms directly into an SNR decrease according to (4.90).

Linear Channel Inversion and Joint Transmission

We assume n_T transmit and m_R receive antennas and a flat fading channel matrix $\mathbf{H}$ of size $m_R \times n_T$. With rank($\mathbf{H}$) $= min(n_T, m_R) = m$, data transmission can be performed over m parallel sub-channels. Furthermore we assume the transmitter at the BS has more or an equal number of antennas than the MT. With the given transmission equation $\mathbf{y} = \mathbf{Hx} + \mathbf{n}$ and channel state information (CSI) at the transmitter, the channel can be pre-compensated prior to the transmission of the data symbols which we then will call **Channel Inversion** (CI). Under the assumption of channel reciprocity the Tx can obtain CSI prior to the transmission from a channel measurement into the opposite direction.

Several groups use slightly different terms to distinguish between variations of the CI. Since the terms Linear Channel Inversion and Joint Transmission are widely used a short explanation will be given.

Linear Channel Inversion (LCI) is a pre-equalization technique which fully inverts the phases and the attenuation of the channel at the transmit side. The resulting transmission channel behaves like an AWGN channel. LCI compensates the path loss and even ill conditioned channels to full extend. This can require a huge dynamic range at the Tx depending on the channel statistics and the number of Tx and Rx antennas. Therefore even the expectation value of the required transmit power might be unlimited.

Joint Transmission (JT) works in principle the same way as LCI but with the important difference that the transmit power is held constant. Therefore all channel variations which require more than the maximum transmit power translate directly into SNR loss at the Rx. Therefore the BER shows the same diversity order as known from the ZF-detector at the Rx [BMWT00].

For convenience we will refer to the linear pre-coding as **Channel Inversion** (CI) throughout the following paragraphs.

Depending on the number of Tx and Rx antennas, CI is performed at the Tx by partly or fully inverting the channel before transmission using either the equivalent of the ZF or MMSE solution known from the Rx detection schemes.

The ZF pre-coding matrix is easily obtained by inverting the channel matrix, while the construction of the optimum MMSE transmit solution is more involved, since the solution with the minimum transmit power has to be found. Based on the duality of the up-link and down-link channel [SB02b, BS02a, Sch02] and also [VT03] showed the equivalence of the MMSE detection solution and the MMSE transmit filter, which minimize the MSE at the Rx while transmitting with minimum transmit power.

Now, for the following we will write $\mathbf{H}^\dagger$ for convenience which can be the ZF or MMSE solution, depending on knowledge at the Tx about the receiver noise power. We then can formalize the linear pre-coding according to the following formulas:

Down-link Channel Inversion ($n_T > m_R$)

Tx pre-processing: $\mathbf{x}^{CI}_{Down} = \mathbf{H}^\dagger \mathbf{d}$.

Reconstructed data at Rx

$$\mathbf{d}' = \mathbf{H}\mathbf{x}_{Down}^{CI} + \mathbf{n} = \mathbf{d} + \mathbf{n}. \tag{4.93}$$

The signals to transmit ($\mathbf{d}$) are pre-processed by a multiplication with $\mathbf{H}^\dagger$ denoting the Moore-Penrose pseudo-inverse of $\mathbf{H}$, thus equalizing the SNR in all transmission channels. In this way each data signal is fully reconstructed at one Rx antenna and forced-to-zero at all other antennas by destructive interference thus creating parallel channels at the Rx antennas without post-processing for signal separation.

Up-link Channel Inversion ($n_T < m_R$)

For the case of less transmit than receive antennas the number of streams which can be multiplexed is limited to n_T. Then two principle options exist, first performing channel inversion on a reduced set of receive antennas or second using the well known SVD approach, but instead of water-filling we compensate the attenuation on the eigen-channel such that all parallel channels will have the same SNR. The first case is equivalent to the previously described method, but due to the same numbers of transmit and receive antennas we have to expect a high increase in transmit power.

The second approach uses the following Tx pre-processing: Tx: $\mathbf{x}_{Up}^{CI} = \mathbf{V}\mathbf{D}^{-1}\mathbf{d}$.

Reconstructed data at Rx

$$\mathbf{d}' = \mathbf{U}^H(\mathbf{H}\mathbf{x}_{Up}^{CI} + \mathbf{n}) = \mathbf{d} + \mathbf{U}^H\mathbf{n}. \tag{4.94}$$

(The matrices $\mathbf{U}, \mathbf{D}, \mathbf{V}$ are derived from singular value decomposition of $\mathbf{H}$ ($\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^H$). $[\cdot]^H$ means Hermitian Conjugate, $[\cdot]^{-1}$ means inverse matrix.)

Since the receiver has more antennas than data streams to be detected, post processing is necessary at the BS. Since $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices no noise enhancement is observed after the linear operation in (4.94).

Note, that LCI has the advantage that all channels are equalized to the same link quality which allows the same rate and coding for all multiplexed streams. If a post-processing is given at the Rx as in (4.94), then a channel inversion with $\mathbf{x} = \mathbf{V}\mathbf{D}^{-1}\mathbf{d}$ is simple towards common rate and coding but already at reasonable SNR an independent rate control and uniform power allocation per data stream is the SVD-MIMO solution which achieves the optimum sum rate. Therefore channel inversion makes only sense when the signal processing is constrained to the transmitter/BS alone.

Qualitatively, CI performs like data transmission over a channel with additive white Gaussian noise (AWGN). To illustrate the BER performance we compare standard Rx detection schemes (ZF, VBLAST)with CI. The simulation assumes perfect CSI and a Rayleigh block fading channel.

Fig. 4.11 shows the simulated BER of MIMO systems with perfect CSI and BPSK modulation which use Zero-Forcing (ZF), VBLAST or CI. Taking into consideration that for CI the average transmit power depends on the number of Tx and Rx antennas [JWHJ01] the abscissa of Fig. 4.11 shows the average transmitted sum power per noise power at one receive antenna to give a fair comparison of the transmission techniques. CI performs better than ZF with BERs of less than 10^{-3} but performs worse than VBLAST until a BER of 10^{-5} or below where the curves cross each

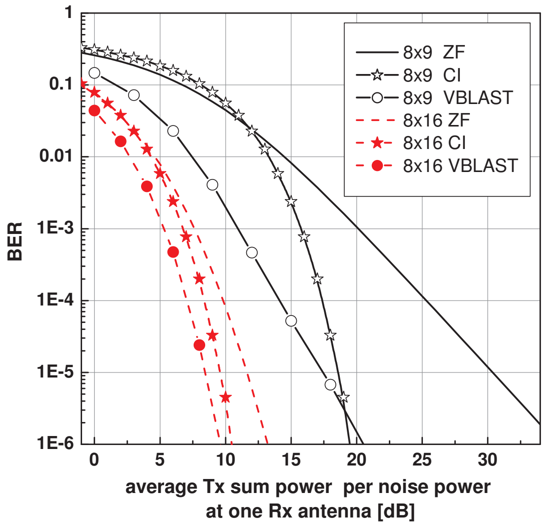

Figure 4.11: Up-link BER performance for two MIMO systems using ZF (no symbols), VBLAST without error propagation (○) and CI (⋆) with 8 Tx antennas and 9 (–) / 16 (- -) Rx antennas, perfect CSI is assumed.

other. If the antenna diversity is sufficiently high (8 Tx / 16 Rx), all techniques perform quite similar (dashed curves).

Adaptive Channel Inversion

A very important aspect towards high data throughput and reliable low BERs is the adaptation of the transmission scheme to the actual channel realization thus avoiding data loss when transmitting over bad channels. This strategy requires at least little channel knowledge at the Tx or some CSI which has to be fed back from the receiver but fortunately with LCI CSI is already available at the Tx. The concept of channel adaptive transmission can be combined with adaptive modulation and coding to maximize throughput at a desired BER with a common modulation and coding scheme on all parallel links.

Towards Channel Inversion channel aware means step-wise switching off multiplexed channels until a reasonable condition number of the reduced channel matrix is reached or the transmit power enhancement is limited to a value which guarantees a satisfactory SNR at the Rxs. Following this strategy we can provide a reliable linearly pre-coded transmission which can meet required BER targets simply by reducing the number of simultaneously multiplexed links according to the rank of the actual channel.

We call a transmission scheme using Channel Inversion with a variable number of multiplexed data streams ***Adaptive Channel Inversion*** (ACI).

A more sophisticated version of this strategy is realized when ACI is combined with adaptive bit-loading. Since the whole transmission channel is pre-equalized at the Tx all data streams will experience the same SINR at the Rx. Therefore bit-loading offers the possibility of trading BER performance and sum rate against each other. Example: with four antennas a maximum of

4 data streams can be multiplexed. Assuming that for a certain channel realization 4 $\times$ QPSK modulation (8 bit/s/Hz) can meet a certain maximum BER target. If the reduction to three parallel links allows a 16-QAM modulation then this mode can be chosen to increase the sum throughput by 50%.

4.2.4 Non-linear Pre-coding

The equivalent to non-linear decoding strategies e.g. VBLAST can be found in non-linear pre-coding. Due to a calculation of a pre-coded transmission signal over several iterations the required transmit power enhancement can be reduced significantly especially if the channel is mal-conditioned [HSB03].

The effect of transmit power enhancement which appears when pre-coding is applied is the equivalent of the noise power enhancement known from decoding at the receiver side[HPJ+02].

Since all signals to be sent are perfectly known to the transmitter the effect of error propagation known from iterative decoding structures at the Rx has no equivalent counter-part. Therefore, a pre-coding with a DFE[3] similar structure can achieve a BER performance which can be reached at the Rx only with a genie-aided DFE structure[Win04]. Furthermore, if full Tx pre-coding is applied like for a down-link scenario with decentralized Rxs [FWLH02], then the pre-coding order can be chosen freely because it is not visible to the Rxs.

The technique of non-linear pre-coding can be of great benefit e.g. for the down-link transmission towards distributed MTs if applied correctly. The potential of transmit power reduction when using non-linear pre-coding instead of linear pre-coding like ACI will be discussed in more detail in section 4.2.5.

The Principle of Costa Precoding

A basic result from information theory is Costa's "writing on dirty paper result" [Cos83], which can be informally summarized as follows:

When transmitting over a channel, any interference which is known a priori to the transmitter does not affect the channel capacity. That means, by appropriate coding, transmission at a rate equal to the capacity of the channel without this interference is possible.

An illustration of the setup used for the prove this is shown in Fig. 4.12 with the original denomi-

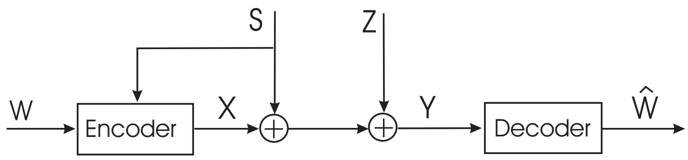

Figure 4.12: Block diagram of Costa's "Writing on Dirty Paper" Pre-coding.

nations of [Cos83], where the variable W is the message to be transmitted, S is interference known to the encoder, X is the channel input (which has to satisfy some power constraint), Z additive

[3] Decision Feedback Equalizer

noise, and Y the channel output used by the decoder to produce the estimate $\hat{W}$. Using a random coding argument it is shown that the capacity of this channel, i.e., the maximum mutual information $I(W;\hat{W})$, is equal to that of the AWGN channel with the same noise power (equivalently, the setup in Fig. 4.12 with $S = 0$).

We can easily extend this to a scheme with multiple interfering sub-channels: Considering these sub-channels in some arbitrary order, the encoding for the first sub-channel has to be performed accepting full interference from the remaining channels, since at this point the interference is unknown. For the second sub-channel, however, if the transmitter is able to calculate the interference from the first sub-channel, "Costa pre-coding" of the data is possible such that the interference from the first sub-channel is taken into account. Generally, in the k-th sub-channel considered, Costa pre-coding is possible such that interference from sub-channels 1 to $k-1$ is ineffective. (Interference from the k-th sub-channel back to the previous sub-channels decreases the reliability, or equivalently the capacity, of those channels, and has to be accounted for by additional measures).

We can apply this result to the present setting of a MIMO channel (cf. also [SB02a]): If the pre-coding operation contains a Costa pre-coder, no interference can be observed from lower number sub-channels into higher number sub-channels. This is the same as saying that we can simply disregard the part of the channel matrix that is below the main diagonal. (In fact, results concerning the capacity of (MIMO) broadcast channels [WJ01, VJG03, JG04, VT03, CS03] have shown that Costa pre-coding is an essential ingredient to achieve the capacity of these channels).

In section 4.1.3 we have already made use of the fact that it is possible to subtract already decided signals before a next decision is made (successive interference cancellation). For a realization often the channel matrix $\mathbf{H}$ is transformed into a lower triangular matrix with an orthonormal operation (see QLD in sec. 5.2.3). In this way interference from lower-index sub-channels into higher-index sub-channels is completely eliminated, and together with Costa pre-coding adjusted to this modified transmission channel matrix, effectively only a diagonal matrix remains for the transmission.

It turns out that a simple scheme for Costa pre-coding works analog to the feed-back part of a decision-feedback-equalizer, used here at the transmitter side and with the nonlinear decision device replaced by a modulo-operation. This scheme is also known as Tomlinson-Harashima precoding (THP) [Tom71, HM72], and the link between THP and Costa pre-coding was first explored in [ESZ00].

As a prominent example of non-linear pre-coding for the multiple antenna broadcast channel with decentralized receivers THP will be visited in more detail in the following subsection.

Tomlinson-Harashima Pre-coding

Tomlinson-Harashima Pre-coding can be seen as the simplest implementation of Costa pre-coding possible; in [Egg01] this scheme is called "scalar Costa scheme" and was applied to digital watermarking of arbitrary host signals, e.g., images. In the digital transmission context, however, this has long been known as Tomlinson-Harashima pre-coding [Tom71, HM72] for inter-symbol-interference channels [Fis02]. Tomlinson-Harashima pre-coding for MIMO Channels was proposed in [FWLH02], and independently in [GC02] under the name "vectored transmission" for DSL systems; similar ideas can also be found in [CS03].

To give a intuitive introduction on how the pre-coding actually works we will assume a complex-valued transmission model of the type

$$\tilde{\mathbf{y}} = \mathbf{B}\mathbf{x} + \tilde{\mathbf{n}}, \tag{4.95}$$

where $\mathbf{B}$ is the effective transmission matrix, i.e., a transformed version of $\mathbf{H}$ such that the interference which Costa pre-coding, and equivalently THP taken into account is minimized (i.e., the upper triangular part of the matrix $\mathbf{B}$ is minimized). We additionally assume that the effective transmission matrix is scaled such that it has unit diagonal, i.e., $b_{kk} = 1, k = 1, ..., K$. For example, using the QL decomposition (sec. 5.2.3) of $\mathbf{H}$ into

$$\mathbf{H} \quad = \quad \mathbf{F}^H \cdot \mathbf{S} \tag{4.96}$$

with ($\mathbf{F}$ orthonormal, $\mathbf{S}$ lower triangular) we can force the upper triangular part to be zero, obtaining $\mathbf{B} = \text{diag}(1/s_{11}, ..., 1/s_{KK})\mathbf{S}$. We then obtain $\tilde{\mathbf{y}} = \mathbf{B}\mathbf{x} + \tilde{\mathbf{n}}$ as the equivalent transmission channel.

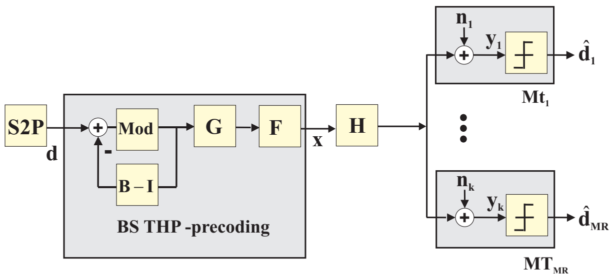

Figure 4.13: Block diagram of THP for decentralized receivers. $\mathbf{d}$ is the parallelized data vector, $\mathbf{x}$ the sent transmission signal, $\mathbf{G}$ is the diagonal power allocation or gain matrix and $\mathbf{F}$ is the beam-forming or feed forward matrix, $\mathbf{B} - \mathbf{I}$ is the feed-back matrix.

Now, the pre-coding procedure analogous to Fig. 4.13 works the following.

Assuming a 4-user example, the pre-coding will be started with the signal for user 1. Then the following symbol-by-symbol pre-coding is used:

For user 1 we simply send the plain message d_1. For user 2 we do not send the plain message d_2, instead we will send the message d_2 subtracted by the interference which user 2 will be exposed from the message sent to user 1. User 3 will have a prepared message which consists of the desired signal d_3 subtracted by the interference from the signals which are sent to user 1 and user 2 and so on. Since the symbols to be sent are perfectly known to the Tx error propagation is not an issue here, at least with perfect CSI. Note, that the signals to be send are not straight forward $x_1, ..., x_4$ instead we will use $\tilde{x_1}, ..., \tilde{x_4}$ which are the equivalent representatives in the corresponding Voronoi-cell which have the minimum transmit energy.

$$x_1 \quad = \quad d_1 \tag{4.97}$$
$$x_2 \quad = \quad d_2 - B_{21}x_1 \tag{4.98}$$
$$x_3 \quad = \quad d_3 - B_{31}x_1 - B_{32}\tilde{x_2} \tag{4.99}$$
$$x_4 \quad = \quad d_4 - B_{41}x_1 - B_{42}\tilde{x_2} - B_{43}\tilde{x_3} \tag{4.100}$$

This so-called lattice reduction technique is represented by a modulo operation (Mod) in the block diagram of Fig. 4.13. The darker region in Fig. 4.14 shows the Voronoi-region when 16-QAM

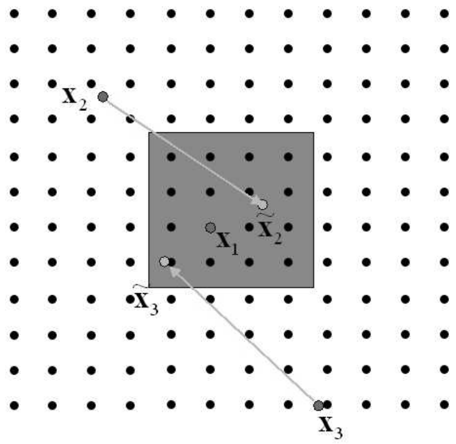

Figure 4.14: Exemplary pre-coding with THP using modulo operations to map signals back into the basic lattice (Voronoi region).

constellations are used for transmission. We clearly see that the pre-coding with a subtraction of several values can produce symbols to be sent which are outside the lattice basic cell (dark), therefore a modulo operation maps those symbols back into the basic cell. This operation is in principle what makes Costa's pre-coding of causal known interference possible without enhancing the transmit power. Note, that when using a modulo operation a slight increase in average transmit power is experienced since now basically the whole Voronoi region becomes a valid transmit signal space, therefore

$$\mathbb{E}[\tilde{\mathbf{x}}^H \tilde{\mathbf{x}}] \geq \mathbb{E}[\mathbf{d}^H \mathbf{d}]. \tag{4.101}$$

The example in Fig. 4.15 shows 2000 random non-linear pre-coded signals vectors (16-QAM) for 4 users / Rx-antennas. The belonging exemplary channel $\mathbf{H}$ is given by

$$\mathbf{H} = \begin{pmatrix} -0.813 + 0.302i & 1.735 - 1.602i & -0.937 + 0.865i & 0.417 + 0.181i \\ -0.817 - 0.452i & -0.047 + 0.289i & -0.481 + 0.229i & 1.720 + 0.351i \\ 0.346 - 0.770i & -0.177 - 0.421i & 1.278 - 0.659i & -0.847 - 0.659i \\ -0.125 + 0.788i & -0.101 - 0.049i & 0.435 + 0.797i & -0.267 + 0.471i \end{pmatrix},$$

which results in the equivalent effective channel $\mathbf{B}$

$$\mathbf{B} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0.056 - 0.145i & 0 & 0 & 0 \\ 0.645 + 0.034i & -0.751 - 0.110i & 0 & 0 \\ -0.260 + 0.108i & 0.187 - 0.036i & -0.290 + 0.438i & 0 \end{pmatrix}$$

used for the pre-coding.

The next step is a signal scaling by the gain matrix $\mathbf{G}$ and a multiplication with the beam-forming matrix $\mathbf{F}$, before the signals are jointly transmitted from the Tx antennas. In the reality of an application the scaling with $\mathbf{G}$ will be performed at the Rx side to avoid an increase in transmit

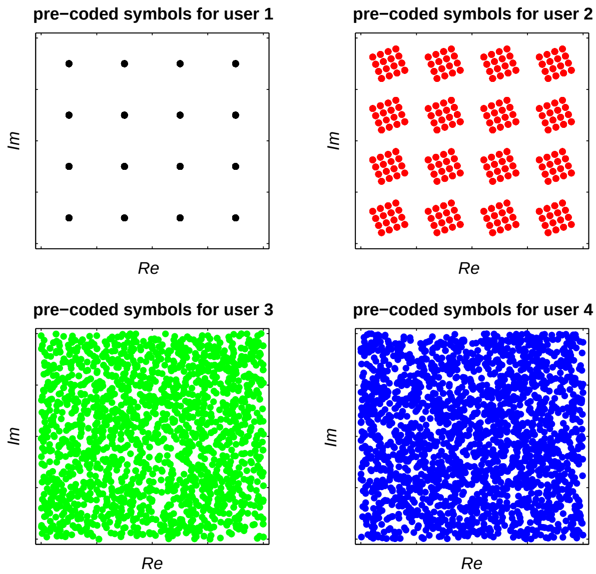

Figure 4.15: Pre-coded signals with THP using modulo operations for a 4 data stream example for a fixed channel. The pre-coded signals are all limited to the basic lattice (Voronoi region).

power which will cause clipping by the Tx amplifiers. But, the principle remains unchanged. Furthermore the modulo operation at the Tx has to be cancelled by another modulo operation at the Rx. Then the received signals are ready for detection, free of cross channel interference due to the pre-coding.

In principle, now all non-linear decoding techniques as ZF-VBLAST or MMSE-VBLAST can find their counterpart as pre-coding technique as well. In [Win04] it was shown that MMSE-Tomlinson-Harashima Pre-coding with VBLAST ordering gives a superior performance compared to the standard THP which is equivalent to a ZF pre-coding. The VBLAST ordering is a close to optimum pre-coding order as observed by [HSB03, Win04, Joh04]. A further performance increase can be obtained by so-called inflated lattice techniques [Win04] which we mention only for the sake of completeness.

4.2.5 Transmit Power Reduction Strategies

Transmit power reduction is of eminent importance whenever pre-coding techniques are used. Since this is mainly an option for a down-link scenario from the BS to the MTs, where the BS has sufficient signal processing power and CSI while several MTs can be supported simultaneously by spatial multiplexing. This scenario is often denoted as a multi-user down-link or braodcast scenario with decentralized receivers, implying that the MTs are not able to perform any joint signal processing.

Multi-User Down-link Strategies

We consider a down-link transmission system with a BS using n_T transmit antennas and m_R MTs with one receive antenna each. To have a fair comparison between all schemes we limit ourselves to the case of $(n_T \geq m_R)$. It has to be noted that the fully decorrelating schemes of LCI, JT, RKI and THP can only support a maximum number of MTs which equals $min(n_T, m_R)$ while JCBF is not limited this way because of its non-fully decorrelating approach.

The channel is assumed to be perfectly known at the BS only. Each of the MTs will be supplied with an independent data stream and the average signal-to-noise-plus-interference-ratio requirements (SINR) are assumed to be the same for all MTs.

The complex valued transmission equation reads

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{4.102}$$

where $y \epsilon \mathbb{C}^{m_R}$ is the receive signal vector at the Rx antennas, $\mathbf{H} \epsilon \mathbb{C}^{m_R \times n_T}$ is the channel matrix which is assumed to be a flat block fading channel, $\mathbf{x} \epsilon \mathbb{C}^{n_T}$ is the pre-coded and transmitted data vector $\mathbf{d} \epsilon \mathbb{C}^{m_R}$ and $\mathbf{n} \epsilon \mathbb{C}^{m_R}$ is the independently distributed additive white Gaussian noise at the Rx with the same variance σ_N^2 for all MTs. The transmitted signal vector $\mathbf{x}$ is given by

$$\begin{bmatrix} x_1 \\ x_1 \\ \vdots \\ x_{n_T} \end{bmatrix} = \begin{pmatrix} \cdot & \cdot & \cdot \\ \vdots & \mathbf{BF} & \vdots \\ \cdot & \cdot & \cdot \end{pmatrix} \cdot \begin{pmatrix} g_{11} & \cdot & 0 \\ \vdots & \mathbf{G} & \vdots \\ 0 & \cdot & g_{m_R m_R} \end{pmatrix} \cdot \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_{m_R} \end{bmatrix} \tag{4.103}$$

where $\mathbf{G}$ is a diagonal power allocation matrix to weigh each data stream individually and the columns of $\mathbf{BF}$ are the unitary beam-forming vectors of each data stream. Therefore if $\mathbf{G}$ and $\mathbf{BF}$ are once calculated for the different down-link transmission schemes we are interested in trace($\mathbf{G}$) which gives the required sum transmit power for a given channel realization and SINR requirements.

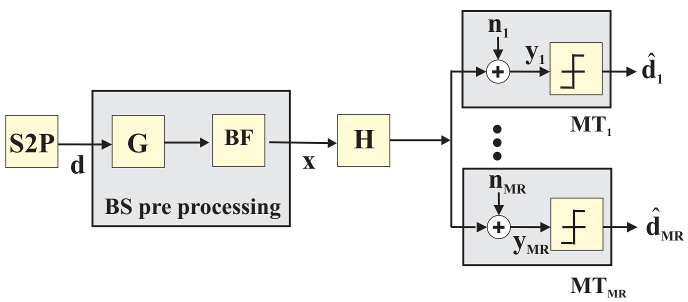

Figure 4.16: Block diagram of joint Tx pre-processing for all data streams. $\mathbf{d}$ is the parallelized data vector, $\mathbf{x}$ the sent transmission signal, $\mathbf{G}$ is the diagonal power allocation matrix and $\mathbf{BF}$ is the beam-forming matrix.

The transmission block to illustrate the principle of all investigated down-link transmission scheme is shown in Fig. 4.16. Therefore, in the following we will keep this decomposition of the transmit preprocessing and discuss the structure of $\mathbf{G}$ and $\mathbf{BF}$.

In the following we will give a comparison of several pre-coding strategies towards the transmit power enhancement. To have a fair comparison between all schemes we limit ourselves to $n_T > m_R$ otherwise the power enhancement factor (PEF) might not be limited.

We define the following measure as **Power Enhancement Factor**(PEF):

$$\mathbf{PEF} = \frac{\mathbb{E}[\mathbf{x}^H\mathbf{x}]}{\mathbb{E}[\mathbf{d}^H\mathbf{d}]} \tag{4.104}$$

where $\mathbf{x}$ is the transmitted symbol vector and $\mathbb{E}[\mathbf{x}^H\mathbf{x}]$ the expectation of the transmit power. $\mathbf{d}$ is the data symbol vector where d_i corresponds to the message to be sent to the i-th MT.

Linear Channel Inversion (LCI) A linear down-link strategy is LCI [JHJvH01, IHRF04] as introduced in section 4.2.3. Here all pre-coding is done with a one-step pre-distortion of the expected known transmission channel $\mathbf{H}$. Without knowledge about the actual noise level at the Rx antennas the solution with minimum Tx power is the Moore-Penrose-pseudo-inverse $\mathbf{x} = \mathbf{H}^\dagger \cdot \mathbf{d}$. Therefore no separate power allocation and beam-forming is performed

$$\mathbf{BF} \cdot \mathbf{G} = \mathbf{H}^\dagger. \tag{4.105}$$

$$\hat{\mathbf{d}} = \mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} = \mathbf{H}\mathbf{H}^\dagger\mathbf{d} + \mathbf{n} = \mathbf{d} + \mathbf{n} \tag{4.106}$$

The new over all channel in (4.106) behaves like fully decorrelated parallel AWGN channels with the same signal power at all Rx.

The transmission is done in a way that the signal power at the desired receive antenna is one and at all other antennas spatial nulls have to be placed. This spacial nulling is very power consumptive if the channel is bad conditioned [4]. Therefore LCI may encounter feasibility problems in many macro cellular scenarios with a large difference in distance between the MTs and the BS or a high Rician factor. For more details see the discussion on limited transmit dynamics in section 4.3.2.

Besides the actual noise level at the Rx and the given SINR requirements, the necessary transmit power is determined by the power enhancement factor $\mathbf{PEF}_{LCI}$ which describes the amount of additional transmit power which is necessary when compared to parallel AWGN transmission.

$$\mathbf{PEF}_{LCI} = \sum_{i=1}^{\mathrm{rank}(\mathbf{H})} \frac{1}{\lambda_i^2} = \mathrm{trace}(\mathbf{D}^{-2}) \tag{4.107}$$

$\mathbf{PEF}_{LCI}$ has to be computed, in general. λ_i are the singular values (SVs) of $\mathbf{H}$ obtained by singular value decomposition (SVD). Especially, small SVs lead to an increase in the required transmit power.

For the special case of a pure Rayleigh channel the average of $\mathbf{PEF}_{LCI}$ can be given as closed form solution [JWHJ01]

[4] A bad conditioned channel is close to singular, then the ratio between highest and lowest singular value $\lambda_{max}(H)/\lambda_{min}(H) \gg 1$. Theoretically, best condition is given for $\lambda_{max}(H) = \lambda_{min}(H)$ and therefore the least power is required for down-link transmission.

$$\text{Rayleigh: } \mathbf{PEF}_{LCI} = \mathbb{E}\left[\sum_{i=1}^{rank(\mathbf{H})} \frac{1}{\lambda_i^2}\right] = \frac{m_R}{n_T - m_R} \tag{4.108}$$

This expression is defined only for $n_T > m_R$ otherwise the average power is infinite.

Ranked Known Interference (RKI) and Tomlinson Harashima Pre-coding (THP) The recently proposed non-linear pre-coding scheme of Ranked Known Interference (RKI) [CS01] is reminiscent of Tomlinson-Harashima Pre-coding (THP) for decentralized receivers [FWLH02]. Both are more sophisticated schemes which base on iterative cancellation of non-causally known interference at the Tx *à* la Costa [Cos83]. The required transmit power is reduced because less spatial nulls have to be placed than with LCI.

RKI and THP have the following two important differences: 1) as the block length goes to infinity, the shaping loss can be made arbitrarily small by choosing a sequence of optimal lattice quantizers, and 2) the modulo-loss can be made arbitrarily small by optimizing the lattice inflation factor. With standard THP always a pre-coding and shaping-loss is apparent while it can be eliminated for the standard RKI as shown in [CS01]. For this paper we assume perfect conditions (no pre-coding & shaping-loss) therefore RKI and THP will result in the same $\mathbf{PEF}_{THP/RKI}$. A modified version of RKI which was shown to be optimal in [CS01] for the case $n_T \geq m_R$ is equivalent to the JCBF scheme, proposed by [SB02a].

RKI and THP both base on a QL-type decomposition of the transmission channel

$$\mathbf{H} = \mathbf{F} \cdot \mathbf{S} \tag{4.109}$$

where $\mathbf{F}$ is a unitary beam-forming matrix and $\mathbf{S}$ is a lower triangular matrix. In Fig. 4.13 $\mathbf{S}$ is represented by $\mathbf{G}$ and the iterative loop with $\mathbf{B} - \mathbf{I}$. The modulo operation satisfies that the expected interference can be subtracted without an increase of the transmission power. Let $\mathbf{G} = (\mathrm{diag}\ \mathbf{S})^{-1}$, then the power enhancement factor is given by

$$\mathbf{PEF}_{THP/RKI} = \mathrm{trace}(\mathbf{G}^2). \tag{4.110}$$

The transmission is also like parallel AWGN channels with the same signal power at the MTs. In general, the following holds

$$\mathrm{trace}(\mathbf{G}^2) \leq \mathrm{trace}(\mathbf{D}^{-2}). \tag{4.111}$$

Since we have no analytical expression for the expectation of $\mathrm{trace}(\mathbf{G}^2)$ we have to simulate $\mathbf{PEF}_{THP/RKI}$.

For convenience in comparing the linear and non-linear pre-coding we define the ratio

$$PG_{LCI}^{RKI/THP} = \frac{\mathbf{PEF}_{LCI}}{\mathbf{PEF}_{RKI/THP}} \tag{4.112}$$

which reads as power reduction gain of RKI/THP over LCI which is the duality equivalent of the sensitivity gain at the Rx (how much less power is needed for perfect decision feedback equalization (DFE) (without error propagation) compared to zero-forcing at the Rx.

It has to be noted that trace($\mathbf{G}^2$) is a function of the actual channel $\mathbf{H}$ and the pre-coding order at the transmitter.

Joint Costa Beam-forming (JCBF) A more general approach which is not confined to $n_T \geq m_R$ is the JCBF algorithm [SB02a]. Here, the solution is found by solving a given SINR optimization task. The signal received at the i-th mobile terminal is given by

$$y_i = \mathbf{h}_i \mathbf{u}_i \sqrt{p_i} s_i + \sum_{k=1, k \neq i}^{m_R} \mathbf{h}_i \mathbf{u}_k \sqrt{p_k} s_k + n_i, \forall i \tag{4.113}$$

where $\mathbf{u}_i$ is the i-th column of the unitary beam-forming matrix $\mathbf{BF}$ and p_i the belonging transmit power for signal s_i when $\mathbb{E}[s_i s_i^H] = 1$. $\mathbf{h}_i$ is the i-th row of the down-link transmission channel $\mathbf{H}$ from the n_T antennas at the BS to the m_R mobiles. The first part in (4.113) is the actually desired signal for user i while the rest consists of the co-channel interference and the user specific receiver noise n_i. We use the notation $\mathbf{U} = [\mathbf{u}_1, ..., \mathbf{u}_{m_R}]$ and $\mathbf{p} = [p_1, ..., p_{m_R}]^T$. Finally, $\sigma_i^2 = \mathbb{E}[n_i^2]$ denotes the user specific noise power. Defining $\mathbf{R}_i = \mathbf{h}_i^H \mathbf{h}_i$, the instantaneous SINR seen by user i becomes

$$SINR_i^{DL}(U, p) = \frac{p_i \mathbf{u}_i^H \mathbf{R}_i \mathbf{u}_i}{\sum\limits_{\substack{k=1 \\ k \neq i}}^{m_R} p_k \mathbf{u}_k^H \mathbf{R}_i \mathbf{u}_k + \sigma_i^2}, \forall i. \tag{4.114}$$

Assuming an encoder order $1...m_R$, the interference transmitted from the i-th user to the user $k > i$ is known, thus it can be pre-subtracted at the transmitter prior to submission. This can be done by quantization-based pre-coding schemes, which have been proposed, e.g. in [WJ01, ESZ00]. Without increasing transmit power the pre-coded SINR of the i-th user becomes

$$SINR_i^{DL,coded}(\mathbf{U}, \mathbf{p}) = \frac{p_i \mathbf{u}_i^H \mathbf{R}_i \mathbf{u}_i}{\sum\limits_{k=i+1}^{m_R} p_k \mathbf{u}_k^H \mathbf{R}_i \mathbf{u}_k + \sigma_i^2}, \forall i. \tag{4.115}$$

Now, the task is to find $\mathbf{U}$ and $\mathbf{p}$ which satisfy $\gamma_i \leq SINR_i, \forall i$ with γ_i denoting the target SINR of user i.

The algorithm exploits the duality between the multi-user broadcast channel (BC) and the multi-user medium access channel (MAC)[BS02a]. The more complex down-link problem is solved by solving the belonging "virtual" uplink problem instead. The algorithm is initialized with the target SINRs γ_i for each MT and therefrom it calculates the optimal power allocation $\mathbf{p}$ and the beam-forming matrix $\mathbf{U}$ to meet the above requirements. To be conform with the general notation used in Fig. 4.16 the general power allocation matrix $\mathbf{G}$ is given by $diag(\mathbf{G}) = \mathbf{p}$ and the unitary beam-forming matrix $\mathbf{BF} = \mathbf{U}$.

The JCBF algorithm finds the best MMSE solution therefore it is not restricted to $n_T \geq m_R$. In [SB02a] it was proven that the JCBF solution is the solution with the minimum power and therefore it is optimal regarding power efficiency.

The RKI solution which was shown to be optimal for $n_T \geq m_R$ is asymptotically identical with the solution of the optimal JCBF algorithm for SINR$\rightarrow \infty$ except for the signs of the real and imaginary values of the beam-forming vectors in BF.

If we assume SINR requirement which are relaxed (0 dB or less) which may be realistic for CDMA systems, then we observe differing solutions. A full decorrelation is not aimed with JCBF therefore it may achieve a considerable power reduction compared to the fully decorrelating algorithms of RKI and THP especially for bad conditioned channels

As we have seen already for RKI/THP the power enhancement factor $\mathbf{PEF}_{JCBF}$ also depends on the actual channel $\mathbf{H}$ and the pre-coding order.

System Modell for the Numerical Simulations A good case scenario is assumed to be e.g. a Rayleigh channel which could represent an indoor environment without a line of sight (LOS). The Rayleigh channel is modelled by $\mathbf{H}$ with identically and independently distributed random complex entries with zero mean and unit variance.

A bad case scenario might be found in a rural environment with little or no multi-path propagation and only a few local scatterers close to the mobiles. The rural area scenario used in the following is based on a geometrical model with a $\lambda/2$-array at the BS (120° aperture) and the MTs are distributed as follows: 1/3 of the users at a distance of 500 m $\pm 10\%$, 1/3 at a distance of 5000 m $\pm 10\%$ and the rest is distributed randomly in between. All MTs are surrounded by scatterers while the BS is without scatterers nearby (e.g. a scenario with the BS positioned on an aerial mast on top of a hill).

For comparison of the three down-link transmission schemes we distinguish between three scenarios. Scenario I: no limits are put onto the angular distribution of the users within the sector seen by the BS. This might produce a close to singular channel matrix when two users are situated inline seen from the BS. Scenario II: we limit the worst case by allowing no users to be active which have an angular separation less than 2 deg seen from the BS. In scenario III we restrict $\Delta\phi > 5$ deg. Since LCI and RKI/THP are fully decorrelating schemes bad conditioned channels would result in a high increase in transmit power. Therefore, in field applications users which cause the channel to be bad conditioned could be separated by time, frequency or code instead of spatial multiplexing.

An exemplary outdoor scenario II is depicted in Fig. 4.17 with 15 MTs and 16 Tx antennas at the BS. The minimum angular separation as seen from the BS is more than 2 deg.

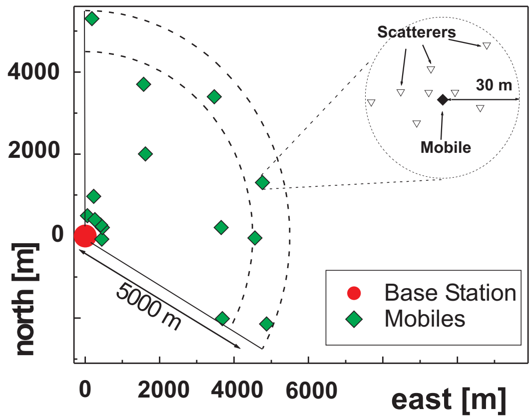

Figure 4.17: Exemplary setup of 15 mobiles in an outdoor scenario with no multi-pathes from the environment. The base station has a linear array of 16 Tx antennas and the sector is 120 deg. See magnification of users vicinity with 8 local scatterers randomly positioned around each mobile within 30 m radius.

Simulation Results

To compare the linear pre-coding of LCI with the more sophisticated non-linear pre-coding schemes of RKI/THP and JCBF we do Monte Carlo simulations with random channels according to the Rayleigh and rural area scenario. We normalize $\mathbf{H}$ that $\mathrm{trace}(\mathbf{HH}^H) = n_T \cdot m_R$. The statistics over the needed sum Tx power is then taken and the average PEF is compared. Without loss of generality we set all SINR requirements to be identical, $n_T \geq m_R$ and the noise at all MTs is set to one.

Stream Ordering Remember the QL-type decomposition of $\mathbf{H} = \mathbf{F} \cdot \mathbf{S}$ with $\mathbf{S}$ being a lower triangular matrix. When the channel matrix is permuted in its columns then the resulting power enhancement $\mathbf{PEF} = \mathrm{trace}(\mathbf{G}^2)$ with $\mathbf{G} = (\mathrm{diag}(\mathbf{S}))^{-1}$ is depending on the actual permutation since the values on $\mathrm{diag}(\mathbf{G})$ are changing. To have a fair comparison of the pre-coding schemes we apply the same stream ordering.

From [VWFin] we know that the VBLAST algorithm finds the permutation P_{VBLAST} where the lowest entry in $\mathrm{diag}(\mathbf{S})$ will be maximized.

$$P_{VBLAST} = \mathrm{argmax}\ \min\{|s_{11}|^2, ..., |s_{m_R,m_R}|^2\} \tag{4.116}$$

Simulations gave rise to assume that the VBLAST sorting is a close to optimum ordering [WVF02] even if it not finds the optimal permutation

$$P_{opt} = \mathrm{argmax}\ \sum\{|s_{11}|^2, ..., |s_{m_R,m_R}|^2\} \tag{4.117}$$

which achieves the lowest sum transmit power. Independently, this result towards the suboptimum ordering of the VBLAST algorithm was also found by [Joh04]. Therefore the VBLAST ordering will be used for all simulations.

Fig. 4.18 and Fig. 4.19 show the **PEF**s for the pre-coding techniques LCI/JT, RKI/THP and JCBF. For the good case (Rayleigh channel, Fig. 4.18) the non-linear pre-coding schemes need

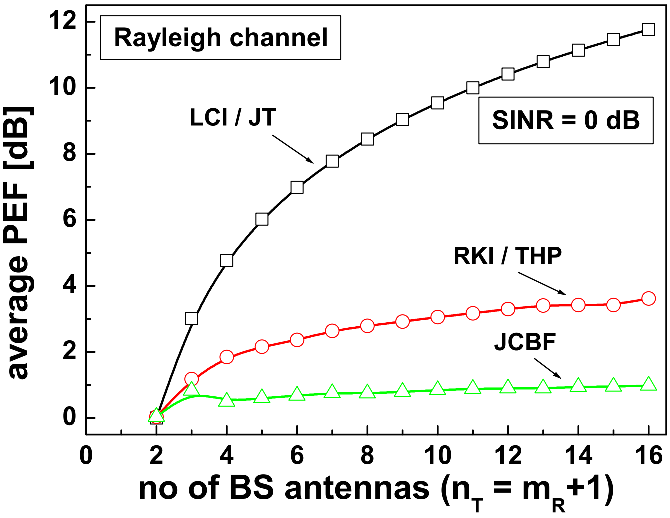

Figure 4.18: Power enhancement factor **PEF** for LCI, RKI/THP and JCBF compared to parallel AWGN channels for the good case of a Rayleigh channel. $n_T = m_R + 1$, SINR = 0 dB at each mobile

significant less transmit power with rising n_T. If $n_T - m_R > 1$ and rising then the advantage of non-linear pre-coding decreases because $\mathbf{PEF}_{LCI}$ decreases itself. Therefore, we can outline that for the Rayleigh channel (e.g. indoor scenario without LOS) the gain of non-linear pre-coding is maximum when $n_T - m_R$ is minimum. The dual result is known when changing from a linear to a non-linear detector with ZF, then the maximum SNR gain is achieved when $n_T = m_R$.

In a rural area scenario (Fig. 4.19) standard RKI and THP need much less Tx power than LCI/JT due to the Costa pre-coding. This power advantage can rise up to several orders of magnitude depending on the channel conditions (see scenario I and II), while LCI/JT seems not feasible here. As said before, the probability of channel singularity decreases with a sufficient surplus of transmit antennas in the Rayleigh channel, this does not hold here, especially in the outdoor scenario I and II where the channel condition is mainly determined by angular distribution of the users instead of the no of Tx antennas. Therefore the use of more transmit antennas is not a profitable option here.

All three scenarios clearly show that the JCBF algorithm ($\triangle$) performs best. It can benefit from its SINR optimization with the MMSE criterion and requires therefore the smallest Tx power. If the channel is well conditioned, e.g. the users are well distinguishable regarding their angular distribution (scenario III, $\Delta\phi_{ij}$ >5 deg) then the difference between RKI/THP and JCBF is quite small.

We expected from theory that for high SINR requirements the performance and the individual solutions of RKI/THP and the JCBF algorithm must converge. From our simulations we can confirm that the calculated power allocation matrices are identical and the beam-formers for each user in **BF** are also identical except the signs of the real and the imaginary part. The reason is seen in the JCBF algorithm itself, which uses only the power of the signals, interferences and noise , therefore the signs of the beam-formers are not relevant.

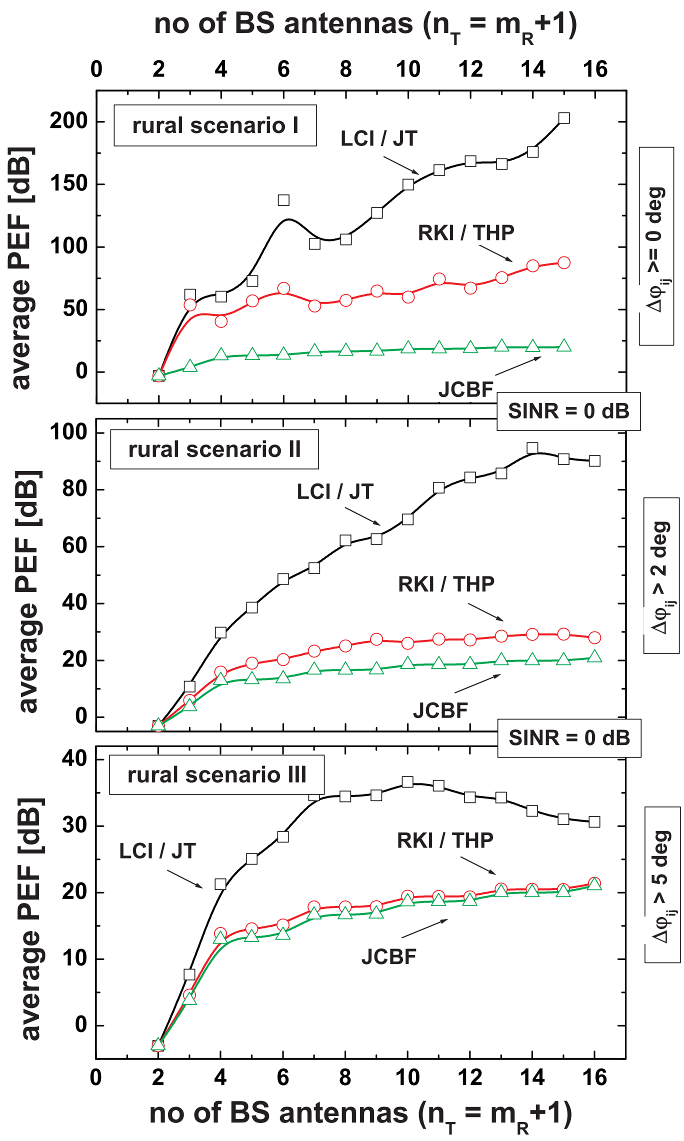

Figure 4.19: Power enhancement factor **PEF** for LCI, RKI/THP and JCBF compared to parallel AWGN channels for the bad case of rural scenarios. Scenario I: no limits on the angular distribution of the users, $\Delta\phi_{if} \geq 0$ deg; Scenario II: $\Delta\phi_{if} > 2$ deg; Scenario III: $\Delta\phi_{if} > 5$ deg - note different dB-scale at the ordinates. $n_T = m_R + 1$, SINR $= 0$ dB at each mobile

Fig. 4.20 shows the transmit power gain of JCBF against RKI/THP for the example of $n_T = 10; m_R = 9$ plotted over the required SINR at the MTs. In the left part of Fig. 4.20 we see that the gain vanishes for SINR $\to \infty$ because the solutions of RKI/THP and JCBF converge asymptotically. On the right part of Fig. 4.20 we see the convergence already at approx. 30 dB. This is due to the the fact that channel singularity is very unlikely for these scenarios and therefore the MMSE solution of JCBF can't gain much in the high SINR region.

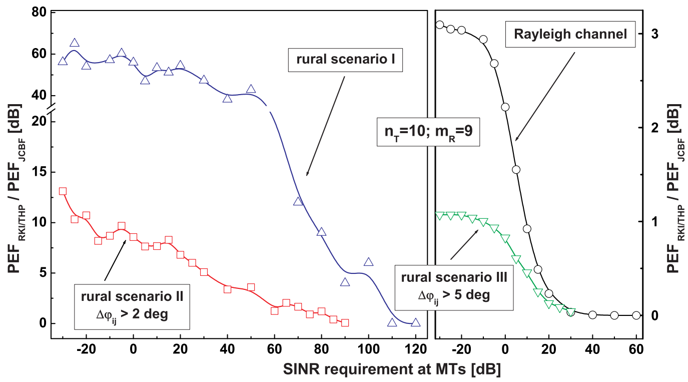

Figure 4.20: Power reduction gain of optimal JCBF pre-coding against RKI/THP pre-coding for various SINR requirements at the MTs (left: rural scenarios I and II, right: rural scenario III and Rayleigh channel). n_T=10, m_R=9

4.3 Performance Degradation

4.3.1 Imperfect Channel Estimation

One very important issue for the achievable performance of a real MIMO transmission system is the available accuracy of the channel knowledge which is needed at the Tx, Rx or at both sides of the link. Many transmission schemes require accurate channel information e.g. about the degrees of freedom of the channel or the achievable SINR at the receiver. If those information is affected by estimation errors, then the performance of the transmission scheme can be severely degraded. Therefore, the reliability of the achieved channel estimates is of great importance for real world applications and has therefore to be acquired in an appropriate way.

In the following we focus on the single user MIMO case and investigate how received signals distorted by receiver noise affect the estimation of the channel rank, the singular values and resulting from this the predicted channel capacity. Furthermore we show how the BER performance with fixed modulation is degraded by imperfect CSI.

We will start with the affect on the Rx detection with the example of ZF. In the second part we give a more detailed discussion on the BER performance degradation with erroneous linear transmit pre-processing due to imperfect channel knowledge at the transmitter.

We assume a channel measurement based on the correlation of orthogonal pilot sequences to identify each transmit antenna. Suitable sequences might be Hadamard- or Goldsequences as used for the real-time transmission experiments in chapter 5. For the following let us assume a correlation based measurement with sequences of length L in a separate time slot before data transmission.

Distribution of the Singular Values:

Using the orthogonal pilot sequences the estimate of the channel matrix is given by [HJJ+01a]

$$\hat{\mathbf{H}} = \mathbf{H} + \sigma_{err} \cdot \mathbf{N} \quad \text{with} \quad \sigma_{err} = \sqrt{\frac{n_T}{L \cdot SNR}}. \tag{4.118}$$

The channel matrix $\mathbf{H}$ and the noise like matrix $\mathbf{N}$ are of $m_R \times n_T$ shape each and with i.i.d. complex Gaussian entries, L: sequence length in symbols, SNR: average SNR at one receive antenna.

Let $\mathbf{A} = \mathbf{H} \cdot \mathbf{H}^H$ be the covariance matrix of the channel matrix $\mathbf{H}$ and let channel estimation errors cause an estimate of $\mathbf{H}$ of

$$\hat{\mathbf{H}} = \mathbf{H} + \Delta\mathbf{H} \tag{4.119}$$

then we estimate the perturbation of the covariance matrix $\mathbf{A}'$ in first order as

$$\begin{aligned} \mathbf{A}' &= (\mathbf{H} + \Delta\mathbf{H}) \cdot (\mathbf{H} + \Delta\mathbf{H})^H \\ \mathbf{A}' &= \mathbf{H} \cdot \mathbf{H}^H + \Delta\mathbf{H} \cdot \Delta\mathbf{H}^H \end{aligned} \tag{4.120}$$

The shift of the SVs of $\mathbf{H}$ is then determined by $\Delta\ \mathbf{A} = \Delta\mathbf{H} \cdot \Delta\mathbf{H}^H$ depending on the origin and structure of $\Delta\mathbf{H}$. If $\Delta\mathbf{H}$ is a statistically uncorrelated additive error for each h_{ij} it can be written as $\Delta\mathbf{H} = \sigma_{err} \cdot \mathbf{N}$ as we introduced above in (4.118).

Then the expectation of the perturbed SVs $\hat{\lambda}_i$ is given by

$$\mathbb{E}\left[\hat{\lambda}_i^2\right] = \mathbb{E}\left[\lambda_i^2\right] \cdot (1 + \sigma_{err}^2) \tag{4.121}$$

Equation (4.121) shows that the average received signal power which is

$$\text{trace}(\mathbf{H}^2) = \sum_{i=1}^{min(n_T, m_R)} \hat{\lambda}_i^2 \tag{4.122}$$

is overestimated if the system suffers from noteworthy channel estimation errors represented by σ_{err}. This means a required average SNR at the Rx is not given or adaptive transmission schemes may suffer degradation because more parallel channels are used than there are actually available.

Another very important issue is that the matrices $\mathbf{U}$ and $\mathbf{V}$ obtained from SVD in (3.9) also suffer from the channel estimation errors in a way that the projection of the input signal vector onto the parallel sub-channels at the Tx and the decorrelation from those sub-channels at the Rx does not match correctly, therefore co-channel interference occurs, which decreases the actual SINR and the BER performance. Therefore, if a SVD-MIMO based transmission is realized, appropriate channel estimation has to be acquired. Furthermore, in case the pre-coding and decoding matrices are calculated at the Tx and the Rx separately based on independent CSI then ordered SVD can cause a change in eigenvectors due to the independent disturbance of the SVs, which was discussed by [LGF02, TLF03]. A successful concept of mitigating those effects is proposed in section 6.2.2 together with the experimental results of measurements with a SVD-MIMO scheme.

A different approach was proposed by [QUA04]. Here, the transmit pre-coding is performed for the data and the pilot sequences with a unitary matrix $\mathbf{V}$ which might be out-dated or erroneous. The Rx is equalizing the actual composed channel realization $\mathbf{H} \cdot \mathbf{V}$ with an MMSE detector at the Rx to combat the inter-channel-interference due to the imperfect pre-coding. So, the CSI at the Rx has to be reasonably well. This quite straight forward concept is discussed for the 802.11n standard for WLANs.

Capacity and Channel Estimation Errors

If we focus on the effect of the channel estimation errors on the average capacity we can conclude directly from the distribution of the disturbed SVs given in (4.121).

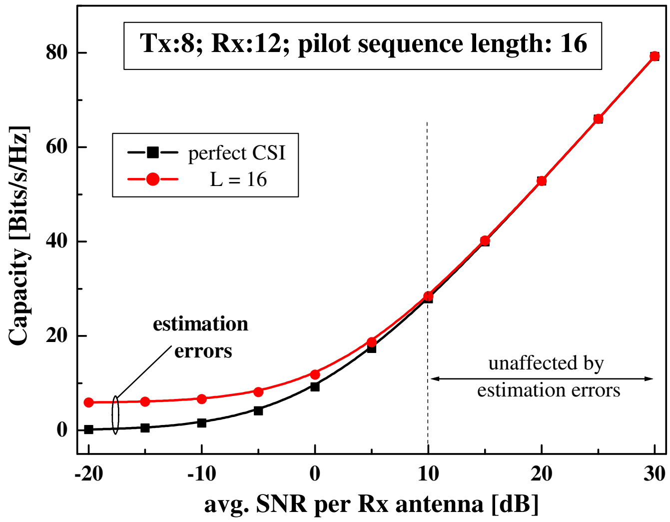

Figure 4.21: Average capacity in case of channel estimation errors caused by a correlation based channel measurement in an i.i.d. Rayleigh channel.

Fig. 4.21 shows that at low SNR where σ_{err} is high, the average capacity will be estimated much higher than it is actually available by the real channel. This can be explained by the fact that the channel estimation errors add artificially produced uncorrelated matrix entries which improve apparently the average capacity. This effect can be reduced significantly by the use of longer pilot sequences for the channel estimation.

For a given SNR of more than 10 dB the effect of channel estimation errors caused by the correlation based channel measurement becomes negligible for the contemplated MIMO system with 8 Tx and 12 Rx antennas.

Bit Error Rates and Channel Estimation Errors

For BER simulations we assumed perfect CSI and ZF at the Rx like in (3.51). The degradation of the BER performance in Fig. 4.22 is then mainly caused by the imperfect data de-correlation at the Rx. This co-channel interference equals a lower SNR so the BER curves are expected to be shifted. In [HJJ+01a] it was shown that under the assumption of uncorrelated data streams this co-channel interference equals additional noise at the Rx and the data estimate after imperfect de-correlation can be written as

$$\mathbf{x}' = \mathbf{x} + \mathbf{H}^{\dagger}\left(\mathbf{n} + \sqrt{\frac{n_T}{L}} \cdot \tilde{\mathbf{n}}\right). \tag{4.123}$$

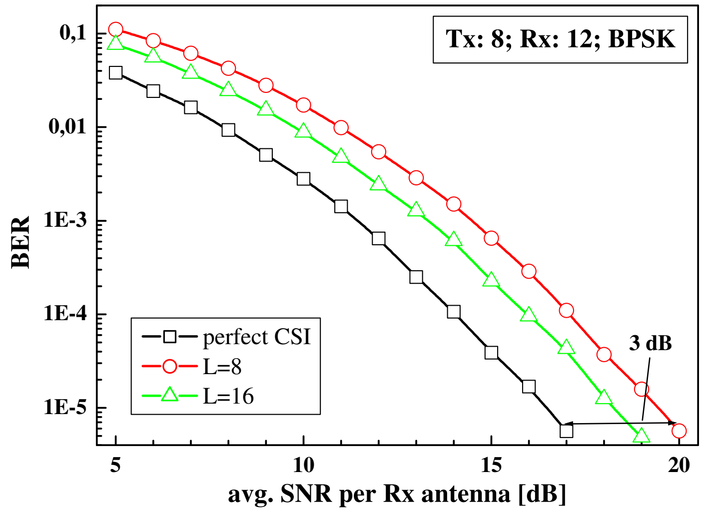

Figure 4.22: BER performance for a 8x12 MIMO system with channel estimation errors caused by a correlation based channel measurement with a varying length of the pilot sequence (BPSK modulated).

The impact of the channel estimation error is similar to additional noise $\tilde{\mathbf{n}}$ at the receiver which has the same statistical properties like $\mathbf{n}$ but the genuine Rx noise and the noise-like co-channel interference $\tilde{\mathbf{n}}$ are not correlated. The worst case of a correlation based channel knowledge is when the pilot sequence $L = n_T$. This produces a shift of the BER curve of 3 dB as to be seen in Fig. 4.22.

The effect of channel estimation errors on MIMO systems exploiting CSI at the Tx is investigated in the following paragraph where channel inversion [HPJ+02] at the Tx is assumed. As expected the degradation effect is much higher with imperfect pre-procession at the Tx (see Fig. 4.23) than we find in Fig. 4.22 with ZF at the Rx.

In [Joh04] it was shown that with perfect CSI ZF/MMSE pre-coding at the Tx and ZF/MMSE detection at the Rx perform more or less the same with a slight BER performance advantage for the Tx pre-processing. With imperfect CSI which has to be assumed for any real transmission system this performance equality is easily lost when insufficient effort is put into a good channel measurement, e.g. longer pilot sequences may help.

Bit Error Rates and Channel Estimation Errors for Linear Pre-coding

If $\mathbf{H}$ is the channel matrix for the up-link ($n_T < m_R$) then we expect a correlation measurement based channel estimation error like in (4.118) with

$$\Delta\mathbf{H} = \sigma_{err} \cdot \mathbf{N} \text{ with } \sigma_{err} = \sqrt{\frac{n_T}{L \cdot SNR}} \tag{4.124}$$

Based on this channel estimation we perform CI in the down-link (the transmission channel is now the transposed channel $\mathbf{H}^T$). The estimate of the transmitted data symbols can then be written as:

$$\mathbf{x}' = \mathbf{H}^T\tilde{\mathbf{x}} + \mathbf{n} \quad \text{with:} \quad \tilde{\mathbf{x}} = (\mathbf{H}^T + \Delta\mathbf{H}^T)^\dagger\mathbf{x} \tag{4.125}$$

Using linear Taylor expansion for the estimated transmit matrix we obtain the following expression for the Down-link CI

$$\mathbf{x}' = \mathbf{x} + \mathbf{n} + \sigma_{err}\mathbf{N}^T(\mathbf{H}^T)^\dagger\mathbf{x}. \tag{4.126}$$

It can be shown that the expectation values of the two following expressions are equal (for a sufficient number of antennas).

$$\mathbb{E}\left[\mathbf{N}^T(\mathbf{H}^T)^\dagger\mathbf{x}\right] = \mathbb{E}\left[\mathbf{H}^\dagger\mathbf{N}\mathbf{x}\right] \tag{4.127}$$

Using (4.127) we can give an analytical expression for the data estimate at the Rx

$$\mathbf{x}' = \mathbf{x} + \mathbf{n} + \sigma_{err}\mathbf{H}^\dagger\mathbf{N}\mathbf{x}. \tag{4.128}$$

The last term of (4.128) is cross-talk which is similar to additional noise that is enhanced by $\mathbf{H}^\dagger$

$$\mathbf{x}' = \mathbf{x} + \mathbf{n} + \mathbf{H}^\dagger\tilde{\mathbf{n}} \tag{4.129}$$

with $\tilde{\mathbf{n}}$: being something which behaves like additional noise but which is independent on noise $\mathbf{n}$ caused by the Rx.

This means that for a pilot sequence of the length $L = n_T$ (worst case) and for high SNR (when the last term is the dominating error term) CI performs with an identical slope like ZF. With a sufficient long training sequence ($L > 32$ for the 10x8 example) CI performs better than ZF but for high SNR there is always a bend in the curve for the BER because of the term $\mathbf{H}^\dagger\tilde{\mathbf{n}}$ in (4.129) which still depends on the antenna diversity gain. At high SNR the second term $\mathbf{n}$ can be neglected and the slope of the curve is dominated by the $m_R - n_T + 1$-branch diversity of

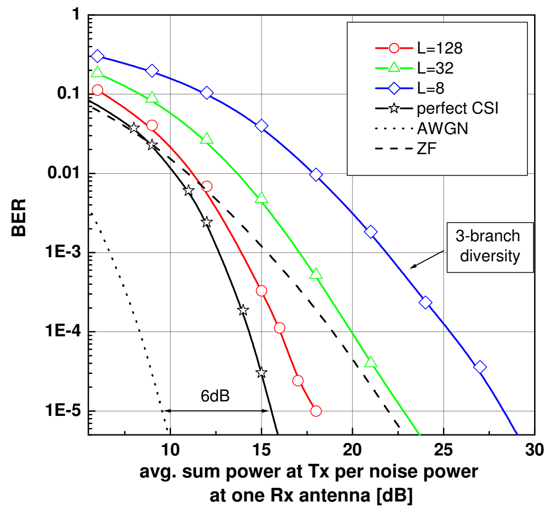

Figure 4.23: Down-link-CI, MIMO 10x8, BPSK, length of sequence L variable.

the MIMO antenna configuration (see Fig. 4.23). The curve for perfect CSI is shifted against the AWGN curve because of the fact that more average Tx power is needed given by the factor $\frac{m_R}{n_T - m_R} = 4$ (see (4.108)) [JWHJ01].

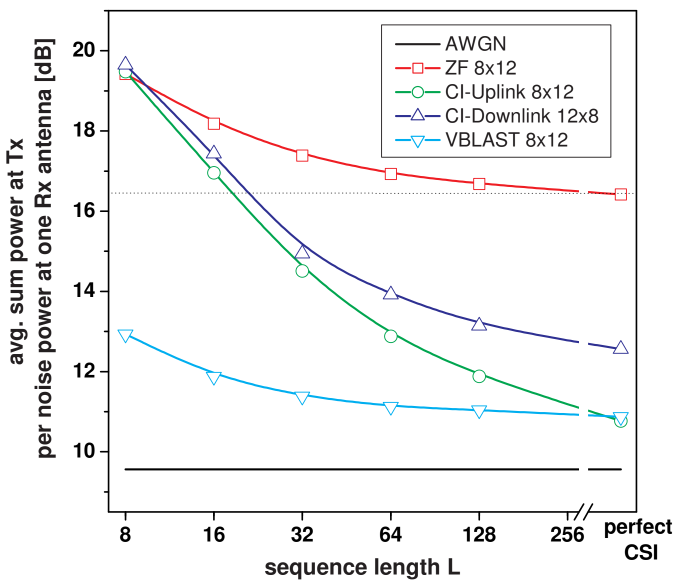

Figure 4.24: Sensitivity against channel estimation errors for a BER performance of 10^{-5} (BPSK) using ZF, VBLAST and CI, variable sequence length L.

The sensitivity against channel estimation errors for various transmission schemes (ZF, VBLAST, Down-link-CI, Up-link-CI) is compared in Fig. 4.24. To be consistent we plotted the average transmitted sum power per noise at one Rx antenna needed to obtain a BER of 10^{-5} over the length of the training sequence. It is found that for a 12x8 antenna configuration we need a training sequence with $L > 16$ to achieve the same performance like ZF with perfect CSI (see dotted line). Up-link-CI and down-link-CI show the same BER performance in front of the decision unit (when the receive antenna gain is already exploited). Since CI is based on CSI at the Tx in general a better channel estimation is required than with ZF or VBLAST. Given this and a reasonable amount of antenna diversity, CI shows a good BER performance with all advantages of its simple linear algebraic structure.

4.3.2 Limited Transmitter Dynamics

The limitation of transmit power either due to government regulations or in most cases due to a limited dynamical range of the transmit amplifiers is a very important issue for all transmission schemes using pre-coding. If the transmit signal exceeds the linear dynamic range of the amplifier, then non-linear distortions of phase and amplitude of the signal are observed. In the following we discuss the problem of a limited dynamic range under the simplified assumption of a amplitude clipping to a maximum value without any phase distortion. In reality the situation is more complex and there are already proposals on how pre-coding can be made feasible in the non-linear range of amplifiers and I/Q-modulators [GOdMV04]. Nevertheless, the principle of performance degradation due to a limited transmitter dynamics can be seen exemplarily with linear channel inversion or ZF-Tx pre-coding.

Channel Inversion (CI) requires antenna diversity, otherwise the transmitted power is not limited. This is easily understood when the expectation value of the term $\mathbf{x} = \mathbf{H}^{\dagger}\mathbf{d}$ is discussed. If the expectation value of the symbols of each data stream $\mathbb{E}[d_i^2] = 1$ and $n_T > m_R$ (down-link-CI) then the expectation value of the transmitted power is

$$\mathbb{E}\left[\mathbf{x}^H\mathbf{x}\right] = \mathbb{E}\left[\mathbf{d}^H\mathbf{H}^{\dagger H}\mathbf{H}^{\dagger}\mathbf{d}\right] \quad = \quad \sum_{i=1}^{m}\frac{1}{\lambda_i^2}\|d_i\|^2 \tag{4.130}$$

$$\mathbb{E}\left[\sum_{i=1}^{m}\frac{1}{\lambda_i^2}\right] = \frac{1}{n_T - m_R} \qquad \Rightarrow \qquad \mathbb{E}\left[\mathbf{x}^H\mathbf{x}\right] = \frac{m_R}{n_T - m_R} \tag{4.131}$$

with λ_i^2 being the Eigenvalues of $\mathbf{H}$. This result can be easily checked by simulations for various antenna number configurations. It is a very important relation which was also studied by [KSS99] and [TH99] and its general validity could be shown by [JWHJ01] for the case of m_R and n_T going to infinity with $\frac{m_R}{n_T} = const$ using results from [SB95]. If $m_R = n_T$ then the average transmit power in (4.131) goes to infinity. In [JHJvH02] it was shown that (4.131) is valid for any positive integer number m_R, n_T with $n_T \neq m_R$. Practically this means that CI works only with a reasonable amount of transmit antenna diversity. The additional antennas reduce the necessary dynamic range (DR) of the transmitted power, which we define as the peak to average Tx power per transmit antenna. A small DR is favorable with regard to amplifier requirements and to peak power. If the input power at the amplifier exceeds a certain level, amplitude clipping and phase distortions are observed.

Fig. 4.25 shows the simulated pdf of the transmitted power from one Tx antenna if CI is used. The black curve represents the emitted power distribution for a MIMO system with 12 Tx and

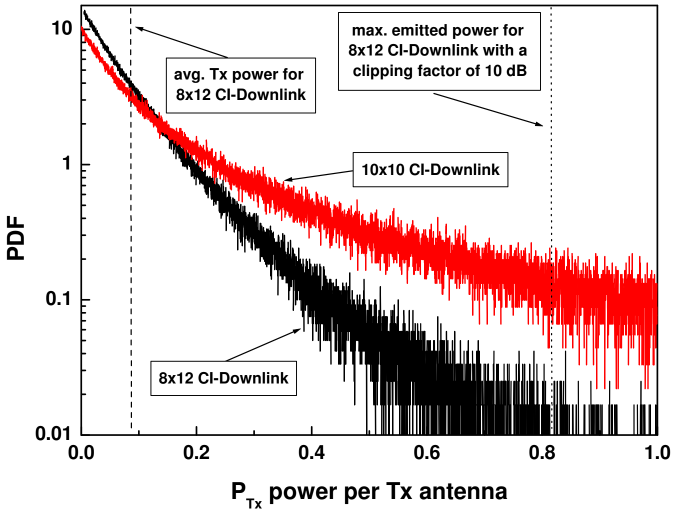

Figure 4.25: PDF of the transmitted power per Tx antenna for MIMO systems using Down-link-CI (12 / 10 Tx and 8/10 Rx antennas)

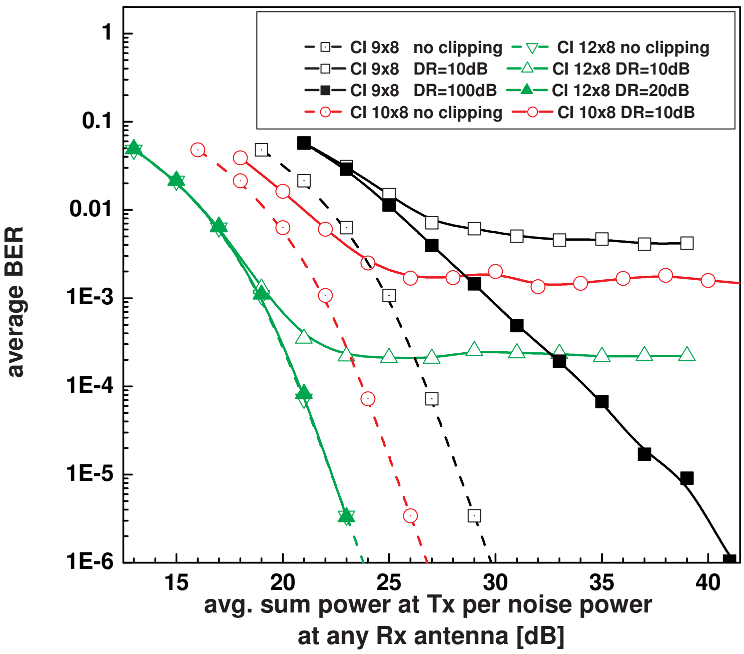

Figure 4.26: Power clipping for MIMO using Up-link-CI (16-QAM, L=128), various sets of dynamic range (DR) and antenna diversity

8 Rx antennas, while the red curve is a system with 10 antennas on each side. It is obvious that the integral above the power clipping line which is also depicted in Fig. 4.25 is significantly smaller when more Tx antennas are used.

Fig. 4.26 illustrates how clipping at the Tx causes error floors depending on antenna diversity and on the DR of the amplifier. The floor can be estimated from the DR and from the power distribution per antenna like in Fig. 4.25. In general, a dynamic range of $P_{max}/\mathbb{E}[P] = 20$ dB seems to be sufficient when 12 Tx and 8 Rx antennas are used.

4.3.3 Co-Channel Interference

BER Degradation

In reality, interference with devices operating in the same frequency band like the MIMO system have to be considered. The received interference power caused by simultaneously transmitted symbols $\mathbf{z}$ is assumed to follow Rayleigh or Rician statistics of the interference channel $\mathbf{J}$

$$\mathbf{i} \quad = \quad \mathbf{J} \cdot \mathbf{z}. \tag{4.132}$$

Since the interference channel is not known to the Rx a priori, this means that we have to expect interference power levels that may vary significantly at different Rx antennas. Furthermore we assume the interfering signals to be perfectly synchronized in frame and symbol clock with the data transmission clocks.

The reconstructed symbols can be described as follows

$$\text{Up-link ZF} \quad (n_T < m_R) \quad \mathbf{d}' = \mathbf{d} + \mathbf{H}^{\dagger}(\mathbf{n} + \mathbf{i}) \tag{4.133}$$

$$\text{Up-link CI} \quad (n_T < m_R) \quad \mathbf{d}' = \mathbf{d} + \mathbf{U}^{H}(\mathbf{n} + \mathbf{i}) \tag{4.134}$$

$$\text{Down-link CI} \quad (n_T > m_R) \quad \mathbf{d}' = \mathbf{d} + (\mathbf{n} + \mathbf{i}). \tag{4.135}$$

For high SNR ($\mathbf{n}$ is negligible compared to $\mathbf{i}$) the performance will be dominated by the signal-to-interference-ratio (SIR) and an error floor is expected for the BER.

Individual Power Control for Channel Inversion

Since the interference power per Rx antenna may vary significantly and the channel between the interference sources and the Rx antennas is generally unknown a priori, little can be done by the Rx. But a simple strategy can help out, at least partially. We propose to measure the relative noise-plus-interference power per data stream in front of the decision unit at the Rx unit and to feed back this information to the Tx. With this additional information, the Tx adapts the power in each data stream at the Tx to equalize the SINR of all data streams.

This Individual Power Control (IPC) requires little data load in the feed-back channel but it allows reduction of co-channel interference. The best results are reached with up-link-CI (see Fig. 4.27). Down-link-CI suffers more degradation because no receive diversity is available at the Rx.

Individual Power Control for ZF and VBLAST

Analyzing the effect of IPC for CI, we find that it can be used far more generally.

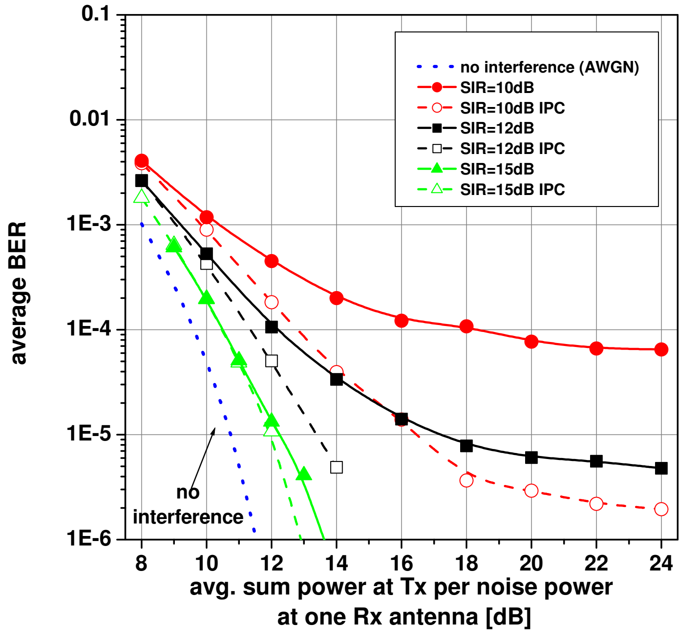

Figure 4.27: IPC to combat co-channel interference for a 8x12 MIMO system using up-link CI, BPSK, two BPSK interference sources. Achievable low BER for SIR > 12dB.

Systems which use e.g. ZF or VBLAST at the Rx can also benefit from this technique. Like for CI the SINR of all data streams can be equalized. This means for the special case of no interference

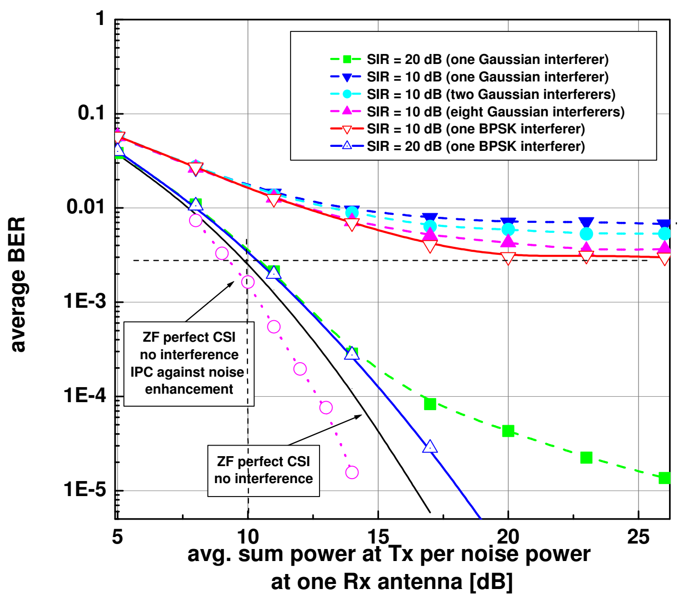

Figure 4.28: Bit Error floors for 8x12 MIMO systems with co-channel interference (data BPSK modulated, ZF at Rx, interference: BPSK or Gaussian)

sources IPC can balance the effect of noise enhancement caused by $\mathbf{H}^{\dagger}\mathbf{n}$. ZF will benefit directly from an equalization of the noise enhancement (see dotted line in Fig. 4.28) by redistributing the transmitted power per data stream. A system using VBLAST can be tuned to specific needs. For example, if we have a very good SNR, the BER can be reduced by transmitting more power into the data stream that is first to be detected (lowest noise enhancement). This first detected data stream dominates the BER at high SNR while all latter detected data streams are nearly free of errors. Note, that this power allocation must be consistent with the dynamic range of the Txs, as discussed in sec. 4.3.2.

So the technique of IPC is applicable to all MIMO transmission schemes at little expense with common modulation and coding on all data streams.

The efficiency of IPC depends on the number of antennas, the sort of interference and the number of interference sources. Generally, IPC works the better the more parallel data streams are to be balanced. Another important issue is the distribution of the data streams over several Rx antennas. A high receive diversity improves the efficiency of IPC.

To study how the performance of IPC depends on the interference symbols we did simulations for the BER (Fig. 4.28). We used interference sources with a Gaussian symbol alphabet or BPSK / M-QAM as it was used for the MIMO data transmission. We assume the interfering symbols to be perfectly synchronized for the simulation. We find in Fig. 4.28 that one interference source using a Gaussian alphabet causes an error floor three times higher than one or many interference sources using BPSK or M-QAM modulation. A sufficient number of Gaussian interferers behave like an interferer using BPSK or M-QAM. This can be explained by the multiplication of only one Gaussian distributed symbol stream with the Gaussian distributed entries of the interference matrix resulting in a 2^{nd}-order Bessel distribution for the received interference power. BPSK or M-QAM symbols transmitted over the same interference channel have a power distribution at the Rx which is χ^2. The superposition of several (8-10) Gaussian interference signals performs similar like BPSK or M-QAM interference. This makes clear that only for Gaussian interference the number of sources is of importance. For interference caused by sources using BPSK or M-QAM alphabets the average SIR per Rx antenna alone determines the error floor. Then in the simple case of ZF the error floor can be predicted from the curve without interference. We just look at the BER of an system free of interference at a SNR of the same value as the SIR (see Fig. 4.28). Example: We assume a SIR of 10 dB, then the bit error floor with interference from synchronous M-QAM sources will be at the same level as the BER without any interference and an equivalent SNR of 10 dB which is approximately $3 \cdot 10^{-3}$ with ZF and 8 Tx and 12 Rx antennas.

4.4 Channel Adaptive Bit-loading

4.4.1 Single-user MIMO Bit-Loading

We consider a single user MIMO system with n_T transmit antennas and m_R receive antennas ($n_T \leq m_R$).

Full CSI at Tx: If perfect CSI is available at the Tx then transmission of one data stream per channel sub-space was shown to be optimum in sec. **??**. The optimum power allocation to minimize the sum MSE from (4.22) has a water-filling-like solution. The data streams are coupled into the spatially orthogonal sub-channels which results in a bit-loading strategy very similar to OFDM. If the modulation level and/or the power allocation of one data stream is changed it has no effect on to the SNR of the other data streams. In this way joint bit-loading and power allocation becomes quite simple.

The bit-loading strategy (BLS) uses finite modulation alphabets and considers a minimum transmission quality (e.g. maximum BER). The BLS finds then the best match of the available SNRs per data stream and the finite data symbols. In this way we achieve the highest data throughput under the given constraints.

1. The BLS starts with the optimum solution for the power allocation from (4.22) and computes the SNR for every sub-stream. Each of the sub-streams is then given the highest modulation satisfying the BER constraint.
2. The allocated Tx power per data stream is reduced to achieve just the necessary SNR for each modulation.
3. Test, if taking one bit/Hz/s from one data stream and giving it to another saves power. If not, proceed to next step.
4. Sum up the remaining power and give it to the stream which can support one more bit/Hz/s (a higher modulation scheme) while it needs least from the rest power.
5. Continue from 1. to 4. until no positiv rest power is available after an intended modulation step.
6. Distribute the remaining power in a way that all channels have the same SNR increase.

No CSI at Tx: If no CSI is available at the Tx then the receive signals of all data streams are entangled (see (4.29)) and uniform power allocation and the transmission of one data stream per Tx antenna is optimum. Therefore the adaptation of the transmission strategy reduces to power and rate allocation per antenna (PARC) which is in principle the same as for a Multi-User SIMO system described in the next subsection 4.4.2.

The power allocation and the belonging bit-loading vector can be calculated at the Rx and fed back to the Tx via a low bit-rate feedback channel. Alternatively, the bit-loading could be calculated at the Tx itself, based on CSI attained from a channel measurement into the opposite direction and further information about the noise level at the Rx. But then the noise level at the Rx has to be transmitted to the Tx and the allocated modulation has to signalled to the Rx in an appropriate way. So in the following we always assume a calculation of the bit-allocation at the Rx or BS.

4.4.2 Multi-user SIMO Bit-Loading

In real world transmission systems we have to do adaptive bit-loading to achieve as much throughput as possible. Therefore we propose an algorithm for MU SIMO bit-loading that follows mainly the same strategy like in the single user MIMO case with CSI at the Tx and eigenspace signalling [HB03] as discussed in the last subsection. Additional effort is needed to obtain the cost and gain functions for one bit upgrade or downgrade of a data stream, since the reconstructed data streams are still entangled due to the MMSE solution. In case of a ZF detector at the BS we can follow the proposed BLS without iterations for the recalculation of the SINR.

Here, more power put on data stream d_i directly improves the corresponding SINR at the decision unit, but all other data streams face a decreasing SINR because of rising cross talk from d_i. The final solution can be found only after a few iterations. Remember that in the single-user case [HB03] this was a single step task because the data streams were transmitted over the separated eigenvectors of $\mathbf{H}$ and cross talk was not apparent.

Nevertheless the crucial point is to find the optimal power allocation to initialize the algorithm which is not easy if the sum rate functional is not concave in $\mathbf{p}$.

Multi-user SIMO Bit-loading Algorithm

1. Start bit loading with the optimum solution from (4.36) or (4.43).
2. Compute the MSE_i / $SINR_i$ for each sub-stream i from the actual power allocation $\mathbf{p}$ and load it with the highest modulation satisfying the targeted BER constraint.
3. Find the minimum allocated Tx power per data stream to maintain the above computed bit allocation.
4. Compute cost and gain function for upgrading or downgrading each data stream by one bit.
5. Sum up the remaining power and give it to the stream which supports one more bit/Hz/s (a higher modulation) while it needs the least from the rest power. This is done by adjusting the transmit power allocation for all users.
6. Continue from 2. to 6. until no positiv rest power is available after an intended modulation step.
7. Step over, if already been here with the same bit allocation. Otherwise test, if taking one bit/Hz/s from one data stream and giving it to another saves power. Downgrade the most profitable data stream and proceed to step 2.
8. Distribute the remaining power in a way that all channels have the same SNR increase.

The described algorithm can be used in different ways, bit-loading can be performed with and w/o SIC and a MMSE or ZF detector.

A practical Bit-loading example

For a 2 user example with ($n_T = 2, m_R = 4$, $\sigma_N^2 = 1$) and the channel from (4.50) we perform bit-loading with and without SIC. The very important aspect of error propagation in bit loading schemes with SIC was ignored for this example. If this problem is tackled properly, error propagation can be considered within the required SINR for each modulation level and therefore decreases the sum rate gain achieved with SIC detection. For our example, we therefore assume error free cancellation of the first layer. The bit-loading algorithm described in 4.4.2 is initialized either with the optimum power allocation or all power allocations are conducted by stepwise increasing p_1 on the cost of p_2. The second strategy which is time consuming and would not be practicable in real applications was done simply to make sure, that all starting values were used as initialization for the bit-loading. Otherwise we could get stuck in a local maximum or furthermore we would try to support two users from the optimum solution while the achievable sum rate with discrete bit-loading may be higher if only one user is supported. The resulting maximum sum rate with and without SIC is depicted in Fig. 4.29.

We can state that the optimum power allocation with and without SIC is well suitable as a starting point for the bit-loading algorithm at least for the two user case. Then we achieve the maximum sum rate with discrete bit-loading in most of the cases. This optimum starting point can be calculated straight forward for SIC while more effort is necessary to calculate the theoretical optimum power allocation without SIC. In some cases another initial power allocation can achieve a higher sum rate than with the optimum solution as a starting point (dotted line in right part of Fig. 4.29). This might be explained by the non equidistant power steps in QAM which means that the bit-loading algorithm can get stuck with certain QAM levels while it finds a higher sum rate when the initial QAM symbol distribution is different due to a different starting point. Such deadlocks might be considered in a more complex bit loading algorithm.

The achieved sum rate with uncoded M-QAM bit-loading is depicted in Fig. 4.29. The SINR penalty compared to the theoretically achievable sum rate is exactly 5.2 dB which corresponds to the well known SNR gap for uncoded M-QAM with a BER$\leq 10^{-3}$.

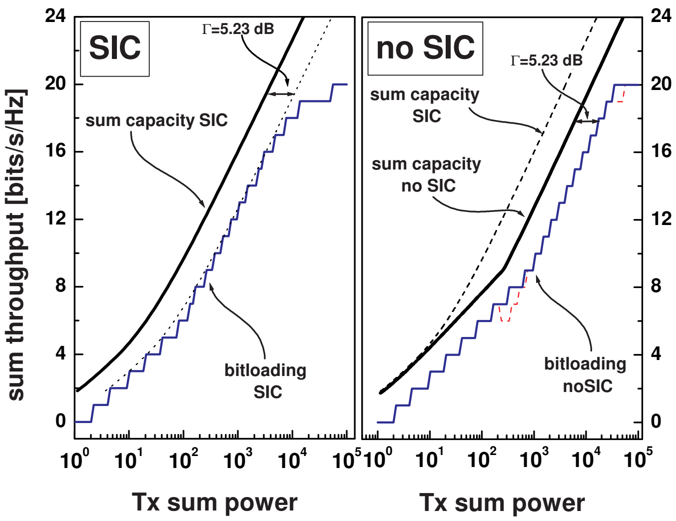

Figure 4.29: Theoretical sum capacity (solid lines) with MMSE receiver. Throughput with bit-loading (uncoded M-QAM with / without SIC) 2 user case, fixed channel, various sum powers P_{total} The rate is limited due to 10 bits as highest modulation level.

4.4.3 Multi-User Down-link

Bit-loading for multi-user down-link scenarios follow the same principle of bit and power allocation as the single user or MAC schemes.

First, the optimum or suboptimum power allocation has to be computed. Second, the resulting SINRs at the MTs can be calculated and be used as input parameter for the bit-loading algorithm. A further assumption is that the MTs do not suffer significantly from interference caused from other cells. In reality this external interference level has to be measured by the MTs and transferred to the BS to avoid rate overload. If the interference level is known to the BS and the interference is well behaved than this can be included into the power and rate allocation per user as discussed in section 4.3.3. A more detailed overview on which parameters are used in particular can be found in table 6.1.

A special form of adaptive bit-loading with common modulation for all data streams occurs when adaptive channel inversion is used as investigated in experiments in section 6.3.

4.5 Scheduling

Mobile communication often relies on a cellular structure e.g. GSM, UMTS which means that several users in a cell are supported by one base station. Since the number of users can be large, user scheduling becomes mandatory to allocate e.g. frequency and time slots in a reasonable fashion. This task is usually performed by the base station as the central communication point. This makes clear that multi user scheduling is a crucial point towards throughput optimization especially if QoS parameters like service quality, availability, delay are to be met by a service provider.

Transmission systems using several antennas at the BS and the MTs need a more sophisticated user scheduling since several users can be supported simultaneously by spatial multiplexing. A theoretical frame work using cross-layer optimization techniques was studied in e.g. [HBH03, BW02, BW03] and allows an efficient solution of the challenging optimization task of the multi user scheduling with multiple antennas.

In the following we will focus on the high speed up link essential to meet the increasing throughput demands for the so-called "new services" e.g. multi media messages (MMS) which allow the transmission of pictures and video streams from MTs over the internet or from MTs to MTs. A detailed discussion of the theoretical framework, which allows a computational efficient optimization in a cross-layer approach is given in [BW04a, WB04a, BW04c, BW04b]. We will only revisit a few results from this powerful framework which are relevant to derive the fair scheduling algorithm applied in the experiments in section 6. For the proofs and further results we refer to the references given above.

The high speed uplink corresponds in terms of information theory to the multiple-access channel (MAC). We assume K MTs indexed by numbers $k = 1, 2, \ldots, K$, each equipped with multiple transmit antennas. The BS is assumed to have full channel state information (CSI) and to utilize the MMSE (*minimum mean square error*) detector with successive interference cancelation (SIC). We concentrate on the scheduling policy originating from a cross-layer optimization problem including the physical layer (PHY) and the data link layer (DLL). In the PHY-layer we denote the vector of instantaneous data rates as $\mathbf{R}(n) = (R_1(n), R_2(n), \ldots, R_K(n))$ and group the MIMO channel states in the matrix set $\mathcal{H}(n) = \{\mathbf{H}_1(n), \mathbf{H}_2(n), \ldots, \mathbf{H}_K(n)\}$. Similarly, we group the instantaneous transmit covariance matrices in the matrix set $\mathcal{Q}(n) = \{\mathbf{Q}_1(n), \mathbf{Q}_2(n), \ldots, \mathbf{Q}_K(n)\}$. The $K!$ possible SIC-orders are denoted by permutation symbols $\pi_k = \pi_k(1) \leftarrow \pi_k(2), \ldots, \leftarrow \pi_k(K)$, $k = 1, 2, \ldots, K!$, where $\pi_k(1)$ corresponds to the last decoded link signal, ... and $\pi_k(K)$ to the first decoded link signal. In the DLL-layer we assume the K

processes of bit arrivals into the buffers to be Poisson. The arrival rates are grouped in the vector $\boldsymbol{\rho} = (\rho_1, \rho_2, \ldots, \rho_K)$ and the numbers of bit arrivals in n-th time slot $[(n-1)T, nT]$ are grouped in the vector $\mathbf{a}(n) = (a_1(n), a_2(n), \ldots, a_K(n))$. Similarly, we denote the vector of instantaneous buffer occupancies (bit-queue lengths) as $\mathbf{q}(n) = (q_1(n), q_2(n), \ldots, q_K(n))$.

We associate the scheduling policies with mappings of the form

$$\{\mathcal{H}(n), \mathbf{q}(n), n\} \longrightarrow \phi(\{\mathcal{H}(n), \mathbf{q}(n), n\}) = \{\mathcal{Q}(n), \pi_k(n)\} \tag{4.136}$$

and use also the splitted notation $\phi_{\mathcal{Q}}(\{\mathcal{H}(n), \mathbf{q}(n), n\}) = \mathcal{Q}(n)$

$$\{\mathcal{H}(n), \mathbf{q}(n), n\} \longrightarrow \phi_{\mathcal{Q}}(\{\mathcal{H}(n), \mathbf{q}(n), n\}) = \mathcal{Q}(n) \tag{4.137}$$

and $\phi_{\pi}(\{\mathcal{H}(n), \mathbf{q}(n), n\}) = \pi_k(n)$

$$\{\mathcal{H}(n), \mathbf{q}(n), n\} \longrightarrow \phi_{\pi}(\{\mathcal{H}(n), \mathbf{q}(n), n\}) = \pi_k(n) \tag{4.138}$$

to access separately the assigned transmit covariance matrices and the SIC-order. The principle of the policy computation is depicted in Fig. 4.30. With the above policy notion and i.i.d.-property of the fading processes over time it can be easily shown, that the queue system evolves according to the *Discrete Time Markov Chain* (DTMC) and the evolution can be described by

$$q_k(n+1) = [q_k(n) - R_k(\phi, \mathcal{H}(n))T]_+ + a_k(n). \tag{4.139}$$

The objective of our desired policy is associated with the notion of stability of the queue system in the uplink. In broad terms, the system is called stable if no queue in the system blows up to infinity during the system evolution. Stability can be characterized by several different definition, like e.g. weak stability, strong stability, non-evanescence [WB04b] etc., where a special role is played by so called observation-based stability notion.

Definition 3 (Observation-Based Stability): *The system of K queues is called stable, if for all $k = 1, 2, \ldots, K$ holds*

$$\lim_{M\to\infty} g_k(M) = 0, \tag{4.140}$$

with

$$g_k(M) = \limsup_{t\to\infty} \frac{1}{t} \int_0^t \mathbf{1}_{\{q_k(\tau)\geq M\}}\ \tau d\tau, \tag{4.141}$$

and

$$\mathbf{1}_{\{q_k(\tau)\geq M\}} = \begin{cases} 1 & q_k(\tau) \geq M \\ 0 & \text{elsewhere.} \end{cases} \tag{4.142}$$

The above stability notion gives rise to the definition of the stability region.

Definition 4 (Stability Region): *The stability region $\mathcal{D}$ of the system of K queues is the set of all arrival rate vectors $\boldsymbol{\rho}$, such that there exists a policy achieving stability in the observation-based sense for all $\boldsymbol{\rho}$ lying in the interior of $\mathcal{D}$.*

The scheduling policy achieving the largest stability region is the desired policy in this work. It can be concluded easily, that such policy leads to the best stability behavior in the intuitive sense. Moreover, a large stability region is also attractive from the system and operators point of view. It reduces the risk of buffer overflows[5] to the minimum, allowing in this way for well-behaved system operation and reducing operators utility loss, which results from dropping the service for overflowed links.

It was shown in [BW02], that the largest achievable stability region in the MAC corresponds to its ergodic capacity region.

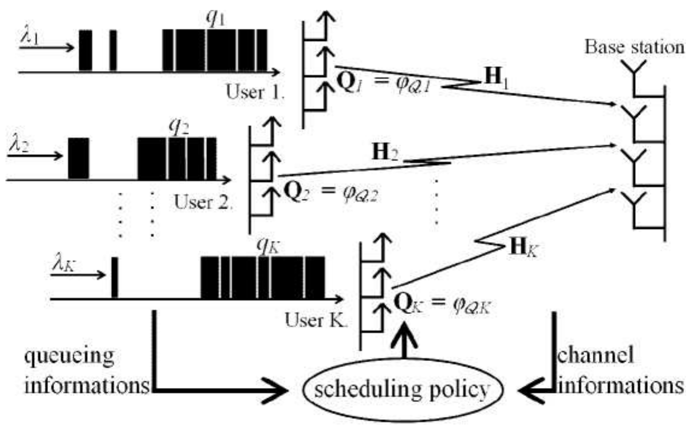

Figure 4.30: The routine of computation of a scheduling policy in the MIMO-MAC.

The corresponding scheduling policy is given by the following Theorem [WB04b].
Theorem 5: *The largest stability region in the MIMO-MAC with K links is achieved by the scheduling policy $\hat{\phi}$ satisfying*

$$\hat{\phi} = \arg\max_{\phi \in \mathcal{M}} \sum_{k=1}^{K} q_k(n) R_k\left(\phi, \mathcal{H}(n)\right), \tag{4.143}$$

for all $n \epsilon \mathbb{N}$.

This result appears not surprising for readers familiar with system control and dynamic system theory since the principle of weighted sum, as in (4.143), occurs in the context of stability optimal policies in many fields.

However, without specification of the optimal queue system and channel state dependent SIC-order the realization of policy from the above theorem results in $K!$-fold computation of optimal transmit covariance matrices for every possible SIC-order and a final comparison. Such principle is not practicable for common user numbers in the cell. The following statement provides more insights in the realization principle.
Theorem 6: *The largest stability region in the MIMO-MAC with K links is achieved by the spatial scheduling policy $\hat{\phi}^S$ satisfying*

$$\hat{\phi}^S_{\mathcal{Q}} = \arg\max_{\phi_{\mathcal{Q}}:\phi \in \mathcal{M}} \sum_{k=1}^{K} q_k(n) R_k\left(\phi_{\mathcal{Q}}, \mathcal{H}(n)\right) \tag{4.144}$$

[5] The buffers are assumed here to be large. Hence, the overflow is assumed to occur when the queue length approaches infinity.

and

$$q_{\pi(1)}(n) \geq q_{\pi(2)}(n) \geq \ldots, \geq q_{\pi(K)}(n) \geq 0, \tag{4.145}$$

with $\pi = \phi^S_\pi$ *for all* $n \epsilon \mathbb{N}$.

Using the notion of *S-rate region* $\mathcal{S}_\pi$ as special region of rates achievable with fixed SIC-order π, Theorem 6 is easily interpretable in terms of optimization over rates. Precisely, the stability optimal instantaneous rate vector solves

$$\max_{\mathbf{R} \in \mathcal{S}_\pi(\mathcal{H}(n))} \mathbf{q}^T(n)\mathbf{R}(n), \tag{4.146}$$

with π characterized by (4.145). A geometric view is provided with Fig. 4.31, where an exemplary instantaneous capacity region and S-rate regions of a two-link MIMO-MAC are plotted. The

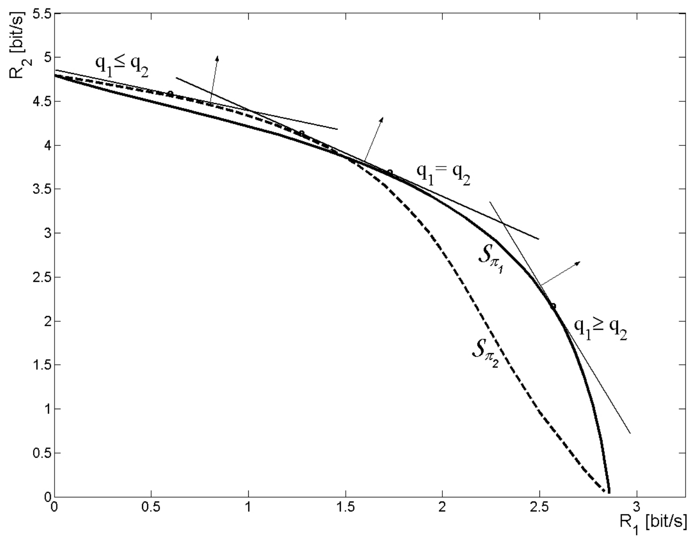

Figure 4.31: Exemplary instantaneous capacity region of a 2×2 sum-power constrained MIMO-MAC with two links (channel matrices generated randomly from the element wise uniform distribution with average SINR = 6 dB, SIC orders $\pi_1 = 1 \leftarrow 2$, $\pi_2 = 2 \leftarrow 1$). Three exemplary optimization objectives are modelled by lines with normal vectors corresponding to queue system states.

optimization objectives are associated with queue system states satisfying (4.145) for two different SIC-orders and both of them. Given any such $\mathbf{q}(n)$, the stability optimal rate vector represents the point of support at the boundary of the capacity region, which pertains to S-rate region $\mathcal{S}_{\pi_k}(\mathcal{H}(n))$ if $\phi^S_\pi = \pi_k$ satisfies (4.145). For the symmetric queue system state both SIC orders are stability optimal and the associated hyperline supports both S-rate regions. Fig. 4.31 makes plausible, that the boundary part of any S-rate region $\mathcal{S}_{\pi_k}(\mathcal{H}(n))$, which is supportable by hyperplanes satisfying (4.145) with $\pi = \pi_k$ is convex, whereas the complementary boundary parts are in general non-convex. Moreover, the independence of channel states in (4.145) shows, that the geometric "positions" of convex boundary parts of all S-rate regions do not depend on fading states.

4.5.1 Multi-user SIMO MAC - Scheduling

In the special case of MTs with only one single antenna the in general iterative calculation of the power allocation and transmit covariance matrices per user [E.J04] reduces to a per user power allocation. Furthermore we know from section 4.1.3 that under the assumption of a MMSE receiver at the BS, perfect channel knowledge and SIC at the receiver the sum capacity functional shows a concave behavior. This guarantees for a simple search algorithms to maximize the sum rate. As stated in (4.43) the maximum sum rate is independent on the detection order used for SIC. The same is valid for the optimum power allocation belonging to the maximum sum rate.

These two facts have an important consequence: By choosing a certain detection order we can influence the data rates of the individual users while the sum rate remains unchanged. The later a user is detected the higher is the individual data rate he receives. This behavior is used by strategies like [BW03, BW04a, YC03], well known from queueing theory. Simulations about individual achievable rates versus the position of the user in the detection order are shown in Fig. 4.32 and Fig. 4.33.

Let's now try to answer the question of practical relevance of fair scheduling for a cellular mobile network. We assume sets of K mobile users which are chosen out of L users within one cell; with $L > K$ and a MMSE receiver with SIC at the BS. Then we find:

- *Operators interest:* find a user set and a belonging power allocation which assure the highest sum transmission rate (compare C in Fig. 4.37). More transmitted MBytes in the cell mean more money to be billed.
- *Users interest:* small and more or less predictable delays according to certain QoS requirements[16](compare D in Fig. 4.37). Reliability in QoS can help to increase the acceptance of new applications.
- *Operators and users interest:* QoS means a new range of applications. The higher the QoS the more money the user might be willing to pay.

With multiple antennas at the BS we have the advantage of spatial multiplexing which means we have the opportunity of fair scheduling without loosing the maximum rate option. This fact is of great practical relevance. Imagine all users requested a certain quality of service (QoS) which very often simply means certain limits to real time delays. Here it becomes clear that an operator is also interested in delay optimized scheduling.

The last and only task is to find the best active user set, because the belonging optimum power allocation is easy to compute and the detection order results directly from the queue length of the users. This more general optimization problem including fair scheduling can be formulated with the following maximization task [BW03] in analogy to (4.144):

$$\hat{f} = argmax \sum_{i=1}^{K} q_i R_i(f, \mathbf{H}) \tag{4.147}$$

where q_i are the queue states, R_i are the instantaneous rates, $\mathbf{H}$ is the actual SIMO channel

[6]Remark: A systems like QUALCOM CDMA2000 1ev DO which is more or less a TDD system where only the user with the best channel is served at one time. This can cause tremendous delays for other users which are e.g. shadowed or far away from the BS. With Multiuser-SIMO fair scheduling appropriate delays and high throughput can be obtained at the same time.

matrix and f a queueing policy. The optimum queueing policy $\hat{f}$ achieves stability of all queues in the SIMO-MAC. To find the optimum solution still remains time consuming because of the high number of possible combinations of all users and their detection orders.

Since an exhausting search to maximize the above equation is not always a practical option we propose a suboptimum but quite straight forward solution as an applicable compromise.

Scheduling Proposal for K users out of L

Let us consider the following scenario. We assume L users in the cell and the BS can support K users simultaneously. Furthermore we assume that all users $(1...L)$ in the cell have a known data queue state q_i. Now, data throughput maximization aims to find the set of users which receives maximum throughput but disregards any queueing states. This may lead to queue instability when the individual queues grow very unequal.

On the other hand a delay optimum approach would favor a set of users with the longest queue states. This approach tries to minimize the actual queue length of all users but does not consider the sum rate.

Having said that we see the need to trade both optimizations against each other. One promising and quite pragmatic approach is to choose one or two users with the longest queue states and add those users which maximize the sum rate. The detection order to choose is obvious: the user with the longest queue is detected last and the user with the shortest queue is detected first like in (4.145). By choosing this scheduling scheme we can support delay limitation (fairness) and still have a good rate performance. This suboptimum but easy to compute solution provides little less sum rate than the sum rate optimum approach while keeping the maximum delay limited. For further details please compare scheduling policies C,D and E in Fig. 4.37 in section 4.5.2.

A variation of this fair scheduling scheme with a linear MMSE detector was implemented in the real-time MIMO test-bed and the experiments clearly show the advantage and the applicability of the fair scheduler (see sec. 6.4 for more details).

4.5.2 Numerical Simulations - Scheduling Examples

Detection Order and Achievable Rates

We assume a cell containing 6 users. The channel for a certain time slot is perfectly known to the BS with 5 Rx antennas. 5 users can be supported simultaneously. So there are 6 sets consisting of 5 users each. Each set of users can have $5! = 120$ detection order permutations.

We randomly choose a user i which is then held fixed after being selected. Next we choose a certain permutation of the four remaining users. We are now interested in the following: what happens to his individual rate if he is detected at a certain position before, between or after the other four co-users? Fig. 4.32 shows the achievable individual capacity of the elected user i if he is detected first, second,..., last. We clearly see that the individual rate is higher if a certain user is detected later within a fixed set of users. Per set there are $4! = 24$ possible detection orders of the 4 co-users and then 5 positions for user i.

Fig. 4.33 shows the same but the achievable rate after bit-loading with M-QAM is done. For

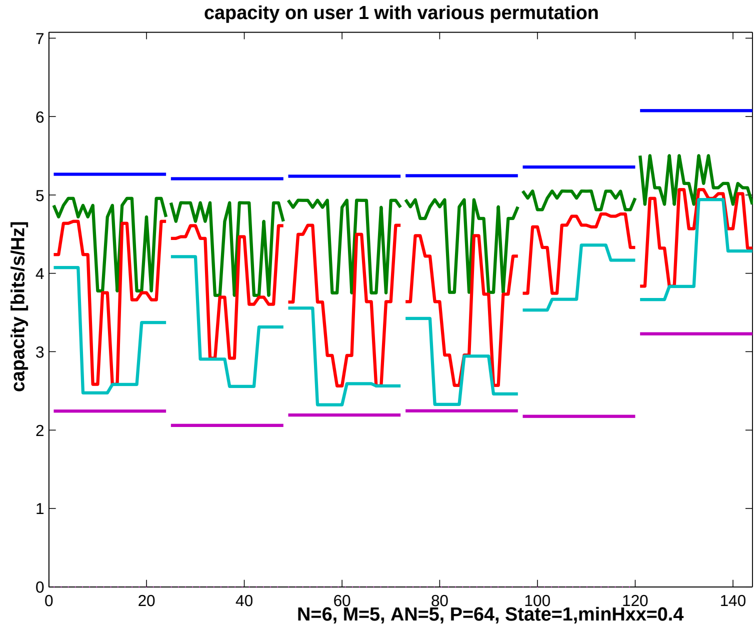

Figure 4.32: ordered capacities with optimum power allocation. User set: 5 out of 6 users in a cell. All order permutations are depicted.

the bit-loading we assumed uncoded M-QAM modulation and a BER$< 10^{-3}$. Compared with Fig. 4.32 we see a smaller rate due to bit-loading but in principle the rule still holds. If user i is detected later then his achievable rate with bit-loading is higher . The exceptions from this rule can be explained by the fact that practical bit-loading can only achieve integer valued rates. This means that the remaining power goes sometimes to other users or in reverse is given from other users to user i and user i can load one or two more bits. Therefore the rule is broken sometimes.

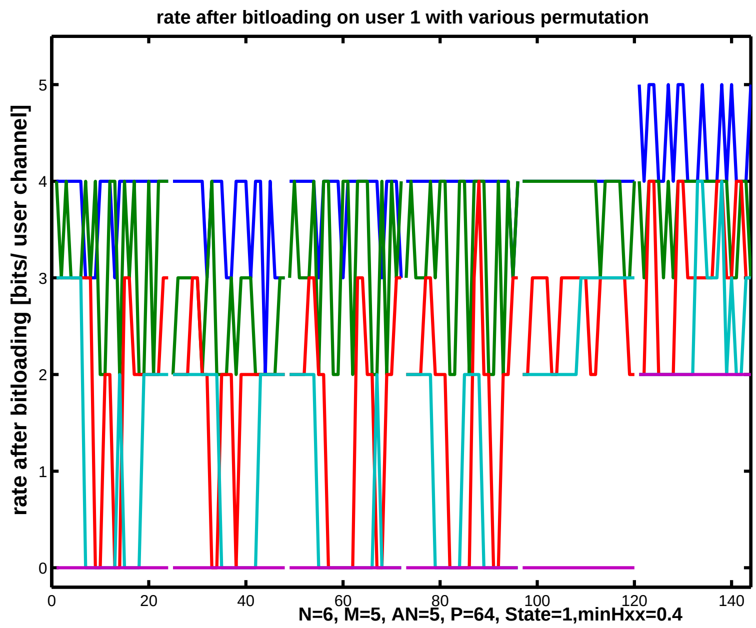

Figure 4.33: ordered rates achievable after bit-loading with optimum power allocation as starting point. User set: 5 out of 6 users in a cell. All order permutations are depicted.

Fig. 4.34 shows the sum rate for the scenario. We clearly see that the sum rate is independent on the detection order and is only a function of the chosen user set. With bit-loading the achievable sum rate is nearly constant. Some detection orders achieve one bit more or less in sum rate. This means that in principle, assuming optimal power allocation, the constant sum rate rule still holds with practical bit-loading.

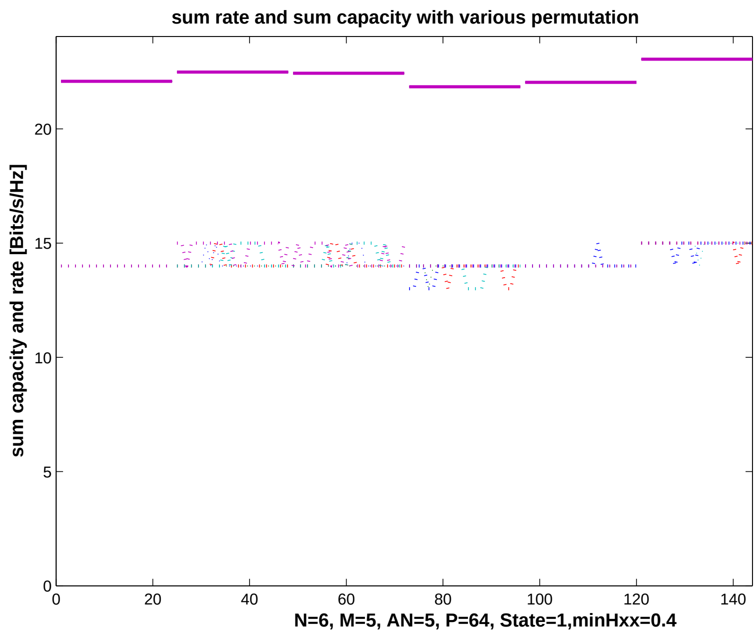

Figure 4.34: sum capacities and achievable sum rates after bit-loading with optimum power allocation as starting point. User set: 5 out of 6 users in a cell. All order permutations are depicted.

Bit-loading with fixed Detection Order - 3 user example

Fig. 4.35 shows the achievable sum rate with bit-loading in a 3 user scenario with fixed detection order. The sum rate is plotted versus the transmit power for fixed channel (top) and versus 100 random channel realizations with fixed transmit power. The noise power per antenna is held fixed for both scenarios. The normalized correlation[7] between all user channels is > 0.5. With lower correlation the impact of SIC decreases and bit-loading becomes much easier. The extreme is to be found when all channel vectors are orthogonal then SIC has no impact and the power allocation to one user does not influence the SINR of the other users.

To characterize the performance of the bit-loading we start the algorithm first with the easy to calculate optimum PA (blue bullets) and second with all possible PA within a discrete power steps size (green bullets). The later technique is simply to guarantee that we find the maximum rate with bit-loading but it would be totally inapplicable in real time. This methods allows to evaluate the rate performance of our new proposed bit-loading algorithm compared to the maximum achievable rate.

The blue bullets in Fig. 4.35-Top show the achievable sum rate when the bit-loading algorithm is

[7]The correlation of the entangled channel vectors $\mathbf{a}$ and $\mathbf{b}$ is used here in the sense of the normalized scalar product. $0 \leq \frac{\mathbf{a}}{|\mathbf{a}|} \cdot \frac{\mathbf{b}}{|\mathbf{b}|} \leq 1$, where 0 means no correlation or $\mathbf{a} \perp \mathbf{b}$ and 1 means full correlation or $\mathbf{a} \parallel \mathbf{b}$.

started with the optimum power allocation obtained by the sum capacity maximization. In some cases there exist another PA to start with which achieves one bit more. This is due to the fact that the power steps between each modulation type are not all equal. Therefore the algorithm can get stuck when started from a certain side. In average the loss is small as we can see in the Fig. 4.35-Bottom. Here the transmit and noise power are kept constant and 100 channels are randomly generated. The solid black line depicts the achievable capacity with optimum PA. The blue bullets show the achievable rate after bit-loading (optimum PA as starting point). Some green bullets above the blue ones indicate that sometimes we can load one or two bits more if the loading algorithm is started with some other initialization. The averaged loss over 100 channels is 0.1 bits/s/Hz which is below 1% because the average sum throughput is about 11 bits/s/Hz.

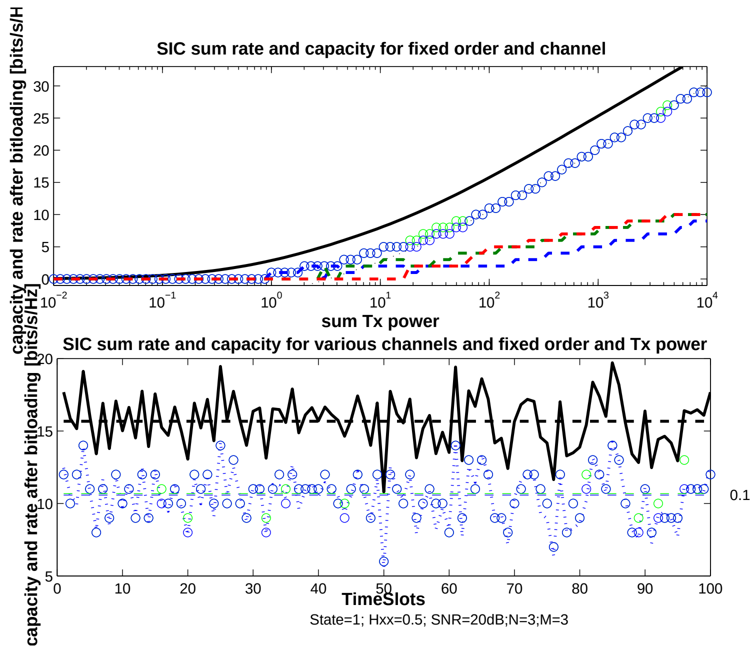

Figure 4.35: top: sum capacity and achievable sum rate for 3 users and fixed channel and SIC order; transmit power is varied. bottom: sum capacity and achievable sum rate for 3 users, fixed transmit power and SIC order; 100 channel realization with correlation of all user above 0.5

Optimal Power Allocation for fixed Detection Order - 3 user example

We assume a 3 user scenario with fixed channel and fixed detection order for the bit-loading algorithm. The noise power at the 3 receivers at the BS is identical and kept constant. The solid lines in Fig. 4.36 show the relative optimal power allocation versus the transmitted sum power. We see that one user after the other is switched on. At high SNR (high Tx power) the optimum solutions tends to uniform power allocation as expected from theory. At very low SNR (low Tx power) only one user is active - single user range as discussed in [BJ02].

The dotted lines represent the allocated power after bit-loading. The zig-zag is due to the integer valued rate allocation with practical bit-loading. The remaining power after bit-loading is depicted in dashed black. For high SNR it is about 10% in average. This value is expected to be smaller when the number of active users increases.

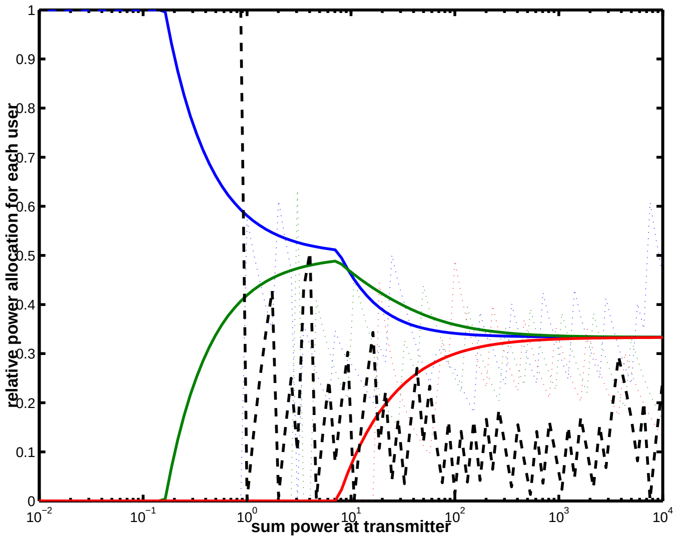

Figure 4.36: Optimum power allocation for a set of 3 users, fixed channel and fixed SIC order. Solid lines relative optimum power allocation to reach highest sum capacity - used as starting value for bit-loading. Dotted lines relative allocated power for user after bit-loading. The black dashed line depicts the remaining power after bit-loading.

Comparison of various schedulers with respect to delay and throughput

We assume a set of 5 users in a radio cell. 3 of them can form a set of active users which can be supported simultaneously by the BS. All user have instantaneous data-packet queues which grow due to independent Poisson distributed arrival processes and which decrease with the transmitted bits assigned by bit-loading. This is done over 1000 independent random channels (one channel per timeslot). The normalized correlation between all channel vectors is > 0.5 otherwise the right detection order is of less importance.

We compare 5 different scheduling policies (A,...,E) towards capacity and throughput after bit-loading (left column) and the delays and queueing states (right column) of all users in the cell. The red line in the left column always represents the sum capacity while the blue line represents the sum throughput after bit-loading (uncoded M-QAM, BER$< 10^{-3}$).

A: best user only scheduler (similar to CDMA 2000 1ev DO) Scheduler A supports only one user. This simply considers the channel vectors of all users in the cell. Users which are shadowed or far from the BS would experience no connection over a long period of time because always someone else has a better channel. Since the channels are randomly generated all users are more or less equally supported over the time of 1000 time slots. The achievable rate is the smallest compared to all other scheduling schemes because here only receive diversity is exploited and no multiplexing gain can be achieved. In consequence the queueing states grow constantly which means this scheduler is not stable, queue-wise seen.

B: 3 of 4 cyclic scheduler Scheduler B supports always 3 users simultaneously. We start with user set 1,2,3 next comes user set 2,3,4 and so on. This scheduler does not exploit any knowledge about the actual channel state nor it considers any packet-queues at the user nodes to form a user set. The instantaneous queueing states are only used to determine the optimum detection order. Still this is not enough to keep the queueing states similar in length nor to stabilize the queue length. The queue buffer fill up to over 200 packets. The sum rate is higher then with scheduler A due to the multiplexing gain.

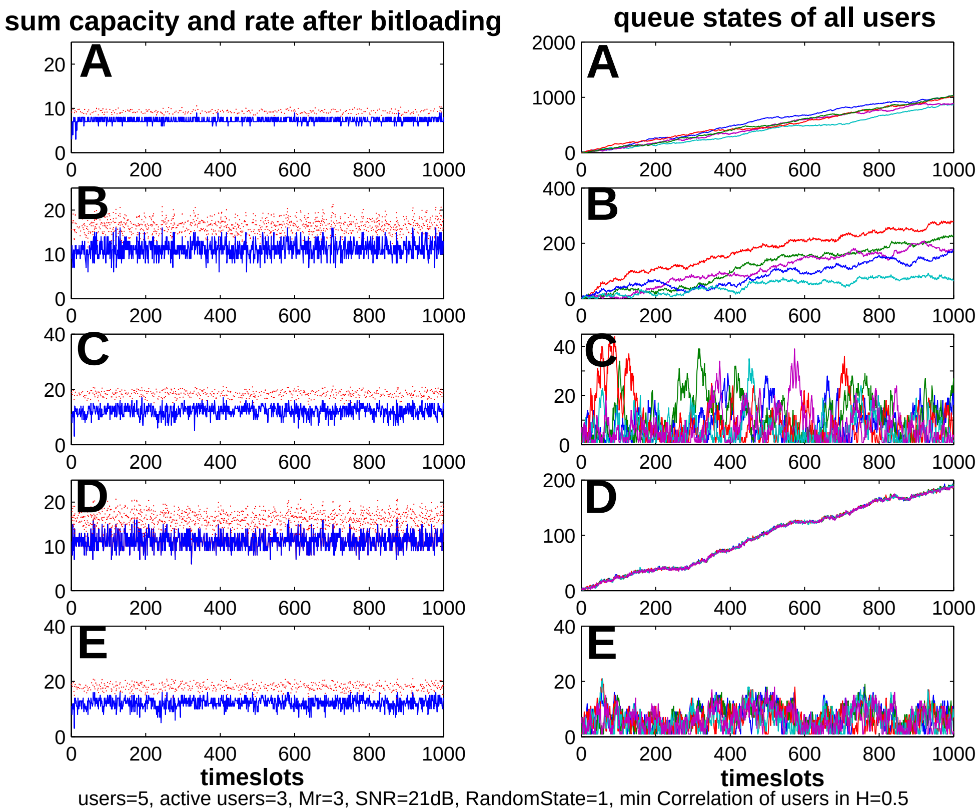

Figure 4.37: Comparison of various scheduling policies regarding maximum sum capacity and queueing states of the users. A maximum of 3 users out of 5 users in the cell are supported in one time slot. Schemes: A) best user only (capacity criterion), B: 3 of 4 cyclic scheduler, C: maximum sum rate approach and delay considering SIC ordering, D: Minimum delay approach, E: User with longest queue plus 2 users which maximize the sum capacity. Left column: Achievable capacity and rate after bit-loading. Right column: Queueing states of all 5 users versus channels realizations / time slots.

C: max sum rate scheduler Scheduler C selects always those users which form the user set with the highest sum rate. The queueing states of the users are not important for this selection. Only after the user set is chosen this knowledge determines the detection order. As expected the sum capacity and sum rate after bit-loading are the highest of all schedulers and the queues at all users keep limited. This is due to the high throughput on the one hand and due to the random channels on the other hand. Still we see queueing states of one user which remain above 30 bits over a period of about 50 time slots which is equivalent to a noteworthy delay. If a generic channel model would be used those delays can increase even more.

D: min delay scheduler Scheduler D selects always those three users which have the longest queue regardless of the actual user channels. In consequence this works like a delay equalizer, all queues sizes are kept similar. Because the sum rate is not a parameter to form the user set the average throughput is to small to keep the system within a stable region. Therefore all queues grow constantly.

E: new proposed fair scheduler Scheduler E selects always the user which has the longest instantaneous queue length. Next those two users are added which achieve the maximum sum rate together with the first selected user.

We see that the sum rate is slightly lower than that of scheduler C but the queues can be kept much shorter for all users. Therefore this new fair scheduler outperforms all other schedulers compared here regarding queue length.

Summarizing the information of this comparing figure we clearly see the necessity of a combined approach for a fair and stable scheduler. Since an optimization over all possible combinations of users and their detection order is to exhaustive in time we proposed a quite simple scheduling policy which was exemplarily presented in form of scheduler E.

5 Real-Time Algorithms for Multiple-Antenna Systems

5.1 MIMO Algorithms and Optimization

5.1.1 Basic Algorithmic Strategies for Real-Time Multi-Antenna Systems with high Data Rates

With the perspective of real-time capable algorithm implementation for very high data rates the complexity of algorithms often becomes a limiting factor. Therefore it is reasonable to search for solutions which have a high performance and match the capability of a dedicated hardware.

The hybrid FPGA / DSP architecture of the test-bed gives a high flexibility over algorithms used for data stream separation at the Tx and/or the Rx, rate and power control. Those algorithms are run on the DSP while the fixed part (e.g. channel estimation, data separation, Mod/Demod, BER) is performed by the FPGA. The DSP works fully asynchronous and refreshes e.g. the necessary MMSE weights and/or the bit-loading vector at the Tx-FPGA within a millisecond or less.

Following this divide and rule strategy we are able to support high data rates in a MIMO transmission and still have the flexibility towards algorithms.

To realize this ambitious approach we implemented the high speed matrix vector multiplications for the reconstruction of the data streams in VHDL on the FPGA and the DSP performs the calculation of the required matrices. The complexity which can be implemented in the FPGA is mainly limited by the number of dedicated multipliers, RAM etc. and particularly by the maximum clock rate at which the design can be routed within the required delay limits. The more resources are used from the FPGA (70% or more) the more difficult the place & route procedure becomes. The limiting factor for high speed signal processing in the FPGA is determined by the ADC, DAC and FFT/IFFT blocks (e.g. OFDM) which run at the highest clock rates which is limited to 150-200 MHz in reality (Virtex II Pro 100), which equally limits the usable signal bandwidth to be used for transmission. This means that for high data rates of several 100 Mbit/s to 1 Gbit/s or more, higher modulation levels and spatial multiplexing are a necessity for spectrally efficient transmission.

A recent FPGA implementation of MIMO-OFDM at a clock rate of 100 MHz [JFH+05a] allowed a reliable low mobility transmission with a gross data rate of 1 Gbit/s with 3 Tx and 5 Rx antennas using 48 active OFDM carriers and 100 MHz bandwidth at 5.2 GHz.

If the data transfer on the parallel bus between DSP and FPGA is optimized, then the calculation of the detection matrices itself can become the most time consuming part. The received signals of the current MIMO-OFDM system with 3 Tx and 5 Rx antennas and 48 carriers, are again treated as real-valued e.g. due to remaining I/Q-imbalances[1]. Therefore the DSP calculates 48 MMSE solutions where each matrix has size 10×6. If we remember that matrix inversions have roughly

[1] The I/Q mismatch is below 1 degree with a newly improved RF frontend.

a complexity $\sim N^3$ for square matrices it becomes clear that the optimization of DSP code is crucial. If the number of sub-carriers is high (256 or 1024) we will use DSP clusters which can work in parallel to perform the calculation task still within the channel coherence time. In many transmission scenarios the channel has only a a few taps (10 or less), hence bout three times as many sub-carriers would be sufficient to equalize the channel. But for the sake of spectral efficiency many more OFDM sub-carriers are normally used which now carry redundant information. This redundancy can be exploited to reduces the MIMO signal processing significantly. A promising approach is the calculation of an exact solution (e.g. ZF-pseudo-inverse as proposed by [BB04b]) on $(L-1)(N_T-1)+1$ sub-carriers only and to interpolate the filter solutions in between [2]. If this is done in an appropriate trigonometrical fashion [HMW05] the interpolated filter matrices can reconstruct the multiplexed data streams with high accuracy. The savings in time for the calculation of the MMSE solutions have to be traded carefully against the additional effort for the interpolation.

MIMO transmission schemes require specific algebraic procedures to be performed in order to pre-code or decode the data appropriately. Some useful algorithms are discussed in the following paragraphs. Most of them were implemented on the DSP in C language and used for the calculation of the MIMO-filter matrices in the transmission experiments.

5.1.2 DSP - Architecture and Optimization

One of the initial questions to be asked is: what to use - floating point or fixed point arithmetic? Fixed point DSPs are offered on the market at much higher clock rates (e.g. 1 GHz) than floating point DSPs (300 Mhz), so one might say let's take the faster one. But this is only true if all calculations are performed in the integer domain and the dynamic range is fixed and well known. If floating types like float or double are used, the mapping to integer numbers is performed automatically by the compiler. A simple test showed that e.g. a matrix inversion on a 16-bit fixed-point TI-DSP (1 GHz) performs slower than the 300 MHz 32-bit floating-point DSP (TI6713) by a factor of 10. A way out is to optimize the mapping by hand using additional knowledge about the dynamic range etc. A major drawback of this approach is that hand optimized program code is hard to read and therefore very error-prone and not very flexible to code changes, not to mention a lot of overhead may occur when different people are contributing to the same algorithm library without necessarily knowing all details on dynamic range of the possible input and output values. Furthermore there are no floating point instructions available when it comes to assembler programming.

Therefore we chose a floating-point architecture (TI6713) with 225 MHz for the test-bed to have as much algorithmic flexibility as possible.

[Sch04] investigated several MIMO-algorithms in great detail regarding general C-code and assembler optimization. We will limit ourselves to some principle steps which can help to speed up the

[2] L denotes the order of the channel and n_T the number of transmit antennas. The classical approach of interpolation of the frequency channel estimates by a transfer into time domain, appropriate windowing and a back transformation to the required number of sub-carriers in the frequency domain improves the accuracy of the channel estimation but does not help to reduce the calculation effort at all. Note, that the filter envelopes of analogue or digital filters which are used for image band suppression have to be measured carefully before interpolation techniques can be exploited. This is important in particular when more than 80% of the OFDM sub-carriers are used, which can be done with channel adaptive bit-loading.

programs. The optimization was tailored for a TI6713 floating-point DSP, but if the designated features of other DSP types are considered properly then the optimization process can be followed in the same manner.

- load program and data into internal memory
- optimize compiler settings (big impact on performance)
- use only one-dimensional arrays
- use optimized library 'fastrts67x.lib' for divisions, reciprocals, roots etc.
- avoid divisions if possible
- avoid if-then-else branches
- the length of program loops should be constant (especially inner loops)
- place long loops as inner loops (short loops as outer loops)
- optimize cache
- write assembly code

5.2 Matrix Inversion and Decompositions

Mainly all MIMO pre-coding and reception techniques are based on matrix $\times$ vector multiplications either in a linear sense or a non-linear sense which means repeating matrix $\times$ vector operations with decisions in between. The required matrices are mostly obtained by matrix decompositions or matrix inversions, so the following paragraphs are dedicated to those very important algebraic algorithms. Since real-time capability is mandatory for high data rate MIMO applications, speed and numerical stability are of great importance.

5.2.1 The Inverse of a Matrix and the Pseudo-Inverse

In multiple antenna systems the signals coming from all Tx antennas are superimposed at the Rx antennas. For the separation of these signals e.g. a linear filter can be used. A simple realization can be achieved with a zero-forcing (ZF) filter while the minimum mean square error (MMSE) is more complex but considers the noise from the Rx and outperforms ZF especially in the low SNR region. Both solutions require one matrix inversion.

This linear equalization at the Rx which corresponds to a multiplication of the receive vector $\mathbf{y}$ with a matrix $\mathbf{H}^\dagger$ the estimates for the transmitted data read

$$\hat{\mathbf{x}} = \mathbf{H}^\dagger \mathbf{y} = \mathbf{H}^\dagger \mathbf{H} \mathbf{x} + \mathbf{H}^\dagger \mathbf{n} = \mathbf{x} + \mathbf{H}^\dagger \mathbf{n} \tag{5.1}$$

where the ZF-pseudo-inverse of $\mathbf{H}$ if $m_R \neq n_T, m_R > n_T$ is

$$\mathbf{H}^\dagger_{ZF} = (\mathbf{H}^\mathbf{H}\mathbf{H})^{-1}\mathbf{H}^\mathbf{H}, \tag{5.2}$$

or if we considering the receiver noise, additionally, the belonging MMSE filter reads

$$\mathbf{H}^{\dagger}_{MMSE} = \mathbf{H^H}(\mathbf{HH^H} + \sigma_N^2 I)^{-1} \tag{5.3}$$

where the noise variance σ_N^2 is assumed to be the same for all receivers for a more convenient notation. Note, that in general we have to expect different noise variances for each receiver if e.g. independent automatic gain controls (AGC) are used. Then we have to use

$$\begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_{m_R}^2 \end{pmatrix} \text{ instead of } \sigma_N^2 I.$$

$\mathbf{H}^{\dagger}\mathbf{n}$ denotes the noise enhancement due to the linear filter.

Matrix-Inverse By definition, the inverse of a matrix only exist for matrices with the same number of rows and columns. Let $\mathbf{A}$ be a matrix of size $m_R \times n_T$ with $m_R = n_T$. Then we define $\mathbf{A}^{-1}$ the inverse of matrix $\mathbf{A}$ if holds

$$\mathbf{I}_{n_T} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} \tag{5.4}$$

where $\mathbf{I}_{n_T}$ is the unity matrix of size $n_T \times n_T$.

Pseudo-Inverse If $\mathbf{A}$ is of rectangular shape $m_R \times n_T$ with $m_R \geq n_T$ then an inverse is not defined. Therefore a so-called pseudo-inverse has to be computed instead.

$$\mathbf{A}^{\dagger} = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H \tag{5.5}$$

where $(\mathbf{A}^H\mathbf{A})^{-1}$ has square shape and standard algorithms for matrix inversion are applicable. $\mathbf{A}^{\dagger}$ then satisfies $\mathbf{I}_{n_T} = \mathbf{A}^{\dagger}\mathbf{A}$ similar like in (5.4).

Greville's Method: One straight forward approach to implement the calculation of the inverse and/ or pseudo-inverse of a matrix especially if the matrix is not necessarily of quadratic shape is using Greville's method (page 48 in [Gan86]). This algorithm provides full flexibility in the number of Tx and Rx antennas and even some columns or rows can be zero vectors.

While the ZF filter from (5.2) can be calculated directly from $\mathbf{H}$ instead inverting $\mathbf{H}^H\mathbf{H}$, the MMSE filter from (5.3) requires two extra matrix multiplications and the inversion of $(\mathbf{H}\mathbf{H}^H + \sigma_N^2\mathbf{I})$ which is of size $m_R \times m_R$.

Keeping in mind that the computational effort of multiplications and inversions increases by $\sim N^3$ with $N = max(n_T, m_R)$ we choose a dimension reduced formulation of the MMSE.

$$\text{reduced MMSE: } \mathbf{H}^{\dagger}_{MMSE} = (\mathbf{H}^H\mathbf{H} + \tilde{\sigma}_N^2\mathbf{I})^{-1}\mathbf{H}^H, \tag{5.6}$$

where $\tilde{\sigma}_N^2$ is the effective noise power per data stream.

This "short form" of the MMSE filter needs about half the computational effort compared to the classical MMSE solution. For instance, if $\mathbf{H}$ has size 4x3 then $\mathbf{H}^H\mathbf{H}$ is of size 3x3 and we need about $3^3 = 27$ operations while $\mathbf{H}\mathbf{H}^H$ is of size 4x4 and $4^3 = 64$ operations are required for the multiplication.

Furthermore the range issue of the data is very important in the conjunction with algorithms for the calculation of the pseudo-inverse, since a calculation of $\mathbf{H}^H\mathbf{H}$ doubles the binary range from e.g. 12 bits to 24 bits which often can decrease the algorithmic stability. This range extension is not required when Greville's method is used, so this may be an algorithm of choice for fixed point implementation.

Frobenius Decomposition: Another algorithm which can be used is based on the a modification of the Frobenius formula (page 73 in [Gan86]) where the calculation of a pseudo-inverse can be performed by the calculation of pseudo-inverses of sub-matrices.

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{K}^{-1} & -K^{-1}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}\mathbf{K}^{-1} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}\mathbf{K}^{-1}\mathbf{B}\mathbf{D}^{-1} \end{pmatrix} \tag{5.7}$$

where $\mathbf{K} = \mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}$. If the sub-matrices are regular and of square shape (e.g. $\mathbf{A}$) then inversion can be performed by calculating the elements of the inverse matrix $\mathbf{A}^{-1}$ directly with

$$a_{ik}^{(-1)} = \frac{\mathbf{A}_{ki}}{\|\mathbf{A}\|}. \tag{5.8}$$

The implementation of (5.8) is quite straight forward up to a matrix size of 4x4 real values. For instance if the matrix $\mathbf{H}$ is of size 6x6 or 8x8 then a decomposition into 3x3 or 4x4 sub-matrices is advised, respectively.

Example: An implementation carefully matched to the internal structure of the DSP in our test-bed (TI6713) has reduced this value down to 10 μs (4Tx x 4Rx, real valued). So, in principle a MIMO system (2Tx x 2Rx, complex valued) using e.g. 48 carriers actively like in 802.11a/g can be tracked in a total time of less than a millisecond with a single DSP, which is expected to be sufficient for indoor and pedestrian applications.

If $\mathbf{A}$ is of rectangular shape then we can reformulate the expression according to (5.5) $\mathbf{A}^\dagger = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ where $(\mathbf{A}^H\mathbf{A})^{-1}$ has square shape and (5.8) can be applied again for the matrix inversion. Additional matrix multiplications are a small price to pay for the reuse of the very fast algorithm to compute the inverse of a small square matrix. Note, that it is very important to keep in mind that rounding errors can cause instability of the algorithm if only 32-bit floating point numbers are used. This can be mitigated by reasonably scaling of the matrix entries before calculations or using the double format (64-bit) instead.

Alternatively, all sub-matrices can be pseudo-inverted using Greville's method like discussed before.

Gauss-Jordan-Elimination: For the special case of the Inversion of a square matrix with full rank, which is true for the MMSE solution with non-zero noise in (5.3) and (5.6) there is another option to obtain an matrix-inverse. Following the outline of page 36 in [PTVF92] Gauss-Jordan elimination has the advantage of a high numerical stability, especially when full pivoting is used. Furthermore the structure of the algorithm allows a very efficient manual optimization of the C-code.

5.2.2 The Singular Value Decomposition

For the case of perfect CSI at the Tx we know that the transmission into the direction of the eigenvectors of the channel feeds the data streams perfectly into the spatial sub-channels. A power allocation per data stream e.g. water-filling can be used to meet certain optimization targets as discussed in sec.4.2.2. The necessary matrices are calculated with a singular value decomposition (SVD).

Another very important information about the degree of singularity of the channel $\mathbf{H}$ can be derived from the condition number of the matrix. If only an inversion of a matrix is intended then SVD solution gives the closest to the null space if the channel matrix is singular or very close to singular. For a more detailed discussion see [PTVF92, GL96].

The SVD methods are based on a theorem of linear algebra, whose proof is beyond our scope: Any $M \times N$ matrix $\mathbf{H}$ whose number of rows M is greater than or equal to its number of columns N, can be written as the product of an $M \times N$ column-orthogonal matrix $\mathbf{U}$, an $N \times N$ diagonal matrix $\mathbf{D}$ with positive or zero elements, and the hermitian transpose of an $N \times N$ orthogonal matrix $\mathbf{V}$. The various shapes of these matrices will be made clearer by the following tableau:

$$\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^H \tag{5.9}$$

$$\begin{pmatrix} \\ \\ \mathbf{H} \\ \\ \\ \end{pmatrix} = \begin{pmatrix} \\ \\ \mathbf{U} \\ \\ \\ \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_{n_T} \end{pmatrix} \cdot \begin{pmatrix} & \mathbf{V}^H & \end{pmatrix} \tag{5.10}$$

The matrices $\mathbf{U}$ and $\mathbf{V}$ are each unitary in the sense that their columns are orthonormal. The matrix $\mathbf{D}$ has a diagonal structure and holds

$$\text{trace}(\mathbf{D}^2) = \text{trace}(\mathbf{H}\mathbf{H}^H) \tag{5.11}$$

which represents the transmitted power over the transmission channel. The matrix entries λ_i in $\mathbf{D}$ are called **singular values** (SV) and correspond to the transmission quality of each sub-channel, respectively. They are the relevant input parameters for power and bit-allocation algorithms e.g. water-filling or SVD-MIMO bit-loading.

We define the condition number of a matrix to be $Cond = \frac{\lambda_{max}}{\lambda_{min}}$. The "best" condition number is therefore 1 meaning all singular values are the same or if $Cond = \infty$ than at least one singular value is zero and one greater than zero. The size of the condition number gives an indication on possible rank deficiency due to a close to singular structure of the matrix. One practical option is e.g. not to load data on weak sub-channels or decrease the number of parallel transmitted data streams to avoid overloading a rank deficient channel.

The algorithm used for the implementation on the DSP is an adaptation of the *svdcmp* algorithm (page 67 of [PTVF92]) which is in fact based on an algorithm from Golub and Reinsch [WR71].

Performing a full SVD has a complexity of approx. $\mathcal{O}(22N^3)$ which is about one order of magnitude higher than a QR-decomposition (QRD) which has complexity of approx. $\mathcal{O}(\frac{2}{3}N^3)$. A recent proposal by [HK04] suggested the use of iterative QRD based pre-coding in a TDD system with assumed reciprocal base band channel. Before emitting the data over the air each side performs a linear pre-coding using a unitary pre-coding matrix $\mathbf{Q}^H$ obtained from a QRD of the transmission channel measured from the previous and opposite transmission phase. Following this strategy each side will measure the composed channel of $\mathbf{H} \cdot \mathbf{Q}_i^H$ at iteration step (i) and then decompose this new channel with QRD. The authors claim, that the obtained unitary matrices $\mathbf{Q}_i$ soon converge to the desired matrices $\mathbf{V}$ and $\mathbf{U}$ from a full SVD. When convergence is achieved after a few iterations in a ping pong like manner, tracking of time-variations of the channel performs quite well. This approach seems promising to reduce the required time to calculate the matrices for the SVD scheme, significantly, but further investigations have to be made towards imperfect base band channel reciprocity.

5.2.3 The QR- and QL-Decomposition

As discussed in sec. 4.1.3 non-linear decoding and pre-coding can achieve a higher performance due to a SNR gain or a transmit power reduction, respectively.

A successive interference cancellation (SIC) detector e.g. VBLAST can be reformulated as a decision feedback equalizer (DFE) which can be composed from matrices derived from a **QR decomposition** (QRD) or **QL decomposition** (QLD). Whether a QRD or a QLD has to be chosen depends mainly on the fact if the user 1 or user k out of a number of ordered users $(1, 2 ... k-1, k)$ is to be detected first or last.

QRD decomposes e.g. the channel matrix $\mathbf{H}$ into a unitary matrix $\mathbf{Q}$ and an upper right triangular matrix $\mathbf{R}$.

$$\mathbf{H} = \mathbf{Q} \cdot \mathbf{R} \tag{5.12}$$

$$\left(\quad \mathbf{H} \quad \right) = \left(\quad \mathbf{Q} \quad \right) \cdot \begin{pmatrix} r_{11} & \cdots & r_{1k} \\ 0 & \ddots & \vdots \\ 0 & 0 & r_{kk} \end{pmatrix} \tag{5.13}$$

while **QLD** decomposes $\mathbf{H}$ into a unitary matrix $\tilde{\mathbf{Q}}$ and a lower left triangular matrix $\mathbf{L}$.

$$\mathbf{H} = \tilde{\mathbf{Q}} \cdot \mathbf{L} \tag{5.14}$$

$$\left(\quad \mathbf{H} \quad \right) = \left(\quad \tilde{\mathbf{Q}} \quad \right) \cdot \begin{pmatrix} l_{11} & 0 & 0 \\ \vdots & \ddots & 0 \\ l_{k1} & \cdots & l_{kk} \end{pmatrix} \tag{5.15}$$

The columns of $\mathbf{Q}$ and $\tilde{\mathbf{Q}}$ or orthonormal.

Example: The construction of the two matrices is algorithmically straight forward. For conveniences we denote $\mathbf{h}_1$ to be the first column of $\mathbf{H}$. For QRD we start from the left. The first column in $\mathbf{Q}$ is the normalized first column from $\mathbf{H}$. The normalization factor r_{11} is given by $|\mathbf{h}_1| = (\mathbf{h}_1^H \cdot \mathbf{h}_1)^{1/2}$ which denotes the norm of $\mathbf{h}_1$.

$$\mathbf{h}_1 = \mathbf{q}_1 r_{11} \tag{5.16}$$

$$r_{11} = |\mathbf{h}_1| \tag{5.17}$$

Next $\mathbf{h}_2$ is projected onto $\mathbf{q}_1$ and subtracted from $\mathbf{h}_2$. The result obtained from the normalization is then $\mathbf{q}_2$.

$$\mathbf{h}_2 - (\mathbf{h}_2^H \cdot \mathbf{q}_1)\mathbf{q}_1 = \mathbf{q}_2 r_{22} \tag{5.18}$$

$$r_{12} = \mathbf{h}_2^H \cdot \mathbf{q}_1 \tag{5.19}$$

$$r_{22} = |\mathbf{h}_2 - (\mathbf{h}_2^H \cdot \mathbf{q}_1)\mathbf{q}_1| \tag{5.20}$$

Next we project the third column $\mathbf{h}_3$ onto $\mathbf{q}_1$ and $\mathbf{q}_2$ and normalize the third column in $\mathbf{Q}$.

$$\mathbf{q}_3 r_{33} = \mathbf{h}_3 - (\mathbf{h}_3^H \cdot \mathbf{q}_1)\mathbf{q}_1 - (\mathbf{h}_3^H \cdot \mathbf{q}_2)\mathbf{q}_2 \tag{5.21}$$

$$r_{33} = |\mathbf{h}_3 - (\mathbf{h}_3^H \cdot \mathbf{q}_1)\mathbf{q}_1 - (\mathbf{h}_3^H \cdot \mathbf{q}_2)\mathbf{q}_2| \tag{5.22}$$

$$r_{13} = \mathbf{h}_3 \cdot \mathbf{q}_1 \tag{5.23}$$

$$r_{23} = \mathbf{h}_3 \cdot \mathbf{q}_2 \tag{5.24}$$

If $\mathbf{H}$ has more than 3 columns the procedure has to be continued further on.

Sorted QLD: For many non-linear detection or pre-coding strategies it is mandatory to control the detection/pre-coding order, properly. Therefore the appropriate ordering, e.g. V-BLAST ordering, has to be obtained by a sorting algorithm. This can be done in advance of the QLD or together with the QLD itself. In principle, a calculation of the whole V-BLAST algorithm has to be performed despite the fact that only the detection order is of interest. Since the classical V-BLAST detection requires one matrix inversion per decoding step it has a total complexity which grows with

$$\sim n_T \cdot m_R^3. \tag{5.25}$$

A pre-sorting algorithm proposed by [WBR+01] avoids inversions and performs much better than random or fixed ordering for little numbers of detection layers. The main drawback of this algorithm compared to the optimum ordering is to be seen in the fact that the ordering is calculated more or less backwards and based on the columns of $\mathbf{H}$ instead of a forward calculation based on the rows of $\mathbf{H}^\dagger$ as done in the original V-BLAST algorithm. This results in the very undesired effect that the first layer to be detected is not necessarily the layer which has the best SINR and therefore error propagation is more likely.

To avoid suboptimum ordering while still exploiting the low complexity of the QLD, a post-sorting algorithm from [WBKK03] can solve the problem. The degree of sorting of the diagonal entries in the lower triangular matrix L gives an indication of how far the pre-sorting differs from the optimum V-BLAST sorting (optimum ordering: the diagonal entries in L appear in ascending order). If suboptimum sorting is detected the post-sort algorithm rearranges the sorting in the QLD and guaranties optimum ordering in the end. Since a post-sorting is not always required and mainly the first layers have to be detected in the right order, the combined pre-sort and post-sort algorithm is a good alternative to achieve V-BLAST ordering without performing the whole V-BLAST algorithm.

The entries of $(\mathrm{diag}(\mathbf{L}))^{-1}$ represent the noise enhancement for each detection layer with the SIC detector and can be used directly for the power and bit-loading algorithms.

To complete the calculation of all matrices required for a QLD-based DFE structure we further have to calculate the weight feed-forward matrix $\mathbf{GF}$ and feed-back matrix $\mathbf{B}-\mathbf{I}$

$$\mathbf{GF} = (\mathrm{diag}(\mathbf{L}))^{-1} \cdot \mathbf{Q}^H \tag{5.26}$$

$$\mathbf{B}-\mathbf{I} = (\mathrm{diag}(\mathbf{L}))^{-1} \cdot \mathbf{L} - \mathbf{I} \tag{5.27}$$

For the experiments described in 6.2.2 a sorted QLD algorithm was used including the sorting routine motivated by the above algorithms from [WBKK03]. It has to be noted that if MMSE-SIC is to be used the original matrix $\mathbf{H}$ has to be expanded by an $n_T \times n_T$ noise weighted identity matrix and the QLD has to be performed on this expanded matrix $\tilde{\mathbf{H}}$ which has size $(m_R+n_T)\times n_T$ with $n_T = 3$ and $m_R = 4$ in the following example:

$$\tilde{\mathbf{H}} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \\ h_{41} & h_{42} & h_{43} \\ \sigma_N & 0 & 0 \\ 0 & \sigma_N & 0 \\ 0 & 0 & \sigma_N \end{pmatrix} \tag{5.28}$$

Beside the given examples many more algorithms were optimized, implemented on the test-bed and evaluated towards numerical stability and speed [Sch04]. Fig. 5.1 gives a short overview including QR and QL decomposition.

MIMO-Detection Schemes

- linear
  - **ZF**
    - #transmitter = #receiver
      - **Inverse (I)**
        - ***LU-Decomposition (LUD)***
          - *Crout*
          - *Doolittle*
          - *Gauss-Algorithm*
        - ***Inverse (I)***
          - *Gauss-Jordan (GJ)*
          - *LU-Decomposition + Forward-and Backsubstitution*
          - *Gauss-Algorithm. + Backsubstitution*
    - #transmitter ≤ #receiver
      - **Pseudoinverse (PI)**
        - *Moore-Penrose (MP)*
          - *Gauss-Jorden for symmetric. positive definite matrices (GJsym) + matrix multiplication. (symmetric) + matrix multiplication*
          - *Choleski-Decomposition + Forward-and Backsubstitution + matrix multiplication. (symmetric) + matrix multiplication*
        - *Greville*
  - **MMSE**
    - #transmitter ≤ #receiver
      - **Pseudoinverse (PI)**
        - *Moore-Penrose (MP), see above*
        - *with QR-Decomposition (QRD)*
          - *Gram-Schmidt-QRD + matrix multiplication (triangular matrix)*
- non linear
  - **SIC**
    - ZF
      - ***QR-Decomposition (QRD)***
        - *Householder (Ho)*
        - *Gram-Schmidt (GS)*
    - MMSE
      - ***QR-Decomposition (QRD)***
        - *Householder (Ho)*
        - *Gram-Schmidt (GS)*
  - **V-BLAST**
    - ZF
      - ***with Pseudoinverse (PI)***
        - *Moore-Penrose (MP), see above*
      - ***with QR-Decomposition (QRD)***
        - *Householder-QRD + Inverse (triangular matrix)*
    - MMSE
      - ***with Pseudoinverse (PI)***
        - *Moore-Penrose (MP), see above*
      - ***with QR-Decomposition (QRD)***
        - *Gram-Schmidt-QRD + Inverse (triangular matrix)*

Figure 5.1: Algorithms and detection schemes implemented on a TI6713 DSP.

5.2.4 Complexity and Performance Analysis

To evaluate and compare algorithms we have to characterize the complexity or the computationally required effort. Very often the measure is given in flops (floating point operations), where the definitions are varying among different authors. Instead we will compare all algorithms by the amount of required multiplications. Since additions mostly occur in pairs with multiplications we only have to count the latter.

Reciprocal values ($1/\mathbf{X}$), square roots ($\sqrt{\mathbf{X}}$) and reciprocal square roots ($1/\sqrt{\mathbf{X}}$) are counted separately, since their computation needs more cycles on the DSP. In the algorithmic optimization process the minimization of those operations has a high priority. Unavoidable divisions will always be replaced by reciprocal values. All algorithms are used on matrices of size $m \times n$ and

$$mn^3 + n^2| \quad n \quad 1/\mathbf{X}, n \quad 1/\sqrt{\mathbf{X}} \tag{5.29}$$

denotes an algorithms consisting from $mn^3 + n^2$ multiplications (additions), n reciprocal values and n reciprocal roots. In the table 5.1 [Sch04] the complexity of several algorithms is summarized.

Algorithm	**Multiplications (Additions)**	$1/\mathbf{X}$	$\sqrt{\mathbf{X}}$	$1/\sqrt{\mathbf{X}}$
ZF-I-LUD	$\frac{1}{3}n^3 - \frac{1}{3}n$	n		
ZF-I-GJ	$n^3 - n$	n		
ZF-PI-Greville	$\frac{3}{2}mn^2 + \frac{1}{2}mn$	n		
ZF/MMSE-PI-MP	$\frac{3}{2}mn^2 + \frac{1}{2}n^3 + \frac{1}{2}mn + \frac{1}{2}n^2 - n$	n		
MMSE-PI-QRD-GS	$\frac{3}{2}mn^2 + \frac{1}{3}n^3 + \frac{3}{2}mn + \frac{7}{6}n$			n
ZF-SIC-QRD-Ho	$mn^2 - \frac{1}{3}n^3 + mn + \frac{1}{3}n$	n	n	
ZF-SIC-QRD-GS	$mn^2 + mn$	n		n
MMSE-SIC-QRD-Ho	$mn^2 + n^2 + n$	n	n	
MMSE-SIC-QRD-GS	$mn^2 + \frac{1}{3}n^3 + mn + n^2 + \frac{2}{3}n$	n		n
ZF/MMSE-VBLAST-PI	$\frac{1}{8}n^4 + mn^2 + \frac{5}{12}n^3 + mn - \frac{1}{8}n^2 - \frac{5}{12}n$	$\frac{1}{2}n^2 + \frac{1}{2}n$		
ZF-VBLAST-QRD-Ho	$3mn^2 - \frac{5}{6}n^3 + 3mn + n^2 - \frac{1}{6}n$	n	$3n$	$3n$
MMSE-VBLAST-QRD-GS	$2mn^2 + \frac{3}{2}n^3 + 2mn + 3n^2 + \frac{3}{2}n$	n	n	$2n$
MMSE-VBLAST-QRD opt.	$\frac{3}{2}mn^2 + n^3 + \frac{1}{2}mn + \frac{7}{2}n^2 - \frac{1}{2}n$	$2n$	$2n$	$2n$

Table 5.1: Complexity of some basic MIMO algorithms, depending on the number of n Tx and m Rx antennas

Fig. 5.2 illustrates a complexity comparison of typical MIMO algorithms based on real multiplications. It is clearly to be seen that complex calculations [3] reduce the complexity significantly but can only be exploited when the I/Q-imbalance is negligible. On the other hand, real valued SIC detection offers exploitable performance gains even without I/Q-imbalance as shown in [FW03]. In graph(3) we can see that the classical V-BLAST algorithm (green triangles) based on ZF- or MMSE-matrix inversions, which is in principle an $\mathcal{O}(N^4)$ algorithm, will be outperformed by the QRD pre- and post-sort approach (red bullets) proposed by [WBR$^+$01] only for large numbers of antennas $N \geq 10$ when a complex calculation would be performed. For the real valued signal processing a comparable complexity is achieved at about 6 Tx and Rx antennas. So the computational gain is more to be seen in a sense that the post-sorting algorithm has to be run only when the detection order has to be tracked permanently e.g. with fixed rate transmission. In case of adaptive bit-loading, the detection order is only once computed for every bit-loading procedure and is then held fixed till the next bit-loading, hence most of the time QRD is sufficient for track-

[3] when a complex valued channel matrix is transferred to the real-valued equivalent, the number of rows and columns double. Matrix inversion have complexity of order $\mathcal{O}(N^3)$ where N is the number of Tx antennas. The real representation needs $2^3 \cdot n^3$ real multiplications while the complex valued inversion needs N^3 complex multiplications which equals $4 \cdot N^3$ real multiplications. Therefore, the total complexity difference is a factor of 2 which can be seen in the graphs of Fig. 5.2.

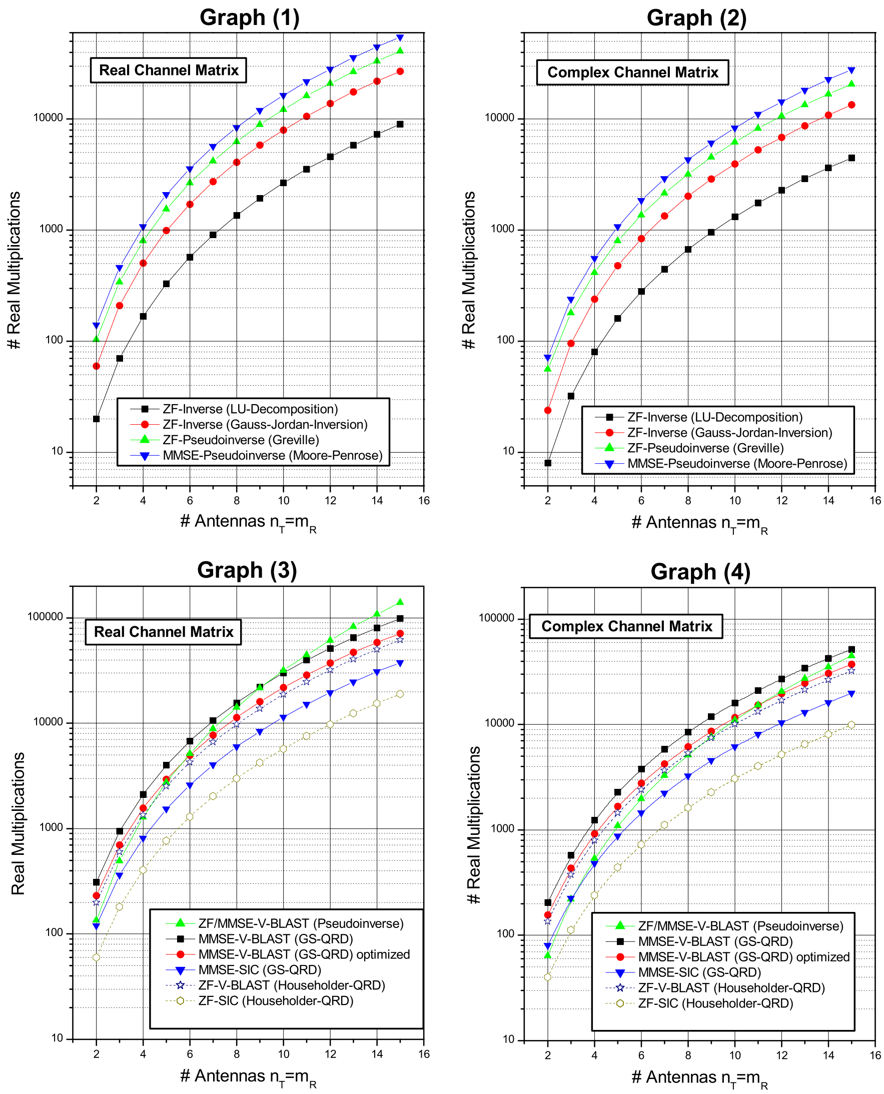

Figure 5.2: Computational complexity of several algorithms used for linear (graph 1/2) and non-linear(graph 3/4) MIMO processing. Left: matrices are real valued, Right: matrices are complex valued. All multiplications are counted as real valued multiplications.

ing the channel. Therefore the additional expenses for the V-BLAST ordering now and then are less burden to the time budget.

So by carefully counting all necessary operations a principle performance prediction with e.g. rising matrix size can be given. A implementation of the algorithms on a DSP might give different results since every dedicated DSP architecture supports some algorithmic structures better than others. Therefore the experienced programmer matches the algorithm implementation to the computational strength of a specific DSP type. Still limitations like a certain number of possible parallel assembly instructions or a limited cache size can cause that even slight changes in the code (e.g. loop length or matrix size) can change the number of required cycles significantly.

Fig. 5.3 shows the algorithm speed implemented on the TI6713 DSP for a single carrier system

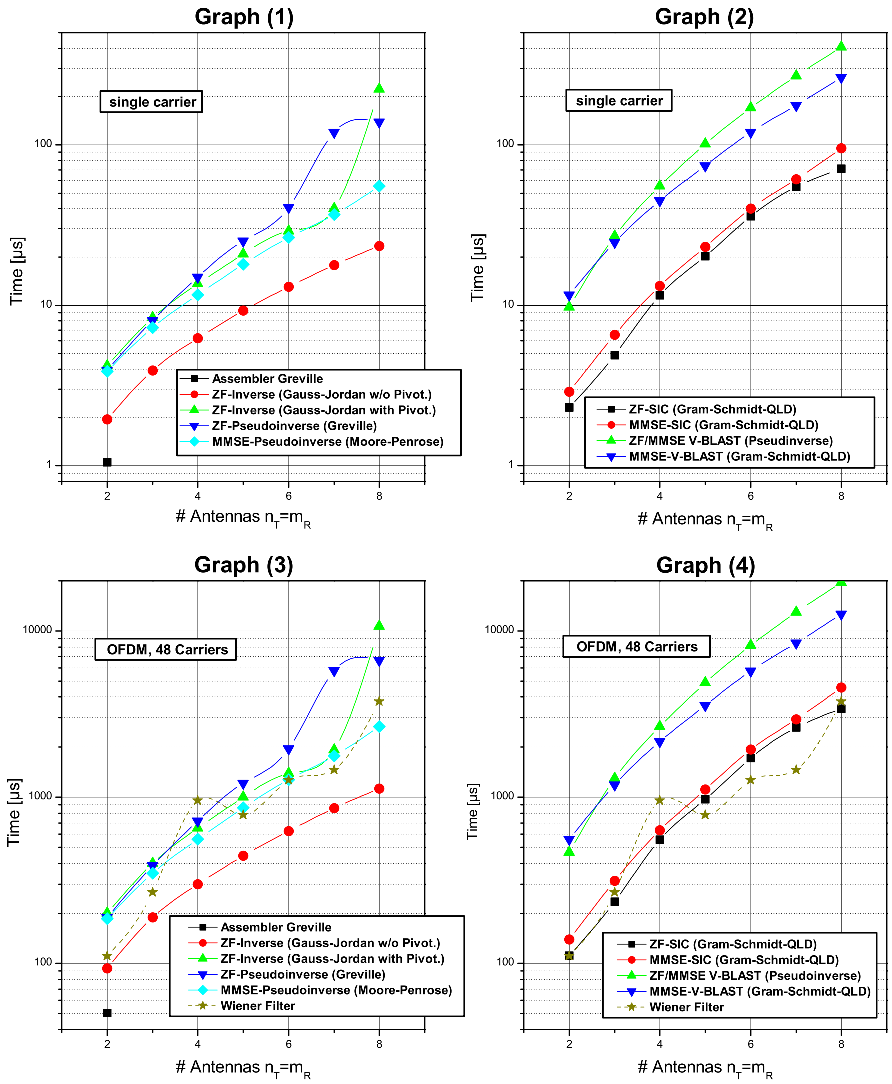

Figure 5.3: Measured cycles on TI6713 DSP displayed in μs for linear (left) and non-linear (right) MIMO algorithms. Top: Single Carrier system; bottom: OFDM system with 48 active sub-carriers.

(top) and an OFDM system where 48 sub-carriers (bottom) are active, hence 48 channel matrices have to be inverted. Several linear detection algorithms are depicted in the left figures while the right figures show the performance of some algorithms used for non-linear detection. All algorithms are performed with real valued calculation. For a 48 sub-carrier OFDM the run time exceeds the 1 ms (indoor environment) level already for small numbers of antennas ($N < 6$) even for the linear schemes. This shows that further acceleration including assembly programming, multiple DSP and / or interpolation techniques are inevitable.

The dashed curve in Fig. 5.3 (bottom) marked with stars denotes the measured performance for an implementation of a Wiener Filter as proposed in [Wie49] to smooth a noise distorted frequency response with 48 sub-carriers. Here, the measured channel estimates for each matrix element of the channel matrix are transformed into the time domain then appropriately windowed followed by a transformation back into the frequency domain. A similar computational effort is to be expected when a filter interpolation is desired to save time for the filter calculation for each sub-carrier separately. Then the calculation of the filter matrices might be dominated by the interpolation procedure instead of the MIMO algorithm, which can offer gain in computation time especially when many OFDM sub-carriers are used.

The frequency selective channel matrices $\mathbf{H}$ satisfy FIR property and can be interpolated by a polynomial of degree L, where L is the number of relevant channel taps in the time domain. Recent work from [BB04b, CBB$^+$05] proposed to exploit this property for the interpolation of MIMO filters, namely ZF and QR decomposition. In [BB04b] the ZF filter matrices were constructed using Cramer's rule and the interpolation was performed on the adjoints and determinants which also fulfill the FIR property but with higher order ($\sim L \cdot n_T$). In [CBB$^+$05] the idea was extended to QRD which is an essential part for e.g. SIC or sphere decoding. In general, the matrices $\mathbf{Q}$ and $\mathbf{R}$ are IIR due to divisions in the construction procedure. Nevertheless, there exist an invertible mapping to matrices $\tilde{\mathbf{Q}}$ and $\tilde{\mathbf{R}}$ which satisfy the FIR property (degree $\sim 2 \cdot L \cdot n_T$) and can therefore be interpolated.

There is still further investigation needed if there are other ways to perform interpolation on the MIMO filters themselves. In general, those MIMO filters (e.g. ZF, MMSE, QRD) have IIR property and can not be interpolated by a polynomial of finite degree. Nevertheless, when the channel is enhanced by e.g. receive antenna diversity (channel hardening) the bandwidth in which singularities occur is in general distrbuted over several sub-carriers. When the system is using adaptive bit-loading we only need knowledge about the sub-carriers which have noise enhancement above a certain threshold and which then are not used for transmission. Therefore, we don't need a proper detection filter for those sub-carriers, thus filter interpolation has not to be correct at those tones. It would be of great practical value to know how to choose a suitable sample grid to prevent overseeing filter singularities.

The black square in the left figures depicts the performance which was achieved with an exemplary assembly code optimization for 2 Tx and 2 Rx antennas (4×4 matrix real valued). This measurement together with an assembly design for an 8×8 real-valued matrix were used to predict the assembler performance for some MIMO algorithms. The estimated run-times (in μs) for an OFDM system with 48 sub-carriers are collected in table 5.2 from [Sch04].

	no of antennas $n_T \times m_R$						
	2×2	3×3	4×4	5×5	6×6	7×7	8×8
ZF-I-LUD	25	48	86	140	220	330	460
ZF-I-GJ	36	88	180	330	550	840	1200
ZF-PI-Gr	49	130	270	490	820	1300	1900
MMSE-PI-MP	90	160	350	640	1100	1700	2500
MMSE-PI-QRD-GS	66	170	360	660	1100	1700	2400
ZF-SIC-QRD-Ho	55	110	190	310	480	710	1000
MMSE-SIC-QRD-GS	53	130	270	490	800	1200	1800
ZF/MMSE-VBLAST-PI	86	240	540	1000	1800	3000	4700
ZF-VBLAST-QRD-Ho	170	350	620	1000	1600	2300	3300
MMSE-VBLAST-QRD-GS	140	310	600	1000	1600	2500	3500

Table 5.2: Assembler program run-time (prediction) in μs for some detection schemes (real valued) for a MIMO-OFDM system with 48 active sub-carriers

Assuming an OFDM frame length of 2 ms which is adapted to a nomadic indoor environment with small and medium sized office rooms we define 1 ms to be the critical computational time which should not be exceeded to guarantee that the next frame can be detected with a new filter based on the channel estimation in the actual frame. We can expect that for quadratic antenna configurations ZF filters with up to 8×8 antennas and MMSE- pseudo inverses up to 5×5 antenna configuration can be calculated with an optimized assembler implementation in one DSP. Non-linear detection seems to be feasible with up to 6×6 antennas without optimum ordering. If, additionally a V-BLAST ordering is required for every filter, then the matrix size is limited to a 4×4 antenna configuration.

MIMO-OFDM configurations with higher antenna numbers can be supported with one TI6713 DSP only when the channel coherence time is much longer (quasi static scenarios) or alternatively a DSP cluster must be used to partition the calculation effort sub-carrier-wise and work in parallel.

5.2.5 A Practical Bit-loading Algorithm

Channel adaptive bit-loading is a key element to achieve a reliable data throughput when CSI is available. Since, in general bit-allocation information has to be fed back to the transmitter, which costs time in itself, we need an appropriate bit-loading algorithm which calculates the bit-allocation at low computational cost. This can be ideally done when certain values e.g. $\mathbf{H}^{\dagger}$ from the channel tracking can be reused for the calculation of the bit-allocation, especially if many carriers have to be loaded with data as e.g. for a MIMO-OFDM system. In accordance with the approach from 4.4 we calculate the optimum power allocation for the maximum sum throughput. Except for the case of fully decorrelating schemes like SVD, ZF or ZF-SIC the achievable SINRs of the parallel data streams are entangled and the optimization algorithm may require several iterations to achieve the optimum power allocation. Those additional iterations can be traded against computation time, if needed.

For the following bit-loading algorithms we assume a rate control per antenna or per user/data stream rate with an individual power constraint per antenna which can be motivated by the limitation of the transmit amplifier at each antenna. We distinguish between simple bit-loading without power control and bit-loading with power control. When the latter is used optimum or close to optimum power allocation is calculated before the bit-allocation is performed.

Power allocation: The calculation of an optimum power allocation for the initialization of the bit-loading algorithm as proposed in 4.1.2 results in a uniform power allocation (UPA) for the high SNR region which perfectly fits with the individual power constraints induced by the limited amplifier power per Tx antenna. This can be motivated by the fact that at high SNR the optimum power allocation even with water-filling is UPA. Furthermore we derived in 4.1.5 that with SIC detection and individual power constraint the sum rate is always maximized by transmitted with the maximum individual power of each user.

Bit-loading for linear detection and pre-coding

For the case of a linear ZF or MMSE detector / pre-coding the achievable SNR or SINR after the MIMO-detector is given by the rows of the pseudo-inverse $\mathbf{H}^\dagger$. A comparison with look-up tables gives the number of bits which can be transmitted with this antenna under a desired BER constraint.

If one or more channels are found to be allocated with zero bits, then stepwise data streams / Tx antennas will be switched off (for simplicity in a practical implementation we choose the row of $\mathbf{H}^\dagger$ with biggest norm) and a reduced $\mathbf{H}^\dagger$ has to be calculated.

We proceed until all remaining data streams are loaded with at least one bit each.

Bit-loading with eigenvector signalling (SVD-MIMO) means the specific loading data streams instead off Tx antennas. The loading parameter is given by the inverse of the eigenvalues of $\mathbf{H}$ and additional power weights from a water-filling solution.

The obtained bit-loading vector is transmitted to the Tx via a feed-back link.

A recent implementation of channel adaptive bit-loading per antenna and per OFDM sub-carrier in a Gigabit-Test-bed allowed additional power redistribution among the OFDM tones always keeping in mind a reasonable power back-off for the OFDM crest-factor to prevent transmit signal degradation. Furthermore it was found that in combination with adaptive bit-loading more tones can be loaded with data. In practice, virtual tones at the band edges (20 % of the OFDM tones e.g. in standard 802.11) are in general not be used for data transmission which helps to reduce out-of-band interference and relax the pulse-shaping requirements. With channel adaptive bit-loading these tone can also be used for data transmission and the lower SINR due to e.g. analog filter characteristics is considered automatically in the modulation scheme.

Bit-loading for non-linear detection and pre-coding

We assume the detection or pre-coding order to be held fixed or to be calculated with the V-BLAST algorithm.

The expected noise enhancement or SINR value for each layer is given by the entries in $(\mathrm{diag}(\mathbf{L}))^{-1}$ which are already available from the QLD. A comparison with the same look-up tables used with the MMSE detector gives the number of bits which can be transmitted with this antenna under a desired BER constraint.

If one or more channels are found to be allocated with zero bits, then stepwise channels will be switched off (we choose the column of $\mathbf{H}$ which has no bit allocation and is detection earlier than the other streams), then a reduced QLD has to be calculated.

We proceed until all remaining data streams are loaded with at least one bit each.

The obtained bit-loading vector is again transmitted to the Tx via a feed-back link.

Channel measurements in static and dynamic MIMO channels [JPN$^+$02] showed that the channel has to be tracked within a time span of the order of the inverse doppler spread while the bit-loading can happen on a much slower time scale e.g. every 100 ms seemed to be sufficient in indoor scenarios. This is due to the fact that the movement of antennas themselves or of close and dominant reflectors nearby cause mainly phase variations which are compensated by the channel tracking, while the rank of the channel and the SINR of each channel remain almost unchanged.

A new bit-allocation requires a change in the SINR of a channel of some dBs, remembering that the difference between two QAM modulations steps is about 6 dB. Considering these facts, bit-loading can be performed on a much slower time scale than the channel tracking.

The finally implemented bit-loading algorithm for the transmission experiments roughly works as follows.

1. Calculate the the DFE-matrices / pseudo-inverse corresponding to the channel matrix $\mathbf{H}$.
2. Calculate the noise enhancement per data stream
   SIC:$P_i^N = 1/l_{ii}^2$ obtained from QL-decomposition of $\mathbf{H}$
   MMSE:$P_i^N = \mathbf{h}_i^\dagger \mathbf{h}_i^{\dagger H}$) ($\mathbf{h}_i^\dagger$ is the i-th row of $H^\dagger$
   SVD: $P_i^N = D_{ii}^{-2}$ obtained from SVD of $\mathbf{H}$
   ACI/JT: $P_i = \alpha^2$ given by scaling for 12 bit ADC.
3. For all streams do:

   Find max P_i^N (maximum noise enhancement).

   Decide if P_i^N is below a certain "On-Level"

   if NO: switch off modulation and cancel belonging column from $\mathbf{H}$ proceed to 1.)

   if NO and ACI/JT: find maximum row norm in $\mathbf{H}^\dagger$ and cancel belonging column from $\mathbf{H}$ proceed to 1.)

   if YES: allocate BPSK and proceed to next PAM-level if $P_i^N <$"On-Level"/4 $\rightarrow$ 4-PAM, $P_i^N <$"On-Level"/16 $\rightarrow$ 8-PAM, $P_i^N <$"On-Level"/64 $\rightarrow$ 16-PAM
4. If all streams have a bit-allocation, transmit bit-loading vector to FPGA and exit.

The parameter "On-Level" is adjustable to the targeted maximum average BER. With additional power control the BERs could be controlled more precisely but this was not implemented for these experiments since a sum power constraint was not investigated. The step size for the modulation levels is motivated by the fact that e.g. 2-PAM and 4-PAM achieve the same BER at an average SNR about 6 dB apart. Implementation penalties can be estimated based on BER measurements and this can be taken into account to fine-tune the actual step size. Those parameters are made available in a look-up table for certain BER targets.

5.2.6 A new Fair Scheduling Algorithm

The scheduling problem arises if many users (e.g. K) have to be supported exceeding the number of BS antennas. Since, pure spatial multiplexing can not solve the task a scheduling in space and time is needed in a single-carrier system [4].

The fair scheduler described in the following has knowledge about the actual channel state (channel estimator, PHY level) and required QoS e.g. average data rate of each user (DLL). We assume

[4] In multi-tone transmission schemes like OFDM the multi-user scheduler can be extended to the frequency domain (OFDM-A). For CDMA systems the code domain offers an additional degree of freedom.

linear MMSE or SIC decoding or pre-coding. We try to achieve a high sum throughput and a limited queue length for all users at the same time.

Since, the optimum scheduling policy may require time consuming calculations we proposed and implemented a very pragmatic approach, which looses little against the optimum solution.

At high SINR we choose k users with the longest queue state (e.g. $k = 1$ or 2) and add further $m_R - k$ users to form a set of users which will be supported simultaneously by spatial multiplexing. The $m_R - k$ users are chosen to achieve maximum sum rate with the k users chosen first. This approach considers the longest queues, the sum throughput given by the channel and reduces the search space significantly.

If SIC detection or non-linear pre-coding is performed the detection / pre-coding order is chosen according to (4.145) to ensure higher individual rates for users with longer queues.

Given this set of users, bit-loading is performed as described before. When the bit-allocation is calculated the bit-loading vector is transmitted to the users, including the "no-transmit" signals for all users not supported at this frequency in the same time slot.

The scheduling algorithm scales with $\sim \frac{(L-k)!}{(m_R-k)!}$ in general, which means that in reality it is applicable mainly for a reasonable small number of users exceeding the number of BS antennas. Here, a pre-selection of users can help to reduce the computational effort. For instance, some users are allocated to frequency f_1 while other users are allocated to frequency f_2. Now, the scheduling tasks can be solved independently and in parallel.

6 Real-Time Multi-Antenna Transmission Experiments

6.1 The Real-Time MIMO Test-bed

The real-time MIMO test-bed described here was developed in the German *HyEff* project. The goal was to show the feasibility of MIMO in real-time on a single carrier link, and to speed-up the signal processing in this first step beyond the natural limits set by the temporal dispersion found in typical indoor channels. Various architectures were evaluated therefore and a promising approach is implemented and fully operational now (see Fig. 6.1). This prototype has been shown with real-time transmission experiments at the Globecom conference in San Francisco in December 2003.

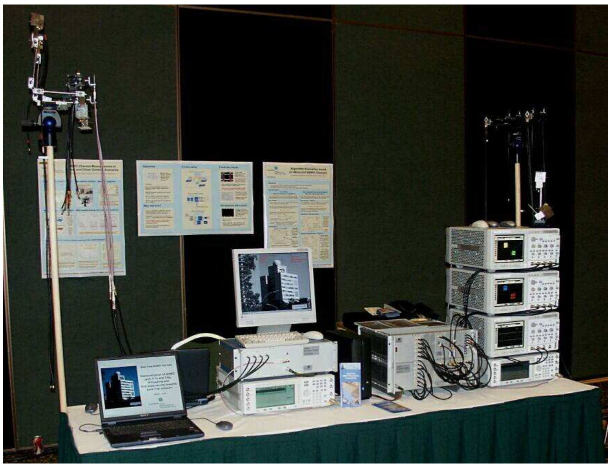

Figure 6.1: Real-time MIMO test-bed at a presentation at Globecom 2003.

6.1.1 General Concept of the Multi-Antenna Test-bed

To exploit the multiplexing and diversity potential of multi-antenna systems a higher effort of base band signal processing is a prerequisite. To match those signal processing requirements a hybrid design was chosen for the test-bed. The main base band signal processing units consist of a FPGA for very fast matrix vector multiplications and a DSP for a flexible implementation of more sophisticated algorithms. This base band design concept unites real-time high data rate capability and a high flexibility regarding the detection and pre-coding algorithms under investigation. The D/A and A/D converters use duplex mode and are integrated on a special board which is plugged on to the FPGA board.

The RF frontend uses direct up- and down-conversion (DUC/DDC) and uses a center frequency of 5.2 GHz for the local oscillator (LO). Further details are given in the following paragraphs.

6.1.2 Description of the Transmitter and Receiver

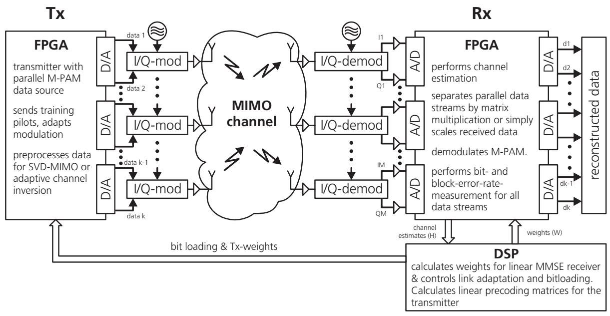

Figure 6.2: Principle of the real-time MIMO test-bed.

Transmitter: In the setup under investigation we use four transmit antennas. The 5.2 GHz radio hardware has a bandwidth of roughly 100 MHz and it performs direct analog up-conversion using four I/Q mixers each followed by +20 dBm power amplifier (ZRON-8G, Mini Circuits). The BB/RF transmit chain is depicted in Fig. 6.3.

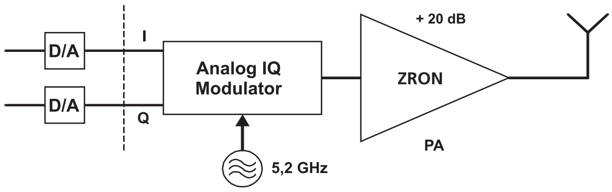

Figure 6.3: Base band to RF transmitter chain.

Up to four independent complex valued data streams may so be transmitted over the air. The data generation and the modulation are realized within a Xilinx Virtex 2 FPGA with 8 million gates. The output signals are DA converted with 12-bit resolution and used to modulate the carrier. The reason to use FPGAs instead of DSPs is the need to process multiple data streams in one single unit, particularly at the Rx. The limited number of in- and output ports of current DSPs may not allow multiple high data rate streams in parallel. Due to the FPGA realization, all the signal processing must be carefully programmed in VHDL to allow a proper timing control. The periodically transmitted signal consists of a pre-amble and a data block. Each I and Q branch of the Tx antennas is tagged with a different 127-bit Gold sequence transmitted in BPSK format in the preamble. The length of the pilots is intentionally oversized in the experimental system to get precise channel estimates. The pilots are followed by a pseudo-random data block with 1024 symbols on each stream. The modulation of the data is independently set on each I and Q branch with up to 16 PAM levels. The modulation is individually controlled via a binary vector, where the branches may be switched off if needed.

Receiver: The received signals from 5 antennas are directly down-converted using analog I/Q demodulators and digitized using 12-bit AD converters. The BB/RF transmit chain is depicted in Fig. 6.4.

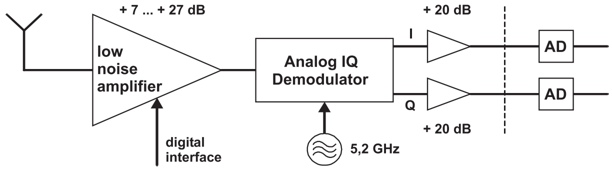

Figure 6.4: RF to Base band receive chain.

The analog design creates a severe I/Q imbalance which has to be taken into account in the entire system concept. In principle, we treat the complex-valued MIMO base-band system with 4 Txs and 5 Rxs as a real-valued system having 8 Txs and 10 Rxs. This strategy can be omitted in case of digital up- and down-conversion.

Alternatively, I/Q imbalance can be compensated at each transmit and receive antenna after a careful calibration is done. This is of ever greater importance for OFDM schemes due to additional cross talk between the image frequencies. For the SISO OFDM case [LT04] proposed an automatic estimation of the IQ-imbalances but this concept is not applicable straight forward for multiple antennas. Therefore our concept of real valued data separation can be used here as well but now the symbols on sub-carrier f_i have to be reconstructed together with the symbols from sub-carrier $-f_i$ [Yla03] which expands the detector matrix e.g. MMSE filter by a factor of 4. For a MIMO-OFDM system with 4 Tx and 5 Rx antennas this means a real valued matrix with $2(2n_T) \times 2(2m_R) = 320$ entries has to be computed and processed in real-time with the received data vector. In case that the number of multipliers in the FPGA is limited, then an I/Q pre-equalization at the Tx antennas and an I/Q equalization at the Rx antennas is a reasonable alternative, but careful calibration is needed in advance. For a small base band signal bandwidth digital up- and down-conversion is another favorable option.

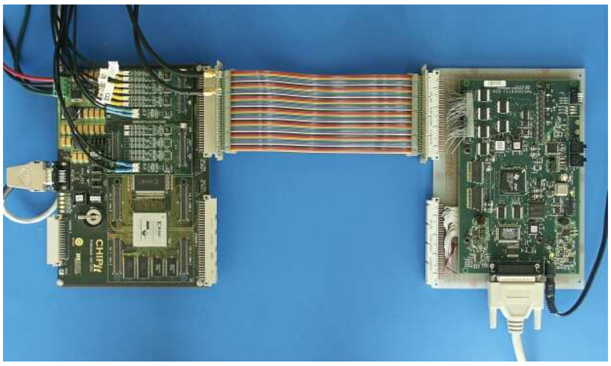

Figure 6.5: MIMO signal processing at the receiver. The A/D converters are directly connected to the FPGA hidden behind the A/D boards (left). A DSP (right) is connected to the FPGA via a parallel bus.

6.1.3 FPGAs - for High Speed Parallel Signal Processing

Channel estimation: In the Rx FPGA (8 Mio. gates), 80 correlation circuits (CC) are implemented using the known training sequences. Since binary pilot sequences are used, the CCs need no multiplication. The next bit in the sequence may eventually change the sign of the signal to be accumulated and then the CC switches from addition and subtraction. Additional CCs based on unused sequences are used to estimate the noise variance of each receive branch. The new channel estimates are immediately available after the last bit in the training sequence and stored in dedicated registers. These registers are read-out by a separate DSP (Texas Instruments 6713) connected to the FPGA via a parallel bus (24 bit flat ribbon cable). The DSP is used to calculate the coefficients of e.g. a linear MMSE filter which are then sent back to the dedicated weight registers in the FPGA via the same link. The read and write operations of the DSP are fully asynchronous to the transmitted frame structure, enabled by back-up register pages.

MIMO detection : Two linear detection schemes, ZF and MMSE are implemented in the Rx-FPGA as a matrix-vector multiplication unit to separate the spatially multiplexed data streams. Note that for a 4×5 MIMO system this unit consumes 80 dedicated multipliers, which sets an upper limit to the numbers of antennas depending on the FPGA size (Virtex II, Virtex II Pro 70/100 etc.). If a matrix $\times$ vector multiplication of bigger size has to be performed, then e.g. a row-wise multiplication of $\mathbf{H}^{\dagger} \cdot \mathbf{y}$ can help to overcome the limited number of multiplier units.

For non-linear detection like SIC and V-BLAST a decision feed-back equalizer (DFE) structure was implemented. The feed-forward matrix **GF** uses the same matrix block as for the linear equalization and after symbol decision the decided symbols are feed-back by a multiplication with a triangular feed-back matrix $\mathbf{B} - \mathbf{I}$. The DFE design was implemented such that for the detection of one symbol vector the DFE loop is passed several times until the last element of the symbol vector is detected. With 8 real valued data streams the maximum symbol rate of this DFE design is limited to 1 MSymbol per second, due to 25 MHz FPGA system clock. A way out to support higher symbol rates the DFE detection unit can be run at a higher system clock rate (100-150 MHz) and the structure can be set up in parallel at the cost of more multiplication units.

The DFE design enabled a fair comparison of several detection schemes by simply loading different solution into the matrices e.g. for ZF and MMSE the feed-back matrix $\mathbf{B} - \mathbf{I}$ is loaded with zeros.

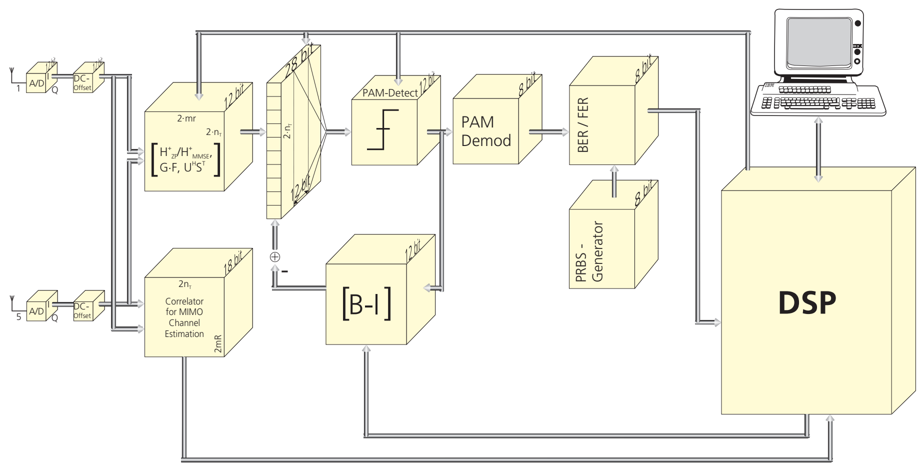

Figure 6.6: Block diagram of DFE structure inside the Rx-FPGA including channel estimation (correlation unit), MIMO detector (DFE), PAM demodulator and a BER/FER unit.

MIMO pre-coding: Several MIMO transmission schemes like SVD-MIMO or Joint Transmission / Linear Channel Inversion require spatial pre-coding at the transmitter. The spatial pre-coding was implemented in the Tx-FPGA after the parallel PAM modulation block with a matrix multiplication unit similar to that from the Rx but using only 64 dedicated multipliers. The matrix entries are calculated by the DSP as well and loaded via the 24 wide DSP-FPGA parallel bus.

Demodulation: The separated streams are demodulated using hard decisions in each I- and Q-branch. The eye pattern in a single I-branch after multiplication of the receive vector with the weight matrix is shown in Fig. 6.7. The symbol rate was 5 Mio. symbol vectors per second. The jitter-like artefact of 40 ns width is caused by unsynchronized DA conversion with a sample rate of 25 MHz.

The decoupled data streams are shown in the right of Fig. 6.7. The upper two streams were allocated to BPSK and the 3rd and the 4th data stream carry 4-PAM. The belonging I/Q constellation diagrams are depicted in the middle.

The temporal dispersion in the multi-path indoor channel obviously sets the upper limit to the maximal symbol rate, which refers to an overall data rate of 40 Mbit/s with QPSK and 120 Mbit/s with 64-QAM modulation on all four Tx antennas (8 bit/s/Hz and 24 bit/s/Hz). The channel impulse response can be seen in the spikes of Fig. 6.7 when changing from symbol to symbol.

The signal processing itself could support even higher rates and more complex schemes like e.g. MIMO-OFDM can be implemented on the reconfigurable signal processing platform as well.

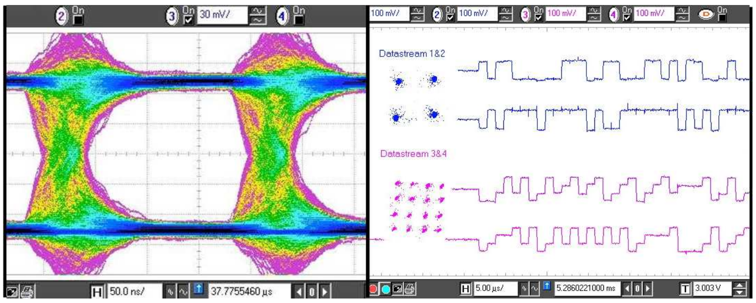

Figure 6.7: Left: Eye pattern after the separation of streams with 3 Tx and 4 Rx (QPSK modulation on all streams, 30 Mbit/s)). Middle: Complex signal constellations. Right: Reconstructed streams at the Rx.

Bit Error Measurements are performed automatically on all data streams based on a comparison of the separated and demodulated signals at the Rx and the data coming from the PRBS-data generator also programmed inside the Rx-FPGA. The error measurement is performed on bit and frame level as well. Finally, the DSP reads out the bit and frame error rates (BER and FER, respectively) and displays them on an external PC monitor.

Synchronization between Tx and Rx was realized by two cables, one for the symbol clock and one for the frame clock. The synchronization between the two clock signals was roughly adjusted by comparing the signal edges at an oscilloscope and then fine tuned inside the FPGA with help of adjustable delays. Since the channel impulse response causes spikes with exponential decay when changing from symbol to symbol the symbol are sampled at about 70% to 80% of its length. By this adjustment a reliable channel measurement could be achieved up to symbol rates of close to 10 Msymbols/s.

Synchronization will be performed over the air interface in future but was not the central part in this very basic transmission setup.

6.1.4 DSPs - Exploiting Flexibility

Channel tracking: With respect to higher mobility, it becomes critical in general to track the MIMO channel sufficiently fast. The most challenging part becomes the weight calculation when there are a few dozens of carriers and for each of them a weight matrix has to be calculated. Appropriate algorithms for the implementation on a DSP are discussed in section 5.2.4. If those weights are available within one or a few milliseconds channel tracking is expected to be fast enough for indoor and pedestrian applications.

Bit-loading or Rate Control is calculated at the Rx. The DSP calculates the actual possible PAM constellation based on the expected noise enhancement after the decorrelation. This is equivalent to the SINR in front of the demodulator. Here, different noise enhancement in I and Q are caused by I/Q imbalances. Therefore, we control the modulation independently for the I- and Q-part of each symbol by using PAM instead of M-QAM. This higher channel adaptivity translates directly into a higher throughput and link reliability.

Feed-back link: Based on the channel estimates, the DSP may calculate the optimal modulation in each stream. Note, that at the time of the experiments described in this thesis the test-bed was operating in simplex mode. So the loading vector was sent back to the Tx FPGA via a parallel bus, thus realizing an ideal feed-back link.

6.1.5 Transmit and Receive Configurations

Thanks to the reconfigurability of the test-bed we can run a wide range of transmission schemes on the platform, we simply calculate different solutions for the transmit pre-coding or/and the receive decoding in the DSP and load the matrices to the Tx- and the Rx-FPGA. So, the flexible algorithmic part is performed by the DSP while the FPGAs simple do always the same straight forward matrix $\times$ vector multiplications with the actually loaded solutions from the DSP.

To bring more transparency into all possible transmit and receive configurations the following table will help. Tab. 6.1 has to be read in the following way. The first column gives the transmission scheme under investigation and the belonging up-link (UP) or down-link (DL) scenario where it can be applied to. The next two columns contain the matrices which are loaded into the Tx- and the Rx-FPGA. The column modulation contains the modulation levels which are assigned e.g. per antenna, per data stream etc. The last column contains the parameter for the bit-loading which is specific for all schemes. This parameter represents the expected noise enhancement or SINR in front of the decision unit which is used for the bit-allocation. The scaling parameter α used for the Adaptive Channel Inversion (ACI) is necessary to limit the transmitted signals to the 12-bit ADC range.

Transmission-Scheme	Transmit-Processing	Receive-Processing	Modulation Alphabet	Bit-loading Parameter
PARC (UL)	$\mathbf{I}$	ZF / MMSE: $\mathbf{H}^\dagger$ SIC: $\mathbf{GF}, \mathbf{B}-\mathbf{I}$	Mod per Antenna 0-/2-/4-/8-/16-PAM	$\text{diag}(\mathbf{H}^\dagger \cdot \mathbf{H}^{\dagger T})$ $(\text{diag}(\mathbf{L}))^{-2}$ (QLD)
SVD-MIMO (UL/DL)	$\mathbf{V}$	$\mathbf{U}^T \cdot \mathbf{D}^{-1}$	Mod per data stream 0-/2-/4-/8-/16-PAM	$\text{diag}(D^{-2})$ (SVD)
ACI / JT (DL)	$\mathbf{H}^\dagger/\alpha$	$\alpha \cdot \mathbf{I}$	same Mod for all active streams 0-/2-/4-/8-PAM	α^2 from 12-bit-DAC scaling
Multi-User-Scheduling (UL)	$\mathbf{I}$	ZF / MMSE: $\mathbf{H}^\dagger$	Mod per User 0-/2-/4-/8-PAM	$\text{diag}(\mathbf{H}^\dagger \cdot \mathbf{H}^{\dagger T})$

Table 6.1: Transmit- and Receive Configurations of the Multi-Antenna Test-bed

6.2 Spacial Multiplexing and Antenna Diversity

We remember the definitions of the antenna diversity gain d from (3.45) and spatial multiplexing gain r from (3.44) and replace SNR$\rightarrow \infty$ with high SNR for practical reasons. In the following subsections we investigate the achieved antenna diversity read from the measured BERs and the sum throughput and the spatial multiplexing gain obtained from the experimental data in the high SNR region.

6.2.1 Measurements on Antenna Diversity

All measurements were done in our lab with the dimensions HxLxW: 3m $\times$ 7m $\times$ 5m. The Tx antennas were mounted on a pole about the middle of the room 1m below the ceiling. The Rx antennas were mounted on a tripod which can be driven along a rail across the room.

The ***MIMO Antennas*** used throughout all experiments were self made triple antennas as proposed in [JPN$^+$02] and depicted in Fig. 6.8(left). We used 4 Tx antennas and 4 or 5 Rx antennas chosen from the 4 antenna triples at the Tx and Rx.

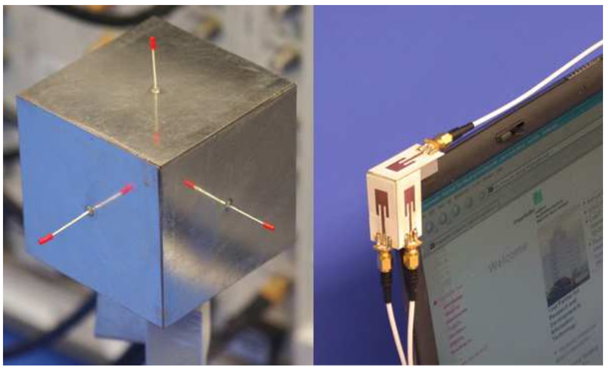

Figure 6.8: Left: close-up of MIMO Triple antenna for 5.2 GHz used in the experiments, right: triple patch antenna suitable for device corners.

The receive antennas were mounted on a tripod and driven by a little electric motor thus enabling varying speed and reproducible channel statistics with high accuracy. The trek was 5 meters long and the min/max distance between Tx and Rx was 1.5 m and 3 m respectively.

Channel Statistics: We performed channel measurements over the whole length of the trek across the room. About $2 \cdot 10^3$ channel realizations were monitored and file logged. The 10×8 channel matrices were normalized to unit channel gain and decomposed by singular value decomposition (SVD).

Fig. 6.9 shows the singular value (SV) distribution. The distribution (right) clearly reflects the I/Q-imbalance discussed in section 6.1. The 8 SVs of the real-valued channel matrix $\mathbf{H}$ should be pairwise degenerated, e.g. SV 1 and SV 2 should be identical. Due to the I/Q imbalance this degeneracy is split up and we find 8 different SVs instead of 2×4. This effect can be modelled as an additive noisy channel estimation error $\Delta\mathbf{H}$ on the real-valued channel matrix $\mathbf{H}$. Furthermore we see the increase of the smallest SVs when we add one more antenna. This shift in the SV distribution was discussed in detail by [HJJ+01a, JWHJ01] and explains the improvement of the BER performance generally described by a rising diversity order for high SNR (see also discussion on antenna diversity in section 3.5.1).

We can conclude from these results that a sufficient channel statistics is found in the chosen environment which is in accordance to the extensive channel measurements conducted at 5.2 GHz (bandwidth 120 MHz) [JPN+02]. Now, further experiments to study the diversity gain and the effect of channel adaptive rate control could be conducted.

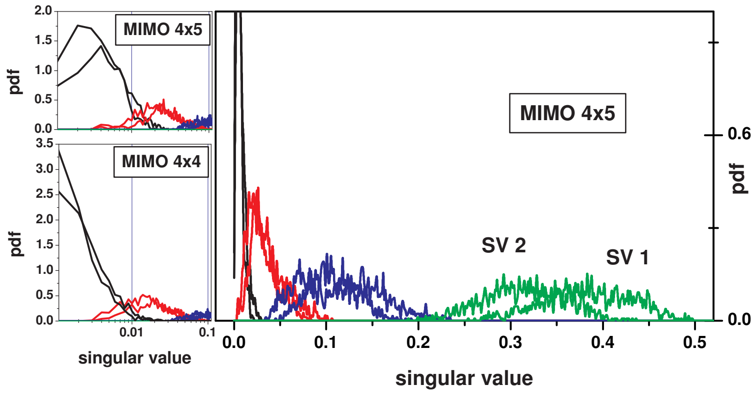

Figure 6.9: Channel statistics along 5 m trek across the lab. The distribution of the singular values is shown for a 4x5 MIMO configuration at the right. The shift of the smallest SVs can be seen in logarithmic scale on the left.

The BER Diversity Order for a MIMO system with n_T transmit and m_R receive antennas is expected to follow [TV01]

$$\frac{\log \mathrm{BER}}{\log \mathrm{SNR}} = -\beta, \text{ for SNR } \to \infty \tag{6.1}$$

where $\beta = m_R - n_T + 1$ when linear receive filters e.g. ZF or MMSE [J.G00, JJR94] are used or their non-linear extensions like VBLAST [PV02].

To check the achievable diversity order during transmission we performed BER measurements averaged over a sufficient set of channel realizations. Each BER curve was measured with a certain fixed number of Tx and Rx antennas, such that the configuration with the smallest/highest diversity were 4x4 and 2x5 antenna configurations. The Rx antenna-set is driven cross the room with a constant speed of 4 cm/s. The start and the end were marked on the floor and the antennas looked into fixed directions in all experiments. The automatic gain control was switched off at the receivers, thus contributing a constant noise figure to the receive branches. The receive SNR was then adjusted by additional attenuators at the transmit antennas (see attenuator values at x-axis). The behavior of the averaged uncoded bit error rate (BER) is depicted in Fig. 6.10. The

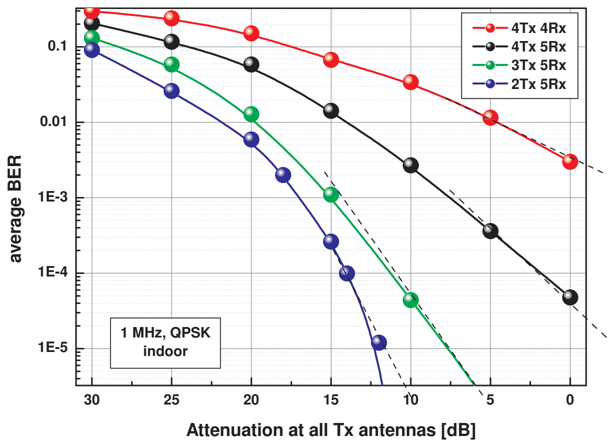

Figure 6.10: Uncoded BERs for various Tx/Rx configurations in the Lab.

slope of the curves at high SNR represents the diversity order of (6.1) which agrees very well with simulation results in i.i.d. Rayleigh and Rician channels [HJJ+01a]. During all measurements only one person was operating the system avoiding unnecessary movements about 2 m away from the Rx antennas. By this means all measurements could be reproduced very accurately. Fig.6.11

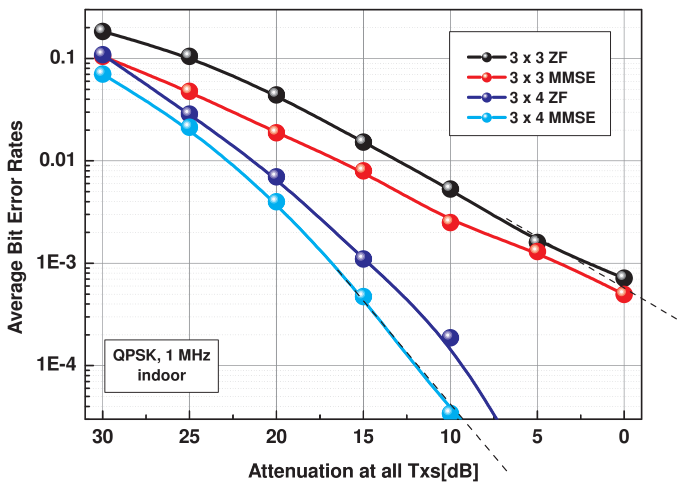

Figure 6.11: Measured uncoded BERs with ZF and MMSE detection with 3 Tx antennas and 3 or 4 Rx antennas.

shows the achievable BER performance with the two linear detection schemes ZF and MMSE. The MMSE solution, which considers the measured and therefore known receiver noise outperforms the ZF solution by 3-5 dB on the SNR scale. Algorithmically this means very little additionally effort in the DSP (30% extra time for the calculation of the MMSE solution) but the gain regarding SNR or transmit power is significantly. As we see in the figure the diversity order is the same for ZF and MMSE and follows (6.1) as expected from the theory.

Linear and non-linear detection schemes were compared in the following experiments. The Tx was transmitting QPSK modulated pilots and data streams, one from each Tx antenna and the receiver used a linear or non-linear detection unit in the FPGA.

For the linear detection schemes only the feed-forward matrix of the DFE was loaded with the ZF- or MMSE-solution while all elements of the feed-back matrix were set to zero. The non-linear detection schemes used both the feed-forward and the feed-back matrix loaded with the appropriate solutions calculated by the DSP. In this way we can directly compare different detection algorithms without changing the FPGA design which is favorable towards comparability between different experimental setups. Even a very carefully designed VHDL-code can behave differently if place-and-route constraints are slightly varied. Fig.6.12 shows the measured BER performance of the

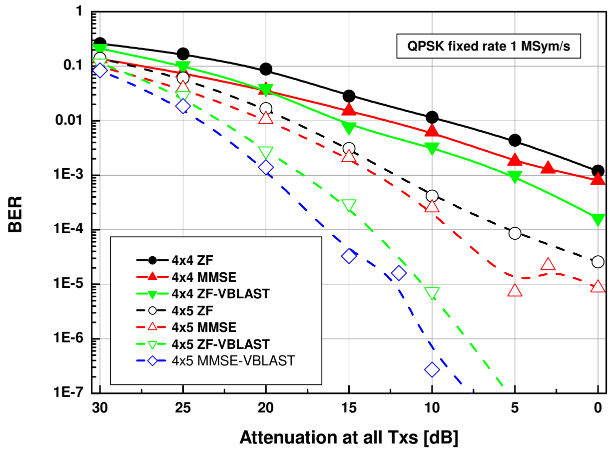

Figure 6.12: Measured uncoded BERs with ZF, MMSE, ZF-VBLAST and MMSE-VBLAST detection with 4 Tx antennas and 4 or 5 Rx antennas.

linear detection schemes ZF and MMSE and their non-linear extension of ZF-SIC and MMSE-SIC each using the VBLAST ordering which is optimum regarding BER performance at high SNR. The results show that at very low SNR the non-linear scheme performs similar like its linear counterpart since error propagation is severe in this region and the interference cancellation which is the beneficial element of the non-linear schemes can not improve the BER performance of the layers detected later.

At high SNR we see a significant improvement of the BER performance especially to be seen in a higher diversity decay. Theoretically this performance improvement increases with the number of detection layers which are included in the successive interference cancellation process and is less prominent the more diversity is already in the system. This means that in a rich multi-path environment SIC can gain most for quadratic antenna configurations ($n_T = m_R$) with a high number of detection layers.

Theory further predicts that at very high SNR the slope of the BER with SIC follows also (6.1) because here only the first detected layer causes the error and all layers detected after are error-free. The first layer has no diversity gain yet, which explains the final slope of the BER curve. This behavior can not be seen in the measurement results since this effect is expected to be evident at very high SNR or at very low BER respectively. This region was not covered by the experimental setup, because any further increase of transmit power would result in exceeding the dynamic input range either of the RF-chain or of the ADCs.

Note that in general precise measurements of very low BERs are very difficult because the bit errors are mainly caused by the fading statistics instead of the noise statistics of the receiver amplifiers. On the one hand, in practical measurements it is very hard to reproduce the exact fading behavior during each run across the room because already small changes in the room (e.g. a small change of the position of the person conducting the measurements) do not change the fading statistics in general but can cause minor changes in the depth of certain fading coefficients which can cause a slightly different number of errors. On the other hand we have only a rather limited number of transmitted bits which are about 200 - 1000 measured BER blocks containing 100 frames with 1000 symbol vectors each. The number of transmitted bits with 4 QPSK modulated data streams is then $1.6 \cdot 10^8$ or $8 \cdot 10^8$. This means that one single bit error results in a BER of $6.25 \cdot 10^{-9}$ or $1.25 \cdot 10^{-9}$, respectively. The confidence interval of the BER (error bars) reduces with the number of independent error events κ by $\sqrt{\kappa}$, this means that we need at least 100 bit errors (distributed over different data streams or/and frames, assuming Poisson distribution) to decrease the error bar down to 10 %. Keeping this in mind we have to conclude that measured BERs below 10^{-6} may have a low or limited confidence towards interpretation.

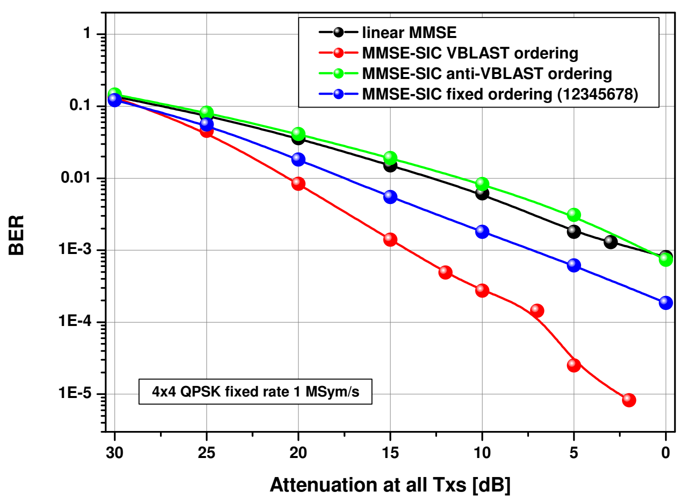

Figure 6.13: Measured uncoded BERs with MMSE-SIC detection with 4 Tx antennas and 4 Rx antennas. The detection order is varied.

Fig.6.13 and Fig.6.14 show the dependence of the BER performance on the detection order. The best order (red), VBLAST proposed by [Fos96] detects always the data stream with the best SINR

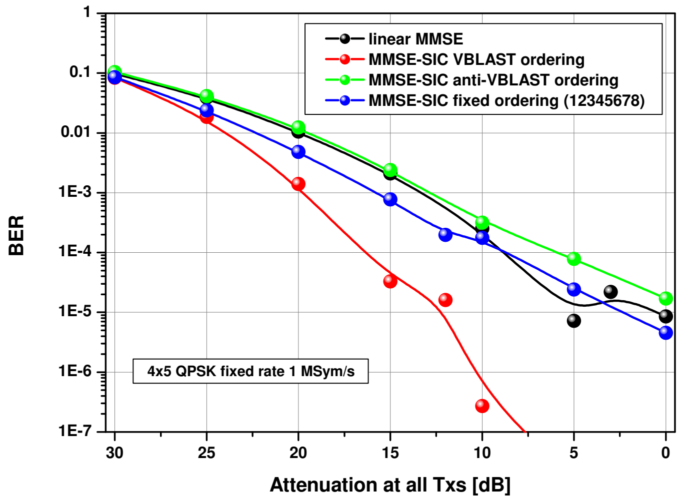

Figure 6.14: Measured uncoded BERs with MMSE-SIC detection with 4 Tx antennas and 5 Rx antennas. The detection order is varied.

in each detection step. This guarantees a minimum possible BER for each detection layer and in direct consequence also minimum error propagation.

The anti-VBLAST ordering (green), which always picks the stream with the worst SINR achieves the worst BER in each layer and the worst error propagation as well. This can result in a weaker average BER performance than the corresponding linear detection scheme as to be seen in both figures. A fixed ordering (e.g. [1,2,3,4,5,6,7,8], blue curves in both figures) will always result in a better performance than the underlying linear detection scheme itself because the first layer has the same error probability like the linear scheme and all layers detected after will profit from the stepwise diversity increase with each layer. Since the first layer will not be necessarily the worst channel all the time due to channel variations, we still find a better performance despite the additional bit errors caused by error propagation.

These experimental results with fixed modulations verify that SIC with fixed ordering or preferably random ordering is always beneficial compared to the corresponding linear scheme even if we have to face error propagation from layer to layer. An ordering which will always loose (anti-VBLAST) over a sufficient channel statistics requires channel knowledge and can be avoided easily.

In case of SIC, the detection ordering looses its importance when channel aware bit-loading is applied on top, since the BER for each layer is assured by the bit-loading algorithm. Each layer is loaded according to its achievable SINR and the achievable sum rate under sum power constraint is constant as shown in [HLB03]. Here, the detection order offers a new dimension towards controlling the individual rates of each data stream without loosing sum throughput. This can be exploited for multi-user scenarios with one transmit antenna at each terminal as shown in the experiments described in section 6.4.

6.2.2 Measurements on Channel Adaptive Rate Control

Bit-Loading Algorithm: Bit-Loading is an important means to reach a reliable channel throughput. The criterion for the allocated modulation level per data stream is based on the noise enhancement seen after the data separation using either a linear or non-linear detection scheme. This SINR criterion is used to predict the BER for a certain modulation. The algorithm described in section 4.4.2 [HLB03] calculates the optimum power allocation to maximize the sum rate under the assumption of sum power constraint and successive interference cancellation. For the experiments this algorithm was adapted to individual power constraint and various detection schemes e.g. linear MMSE or MMSE with SIC. For a detailed description of the algorithms be referred to sec. 5.2.5.

Adaptive Bit-Loading with linear detection: The impact of *adaptive bit-loading* without transmit pre-coding which is also commonly known as *per antenna rate control* (PARC) [CLH+03, HFG+04c] was investigated in the following scenario.

We chose a fixed transmit power for all Tx antennas of +12dBm $\pm 1.2dB$. Then fixed modulation (8x BPSK) and Adaptive Bit-loading (Off, BPSK, 4-PAM) are compared for a fixed amplifier gain at the Rx (AGC is switched off). The antennas are adjusted to the antenna post in random directions which are not changed over the whole experiment. At the receiver we use either 5 or 4 Rx antennas. Each experiment is conducted along the same trek, over the same distance and with the same speed of 4 cm/s. The transmit power was adjusted with variable attenuators.

Adaptive bit-loading is performed every 100 ms which appears to be sufficient at the given speed. The FPGA counts the transmitted bits per data stream, bit errors and frame errors over a set of 100 frames with 1000 data symbols each (pilot symbols of 128 bits are omitted for the BER calculation). The results are stored and transferred to the DSP and file-logged on the hard drive. Since we average over 100 frames the minimum temporal resolution for error bursts is about 125 ms @1 MSym/s and 25 ms @5 MSym/s.

Fig. 6.15 shows the achievable spectral efficiency versus the position along the trek using PARC. It is obvious that the spectral efficiency is increased when the Tx and Rx are close together (about position 2500 - 3000). At the beginning of the trek channel realizations are found which do not allow a transmission of 8 bits/s/Hz at a BER$< 10^{-3}$, therefore data streams are switched off and the sum efficiency drops down to 4 bits/s/Hz. If 5 Rx antennas are used the average spectral efficiency rises from 11.0 bits/s/Hz to 13.2 bits/s/Hz while the constant QPSK modulation system transmits only 8 bits/s/Hz (dashed line). The inset shows the empirical cumulative density function of the spectral efficiency of the scenarios. The highest achievable payload data rate was 65.4 Mbits/s with a symbol rate of 5 MSym/s.

Fig. 6.16 shows BERs along the measurement trek for the 4x4 MIMO system using BPSK only (top) and PARC (bottom). In the upper figure we clearly see error-free sections around position 2500 where the Tx and the Rx are very close together. Around position 200 (maximum distance) we find the channel in a rather singular state causing very high BERs over an extended time. This effect can be mitigated by PARC by simply switching one or two data streams off resulting in a diversity gain which improves the BER immediately. If the BER and FER is averaged over the whole scenario we reach a very comparable average BER of $5.1 * 10^{-3}$ for PARC and $6.5 * 10^{-3}$

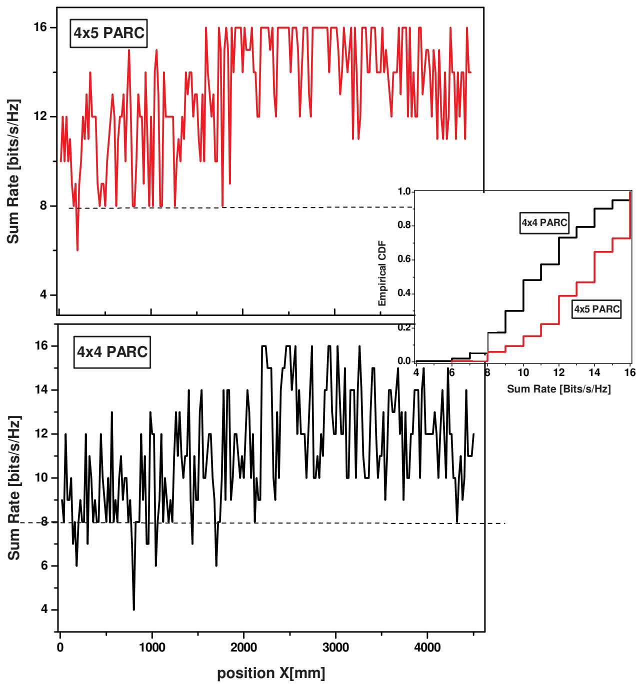

Figure 6.15: Achieved spectral efficiency with PARC and antenna configurations 4x4 or 4x5. Modulation levels: Off, 2-/4-PAM

Conf.	4x4	4x4	4x5	4x5
Mod.	BPSK	PARC	BPSK	PARC
bits/s/Hz	8	11.01	8	13.22
BER	$6.5 * 10^{-3}$	$5.1 * 10^{-3}$	$3.2 * 10^{-5}$	$2.5 * 10^{-4}$
FER	0.094	0.124	0.047	0.056

Table 6.2: Achieved rates and BERs with various antenna configurations and fixed rate (QPSK) or PARC with 0-/2-/4-PAM

for BPSK only. The measurement was repeated with 5 Rxs and the results are summarized in Tab. 6.2.

The FER seems to be significantly higher than the BER due to the frame length of 1000 symbols. A shorter frame length translates directly into a lower FER.

Fig. 6.17 gives more insight into the fact that the BER of the 4x4 case with PARC is slightly better than with BPSK only. If the cdf of the BERs is used we see the curves crossing (marked with circle). Here, the effect of extended exposure to bad channels dominates the BER of the 4x4 BPSK system and error free sections can not compensate the accumulated errors. The PARC scheme avoids high data rate transmission over bad channels and can profit from this strategy. The 4x5 case shows clearly that the BPSK system profits significantly from the diversity gain while the PARC system benefits as well but if a certain SNR level is reached the modulation scheme is switched to 4-PAM to trade better BER performance against higher throughput.

The experimental results show that already adaptation with few modulation levels (e.g. OFF,

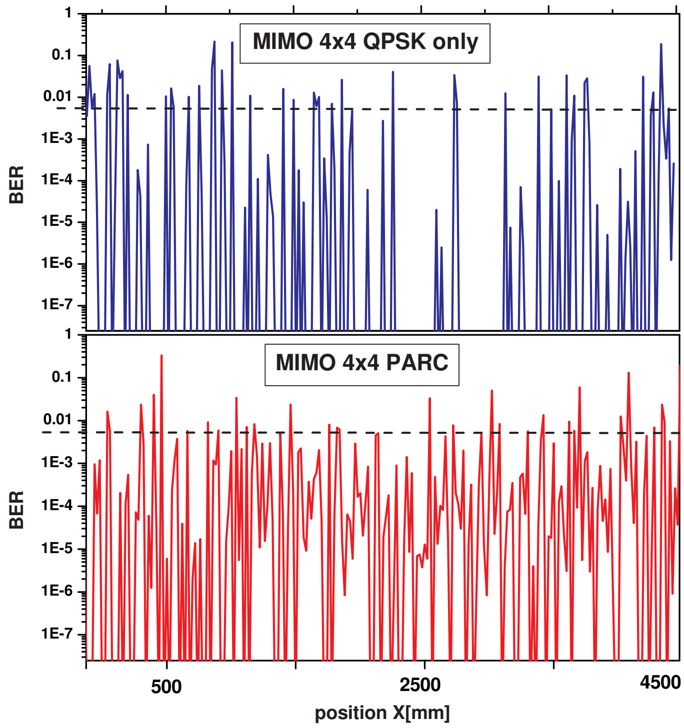

Figure 6.16: Uncoded BERs with and w/o PARC(modulation levels: Off, 2-/4-PAM),average BER (dotted)

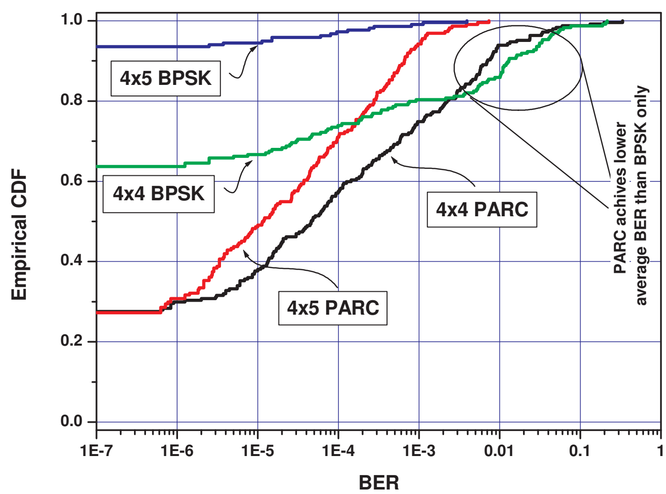

Figure 6.17: Empirical CDF of BERs with and w/o PARC.

BPSK, 4-PAM) can achieve a significant improvement towards sum throughput while the BER is kept comparable to the fixed rate transmission scheme. A further increase of the sum throughput is realized when higher modulation levels (8-PAM, 16-PAM) are included in the bit-loading algorithm which was measured in later experiments where channel adaptive bit-loading was combined with

ACI/JT (Fig.6.23) or MMSE, VBLAST and SVD (Fig.6.19).

The initial transmission experiments with adaptive per antenna rate control (OFF, BPSK, 4-PAM) already achieved peak data rates up to 80 Mbits/s and a peak spectral efficiency of 16 bits/s/Hz. The measurement data show that in stationary scenarios or with low mobility extensive periods of rank degradation of the channel can occur which may produce heavy error bursts when full spatial multiplexing is performed with fixed modulation. This effect can be mitigated by adapting the allocated modulation per Tx antenna / user according to the actual channel quality or by exploiting a significant number of diversity antennas at the receiver which cause a kind of channel "hardening". Exploiting adaptive bit-loading with 4 Tx and 5 Rx antennas we measured an average throughput improvement of about 65%.

Furthermore we find that the effects of imbalanced I/Q caused by analog direct up- and down-conversion are well compensated by processing the channel as real-valued. In consequence, rate adaptation has to be performed independently for I and Q (see left in Fig.6.18) to guarantee certain BER targets in each branch and as a side effect we have bit-wise steps for the sum rate instead of 2-bit steps as with complex QAM signal constellations.

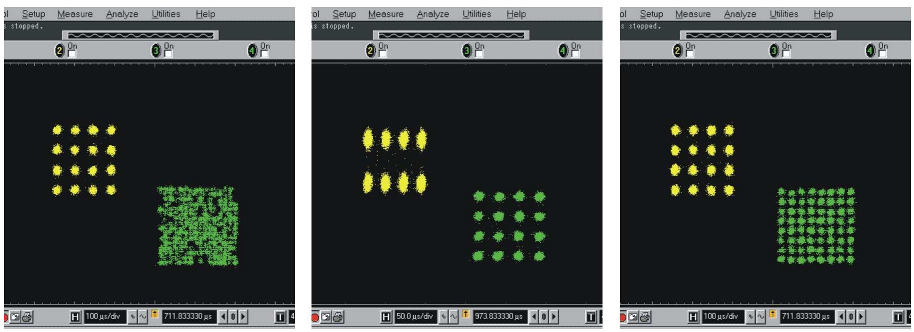

Figure 6.18: Screen shots of reconstructed data symbols from Tx 1 (yellow) and Tx 2 (green). Modulations were 2-16 PAM, left: 16-QAM and 256-QAM as highest modulation level. middle: I/Q imperfections can require different modulations in I and Q. right: 16-QAM and 64-QAM. MIMO transmission 2×4 in narrowband channel of approx. 1 MHz in indoor scenario.

Adaptive Bit-Loading with Successive Interference Cancellation: In order to obtain more insight on how much throughput gain can be achieved in reality by applying successive interference cancellation techniques we conducted throughput measurements with linear MMSE and MMSE with SIC and furthermore with SVD-MIMO which will be described in the next paragraph. To allow a fair comparison all schemes were used with the same antenna configuration of 4 Tx and 5 Rx antennas. Furthermore the Rx antennas were moved always the same path through the room to guarantee the same channel realizations for all three schemes. The transmitter had a per antenna power constraint and the receiver used the DFE-MIMO detector implemented in the FPGA. All transmit and detection matrices as well as the bit-loading were computed by the DSP.

In contrast to the observations with the fixed rate BER measurements where significant error propagation was observed especially in rank deficient channels, we now find no such behavior any more. Error propagation is still an issue but strongly limited due to the target BER assured by the bit-loading algorithm.

Furthermore we expect that in combination with adaptive bit-loading the SIC order is of little or no importance as long as it is matched to the actual bit-allocation. The theoretical result from 4.43 predicted that we can use any SIC order without loosing sum capacity. This ***conservation law of the sum capacity*** was found to hold at least in principle also in practice for the low and mid SNR range, meaning that the bit-loading algorithm operates in the mid range of the available modulation formats. When the system is operated at the highest modulation levels we observed a decrease in throughput when not using the V-BLAST ordering. The simple explanation for this effect is to be seen in the fact that the first detected data stream has the best SNR of all streams at this detection step but could achieve a further improved SNR when detected later. In case the best stream will be allocated already the highest possible signal constellation e.g. 256-QAM when detected first, we would eventually have to allocate 512-QAM if detected later to conserve the overall sum rate which is exceeding the available QAM range. Having said this we can conclude, that the SIC ordering can be used to control individual antenna or user rates but at high SNR the V-BLAST ordering still provides the highest avarage throughput.

SVD-MIMO transmission: As discussed in section 3.2 the channel capacity with CSI at the Tx and the Rx is higher than with CSI only at one end of the link. The appropriate transmission strategy for joint signal processing at the Tx and the Rx is realized with the *SVD-MIMO transmission* or *Eigenvalue signalling*. In the SVD-MIMO mode each data stream is transmitted from all antennas and feed into one of the orthogonal sub-spaces by sending into the direction of the Eigenvectors of $\mathbf{H}$ as discussed in section 4.1.1. At the Rx the mixed signals at the Rx antennas are decoupled from the sub-spaces by a multiplication with $\mathbf{U}^T$ obtained from SVD. Now, bit-loading is not performed on a per antenna basis but on a per data stream basis and the channel quality (SNR) is determined by the reciprocal value of the eigenvalues of $\mathbf{HH}^H$.

This means that a bad conditioned MIMO channel can have one or two small Eigenvalues which allow no transmission satisfying certain BER targets. Nevertheless the other Eigenvalues are not affected by this and the transmission over their eigenspaces still works very reliable. This allows a significant higher throughput especially at low SNR.

The SVD transmission experiments use the unitary matrix $\mathbf{V}$ as preprocessing matrix at the Tx which contains the eigenvectors of the channel $\mathbf{H}$. At the Rx $\mathbf{U}^T$ is used and a rescaling with the reciprocal value of the belonging eigenvalue is performed.

The pre-coding matrix $\mathbf{V}$ is calculated by the DSP at the Rx by SVD and send to the Tx FPGA via cable thus enabling a perfect fit of the pre-coding matrix $\mathbf{V}$ and its belonging decoding counterpart $\mathbf{U}$.

When $\mathbf{V}$, $\mathbf{D}$ and $\mathbf{U}$ are calculated separately at different sides of the link based on an independent channel measurement, precautions have to be taken to assure that all transmit and receive vectors and the belonging eigenvalues fit together. Otherwise, e.g. if a sorted SVD (the order of the eigenvalues in $\mathbf{D}$ is in rising or decreasing order) is used, signals can be scaled in a wrong way or the sign can be changed. This observation was made by [LGF02, TLF03] based on numerical simulations with sorted SVD calculated independently at each side of the link, but this can not be assessed as a critical problem from our point of view.

A change in the sign can be prevented by forcing the first row in $\mathbf{V}$ to be of non-negative elements.

Furthermore we don't need a strict ordering towards the size of the eigenvalues, instead all valid eigenvalues (those which will carry data) are shifted to the left and transmission is performed over the leftmost valid eigenvalue and the belonging eigenvectors.

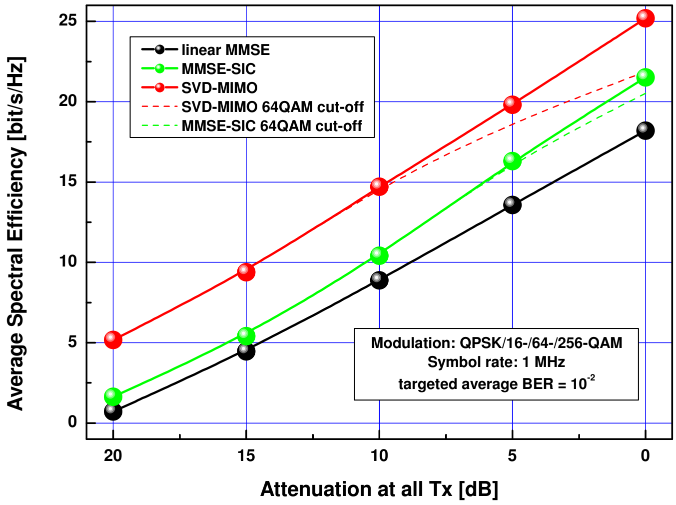

Figure 6.19: Comparison of the achieved average sum rate with 4 Tx and 5 Rx antennas with linear MMSE or MMSE-SIC and SVD-Eigenvalue transmission. Since the modulation level was loaded independently in the real and imaginary part also intermediate levels between e.g. QPSK and 16-QAM were used.

Fig. 6.19 shows the measured sum throughput with a BER$\leq 10^{-2}$ with three transmission schemes: red: SVD-MIMO, green: MMSE-VBLAST at Rx and black: linear MMSE at Rx. At very low SNR the latter two schemes achieve similar low throughput which can be explained that with both schemes most of the time only one or two data streams are switched on and SIC can not gain much. At high SNR SIC gains upto 3 bits additional throughput compared to the linear MMSE due to the SINR increase for later detected layers. The SVD scheme outperforms the other two other schemes by a higher throughput even at high SNR values. Note that from theory we would expect a similar throughput performance for SVD-MIMO and MMSE-SIC, which is known to be capacity achieving as well [VG97] (see also Fig. 3.8 in chapter 3.2). The observed difference at high SNR is to be explained by error propagation which can become significant due to symbol by symbol decisions and furthermore due to the uncoded transmission. Since we perform adaptive bit-loading in such a manner that all layers meet a certain BER target, we have to consider the effect of error propagation in the bit-loading algorithm. The weaker the BER decay (diversity slope) the more extra transmit power is necessary to fulfill the target. As an example let's assume a BER target of 10^{-3} for all layers. Since all layers including the last layer shall meet this BER target, we have to set the BER target for each layer lower such that including error propagation we will satisfy the targeted BER. Assuming 4 Tx and 5 Rx antennas and a multiplexing of 4 data streams we can expect a BER diversity order $\sim SNR^{-2}$. If we had a 100% error propagation then as a rule of thumb the last layer would suffer from 3/4 of possibly propagated errors and 1/4 of own decision errors meaning that we should set the target BER to $\frac{1}{4} \cdot 10^{-3}$. At the given diversity slope this corresponds to an SNR loss of approximately 3-4 dB, something comparable to the measurements. This SNR loss is expected to increase to about 6-8 dB with 4 Tx and 4 Rx antennas.

Generally, this means that the SNR loss against the water-filling or SVD-MIMO scheme increases with the number of layers / transmit antennas and decreases with the number of extra receive antennas / degree of receive diversity. Furthermore the correlation of the data streams influences the error propagation, e.g. orthogonal transmit channel vectors don't propagate errors from one detection layer to another. So in reality the SNR margin has to be found by averaging over a statistical ensemble of channels and can later be adopted automatically if the channel entanglement is changing in different deployments.

At low SNR SVD-MIMO achieves a tremendous relative gain compared to MMSE and MMSE-SIC. This high throughput advantage can be explained that with SVD one data stream is coupled into one eigenmode of the channel. The other two schemes couple each data stream into all eigenmodes depending on the actual channel realization, which means in average 1/4 of each data stream. At very low SNR when only one complex stream is transmitted in all schemes MMSE and SIC transmit only 1/4 of their one and only stream over the best eigenmode. In average this should result in a disadvantage of about 6 dB on the SNR scale which is roughly the measured value at low SNR.

The dashed lines in Fig. 6.19 show the behavior when the maximum modulation level is limited to 8-PAM or 64-QAM, respectively. The maximum rate saturates already within our measurement range and shows that the achievable maximum slope for the average throughput which means maximum achieved spatial multiplexing gain is determined by limited modulation levels. With a M-ary QAM level of 1024 (if implementable in multi-antenna schemes) a smaller gap between theory and practice towards the spatial multiplexing gain might be achievable in principle.

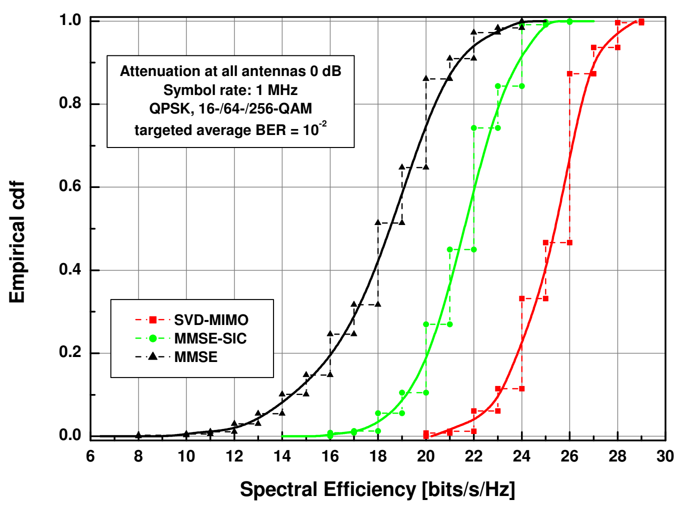

Figure 6.20: Empirical cdf of the achieved average sum rate with 4 Tx and 5 Rx antennas with linear MMSE, MMSE-SIC and SVD-Eigenvalue transmission. Attenuation at all Tx antennas = 0 dB, meaning maximum transmit power over the air.

Fig. 6.20 shows the empirical cumulative distribution function of the measured sum throughput at the highest possible SNR point. We see that the fitted curve is steepest for the SVD-MIMO and has the longest tail at low rates for the linear MMSE. This is in good accordance with capacity simulations from the measured channels. Especially at low outage probabilities the three schemes have a huge difference in throughput. Example: Outage = 0.01 MMSE: 11 bit/s/Hz, MMSE-SIC: 17 bit/s/Hz and SVD-MIMO: 21 bit/s/Hz.

6.3 Adaptive Channel Inversion (Down-link)

The main focus of the following experiments was to show that an implementation of a down-link transmission scheme with pre-coding is feasible and what can be achieved in terms of BER and throughput performance in this very first experiment. To make live easier and to overcome the reality of non-reciprocity in the baseband channel with the actually used RF-chain, we measured the channel from the BS (Tx) to the MTs (Rx) during a training sequence. Next the DSP calculates the linear pre-coding matrix (ZF or MMSE) and loads the matrix over a parallel bus into the Tx FPGA. This means that the pre-coding solution is calculated at the receive side, which would not be an option in a real communication system, since a transmission of the matrix values over a feed back channel would require a huge bandwidth and the transmission of the matrix could be erroneous due to the fading channel. Nevertheless this quite ideal experimental configuration gives us an upper performance limit of what can be achieved with pre-coding and spatial multiplexing in the broadcast channel.

To get closer to real world field applications there are still some obstacles to be taken, namely first the CSI has to be obtained by a measurement into the opposite direction (MTs to BS). This will be automatically solved when the test-bed is upgraded to TTD duplex mode. Second, we have to assure reciprocity in the base band channel, which will be solved with a reciprocal transceiver design, which was proposed only recently [JKI+04]. The reciprocal transceivers [HHI+04] are currently calibrated and the upgrade of the test-bed with reciprocal transceivers at both sides is expected to be operational in the next few months. The new test-bed will allow a direct measurement of how much performance will be lost due to the imperfections of reciprocity in the base band channel.

All measurements were done in our lab with the dimensions HxLxW: 3m × 7m × 5m. The Rx antennas were mounted on a pole about the middle of the room 1m below the ceiling. The BS/Tx antennas were mounted on a tripod which can be driven along a rail across the room.

Since we have only one Rx FPGA available at the moment, 4 Rx units were implemented in parallel and we perform channel estimation and receive signal scaling in one unit as in a single user MIMO system. Nevertheless we can treat the Rx antennas like decentralized receivers which are distributed over the lab.

In a practical transmission system the CSI will have to be acquired by a pilot transmission into the up-link and relying on the reciprocity of the channel. The ideal feedback of CSI to the Tx in our test-bed shows the potential of transmit pre-processing for the first time in an over the air experiment to our knowledge. The results can be understood as a upper performance bound due to the best achievable CSI at the Tx.

The transmitter was driven by an electric motor thus enabling varying speed and reproducible channel statistics with high accuracy. The trek was 5 meters long and the min/max distance between Tx and Rx was 1.5 m and 3 m respectively.

Channel Statistics: We performed channel measurements over the whole length of the trek across the room. About $2*10^3$ channel realizations were monitored and file logged. The measurement was done twice with maximum power and without transmitting any signal, respectively, to measure the noise ground at the ADCs for the estimation of the actual system SNR. With maximum transmit power and averaging over the whole trek we found an average measurement SNR per Rx antenna of 38 dB due to the high correlation gain from 128 pilot symbols which can be exploited for the very precise channel estimation.

Required Transmit Power: The measured channel matrices are then pseudo-inverted and all pre-coding matrices are evaluated to study the transmit power enhancement. This is in analogy to the noise enhancement per data stream which determines the SNR for a classical MIMO systems in the up-link. We use the same measure of the power enhancement factor (PEF) defined in (4.104) [HSB03] by normalizing the transmit power per antenna by the factor of the

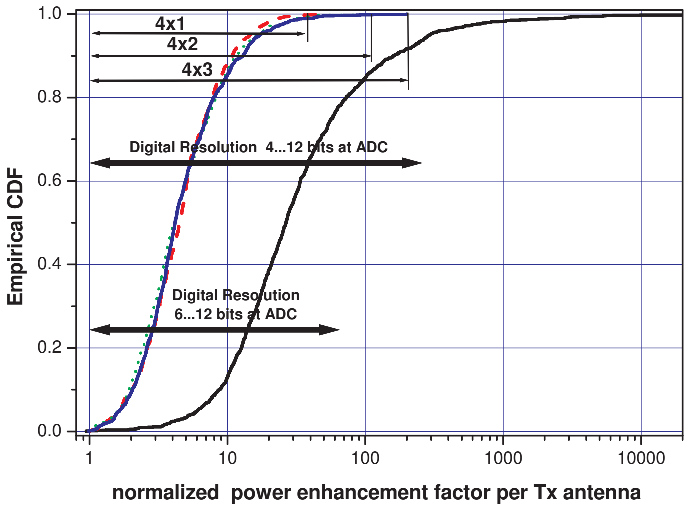

Figure 6.21: Required transmit power statistics for LCI along 5 m trek across the lab. The CDF of the power enhancement factor depending on the Tx antenna diversity (4x1...4x4) is depicted from left to right. Normalization is performed by averaging the required Tx power over the whole trek and all antennas. The arrows indicate the digital range when 4...12 or 6...12 bits resolution for the pre-coding matrix are used at the DACs.

Tx power averaged over all antennas and all channels. This measure allows to predict the system performance under the constraint of a limited dynamic range of the Tx amplifiers and of the DACs at the Tx.

Fig. 6.21 depicts the CDFs for different transmit diversity. If a 4x4 MIMO system is chosen to be run with LCI we would need DACs with a minimum of 16 bits to provide the required digital dynamics of 43 dB necessary for this curve (outage 0.1%).

Since all real-world transmission systems will have a limited transmit power range the achievable SNR is always determined by the power limit of one antenna. This limitation causes a similar behavior for the slope of the BER as we found from simulations with fixed sum transmit power. This explains that the uncoded BER curves in Fig. 6.22 for 1 to 4 data streams with LCI are

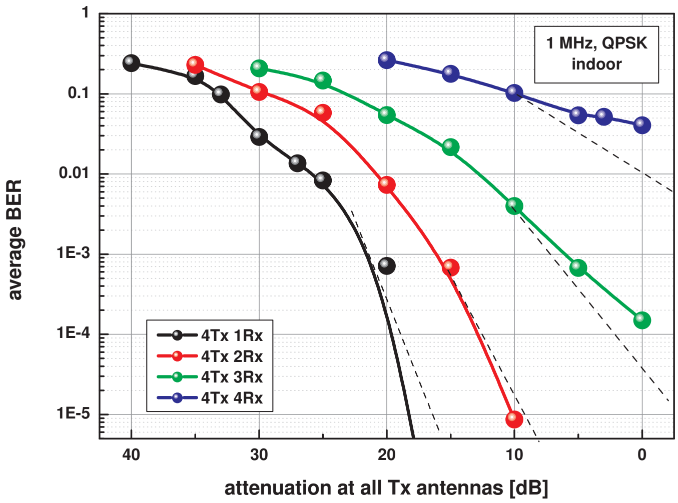

Figure 6.22: Uncoded BERs for Linear Channel Inversion with fixed numbers of 4 Tx and varying number of Rxs (1...4). Dotted lines indicate expected diversity for fixed transmit power constraint.

shifted along the x-axis (Tx antenna attenuation) and have a slope according to a $n_T - m_R + 1$ diversity gain as discussed in section 3.4. The dotted lines in Fig. 6.22 indicate the expected diversity order for high SNR. Those BER measurements are found to be in very good agreement with analytical predictions and simulation results. A higher probability of exceeding the maximum transmit power with reduced transmit diversity translates directly into a weaker slope for the BER, which gives advantage to a configuration with more transmit diversity. If we would assume only an average sum power limit then we would expect AWGN like BER curves which are only shifted on the SNR axis by 3 dB if the number of streams is doubled. We clearly see transmit diversity is the crucial element for the application of LCI/JT to avoid system degradation due to the limited dynamic range of the transmitter.

Adaptive Channel Inversion: The *channel-aware bit-loading* combines transmit diversity and common modulation level control. In Fig.6.23 we see that the adaptive bit-loading algorithm avoids overloading weak down link channels thus ensuring reliable BER performance for each transmitted data stream. This is done by controlling the number of supported users and the modulation level. By this strategy we find a spectral efficiency which ranges from 4-18 bit/s/Hz with an average of 9.6 bit/s/Hz. This means a payload data throughput of 18-80 Mbit/s (5 MSymbols/s) with an average uncoded BER below 10^{-3}.

The experimental results show that ACI is a feasible and suitable transmission scheme for the broadcast channel with decentralized receivers when there is only limited signal processing capability is available at the MTs due to only one Rx antenna per terminal. This allows for very cost efficient MTs since the SP is performed by the BS.

The crucial point is a reliable CSI at the BS which has to be acquired with a channel measurement into the opposite direction. A further requisite is the reciprocity of the base band channel which can not be taken for granted for a standard RF-chain where different amplifiers and I/Q-mixers are

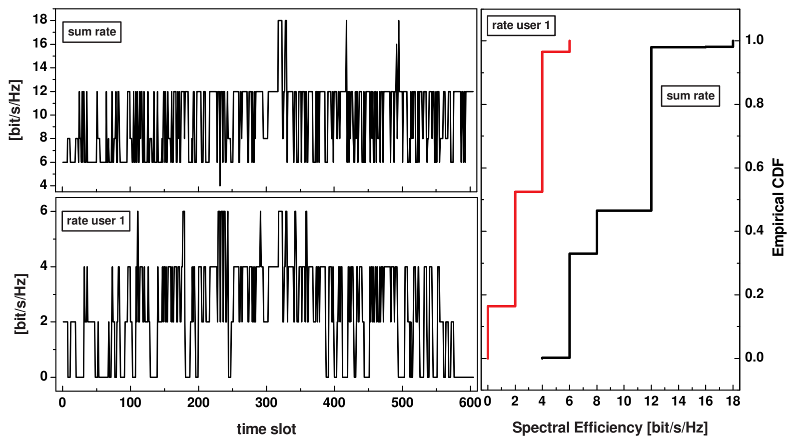

Figure 6.23: Spectral efficiency with Adaptive Channel Inversion. Top: Sum Rate, bottom: rate for user 1, right: CDFs of the spectral efficiency

used for transmission and reception. The recently proposed transmission design with reciprocal transceivers [JKI+04] allows a pre-coding for a TTD broadcast scenario especially for WLAN applications which are of great demand for hot spot applications assuring perfect downwards compatibility with MTs with only one antenna.

6.4 Multi-User SIMO Scheduling (Up-link)

The scheduling policies discussed before are implemented in the real-time demonstration test bed at HHI. We will present the experimental setup of the test bed (Fig. 6.24) which is based on a hybrid setup of FPGAs and a DSP at the BS.

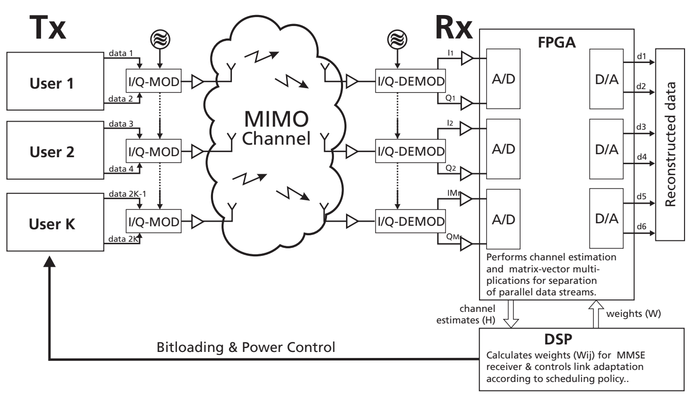

Figure 6.24: Setup of the MIMO test-bed for the MU SIMO MAC scenario.

6.4.1 Measurement Results

In the following we show measurement results on achievable data throughput, delay and buffer size with real-time channel adaptive bit-loading and scheduling. The performance of several scheduling policies was measured in an experiment and was then evaluated with regard to sum throughput, delay (queueing state) under certain QoS (BER and average rate) requirements of the individual users. We show the pros and cons of each scheduler depending on the available SNR at the BS. The real-time data transmission was performed with up to 5 MSymbols/s and up to 64-QAM modulation.

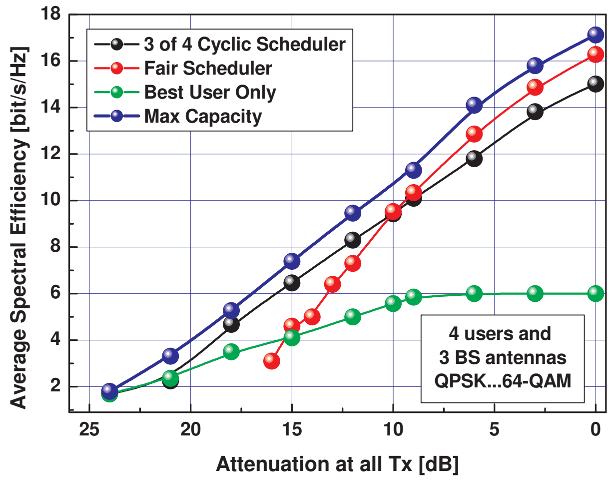

Figure 6.25: Achieved average spectral efficiency with different scheduling schemes. BS: 3 Antennas and 4 users with one antenna each.

The transmission scenario consists of 4 users which are distributed in the lab, assuming the same average individual rate request during the measurement. The BS is equipped with only 3 antennas, meaning that spatial multiplexing can be performed with up to 3 users maximum. The BS is moved over 5 meters (speed approx. 5 cm/s) across the room on a railway-like construction to ensure the same channel realizations for all experiments. This reproducibility is viable for the comparison of different schemes.

Fig. 6.25 shows the achievable average sum rate along the 5m trek across the room. The 3 of 4 cyclic scheduler (black) is outperformed by the fair scheduler (red) and the max. capacity approach (blue) in the high SNR region. With decreasing SNR the fair scheduler degrades below the cyclic scheduler since the sum rate is here dominated by the user which has the worst average channel. The best user only scheme shows the lowest cut-off rate while the other schemes tends to reach three times as much at high SNR.

Fig. 6.26 displays the possible average throughput per user. The filled symbols represent the averaged rate of the best user and the open symbols the average rate of the worst user. Here, over the whole SNR range the newly proposed fair scheduler achieves the highest QoS (average rate)-highest minimum average rate per user. This rate, at least can be assured (open circles) to all users. This clearly shows that already with today's hardware simple but efficient fair scheduling algorithms can be implemented.

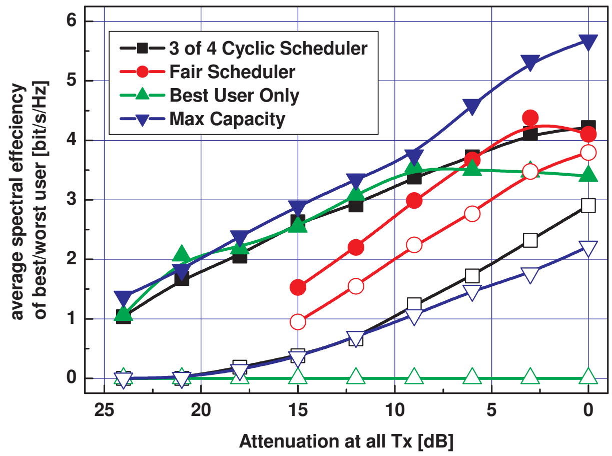

Figure 6.26: Average spectral efficiency per user with different scheduling schemes. filled symbols: user with best rate, open symbols: user with lowest rate. Note, that the min rate can be assured to all users as a QoS.

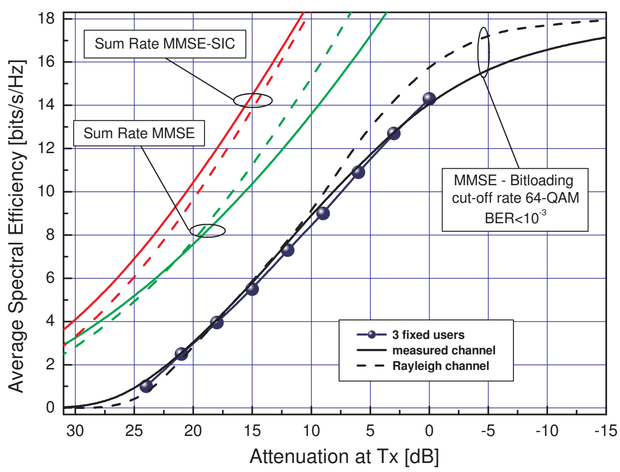

Figure 6.27: Comparison of theoretical and experimental results. Simulated and measured sum throughput with bit-loading. Simulation on Rayleigh channels (dashed lines), simulation on the measured channels (solid lines), measured throughput in the experiment (•).

Fig. 6.27 depicts the comparison of the sum throughput achieved in the experiment (circles) with the expected throughput on the measured channel along the 5m trek in the lab (solid lines) and a simulated Rayleigh channel (dotted lines). The slope of 3 bit/s/Hz per 3 dB SNR increase is not found in the experiment which is due to the fact that before full spatial multiplexing can be exploited the sum rate is cut-off due to the limited level of the QAM modulation. A similar behavior was observed in the sum rate experiments depicted in Fig. 6.19 where the cutoff rate limited the slope of the throughput curve. Therefore the measurement results coincide very well with what can be expected from the theory to be seen in Fig. 3.10 in section 3.4.

The experimental results shows clearly that channel aware bit-loading and scheduling are key factors to exploit the high capacity of the multi-path channel in a multi-user scenario efficiently. We could show that multi-antenna techniques are applicable already on today's hardware by implementing channel aware bit-loading and scheduling on a real-time experimental test-bed. We achieved an average spectral efficiency of 17 bit/s/Hz with 3 BS antennas and 4 users which means 85 Mbit/s average payload data rate in the up-link with an assured uncoded average BER$\leq 10^{-3}$.

7 Conclusions and future research

7.1 Conclusions

In this thesis, real-time capable transmission schemes for single-user and multi-user multiple-antenna wireless systems were studied where the emphasis was put on the single carrier flat-fading case. We started with a general optimization of several transmission strategies and developed algorithms for multi-antenna systems including channel aware bit-loading and multi-user scheduling. A selection of basic multi-antenna transmission schemes was implemented in a real-time MIMO test-bed and the performance was evaluated from the measured experimental results.

The following topics were covered and the following results were derived:

- In chapter 3 we made basic considerations for the implementation of multiple antenna techniques in WLAN systems. We pointed out that a reasonable antenna diversity might drastically improve the BER performance and makes the wireless connection over the fading channel more reliable. In scenarios with a strong LOS component we have to expect performance degradations in general due to rank deficiency of the MIMO channel. The performance loss can be limited by an adaptation of the transmission to the actual number of degrees of freedom.

- In chapter 4 we discussed optimum transmission schemes for the single-user MIMO scenario, the multi-user multiple access channel and the multi-user broadcast channel. We could show that channel knowledge should be exploited for an appropriate adaptation of the transmission scheme to the instantaneous channel states. Several sources of performance degradation observed in real-world applications were discussed and strategies to limit or combat the performance loss were proposed. We developed real-time capable bit-loading and multi-user scheduling strategies which are key elements for an efficient exploitation of the available bandwidth in real systems.

- Chapter 5 discussed the basic algorithms for MIMO base band signal processing and their potential towards optimization on standard DSPs.

- Chapter 6 was dedicated to the real-time transmission experiments. We introduced the reconfigurable MIMO test-bed and some configurations of relevance for the experiments.

  We showed measurement results on antenna diversity with fixed rates using linear and non-linear detection which coincide very well with numerical simulation based on statistical fading models. Furthermore, we showed the throughput gains which can be obtained from channel aware transmission using adaptive bit-loading and power control. A sum rate comparison of linear and non-linear detection schemes and SVD-MIMO transmission was given from the measurement results.

  Furthermore, we showed a first implementation of adaptive channel inversion for a broadcast scenario with decentralized receivers and an implementation of a fair multi-user scheduling scheme in space and time for the multiple access channel.

The measurements indicated that channel aware transmission schemes are key factors for a spectrally efficient exploitation of the radio channel. Pre-coding techniques based on CSI at the transmitter increase the system performance dramatically, which could be shown in the experiments on adaptive channel inversion and SVD-MIMO transmission. Finally, we could show that implementations of advanced multi-user scheduling algorithms can be run in real-time, offering high sum throughput and limited data queues for the individual users.

7.2 Open problems and future work

7.2.1 MIMO transmission in frequency selective channels

A recent transmission of 1 Gbit/s with low mobility in a standard office environment [JFH+05a] proved that MIMO algorithms developed in this thesis for the flat fading case could be transferred to a MIMO-OFDM system at little time expense. Minor optimization of MIMO algorithms matched to the computational capabilities of the MIMO-OFDM test-bed allowed such high rates (1 Gbit/s) with spatial multiplexing of three data streams.

A further increase in the number of OFDM-tones will linearly increase the complexity for the MIMO filter calculation. Therefore, parallel computing with several DSPs or even a filter calculation inside the FPGA are approaches at hand. The first option is mainly hardware limited (number of DSPs which can be connected to a DSP star architecture with reasonable effort) while the second requires fixed point implementation in most cases which is the right way to go for a final product implementation but it is very time consuming and error-prone for an experimental test-bed where algorithmic flexibility is highly desired.

For TDD systems exploiting MIMO-OFDM and multi-antenna capabilities at the MTs as well, the optimum transmission strategy of SVD-MIMO with water-filing / adaptive bit-loading can be exploited. To reduce the high computational effort of a full SVD at both ends of the link, including ambiguity of the calculated beam-forming vectors at each side an iterative ping-pong-like strategy [HK04] could be applied. The proposal of [HK04] suggests to pre-code in the i-th step with $\mathbf{Q}_i^H$ obtained from a QRD of $\mathbf{H}\mathbf{Q}_{i-1}^H$ which was received from the opposite direction. After a few TDD frames the $\mathbf{Q}_i$ converge to the desired $\mathbf{V}$ and $\mathbf{U}$ from SVD at much lower computational cost which is very important for MIMO-OFDM with many sub-carriers.

The MIMO decoding at each receiver side can be performed with a linear MMSE detector which is optimum when the pre-coding matrix converges to the input eigenvectors of the channel matrix $\mathbf{H}$. Since, in general the compound channel $\mathbf{H}\mathbf{Q}_{i-1}^H$ is much easier to decode with a linear MMSE than $\mathbf{H}$ alone, due to reduced noise enhancement, this strategy appears to be always advantageous even if pre-coding is only possible at one side of the link.

If joint signal processing at each link side is not possible for some reason then simple detection schemes with a good BER performance should be used. Proposals from [WF03, WBKK04a, WBKK04b] using lattice-reduction aided detectors show convenient performance with parallel detection. Nevertheless, it is still an open problem how to find the optimum or close-to-optimum matrix decomposition needed for this scheme under real-time constraints.

Furthermore channel interpolation techniques might be exploited to improve the accuracy of the channel estimates [SCP+04, HSJ+05] or to calculate MIMO filters more efficiently [HSJ+05, BB04b, CBB+05] since adjacent OFDM-tones are highly correlated.

7.2.2 Achieving reciprocity in base band

A prerequisite for an efficient exploitation of pre-coding techniques at the transmitter side is accurate channel knowledge at the transmitter. In static or quasi-static channels this might be achievable even with a transmission of the channel estimates or filter coefficients from the receiver via a feed-back link. But with shorter channel coherence times the required capacity of this feed-back will not be negligible at one hand and the CSI can easily become outdated on the other hand.

Therefore, a TDD multiplexing scheme offers the utilization of the channel measurement already available at the transmitter from a previous transmission in the opposite direction, at least in principle. Despite the fact that the radio channel between the antenna pairs is reciprocal (the uplink channel is the transpose of the down-link channel), this does not hold necessarily for the corresponding base band equivalent. The obvious reasons are different amplifiers, mixers and filters for the uplink and down-link transmission.

To achieve a sufficient reciprocity also for the base band channel a careful and desirable fully automatic hardware calibration is needed which allows to compensate the reciprocity-mismatch by appropriate base band signal processing. Another approach [JKI+04] proposed reciprocal transceivers which reduce the calibration effort to a reasonable minimum.

When reciprocity can be guaranteed also in the base band then full duplex TTD mode operation with pre-coding for Eigenmode Signalling and / or the down-link broadcast channel can be exploited.

7.2.3 Combatting Dirty RF effects with signal processing

In practical transmission systems often low-cost of-the-shelf hardware components are used. The ideally assumed complex valued signal model with perfect frame and symbol synchronization will then be severally violated. In direct consequence the performance can degrade significantly if no counter measures are taken. In the flat-fading case we simply could perform all BB SP as real-valued but this is not sufficient e.g. with OFDM when I/Q imbalances can cause severe cross talk between image carriers. Those and many other Dirty-RF effects [GFZ04] can be compensated by BB processing when an accurate mismatch estimation is performed in advance. For the case of OFDM [WF04a, WF04c, WF04b] proposed several approaches while [BB04a] discussed the degrading effect of phase noise on the channel estimation.

Having said this, it becomes clear that the compensation of "dirty" effects caused by low-cost hardware will become increasingly more important in base band signal processing.

7.2.4 Channel coding for adaptive Transmission with multiple antennas

A very important aspect towards the implementation of adaptive transmission schemes is an appropriate channel coding to fulfill the BER and delay requirements of various applications. In static or slow time variant channels, channel adaptive transmission allows to decompose the MIMO channel in parallel AWGN channels with predictable SNR. Therefore, we need channel coding more suited for AWGN channels than for channels with slow or fast fading. Furthermore the particular code design depends heavily on the overall used error correction scheme implemented in higher layers e.g. ARQ, Hybrid-ARQ or no ARQ. Let us give an example for illustration: ARQ works optimum at a frame error rate of about 10^{-1}, hence the FEC has to be chosen to perform well at this working point. If no ARQ scheme is used then the FEC must correct as many errors as possible without help from replicas of lost packets. Let us assume a raw BER of 10^{-3} and a targeted BER of 10^{-6} after the FEC for a certain application, then e.g. LDPC [Gal62, LTS00, ZPBF04] codes can be used if the block length is reasonably large (1000 - 5000 bits) even with high codes rates (0.8 or 0.9). The advantage of LDPC codes is to be seen in a high code rate achievable which improves directly the spectral efficiency for the data payload. This concept of combining channel adaptive MIMO transmission with LDPC codes seems very promising for data streaming applications which allow a reasonable amount of processing latency.

Other applications like voice over IP or video conferences may have quite stringent delay requirement, therefore FEC needs to perform according to the required BER for the application and at the same time meet the delay constraints. In reality, this means often a tradeoff between decoding complexity of the FEC algorithm (the faster the better towards short latency) and the achievable code rate. Therefore, often convolutional codes with low code rates e.g. 1/2 are used which means complex Viterby decoding at the receiver but this is a standard technique, implemented in many commercial systems e.g. UMTS.

Other proposals prefer Reed-Solomon-codes or product accumulate codes e.g.[LNG04] due to their low decoding complexity and their easily adjustable code rate between 1/2 and 1 which is very favorable for an efficient and fine granular implementation of adaptive modulation and coding (AMC).

In MIMO systems, the new dimension space allows for space-time codes (STC) in stead of the classical coding with interleaving in the time domain. STCs have in general moderate code rate but many of them are very easy to decode, which is favorable for limited latency and processing complexity.

Multi-antenna systems in frequency selective channels can exploit additional diversity from the frequency domain, which can be incorporated in the codes design. So a joint optimization over three dimension has to be performed in order to develop low complexity space-time-frequency-codes well suited for different working points (e.g. SNR, BER/FER targets etc.) and radio channel propagation environment with varying parameters like e.g. delay spread and doppler.

7.2.5 Channel Tracking for Higher Mobility

The ITU-R Recommendation M.1645 states that 'potential new radio interface(s) will need to support data rates of up to approximately 100 Mbit/s for high mobility such as mobile access and up to approximately 1 Gbit/s for low mobility such as nomadic/local wireless access'. Both aims are only achievable with a wider bandwidth and MIMO techniques. MIMO-OFDM is supposed to be a good candidate for 4G due to its relatively easy base band signal processing. Nevertheless, signal processing at high mobility in itself is challenging with MIMO and more challenging towards processing capability of the hardware when many OFDM-tones are used.

Due to the long coherence time of the channel in stationary scenarios the OFDM frame length can be chosen within some ms, and still the overhead from pilots for the measurement of the MIMO channel is acceptably low. This changes dramatically, when the channel is changing much faster due to high mobility of the user.

Still, all MIMO filters have to be updated within the now much shorter channel coherence time, but the same amount of pilot symbols transmitted now more often decreases the spectral efficiency significantly. One possible option are pilots scattered over the frame and the sub-carriers, allowing MIMO channel estimates at different time instances at different sub-carriers. As long as the phase evolution over time increases linearly, the separated data symbols after the MIMO detector can be phase compensated after.

Alternatively, blind channel and phase tracking based on detected data might help to reduce the pilot overhead, but short feed-back rates are a stringent requirement.

P Appendix

P.1 Properties of the Real-Valued Model

This appendix gives an overview on the properties of the mappings between the complex-valued model and the real-valued model.

$$\mathbf{H}_c \longmapsto \mathbf{H} = \begin{bmatrix} \Re\mathbf{H}_c & -\Im\mathbf{H}_c \\ \Im\mathbf{H}_c & \Re\mathbf{H}_c \end{bmatrix} \tag{P.1}$$

$$\mathbb{C}^{K_c \times K_c} \longrightarrow \mathbb{R}^{2K_c \times 2K_c} = \mathbb{R}^{K \times K} \tag{P.2}$$

and

$$\mathbf{x}_c \longmapsto \mathbf{x} = \begin{bmatrix} \Re\mathbf{x}_c \\ \Im\mathbf{x}_c \end{bmatrix} \tag{P.3}$$

$$\mathbb{C}^{K_c} \longrightarrow \mathbb{R}^{2K_c} = \mathbb{R}^{K} \tag{P.4}$$

For details on the transformation from complex-valued vectors and matrices see also [Tel99]:

$$\mathbf{H}_c^H \Longleftrightarrow \mathbf{H}^T \tag{P.5}$$

$$\| \mathbf{x}_c \|^2 = \sum_{k=1}^{K_c} |x_{c,k}|^2 = \sum_{k=1}^{K_c} \left((\Re x_{c,k})^2 + (\Im x_{c,k})^2 \right) = \sum_{k=1}^{K} x_k^2 = \| \mathbf{x} \|^2 \tag{P.6}$$

From this follows directly

$$\mathrm{trace}(\mathbf{x}_c \mathbf{x}_c^H) = \mathrm{trace}(\mathbf{x}\mathbf{x}^T) \tag{P.7}$$

$$\mathbf{H}_c \mathbf{G}_c = \mathbf{F}_c \Longleftrightarrow \mathbf{H}\mathbf{G} = \mathbf{F} \tag{P.8}$$

$$\mathbf{y}_c = \mathbf{H}_c\mathbf{x}_c + \mathbf{n}_c \Longleftrightarrow \mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{P.9}$$

$$\text{trace}(\mathbf{H}\mathbf{H}^T) = 2\text{trace}(\mathbf{H}_c\mathbf{H}_c^H) \tag{P.10}$$

If $\Re\mathbf{x}_c$ and $\Im\mathbf{x}_c$ are independent vectors of independent random variables with variance $\sigma^2 = \sigma_c^2/2$ each,

$$\mathbb{E}[\mathbf{x}_c\mathbf{x}_c^H] = \sigma_c^2\mathbf{I} \Longleftrightarrow \mathbb{E}[\mathbf{x}\mathbf{x}^T] = \frac{\sigma_c^2}{2}\mathbf{I} \tag{P.11}$$

Due to the factor of 2 in (P.10) and (P.11) or(P.14) we still have

$$\begin{aligned}\mathbb{E}[\text{trace}(\mathbf{x}_c\mathbf{x}_c^H)] = \sigma_c^2\text{trace}(\mathbf{I}) = \mathbb{E}\{trace(\mathbf{x}_c^H\mathbf{x}_c)\} = K_c\sigma_c^2 \\ = K\sigma^2 = \mathbb{E}[\text{trace}(\mathbf{x}^T\mathbf{x})] = \mathbb{E}[\text{trace}(\mathbf{x}\mathbf{x}^T)] = \sigma^2\text{trace}(\mathbf{I})\end{aligned} \tag{P.12}$$

where the dimensions of the identity matrices are K_c and $K = 2K_c$, respectively.

P.2 A mathematical measure of correlation

Following the outline of sec.2.2.2. in [E.J04] we take two arbitrarily chosen transmit correlation matrices $\mathbf{R}_T^1$ and $\mathbf{R}_T^2$ with the constraint that $\text{trace}(\mathbf{R}_T^1) = \text{trace}(\mathbf{R}_T^2) = n_T$ which is equivalent to

$$\sum_{l=1}^{n_T} \lambda_l^{T,1} = \sum_{l=1}^{n_T} \lambda_l^{T,2}, \tag{P.13}$$

with $\lambda_l^{T,1}$, $1 \leq l \leq n_T$, and $\lambda_l^{T,2}$, $1 \leq l \leq n_T$, are the eigenvalues of the covariance matrix $\mathbf{R}_T^1$ and $\mathbf{R}_T^2$, respectively.

This constraint regarding the trace of the correlation matrix $\mathbf{R}_T$ is necessary because the comparison of two transmission scenarios is only valid if the average path loss is equal. Without receive correlation, the trace of the correlation matrix can be written as

$$\text{trace}(\mathbf{R}_T) = \sum_{i=1}^{n_T} \left(\mathbb{E}\left[\mathbf{H}\mathbf{H}^H\right]\right)_{ii} = \sum_{i=1}^{n_T} \mathbb{E}\left[|\mathbf{h}_i|^2\right]. \tag{P.14}$$

However, the RHS of (P.14) is the sum of the average path loss from the transmit antenna $i = 1...n_T$. In order to study the impact of correlation on the achievable capacity separately, the average path loss is kept fixed by applying the trace constraint on the correlation matrices $\mathbf{R}_T^1$ and $\mathbf{R}_T^2$.

We will say that a correlation matrix $\mathbf{R}_T^1$ is more correlated than $\mathbf{R}_T^2$ with descending ordered eigenvalues $\lambda_1^{T,1} \geq \lambda_2^{T,1} \geq ... \geq \lambda_{n_T}^{T,1} \geq 0$ and $\lambda_1^{T,2} \geq \lambda_2^{T,2} \geq ... \geq \lambda_{n_T}^{T,2} \geq 0$ if

$$\sum_{k=1}^{m} \lambda_k^{T,1} \geq \sum_{k=1}^{m} \lambda_k^{T,2} \quad 1 \leq m \leq n_T - 1. \tag{P.15}$$

The measure of correlation which we will introduce is defined in a natural way: the larger the first m eigenvalues of the correlation matrices are (with the trace constraint in (P.14)), the more correlated is the MIMO channel. As a result, the most uncorrelated MIMO channel has equal eigenvalues, whereas the most correlated MIMO channel has only one non-zero eigenvalue which is given by $\lambda_1 = n_T$.

Before proceeding with our definition of 'more correlated' in terms of the eigenvalue distribution of the channel covariance matrix, we give the necessary definitions we will need in the following.
Definition 5: For two vectors $\mathbf{x}, \mathbf{y} \in R^n$ with descending ordered components $x_1 \geq x_2 \geq ... \geq x_n \geq 0$ and $y_1 \geq y_2 \geq ... \geq y_n \geq 0$ one says that the vector $\mathbf{x}$ majorizes the vector $\mathbf{y}$ and writes

$$\mathbf{x} \succ \mathbf{y} \;\text{ if }\; \sum_{k=1}^{m} x_k \geq \sum_{k=1}^{m} y_k \;\;, m = 1, ..., n-1. \;\text{ and }\; \sum_{k=1}^{n} x_k = \sum_{k=1}^{n} y_k.$$

The next definition describes a function Φ which is applied to the vectors $\mathbf{x}$ and $\mathbf{y}$ with $\mathbf{x} \succ \mathbf{y}$:
Definition 6: A real-valued function Φ defined on $\mathcal{A} \subset R^n$ is said to be *Schur-convex* on $\mathcal{A}$ if

$$\mathbf{x} \succ \mathbf{y} \;\text{ on }\; \mathcal{A} \Rightarrow \Phi(\mathbf{x}) \geq \Phi(\mathbf{y}).$$

Similarly, Φ is said to be *Schur-concave* on $\mathcal{A}$ if

$$\mathbf{x} \succ \mathbf{y} \;\text{ on }\; \mathcal{A} \Rightarrow \Phi(\mathbf{x}) \leq \Phi(\mathbf{y}).$$

Example: Suppose that $\mathbf{x}, \mathbf{y} \in R_+^n$ consists of positive real numbers and the function Φ is defined as the sum of the squared components of the vectors, i.e. $\Phi_2(\mathbf{x}) = \sum_{k=1}^{n} |x_k|^2$. Then, it is easy to show that the function Φ_2 is Schur-concave on R_+^n, i.e. if $\mathbf{x} \succ \mathbf{y} \Rightarrow \Phi_2(\mathbf{x}) \leq \Phi_2(\mathbf{y})$.

We will need the following lemma (see [MO79, Theorem3.A.4]) which is sometimes called Schur's condition. It provides an approach for testing whether some vector valued function is Schur-convex or not.
Lemma 1: Let $\mathcal{I} \subset \mathbb{R}$ be an open interval and let $f : \mathcal{I}^n \to \mathbb{R}$ be continuously differentiable. Necessary and sufficient conditions for f to be Schur-convex on $\mathcal{I}^n$ are

$$f \;\text{ is symmetric on }\; \mathcal{I}^n \tag{P.16}$$

and

$$(x_i - x_j)\left(\frac{\partial f}{\partial x_i} - \frac{\partial f}{\partial x_j}\right) \geq 0 \;\;\text{ for all }\; 1 \leq i, j \leq n. \tag{P.17}$$

Since $f(\mathbf{x})$ is symmetric, Schur's condition can be reduced as in [MO79, p. 57]

$$(x_1 - x_2)\left(\frac{\partial f}{\partial x_1} - \frac{\partial f}{\partial x_2}\right) \geq 0. \tag{P.18}$$

From Lemma 1 follows that $f(\mathbf{x})$ is a Schur-concave function on $\mathcal{I}^n$ if $f(\mathbf{x})$ is symmetric and

$$(x_1 - x_2)\left(\frac{\partial f}{\partial x_1} - \frac{\partial f}{\partial x_2}\right) \leq 0. \tag{P.19}$$

The definition of Schur-convexity and Schur-concavity can be extended if another function $\Psi : R \rightarrow R$ is applied to $\Phi(\mathbf{x})$. Assume that Φ is Schur-concave, if the function Ψ is monotonically increasing then the expression $\Psi(\Phi(\mathbf{x}))$ is Schur-concave, too. If we take for example the function $\Psi(n) = \log(n)$ for $n \in R_+$ and the function Φ_p from the example above, we can state that the composition of the two functions $\Psi(\Phi_p(\mathbf{x}))$ is Schur-concave on R_+^n. This result can be generalised for all possible compositions of monotonically increasing as well as decreasing functions, and Schur-convex as well as Schur-concave functions. For further information about majorization theory see [MO79].

The following definition provides a measure for comparison of two correlation matrices.
Definition 7: The transmit correlation matrix $\mathbf{R}_T^1$ is more correlated than $\mathbf{R}_T^2$ if and only if

$$\sum_{l=1}^{m} \lambda_l^{T,1} \geq \sum_{l=1}^{m} \lambda_l^{T,2} \quad \text{for} \;\; m = 1...n_T, \;\; \text{and} \;\; \sum_{l=1}^{n_T} \lambda_1^{T,1} = \sum_{l=1}^{n_T} \lambda_2^{T,2}. \tag{P.20}$$

One says that the vector consisting of the ordered eigenvalues $\boldsymbol{\lambda}_1^T$ majorizes $\boldsymbol{\lambda}_2^T$, and this relationship can be written as $\boldsymbol{\lambda}_1^T \succ \boldsymbol{\lambda}_2^T$ like in Definition 1.

Remark I: It can be shown that vectors with more than two components cannot be totally ordered. So there are examples of correlation vectors that cannot compared using our Definition 7, e.g. $\boldsymbol{\lambda}_1 = [0.6, 0.25, 0.15]$ and $\boldsymbol{\lambda}_2 = [0.55, 0.4, 0.05]$. This is a problem of all possible orders for comparing correlation vectors. Majorization induces only a partial order.

Note that our definition of correlation in Definition 7 differs from the usual definition in statistics. In statistics a diagonal covariance matrix indicates that the random variables are uncorrelated. This is independent of the auto-covariances on the diagonal. In our definition, we say that the antennas are uncorrelated if in addition to statistical independence, the auto-covariances of all entries are equal. This difference to statistics occurs because the direction, i.e. the unitary matrices of the correlation have no impact on our measure of correlation. Imagine the scenario in which all transmit antennas are uncorrelated, but have different average transmit powers because of their amplifiers. In a statistical sense, one would say the antennas are uncorrelated. Our measure of correlation says that the antennas are correlated, because they have different transmit powers. The measure of correlation in Definition 7 is more suitable for the analysis of the performance of multiple antenna systems, because different transmit powers at the antennas obviously have a strong impact on the performance. In this thesis, these effects are considered.

This measure of correlation allows us to analyze the impact of correlation on the various performance metrics introduced in chapter 2 in single-user MIMO systems under different types of CSI. In the following, the measure of correlation is applied to transmit correlation matrices $\mathbf{R}_T$ and to receive correlation matrices $\mathbf{R}_R$ as well.

Remark II: As mentioned above, the case in which the transmit antennas are fully correlated corresponds to $\lambda_1^T = n_T$, $\lambda_2^T = ... = \lambda_{n_T}^T = 0$. The case in which the transmit antennas are fully uncorrelated corresponds to $\lambda_1^T = \lambda_2^T = ... = \lambda_{n_T}^T = 1$. This illustrates that the expression in

(P.20) can be used as a measure for correlation.

Example: At this point, we give another example for the measure of correlation. Assume the situation in figure (P.1). We have two different correlation scenarios. In scenario A and B the largest two eigenvalues ($\lambda_1^A = \lambda_1^B$ and $\lambda_2^A = \lambda_2^B$) are equal. The smallest three eigenvalues in scenario B are equal ($\lambda_3^B = \lambda_4^B = \lambda_5^B$) but in scenario A the smallest three eigenvalues are unequal ($\lambda_3^A > \lambda_4^A > \lambda_5^A$). In addition to this, the sum of all eigenvalues in scenario A and B is equal. Applying the order which is introduced in Definition 3, eigenvalue vector A majorizes eigenvector B ($\boldsymbol{\lambda}^A \succ \boldsymbol{\lambda}^B$).

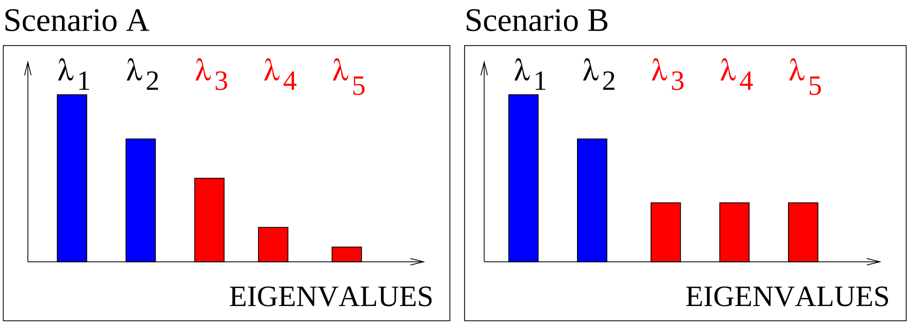

Figure P.1: Example correlation matrix eigenvalue distribution.

Scenario B applies for all eigenvalue distributions $\boldsymbol{\lambda}$ with fixed λ_1 and λ_2 and equal trace the 'smallest' eigenvalue distribution, i.e. $\boldsymbol{\lambda}^B \prec \boldsymbol{\lambda}$ for all $\boldsymbol{\lambda}$ with

$$\lambda_1 + \lambda_2 + \sum_{k=3}^{n_T} \lambda_k = 1$$

and for $\boldsymbol{\lambda}^B$ with

$$\frac{1 - \lambda_1^B - \lambda_2^B}{n_T - 2} = \lambda_3^B = \ldots = \lambda_{n_T}^B.$$

Publication List

[HFG$^+$05b] T. Haustein, A. Forck, H. Gäbler, V. Jungnickel, and S. Schiffermüller. Real Time Signal Processing for Multi-Antenna Systems: Algorithms, Optimization and Implementation on an Experimental Test-bed. *will appear in EURASIP - special issue on Implementation Aspects and Testbeds for MIMO Systems*, 2006.

[HJZ$^+$06] T. Haustein, V. Jungnickel, W. Zirwas and E. Schulz. MIMO-OFDM for Wireless High Data Rate. In *ICT, Funchal, Madeira island, Portugal*, Mai 9-12 2006.

[HFG$^+$06] T. Haustein, A. Forck, H. Gäbler, V. Jungnickel, and S. Schiffermüller. Real-Time Experiments on Channel Adaptive Transmission in the Multi-User Up-link at very high Data Rates using MIMO-OFDM. In *EUSIPCO, Florence, Italy*, Sept. 4-8, 2006.

[HSJ$^+$05] T. Haustein, S. Schiffermüller, V. Jungnickel, M. Schellmann, T. Michel, and G. Wunder. Interpolation and Noise Reduction in MIMO-OFDM - A Complexity Driven Perspective. In *ISSPA, Sydney, Australia*, Aug. 28th - 1. Sept. 2005.

[HFG$^+$05a] T. Haustein, H. Boche, M. Wiszanowski, and E. Schulz. Real-Time Implementation of Cross-Layer Optimization: Multi-Antenna High Speed Uplink Packet Access. In *IEEE ICASSP, Philadelphia, U.S.A.*, March 19-23th 2005.

[HFGS05] T. Haustein, A. Forck, H. Gäbler, and S. Schiffermüller. From Theory to Practice: MIMO Real-Time Experiments of Adaptive Bit-loading with linear and non-linear Transmission and Detection Schemes. In *IEEE VTC Spring, Stockholm, Sweden.*, May 2005.

[JFH$^+$05a] V. Jungnickel, A. Forck, T. Haustein, S. Schiffermüller, C. von Helmolt, F. Luhn, M. Pollock, C. Juchems, M. Lampe, W. Zirwas J. Eichinger, and E. Schulz. 1 Gbit/s MIMO-OFDM Transmission Experiments. In *IEEE VTC Fall, Dallas,TX, U.S.A.*, Sept. 25th-28th 2005.

[HFG$^+$04a] T. Haustein, A. Forck, H. Gäbler, C.v. Helmolt, V. Jungnickel, and U. Krueger. Implementation of Adaptive Channel Inversion in a Real-Time MIMO System. In *IEEE PIMRC, Barcelona, Spain*, Sept. 2004.

[HFG$^+$04b] T. Haustein, A. Forck, H. Gäbler, C.v. Helmolt, V. Jungnickel, and U. Krueger. Implementation of Channel Aware Scheduling and Bit-Loading for the Multiuser SIMO MAC in a Real-Time Demonstration Test-bed at high Data Rate. In *IEEE VTC Fall, Los Angeles, C.A., U.S.A.*, Sept. 2004.

[HFG$^+$04c] T. Haustein, A. Forck, H. Gäbler, C.v. Helmolt, V. Jungnickel, and U. Krueger. Real-Time MIMO Transmission Experiments with Adaptive Bitloading. In *IASTED WOC, Banff, Canada*, July 8-10th 2004.

[HHI$^+$04] T. Haustein, C.v. Helmolt, G. Istoc, V. Jungnickel, and U. Krüger. Reziproke HF-Sendeempfangsstufe und Anwendung davon. Patent 10 2004 029 893.9, 2004.

[JKI$^+$04] V. Jungnickel, U. Krüger, G. Istoc, T. Haustein, and C.v. Helmolt. A MIMO System with Reciprocal Transceivers for the Time-Division Duplex Mode. In *IEEE APS, special session: Antennas and Propagation in MIMO System, Monterrey, CA, U.S.A.*, March 2004.

[JHF$^+$04] V. Jungnickel, T. Haustein, A. Forck, S. Schiffermueller, H. Gaebler, C. v. Helmolt, W. Zirwas, J. Eichinger, and E. Schulz. Real-Time Concepts for MIMO-OFDM. In *IEEE Global Mobile Congress*, Shanghai, China, Oct. 11-13 2004.

[HB03] T. Haustein and H. Boche. Optimal Power Allocation for MSE and Bitloading in MIMO Systems and the Impact of Correlation. In *IEEE ICASSP Hongkong*, volume IV, pages 405–409, April 2003.

[HSB03] T. Haustein, M. Schubert, and H. Boche. On Power Reduction Strategies for the Multi-User Downlink with Decentralized Receivers. In *IEEE VTC-Spring , Korea*, April 2003.

[HK03] T. Haustein and U. Krueger. Smart Geometrical Antenna Design exploiting the LOS component to enhance a MIMO System based on Rayleigh-fading in Indoor Scenarios. In *IEEE PIMRC, Beijing, China*, 7-10th Sept. 2003.

[HLB03] T. Haustein, G. Lehmann, and H. Boche. Bitloading for the SIMO Multiple Access Channel. In *IEEE PIMRC, Beijing, China*, 7-10th Sept. 2003.

[HBH03] E. Jorswieck H. Boche and T. Haustein. Channel Aware Scheduling for Multiple Antenna Multiple Access Channels. In *Asilomar 2003, Stanford, USA.*, Nov 2003.

[JHPvH03] V. Jungnickel, T. Haustein, V. Pohl, and C. v. Helmolt. Link Adaption in Multi-Antenna Systems. In *IEEE VTC-Spring 2003, Jeju, Korea.*, April 22-25th 2003.

[HPJ+02] T. Haustein, V. Pohl, V. Jungnickel, E. Jorswieck, and C. v. Helmolt. Performance of MIMO Systems with Channel Inversion. In *IEEE VTC-Spring 2002, Birmingham, Alabama, U.S.A.*, May. 6-10th 2002.

[JPN+02] V. Jungnickel, V. Pohl, H. Nguyen, U. Krueger, T. Haustein, and C.v. Helmolt. High Capacity Antennas for MIMO Radio Systems. pages 407–411, 2002.

[JHJvH02] V. Jungnickel, T. Haustein, E. Jorswieck, and C. von Helmolt. On Linear Pre-Processing in Multi-Antenna Systems. In *IEEE Globecom, Taipeh, Taiwan*, 2002.

[PJHvH02] V. Pohl, V. Jungnickel, T. Haustein, and C. v. Helmolt. Antenna Spacing in MIMO Indoor Channels. In *IEEE VTC-Spring 2002, Birmingham, Alabama, U.S.A.*, May. 6-10th 2002.

[HJJ+01a] T. Haustein, E. Jorswieck, V. Jungnickel, U. Krueger, V. Pohl, and C. v. Helmolt. Bit Error Rates for a MIMO System in Rayleigh and Rician Channels. In *IEEE VTC-Fall, Atlantic City, N.J. U.S.A.*, Oct. 2001.

[JWHJ01] E. Jorswieck, G. Wunder, T. Haustein, and V. Jungnickel. Inverse Eigenvalue Statistics for Rayleigh and Rician MIMO Channels. In *IEE Professional Network on Antennas and Propagation*, pages 3/1–3/6, Dec. 2001. Seminar.

[JHJvH01] V. Jungnickel, T. Haustein, E. Jorswieck, and C. v. Helmolt. A MIMO WLan based on Linear Channel Inversion. In *IEE Professional Network on Antennas and Propagation*, Dec. 2001. Seminar.

References

[AC02] A.G. Armada and J.M. Cioffi. Multi-user constant-energy bit loading for m-psk-modulated orthogonal frequency division multiplexing. In *IEEE Wireless Communications and Networking Conference (WCNC)*, volume 3, pages 450–454, March 2002.

[ACV$^+$04] A. Agustin, E. Calvo, J. Vidal, O. Muñoz, and M. Lamarca. Hybrid Turbo FEC/ARQ Systems and distributed Space-Time coding for cooperative Transmission in the Downlink. In *Personal Indoor Mobile Radio Communications (PIMRC), Barcelona, Spain*, 5th-8th September 2004.

[Ala98] S.M. Alamouti. A simple transmit diversity technique for wireless communications. *IEEE Trans. on Information Theory*, 16(8):1451–1458, 1998.

[AMdC01] M.R. Andrews, P.P. Mitra, and R. de Carvalho. Tripling the capacity of wireless communications using electromagnetic polarization. *Nature*, 409:316–318, 2001.

[ATG99] M.S. Alouini, X. Tang, and A.J. Goldsmith. An Adaptive Modulation Scheme for Simultaneous Voice and Data Transmission over Fading Channels. *IEEE Journal on Selected Areas in Communications*, 17(5):837, May 1999.

[AVCM04] A. Agustín, J. Vidal, E. Calvo, and O. Muñoz. Evaluation of Turbo H-ARQ Schemes for cooperative MIMO Transmission. In *International Workshop on Wireless Ad-Hoc Networks (IWWAN),Oulu ,Finland*, May 2004.

[BB04a] D. Baum and H. Bölcskei. Impact of Phase Noise on MIMO Channel Measurement Accuracy. In *Proc. IEEE VTC'04 Fall*, Sep. 2004.

[BB04b] M. Borgmann and H. Bölcskei. Interpolation-Based Efficient Matrix Inversion for MIMO-OFDM Receivers. In *Proc. 38th Asilomar Conf. Signals, Systems, Computers, Pacific Grove, CA, (invited paper)*, Nov. 2004.

[BCK03] A. Bourdoux, B. Come, and N. Khaled. Non-reciprocal Transceivers in OFDM/SDMA Systems: Impact and Mitigation. *RAWCON,Boston, U.S.A.*, Aug. 2003.

[BF01] A.N. Barreto and G. Fettweis. Capacity Increase in the Downlink of Spread Spectrum Systems through Joint Signal Precoding. In *IEEE Conference on Communications (ICC), Helsinki, Finland*, volume 4, pages 1142–1146, 11th-14th June 2001.

[BGT93] C. Berrou, A. Glavieux, and P. Thitimajshima. Near shannon limit error-correcting coding and decoding: turbo-codes. *IEEE International Conference on Communications*, pages 1064–1070, May 1993.

[BHSN00] C. Brunner, J.S. Hammerschmidt, A. Seeger, and J.A. Nossek. Space-Time Eigenrake and Downlink Eigenbeamformer: Exploiting Long-Term and Short-Term Channel Properties in WCDMA. In *IEEE Globecom, San Francisco, CA*, 2000.

[Bin90] J.A.C. Bingham. Multicarrier modulation for data transmission: An idea whose time has come. *IEEE Communications Magazine*, 28(5):5–14, May 1990.

[BJ02] H. Boche and E.A. Jorswieck. Sum Capacity Optimization of the MIMO Gaussian MAC. In *WPMC, Honolulu, USA*, 2002.

[BMWT00] P.W. Baier, M. Meurer, T. Weber, and H. Tröger. Joint Transmission (JT, an alternative rationale for the downlink of Time Division CDMA using multi-element transmit antennas). *proc. IEEE ISSSTA, NJIT, New Jersey, U.S.A.*, pages 1–5, 6-8th September 2000.

[Bor] S. Borst. User-Level Performance of Channel-Aware Scheduling Algorithms in Wireless Data Networks.

[BS02a] H. Boche and M. Schubert. A General Duality Theory for Uplink and Downlink Beamforming. *IEEE VTC, Vancouver, Canada*, Sept. 2002.

[BS02b] H. Boche and M. Schubert. Solution of the SINR Downlink Beamforming Problem. In *Proc. Conf. on Information Sciences and Systems (CISS), Princeton, USA*, March 2002.

[BSJ03] H. Boche, M. Schubert, and E.A. Jorswieck. Throughput Maximization for the Multiuser MIMO Broadcast Channel. In *ICASSP, Hong Kong*, April 2003. (invited).

[BW02] H. Boche and M. Wiczanowski. Stability Region of Arrival Rates and Optimal Scheduling for MIMO-MAC - A Cross-Layer Approach. In *Proc. IEEE International Zurich Seminar, Zurich*, Feb. 2002.

[BW03] H. Boche and M. Wiczanowski. Queueing Theoretic Optimal Scheduling for Multiple Input Multiple Output Multiple Access Channel. In *IEEE Symposium on Signal Processing and Information Technology (ISSPIT), Invited Paper*, 2003.

[BW04a] H. Boche and M. Wiczanowski. Optimal Scheduling for High Speed Uplink Packet Access - A Cross-Layer Approach. In *IEEE Vehicular Technology Conference (VTC), Spring 04*, March 2004.

[BW04b] H. Boche and M. Wiczanowski. Optimal Scheduling for High Speed Uplink Packet Access - A Cross-Layer Approach. Technical report, available at Heinrich Hertz Chair for Mobile Communications, Technical University of Berlin, 2004.

[BW04c] H. Boche and M. Wiczanowski. Scheduling and Signal Processing for MIMO Multiple Access Channel:a Geometric view. Technical report, available at Heinrich Hertz Chair for Mobile Communications, Technical University of Berlin, 2004.

[Cam98] J. Campello. Optimal discrete bit loading for multicarrier modulation systems. In *IEEE ISIT*, page 193ff, Aug. 1998.

[Cam99] J. Campello. Practical bit loading for DMT. In *IEEE ICC*, volume 2, pages 801–805, 1999.

[CBB+05] D. Cescato, M. Borgmann, H. Bölcskei, J. Hansen, and A. Burg. Interpolation-Based QR Decomposition in MIMO-OFDM Systems. In *IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC) (invited paper)*, June 2005.

[CCB95] P.S. Chow, J.M. Cioffi, and J.A.C. Bingham. A Practical Discrete Multitone Transceiver Loading Algorithm for Data Transmission over Spectrally Shaped Channels. *IEEE Trans. on Communications*, 43(2/3/4):773–775, Feb./ Mar./ April 1995.

[CFV+00] C.N. Chuah, G.J. Foschini, R.A. Valenzuela, D. Chizhik, J. Ling, and J.M. Kahn. Capacity Growth of Multi-Element Arrays in Indoor and Outdoor Wireless Channels. *IEEE WCNC, Chicago*, Sept. 23.-28. 2000.

[CG01] S.T. Chung and A.J. Goldsmith. Degrees of Freedom in Adaptive Modulation: A Unified View. *IEEE Transactions on Comunications*, 49(9):1561, September 2001.

[CLH+03] S.T. Chung, A.L. Lozano, H.C. Huang, A. Sutivong, and J.C. Cioffi. Approaching the MIMO capacity with a low-rate feedback channel in V-BLAST. In *EURASIP JASP* vol.5:762–771. 2004.

[CLW+02] D. Chizhik, J. Ling, P. Wolniansky, R. Valenzuela, N. Costa, and K. Huber. Multiple Input Multiple Output Measurements and Modelling in Manhattan. In *IEEE VTC-Fall, Vancouver, Canada*, Oct. 2002.

[Cos83] M. Costa. Writing on dirty paper. *IEEE Trans. on Information Theorie*, 29(3):439–441, May 1983.

[CS01] G. Caire and S. Shamai. On the multiple-antenna broadcast channel. In *36th Asilomar Conf. on Signals, Systems and Computers, Asilomar, Pacific Grove, CA*, 3-6 Nov. 2001.

[CS03] G. Caire and S. Shamai. On the Achievable Throughput of a Multi-Antenna Gaussian Broadcast Channel. *IEEE Trans. on Inform. Theory,*, 49(7):1691–1706, July 2003.

[CTK98] C.N. Chuah, D. Tse, and J.M. Kahn. Capacity of Multi-Antenna Array Systems in Indoor Wireless Environment. In *IEEE Globecom, Sydney, Australia*, Nov. 1998.

[Dar04] D. Dardari. Ordered Subcarrier Selection Algorithm for OFDM based high-speed WLANs. *In IEEE Trans. on Wireless Communications*, Volume: 3 , Issue: 5 , pp 1452 - 1458, Sept. 2004.

[DF99] P.F. Driessen and G.J. Foschini. On the Capacity Formula for Multiple Input - Multiple Output Wireless Channels: A Geometric Interpretation. *IEEE Transactions on Communication*, 47(2):173–176, February 1999.

[EB99] F. Boixadera Espax and J.J. Boutros. Capacity Considerations for Wireless MIMO Channels. In *Workshop on Multiaccess, Mobility and Teletraffic for Wireless Communications, Venice, Italy*, 6th-8th Oct. 1999.

[Ede89] A. Edelman. *Eigenvalues and Condition Numbers of Random Matrices*. PhD thesis, Department of Mathematics MIT, Cambridge, MA, 1989.

[Egg01] J.J. Eggers. *Information Embedding and Digital Watermarking as Communication with Side Information*. PhD thesis, University Erlangen-Nuremberg, 2001.

[E.J04] E.Jorswieck. *Unified Approach for the Optimisation of Single-user and Mmulti-user Multiple-Input Multiple-Output Wireless Systems*. PhD-Thesis, Technical University of Berlin, 2004.

[ESA04] C.H.Y. Eugene, K. Sakaguchi, and K. Araki. Experimental and Analytical Investigation of MIMO Channel Capacity in an Indoor Line-of-Sight (LOS) Environment. In *IEEE PIMRC, Barcelona, Spain*, 5/8-th Sept. 2004.

[ESZ00] U. Erez, S. Shamai, and R. Zamir. Capacity and Lattice-Strategies for Cancelling Known Interference. In *ISIT 2000*, Honolulu, Hawaii, U.S.A., 5-8 Nov. 2000.

[FG98] G.J. Foschini and M.J. Gans. On the Limits of Wireless Communications in a Fading Environment When Using Multiple Antennas. *Wireless Pers. Commun.*, page pp. 315ff, 1998.

[FH96] R. Fischer and J. Huber. A New Loading Algorithm for Discrete Multitone Transmission. In *IEEE Global Telecommunications Conference (GLOBECOM '96), London*, pages 724–728, Nov. 1996.

[FH03] R. Fischer and J. Huber. Signal Processing in Receivers for Communication over MIMO ISI Channels. In *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT 2003), Darmstadt, Germany*, Dec. 2003.

[Fis02] R.F.H. Fischer. *Precoding and Signal Shaping for Digital Transmission*. John Wiley & Sons, Inc., New York, NY, USA, 2002.

[Fos96] G.J. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multiple antennas. *Bell Labs Tech. J.*, pages pp. 41–59, 1996. 1,(2).

[FSH04] R. Fischer, C. Stierstorfer, and J. Huber. Precoding for Point-to-Multipoint Transmission over MIMO ISI Channels. In *International Zurich Seminar (IZS '04), Zurich, Switzerland*, pages 208–211, Feb. 2004.

[FS98] S. Falahati and A. Svensson. Hybrid type-II ARQ Schemes for Rayleigh Fading Channels. In *International Conference on Telecommunications, Porto Carras, Greece*, pages 39–44, June 1998.

[FW03] R. F. H. Fischer and C. Windpassinger. Real- vs. Complex-valued Equalisation in V-BLAST Systems. *Electronics Letters*, 39(5):470–471, March 2003.

[FWLH02] R. Fischer, C. Windpassinger, A. Lampe, and J. Huber. MIMO Precoding for Decentralized Receivers. *ISIT, Lausanne*, page 496, June 2002.

[Gal62] R. G. Gallager. Low-density parity-check codes. *IRE Trans. Inform. Theory*, IT-8:21–28, Jan. 1962.

[Gan86] F.R. Gantmacher. *Matrizentheorie*. Springer, Berlin Heidelberg New York Tokyo, 1986.

[GAY+02] J. Gans, N. Amitay, Y.S. Yeh, H. Xu, T.C. Damen, R.A. Valenzuela, T. Sizer, R. Storz, D. Taylor, W.M. MacDonald, C. Tran, and A. Adamiecki. Outdoor BLAST Measurement System at 2.44 GHz: Calibration and Initial Results. *IEEE JSAC*, 20(3):570–583, 2002.

[GBGP00] D. Gesbert, H. Bölcskei, D. Gore, and A. Paulraj. Mimo Wireless Channels: Capacity and Performance Prediction. In *IEEE GLOBECOM 2000*, pages 1083–1088, Nov. 2000.

[GC02] G. Ginis and J.M. Cioffi. Vectored Transmission for Digital Subscriber Line Systems. *IEEE Journal on Selected Areas in Communications*, 20(5):1085–1104, June 2002.

[GFZ04] D. Petrovic M. Windisch P. Zillmann G. Fettweis, M. Löhning and E. Zimmermann. Dirty RF. In *11th Wireless World Research Forum (WWRF11), Oslo, Norway*, June 2004.

[GL96] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, MD, USA, 3rd edition edition, 1996.

[GOdMV04] P. García, A. Ortega, J. de Mingo, and A. Valdovinos. Nonlinear Distortion Cancellation in OFDM Systems Uing an Adaptive LINC Structure. In *IEEE PIMRC, Barcelona, Spain*, 5/8-th Sept. 2004.

[Goe99] D.L. Goeckel. Adaptive coding for time-varying channels using outdated fading estimates. *IEEE Trans. Commun.*, 47:844–855, June 1999.

[GVK02] N. Geng, I. Viering, and M. Kiessling. Multi-User MIMO-OFDM Cell Throughput under Real-World Propagation Conditions. In *IEEE VTC Fall, Vancouver, Canada*, 2002.

[H. 02] H. Boche and E. Jorswieck . On the Schur-concavity of the Ergodic and Outage Capacity with respect to Correlation in Multi-antenna Systems with no CSI at the Transmitter. In *Allerton 02*, October 2002.

[Hay01] S. Haykin. *Communication Systems*. John Wiley & Sons, Inc., New York, NJ, U.S.A., 4th edition, 2001.

[HBD00] J.S. Hammerschmidt, C. Brunner, and C. Drewes. Eigenbeamforming - A Novel Concept in Space and Space-Time Processing. In *European Wireless 2000, Dresden, Germany*, 2000.

[HMW05] O. Henkel, T. Michel, and G. Wunder. Moderate Complexity Aproximation to MMSE for MIMO-OFDM Systems. In *VTC-Spring, Stockholm, Sweden*, 2005.

[HD03] T. Hunziker and D. Dahlhaus. Optimal Power Allocation for OFDM Systems with Ideal Bit-Interleaving and Hard-Decision Decoding. In *IEEE International Conference on Communications (ICC 2003), Anchorage, AK*, volume 5, pages 3392–3397, May 2003.

[HK04] Kwang. H. Han and Chung G. Kang. Bidirectional Iterative Tracking (BIT) Scheme for Estimating Spatial Sub-Channel Parameters in Time-Varying Reciprocal MIMO Channels. In *IEEE Global Mobile Congress*, Shanghai, China, Oct. 11-13 2004.

[HKM+04] K. Higuchi, H. Kawai, N. Maeda, M. Sawahashi, T. Itoh, Y. Kakura, A. Ushirokawa, and H. Seki. Likelihood Function for QRM-MLD Suitable for Soft-Decision Turbo Decoding and Its Performance for OFCDM MIMO Multiplexing in Multipath Fading Channel. In *IEEE PIMRC, Barcelona, Spain*, Sept. 2004.

[HL02] K.J. Hwang and K.B. Lee. Transmit Power Allocation with small Feedback Overhead for a Multiple Antenna System. *IEEE Vehicular Technology Conf., Vancouver, BC, Canada*, pages 2158–2162, Sept. 2002.

[HM72] H. Harashima and H. Miyakawa. Matched-transmission technique for channels with intersymbol interference. *IEEE Transactions on Communications*, 20(4):774–780, Aug. 1972.

[HPB+04] S. Häne, D. Perels, D. Baum, M. Borgmann, A. Burg, N. Felber, W. Fichtner, and H. Bölcskei. Implementation aspects of a real-time multi-terminal MIMO-OFDM testbed. In *IEEE Radio and Wireless Conference (RAWCON)*, Sep. 2004.

[Hun02] Thomas Hunziker. *Multicarrier Modulation Techniques for Bandwidth Efficient Fixed Wireless Access Systems*, volume 11 of *Series in Broadband Communications*. Hartung-Gorre Verlag Konstanz, 1 edition, 2002.

[Irm05] Ralf Irmer. *Multiuser Transmission in Code Division Multiple Access Mobile Communications Systems*. PhD thesis, Technical University Dresden, Germany, April. 2005.

[IBW04] I.J. Wassell I. Berenguer, J. Adeane and X. Wang. Lattice-Reduction-Aided Receivers for MIMO-OFDM in Spatial Multiplexing Systems. In *IEEE PIMRC, Barcelona, Spain*, 5/8-th Sept. 2004.

[IHRF03] R. Irmer, R. Habendorf, W. Rave, and G. Fettweis. Nonlinear Multiuser Transmission using Multiple Antennas for TDD-CDMA. In *IEEE WPMC*, volume 3, pages 251–255, Oct. 2003.

[IHRF04] R. Irmer, R. Habendorf, W. Rave, and G. Fettweis. Nonlinear chip-level multiuser tranmission for TDD-CDMA with frequency-selective MIMO channels. In *ITG Conference on Source and Channel Coding (SCC). Erlangen, Germany*, 14th-16th January 2004.

[Jak74] W.C. Jakes. *Microwave Mobile Communication*. New York: IEEE Press, 1974.

[J.G00] J.G.Proakis. *Digital Communications*. McGraw-Hill, New York, NY, U.S.A., 4th edition, 2000.

[JG04] N. Jindal and A. Goldsmith. Dirty Paper Coding vs. TDMA for MIMO Broadcast Channels. *Draft to IEEE Trans. on Information Theory*, submitted June 2004.

[JJR94] J.H.Winters, J.Salz, and R.D.Gitlin. The Impact of Antenna Diversity on the Capacity of Wireless Communication Systems. *IEEE Trans. on Commun.*, 42(2), Feb./Mar./Apr. 1994.

[Joh04] M. Joham. *Optimization of Linear and Nonlinear Transmit Signal Processing*. Shaker Verlag GmbH, Aachen, June 2004. PhD Thesis,Technical University Munich,ISBN: 3-8322-2913-2.

[KH95] R. Knopp and P.A. Humblet. Information Capacity and Power Control in Single-Cell Multiuser Communications. In *IEEE Int. Conf. on Communications (ICC), Seattle*, pages 331–335, June 1995.

[KMJ+00] J.P. Kermoal, P.E. Mogensen, S.H. Jensen, J.B. Andersen, F. Frederiksen, T. B. Sorensen, and K.I. Pedersen. Experimental Investigation of Multipath Richness for Multi-Element Transmit and Receive Antenna Arrays. In *IEEE Vehicular Technology Conference VTC 2000 Spring, Tokyo, Japan*, volume 3, pages 2004–2008, May 2000.

[KSM+05] T. Koike, Y. Seki, H. Murata, S. Yoshida, and K. Araki. FPGA Implementation of 1Gbps Real-time 4×4 MIMO-MLD. In *IEEE VTC Spring, Stockholm, Sweden.*, May 2005.

[KSS99] H. R. Karimi, M. Sandell, and J. Salz. Comparison between transmitter and receiver array processing to achieve interference nulling and diversity. In *IEEE PIMRC*, pages 997–1001, 1999.

[KWV00] P. Kyritsi, P. Wolniansky, and R. Valenzuela. Indoor BLAST Measurements: Capacity of Multi-element Antenna Systems. *MMT*, 2000.

[Lam04] Matthias Lampe. *Adaptive Techniques for Modulation and Channel Coding in OFDM Communication Systems*. PhD thesis, Technical University Hamburg-Harburg, Germany, April 2004.

[Lap96] A. Lapidoth. Nearest Neighbor Decoding for Additive non-Gausssian Noise Channels. *IEEE Trans. on Information Theory*, 42(5):1520–1529, Sept. 1996.

[LCW+01] J. Ling, D. Chizhik, P. Wolniansky, R. Valenzuela, N. Costa, and K. Huber. Multiple transmit multiple receive (mtmr) capacity survey in manhattan. *IEE Electronic Letters*, 37(16):1041–1042, Aug 2001.

[LFV01a] A. Lozano, F.R. Farrokhi, and R.A. Valenzuela. Lifting the Limits on High-Speed Wireless Data Access Using Antenna Arrays. *IEEE Communications Letters*, Sept. 2001.

[LFV01b] A. Lozano, F.R. Farrokhi, and R.A. Valenzuela. Link Spectral Efficiency of BLAST in Ricean Wireless Channels. *IEEE Communications Letters*, 2001.

[LGF02] G. Lebrun, J. Gao, and M. Faulkner. MIMO Transmission over a Time-Varying Channel Using SVD. In *IEEE Globecom, Taipei, Taiwan*, 2002.

[LNG04] J. Li, K.R. Narayanan, and C.N. Georghiades. Product Accumulate Codes: A Class of Capacity Approaching Codes with Low Decoding Complexity. *IEEE Tran. Info Theory*, Jan. 2004.

[LT04] J. Lin and E. Tsui. Joint Adaptive Transmitter/Receiver IQ Imbalance Correction for OFDM Systems. In *IEEE PIMRC, Barcelona, Spain*, 5/8-th Sept. 2004.

[LTS00] B. Levine, R.R. Taylor, and H. Schmit. Implementation of Near Shannon Limit Error-Correcting Codes Using Reconfigurable Hardware. In *8th IEEE Symposium on Field-Programmable Custom Computing Machines (FCCM 2000), Napa Valley, CA, U.S.A.*, 17th-19th April 2000.

[LZG04] Q. Liu, S. Zhou, and G. B. Giannakis. Cross-Layer Combining of Adaptive Modulation and Coding with Truncated ARQ over Wireless Links. *to appear in IEEE Transactions on Wireless Communications*, 2004.

[Man05] Ioannis Maniatis. *Joint Channel Estimatiom in Service Area based OFDM Air Interfaces for beyond 3G Mobile Radio Systems*. PhD thesis, Technical University Kaiserslautern, Germany, Feb. 2005.

[MBFK00] D.P. McNamara, M.A. Beach, P.N. Fletcher, and P. Karlsson. Capacity Variations of Indoor Multiple-Input Multiple-Output Channels. *Electronic letters*, 36(24):2037–2038, Nov. 2000.

[MDHF03] C. Mutti, D. Dahlhaus, T. Hunziker, and M. Foresti. Bit and Power Loading Procedures for OFDM Systems with Bit-interleaved Coded Modulation. In *IEEE COST273 workshop, Paris, France*, pages 1422–1427, 2003.

[Min04] Mobilkommunikation und Rundfunk der Zukunft: Konzepte zur Minimierung der Exposition der Bevölkerung durch elektromagnetische Felder. Technical report, Herausgeber: DLR in Auftrag des BMBF, 2004.

[MO79] A.W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Application*, volume Vol. 143 of *Mathematics in Science and Engineering*. Academic Press, Inc. (London) Ltd., 1979.

[NMR03] M.J. Neely, E. Modiano, and C.E. Rohrs. Power Allocation and Routing in Multibeam Satellites with Time-Varying Channels. *IEEE Trans. on Networking*, 11(1):138–152, Feb. 2003.

[NWK+05] S. Nanda, R. Walton, J. Ketchum, M. Wallace, and Inc. S. Howard, Qualcomm. A High Performance MIMO OFDM Wireless LAN. *IEEE Communications Magazine*, Feb. 2005.

[PC04] J. Penketh and M. Collados. Performance and Implementation Complexity of Receiver Algorithms for MIMO-OFDM Based Wireless LAN Systems. In *IEEE PIMRC, Barcelona, Spain*, 5/8-th Sept. 2004.

[PCL03] D.P. Palomar, J.M. Cioffi, and M.A. Lagunas. Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization. In *IEEE Transactions on Signal Processing*, volume 51, pages 2381–2401, Sept. 2003.

[PFL00] D.P. Palomar, J.R. Fonollosa, and M.A. Lagunas. MMSE Joint Detection in frequency-selective wireless communication channels for DS-CDMA systems. In *IEEE 6th Int. Symp. on Spread-Spectrum Tech. & Appli., NJIT, N.J., U.S.A.*, 6-8 Sept. 2000.

[PTB03] A.S.Y. Poon, D.N.C. Tse, and R.W. Brodersen. An Adaptive Multi-Antenna Transceiver for Slowly Flat Fading Channels. In *IEEE Trans. Communications*, volume 51, pages 1820–1827, Nov. 2003.

[PTVF92] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C, The Art of Scientific Computing.* Cambridge University Press, 2nd edition, 1992.

[PV02] N. Prasad and M.K. Varanasi. Analysis of the Decision Feedback Detection for MIMO Rayleigh Fading Channels and Optimum Allocation of Transmitter Powers and QAM Constallations. *Draft*, March 2002.

[QUA04] QUALCOMM. Proposal for SVD based Channel Tracking In *Standardization IEEE802.11n Interim meeting Berlin*, Sept. 2004.

[RC98] G.G. Raleigh and J.M. Cioffi. Spatio-temporal coding for wireless communication. *IEEE Trans. on Communications*, 46(3):357–366, March 1998.

[Sal85] J. Salz. Digital Transmission Over Cross-Coupled Linear Channels. *AT&T Technical Journal, U.S.A.*, 64(6), 1985.

[SB95] J.W. Silverstein and Z.D. Bai. On the Empirical Distribution of Eigenvalues of a Class of Large Dimensional Random Matrices. *Journal of Multivariate Analysis*, 54(2):175–192, 1995.

[SB02a] M. Schubert and H. Boche. Joint 'Dirty Paper' Pre-Coding and Downlink Beamforming. *IEEE ISSSTA, Prague, Czech Republic*, September 2002.

[SB02b] M. Schubert and H. Boche. A Unifying Theory for Uplink and Downlink Multi-User Beamforming. *International Zürich Seminar*, 2002.

[SBH+02] D.D. Stancil, A. Berson, J.P. Van't Hof, R. Negi, S. Sheth, and P. Patel. Doubling wireless channel capacity using co-polarized, co-located electric and magnetic dipoles. *Electronics Letters*, 38(14):746–747, 2002.

[Sch02] M. Schubert. *Power-Aware Spatial Multiplexing with Unilateral Antenna Cooperation.* PhD thesis, Technical University of Berlin, Dec. 2002.

[Sch04] S. Schiffermüller. Effiziente Implementierung von MIMO-Algorithmen für die Echtzeitübertragung in mobilen Funksystemen. Master's thesis, Technical University of Berlin, 2004.

[SCP+04] J. Siew, J. Coon, R. Piechocki, A. Dowler, A. Nix, M. Beach, S. Armour, and J. McGeehan, "A Channel Estimation Algorithm for MIMO-SCFDE," *IEEE Communication Letters*, vol. 8, no. 9, pp. 555–557, Sept. 2004.

[Sze05] Aydin Szezgin. *Space-Time Codes for MIMO Systems: Quasi-Orthogonal Design and Concatenation.* PhD thesis, Technical University Berlin, Germany, June 2005.

[SFGK99] D. Shiu, G.J. Foschini, M.J. Gans, and J.M. Kahn. Fading Correlation and Its Effect on the Capacity of Multi-Element Antenna Systems. *IEEE Intl. Conf. on Universal Pers. Commun.*, pp. 429-433, Florence, Italy, October 5-9, 1998.

[Sha48] C. E. Shannon. *A mathematical theory of communication.*, volume 27, chapter Bell System Technical Journal, pages 379–423 and 623–656. July / Oct. 1948.

[SJZO05] D. Samuelsson, J. Jaldén, P. Zetterberg, and B. Ottersten. Realization of a Spatially Multiplexed MIMO System. *EURASIP Journal on Applied Signal Processing*, March 2005.

[SL03] H. Shin and J.H. Lee. Capacity of Multiple-Antenna Fading Channels: Spatial Fading Correlation, Double Scattering, and Keyhole. *IEEE Trans. on Information Theory*, 49(10):2636–2647, Oct. 2003.

[SO00] R. Stridh and B. Ottersen. Spatial Characterization of Indoor Radio Channel Measurements at 5 GHz. *First IEEE Sensor Array and Multichannel Signal Processing Workshop*, March 2000.

[SS04] A. Seyedi and G.J. Saulnier. A CDM based Robust Bit-Loading Algorithm for Wireless OFDM Systems. In *IEEE VTC Fall, Los Angeles, U.S.A.*, Sept. 2004.

[TAG99] X. Tang, M.S. Alouini, and A.J. Goldsmith. Effect of Channel Estimation Errors on M-QAM BER Performance in Rayleigh Fading. *IEEE Trans. Communication*, 47:1856–1864, Dec. 1999.

[Tel95] E. Telatar. Capacity of multi-antenna Gaussian channels. Technical report, AT & T Bell Labs Internal Technical Memorandum, June 1995.

[Tel99] E. Telatar. Capacity of multi-antenna Gaussian channels. *European Transactions on Telecommunications (ETT)*, 10(6):585–595, Dec. 1999.

[TH98] D.N.C. Tse and S.V. Hanly. Multiaccess Fading Channels-Part I: Polymatroid Structure, Optimal Resource Allocation and Throughput Capacities. *IEEE Trans. on Information theory*, 44(7):2796–2814, Nov. 1998.

[TH99] D. Tse and S. Hanly. Linear multiuser receivers: Effective interference, effective bandwidth and user capacity. *IEEE Transactions on Information Theory*, 45(2):641–657, March 1999.

[TLF03] Y. Tan, G. Lebrun, and M. Faulkner. An Adaptive Channel SVD Tracking Strategy in Time-varying TDD System. In *IEEE VTC Spring, Jeju, South Korea*, April 2003.

[Tom71] M. Tomlinson. New Automatic Equaliser Employing Modulo Arithmetic. *Electronics Letters*, 7:138–139, March 1971.

[Tre71] H.L. Van Trees. *Detection, Estimation, and Modulation Theory*, volume I. John Wiley & Sons, Inc., New York, NY, U.S.A., 1971.

[TT98] E. Telatar and D. Tse. Capacity and Mutual Information of Broadband Multipath Fading Channels. In *IEEE ISIT, Cambridge, MA. USA*, Aug. 1998.

[TV01] D. Tse and P. Viswanath. Downlink-uplink duality and effective bandwidth. Oct. 21th 2001.

[TV04] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication.* to be published by Cambridge University Press, electronic preprint on http://degas.eecs.berkeley.edu/ dtse/ edition, 2004.

[VBW98] L. Vandenberghe, S. Boyd, and S.-P. Wu. Determinant Maximization with linear Matrix Inequality Constraints. *SIAM Journal on Matrix Analysis and Applications*, 19(2):499–533, 1998.

[Ven02] T. Vencel. Tomlinson-Harashima Vorcodierung für Mehrkanal- und Mehrbenutzersysteme. Diploma thesis (in German), University Erlangen-Nürnberg, March 2002.

[Ver02] S. Verdú. Spectral efficiency in the wideband regime. *(Invited Paper) IEEE Trans. Information Theory, Special Issue on Shannon Theory: Perspective, Trends and Applications*, 48(6):1319–1343, June 2002.

[VG97] M.K. Varanasi and T. Guess. Optimum decision feedback multiuser equalization with successive decoding achieves the total capacity of the gaussian multiple-access channel. In *Proc. Asilomar Conf. on Signals, Systems and Computers, Monterey, USA*, pages 1405–1409, Nov 1997.

[VJG03] S. Vishwanath, N. Jindal, and A. Goldsmith. Duality, Achievable Rates and Sum-Capacity of Gaussian MIMO Broadcast Channels. *IEEE Trans. on Information Theory*, 49(10):2658–2668, Oct. 2003.

[VT03] P. Viswanath and D. Tse. Sum Capacity of the Multiple Antenna Gaussian Broadcast Channel and Uplink-Downlink Duality. *IEEE Trans. on Information Theory*, 49(8):1912–1921, August 2003.

[VWFin] T. Vencel, C. Windpassinger, and R. Fischer. Sorting in the VBLAST Algorithm and Loading. *IASTED International Conference on Communication Systems and Networks*, 9-12 September 2002, Malaga, Spain.

[WB04a] M. Wiczanowski and H. Boche. Optimal Transmit Covariance Matrices for High Speed Uplink Packet Access. In *Wireless Communications and Networking Conference (WCNC) 2004*, March 2004.

[WB04b] M. Wiczanowski and H. Boche. Optimal Transmit Covariance Matrices for MIMO High Speed Uplink Packet Access. In *IEEE WCNC, Atlanta, U.S.A.*, 21-25 March 2004.

[WBKK03] D. Wuebben, R. Boehnke, V. Kuehn, and K.D. Kammeyer. MMSE Extension of V-BLAST based on Sorted QR Decomposition. In *IEEE VTC2003-Fall, Orlando, Florida, USA*, 2003.

[WBKK04a] D. Wübben, R. Böhnke, V. Kühn, and K.D. Kammeyer. Near-Maximum-Likelihood Detection of MIMO Systems using MMSE-Based Lattice Reduction. In *IEEE ICC ,Paris, France*, pages 798–802, 20th-24th June. 2004.

[WBKK04b] D. Wübben, R. Böhnke, V. Kühn, and K.D. Kammeyer. On the Robustness of Lattice-Reduction Aided Detectors in Correlated MIMO Systems. In *IEEE VTC-Fall ,Los Angeles, California, USA*, 26th-29th Sept. 2004.

[WBR+01] D. Wübben, R. Böhnke, J. Rinas, V. Kühn, and K.D. Kammeyer. Efficient algorithm for decoding Layered Space-Time Codes. *IEE Electronic letters*, 37(22):1348–1350, Oct. 2001.

[WF03] C. Windpassinger and R.F.H. Fischer. Low-Complexity Near-Maximum-Likelihood Detection and Precoding for MIMO Systems using Lattice Reduction. In *IEEE Information Theory Workshop, Paris, France*, 2003.

[WF04a] M. Windisch and G. Fettweis. Adaptive I/Q Imbalance Compensation in Low-IF Transmitter Architectures. In *VTC'04 Fall,Los Angeles, U.S.A.*, 26th-29th Sept. 2004.

[WF04b] M. Windisch and G. Fettweis. Blind I/Q Imbalance Parameter Estimation and Compensation in Low-IF Receivers. In *1st International Symposium on Control, Communications and Signal Processing (ISCCSP), Hammamet, Tunisia*, 21.-24th March 2004.

[WF04c] M. Windisch and G. Fettweis. Standard-Independent I/Q Imbalance Compensation in OFDM Direct-Conversion Receivers. In *9th International OFDM Workshop (InOWo), Dresden, Germany*, 15th-16th Sept. 2004.

[WFGV98] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela. V-BLAST:an architecture for realising very high data rates over the rich-scattering wireless channel. In *Proceedings URSI International Symposium on Signals, Systems, and Electronics., (IEEE, New York, NY, USA)*, pages 295–300, Sept. 1998. invited paper.

[Wie49] Norbert Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications*. The Technology Press of the Massachusetts Institute of Technology and John Wiley & Sons, Inc., 1949.

[Win98] J.H. Winters. The Diversity Gain of Transmit Diversity in Wireless Systems with Rayleigh Fading. 47(1):119–123, Feb. 1998.

[Win04] C. Windpassinger. *Detection and Precoding for Multiple Input Multiple Output Channels*. PhD Thesis, University of Erlangen, 2004.

[WJ01] W.Yu and J.M.Cioffi. Trellis Precoding for the Broadcast Channel. In *IEEE Globecom, San Antonio, USA*, Nov. 2001.

[WM03] T. Weber and M. Meurer. Optimum Joint Transmission: Potentials and Dualities. In *IEEE WPMC*, volume 1, pages 79–83, Oct. 2003.

[WR71] J.H. Wilkinson and C. Reinsch. *Linear Algebra, Handbook of Automatic Computation*, volume 2. New York: Springer Verlag, 1971.

[WVF02] C. Windpassinger, T. Vencel, and R. Fischer. Optimizing MIMO DFE for Systems with spatial Loading. *IEE Elec. Letters*, 38(24):1591–1592, 21. Nov. 2002.

[Xio00] Fuqin Xiong. *Digital modulation techniques.* II. Artech House, Inc., Norwood, MA 02062, 2000.

[YC03] E.M. Yeh and A. Cohen. Throughput and Delay Optimal Resource Allocation in Multiple Fading Channels. In *DIMACS Network Information Theory Workshop*, 18th March 2003.

[Yla03] T.M. Ylamurto. Frequency Domain IQ Imbalance Correction Scheme for Orthogonal Frequency Division Multiplexing (OFDM) Systems. In *IEEE WCNC, New Orleans, USA*, pages 20–25, March 2003.

[ZLH42] H. Zheng, A. Lozano, and M. Haleem. Multiple ARQ Processes for MIMO Systems. *EURASIP JASP*, 5:772–782, 2004 DOI: 10.1155/S1110865704311042.

[ZPBF04] E. Zimmermann, P. Pattisapu, P.K. Bora, and G. Fettweis. Reduced Complexity LDPC Decoding using Forced Convergence. In *7th International Symposium on Wireless Personal Multimedia Communications (WPMC04), Abano Terme, Italy*, Sept. 2004.

[ZT03] L. Zheng and D.N.C. Tse. Diversity and Multiplexing: A Fundamental Tradeoff in Multiple Antenna Channels. *IEEE Trans. on Information Theory*, 49(5), May 2003.