Space-Time Codes for MIMO Systems: Quasi-Orthogonal Design and Concatenation

von Diplom-Ingenieur
Aydin Sezgin
aus Kemah

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. Thomas Sikora
Gutachter: Prof. Dr. Dr. Holger Boche
Gutachter: Prof. Dr. Arogyaswami Paulraj (Stanford University)

Tag der wissenschaftlichen Aussprache: 15.Juni 2005

Berlin 2005
D 83

Zusammenfassung

Der Nachfrage an Mobilfunksystemen mit hoher Datenrate und Übertragungsqualität für eine Vielfalt von Anwendungen ist in den letzten Jahren dramatisch gestiegen. Zur Deckung des hohen Bedarfs werden jedoch neue Konzepte und Technologien benötigt, die den Beeinträchtigungen des Mobilfunkkanals entgegenwirken oder sich diese zu Nutze machen und die knappen Ressourcen wie Bandbreite und Leistung optimal ausnutzen. Eine effiziente Maßnahme zur Erhöhung der Performanz stellen Mehrantennensysteme dar. Um das große Potenzial von solchen Mehrantennensystemen auszunutzen, wurden neue Sendestrategien, so genannte Raum-Zeit Codes entworfen und analysiert, die neben der zeitlichen und spektralen auch die räumliche Komponente ausnutzen sollen. In dieser Arbeit wird die Leistungsfähigkeit solcher Raum-Zeit Codes zunächst isoliert und später, im zweiten Teil der Arbeit, in Kombination mit herkömmlichen Kanalcodierungsverfahren untersucht.

Im ersten Abschnitt, d.h. im Fall ohne herkömmliche Kanalcodierung liegt der Fokus auf diversitäts-orientierten Raum-Zeit Codes. Zunächst werden basierend auf den Raum-Zeit Codes mit orthogonaler Struktur (OSTBC) Raum-Zeit Codes mit quasi-orthogonaler Struktur für eine beliebige Anzahl von Sende-und Empfangsantennen entworfen. Aus der Konstruktion resultieren dann zwei Gruppen von Codes.

Die wesentliche Charakteristik der ersten Gruppe ist es, dass sie Verbindungen mit hoher Qualität gewährleistet. Dies wird erreicht, indem räumliche und zeitliche Redundanz eingebracht wird und daraus die volle Diversität (entspricht dem maximalen Abfall der Bitfehlerratenkurve) resultiert. Volle Diversität wird auch von den OSTBC erreicht, die aufgrund ihrer Struktur den matrix-wertigen Kanal für Mehrantennensysteme, so genannte Multiple-Input-Multiple-Output (MIMO)-Kanäle in parallele skalare Ersatzkanäle, so genannte Single-Input-Single-Output (SISO)-Kanäle, transformieren. Die Anzahl der parallelen Ersatzkanäle entspricht dabei der Anzahl der Sendeantennen. Diese Erkenntnis und die Einsicht in die Eigenschaften dieser Ersatzkanäle waren ein wichtiger Meilenstein und ermöglichten es, die Leistungsfähigkeit der OSTBC zu analysieren. Die Bestimmung der Ersatzkanalstuktur ist daher auch hier von zentraler Bedeutung. Im Falle von Raum-Zeit Codes mit quasi-orthogonaler Struktur wird in dieser Arbeit gezeigt, dass der MIMO-Kanal in einen block-diagonalen MIMO-Kanal zerlegt wird, dessen Eigenvektoren konstant und Blöcke identisch sind. Weiterhin konnte gezeigt werden, dass die Eigenwerte von jedem Block voneinander unabhängig sind und einer nichtzentralen Chi-Quadrat-Verteilung mit einer Anzahl von Freiheitsgraden, die dem Vierfachen der Anzahl der Empfangsantennen entspricht, folgen.

Durch Lockerung der Anforderung von voller Diversität an die zu entwerfenden Codes gelangt man zu der zweiten Gruppe der Raum-Zeit Codes mit quasiorthogonaler Stuktur, welche eine Verallgemeinerung der OSTBC darstellen. Insbesondere wird in dieser Arbeit gezeigt, dass nicht nur das Alamouti-Schema, ein OSTBC für zwei Sendeantennen, sondern auch eine verallgemeinerte Version dieses Alamouti-Schemas, die Kapazität im Falle einer Empfangsantenne erreicht. Die in dieser Arbeit entworfenen Raum-Zeit Codes werden schließlich hinsichtlich ihrer Fehlerraten-Performanz und ihrer spektralen Effizienz mit optimalen als auch mit suboptimalen Empfängerstrukturen analysiert.

Im zweiten Teil dieser Arbeit werden verschiedene Raum-Zeit Codes mit herkömmlichen Kanalcodierungsverfahren kombiniert. Dabei werden neue Empfängerstrukturen vorgestellt und die Leistungsfähigkeit der Raum-Zeit Codes mit iterativen Algorithmen zur so genannten Soft-Input-Soft-Output-Decodierung mit Hilfe von neuen Analysetechniken, den so genannten EXIT-Charts, untersucht und optimiert. Im Falle von OSTBC werden zusätzlich Kriterien für die optimale Abbildung von Bitsequenzen auf Sendesymbole hergeleitet.

Abstract

The demand for mobile communication systems with high data rates and improved link quality for a variety of applications has dramatically increased in recent years. New concepts and methods are necessary in order to cover this huge demand, which counteract or take advantage of the impairments of the mobile communication channel and optimally exploit the limited resources such as bandwidth and power. Multiple antenna systems are an efficient means for increasing the performance. In order to utilize the huge potential of multiple antenna concepts, it is necessary to resort to new transmit strategies, referred to as Space-Time Codes, which, in addition to the time and spectral domain, also use the spatial domain. The performance of such Space-Time Codes is analyzed in this thesis with and without conventional channel coding strategies.

In case without conventional channel codes, the focus is on diversity-oriented Space-Time Codes. Based on Space-Time Block Codes from orthogonal designs (OSTBC), the Space-Time Block Codes from quasi-orthogonal designs are developed for any number of transmit and receive antennas. The outcome of this construction are two groups of codes.

The main property of the first group is the support of links with high quality. This is achieved by incorporating spatial and temporal redundancy, which results in full diversity or in other words, in the maximum decay of the bit error rate curves. Full diversity is also achieved by OSTBC, which due to their structure transform the matrix-valued channel for multi-antenna systems, so called multiple-input-multiple-output (MIMO)-channels, into several parallel, scalar single-input-single-output (SISO)-channels. This insight and the understanding of the properties of the equivalent SISO-channels were the key results in order to analyze the performance of the OSTBC. The determination of the structure of the equivalent channel is also a matter of vital importance in this work. To this end, we show that the MIMO-channel in the case of Space-Time Codes from quasi-orthogonal designs is transformed into an equivalent block-diagonal MIMO-channel with identical blocks having constant eigenvectors, independent of the channel realization. Furthermore, we show that the eigenvalues of each block are pairwise independent and follow a non-central chi-square distribution, where the number of degrees of freedom equals four times the number of receive antennas.

By relaxing the requirement of full diversity one arrives at the second group of Space-Time Codes from quasi-orthogonal designs. These codes represent a generalization of Space-Time Codes from orthogonal designs. Particularly, we show in this work, that not only the Alamouti-scheme, a OSTBC for two transmit antennas, but also its generalized version achieves capacity in the case of one receive antenna. The drafted codes are then analyzed with respect to the error rate performance and the spectral efficiency with optimal as well as suboptimal receiver structures.

In the second part of this work the combination of Space-Time Codes with conventional channel coding techniques is considered. New receiver structures are presented and the performance of Space-Time Codes with iterative algorithms for soft-input-soft-output-decoding is analyzed and optimized with the help of new analytical tools, the so called EXIT-charts. Furthermore, some criteria for the optimal

mapping strategy are derived in the case of OSTBC.

Acknowledgements

During my PhD studies, I had the opportunity to meet many interesting and extraordinary individuals. First of all, I would like to thank my doctoral supervisor Professor Holger Boche, who render possible to start my PhD study and for the confidence shown to me throughout the whole time. I would also like to express my gratitude to Professor Arogyaswami Paulraj who served in my thesis committee.

I wish to thank Eduard A. Jorswieck for sharing ideas and room with me during our PhD studies. Working with him was a positive experience on both the professional and personal level. Further, special thanks should be given to Tobias J. Oechtering for the intensive discussions and for being my most challenging reviewer.

My gratitude extends to Thomas Haustein, Oliver Henkel and Peter Jung for discussions, collaborating and co-authoring papers with me. Thanks are also due to all colleagues at the Heinrich-Hertz-Institut (HHI), the Sino-German Lab for Mobile Communications (MCI) and the Technical University of Berlin for stimulating discussions and additional support.

Contents

List of Figures

List of Tables

Abbreviations and nomenclature

$[\cdot]^H$ Matrix Conjugate Transpose, page 3

$[\cdot]^T$ Matrix Transpose, page 3

$[\cdot]^{-1}$ Matrix Pseudo Inverse, page 3

$\chi^2_n(\delta_{nc})$ chi-square distribution with n degrees of freedom and noncentrality parameter δ_{nc}, page 20

$\Im\{\cdot\}$ imaginary part of a complex variable, page 66

$\mathbf{A}(\mathbf{c}, \mathbf{e})$ codeword distance matrix between codewords $\mathbf{c}$ and $\mathbf{e}$, page 7

$\mathbf{B}(\mathbf{c}, \mathbf{e})$ codeword difference matrix between codewords $\mathbf{c}$ and $\mathbf{e}$, page 7

$\mathbf{G}_{n_T}$ transmit matrix, page 17

$\mathbf{h}(\,\cdot\,)$ differential entropy, page 5

$\mathbf{I}_n$ Identity matrix of size n, page 5

$\mathcal{CN}(\cdot, \cdot)$ complex normal distribution with mean and covariance as first and second parameter, respectively, page 17

$\mathcal{N}(\cdot, \cdot)$ normal distribution with mean and covariance as first and second parameter, respectively, page 17

$\mathrm{diag}\,(\cdot)$ (block) diagonal matrix, page 18

$\mathrm{rk}(\,\cdot\,)$ Rank of a Matrix, page 27

$\mathrm{tr}(\,\cdot\,)$ Trace of a Matrix, page 3

$\Re\{\cdot\}$ real part of a complex variable, page 66

ρ average signal-to-noise ratio, page 5

ζ diversity product, page 45

$I(\,\cdot;\cdot)$ mutual information, page 5

n_R number of receive antennas, page 4

n_T number of transmit antennas, page 4

P_{out} outage probability, page 6

$Q(x)$ Gaussian tail integration function, page 7

APP a-posteriori-probability, page 11

AS antenna selection, page 9

AWGN additive white Gaussian noise, page 4

BCH Bose-Chaudhuri-Hocquenghem, page 116

BER bit-error-rate, page 9

BLAST Bell Labs Architecture Layered Space-Time, page 94

CC convolutional code, page 104

CSI channel state information, page 11

DBLAST Diagonal BLAST, page 95

DVB-T Digital Video Broadcasting - Terrestrial, page 141

EMI Ergodic mutual information, page 34

EXIT extrinsic information transfer, page 11

FEC forward error correction code, page 96

GPRS General Packet Radio Service, page 2

GSM Global System for Mobile Communications, page 2

i.i.d. independent identically distributed, page 14

ISIP Iterative Signal Processing, page 95

LOS line of sight, page 15

LR lattice reduction, page 9

MIMO multiple-input multiple-output, page 3

ML maximum likelihood, page 8

MMS Multimedia Message Serive, page 1

MMSE minimum mean-square error, page 11

MRC Maximum-Ratio-Combining, page 99

OFDM Orthogonal Frequency Division Multiplexing, page 142

OMI outage mutual information, page 16

OSTBC orthogonal space-time block codes, page 8

pdf probability density function, page 14

PEP pairwise error probability, page 6

PSK phase-shift keying, page 17

PSP per-survivor processing, page 130

QAM Quadrature Amplitude Modulation, page 42

QoS Quality of Service, page 141

QSTBC quasi-orthogonal space-time block codes, page 8

SCCC Serially concatenated convolutional codes, page 128

SISO single-input single-output, page 4

SISO soft-input soft-output, page 11

SM spatial multiplexing, page 10

SMS Short Message Service, page 1

SNR average signal-to-noise ratio, page 5

SOVA soft-output Viterbi-algorithm, page 11

SPC single parity check code, page 122

STC space-time codes, page 6

SVD singular value decomposition, page 19

TC Turbo code, page 126

UMTS Universal Mobile Telecommunications System, page 2

VBLAST Vertical BLAST, page 95

WSTC-ID WSTC with iterative decoding, page 133

ZF zero-forcing, page 10

1 Introduction

1.1 Motivation

The field of wireless communication systems and networks has experienced explosive growth and wireless communications has become an important part of everyday live. Further, the rapidly increasing number of wireless communication subscribers, the growth of the internet and the quickly increasing use of wireless devices suggest that wireless internet multimedia access will rise rapidly over the next few years. The demand and purchase of wireless telephones is predicted to soon exceed the purchase and use of traditional wired telephones. In some developing countries like China, India and also in many countries in Africa, the infrastructure of wireless communication systems is more sophisticated in comparison to wired communication systems. This comes from the fact that new telephone cables have to be laid at great expense for wired communication systems, which is not necessary for mobile systems. As mentioned above, the market for mobile devices has dramatically increased and continued growth is predicted. Along with this rapid growth comes the costumer demand for more and better applications, improved performance, and increased data rates. They want the ability to communicate on their own terms; to get connected and stay connected in order to send and receive information in any form, let it be voice, text, image, or video. As an example, customers are using mobile telephone applications like Multimedia Message Services (MMS), an extension of text messaging (SMS), that adds pictures and sound elements. In short, they want the ability to rely on a wireless device to improve and add diversity to traditional ways and forms of communication by connecting them to the mobile services they want and need anytime and anywhere.

All these improvements must be accomplished under a considerable number of constraints. Wireless channels are by its nature random and unpredictable and results therefore in uncontrolled reflection, scattering, shadowing and attenuation of the transmitted signal. Due to the constructive and destructive superposition of different signal waves at the receiver, it may be infeasible to detect the transmit signal correctly. These effects can be statistically modeled as a multiplicative random variable and are referred to as fading. The spectrum or bandwidth available to the service provider is often limited and expensive. For example, at the auction of the licences for the new frequencies and radio spectrum of the third generation (3G) mobile technology UMTS (Universal Mobile Telecommunications System) in Germany, the wireless communication systems operators paid about 50 billions euro. Furthermore, the power requirements are such that devices should use as little power as possible to conserve battery life and keep the products small and cheap. This should also apply to next generation handset models of which many have built-in cameras. Furthermore, most models will be multi-band and multi-mode, allowing users to switch seamlessly between different services in various mobile technologies like UMTS, GPRS (General Packet Radio Service), and GSM (Global System for Mobile Communications) in different frequency bands. Designers of wireless systems therefore face a two-part challenge of increasing data rates and improving performance while incurring little or no increase in bandwidth or power, and costs.

This thesis provides an analysis of new transmission schemes referred to as space-time codes in order to guarantee reliable transmission and improved performance

in mobile communication systems. The first part is devoted to uncoded (regarding channel coding) space-time codes or space-time transmission schemes and their performance with respect to data rate and error probability with different optimal and suboptimal detection schemes. In the second part the space-time transmission schemes are combined with conventional channel codes like convolutional or block codes and their performance is analyzed with respect to iterative detection and decoding using information theoretic techniques.

1.2 Notation

Vectors are denoted in bold small letters, e.g. $\mathbf{x}$. Matrices are written in bold capital letters, e.g. $\mathbf{A}$. Transpose is $[\,\cdot\,]^T$, the conjugate transpose is $[\,\cdot\,]^H$. The matrix (pseudo) inverse is denoted by $[\,\cdot\,]^{-1}$. $\mathrm{tr}(\mathbf{A})$ denotes the trace of the $[n \times n]$ matrix $\mathbf{A}$, i.e. $\mathrm{tr}(\mathbf{A}) = \sum_{k=1}^{n} \mathbf{A}_{k,k}$. $\mathrm{rk}(\mathbf{A})$ denotes the rank of the matrix $\mathbf{A}$.$\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary part of a complex variable, respectively. $\mathrm{diag}(\mathbf{x})$ is a matrix with entries of the vector $\mathbf{x}$ on the diagonal. Furthermore, $\mathrm{diag}(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n)$ is a block-diagonal matrix with the matrices $\mathbf{X}_i$, $1 \leq i \leq n$ on its diagonal.

1.3 Diversity

In fading environments, several different diversity sources have been discovered that can improve the performance over wireless channels under varying fading conditions by providing multiple replicas of a transmitted signal to the receiver such that reliable communication is possible. These methods, therefore, reduce the probability that all the replicas are simultaneously affected by severe attenuation and noise. Commonly used diversity techniques are:

- Temporal diversity, in which redundancy is introduced with channel coding and interleaving methods.
- Frequency diversity, in which replicas of the signal are transmitted over different frequencies in order to introduce redundancy in the frequency domain.
- Antenna or spatial diversity, in which multiple antennas at the transmitter and/or receiver are employed in order to independently observe the transmitted signal at the receiver without any loss in bandwidth efficiency.

The latter source of diversity is very promising, since it does not increase the transmit power and the signal bandwidth and can be efficiently utilized in multiple-input multiple output (MIMO) systems, i.e. systems with multiple transmit and multiple receive antennas.

1.4 MIMO

In this section, we describe briefly the performance of single-user, point-to-point wireless links by employing MIMO technology. In multi-antenna systems with n_T transmit and n_R receive antennas, the data is sent simultaneously and synchronously from the transmit antennas. The signal received at each antenna is therefore a superposition of the n_T transmitted signals corrupted by additive noise and multiplicative fading. From the traditional point of view, the interference the received signals experience from each other would be a limiting factor for reliable

communication. From an information-theoretic point of view, one may view the system under consideration as providing not one, but $n_T \times n_R$ potential communication links between the transmitter and receiver, corresponding to each distinct transmit/receive antenna pairing. The improvements in comparison to single-input single-output (SISO) systems is not only the diversity provided by MIMO over fading channels. Information theoretic results have demonstrated that the ability of a system to support high link quality and higher data rates in the presence of Rayleigh fading improves significantly with the use of multiple transmit and receive antennas [Win87, Tel99, FG98] as described in the following subsections.

1.4.1 Mutual Information of MIMO

It is common to represent the input/output relations of time-discrete and flat fading MIMO links by the following notation [Tel99]

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

where $\mathbf{x}$ is the transmit vector of size $(n_T \times 1)$, $\mathbf{y}$ is the receive vector of size $(n_R \times 1)$, $\mathbf{H}$ is the random channel matrix of size $(n_R \times n_T)$, and $\mathbf{n}$ is the additive white Gaussian noise (AWGN) vector of size $(n_R \times 1)$ for a given time instant. Mutual information, a measure of the amount of information that a random variable contains on another variable, can be written as

$$\begin{aligned} I(\mathbf{x};\mathbf{y}|\mathbf{H}) &= h(\mathbf{y}|\mathbf{H}) - h(\mathbf{y}|\mathbf{x},\mathbf{H}) = h(\mathbf{y}|\mathbf{H}) - h(\mathbf{H}\mathbf{x}+\mathbf{n}|\mathbf{x},\mathbf{H}) \\ &= h(\mathbf{y}|\mathbf{H}) - h(\mathbf{n}|\mathbf{x},\mathbf{H}) = h(\mathbf{y}|\mathbf{H}) - h(\mathbf{n}) \, , \end{aligned} \tag{1.1}$$

with $\mathbf{h}(\cdot)$ and $\mathbf{h}(\cdot|\cdot)$ denoting the differential entropy and conditional differential entropy, respecively. The covariance matrices of $\mathbf{x}$ and $\mathbf{n}$ are given as $\mathbf{E}\left\{\mathbf{x}\mathbf{x}^H\right\} = \mathbf{Q}_\mathbf{x}$ $\mathbf{E}\left\{\mathbf{n}\mathbf{n}^H\right\} = \mathbf{Q}_\mathbf{n}$. Note that in this thesis, we assume that the transmitter has no information on the channel. In this case, it is optimal to use a uniform power distribution [Tel99]. The transmit covariance matrix is then given by $\mathbf{Q}_\mathbf{x} = \frac{P_T}{n_T}\mathbf{I}_{n_T}$, where P_T is the sum power at the transmitter. Further, we assume that there is uncorrelated noise at each receive antenna, which is determined by the covariance matrix $\mathbf{Q}_\mathbf{n} = N_0\mathbf{I}_{n_R}$. Then, the covariance matrix of $\mathbf{y}$ becomes

$$\begin{aligned} \mathbf{Q}_\mathbf{y} &= \mathbf{E}\left\{\mathbf{y}\mathbf{y}^H\right\} = \mathbf{E}\left\{(\mathbf{H}\mathbf{x}+\mathbf{n})(\mathbf{H}\mathbf{x}+\mathbf{n})^H\right\} = \mathbf{E}\left\{\mathbf{H}\mathbf{x}\mathbf{x}^H\mathbf{H}^H\right\} + \mathbf{E}\left\{\mathbf{n}\mathbf{n}^H\right\} \\ &= \mathbf{H}\mathbf{E}\left\{\mathbf{x}\mathbf{x}^H\right\}\mathbf{H}^H + \mathbf{E}\left\{\mathbf{n}\mathbf{n}^H\right\} = \mathbf{H}\mathbf{Q}_\mathbf{x}\mathbf{H}^H + \mathbf{Q}_\mathbf{n} \, , \end{aligned}$$

and the probability density function of $\mathbf{y}$ as a complex multivariate normal distributed variable is

$$p_\mathbf{y}(\mathbf{y}) = \frac{1}{\det(\pi\mathbf{Q}_\mathbf{y})} \exp\left(-\mathbf{y}^H\mathbf{Q}_\mathbf{y}^{-1}\mathbf{y}\right) \, .$$

The differential entropy $h(\mathbf{y}|\mathbf{H})$ is given as

$$\begin{aligned} h(\mathbf{y}|\mathbf{H}) = \; & \mathbf{E}\left\{-\log_2 p_\mathbf{y}\right\} = \log_2\det(\pi\mathbf{Q}_\mathbf{y}) + \mathbf{E}\left\{\mathbf{y}^H\mathbf{Q}_\mathbf{y}^{-1}\mathbf{y}\log_2 e\right\} \\ = \; & \log_2\det(\pi\mathbf{Q}_\mathbf{y}) + \log_2 e \, \mathrm{tr}(\mathbf{Q}_\mathbf{y}^{-1}\mathbf{E}\left\{\mathbf{y}\mathbf{y}^H\right\}) = \log_2\det(\pi e\mathbf{Q}_\mathbf{y}) \, . \end{aligned}$$

Similarly $h(\mathbf{n}) = \log_2\det(\pi e\mathbf{Q}_\mathbf{n})$. Using this formula in (1.1) results in

$$I = \log_2 \frac{\det(\frac{P_T}{n_T}\mathbf{H}\mathbf{H}^H + \mathbf{Q}_\mathbf{n})}{\det(\mathbf{Q}_\mathbf{n})} = \log_2\det(\mathbf{I}_{n_R} + \frac{\rho}{n_T}\mathbf{H}\mathbf{H}^H) = \sum_i^r \log_2(1 + \frac{\rho}{n_T}\lambda_i), \tag{1.2}$$

where $r = \min(n_T, n_R)$, λ_i is the ith eigenvalue of $\mathbf{H}\mathbf{H}^H$, and $\rho = \frac{P_T}{N_0}$ is the average signal-to-noise (SNR) ratio at each receive antenna. Based on (1.2), we present two measures for the spectral efficiency of the MIMO channel. These are the ergodic (mean) capacity given as

$$C = \mathbf{E}\left[\log_2 \det(\mathbf{I}_{n_R} + \frac{\rho}{n_T}\mathbf{H}\mathbf{H}^H)\right],$$

where the expectation $\mathbf{E}[\cdot]$ is over all channel realizations $\mathbf{H}$ and the outage probability which is defined as the probability that the mutual information I is smaller than a certain rate R, i.e.

$$P_{out}(R, n_T, n_R, \rho) = \Pr[I < R].$$

If the mutual information falls below this rate R, it is not possible for the transmitted block of information to be decoded with no errors, regardless of the coding scheme employed.

In order to exploit the huge potential of MIMO systems [Win87, Tel99, FG98] with its spatially independent, but mutually interfering links, the design and analysis of new multiple transmit antenna concepts (as a generalization of the work in [Wit91], which was one of the first works proposing a combination of spatial and temporal domain signal processing) have been initiated in [TSC98, GFBK99]. Tarokh et al. coined the term "space-time codes" (STC) to describe the two-dimensional (2-D) signals (time and space) used in wireless multiple transmit antenna systems. The performance criteria derived in [TSC98] are briefly explained in the following section.

1.4.2 Performance criteria of STC

The union bound on the average codeword error probability can be computed as [SA00]

$$P_{err} \leq \frac{1}{|\mathcal{X}|}\sum_{\mathbf{c}\in\mathcal{X}}\sum_{\mathbf{e}\in\mathcal{X}\setminus\{\mathbf{c}\}} P(\mathbf{c}\to\mathbf{e}),$$

where the pairwise error probability (PEP) $P(\mathbf{c}\to\mathbf{e})$ denotes the probability that the codeword $\mathbf{c}$ is transmitted and the receiver decides erroneously in favor of codeword $\mathbf{e}$ with $\mathbf{c}, \mathbf{e} \in \mathcal{X}$ and $|\mathcal{X}|$ denotes the cardinality of $\mathcal{X}$. The PEP is given as

$$\begin{aligned} P(\mathbf{c}\to\mathbf{e}) &= \mathbf{E}\left[Q\left(\frac{\sqrt{E_s}||\mathbf{H}(\mathbf{c}-\mathbf{e})||}{\sqrt{2N_0}}\right)\right] \\ &= \mathbf{E}\left[\frac{1}{2}\mathrm{erfc}\left(\frac{\sqrt{E_s}||\mathbf{H}(\mathbf{c}-\mathbf{e})||}{\sqrt{2}\sqrt{2N_0}}\right)\right] = \mathbf{E}\left[P(\mathbf{c}\to\mathbf{e}|\mathbf{H})\right] \\ P(\mathbf{c}\to\mathbf{e}) &\leq \mathbf{E}\left[\frac{1}{2}\exp\left(-\frac{E_s||\mathbf{H}(\mathbf{c}-\mathbf{e})||^2}{4N_0}\right)\right], \end{aligned} \tag{1.3}$$

where the expectation is taken over the random fading channel $\mathbf{H}$, $||\cdot||$ denotes the Euclidian norm, and (1.3) is the Chernoff-Bound, which is the standard approximation of the Gaussian tail integration function $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty \exp\left(-\frac{t^2}{2}\right)dt$. After some manipulations [TSC98], particularly taking the expectation over all channel

realizations, we arrive at

$$\begin{aligned} P(\mathbf{c} \to \mathbf{e}) &\leq \frac{1}{2} \prod_{j=1}^{n_R} \left(\prod_{i=1}^{n_T} \frac{1}{1 + \frac{E_s}{4N_0} \lambda_i} \right) = \frac{1}{2} \left(\prod_{i=1}^{r} \frac{1}{1 + \frac{E_s}{4N_0} \lambda_i} \right)^{n_R} \\ &\leq \frac{1}{2} \prod_{i=1}^{r} \lambda_i^{-n_R} \left(\frac{E_s}{4N_0} \right)^{-rn_R} \end{aligned} \tag{1.4}$$

for the Rayleigh fading case, i.e. the entries in the channel matrix $\mathbf{H}$ are following a Rayleigh distribution. The $\lambda_1, \lambda_2, \ldots, \lambda_r$ denote the nonzero eigenvalues of the codeword distance matrix $\mathbf{A}(\mathbf{c}, \mathbf{e}) = \mathbf{B}(\mathbf{c}, \mathbf{e})\mathbf{B}(\mathbf{c}, \mathbf{e})^H$, $r \leq n_T$ is the rank of $\mathbf{A}(\mathbf{c}, \mathbf{e})$, and $\mathbf{B}(\mathbf{c}, \mathbf{e})$ is referred to as the codeword difference matrix. Furthermore, rn_R and $\prod_{i=1}^{r} \lambda_i$, which are referred to as the diversity gain and coding gain respectively, are the performance or design criteria which have to be maximized. In a double-logarithmic plot of the average codeword error probability P_{err} versus the average signal-to-noise ratio (SNR), the diversity gain gives the asymptotic slope of the P_{err} versus SNR graph, i.e. the greater the diversity, the greater the decline of the graph is. The coding gain has an effect on the horizontal shift of the graph in comparison to an uncoded system, i.e. the greater the coding gain, the greater the horizontal shift.

This concludes the informal introduction of the performance metrics for the communication over single-user, point-to point wireless MIMO systems. An outline of the following chapters of the thesis is presented next, including some of the contributions.

1.5 Outline of the thesis

The main contributions of this thesis are divided into two parts, which are specified in the following. In Chapter 2, we generalize the space-time block codes from orthogonal designs (or orthogonal space-time block codes, OSTBC) to the space-time block codes from quasi-orthogonal designs (or quasi-orthogonal space-time block codes, QSTBC) and analyze the performance of a MIMO system employing QSTBC. In more details, we show that

- QSTBC are capable of achieving a significant fraction of the mutual information of multiple-input-multiple output (MIMO) wireless communication systems for the case of 2^n transmit antennas and an arbitrary number of receive antennas. We derive an equivalent channel model, which originates from the application of QSTBC to the MIMO system and completely characterize this channel with respect to the probability density function of the eigenvalues. Furthermore, we derive analytical lower bounds for the outage probability achieved with QSTBC and show that one of these bounds is tight for low signal-to-noise-ratios (SNR) values and also for an increasing number of receive antennas. In addition to this, we present some upper bounds, where one of which is tight for high SNR values, and derive analytical expressions for any number of transmit and receive antennas. For the ergodic mutual information, we also derive some closed-form expressions. Furthermore, by exploiting the special structure of the QSTBC, we propose a new transmit strategy which decouples the signals transmitted from different antennas in order to detect the symbols separately with a linear maximum likelihood (ML)-detector rather than joint detection, an advantage which was previously only known for OSTBC.

- QSTBC are a simple and powerful means of achieving higher transmission rates than OSTBC. Although OSTBC exploit MIMO communication systems to obtain full diversity and therefore high link reliability, unfortunately, it is not possible to construct OSTBC with a transmission rate equal one for more than two transmit antennas. We show that it is possible to achieve full diversity as in the case of OSTBC by using QSTBC with a transmission rate equal one by only a small increase in the complexity of the optimal detector. For this purpose, a constellation rotation method is applied, which rotates the constellation of some symbols in the so called transmit matrix in order to improve the performance of QSTBC. We analyze different performance criteria which can be applied to improve the performance and derive analytical results for the performance of QSTBC. Further on, we analytically derive some expressions for the bit-error-rate (BER) performance of different transmit strategies and compare them with different simulation results.

- A nice advantage of both OSTBC and QSTBC is that deploying multiple antennas at the receiving unit is only optional. Nevertheless, the performance of space-time coding schemes can be improved by using multiple antennas at the receiving units in mobile communication systems. However, multiple antenna deployment at mobile handsets requires multiple RF chains (analog-digital converters, low noise amplifiers, downconverters, etc.), which is undesired in systems where the handsets are supposed to remain simple and inexpensive. In order to reduce hardware costs but still benefit from multiple antennas, we use antenna selection (AS) in combination with QSTBC, where only one RF chain is used and this RF chain is concatenated in an adaptive manner according to different criteria with one receive antenna from all available antennas. We analyze the impact of antenna selection (AS) based on different selection criteria on both the mutual information and the bit error rate for different detector schemes of QSTBC with 2^n transmit antennas. Furthermore, we derive an analytical lower bound for the outage probability and an upper bound for the average mutual information achieved with QSTBC and AS.

- Another application of QSTBC presented in this thesis is the following: The performance of recently proposed lattice-reduction (LR) aided linear detectors is parallel to the maximum-likelihood-detector with only a slight loss in power efficiency. In order to reduce this gap, one can employ non-linear schemes at the receiver like successive interference cancelation with higher complexity. In this thesis, instead of non-linear schemes at the receiver, we apply QSTBC at the transmitter, which are inherently linear schemes, in order to close this gap and achieve full diversity. We analyze the performance of QSTBC with LR-aided linear detection, which further reduces the detection complexity of the QSTBC and compare it with spatial multiplexing (SM) schemes.

- By relaxing the constraint of full diversity, we develop QSTBC capable of achieving high transmission rates. These QSTBC are using again OSTBC as building blocks and represent n-times stacked version of an OSTBC. Here, we focus on the Alamouti scheme, an OSTBC for 2 transmit antennas, as building block, since it is well known, that the Alamouti scheme is the only space-time code from orthogonal design achieving the capacity of MIMO wireless communication systems with $n_T = 2$ transmit antennas and $n_R = 1$ receive antenna. In this work, we show that not only the standard Alamouti scheme but also the more general stacked Alamouti-QSTBC for $n_T = 2n$ transmit antennas achieve the capacity in the case of $n_R = 1$ receive antenna. For the more general case of more than one receive antenna, we show that if the

number of transmit antennas is higher than the number of receive antennas we achieve a high portion of the capacity with these QSTBC.

While in chapter 2, we examined the performance of space-time codes without any conventional channel coding, chapter 3 is concerned with the performance of space-time codes in concatenation with conventional channel codes at the transmitter and iterative detection and decoding at the receiver.

In particular, chapter 3 discusses the following issues:

- We analyze a new iterative signal processing algorithm, which can be seen as an extension of many well-known detection algorithms for MIMO channels, e.g. zero-forcing (ZF) and successive nulling and cancelation algorithms. We apply the algorithm in combination with the VBLAST algorithm proposed by [WFGV98] and the ZF algorithm to an uncoded as well as to a channel coded spatial multiplexing scheme . We show that the new algorithm takes advantage of the maximum available diversity level to improve the initial data estimate obtained by the nulling and cancelation (or ZF) algorithm with a minimum additional effort.
- In order to enhance the performance of channel coded space-time codes, we allow the detector and decoder to exchange soft or extrinsic information in the iterative detection and decoding process. At first, we apply a "turbo" coding scheme consisting of the serially concatenation of an outer code and OSTBC as the inner code. We analyze the impact of different mapping strategies on the information transfer of the soft-input soft-output (SISO) space-time detector and derive criteria for the optimum mapping strategy. We show that additional performance gains can be achieved by mapping strategies other than Gray mapping. Furthermore, we use extrinsic information transfer characteristics (EXIT-charts) in order to predict the performance and the behavior of the system.
- Without the assumption of perfect channel state information (CSI) at the receiver, it is not possible to employ OSTBC as inner codes within this concatenated coding scheme. As an alternative, we apply a unitary space-time coded modulation scheme as the inner code, where the detection process works even if there is no information on the varying channel. We analyze this non-coherent case in the same way as the coherent case in order to gain important insights into the performance of the scheme.
- In some cases, it is not feasible to apply the optimum a-posteriori-probability (APP) decoder in combination with space-time codes due to the processing delay. As an example, we study the iterative decoding of a low-complexity space-time architecture called Wrapped space-time codes employing per-survivor-processing at the receiver with the soft-output Viterbi-algorithm (SOVA). With the availability of extrinsic information at the receiver delivered by the SOVA, it is now possible to use a novel receiver scheme, which performs iterative (turbo) decoding in order to improve the performance of the architecture. According to the ZF or minimum mean-square error (MMSE) criteria, linear feedforward and feedbackward filters of the decision-feedback space-time (ST) decoder are derived. Furthermore, the decision metric of the SOVA, employing per-survivor-processing, is developed and the performance of the scheme is analyzed.

Finally, in chapter 4, we conclude the thesis and give directions for further research. The list of publications and the bibliography completes this thesis.

2 Quasi-Orthogonal Space-Time Codes

Space-time block codes from orthogonal design or orthogonal space-time block codes (OSTBC) for multiple-input multiple-output (MIMO) wireless communications systems are a very powerful means of increasing the reliability of communication links. With their ability to transform the MIMO system into a single-input-single-output (SISO) system and their linear optimal maximum-likelihood (ML) detector, they will play an important role in future communications systems and are already a very attractive and possible applicant for practical systems.

In this chapter, we develop the transmission rate one space-time block codes from quasi-orthogonal design (QSTBC) for 2^n transmit antennas and an arbitrary number of receive antennas on the basis of OSTBC. QSTBC are used to increase the code rate which is able to be achieved with OSTBC, since it is not possible to construct OSTBC with a code rate equal one for more than two transmit antennas. Similar to OSTBC, the QSTBC exploit MIMO communication systems in order to obtain diversity for high link reliability. We completely characterize the structure of the resulting equivalent channel evolving from the employment of QSTBC at the transmitter. Based on these achievements, we consider the performance analysis of QSTBC with respect to the bit-error-rate (BER) for different detectors and the mutual information.

In more detail, in section 2.2.1, we introduce the system model. The design of QSTBC for 2^n transmit antennas is shown in section 2.2.2. The complete characterization of the QSTBC such as the important fact about the decoupling of the system model, the structure of the resulting equivalent channel model with its eigenvectors being independent of the current channel realization and the independency and distribution of the eigenvalues and also the preprocessing at the receiver, are described in section 2.2.3. In section 2.3, we analyze the performance of the QSTBC. From the outage behavior point of view, we derive analytical lower bounds for the outage probability achieved with QSTBC and show that one of these bounds is tight for low signal-to-noise-ratio (SNR) values and also for increasing number of receive antennas. We also present upper bounds, whereby one of these bounds is tight for high SNR values, and derive analytical expressions for an arbitrary number of transmit and receive antennas. For the ergodic mutual information, we also derive some closed-form expressions. Furthermore, an expression for the bit-error-rate is derived and analyzed for a system applying QSTBC . Depending on the transmit strategy, we show that there is a tradeoff between the performance and the complexity of the receiver. More precisely, on the one hand it is possible to achieve the full diversity of $n_T n_R$ of the system, on the other hand the property of simple linear ML-decoding can be obtained as in the case of OSTBC. In addition, we analyze the impact of antenna selection (AS) on both the mutual information and the bit error rate, according to different selection criteria . Finally, the bit-error-rate performance with suboptimal detection schemes is also analyzed.

By relaxing the constraint of full diversity, we arrive at QSTBC capable of achieving high transmission rates, which will be discussed in section 2.4. These QSTBC are again using OSTBC as building blocks and are a, in a stacked fashion, generalization of the well known Alamouti scheme. Of course, other OSTBC can also be used as building blocks, however, we have to keep in mind, that they have inherently a loss

of transmission rate. We analyze the performance in terms of achievable mutual information of this stacked Alamouti scheme. We show, that not only the standard Alamouti scheme but also the stacked Alamouti scheme achieves the capacity of a MIMO system in the case of $n_R = 1$ receive antennas. Furthermore, the stacked Alamouti scheme approaches the outage capacity of a MIMO system by increasing the number of transmit antennas for a fixed number of receive antennas. We also provide an upper bound on the mutual information for the stacked Alamouti scheme for arbitrary number of receive antennas and also an lower bound on the outage probability.

2.1 Introduction

In recent years, the goal of providing high speed wireless data services has generated a great amount of interest among the research community. Recent information-theoretic results have demonstrated that the ability of a system to support high link quality and higher data rates in the presence of Rayleigh fading improves significantly with the use of multiple transmit and receive antennas [Win87, Tel99, FG98]. A very important aspect is always the availability of analytical expressions to describe the stochastic nature of the channel under consideration as given in [Tel99, Ede89] for the MIMO channel. This offers an opportunity to obtain, e.g., closed-form analytical formulas for the ergodic capacity or the outage mutual information of such MIMO channels. E.g., in [WG04], the probability density function (pdf) of the random mutual information for independent identically distributed (i.i.d.) MIMO channels was derived in the form of the inverse Laplace transform and a Gaussian approximation of the pdf was presented.

There has been a considerable amount of work on a variety of new codes and modulation signals, called space-time (ST) codes, in order to approach the huge capacity of such MIMO channels. The performance criteria of space-time codes were derived in [GFBK99, TSC98]. One such ST code is Diagonal BLAST (DBLAST), a rate-oriented space-time transmission or spatial multiplexing scheme which theoretically achieves the capacity for such MIMO channels [Fos96]. Vertical BLAST (VBLAST), a simplified and suboptimal version of the BLAST architecture has been proposed in [WFGV98] in order to reduce the high complexity of DBLAST. Basically, the idea is to divide the data stream in multiple substreams, which are then transmitted on different antennas [PK94]. VBLAST achieves a high portion of the capacity at low complexity. Unfortunately, the BER performance of VBLAST is limited due to error propagation. It is possible to improve the performance of VBLAST by using decoding schemes with higher receiver complexity as in [CNC00] for example.

Other approaches, which we refer to as diversity-oriented space-time transmission schemes (in contrast to the rate-oriented schemes like DBLAST), exploit multiple antennas at both the transmitter and receiver in order to obtain transmit and receive diversity and therefore increase the reliability of the system [TSC98, TJC99a, Ala98, FK98, YB00, BBH00, HG00, TJ00]. One scheme of particular interest is the Alamouti scheme [Ala98] for two transmit antennas. Later on, [TJC99a] proposed more general schemes referred to as orthogonal space-time block codes (OSTBC) with the same properties as the Alamouti scheme like, e.g., a remarkably simple maximum-likelihood decoding algorithm. Interestingly, the combination of OSTBC with a MIMO antenna system can be represented equivalently as a single-input-single-output (SISO) system, where the channel gain is equal to the Frobenius norm of the actual MIMO channel. The performance of orthogonal space-time block codes [TJC99b, TH02, Lia03] with respect to mutual information was analyzed

(among others) for the uncorrelated Rayleigh fading case in [SP00, BH02] and for the more general case with different correlation scenarios and line of sight (LOS) components in [NBP02, NBP04]. With the knowledge of the stochastic nature of the resulting equivalent channel due to the employment of OSTBC in a MIMO system, the loss in mutual information of OSTBC in subject to transmission rate, number of receive antennas and channel rank was quantified in [SP00], whereas in [BH02] a comparison of OSTBC with a system applying beam-forming was presented.

Unfortunately, the Alamouti space-time code for two transmit and one receive antennas is the only OSTBC, which, to the best of our knowledge, achieves the maximum possible mutual information of a MIMO system [Tel99], since we cannot construct an OSTBC with a transmission rate equal one for more than two transmit antennas [TJC99a, LX03]. Furthermore, by increasing the number of transmit antennas, the transmission rate of the OSTBC is monotonically decreasing, making them unattractive for systems with a very high number of transmit antennas. One solution to this problem is to divide the n_T transmit antennas into groups, where each group employs an OSTBC. At the receiver, group interference suppression techniques are used in combination with suboptimal detectors [TNSC99, PV01, PV03]. However, there is a significant loss in the diversity order, which results in unsatisfactory error rate performances. In this thesis, we use the rate one Alamouti scheme [Ala98] in order to design rate one QSTBC for a higher number of antennas. However, it is also possible to construct QSTBC with rates lower than one based on other OSTBC [TJC99a, TH02, GS01, SX03]. QSTBC for the special case of $n_T = 4$ transmit and $n_R = 1$ receive antennas have also been analyzed, among others, in [Jaf01, PF03, TBH00]. By properly choosing the signal constellations as done in [SP02a, SP02b, SX02, SX04, SJB03, Tir01, SP04], it is possible to improve the BER performance with ML-detection for the codes given in [Jaf01, PF03, TBH00]. The BER performance of QSTBC with suboptimal detectors has been analyzed in [RMG03, RM02]. The performance of QSTBC with respect to outage mutual information (OMI) for some special cases was analyzed via simulations in [PF03] and [MRG04]. In this thesis, we analyze the general case of $n_T = 2^n, n \geq 2$ transmit antennas and an arbitrary number of receive antennas.

2.2 Complete Characterization of Full Diversity Quasi-Orthogonal Space-Time Codes

After presenting the system model at the beginning of this section, we are going to design QSTBC based on the well known Alamouti scheme. We show that the application of QSTBC results in an equivalent channel model representation, similar to with OSTBC. With OSTBC, the MIMO fading system is transformed into a SISO fading system, where the fading parameter is represented by the Frobenius norm of the fading channel matrix [SP00]. With QSTBC, the equivalent channel is still a matrix-valued channel. Nevertheless, it has some very interesting properties. We show, for example, that, due to the employment of QSTBC, the eigenvalues of the resulting equivalent channel are pairwise independent and each eigenvalue is noncentral chi-square distributed with $4n_R$ degrees of freedom. In addition to this, we show that the eigenvectors of the equivalent channel are independent of each channel realization, i.e. they are constant, making the QSTBC a very attractive candidate for future implementation in communications systems.

2.2.1 System model

We consider a system with $n_T = 2^n$ transmit and n_R receive antennas. Our system model is defined by

$$\mathbf{Y} = \mathbf{G}_{n_T}\mathbf{H} + \mathbf{N}\,, \tag{2.1}$$

where $\mathbf{G}_{n_T}$ denotes the $(T \times n_T)$ transmit matrix, $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_{n_R}]$ the $(T \times n_R)$ receive matrix, $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_{n_R}]$ the $(n_T \times n_R)$ channel matrix, and $\mathbf{N} = [\mathbf{n}_1, \ldots, \mathbf{n}_{n_R}]$ the complex $(T \times n_R)$ white Gaussian noise (AWGN) matrix, respectively. An entry $\{n_{ti}\}$ of $\mathbf{N}$ $(1 \le i \le n_R)$ denotes the complex noise at the ith receiver for a given time instant $t\,(1 \le t \le T)$. The real and imaginary parts of n_{ti} are independent and $\mathcal{N}(0,n_T/(2\mathrm{SNR}))$ distributed. Each entry of the channel matrix is represented by $\{h_{ji}\} \in \mathbf{h}_i$, which describes the complex gain of the channel between the jth transmit $(1 \le j \le n_T)$ and the ith receive $(1 \le i \le n_R)$ antenna. The real and imaginary parts of the channel gains are independent and normal distributed random variables and $\mathbf{h}_i$ is $\mathcal{CN}(\mathbf{m}_i, \mathbf{I})$ distributed, where $\mathbf{m}_i$ is the channel mean or Ricean component. The channel matrix is assumed to be constant for a block of T symbols and changes independently from block to block. The average power of the symbols transmitted from each antenna is normalized to one, so that the average power of the received signal at each receive antenna is n_T and the signal-to-noise ratio (SNR) is ρ. It is further assumed that the transmitter has no CSI and the receiver has perfect CSI.

2.2.2 Code construction

A space-time block code is defined by its transmit matrix $\mathbf{G}_{n_T}$, which is a function of the information vector $\mathbf{x} = [x_1, \ldots, x_p]^T$. The rate R of a space-time block code is defined as $R = p/T$. In this section, we focus on rate one QSTBC with length $n_T = T$, therefore $p = n_T$. Now, let us split the vector $\mathbf{x}$ into two vectors, $\mathbf{x}_\mathrm{o}$ and $\mathbf{x}_\mathrm{e}$, for reasons that will be clear later on. The elements of $\mathbf{x}$ with odd index j are collected in $\mathbf{x}_\mathrm{o}$ and those with even index in $\mathbf{x}_\mathrm{e}$, respectively. Both parts of $\mathbf{x}$ are given as

$$\mathbf{x}_\mathrm{o} = \mathbf{\Gamma}\begin{bmatrix} s_1 \\ \vdots \\ s_{n_T/2} \end{bmatrix} = \mathbf{\Gamma}\mathbf{s}^-, \mathbf{x}_\mathrm{e} = \mathbf{\Gamma}\begin{bmatrix} s_{n_T} \\ \vdots \\ s_{n_T/2+1} \end{bmatrix} = \mathbf{\Gamma}\mathbf{s}^+, \tag{2.2}$$

with $s_1, \ldots, s_{n_T} \in \mathcal{C}$, where $\mathcal{C} \subseteq \mathbb{C}$ denotes a complex modulation signal set with unit average power, e.g. M-PSK. Furthermore, $\mathbf{\Gamma} \in \mathbb{C}^{n_T/2 \times n_T/2}$ is a unitary matrix. More details on $\mathbf{\Gamma}$ and its effect on the detection scheme are discussed in section 2.2.3.7.

Starting with the well known Alamouti scheme [Ala98] for $n_T = 2$ transmit antennas as the building block

$$\mathbf{G}_2(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \\ x_2^* & -x_1^* \end{bmatrix}\,,$$

the generalization of the transmit matrix for the QSTBC with $n_T = 2^n$ $(n_T \ge 4)$ is done in the following recursive way

$$\mathbf{G}_{n_T}\left(\{x_j\}_{j=1}^{n_T}\right) = \begin{bmatrix} \mathbf{G}_{\frac{n_T}{2}}\left(\{x_j\}_{j=1}^{\frac{n_T}{2}}\right) & \mathbf{G}_{\frac{n_T}{2}}\left(\{x_j\}_{j=\frac{n_T}{2}+1}^{n_T}\right) \\ \mathbf{G}_{\frac{n_T}{2}}\left(\{x_j\}_{j=\frac{n_T}{2}+1}^{n_T}\right)\mathbf{\Theta}_{n_T} & -\mathbf{G}_{\frac{n_T}{2}}\left(\{x_j\}_{j=1}^{\frac{n_T}{2}}\right)\mathbf{\Theta}_{n_T} \end{bmatrix}\,, \tag{2.3}$$

where $\{x_j\}_{j=1}^{n_T} = x_1, \ldots, x_{n_T}$ and the diagonal $[n_T/2 \times n_T/2]$ matrix $\mathbf{\Theta}_{n_T}$ is given by

$\mathbf{\Theta}_{n_T} = \mathrm{diag}\left(\{(-1)^{j-1}\}_{j=1}^{\frac{n_T}{2}}\right)$.

Example 2.2.1. *For the case of $n_T = 4$ transmit antennas we have*

$$\mathbf{G}_4(\{x_j\}_{j=1}^4) = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2^* & -x_1^* & x_4^* & -x_3^* \\ x_3 & -x_4 & -x_1 & x_2 \\ x_4^* & x_3^* & -x_2^* & -x_1^* \end{bmatrix} . \tag{2.4}$$

In this work, we use the Alamouti scheme as the building block in order to construct the rate one QSTBC. However, it is also possible to construct QSTBC with rates lower than one based on other OSTBC [TJC99a, TH02]. In the following section, we perform channel-matched filtering as the first stage of preprocessing at the receiver in order to obtain the equivalent channel model, followed by the decoupling of the system model in two parts. Afterwards, we analyze the eigenvalues and the eigenvectors of the resulting equivalent channel, leading to important insights of the properties of QSTBC. Noise pre-whitening as the second stage of preprocessing at the receiver is considered in section 2.2.3.6.

2.2.3 Signal Processing

First of all, we briefly review the usual MIMO fading channel without any coordinated coding and the impact of OSTBC on the MIMO channel in order to provide a better insight into the properties of QSTBC.

2.2.3.1 MIMO channel without any coordinated coding

In this case, after channel matched filtering to (2.1), we have

$$\mathbf{H}\mathbf{H}^H = \mathbf{V}\mathbf{D}\mathbf{D}\mathbf{V}^H , \tag{2.5}$$

where $\mathbf{H} = \mathbf{V}\mathbf{D}\mathbf{U}^H$ is the singular value decomposition (SVD) of $\mathbf{H}$, where the unitary matrices $\mathbf{U}$,$\mathbf{V}$ contain the singular vectors of $\mathbf{H}$. The joint density function of the eigenvalues $\mu_1, \ldots, \mu_m$ of $\mathbf{H}\mathbf{H}^H$ in $\mathbf{D}\mathbf{D}$ in the Rayleigh fading case ($\mathbf{m}_i = \mathbf{0}$) is given as [Tel99, Ede89]

$$p_\mu(\mu_1, \ldots, \mu_m) = \frac{1}{m! K_{m,n}} \, e^{\sum_i \mu_i} \prod_i \mu_i^{n-m} \prod_{i<j} (\mu_i - \mu_j)^2 ,$$

where $K_{m,n}$ is a normalizing factor, $n = \max\{n_T, n_R\}$ and $m = \min\{n_T, n_R\}$. It is obvious, that the eigenvalues are not independent of each other and it is well known that the matrix of eigenvectors $\mathbf{V}$ depend on the actual channel realization.

2.2.3.2 Equivalent channel for OSTBC

In case of OSTBC, the following holds for the transmit matrix [TJC99a]

$$\mathbf{G}_{n_T}^H \mathbf{G}_{n_T} = \sum_{j=1}^{p \le n_T} |x_j|^2 \mathbf{I}_{n_T} .$$

Starting with (2.1), after some manipulations and channel matched filtering one arrives at [TJC99b, SP00]

$$\mathbf{y}'' = \mathbf{H}''_{n_T}\mathbf{x} + \mathbf{n}'' \ ,$$

where

$$\mathbf{H}''_{n_T} = \begin{bmatrix} \sum_{i=1}^{n_R}\sum_{j=1}^{n_T}|h_{ji}|^2 & & 0 \\ & \ddots & \\ 0 & & \sum_{i=1}^{n_R}\sum_{j=1}^{n_T}|h_{ji}|^2 \end{bmatrix} . \tag{2.6}$$

Since there is no interaction between the elements of $\mathbf{x}$, the equation above can be decomposed into p parts. The resulting equivalent channel for each element of $\mathbf{x}$ of the OSTBC is then a single-input-single-output (SISO) channel given as

$$\widetilde{\mathbf{H}}_{\frac{n_T}{p}} = \sqrt{\sum_{i=1}^{n_R}\sum_{j=1}^{n_T}|h_{ji}|^2} \ , \tag{2.7}$$

which is equal to the Frobenius norm of the actual MIMO channel matrix $\mathbf{H}$.

In case of the rate one QSTBC discussed in this section, the actual MIMO channel is also transformed into an equivalent channel given as $\widetilde{\mathbf{H}}_{\frac{n_T}{2}}$. Differently from the OSTBC the equivalent channel of QSTBC is still a MIMO channel, however with very interesting properties like constant eigenvectors and i.i.d. eigenvalues following a noncentral $\chi^2_{4n_R}(\delta_{nc})$-distribution as derived in the following.

Remark 2.2.1. *The property of eigenvectors being constant, i.e. independent of the realization of* $\mathbf{H}$*, is already given for OSTBC, with* $\mathbf{V} = \mathbf{I}$*, which is obvious from (2.6). This property is retained in the transition of OSTBC to QSTBC, however, the eigenvectors have a more complex structure.*

2.2.3.3 Channel-Matched Filtering

After rearranging and complex-conjugating some rows of $\mathbf{Y}$, the system equation in (2.1) can be rewritten as (using (2.3))

$$\mathbf{y}' = \mathbf{H}'_{n_T}\mathbf{x} + \mathbf{n}' \ , \tag{2.8}$$

where an entry of the noise vector $\mathbf{n}'$ is $\mathcal{CN}(0,n_T/(\mathrm{SNR}))$ distributed,

$$\mathbf{H}'_{n_T} = [(\mathbf{H}'_{n_T,1})^T, \ldots, (\mathbf{H}'_{n_T,i})^T, \ldots, (\mathbf{H}'_{n_T,n_R})^T]^T$$

and $\mathbf{H}'_{n_T,i}$, $1 \le i \le n_R$, is given as

$$\begin{aligned} \mathbf{H}'_{n_T,i} &= \mathbf{H}'_{n_T,i}\left(\{h_{ji}\}_{j=1}^{n_T}\right) \\ &= \begin{bmatrix} \mathbf{H}'_{\frac{n_T}{2}}\left(\{h_{ji}\}_{j=1}^{\frac{n_T}{2}}\right) & \mathbf{H}'_{\frac{n_T}{2}}\left(\{h_{ji}\}_{j=\frac{n_T}{2}+1}^{n_T}\right) \\ -\mathbf{\Theta}_{n_T}\mathbf{H}'_{\frac{n_T}{2}}\left(\{h_{ji}\}_{j=\frac{n_T}{2}+1}^{n_T}\right)\mathbf{\Theta}_{n_T} & \mathbf{\Theta}_{n_T}\mathbf{H}'_{\frac{n_T}{2}}\left(\{h_{ji}\}_{j=1}^{\frac{n_T}{2}}\right)\mathbf{\Theta}_{n_T} \end{bmatrix} . \end{aligned} \tag{2.9}$$

Thus it appears that $\mathbf{H}'_{n_T,i}\left(\{h_{ji}\}_{j=1}^{n_T}\right)$, with $\left(\{h_{ji}\}_{j=1}^{n_T}\right) = h_{1i}, \ldots, h_{n_T i}$, is obtained recursively, with the recursion starting at $n_T = 2$,

$$\mathbf{H}'_{2,i} = \mathbf{H}'_{2,i}(h_{1i}, h_{2i}) = \begin{bmatrix} h_{1i} & h_{2i} \\ -h^*_{2i} & h^*_{1i} \end{bmatrix} .$$

In order to perform channel-matched filtering, we multiply $(\mathbf{H}'_{n_T})^H$ from left to (2.8) to get

$$\mathbf{y}'' = \mathbf{H}''_{n_T}\mathbf{x} + \mathbf{n}'' \,, \tag{2.10}$$

where the noise vector $\mathbf{n}'' = (\mathbf{H}'_{n_T})^H\mathbf{n}'$ is spatially colored and $\mathbf{H}''_{n_T}$ is given as

$$\begin{aligned}\mathbf{H}''_{n_T} &= \begin{bmatrix} \mathbf{K}^H\mathbf{K} + \mathbf{L}^H\mathbf{L} & \boldsymbol{\Theta}_{n_T}\mathbf{K}^H\mathbf{L}\boldsymbol{\Theta}_{n_T} - \mathbf{L}^H\mathbf{K} \\ -(\boldsymbol{\Theta}_{n_T}\mathbf{K}^H\mathbf{L}\boldsymbol{\Theta}_{n_T} - \mathbf{L}^H\mathbf{K}) & \mathbf{K}^H\mathbf{K} + \mathbf{L}^H\mathbf{L} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{H}''_{\frac{n_T}{2}} & \boldsymbol{\Theta}_{n_T}\mathbf{K}^H\mathbf{L}\boldsymbol{\Theta}_{n_T} - \mathbf{L}^H\mathbf{K} \\ -(\boldsymbol{\Theta}_{n_T}\mathbf{K}^H\mathbf{L}\boldsymbol{\Theta}_{n_T} - \mathbf{L}^H\mathbf{K}) & \mathbf{H}''_{\frac{n_T}{2}} \end{bmatrix},\end{aligned} \tag{2.11}$$

where $\mathbf{K} = \mathbf{H}'_{\frac{n_T}{2}}\left(\{h_{ji}\}_{j=1}^{\frac{n_T}{2}}\right)$ and $\mathbf{L} = \mathbf{H}'_{\frac{n_T}{2}}\left(\{h_{ji}\}_{j=\frac{n_T}{2}+1}^{n_T}\right)$.

2.2.3.4 Decoupling of the system model

An important property of the QSTBC is that the system in (2.10) can be decoupled into two parts due to the special structure of $\mathbf{H}''_{n_T}$ as described in the following. The decoupling is based on the following lemma.

Lemma 2.2.1. *Let $\mathbf{G}_{n_T}(\mathbf{x})$ be the transmit matrix of the rate one QSTBC as defined in (2.3). Then*

$$\mathbf{G}^H_{n_T}(\tilde{\mathbf{x}}_\mathrm{o}) \cdot \mathbf{G}_{n_T}(\tilde{\mathbf{x}}_\mathrm{e}) + \mathbf{G}^H_{n_T}(\tilde{\mathbf{x}}_\mathrm{e}) \cdot \mathbf{G}_{n_T}(\tilde{\mathbf{x}}_\mathrm{o}) = \mathbf{0} \quad \forall\, \mathbf{x} \,, \tag{2.12}$$

applies for QSTBC, where $\tilde{\mathbf{x}}_\mathrm{o} = \mathbf{x}_\mathrm{o} \otimes [\ 1 \quad 0\ ]^T = [x_1, 0, x_3, 0, \ldots, x_{n_T-1}, 0]^T$ and $\tilde{\mathbf{x}}_\mathrm{e} = \mathbf{x}_\mathrm{e} \otimes [\ 0 \quad 1\ ]^T$, and $\mathbf{x}_\mathrm{o}, \mathbf{x}_\mathrm{e}$ are defined in (2.2).

Proof. The proof is given in Appendix 2.5.1.

The property in (2.12) is very crucial, because this enables a simple maximum-likelihood decoding algorithm. Assuming perfect channel estimation is available, the receiver computes the following decision metric over all possible transmit matrices and decides in favor of the transmit matrix that minimizes the following decision metric based on (2.1):

$$\begin{aligned} ||\mathbf{Y} - \mathbf{G}_{n_T}(\mathbf{x}) \cdot \mathbf{H}||_F^2 &= \mathrm{tr}\{(\mathbf{Y} - \mathbf{G}_{n_T}(\mathbf{x}) \cdot \mathbf{H})^H(\mathbf{Y} - \mathbf{G}_{n_T}(\mathbf{x}) \cdot \mathbf{H})\} \\ &= \mathrm{tr}\{\mathbf{Y}^H\mathbf{Y} - \mathbf{Y}^H\mathbf{G}_{n_T}(\mathbf{x})\mathbf{H} \\ &\quad -(\mathbf{Y}^H\mathbf{G}_{n_T}(\mathbf{x})\mathbf{H})^H + \mathbf{H}^H\mathbf{G}_{n_T}(\mathbf{x})^H\mathbf{G}_{n_T}(\mathbf{x})\mathbf{H}\} \,. \end{aligned} \tag{2.13}$$

After some manipulations, we arrive at

$$\begin{aligned} &\mathrm{tr}\{\mathbf{Y}^H_\mathrm{o}\mathbf{Y}_\mathrm{o} + \mathbf{Y}^H_\mathrm{o}\mathbf{G}_{n_T}(\tilde{\mathbf{x}}_\mathrm{o})\mathbf{H} + \mathbf{H}^H\mathbf{G}^H_{n_T}(\tilde{\mathbf{x}}_\mathrm{o})\mathbf{Y}_\mathrm{o} \\ &+ \mathbf{H}^H\mathbf{G}^H_{n_T}(\tilde{\mathbf{x}}_\mathrm{o})\mathbf{G}_{n_T}(\tilde{\mathbf{x}}_\mathrm{o})\mathbf{H} + \mathbf{Y}^H_\mathrm{e}\mathbf{Y}_\mathrm{e} \\ &+ \mathbf{Y}^H_\mathrm{e}\mathbf{G}_{n_T}(\tilde{\mathbf{x}}_\mathrm{e})\mathbf{H} + \mathbf{H}^H\mathbf{G}^H_{n_T}(\tilde{\mathbf{x}}_\mathrm{e})\mathbf{Y}_\mathrm{e} + \mathbf{H}^H\mathbf{G}^H_{n_T}(\tilde{\mathbf{x}}_\mathrm{e})\mathbf{G}_{n_T}(\tilde{\mathbf{x}}_\mathrm{e})\mathbf{H}\} \,, \end{aligned}$$

where $\mathrm{tr}\{\cdot\}$ is the trace function. $\mathbf{Y}_\mathrm{o}$ and $\mathbf{Y}_\mathrm{e}$ are given as

$$\mathbf{Y}_\mathrm{o} = \mathbf{G}_{n_T}(\tilde{\mathbf{x}}_\mathrm{o})\mathbf{H} + \mathbf{N}_\mathrm{o} \ \text{ and } \ \mathbf{Y}_\mathrm{e} = \mathbf{G}_{n_T}(\tilde{\mathbf{x}}_\mathrm{e})\mathbf{H} + \mathbf{N}_\mathrm{e} \,,$$

respectively. The above decision metric can be decomposed into two parts, one of which

$$\begin{aligned} &\mathrm{tr}\{\mathbf{Y}^H_\mathrm{o}\mathbf{Y}_\mathrm{o} + \mathbf{Y}^H_\mathrm{o}\mathbf{G}_{n_T}(\tilde{\mathbf{x}}_\mathrm{o})\mathbf{H} + \mathbf{H}^H\mathbf{G}^H_{n_T}(\tilde{\mathbf{x}}_\mathrm{o})\mathbf{Y}_\mathrm{o} \\ &+ \mathbf{H}^H\mathbf{G}^H_{n_T}(\tilde{\mathbf{x}}_\mathrm{o})\mathbf{G}_{n_T}(\tilde{\mathbf{x}}_\mathrm{o})\mathbf{H}\} \end{aligned}$$

is only a function of $\mathbf{G}_{n_T}(\tilde{\mathbf{x}}_\mathrm{o})$, and the other one

$$\begin{aligned}&\mathrm{tr}\{\mathbf{Y}_\mathrm{e}^H\mathbf{Y}_\mathrm{e} + \mathbf{Y}_\mathrm{e}^H\mathbf{G}_{n_T}(\tilde{\mathbf{x}}_\mathrm{e})\mathbf{H} + \mathbf{H}^H\mathbf{G}_{n_T}^H(\tilde{\mathbf{x}}_\mathrm{e})\mathbf{Y}_\mathrm{e}\\ &+ \mathbf{H}^H\mathbf{G}_{n_T}^H(\tilde{\mathbf{x}}_\mathrm{e})\mathbf{G}_{n_T}(\tilde{\mathbf{x}}_\mathrm{e})\mathbf{H}\}\ ,\end{aligned}$$

is only a function of $\mathbf{G}_{n_T}(\tilde{\mathbf{x}}_\mathrm{e})$. Thus the minimization of (2.13) is equivalent to the separate minimization of these two parts. Note that, due to the processing at the receiver, the property in (2.12) is projected onto the channel matrix $\mathbf{H}'_{n_T}$ in (2.8). The decoupled parts depend either on $\mathbf{x}_\mathrm{o}$ or $\mathbf{x}_\mathrm{e}$ given in (2.2).

Thus, it is now possible to write a decomposed system model for each part based on (2.10). Equivalently, the decomposed system models can be rewritten in a block-diagonal form as

$$\mathbf{y}_\mathrm{bd} = \begin{bmatrix} \widetilde{\mathbf{H}}_{\frac{n_T}{2}} & \mathbf{0} \\ \mathbf{0} & \widetilde{\mathbf{H}}_{\frac{n_T}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{x}_\mathrm{o} \\ \mathbf{x}_\mathrm{e} \end{bmatrix} + \widetilde{\mathbf{n}}\ . \tag{2.14}$$

For illustration, we present two examples for the cases of $n_T = 4$ and $n_T = 8$ transmit antennas.

Example 2.2.2. *($n_T = 4$ transmit antennas) In this case, $\mathbf{H}'_{4,i}$ in (2.9) is given as*

$$\mathbf{H}'_{4,i} = \begin{bmatrix} h_{1i} & h_{2i} & h_{3i} & h_{4i} \\ -h_{2i}^* & h_{1i}^* & -h_{4i}^* & h_{3i}^* \\ -h_{3i} & h_{4i} & h_{1i} & -h_{2i} \\ -h_{4i}^* & -h_{3i}^* & h_{2i}^* & h_{1i}^* \end{bmatrix}\ .$$

and $\mathbf{H}''_4$ appears in (2.10) as

$$\mathbf{y}'' = \begin{bmatrix} \alpha_1 & 0 & i\alpha_2 & 0 \\ 0 & \alpha_1 & 0 & -i\alpha_2 \\ -i\alpha_2 & 0 & \alpha_1 & 0 \\ 0 & i\alpha_2 & 0 & \alpha_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \mathbf{n}''\ , \tag{2.15}$$

where α_1 and α_2 are given as

$$\alpha_1 = \sum_{i=1}^{n_R}\sum_{j=1}^{4}|h_{ji}|^2 \ \text{ and } \alpha_2 = \sum_{i=1}^{n_R} 2\mathrm{Im}(h_{1i}^* h_{3i} + h_{4i}^* h_{2i}), \tag{2.16}$$

respectively. From (2.15), it is now directly obvious, that the system equation can be decoupled into two parts, which then can be considered separately. For the case considered in this example, the equivalent block-diagonal system model can be written as

$$\mathbf{y}_\mathrm{bd} = \begin{bmatrix} \widetilde{\mathbf{H}}_2 & \mathbf{0} \\ \mathbf{0} & \widetilde{\mathbf{H}}_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_3 \\ x_4 \\ x_2 \end{bmatrix} + \widetilde{\mathbf{n}}\ ,$$

with a non-orthogonal

$$\widetilde{\mathbf{H}}_2 = \begin{bmatrix} \alpha_1 & i\alpha_2 \\ -i\alpha_2 & \alpha_1 \end{bmatrix}\ . \tag{2.17}$$

Example 2.2.3. *($n_T = 8$ transmit antennas) The same procedure applied here*

results in $\widetilde{\mathbf{H}}_4$ *given as*

$$\widetilde{\mathbf{H}}_{\frac{n_T}{2}} = \widetilde{\mathbf{H}}_4 = \begin{bmatrix} \alpha_1 & i\alpha_2 & i\alpha_3 & \alpha_4 \\ -i\alpha_2 & \alpha_1 & -\alpha_4 & i\alpha_3 \\ -i\alpha_3 & -\alpha_4 & \alpha_1 & i\alpha_2 \\ \alpha_4 & -i\alpha_3 & -i\alpha_2 & \alpha_1 \end{bmatrix},$$

where

$$\begin{aligned} \alpha_1 &= \sum_{i=1}^{n_R}\sum_{j=1}^{8}|h_{ji}|^2,\ \alpha_2 = \sum_{i=1}^{n_R} 2\mathrm{Im}(h_{1i}^* h_{3i} + h_{4i}^* h_{2i} + h_{5i}^* h_{7i} + h_{8i}^* h_{6i}), \\ \alpha_3 &= \sum_{i=1}^{n_R} 2\mathrm{Im}(h_{1i}^* h_{5i} + h_{6i}^* h_{2i} + h_{3i}^* h_{7i} + h_{8i}^* h_{4i}),\ \textit{and} \\ \alpha_4 &= \sum_{i=1}^{n_R} 2\mathrm{Re}(h_{1i}^* h_{7i} + h_{8i}^* h_{2i} - h_{3i}^* h_{5i} - h_{6i}^* h_{4i}). \end{aligned} \tag{2.18}$$

The general case of arbitrary $n_T = 2^n$ and very important insights on the eigenvalue decomposition, the eigenvalues themselves and the eigenvectors of the equivalent channel $\widetilde{\mathbf{H}}_{\frac{n_T}{2}}$, which are crucial and necessary for further analysis, e.g. the derivations of the lower and upper bounds, are provided in the following section.

2.2.3.5 Eigenvalues and Eigenvectors of the equivalent channel model

In order to completely characterize the equivalent channel with respect to the eigenvalues and eigenvectors, we first look at the properties of matrices with certain structures. Note that the equivalent channel matrices fulfill this special structure. Let the matrix $\mathbf{M}_N$, where $N = \frac{n_T}{2} = 2^{n-1}$, be defined recursively by

$$\mathbf{M}_N(\alpha_1, \ldots, \alpha_N) = \begin{bmatrix} \mathbf{M}_{\frac{N}{2}}(\alpha_1, \ldots, \alpha_{\frac{N}{2}}) & \mathbf{N}_{\frac{N}{2}}(\alpha_{\frac{N}{2}+1}, \ldots, \alpha_N) \\ -\mathbf{N}_{\frac{N}{2}}(\alpha_{\frac{N}{2}+1}, \ldots, \alpha_N) & \mathbf{M}_{\frac{N}{2}}(\alpha_1, \ldots, \alpha_{\frac{N}{2}}) \end{bmatrix}. \tag{2.19}$$

Similarly

$$\mathbf{N}_N(\alpha_{N+1}, \ldots, \alpha_{2N}) = \begin{bmatrix} \mathbf{N}_{\frac{N}{2}}(\alpha_{N+1}, \ldots, \alpha_{\frac{3N}{2}}) & \mathbf{M}_{\frac{N}{2}}(\alpha_{\frac{3N}{2}+1}, \ldots, \alpha_{2N}) \\ -\mathbf{M}_{\frac{N}{2}}(\alpha_{\frac{3N}{2}+1}, \ldots, \alpha_{2N}) & \mathbf{N}_{\frac{N}{2}}(\alpha_{N+1}, \ldots, \alpha_{\frac{3N}{2}}) \end{bmatrix}, \tag{2.20}$$

where the recursion starts with

$$\mathbf{M}_2(\alpha_1, \alpha_2) = \begin{bmatrix} \alpha_1 & i\alpha_2 \\ -i\alpha_2 & \alpha_1 \end{bmatrix} \text{ and } \mathbf{N}_2(\alpha_3, \alpha_4) = \begin{bmatrix} i\alpha_3 & \alpha_4 \\ -\alpha_4 & i\alpha_3 \end{bmatrix}.$$

Remark 2.2.2. *The matrices* $\mathbf{M}_2$ *and* $\mathbf{N}_2$ *have the following eigenvalue decompositions*

$$\mathbf{M}_2 = \mathbf{V}_2\mathbf{S}_2\mathbf{V}_2^H \ \textit{and}\ \mathbf{N}_2 = \mathbf{V}_2\mathbf{T}_2\mathbf{V}_2^H \tag{2.21}$$

where

$$\mathbf{V}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}$$

and

$$\mathbf{S}_2\left(\{\alpha_l\}_{l=1}^2\right) = \begin{bmatrix} \mu_2^1 & 0 \\ 0 & \mu_2^2 \end{bmatrix} = \begin{bmatrix} \alpha_1+\alpha_2 & 0 \\ 0 & \alpha_1-\alpha_2 \end{bmatrix},$$

$$\mathbf{T}_2\left(\{\alpha_l\}_{l=3}^4\right) = \begin{bmatrix} \nu_2^1 & 0 \\ 0 & \nu_2^2 \end{bmatrix} = i \begin{bmatrix} \alpha_3 - \alpha_4 & 0 \\ 0 & \alpha_3 + \alpha_4 \end{bmatrix} .$$

Immediately the following question follows: Is there any structure for deriving the eigenvalues of the matrices of higher N, i.e. if the eigenvalues of $\mathbf{M}_{\frac{N}{2}}$ are given, how can we compute the eigenvalues of $\mathbf{M}_N$. (Note that deriving the eigenvalues of $\mathbf{N}_{\frac{N}{2}}$ is straightforward if the eigenvalues of $\mathbf{M}_{\frac{N}{2}}$ are given). In order to answer this question we are able to state the following lemma, where the arguments of $\mathbf{M}_N$ and $\mathbf{N}_N$ are omitted.

Lemma 2.2.2. *Let $\mathbf{M}_N, \mathbf{N}_N$ be as given in (2.19),(2.20), then $\mathbf{M}_N, \mathbf{N}_N$ with $N = 2^{n-1}$,$n > 2$ have the following eigenvalue decomposition:*

$$\mathbf{M}_N = \mathbf{V}_N \mathbf{S}_N \mathbf{V}_N^H \quad \text{and} \quad \mathbf{N}_N = \mathbf{V}_N \mathbf{T}_N \mathbf{V}_N^H , \tag{2.22}$$

where

$$\mathbf{V}_N = \left(\mathbf{I}_2 \otimes \mathbf{V}_{\frac{N}{2}}\right) \mathbf{\Pi}_N \left(\mathbf{I}_{\frac{N}{2}} \otimes \mathbf{V}_2\right) , \tag{2.23}$$

$$\begin{aligned} &\mathbf{S}_N = \mathbf{S}_N\left(\{\alpha_l\}_{l=1}^N\right) = \mathbf{\Pi}_N \\ &\cdot \begin{bmatrix} \mathbf{S}_{\frac{N}{2}}\left(\{\alpha_l\}_{l=1}^{N/2}\right) - i\mathbf{T}_{\frac{N}{2}}\left(\{\alpha_l\}_{l=N/2+1}^N\right) & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{\frac{N}{2}}\left(\{\alpha_l\}_{l=1}^{N/2}\right) + i\mathbf{T}_{\frac{N}{2}}\left(\{\alpha_l\}_{l=N/2+1}^N\right) \end{bmatrix} \mathbf{\Pi}_N^H , \end{aligned} \tag{2.24}$$

$$\begin{aligned} &\mathbf{T}_N = \mathbf{T}_N\left(\{\alpha_l\}_{l=N+1}^{2N}\right) = \mathbf{\Pi}_N \\ &\cdot \begin{bmatrix} \mathbf{T}_{\frac{N}{2}}\left(\{\alpha_l\}_{l=\frac{3N}{2}}^{2N}\right) - i\mathbf{S}_{\frac{N}{2}}\left(\{\alpha_l\}_{l=N+1}^{\frac{3N}{2}}\right) & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_{\frac{N}{2}}\left(\{\alpha_l\}_{l=\frac{3N}{2}}^{2N}\right) + i\mathbf{S}_{\frac{N}{2}}\left(\{\alpha_l\}_{l=N+1}^{\frac{3N}{2}}\right) \end{bmatrix} \mathbf{\Pi}_N^H , \end{aligned}$$

and

$$[\mathbf{\Pi}_N]_{ij} = \delta\left[2j - 1 - i\right] + \delta\left[2(j - \frac{N}{2}) - i\right]$$

with $\delta[\cdot]$ denoting the delta function, giving $\delta[l] = 1$ for $l = 0$ and $\delta[l] = 0$ for $l \neq 0$, and $[\mathbf{\Pi}_N]_{ij}$ denotes the (i,j)-element of the $N \times N$ permutation matrix $\mathbf{\Pi}_N$.

Proof. The proof is given in Appendix 2.5.2.
It is important to realize that $\mathbf{S}_N$ and $\mathbf{T}_N$ are constructed with different arguments.

As aforementioned, the from $\mathbf{H}''_{n_T}$ in (2.11) in even and odd block structure resorted equivalent channel matrix $\widetilde{\mathbf{H}}_{\frac{n_T}{2}}$ in (2.14) has exactly the same structure as $\mathbf{M}_N$. Therefore, Lemma 2.2.2 can be directly applied to $\widetilde{\mathbf{H}}_{\frac{n_T}{2}}$. To emphasis the usefulness of the resulting property of the QSTBC, we are able to state the following theorem.

Theorem 2.2.1. *The left and right eigenvectors of the equivalent channel in (2.14) of QSTBC, which fulfill the recursive construction rule of (2.9) are given by eq. (2.23) and therefore constant for any arbitrary channel realization.*

Remark 2.2.3. *Another important aspect of Lemma 2.2.2 is the fact that the eigenvalues in $\mathbf{S}_N$ can be obtained simply by adding the eigenvalues of $\mathbf{S}_{\frac{N}{2}}$ and $\mathbf{T}_{\frac{N}{2}}$ in an appropriate manner as done in (2.82) (cf. Appendix 2.5.2), which is used in the following analysis of the QSTBC.*

Lemma 2.2.3. *Let $\mathbf{S}_N$, $\mu_{n_T}^j$, $\nu_{n_T}^j$ be as in Lemma 2.2.2. The eigenvalues of the equivalent channel matrix $\widetilde{\mathbf{H}}_{\frac{n_T}{2}}$ of the QSTBC are given by the recursive equations (2.24). Let $\mathbf{S}_{\frac{n_T}{2}} = \mathbf{D}_{\frac{n_T}{2}} \mathbf{D}_{\frac{n_T}{2}}$ and $(\tilde{\mu}_{n_T}^j)^2 = \frac{2}{n_T}\mu_{n_T}^j$, where $\mu_{n_T}^j$, $1 \leq j \leq {}^{n_T}/_2$*

are the eigenvalues of $\mathbf{S}_{\frac{n_T}{2}}$. *Then for any* n_T *and* n_R, *the eigenvalues* $(\tilde{\mu}^j_{n_T})^2$ *of* $\frac{2}{n_T}\mathbf{D}_{\frac{n_T}{2}}\mathbf{D}_{\frac{n_T}{2}}$ *are obtained as follows*

$$(\tilde{\mu}^j_{n_T})^2 = \sum_{i=1}^{n_R} \mathbf{h}_i^H \mathbf{A}^j_{n_T} \mathbf{h}_i,\ 1 \leq j \leq \frac{n_T}{2}\ , \tag{2.25}$$

where the matrices $\mathbf{A}^j_{n_T}$ *with* $j = 1, 3 \ldots, {}^{n_T}/_2 - 1$ *and* $n_T = 2^n$,$n \geq 2$ *are given as*

$$\mathbf{A}^j_{n_T} = \frac{1}{2}\begin{bmatrix} \mathbf{A}^{j'}_{\frac{n_T}{2}} & -\mathbf{B}^{j'}_{\frac{n_T}{2}} \\ \mathbf{B}^{j'}_{\frac{n_T}{2}} & \mathbf{A}^{j'}_{\frac{n_T}{2}} \end{bmatrix}, \mathbf{A}^{j+1}_{n_T} = \frac{1}{2}\begin{bmatrix} \mathbf{A}^{j'}_{\frac{n_T}{2}} & \mathbf{B}^{j'}_{\frac{n_T}{2}} \\ -\mathbf{B}^{j'}_{\frac{n_T}{2}} & \mathbf{A}^{j'}_{\frac{n_T}{2}} \end{bmatrix} \tag{2.26}$$

with $\mathbf{B}^{j'}_{n_T} = i\mathbf{\Theta}_{n_T}\mathbf{A}^{j'}_{n_T}$, $j' = \frac{j+1}{2}$ *and* $\mathbf{A}^1_2 = \mathbf{A}_2 = \mathbf{I}$.

Proof. The proof is given in Appendix 2.5.3.

Remark 2.2.4. *In case of OSTBC,* $\mathbf{A}^j_{n_T} = \mathbf{I}$, $\forall j$. *Then the eigenvalues are identical and given in (2.7).*

Theorem 2.2.2 ([Mui82])**.** *If* $\mathbf{h}_i$ *is* $\mathcal{CN}(\mathbf{m}_i, \mathbf{I})$ *and* $\mathbf{P}$ *is an* $n_T \times n_T$ *matrix then* $\mathbf{h}_i^H\mathbf{P}\mathbf{h}_i$ *has a noncentral* $\chi^2_k(\delta_{nc})$ *distribution if and only if* $\mathbf{P}$ *is idempotent (*$\mathbf{P}^2 = \mathbf{P}$*), in which case the degrees of freedom is* $k = 2\mathrm{rk}(\mathbf{P}) = 2\mathrm{tr}\{\mathbf{P}\}$ *(where* $\mathrm{rk}(\mathbf{P})$ *and* $\mathrm{tr}\{\mathbf{P}\}$ *denote the rank and trace of* $\mathbf{P}$*, respectively) and* $\delta_{nc} = \mathbf{m}_i^H\mathbf{P}\mathbf{m}_i$.

Lemma 2.2.4. *The matrices* $\mathbf{A}^j_{n_T}$ *are Hermitian and idempotent.*

Proof. The proof is given in Appendix 2.5.4.
From Lemma 2.2.3 and 2.2.4, it is now possible to prove the following theorem.

Theorem 2.2.3. *Let* $\mathbf{D}_{\frac{n_T}{2}}\mathbf{D}_{\frac{n_T}{2}}$ *defined as in Lemma 2.2.3 be the diagonal eigenvalue matrix of the equivalent channel matrix of an QSTBC, which fulfill the recursive construction rule of (2.9). Then for any* n_T *and* n_R, *the eigenvalues* $(\tilde{\mu}^j_{n_T})^2$ *of* $\frac{2}{n_T}\mathbf{D}_{\frac{n_T}{2}}\mathbf{D}_{\frac{n_T}{2}}$ *are pairwise independent and identical noncentral chi-square distributed with* $4n_R$ *degrees of freedom.*

Remark 2.2.5. *It is important to realize that* $\mathbf{\Theta}^H_{n_T} = \mathbf{\Theta}_{n_T}$ *and* $\mathbf{\Theta}^H_{n_T}\mathbf{A}^j_{n_T}\mathbf{\Theta}_{n_T} = \mathbf{A}^j_{n_T}$, *since the entries on the* l*th-diagonals of the matrices* $\mathbf{A}^j_{n_T}$, *where* $l = \pm\{1, 3, \ldots, n_T - 1\}$ *(*$l = 0$ *represents the main diagonal,* $l > 0$ *above the main diagonal, and* $l < 0$ *below the main diagonal), are equal to zero.*

Remark 2.2.6. *Recall from section 2.2.3.4, that we have decoupled the system into two independent parts. The derivation above holds therefore for both parts, i.e., each eigenvalue appears twice, once for each part.*

Proof. The proof of Theorem 2.2.3 is given in Appendix 2.5.5.

2.2.3.6 Noise pre-whitening

Since $\widetilde{\mathbf{n}}$ in (2.14) is colored noise, the next step to perform is pre-whitening. With the knowledge of the theorems 2.2.1 and 2.2.3 it is easy to compute the pre-whitening filter $\mathbf{F}_{\mathrm{PW}}$ at the receiver now. To this end, we need just the eigenvalue decomposition of $\widetilde{\mathbf{H}}_{\frac{n_T}{2}}$ given as $\widetilde{\mathbf{H}}_{\frac{n_T}{2}} = \mathbf{V}_{\frac{n_T}{2}}\mathbf{S}_{\frac{n_T}{2}}\mathbf{V}^H_{\frac{n_T}{2}}$ with $\mathbf{S}_{\frac{n_T}{2}} = \mathbf{D}_{\frac{n_T}{2}}\mathbf{D}_{\frac{n_T}{2}}$. Therefore the

pre-whitening filter is given as $\mathbf{F}_{\mathrm{PW}} = \mathbf{D}^{-1}_{\frac{n_T}{2}} \mathbf{V}^H_{\frac{n_T}{2}}$. By multiplying $\mathrm{diag}(\mathbf{F}_{\mathrm{PW}}, \mathbf{F}_{\mathrm{PW}})$ from the left to (2.14), we arrive at

$$\widehat{\mathbf{y}}_{\mathrm{bd}} = \begin{bmatrix} \widehat{\mathbf{H}} & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{H}} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{\mathrm{o}} \\ \mathbf{x}_{\mathrm{e}} \end{bmatrix} + \mathbf{w} \,, \tag{2.27}$$

where the entries of $\mathbf{w}$ are mutually i.i.d. Gaussian processes again.

Example 2.2.4. *In the case of $n_T = 4$ transmit antennas, $\widehat{\mathbf{H}}$ in (2.27) is given as*

$$\widehat{\mathbf{H}} = \begin{bmatrix} \tilde{\mu}_4^1 & i\tilde{\mu}_4^1 \\ \tilde{\mu}_4^2 & -i\tilde{\mu}_4^2 \end{bmatrix} \,, \tag{2.28}$$

and

$$\tilde{\mu}_4^1 = \sqrt{\frac{\alpha_1 + \alpha_2}{2}}, \text{ and } \tilde{\mu}_4^2 = \sqrt{\frac{\alpha_1 - \alpha_2}{2}} \,. \tag{2.29}$$

2.2.3.7 Linear maximum likelihood detection

From theorems 2.2.1 and 2.2.3, it is now possible to determine $\mathbf{\Gamma}$ adequately, resulting in an attractive system equation, which allows a very simple but effective ML-decoding. To emphasize this property we formulate the following corollary.

Corollary 2.2.1. *By choosing the matrix $\mathbf{\Gamma}$ in (2.2) as $\mathbf{\Gamma} = \mathbf{V}_{\frac{n_T}{2}}$, (2.27) can be rewritten as*

$$\widehat{\mathbf{y}}_{\mathrm{bd}} = \begin{bmatrix} \mathbf{D}_{\frac{n_T}{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{\frac{n_T}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{s}^- \\ \mathbf{s}^+ \end{bmatrix} + \mathbf{w} \,. \tag{2.30}$$

At this point, the elements of $\mathbf{s}^-$ (and also $\mathbf{s}^+$) are completely decoupled, since they experience no interference from each other. Thus, a linear ML-detector is able to detect the symbols (or elements) transmitted from the antennas separately.

Proof. The matrix $\widehat{\mathbf{H}}$ in (2.27) can be decomposed as

$$\widehat{\mathbf{H}} = \mathbf{D}_{\frac{n_T}{2}} \mathbf{V}^H_{\frac{n_T}{2}} \,, \tag{2.31}$$

where $\mathbf{D}_{\frac{n_T}{2}}$ is a diagonal matrix containing the singular values of $\widehat{\mathbf{H}}$. Since $\mathbf{V}_{\frac{n_T}{2}}$ is constant for all channel realizations, we can set $\mathbf{\Gamma} = \mathbf{V}_{\frac{n_T}{2}}$ without any knowledge of the current channel realization. Using (2.31) in (2.27) results in (2.30). That concludes the proof. □

Example 2.2.5. *For $n_T = 4$ transmit antennas, $\mathbf{V}_2$ and $\mathbf{D}_2$ are given as*

$$\mathbf{V}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}, \mathbf{D}_2 = \sqrt{2} \begin{bmatrix} \tilde{\mu}_4^1 & 0 \\ 0 & \tilde{\mu}_4^2 \end{bmatrix} \,.$$

Example 2.2.6. *For $n_T = 8$ transmit antennas, we have the following $\mathbf{V}_4$*

$$\mathbf{V}_4 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ -i & -i & i & i \\ i & -i & -i & i \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

and $\mathbf{D}_4 = \sqrt{4}\mathrm{diag}(\tilde{\mu}_8^1, \ldots, \tilde{\mu}_8^4)$ *with*

$$\begin{aligned} \tilde{\mu}_8^1 &= \sqrt{(\alpha_1+\alpha_2+\alpha_3-\alpha_4)/4},\ \tilde{\mu}_8^2 = \sqrt{(\alpha_1+\alpha_2-\alpha_3+\alpha_4)/4}\,, \\ \tilde{\mu}_8^3 &= \sqrt{(\alpha_1-\alpha_2+\alpha_3+\alpha_4)/4},\ \textit{and}\ \tilde{\mu}_8^4 = \sqrt{(\alpha_1-\alpha_2-\alpha_3-\alpha_4)/4}\,. \end{aligned} \tag{2.32}$$

Based on these new insights, we provide some performance analysis in the following section, where we focus on the case of Rayleigh fading ($\mathbf{m}_i = \mathbf{0}$).

2.3 Performance analysis of full diversity QSTBC

In this section, we analyze the performance of QSTBC with respect to error probability and mutual information for the following receiver structures.

- Optimal ML-detection (section 2.3.1 and section 2.3.2).
- Optimal ML-detection and linear ZF-detection in combination with antenna selection in section 2.3.3.
- Lattice-Reduction aided linear ZF-detection in section 2.3.4.

2.3.1 Mutual Information

From an information point of view, systems with multiple transmit and receive antennas provide very high data rates in the presence of Rayleigh fading [Win87, Tel99, FG98]. An important aspect is to obtain closed-form analytical formulas for the ergodic capacity or the outage mutual information of such MIMO channels. In [WG04], the probability density function (pdf) of the random mutual information for independent identically distributed (i.i.d.) MIMO channels was derived in the form of the inverse Laplace transform and a Gaussian approximation of the pdf was presented. In [SL03], the impact of MIMO channel rank deficiency and spatial fading correlation on the mutual information is analyzed. Furthermore, the optimal transmit strategy and the impact of correlation on the outage probability were derived in [BJ04] for correlated channels. The performance of OSTBC [TJC99a] with respect to mutual information has been analyzed (among others) for the uncorrelated Rayleigh fading case in [SP00, BH02] and for the more general case in [NBP04] with different correlation scenarios and line-of-sight (LOS) components. In [SP00], the loss in mutual information of OSTBC is quantified subject to the transmission rate, the number of receive antennas and the channel rank, whereas in [BH02] OSTBC are compared with a system applying beamforming. Unfortunately, the Alamouti space-time code for two transmit and one receive antennas is the only OSTBC, which, to the best of our knowledge, achieves the mutual information, which lead us automatically to the question of the rates which can be achieved by applying QSTBC.

The mutual information of a MIMO system with n_T transmit and n_R receive antennas with no CSI at the transmitter and perfect CSI at the receiver by using the optimal transmit strategy is given in eq. (1.2) [Tel99], which we repeat here for convenience

$$I = \log_2 \det\left(\mathbf{I}_{n_R} + \frac{\rho}{n_T}\mathbf{H}\mathbf{H}^H\right)\,.$$

[1] We use the same terminology in this work as in [Tel99], i.e. we use the term capacity only in the Shannon sense and distinguish therefore between the concept of outage mutual information (OMI) and capacity.

The portion of mutual information achieved with QSTBC is

$$I_Q = \frac{2}{n_T}\log_2 \det\left(\mathbf{I}_{n_T/2} + \frac{\rho}{n_T}\mathbf{D}_{\frac{n_T}{2}}\mathbf{D}_{\frac{n_T}{2}}\right) = \frac{2}{n_T}\log_2 \prod_{j=1}^{n_T/2}\left(1 + \frac{\rho\frac{n_T}{2}}{n_T}(\tilde{\mu}_{n_T}^j)^2\right) \tag{2.33}$$

$$= \frac{2}{n_T}\sum_{j=1}^{n_T/2}\log_2\left(1 + \frac{\rho\frac{n_T}{2}}{n_T}(\tilde{\mu}_{n_T}^j)^2\right) . \tag{2.34}$$

Formula (2.34) can be interpreted as a system with $n_T/2$ transmit antennas applying a rate $2/n_T$ space-time code, where each of the $n_T/2$ transmitted signal components is received without any interference from the others by a separate set of $2n_R$ (according to the number of degrees of freedom of the eigenvalues $(\tilde{\mu}_{n_T}^j)^2$) antennas.

2.3.1.1 Outage probability P_{out}

The outage probability P_{out} achievable with QSTBC is defined as the probability that I_Q is smaller than a certain rate R, i.e.

$$P_{out}(R, n_T, n_R, \rho) = \Pr[I_Q < R] .$$

Unfortunately, the exact analysis of P_{out} is not available. We therefore provide lower and upper bounds in the following.

Lower bounds

Proposition 2.3.1. *The outage probability P_{out} is lower bounded by*

$$P_{out}(R, n_T, n_R, \rho) \geq 1 - \exp\left(-\frac{n_T}{\rho}\left(2^R - 1\right)\right)\sum_{k=0}^{n_T n_R - 1}\frac{\left(\frac{n_T}{\rho}\left(2^R-1\right)\right)^k}{k!} . \tag{2.35}$$

Proof. By using the arithmetic mean - geometric mean inequality, i.e.

$$\prod_{l=1}^{L} a_l^{1/L} \leq \frac{1}{L}\sum_{l=1}^{L} a_l,\ a_l \geq 0$$

with equality if and only if $a_1 = a_2 = \cdots = a_L$, we obtain an upper bound for I_Q (and therefore a lower bound on P_{out}) given as

$$I_Q \leq \frac{2}{n_T}\log_2\left(\frac{2}{n_T}\left(\sum_{j=1}^{n_T/2} 1 + \frac{\rho\frac{n_T}{2}}{n_T}(\tilde{\mu}_{n_T}^j)^2\right)\right)^{\frac{n_T}{2}} = \log_2\left(1 + \frac{\rho}{n_T}\alpha_1\right) = I_Q^u , \tag{2.36}$$

where $\alpha_1 = \sum_{i=1}^{n_R}\sum_{j=1}^{n_T}|h_{ji}|^2$ for the general case of arbitrary n_T (compare with (2.16) and (2.18) for $n_T = 4$ and $n_T = 8$, respectively). The lower bound on the outage probability P_{out} can be written as

$$P_{out}(R, n_T, n_R, \rho) = \Pr[I_Q < R] \geq \Pr[I_Q^u < R] = \Pr\left[\alpha_1 < \frac{n_T}{\rho}\left(2^R - 1\right)\right] .$$

Since α_1 is a chi-square distributed random variable with $2n_T n_R$ degrees of freedom, P_{out} is given as [GR83, p.310,3.351(1)] in (2.35). This concludes the proof. □

Corollary 2.3.1. *The lower bound in (2.35) becomes tight for low SNR values or when n_R increases.*

Proof. The proof is given in Appendix 2.5.6.

Another lower bound on the outage probability is given by

$$P_{out}(R, n_T, n_R, \rho) \geq \left(1 - \exp\left(-\frac{2}{\rho}\left(2^R - 1\right)\right) \sum_{k=0}^{2n_R-1} \frac{\left(\frac{2}{\rho}\left(2^R-1\right)\right)^k}{k!}\right)^{\frac{n_T}{2}} , \quad (2.37)$$

which is based on the fact that

$$I_Q \leq \log_2\left(1 + \frac{\rho}{2}\max_j(\tilde{\mu}_{n_T}^j)^2\right) .$$

Upper bounds Using the positive definiteness of $\mathbf{D}_{\frac{n_T}{2}}\mathbf{D}_{\frac{n_T}{2}}$, (2.33) can be lower bounded as

$$I_Q \geq \frac{2}{n_T}\log_2\left(1 + \left(\frac{\rho}{n_T}\right)^{\frac{n_T}{2}} \det\left(\mathbf{D}_{\frac{n_T}{2}}\mathbf{D}_{\frac{n_T}{2}}\right)\right) = I_Q^l .$$

Thus, the upper bound on P_{out} is given as

$$\begin{aligned} P_{out}(R, n_T, n_R, \rho) &= \Pr[I_Q < R] \leq \Pr[I_Q^l < R] \\ &= P\Big[\det\left(\mathbf{D}_{\frac{n_T}{2}}\mathbf{D}_{\frac{n_T}{2}}\right) < \underbrace{\left(\frac{n_T}{\rho}\right)^{n_T/2}\left(2^{\frac{Rn_T}{2}} - 1\right)}_{\tilde{R}}\Big] . \end{aligned}$$

For the special case of $n_T = 4$ transmit antennas, P_{out} is given as

$$P_{out}(R, n_T, n_R, \rho) \leq Pr[4(\tilde{\mu}_4^1)^2(\tilde{\mu}_4^2)^2 < \tilde{R}] .$$

Since $4(\tilde{\mu}_4^1)^2(\tilde{\mu}_4^2)^2$ is a product of two chi-square distributed random variables, both with $m = 4n_R$ degrees of freedom, P_{out} is given by [Spr79, p.365, eq.9.9.34]

$$\begin{aligned} P_{out}(R, m, \rho) &\leq \int_0^{\tilde{R}} \frac{y^{\frac{m}{2}-1}}{\Gamma(\frac{m}{2})^2 2^{m-1}} K_0(\sqrt{y})\, dy \\ &= \frac{\tilde{R}^{\frac{m}{2}}}{\Gamma(\frac{m}{2})^2 2^m}\left(\sum_{k=0}^{\infty} \frac{\left(\ln(\frac{4}{\tilde{R}}) + 2\Psi(k+1) + \frac{1}{\frac{m}{2}+k}\right)(\frac{\tilde{R}}{4})^k}{(\frac{m}{2}+k)(k!)^2}\right), \end{aligned} \quad (2.38)$$

where Ψ is the Psi function [GR83, p.943, eq.8.360] and Γ is the Gamma function [GR83, p.XXXI]. Note that a useful and simple approximation of the outage probability can be obtained for high SNR by retaining only the first term (i.e. $k = 0$) of the series within the upper bound.

Another upper bound on the outage probability for any number of transmit and

receive antennas is given by

$$P_{out}(R, n_T, n_R, \rho) \geq 1 - \left(\exp\left(-\frac{2}{\rho}\left(2^R - 1\right)\right) \sum_{k=0}^{2n_R-1} \frac{\left(\frac{2}{\rho}\left(2^R-1\right)\right)^k}{k!} \right)^{\frac{n_T}{2}}, \quad (2.39)$$

which is based on the fact that

$$I_Q \geq \log_2\left(1 + \frac{\rho}{2}\min_j(\tilde{\mu}_{n_T}^j)^2\right).$$

2.3.1.2 Ergodic mutual information (EMI)

Case $n_R = 1$ Since the eigenvalues are independent, the average mutual information can be written as

$$\begin{aligned} C_Q &= \mathrm{E}\left[I_Q\right] = \mathrm{E}\left[\log_2\left(1 + \frac{\rho}{2}\tilde{\mu}_i^2\right)\right] \\ &= \frac{1}{\ln(2)}\left(1 + e^{\frac{2}{\rho}}\mathrm{E}_1\left(\frac{2}{\rho}\right)\left(\frac{\rho-2}{\rho}\right)\right), \end{aligned} \quad (2.40)$$

where $\mathrm{E}_1(x)$ is the exponential integral function [GR83, p.XXXII]. Since the eigenvalues are independent from the number of transmit antennas, it follows that (2.40) is also independent from the number of transmit antennas. This result is a generalization of [MRG04], where it was shown that the average mutual information for $n_T = 4$ and $n_T = 8$ are equal.

Case n_R arbitrary In the case of an arbitrary number of receive antennas, the ergodic mutual information (EMI) in (2.40) is given as

$$C_Q = \frac{1}{\ln(2)} \sum_{k=0}^{2n_R-1} \left(\frac{2}{\rho}\right)^{2n_R-k-1} e^{\frac{2}{\rho}}\Gamma\left(1-(2n_R-k), \frac{2}{\rho}\right). \quad (2.41)$$

A lower bound on the EMI in (2.41) is given in the following theorem.

Theorem 2.3.1. *The ergodic mutual information by applying a rate one QSTBC as defined in (2.3) is independent of the number of transmit antennas and lower bounded by*

$$C_Q \geq \log_2\left(1 + \rho \exp\frac{\Psi(n_R) + \Psi\left(n_R + \frac{1}{2}\right)}{2}\right). \quad (2.42)$$

Proof. By applying Minkowski's inequality [HJ85](det($\mathbf{I}_n + \mathbf{C}$) $\geq (1 + \det(\mathbf{C})^{\frac{1}{n}})^n$) to (2.33), we arrive at

$$I_Q \geq \log_2\left(1 + \rho \det\left(\frac{1}{n_T}\mathbf{D}^H\mathbf{D}\right)^{\frac{2}{n_T}}\right),$$

which is equal to

$$I_Q \geq \log_2\left(1 + \rho \exp\left(\frac{2}{n_T}\ln\det\left(\frac{1}{n_T}\mathbf{D}^H\mathbf{D}\right)\right)\right).$$

Since $\log_2(1+ce^x)$ is a convex function in x for $c>0$ and by applying Jensen's inequality it holds that $\mathrm{E}\left[\log_2(1+ce^x)\right] \geq \log_2(1+c\exp(\mathrm{E}\left[x\right]))$, we have

$$C_Q = \mathrm{E}\left[I_Q\right] \geq \log_2\left(1+\rho\exp\mathrm{E}\left[\ln\left(\frac{\tilde{\mu}_i^2}{4}\right)\right]\right) .$$

which results in (2.42). That concludes the proof. □

Some simulation results of the performance of QSTBC and their interpretation are presented in the following section.

2.3.1.3 Simulations

In Fig. 2.1, the outage mutual information (OMI) of QSTBC I_Q with our new transmit strategy and our linear detector as derived in section 2.2.3.7 is compared with the nonlinear ML-detector and ZF-detector in [PF03]. Additionally, the OMI of a MIMO system with $n_T = 4$ and $n_R = 1$ is depicted. In the Fig., we can see that our new transmit strategy outperforms the ZF-detector of [PF03] and achieves the same portion of mutual information as the non-linear ML-detector presented in [PF03].

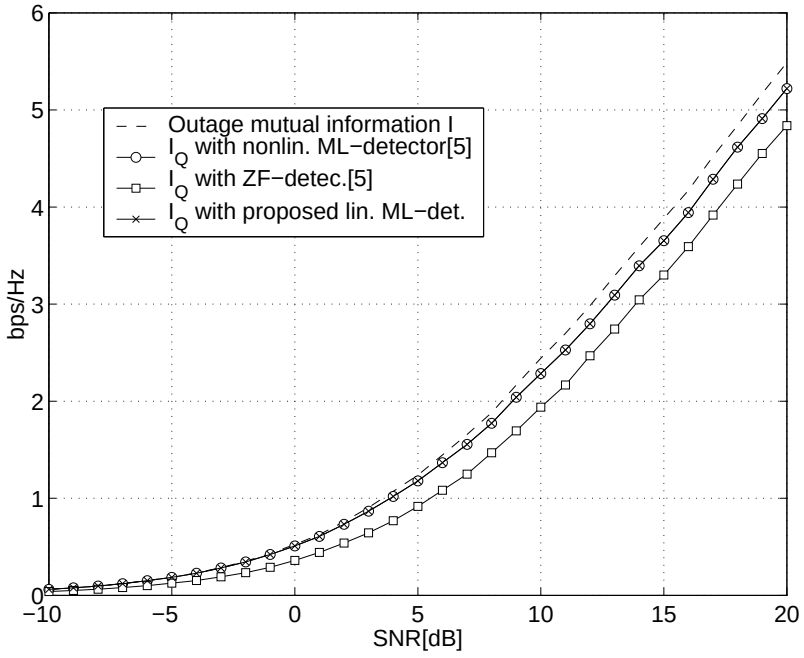

Figure 2.1: 10% Outage mutual information (OMI) of a MIMO system, our new approach, and the ML-detector and ZF-detector from [PF03] with $n_T = 4$ and $n_R = 1$.

In Fig. 2.2, the performance of QSTBC in terms of OMI with $n_T = 4$ and $n_T = 8$ antennas is depicted for $n_R \geq 1$. For $n_R = 1$, the performance with $n_T = 8$ is similar to the case of $n_T = 4$ transmit antennas (depicted in Fig. 2.2), i.e. we achieve a significant fraction of the OMI. However, we observe in Fig. 2.2, that by increasing the number of receive antennas, the performance of QSTBC with $n_T = 8$ as well as with $n_T = 4$ is dramatically reduced in terms of achievable OMI.

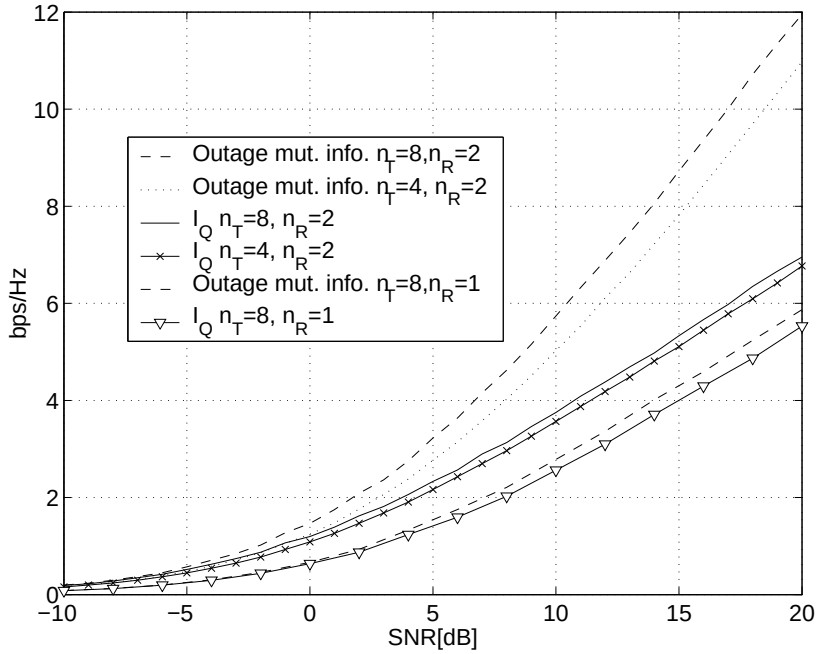

Figure 2.2: 10% Outage mutual information of a MIMO system and our new approach with $n_T = 4$,$n_T = 8$ transmit and $n_R \geq 1$ receive antennas.

In Fig. 2.3, P_{out} of QSTBC with $n_T = 4$ transmit and $n_R = 1$ to $n_R = 3$ and $n_R = 6$ receive antennas is depicted. In addition, the lower and upper bound from (2.35) and (2.38), respectively, are depicted for the given number of receive antennas. For the case of $n_R = 6$ receive antennas, we also depicted the lower and upper bounds from (2.37) and (2.39), respectively. From the Fig. we observe that the lower bound on the performance of QSTBC with respect to P_{out} from (2.35) gets tight by increasing the number of receive antennas. Even the upper and lower bound from (2.38) and (2.37), respectively, perform very well and show to be useful. Only the upper bound from (2.39) is relatively loose, since the decay of the curve is smaller than $n_T n_R$, i.e. the full diversity is not achieved. The performance of QSTBC with respect to P_{out} is depicted for $n_T = 8$ transmit antennas in Fig. 2.4. Similarly to the case of $n_T = 4$ transmit antennas, the lower bound from (2.35) gets tight by using a high number of antennas on the receiver side.

In Fig. 2.5, the EMI of QSTBC C_Q and the closed-form expressions for the EMI as given in (2.40) and (2.42) of a MIMO system with $n_T = 4$ and $n_R = 1$ and $n_R = 4$, respectively, are depicted. From the Fig. we observe, that our lower bound on the performance of QSTBC with respect to C_Q for $n_R = 1$ receive antenna is not tight only for moderate SNR values. By increasing the number of receive antennas the bound gets tight for all SNR values.

The behavior of space-time transmission schemes with respect to mutual information is an important benchmark from an information theoretic point of view. What is also interesting for the practical system designer, however, is the error rate performance, which is analyzed in the following section.

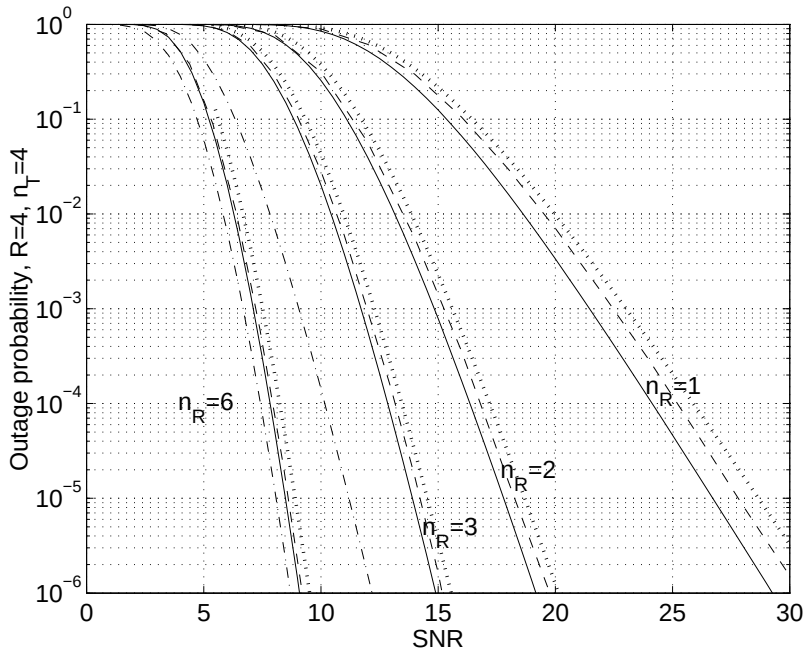

Figure 2.3: Outage probabilities of QSTBC (dashed lines), upper bound from (2.38) (dotted lines), and lower bound from (2.35) (solid lines) for $n_T = 4$ transmit and different numbers of receive antennas n_R, rate=4. For $n_R = 6$, the lower and upper bounds from (2.37) and (2.39), respectively, are also depicted.

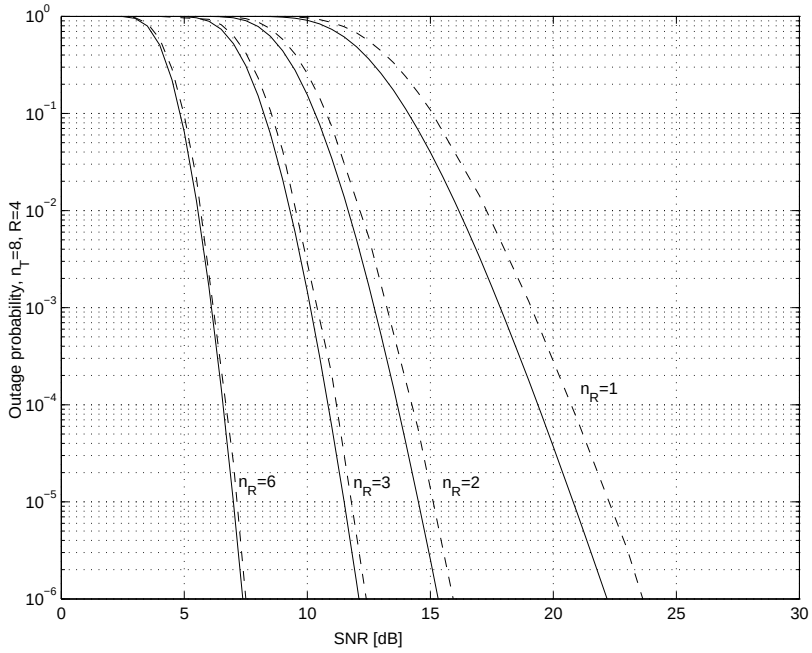

Figure 2.4: Outage probabilities of QSTBC (dashed lines) and lower bound from (2.35) (solid lines) for $n_T = 8$ transmit and different numbers of receive antennas n_R, rate=4.

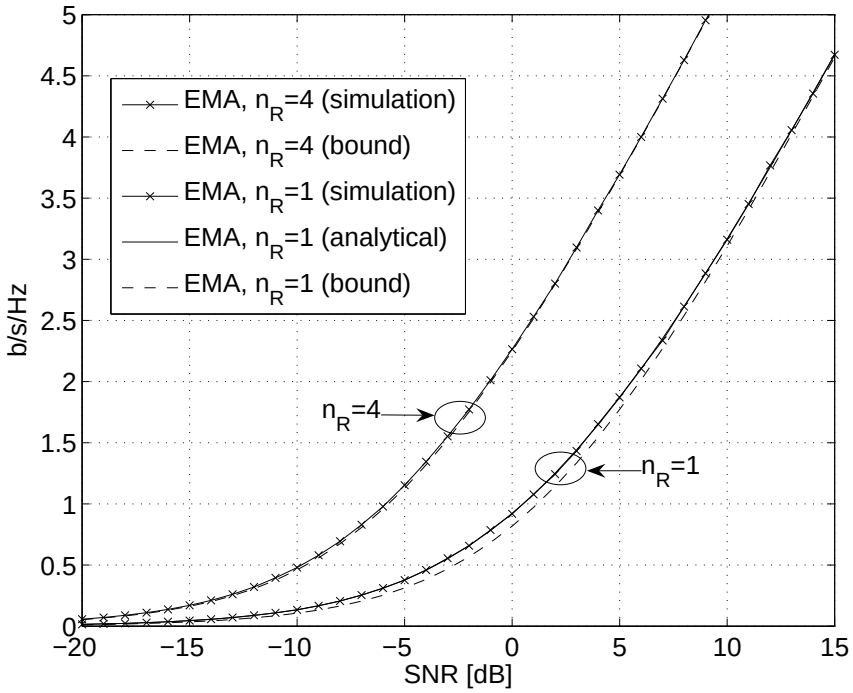

Figure 2.5: Ergodic mutual information (EMI) and lower bounds (dashed lines) with $n_T = 4$ transmit antennas and $n_R = 1$ and $n_R = 4$ receive antennas.

2.3.2 Bit-error rate performance analysis

The bit-error-rate performance of QSTBC was analyzed first in [Jaf01, TBH00]. In the analysis the authors in [Jaf01, TBH00] realized, that the diversity (i.e. the slope of the BER) of their schemes with $n_T = 4$ and $n_R = 1$ was only half of the maximum diversity order of $n_T n_R$. It seemed as if, on the one hand, the QSTBC in [Jaf01, PF03, TBH00] had a higher transmission rate in comparison to OSTBC, on the other hand, the disadvantage of the QSTBC proposed in [Jaf01, PF03, TBH00] was the reduction of transmit diversity, i.e. the slope of the bit error rate (BER) curve was not as steep as in the orthogonal case.

In principle, there is always a tradeoff between different system parameters, e.g., diversity and rate [ZT03] or BER performance and system complexity. In this section, we analyze two different approaches for the QSTBC in [Jaf01, PF03, TBH00], where the aim of the approaches is to

- Improve the BER performance of QSTBC, i.e. enhance the diversity in order to get full diversity of $n_T n_R$(**Approach I**) or
- Reduce the decoding complexity (**Approach II**) by keeping the diversity given in [Jaf01, PF03, TBH00].

We compare these approaches and observe a tradeoff between diversity and decoding complexity for QSTBC with the constraint of a fixed transmission rate.

In **Approach I**, we are interested in optimizing the transmit diversity, i.e, the slope of the BER curve, for the QSTBC. A similar approach was recently made in [Tir01] and also in [SP02a, SX02, SJ03] in order to improve the BER performance. A constellation rotation method was proposed, that either aims at

- Maximizing the minimum Euclidean distance in the constellation of the symbols constituting the code or
- the diversity product of the code.

Both criteria are examined in this section in order to show which one is more important. Interestingly, since [Tir01, SP02a, SX02, SJ03] just optimize the BER with respect to one of the aforementioned criteria, the optimal rotations in these works are in some cases not the global optimum. Furthermore, we obtain analytical BER performance results for QSTBC utilizing the method of constellation rotation.

In **Approach II**, we are interested in reducing the complexity of the QSTBC, since a disadvantage of QSTBC in [Jaf01, PF03, TBH00] is that the ML decoder of these codes works with groups of transmitted symbols instead of single symbols as in the orthogonal case. In contrast to increasing the transmit diversity to $n_T n_R$ as in Approach I and [Tir01, SP02a, SX02, SJ03], here we are interested in simplifying the receiver while maintaining the transmit diversity. To this end, we propose the transmit strategy derived in section 2.2.3.7, which decouples the signals transmitted from different antennas rather than the joint detection of groups of symbols. We obtain analytical performance results in terms of BER for QSTBC employing this new transmit strategy with its linear ML-detector.

A diversity analysis is performed for the QSTBC in the following. Afterwards, we consider Approach I, followed by Approach II and its impact on the complexity and performance of the system. Later on the error rates of both approaches are analyzed.

2.3.2.1 Diversity Analysis

In this section, we consider the performance of the rate one quasi-orthogonal space time block code for $n_T = 4$ transmit antennas with the transmit matrix $\mathbf{G}_4(\mathbf{x})$ in (2.4) specified in section 2.2.2. The generalization of this analysis of QSTBC with higher n_T is relatively straightforward. However, since the generalization provides no additional insight into the error rate analysis, the focus here is on the case of $n_T = 4$ transmit antennas.

To obtain the BER-performance of the scheme in [Jaf01], we choose the matrix $\mathbf{\Gamma}$ given in (2.2) as $\mathbf{\Gamma} = \mathbf{I}_2$, where $\mathbf{I}_2$ is the (2×2) identity matrix. The determinant of $\mathbf{G}_4^H \mathbf{G}_4$ is given as

$$\det(\mathbf{G}_4^H \mathbf{G}_4) = \left(a^2 + b^2\right)^2 \ , \tag{2.43}$$

where

$$a = a\left(\{x_l\}_{l=1}^{n_T}\right) = \sum_{l}^{n_T=4} |x_l|^2 \ , \tag{2.44}$$

$$b = b\left(\{x_l\}_{l=1}^{n_T}\right) = 2i \cdot \mathrm{Im}\{x_1^* x_3 + x_4^* x_2\} \ . \tag{2.45}$$

Note that a is real and b is imaginary. If we replace x_l by $c_l - e_l$, $\mathbf{G}_4$ becomes the codeword difference matrix $\mathbf{B}$ from [TSC98]. In this matrix c_l and e_l are signals (or code-symbols) from the codewords $\mathbf{c} = \{c_l\}_{l=0}^{n_T}$ and $\mathbf{e} = \{e_l\}_{l=0}^{n_T}$, respectively. It is assumed that $\mathbf{c}$ was transmitted and the ML receiver decides erroneously in favor of $\mathbf{e}$, with $\mathbf{c} \neq \mathbf{e}$. The codewords $\mathbf{c}, \mathbf{e}$ are taken from a given constellation set $\mathcal{X}$ with average energy of n_T, where $\mathcal{X} \subseteq \mathbb{C}^{n_T}$.

In order to achieve full diversity, the codeword distance matrix $\mathbf{A} = \mathbf{B}^H \mathbf{B}$ [TSC98, p.749] has to have full rank of $n_T n_R$, which is equivalent to $\det(\mathbf{A}) \neq 0$. Unfor-

tunately, without the assumption of a Gaussian codebook, i.e. by using standard complex modulation signals like M-PSK, it is possible that $b^2 = -a^2$ and therefore $\det(\mathbf{A}) = 0$. This means we have a loss in diversity, since the minimum rank of the codeword distance matrix $\mathbf{A}$ is $n_T n_R/2$ [Jaf01].

Lemma 2.3.1. *There exist two distinct codewords* $\mathbf{c}$ *and* $\mathbf{e}$*, such that the determinant of the codeword distance matrix* $\mathbf{A}(\mathbf{c}, \mathbf{e})$ *is equal to zero, i.e.* $\mathbf{A}(\mathbf{c}, \mathbf{e})$ *becomes rank deficient and consequently the QSTBC in (2.4) does not achieve full diversity.*

Proof. Given (2.43), the determinant of the codeword distance matrix is equal to

$$\left[\left(\sum_{k=1}^{2} |\Delta_{3k-2} + i\Delta_{4-k}|^2\right)\left(\sum_{k=1}^{2} |\Delta_{4-k} + i\Delta_{3k-2}|^2\right)\right]^2, \tag{2.46}$$

where $\Delta_i = c_i - e_i$. Full rank is not achieved, if at least one of the factors in (2.46) is zero. Thus, without loss of generality, we need only to consider the case when, for example, the first factor becomes zero. This factor, in turn, consists of a sum of two nonnegative terms. Hence, the sum is zero only if both terms are zero. After some manipulations we arrive at

$$|(c_3 - ic_1) - (e_3 - ie_1)|^2 + |(c_2 - ic_4) - (e_2 - ie_4)|^2$$

for the first factor. If symmetric complex signal constellations, like M-PSK, M-QAM, are used and $x_1 \dots x_4$ are from the same constellation it is possible that both x_{3k-2} and $x_{4-k} = ix_{3k-2}$, $k = \{1, 2\}$, are constellation-points. In this case, the first factor in (2.46) becomes zero and the determinant of the codeword distance matrix is consequentially zero. □

Remark 2.3.1. *It is possible to get full rank by rotating the signal constellation* $\mathcal{C}_{3,4}$ *of* x_3 *and* x_4 *by a rotation angle* ϕ *with respect to the original constellation* $\mathcal{C}_{1,2}$ *of* x_1 *and* x_2 *properly as discussed in [Tir01, SP02a, SJ03, SX02]. That is, for* x_3 *and* x_4 *it holds that* $x_l = |x_l| \exp(i\theta_{x_l}) = |x_l| \exp(i\theta_{x_l}^{\text{orig}}) \exp(-i\phi), l = \{3, 4\}$*, where* $\theta_{x_l}^{\text{orig}}$ *is the angle of* x_l *in the original (non-rotated) constellation* $\mathcal{C}_{1,2}$*. The issue of rotation is addressed later on in Section 2.3.2.2.*

In order to get full rank for $\mathbf{A}(\mathbf{c}, \mathbf{e})$ for arbitrary $\mathbf{c}$ and $\mathbf{e}, \mathbf{c} \neq \mathbf{e}$, $\mathbf{\Gamma}$ has to be as follows

$$\mathbf{\Gamma} = \mathbf{\Gamma}_{\text{rot}} = \begin{bmatrix} 1 & 0 \\ 0 & \exp(-i\phi) \end{bmatrix}, \tag{2.47}$$

where $\exp(-i\phi)$ describes the phase-shifting of the constellations $\mathcal{C}_{3,4}$ of x_3 and x_4 as depicted in Fig. 2.6 for the case of QPSK. For full rank, some constraints must be imposed on ϕ, which are discussed later.
Hence, assume that $\mathbf{\Gamma} = \mathbf{\Gamma}_{\text{rot}}$ applies for our **Approach I** . However, a joint detection of symbol-pairs is necessary in the case $\mathbf{\Gamma} = \mathbf{\Gamma}_{\text{rot}}$. For this reason $\mathbf{\Gamma}$ should differ from (2.47) in our **Approach II**, where we are interested in reducing complexity at the receiver. As a disadvantage we do not achieve full diversity. Prior to analyzing the Approaches I and II in the following sections 2.3.2.2 and 2.3.2.3, respectively, let us consider the system equation after channel-matched filtering (without loss of generality, we focus on the case with one receive antenna ($n_R = 1$) in the following), which we rewrite for convenience

$$\mathbf{y}'' = \begin{bmatrix} \alpha_1 & 0 & i\alpha_2 & 0 \\ 0 & \alpha_1 & 0 & -i\alpha_2 \\ -i\alpha_2 & 0 & \alpha_1 & 0 \\ 0 & i\alpha_2 & 0 & \alpha_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \mathbf{n}'' . \tag{2.48}$$

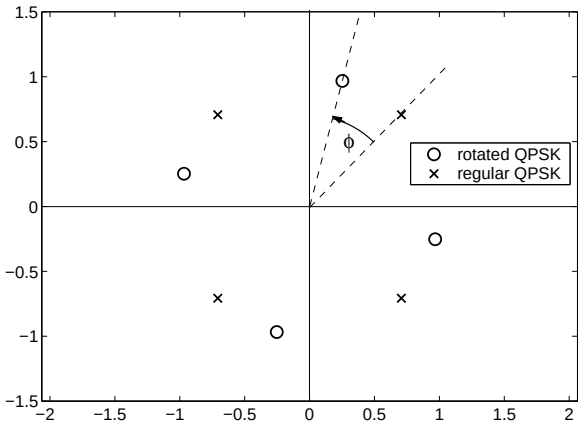

Figure 2.6: Rotation of the constellation (e.g.,QPSK) of x_3 with an angle ϕ.

Remark 2.3.2. *In the case with $\alpha_2 = 0$, we would achieve full diversity and the capacity of a $n_T = 4, n_R = 1$ system.*

Remark 2.3.3. *Note that the product $(\mathbf{H}')^H\mathbf{H}'$ is given as $(\mathbf{H}')^H\mathbf{H}' = \mathrm{diag}(\alpha_1)$, i.e. $\alpha_2 = 0$ in (2.48) if the parameter b in (2.45) is $b = 0$.*

2.3.2.2 Approach I

The aim in Approach I is to improve the BER performance of the QSTBC. For this purpose we introduce and analyze the performance criteria, namely the diversity product ζ and the global minimum Euclidean distance d_E as follows below.

Global minimum Euclidean distance d_E First of all, we need some more notation. To this end, take a look at (2.27), which can be rewritten for $n_T = 4$ as

$$\widehat{\mathbf{y}}_{13} = \begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix} \begin{bmatrix} x_1 + ix_3 \\ x_1 - ix_3 \end{bmatrix} + \mathbf{w} \; .$$

Let us consider (x_1+ix_3) and (x_1-ix_3) as "symbols" in the compound constellations $D^{(+)} = C_{1,2} + iC_{3,4}$ and $D^{(-)} = C_{1,2} - iC_{3,4}$, respectively. The symbols x_1 and x_3 are taken from the constellations $C_{1,2}$ and its rotated version $C_{3,4} = C_{1,2}\exp(-i\phi)$, respectively, where ϕ is the rotation angle. Furthermore, let us denote the minimum Euclidean distance in $D^{(+)}$ and $D^{(-)}$ from the pair (x_1, x_3) to all other pairs of symbols as $d_{\min}^{D^{(+)}}$ and $d_{\min}^{D^{(-)}}$, respectively. Note that, $d_{\min}^{D^{(+)}}$ and $d_{\min}^{D^{(-)}}$ depend on the value of ϕ. To illustrate the impact of this rotation, we take the constellation $D^{(+)}$ and increase the angle ϕ step-by-step from 0.0 rad ($*$) to 0.8 rad ($\square$) in steps of 0.1 rad as depicted in Fig 2.7. Without rotation, there are some constellation points which dwell at the same position. By rotating, these constellation points change their position and the Euclidian distances between them grow, as, for example, the Euclidian distance $d(a, b)$ between the constellation points s_a and s_b, $s_a,\ s_b \in D^{(+)}$. If the angle of rotation is too large however, the Euclidian distances between two neighboring constellation points become smaller again, like, for example, the Euclidian distance $d(b, c)$ between the constellation points s_b and s_c, $s_c \in D^{(+)}$. Therefore, an angle ϕ_{opt} exists with a maximum distance between the nearest constellation points.

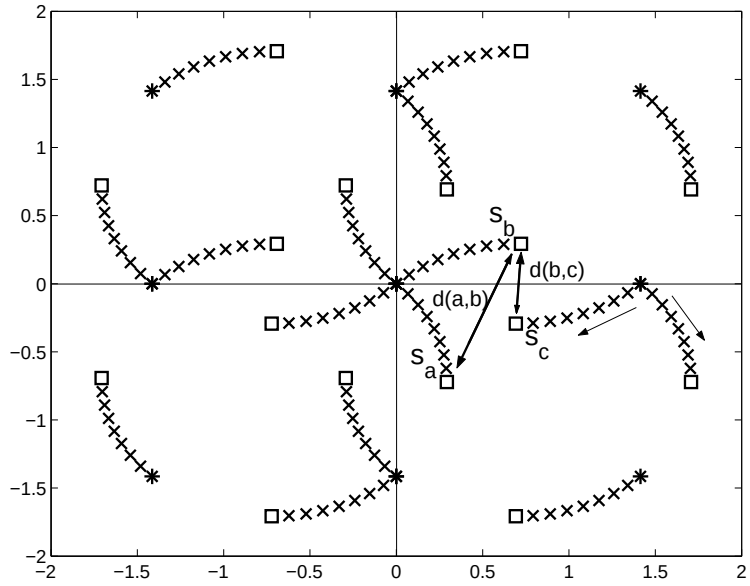

Figure 2.7: Constellation of $D^{(+)}$ for QPSK. The arrows point in the direction of increasing angles ϕ. ϕ is increased from 0.0 rad ($*$) to 0.8 rad ($\square$) in steps of 0.1 rad.

Now, let us discuss the performance criterion proposed in [SP02a, SJ03] and the alternative criterion proposed in [Tir01, SX02]. Denote

$$d_E = \min_{(x_1,x_3)} \left[d_{\min}^{D^{(+)}}, d_{\min}^{D^{(-)}} \right] \tag{2.49}$$

as the global minimum Euclidean distance in the constellations $D^{(+)}$ and $D^{(-)}$. Then, the objective in [SP02a, SJ03] is to maximize d_E by choosing the optimum angle ϕ as $\phi_{\text{opt}} = \arg\max_{\phi} (d_E)$.

Diversity product ζ As opposed to maximizing the Euclidean distance, the objective in [Tir01, SX02] is to maximize the diversity product ζ [HS00, HHSS01], which is given by

$$\zeta = \frac{1}{2\sqrt{n_T}} \min_{\substack{\mathbf{B}=\mathbf{B}(\mathbf{c},\mathbf{e}) \\ \mathbf{c}\neq\mathbf{e}}} |\det(\mathbf{B})|^{\frac{1}{n_T}} , \tag{2.50}$$

where $\mathbf{B}$ is as aforementioned the codeword difference matrix obtained from the transmit matrix $\mathbf{G}_{n_T}$ after replacing x_l by $c_l - e_l$. Further on, ζ is normalized by the factor $1/(2\sqrt{n_T})$ resulting in $0 \leq \zeta \leq 1$. Applying (2.50) to the QSTBC, using (2.43),(2.44) and (2.45), results in

$$\zeta = \frac{1}{4} \min_{\substack{x_l=c_l-e_l, \\ \mathbf{c}\neq\mathbf{e}}} |a^2 + b^2|^{\frac{1}{4}} . \tag{2.51}$$

If ζ is nonzero, we say that the code has full diversity. As mentioned in [HS00], maximizing ζ is fundamentally different from maximizing the Euclidean distance; two signals that have large Euclidean distances can have small diversity product ζ. A natural question which arises is which of the two criteria is more important? Before answering the question in the following we completely characterize the diversity

product of the QSTBC.

Lemma 2.3.2. *For an arbitrary complex signal constellation, the diversity product is upper bounded by*

$$\zeta \leq \zeta_{\text{ub}} = \frac{d_{\min}}{4} \,, \tag{2.52}$$

where $d_{\min}$ is the minimum Euclidean distance between two distinct signal points in the signal constellation. As an example, $d_{\min}$ is given as $d_{\min}^{\text{QPSK}} = \sqrt{2}$ and $d_{\min}^{\text{8PSK}} = \sqrt{2(1-1/\sqrt{2})} = 2\sin(\pi/8)$, for QPSK and 8PSK respectively.

Proof. Given (2.51), it follows that

$$\zeta = \frac{1}{4} \min_{\substack{x_l = c_l - e_l, \\ \mathbf{c} \neq \mathbf{e}}} \left| a^2 - |b|^2 \right|^{\frac{1}{4}} \,. \tag{2.53}$$

Since (2.53) decreases monotonically with respect to $|b|$, ζ is upper bounded at $|b| = 0$ by

$$\zeta \leq \zeta_{\text{ub}} = \frac{1}{4} \min_{\substack{x_l = c_l - e_l, \\ \mathbf{c} \neq \mathbf{e}}} a^{\frac{1}{2}} = \frac{d_{\min}}{4} \,. \tag{2.54}$$

In this case ($|b| = 0$), it is obvious that a achieves the minimum value $a_{\min} = d_{\min}^2$. This happens if all but one of the code symbols of two distinct codewords $\mathbf{c} \neq \mathbf{e}$ are identical, i.e. $c_l = e_l, c_m \neq e_m \forall l, l \neq m$. □

For the next theorem, we need some more notation. Let let $x_l = c_l - e_l$ be the difference between two code symbols of two distinct codewords $\mathbf{c}$ and $\mathbf{e}$, with $\mathbf{x} = \mathbf{c} - \mathbf{e}$ and $\omega := \{l | x_l \neq 0\}$. Further let $|\omega|$ be the cardinality of ω. Then the cardinality $1 \leq |\omega| \leq n_T$ provides the number of $x_l \neq 0$, e.g.,if $x_2 = 0$ and all other $x_l \neq 0, l = 1, 3, 4$, it follows that the cardinality $|\omega| = 3$. Henceforth let $\zeta_{|\omega|}$ denote the diversity product for a given $|\omega|$.

Theorem 2.3.2. *For an arbitrary complex signal constellation the diversity products ζ_2, ζ_3 and ζ_4 are lower bounded (except for ζ_1, which is exactly given) by*

$$\zeta_4 \geq 0 \,, \qquad \zeta_3 \geq \frac{1}{4} 5^{\frac{1}{4}} d_{\min} \,, \qquad \zeta_2 \geq 0 \,, \qquad \zeta_1 = \frac{1}{4} d_{\min}.$$

The proof of this theorem is given in Appendix 2.5.7.

Theorem 2.3.3. *For an arbitrary M-PSK constellation* (M > 2)*, it holds that $\zeta_4 \geq \zeta_2 \quad \forall \phi, 0 \leq \phi \leq 2\pi$, where ϕ is the rotation angle from (2.47).*

The proof of this theorem is given in Appendix 2.5.8.

Corollary 2.3.2. *For an M-PSK constellation* (M > 2)*, the diversity product of the QSTBC as a function of the rotation angle ϕ is given by*

$$\zeta = \frac{d_{\min}}{4} \cdot \begin{cases} \min\left(\sqrt{2\left|\sin(\phi - \frac{2k\pi}{M})\right|}, 1 \right) & \text{for } \phi = \phi_1 \\ \min\left(\sqrt{2\left|\sin(\phi - \frac{2(k+1)\pi}{M})\right|}, 1 \right) & \text{for } \phi = \phi_2 \,. \end{cases} \tag{2.55}$$

where $\frac{2k\pi}{M} \leq \phi_1 < \frac{(2k+1)\pi}{M}$, $\frac{(2k+1)\pi}{M} \leq \phi_2 < \frac{2(k+1)\pi}{M}$ and $k \in \mathbb{Z}$.

Proof. Again, as in the proof of theorem 2.3.3, due to symmetry of the M-PSK constellations it will suffice to analyze the range $0 \leq \phi \leq \frac{\pi}{M}$. By comparing $\zeta_1, \ldots, \zeta_4$ from (2.98), (2.99), (2.100) and (2.101), it is obvious that

$$\zeta = \frac{d_{\min}}{4} \min(\sqrt{2\,|\sin(\phi)|}, 1)\ 0 \leq \phi \leq \frac{\pi}{M}. \tag{2.56}$$

Mapping this to any $k \in \mathbb{Z}$, (2.56) can be further written as given in (2.55). □

Now, after characterizing the diversity product ζ completely, lets turn to the question which one of the performance criteria, the minimum Euclidean distance or the diversity product, is more important.

Impact of the performance criteria d_E and ζ on the BER From (2.55), we observe that the upper bound $^{d_{min}}/_{4}$ of ζ is only achieved for QPSK, but not for higher order modulation schemes. In fact, for QPSK it is possible to find an angle ϕ at which both ζ and d_E can be maximized. For illustration, ζ and d_E are depicted in Fig. 2.8 for QPSK and 8-PSK. From the figure, we observe that for QPSK ζ achieves the upper bound $^{d_{min}}/_{4}$ within the range of $\phi = \pi/6 \ldots \pi/3$. Given this, we are now able to optimize d_E within the range of $\phi = \pi/6 \ldots \pi/3$. For QPSK, the angle optimizing both ζ and d_E is $\phi = \pi/6$. For 8-PSK and higher modulation, both parameters can not be optimized simultaneously. In the case of 8-PSK ,e.g., we observe from (2.55) that the maximum of ζ is achieved at $\phi = \pi/8$. From Fig. 2.8, we notice that the optimum of ζ and d_E are achieved for different ϕ. Moreover ζ achieves the optimum for an angle ϕ at which d_E achieves a local minimum. From [BTT02b, Ion03] we know that it is appropriate to choose the ϕ optimizing ζ. Therefore, we choose $\phi = \pi/8$ for 8-PSK. For even higher modulation schemes, the angle which optimizes ζ can also be obtained from (2.55).

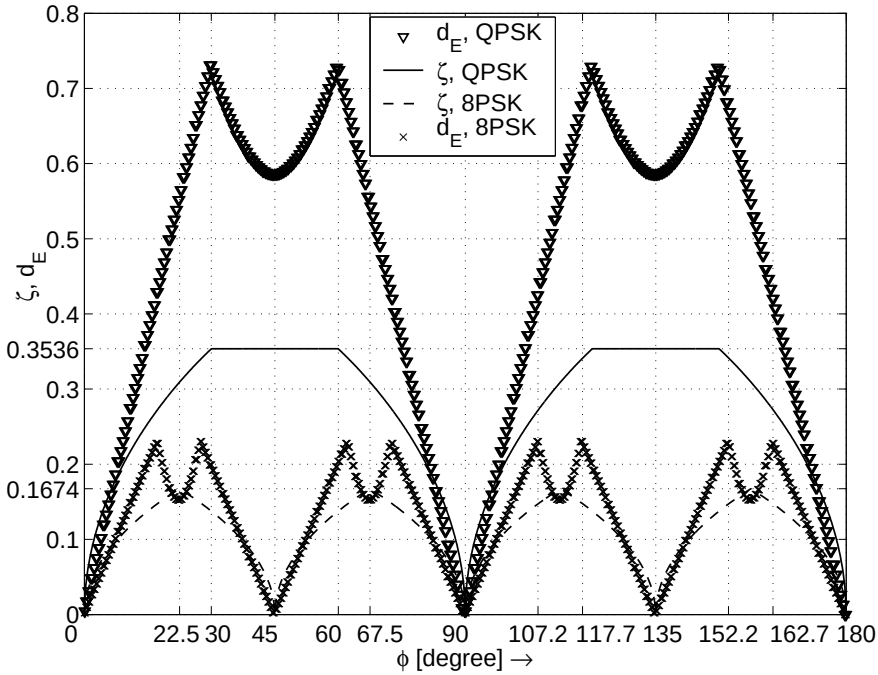

Figure 2.8: Diversity product ζ and the global minimum Euclidean distance d_E versus the rotation angle ϕ.

The impact of ζ and d_E on the BER performance is shown in Fig. 2.9. Within the range of $\phi = \pi/6 \ldots \pi/3$, where the diversity product ζ is constant, the impact of varying d_E on the BER performance in Fig. 2.9(a) for QPSK (left column) and SNR=22 dB is negligible. In contrast to the impact of d_E, the impact of ζ on the BER performance for 8PSK and SNR=27 dB is clearly visible in Fig. 2.9(b). Even though the minimum Euclidean distance d_E increases around the angle of $\phi = \pi/8$, the BER is degraded, nevertheless, by reducing the diversity product ζ. As a result,

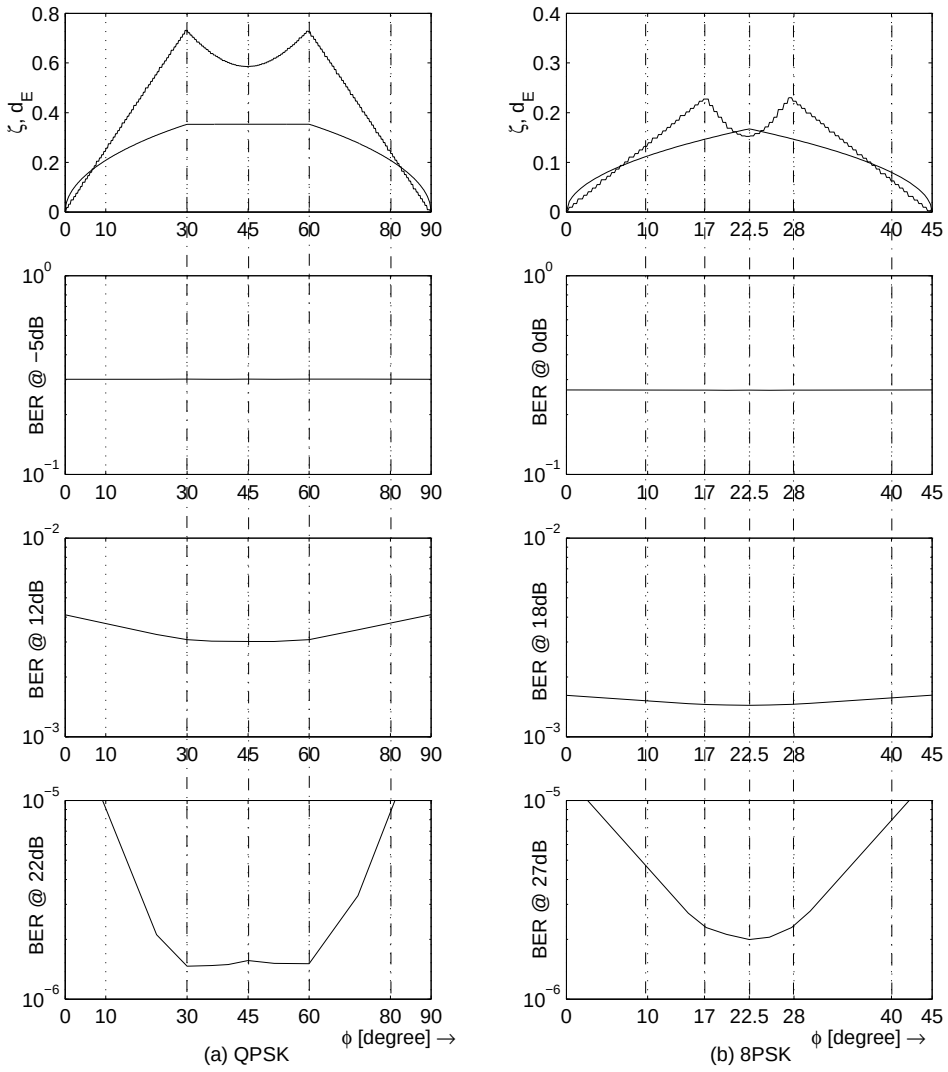

Figure 2.9: The BER performance of M-PSK for different values of ϕ and SNR, the diversity product ζ and the global minimum Euclidean distance d_E.

optimizing the diversity product is more important than optimizing the minimum Euclidian distance d_E, at least for the medium and high SNR region, since the impact of ζ on the BER vanishes for lower SNR as depicted in Fig. 2.9. The impact of d_E on the BER performance may become apparent if the ratio of the SNR to the number of receive antennas n_R is significantly smaller than one [BTT02b]. However, that does not take place here. Hence, to get the best BER-performance, which is the aim of Approach I, it is optimal for QPSK, for example, to choose ϕ in $\mathbf{\Gamma}_{\text{rot}}$ (cf.

(2.47)) as $\phi = \pi/6$ [SJ03].

This does not only provide the optimal minimum Euclidean distance, but also the highest diversity product. For 8-PSK in Approach I it is optimal to choose ϕ as $\phi = \pi/8$. In the following section, we consider Approach II for QSTBC with the objective of reducing the complexity of the system.

2.3.2.3 Approach II

The objective of this approach is to reduce the decoding complexity. We consider the selection of the matrix $\mathbf{\Gamma}$ and its impact on the detection scheme. Recall that in Approach I we chose $\mathbf{\Gamma} = \mathbf{\Gamma}_{\text{rot}}$. By substituting this in (2.27), we get

$$\widehat{\mathbf{y}}_{13} = \widehat{\mathbf{H}}\mathbf{\Gamma}_{\text{rot}}\mathbf{s}_{1,3} + \mathbf{w} \; . \tag{2.57}$$

In order to decode this optimally, the nonlinear ML-detector has to find the pair (s_1, s_3) that minimizes $||\widehat{\mathbf{y}}_{13} - \widehat{\mathbf{H}}\mathbf{\Gamma}_{\text{rot}}\mathbf{s}_{1,3}||_F^2$. However, our objective here in Approach II is to decode the transmitted symbols separately. To this end, instead of $\mathbf{\Gamma} = \mathbf{\Gamma}_{\text{rot}}$, we select $\mathbf{\Gamma}$ as follows:

$$\mathbf{\Gamma} = \mathbf{\Gamma}_{\text{lin}} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ i & i \end{bmatrix} \; . \tag{2.58}$$

With (2.58), (2.27) is given as

$$\bar{\mathbf{y}}_{13} = \widehat{\mathbf{H}} \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} + \mathbf{w} = \widehat{\mathbf{H}}\mathbf{\Gamma}_{\text{lin}} \begin{bmatrix} s_1 \\ s_3 \end{bmatrix} + \mathbf{w} = \sqrt{2} \begin{bmatrix} -\mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_3 \end{bmatrix} + \mathbf{w} \; . \tag{2.59}$$

To get rid of the negative sign of μ_1, we multiply (2.59) from left with the unitary matrix $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2]$,$\mathbf{p}_1 = [-1, 0]^T$,$\mathbf{p}_2 = [0, 1]^T$ to get

$$\bar{\mathbf{y}}_{13} = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_3 \end{bmatrix} = \underbrace{\sqrt{2} \begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix}}_{\bar{\mathbf{H}}} \begin{bmatrix} s_1 \\ s_3 \end{bmatrix} + \mathbf{w} \; . \tag{2.60}$$

At this point, the symbols s_1 and s_3 are completely decoupled, since they experience no interference from each other. In order to decode this optimally, the linear ML-detector in Approach II is able to detect the symbols transmitted from the antennas separately. In the following section we analyze the performance of the systems in (2.57)(Approach I) and (2.60)(Approach II).

2.3.2.4 Analytical BER Performance

In order to verify the simulations in the next session, here we derive analytical BER performance results for Approach I and II. First of all, we derive the exact BER performance for Approach II given in (2.60). Further on, we develop a lower and upper bound on the performance of Approach I given in (2.57). In addition to this, we discuss the case when partial CSI is available at the transmitter.

Approach II With (2.60), the instantaneous or conditional bit error probability of the estimates for s_1 and s_3 is given by $Q(\sqrt{2\gamma})$, where the instantaneous SNR per bit γ is given as $\gamma = \gamma_{\text{lin}} = \frac{(\alpha_1+\alpha_2)E_b}{N_0}$. The average SNR per bit is

$\bar{\gamma} = \bar{\gamma}_{\text{lin}} = 4E_b/N_0$. Since $\alpha_1 + \alpha_2$ (and also $\alpha_1 - \alpha_2$) is a sum of four independent exponential distributed random variables, γ_{lin} has a central-chi square distribution with four degrees of freedom and the following probability density function (pdf): $p_{\gamma_{\text{lin}}}(\gamma_{\text{lin}}) = \frac{\gamma_{\text{lin}}}{(2E_b/N_0)^2} \exp\left(-\gamma_{\text{lin}} N_0/2E_b\right)$. Generally speaking, the BER can be computed by averaging $Q(\sqrt{2\gamma})$ over the respective pdf as [Pro01] $BER = \int_0^\infty Q\left(\sqrt{2\gamma}\right) p_\gamma(\gamma)\, d\gamma$, where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{t^2}{2}\right) dt$ is the standard Gaussian tail integration function. The exact BER for Approach II in (2.60) with $n_T = 4$,$n_R = 1$ and QPSK ($M = 4$) is then given as

$$BER_{\text{lin}} = \frac{1}{2}\left(1 - \frac{1}{\sqrt{1 + \frac{2}{\bar{\gamma}_{\text{lin}}}}}\left(1 + \frac{1}{\bar{\gamma}_{\text{lin}} + 2}\right)\right) .$$

For higher constellations than QPSK, an approximation for the BER with Gray mapping may be obtained via [Pro01]

$$BER \approx \frac{1}{m} \int_0^\infty \Phi p_\gamma(\gamma)\, d\gamma , \tag{2.61}$$

where $m = \log_2(M)$, $\log_2$ is the base two logarithm, $\Phi = 2Q\left(\sqrt{2m\gamma}\sin\frac{\pi}{M}\right)$ for M-ary PSK and

$$\Phi = 1 - \left[1 - 2(1 - 1/\sqrt{M})Q\left(\sqrt{3/(M-1)m\gamma}\right)\right]^2$$

for M-ary QAM ($M > 4$), respectively.

Approach I (lower bound) In (2.57), we observe that the transmitted symbols interfere with each other at the receiver. The exact expression for the probability of error has not been not available up to now. However, in order to get an impression of the performance, we provide a lower bound (lb) on the BER by investigating the signal-to-interference-and-noise-ratio (SINR) Σ after pre-processing.

Lemma 2.3.3. *The SINR Σ after the pre-processing achieves its maximum with respect to α_2 at $\alpha_2 = 0$. In fact, in this case the interference in Σ is canceled out, so that Σ becomes the SNR.*

Proof. First of all, we are interested in a system which is fully balanced, i.e. the interference between the symbols s_1 and s_3 does not alter if we exchange them. To this end, we multiply the following unitary matrix from left with (2.57)

$$\mathbf{\Psi} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

in order to assign the SINR Σ for s_1 (and equally for s_3) as

$$\Sigma = \frac{\alpha_1 + \sqrt{\alpha_1^2 - \alpha_2^2}}{\alpha_1 - \sqrt{\alpha_1^2 - \alpha_2^2} + 2N_0/E_b} . \tag{2.62}$$

Note that the metric of the joint optimal (minimum-distance) ML-detector employed at the receiver is invariant with respect to multiplications with unitary matrices. Thus, the average BER-performance of the system in which we are interested is not altered as opposed to the BER of the substreams s_1 and s_3. Let us now investigate the denominator and the numerator separately. By considering the

denominator, it is obvious, that the dominator increases monotonically with respect to α_2. As a result the reciprocal of the denominator decreases monotonically with respect to α_2. This is also true for the numerator. The product of two monotonically decreasing positive functions is also a monotonically decreasing function. Consequently, the SINR Σ achieves its maximum with respect to α_2 at $\alpha_2 = 0$. If we substitute $\alpha_2 = 0$ into (2.62), the interference is cancelled out completely in Σ, which then becomes the SNR given as $\Sigma_{\alpha_2=0} = \frac{\alpha_1}{N_0} E_b$. □

Utilizing Lemma 2.3.3, we can apply the same procedure as for (2.60) to get a lower bound (lb) on the performance, where the instantaneous SNR per bit γ is given as $\gamma = \gamma_{\text{rot}} = \frac{\alpha_1 E_b}{N_0}$. The average SNR per bit is $\bar{\gamma} = \bar{\gamma}_{\text{rot}} = 4E_b/N_0$. Since α_1 is a sum of eight independent exponential distributed random variables, γ_{rot} has a central-chi square distribution with eight degrees of freedom and the following pdf: $p_{\gamma_{\text{rot}}}(\gamma_{\text{rot}}) = \frac{\gamma_{\text{rot}}^3}{3!(E_b/N_0)^4} \exp\left(-\gamma_{\text{rot}} N_0/E_b\right)$. Averaging $Q\left(\sqrt{2\gamma}\right)$ over this distribution yields the following lower bound (lb) for the BER with $n_T = 4$,$n_R = 1$ and QPSK ($M = 4$):

$$BER_{\text{rot}}^{\text{lb}} = \frac{1}{2}\left(1 - \frac{\frac{1}{32}\sqrt{\bar{\gamma}_{\text{rot}}}\left(\frac{35}{2}\bar{\gamma}_{\text{rot}} + \frac{7}{2}\bar{\gamma}_{\text{rot}}^2 + \frac{1}{4}\bar{\gamma}_{\text{rot}}^3 + 35\right)}{\left(\frac{1}{4}\bar{\gamma}_{\text{rot}} + 1\right)^{7/2}}\right) .$$

For the general case of a QSTBC with $n_T = 2^n$ transmit and n_R receive antennas, the lower bound on the BER is given as

$$BER_{\text{rot}}^{\text{lb}}(n_T, n_R) = \frac{1}{2} - \frac{\sqrt{\frac{\bar{\gamma}_{\text{rot}}}{n_T n_R}} F\left([\tfrac{1}{2}, \tfrac{1}{2} + n_T n_R], \tfrac{3}{2}, -\frac{\bar{\gamma}_{\text{rot}}}{n_T n_R}\right) \Gamma(n_T n_R + \tfrac{1}{2})}{\sqrt{\pi}\Gamma(n_T n_R)} ,$$

where $F([\cdot,\cdot],\cdot,\cdot)$ denotes the hypergeometric function [GR83, p.1039, eq. 9.100]. Again, as explained in the last subsection, for M-ary PSK and M-ary QAM ($M > 4$), the BER is well approximated as given in (2.61).

Remark 2.3.4. *From Remark 2.3.3, we know that $\alpha_2 = 0$ applies if the parameter $b = 0$. Thus, the lower bound derived in this section is tight for codewords $\mathbf{c} \in \mathcal{X}$ such that $b(\mathbf{c}) = 0$.*

Remark 2.3.5. *As we have seen, the case of a nonzero α_2 deteriorates the performance of the system. If the transmitter had partial CSI, we would be able to employ predistortion (by using the phase of the respective channel entry at each transmit antenna) prior to transmitting the signals in order to obtain an α_2 which is always zero. The big advantage of having partial CSI at the transmitter is that we have the full diversity as in Approach I, since $\alpha_2 = 0$, and also the linear ML-detector as in Approach II.*

Approach I (upper bound) The union bound on the bit error probability is given as [SA00]

$$BER_{\text{rot}}^{\text{ub}} \leq \frac{1}{|\mathcal{X}|} \sum_{\mathbf{c}\in\mathcal{X}} \sum_{\mathbf{e}\in\mathcal{X}\setminus\{\mathbf{c}\}} \frac{w}{N} P(\mathbf{c} \to \mathbf{e}) , \tag{2.63}$$

where $|\mathcal{X}|$ denotes the cardinality of $\mathcal{X} \subseteq \mathbb{C}^{n_T}$, w is the number of bit differences between any two distinct codewords $\mathbf{c}$ and $\mathbf{e}$, and N ($w \leq N$) denotes the number of bits in $s_1, \ldots, s_{n_T}$ in a codeword. The pairwise error probability (PEP) may be

upper bounded by eq. (1.4) [TSC98], which we repeat here for convenience

$$P(\mathbf{c} \to \mathbf{e}) \leq \frac{1}{2} \prod_{j=1}^{n_R} \left(\prod_{l=1}^{r} \frac{1}{1 + \frac{\rho}{4n_T}\mu_l} \right), \tag{2.64}$$

where ρ is the SNR and $\mu_1, \ldots, \mu_r$, $r \leq n_T$, are the nonzero eigenvalues of the distance matrix $\mathbf{A}(\mathbf{c}, \mathbf{e})$. With full diversity, the eigenvalues are given as $\mu_{l,l+1} = a \mp |b|$, for $l = \{1, 3\}$, with a and b given in (2.44) and (2.45) respectively. After some manipulations we arrive at

$$P(\mathbf{c} \to \mathbf{e}) \leq \frac{1}{2} \left(\left(1 + \frac{\rho}{4n_T} a \right)^2 - \left(\frac{\rho}{4n_T} |b| \right)^2 \right)^{-2n_R}. \tag{2.65}$$

This is the expression for the PEP after applying the Chernoff-Bound. Unfortunately, the standard Chernoff-Bound is not tight. To this end, in [BV01, SFG02] a new bound was proposed for the PEP without resorting to a Chernoff-Bound, which is asymptotically tighter for higher SNRs. Interestingly, in [BV01, Corollary 1] the asymptotic difference to the Chernoff-Bound is given as $10\log_{10}(4) - \frac{10}{n_T n_R}\log_{10}\binom{2n_T n_R}{n_T n_R}$ in decibels. Applying this corollary, after some manipulations we arrive at

$$P(\mathbf{c} \to \mathbf{e}) \leq \frac{1}{2} \left(\left(1 + \frac{\rho \binom{2n_T n_R}{n_T n_R}^{-\frac{1}{n_T n_R}}}{n_T} a \right)^2 - \left(\frac{\rho \binom{2n_T n_R}{n_T n_R}^{-\frac{1}{n_T n_R}}}{n_T} |b| \right)^2 \right)^{-2n_R}. \tag{2.66}$$

This new upper bound is tighter for low SNR and as tight as the asymptotic bound in [BV01] for higher SNRs. With (2.66) and (2.63) we get an upper bound $BER_{\text{rot}}^{\text{ub}}$ for the BER of Approach I.

A simple approximation of (2.66) is given as follows. Starting at (2.64), for high SNR and full diversity we get

$$P(\mathbf{c} \to \mathbf{e}) \leq \frac{1}{2} \det(\mathbf{A})^{-n_R} \binom{2n_T n_R}{n_T n_R} \left(\frac{\rho}{n_T} \right)^{-n_T n_R}. \tag{2.67}$$

With (2.50), (2.63) and (2.67) we obtain $BER_{\text{rot}}^{\text{ub}} < \frac{1}{8}(|\mathcal{X}| - 1)\binom{2n_T n_R}{n_T n_R}\left(4\zeta^2\rho\right)^{-n_T n_R}$. Thus, the impact of the diversity product ζ is obvious and should be as large as possible.

Remark 2.3.6. *Interestingly, the argument of the mutual information given in the proof of corollary 2.3.1 at the end of this chapter is similar to the argument of the Chernoff-Bound for the PEP in (2.65). In the case $b = 0$, we would minimize the PEP and we would obtain the actual mutual information due to Remark 2.3.3.*

Some simulation results and their interpretation are presented in the following section.

2.3.2.5 Numerical simulations

In Fig. 2.10, the BER of the quasi-orthogonal scheme from Jafarkhani [Jaf01] (ML-detector with $\mathbf{\Gamma} = \mathbf{I}_2$), the BER of Approach I from [Tir01, SP02a, SX02, SJ03] (ML-detector with $\mathbf{\Gamma} = \mathbf{\Gamma}_{\text{rot}}$, given in (2.47)) and the BER of Approach II with its linear detector for a system with $n_T = 4$ transmit and $n_R = 1$ receive antennas and

QPSK modulation are depicted. In addition to this, the analytical results derived

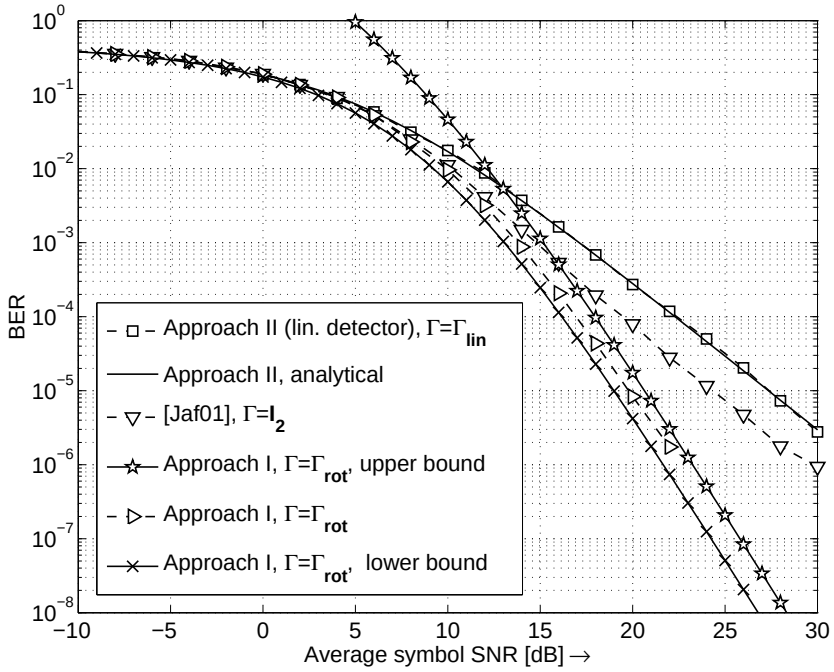

Figure 2.10: BER of the quasi-orthogonal scheme ($n_T = 4, n_R = 1$) with linear and nonlinear detectors, uncoded QPSK modulation.

in Section 2.3.2.4 for the BERs are also depicted in Fig. 2.10. As can be seen in Fig. 2.10, the ML-detector in [Jaf01] and the Approach I (also from [SJ03]) outperforms Approach II with its linear ML-detector. Note that the complexity of the latter is considerably lower in comparison to the detectors in [Jaf01, SJ03]. From this it follows that there is a tradeoff between the receiver-complexity and performance of the schemes. Furthermore, we observe that the analytical results from section 2.3.2.4 fit very well with the simulation of the BER as depicted in Fig. 2.10. Further on, the lower bound on the performance from section 2.3.2.4 is equal to the performance of the system, where the transmitter has partial CSI. Interestingly, the performance is also equal to the rate 3/4 OSTBC for four transmit antennas in [TH02], provided that the OSTBC obeys the energy constraint $E[||G_4||^2] = Tn_T$ ($E[\cdot]$ denote expectation), i.e. the transmit matrix of the OSTBC is multiplied by a factor of $\sqrt{4/3}$.

Up to now, we allowed the receiver to have an arbitrary number of receive antennas. However, multiple receive antennas also means multiple RF chains. But in some cases it may be impractical or even undesirable to have multiple RF chains at the receiver, which is discussed in the following section.

2.3.3 Antenna Selection

It is widely known that in MIMO systems the capacity increases linearly with the minimum number of transmit and receive antennas. However, multiple antenna deployment at mobile handsets requires multiple RF chains (analog-digital convert-

ers low noise amplifiers, downconverters, etc.), which is undesired in systems where the handsets are intended to remain simple and inexpensive. In order to reduce the natural drawbacks of MIMO systems such as hardware costs and increased complexity, antenna selection (AS) was proposed (see [HSP01, GGP03b, GGP03a, GHP02, MW04] and references therein) in combination with spatial multiplexing schemes. For example, by applying antenna selection at the receiver, only the signal of one out of n_R possible receive antennas (in the case of single antenna selection) is fed to the RF chain according to a given selection criterion.

AS was also combined with OSTBC in [GP02, WL03, CVZ03]. In [GP02], the selection criterion was based on the maximization of the channel Frobenius norm. With OSTBC, this criterion is equivalent to minimizing the error probability. In this section, we analyze the impact of AS based on different selection criteria on both the mutual information and the bit error rate of QSTBC. Furthermore, we derive a lower bound for the outage probability and an upper bound on the average mutual information achieved with QSTBC and AS.

2.3.3.1 Impact of AS on the system model

Similar to the system model description in section 2.2.1, we consider a system with n_T transmit antennas. Since there is only one RF chain at the receiver, we are constrained to use only one out of n_R receive antennas. The receive antenna i ($i \in \{1, \ldots, n_R\}$) in use is determined according to an AS criterion applied at the receiver. The system model for each receive antenna is then defined by

$$\mathbf{y}_i = \mathbf{G}_{n_T}\mathbf{h}_i + \mathbf{n}_i \ , \tag{2.68}$$

where $\mathbf{y}_i$ is the receive vector at receive antenna i, $\mathbf{h}_i$ is the channel vector between the transmitter and receive antenna i, and $\mathbf{n}_i$ is the complex white Gaussian noise (AWGN) vector at receive antenna i. As we know from the derivations in section 2.2.3, the eq. (2.68) can be rewritten such that we arrive at an equivalent channel representation. With AS, we have an equivalent channel given as $\widehat{\mathbf{H}}_i$ for each receive antenna i.

2.3.3.2 Selection Criteria (SC)

From all available n_R receive antennas, the antenna in use is determined according to one of the following selection criteria.

Selection Criterion One (SC1)-Min. condition number:
Compute the condition number of $\widehat{\mathbf{H}}_i$ (ratio of the largest singular value of $\widehat{\mathbf{H}}_i$ to the smallest) and choose the receive antenna with the smallest condition number for every receive antenna i. The performance of linear receivers is strongly influenced by the inverse of the channel matrix $\widehat{\mathbf{H}}_i$. If the channel has a very low condition number (near one), i.e. the channel matrix is almost orthogonal, the phase distortion of the noise due to post-processing is negligible resulting in fewer errors.

SC2-Maximum mutual information:
For every receive antenna i compute

$$I_Q^i = \frac{2}{n_T}\log_2 \det\left(\mathbf{I}_{n_T/2} + \frac{\rho}{n_T}\mathbf{D}_i^H\mathbf{D}_i\right) \ , \tag{2.69}$$

which is the portion of the mutual information achieved with QSTBC (cf. eq. (2.33)).

Choose the receive antenna with the largest I_Q^i. The assumption here is that an optimal receiver is used, resulting in a performance loss for suboptimal receivers.

SC3-Maximum smallest eigenvalue:
Compute the eigenvalues μ_{ji} of the equivalent channel for each receive antenna i. Choose the receive antenna with the largest minimum eigenvalue, i.e

$$\max_i \min_j \mu_{ji} \ .$$

The noise amplification depends strongly on the minimum eigenvalue of the channel matrix. By taking the channel with the largest minimum eigenvalue, this amplification is reduced and the performance is improved, especially for linear receivers, where noise amplification is the single most important problem.

2.3.3.3 Impact of AS on mutual information

Unfortunately, the exact analysis of (2.69) with respect to outage and ergodicity is not available. However, by using the trace-determinant inequality $\det(\mathbf{A})^{1/n} \leq \frac{1}{n}\mathrm{tr}(\mathbf{A})$, where $\mathbf{A}$ is a positive semidefinite matrix, we have the following upper bound

$$I_Q^i \leq I_{Q,\mathrm{ub}}^i = \log_2\left(1 + \frac{\rho}{n_T}||\mathbf{h}_i||^2\right) \ .$$

Since $X_i = ||\mathbf{h}_i||^2$, $\forall i$, are independent identically chi-square distributed random variables with $2n_T$ degrees of freedom, i.e. with the following probability density function (pdf)

$$f(x_i) = \frac{x_i^{n_T-1}e^{-x_i}}{(n_T-1)!}$$

and the following cumulative distribution function (cdf)

$$F(x_i) = 1 - \sum_{l=0}^{n_T-1} \frac{x_i^l e^{-x_i}}{l!} \ ,$$

for $Z = \max_i X_i$ the pdf is given as

$$p_Z(z) = n_R F(z)^{n_R-1} f(z) \ . \tag{2.70}$$

Outage probability The outage probability P_{out} achievable with QSTBC and AS is defined as the probability that the maximum of $I_Q^i, \forall i$ is smaller than a certain rate R, i.e.

$$P_{out}(R, n_T, n_R, \rho) = \Pr[\max_i I_Q^i < R].$$

Since, the $||\mathbf{h}_i||^2$, $\forall i$, are pairwise independent, we have

$$P_{out}(R, n_T, n_R, \rho) = \prod_i^{n_R} \Pr[I_Q^i < R] = \Pr[I_Q^i < R]^{n_R} \geq \Pr[I_{Q,\mathrm{ub}}^i < R]^{n_R}.$$

After some manipulations we arrive at a lower bound on the exact P_{out} with AS given as

$$P_{out}(R, n_T, n_R, \rho) \geq \left(1 - \frac{\Gamma\left(n_T, (2^R-1)\frac{n_T}{\rho}\right)}{\Gamma(n_T)}\right)^{n_R} \ ,$$

where $\Gamma(\cdot,\cdot)$ and $\Gamma(\cdot)$ are the incomplete and the complete Gamma function [GR83, p.940,8.350(2)], respectively.

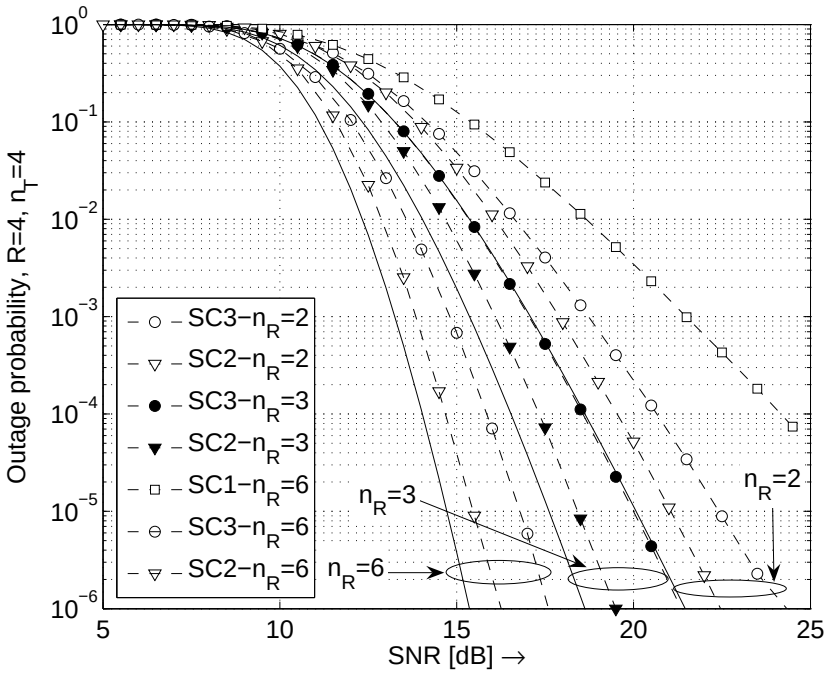

Figure 2.11: Outage probabilities of QSTBC (dashed lines) with antenna selection (AS) according to SC1, SC2 or SC3 and lower bounds (solid lines) for $n_T = 4$ and one out of $n_R = \{2, 3, 6\}$ receive antennas.

In Fig. 2.11, P_{out} of QSTBC with $n_T = 4$ transmit and AS by selecting one out of $n_R = \{2, 3, 6\}$ receive antennas are depicted. In addition the lower bounds are depicted for these three cases ($n_R = \{2, 3, 6\}$). Since the performance improvement for SC1 is negligible, only the curve for $n_T = 6$ is depicted. However, by applying the other selection criteria, SC2 and SC3, the performance of the scheme improves significantly, whereby SC2 performs best. The slope of the outage probability curves indicate that the SC2 is almost the same as if all n_R antennas were used. In the case of SC3, there is a loss of diversity order, whereas with SC3 it does not make sense to apply AS from the outage probability point of view. Furthermore, the lower bounds on the performance of QSTBC with respect to P_{out} perform very well (especially for SC2) and show to be useful.

Ergodic mutual information Using the multinomial expansion, we can rewrite (2.70) as

$$p_z(z) = n_R \sum_{k=0}^{n_R-1} \sum_{\substack{n_0,n_1,\ldots,n_{n_T-1}\geq 0 \\ \sum_i n_i = k}} (-1)^k \binom{n_R-1}{k} \frac{k!}{n_0! n_1! \ldots n_{n_T-1}!} \frac{x^{\beta_1^k + n_T - 1} e^{-x(k+1)}}{\beta_2^k (n_T-1)!} \tag{2.71}$$

where

$$\beta_2^k = \prod_{l=0}^{n_T-1} (l!)^{n_{l+1}} \quad \text{and} \quad \beta_1^k = \sum_{l=0}^{n_T-1} n_{l+1} l \; .$$

Averaging $\max_i I^i_{Q,\mathrm{ub}}$ over the pdf in (2.71) results in an upper bound on the ergodic mutual information given as

$$\begin{aligned} C_Q \leq & \mathbf{E}\left[\max_i I^i_{Q,\mathrm{ub}}\right] \\ = & \frac{n_R}{\ln(2)} \sum_{k=0}^{n_R-1} \sum_{\substack{n_0,n_1,\ldots,n_{n_T-1}\geq 0 \\ \sum_i n_i = k}} \sum_{i}^{\beta_1^k+n_T-1} (-1)^k \binom{n_R-1}{k} \frac{k!}{n_0! n_1! \ldots n_{n_T-1}!} \frac{(\beta_1^k+n_T-1)!}{(k+1)^{i+1}} \\ & \frac{1}{\beta_2^k (n_T-1)!} \left(\frac{n_T}{\rho}\right)^{\beta_1^k+n_T-i-1} e^{\frac{n_T}{\rho}(k+1)} \Gamma\left(1-(\beta_1^k+n_T-i), \frac{n_T}{\rho}(k+1)\right). \end{aligned} \tag{2.72}$$

In Fig. 2.12, the ergodic mutual information achievable with QSTBC with $n_T = 4$ transmit and AS by selecting one out of $n_R = \{2, 6\}$ receive antennas are depicted. In addition the upper bound in (2.72) is depicted for these two cases ($n_R = \{2, 6\}$). Similar to the outage probability, from all SC the SC2 shows the best performance almost achieving the upper bound for both cases $n_R = 2$ and $n_R = 6$.

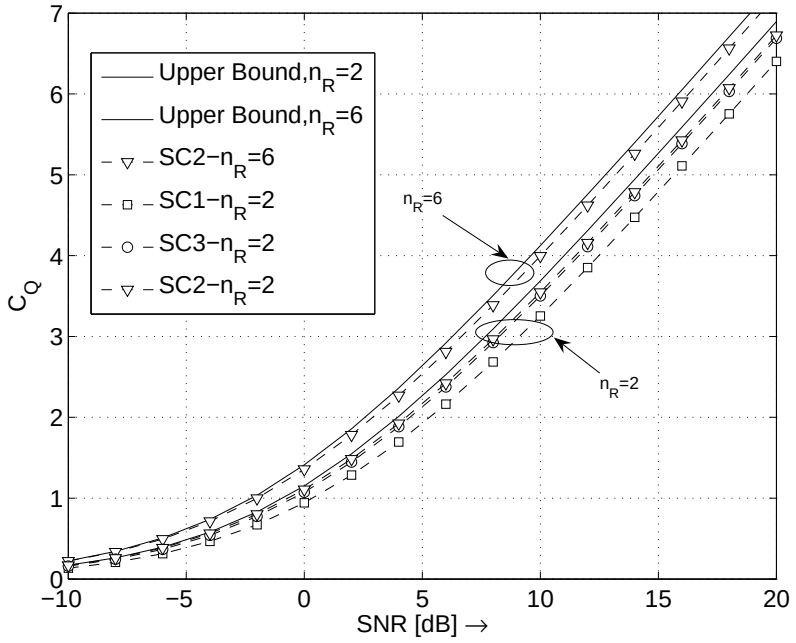

Figure 2.12: Average mutual information of QSTBC (dashed lines) with antenna selection (AS) according to SC1, SC2 or SC3 and upper bounds (solid line) for $n_T = 4$ and one out of $n_R = \{2, 6\}$ receive antennas.

2.3.3.4 Simulations

In this section the BER performance of the three SCs from section 2.3.3.2 are compared for a system with $n_T = 4$ transmit antennas and QPSK modulation. At the receiver, one out of $n_R = 2$ receive antennas is selected according to one of the SCs. In addition to the BERs with AS, the BERs without AS for $n_T = 4$ and $n_R = 1$ are also depicted for comparison.

ZF receiver In Fig. 2.13, the BERs of QSTBC with linear ZF-detection and AS according to one of the SC are depicted. For comparison purposes, also the BER performance of the ZF- and ML-detector without AS are depicted. From the figure, we observe that the diversity gain of the BER with AS for ZF is higher than the BER of ZF without AS, whereby ZF with SC3 achieves the best performance followed by ZF with SC2. The reason for this is that the ZF with SC3 has the lowest noise amplification, which is very crucial for linear detectors. The performance of ZF with SC1 is worse in comparison to the other SCs. Although the phase of the noise is less distorted with this SC obtaining a more "orthogonalized" channel matrix, the noise amplification is not reduced. However, the performance of ZF-SC3 is equal to that of the ML-detector without AS, i.e. with only one extra receive antenna, a ZF-detector with AS can achieve the performance or even outperform the ML-detector without AS with only a fraction of the computational complexity of the ML-detector.

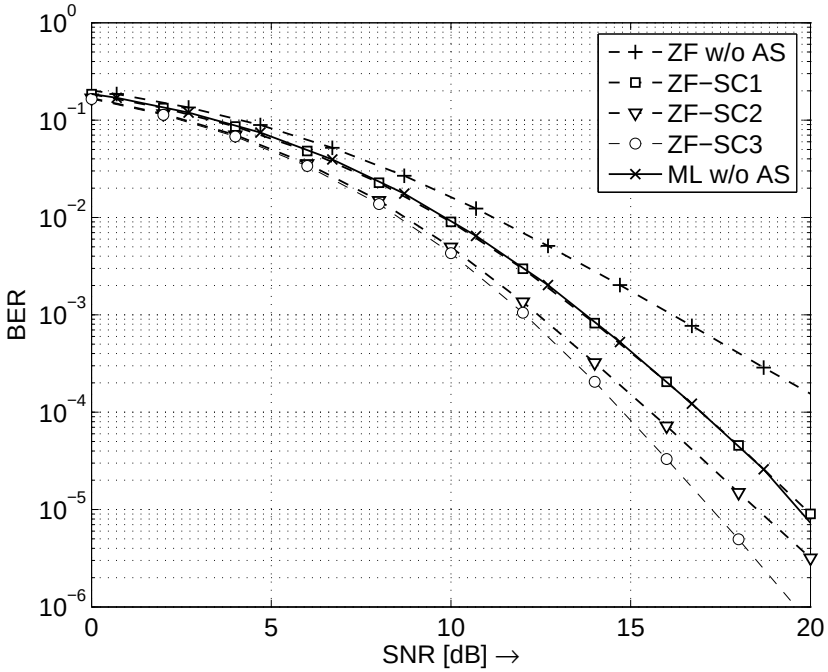

Figure 2.13: BER of QSTBC with and without antenna selection (AS) according to SC1, SC2 or SC3 for ZF-detection.

ML receiver The BERs of QSTBC with ML-detection and AS are depicted in Fig. 2.14. The BER performance of the ML-detector without AS is also depicted. As in the case of linear ZF-detection, we observe that the diversity gain of the BER with AS for ML is higher than the BER of ML without AS. Due to the optimal selection based on the expression for the mutual information, the SC2 achieves, differently from the ZF case, the best performance followed by SC3 in this case of optimal ML-detection. The performance of the ML-detector with SC1 is slightly better than the ML-detector without AS, which shows that the phase distortion on the noise caused by the channel has only small impact on the performance of the ML-detector.

In this section, we have shown that the BER performance of QSTBC employing a suboptimal ZF-detector with antenna selection outperforms an optimal ML-

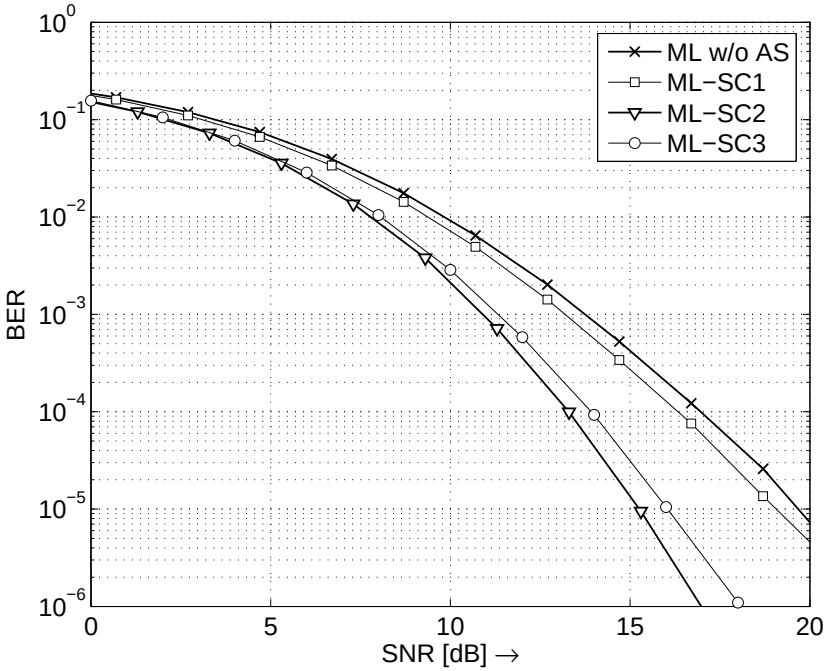

Figure 2.14: BER of QSTBC with and without antenna selection (AS) according to SC1, SC2 or SC3 for ML-detection.

detection scheme without antenna selection by using only one extra antenna. Furthermore, the suboptimal detector reduces hardware costs for multiple RF chains and, most important, computational complexity. The combination of QSTBC with other suboptimal detection schemes and the comparison with spatial multiplexing schemes is provided in the next section.

2.3.4 Suboptimal detection of QSTBC

There is a huge amount of suboptimal detectors with low complexity in the literature, linear detectors like zero-forcing (ZF) or minimum mean square error (MMSE) and nonlinear detectors like e.g. VBLAST [WFGV98]. Unfortunately, these detectors significantly sacrifice performance in terms of the bit-error-rate (BER). Recently, lattice reduction (LR) aided detection in combination with suboptimal detectors and M-QAM modulation has been proposed by Yao and Wornell in order to improve the performance of multi antenna systems [YW02] employing spatial multiplexing (SM) schemes. The lattice reduction algorithm proposed in [YW02] is optimal, but works only for MIMO systems with two transmit and two receive antennas. In [WF03], the work of [YW02] was extended to systems with more transmit and receive antennas, using the sub-optimal LLL [LLL82] lattice reduction algorithm. In [WBKK04], the LR-aided schemes in [WF03] were adopted to the MMSE criterion. Note that the error rate curves of all these LR detectors are parallel to those for maximum likelihood (ML) detection with only some penalty in power efficiency.

Since the power penalty is somewhat higher for linear schemes, the authors in [YW02, WF03, WBKK04] have deployed non-linear schemes in order to reduce this gap between ML and LR-aided detection. In this section, we employ QSTBC at the transmitter without any knowledge of the channel state information (CSI) at the

transmitter in order to reduce this gap instead of non-linear techniques at the receiver. Furthermore, we compare the performance of SM and QSTBC with LR-aided linear ZF and ML detectors respectively. The performance of QSTBC with regular linear ZF and MMSE detectors was analyzed in [RM02, RMG03].

2.3.4.1 Transmission Schemes and LR-aided linear ZF (LR-ZF) detection

In order to apply LR-aided detection, we have to put a constraint on the modulation in use, i.e. we assume that the transmit matrix $\mathbf{G}_{n_T}$ has the entries $x_1, \ldots, x_{n_T} \in \mathcal{C}$, which are elements of the vector $\mathbf{x}$, where $\mathcal{C} \subseteq \mathbb{C}$ denotes a complex M-QAM modulation signal set. Furthermore, we consider the case of $n_T = 4$ transmit antennas, which is the smallest QSTBC. As mentioned earlier in this chapter, the generalization to higher n_T is rather straightforward and does not bring any new insight into the analysis.

Spatial Multiplexing (SM) For SM, the transmit matrix $\mathbf{G}_{n_T}$ is reduced to $\mathbf{x}$, since $T = 1$. In order to apply the suboptimal LR for SM, the system model in (2.1) has to be rewritten as a real model [WF03] of the form

$$\mathbf{y}_E = \mathbf{H}_E^{SM} \begin{bmatrix} \Re\{\mathbf{x}\} \\ \Im\{\mathbf{x}\} \end{bmatrix} + \mathbf{n}_E \ ,$$

where

$$\mathbf{y}_E = \begin{bmatrix} \Re\{\mathbf{y}\} \\ \Im\{\mathbf{y}\} \end{bmatrix} , \mathbf{n}_E = \begin{bmatrix} \Re\{\mathbf{n}\} \\ \Im\{\mathbf{n}\} \end{bmatrix} , \text{ and } \mathbf{H}_E^{SM} = \begin{bmatrix} \Re\{\mathbf{H}\} & -\Im\{\mathbf{H}\} \\ \Im\{\mathbf{H}\} & \Re\{\mathbf{H}\} \end{bmatrix} .$$

In the following, we refer to $\mathbf{H}_E^{SM}$ as the equivalent channel model for the SM scheme.

QSTBC The following strategy is adopted for QSTBC. First we apply the rate one QSTBC for $n_T = 4$ transmit antennas given in (2.4). In addition, in order to get at a higher rate QSTBC, we puncture the last two columns of the transmit matrix of the rate one QSTBC in order to arrive at a rate $n_T/2$ QSTBC. Both codes are shown in the following.

Code Rate One QSTBC Contrary to SM, for the rate one QSTBC it is not necessary to resort to the real system model. With this QSTBC, the system model can be decomposed as shown in the example 2.2.2 in the section 2.2.3.4 such that the iterative optimal algorithm in [YW02] for a system with $n_T = 2$ transmit antennas can be applied. A disadvantage of this QSTBC is that in order to achieve the same transmission rate as SM, we have to compensate the rate loss by using a considerably higher constellation. But recall that higher constellations complicate amplification, synchronization, and detection. For example, a transmission rate of 4 bits/sec/Hz for a system with $n_T = 4$ transmit antennas is achieved by SM with BPSK, whereas 16QAM is required for the code rate one QSTBC. For this reason, QSTBC with higher code rates are necessary.

Code Rate $n_T/2$ QSTBC In order to increase the code rate, we puncture the last two columns of the rate one QSTBC, resulting in the following QSTBC with $T = 2$

$$\mathcal{G}_4(\mathbf{x}) = \begin{bmatrix} x_1 & x_2^* & x_3 & x_4^* \\ x_2 & -x_1^* & x_4 & -x_3^* \end{bmatrix}^T .$$

The equivalent signal model for this QSTBC is given as

$$\mathbf{y}'' = \mathbf{H}_E^{Q2} \begin{bmatrix} \Re\{\mathbf{x}\} \\ \Im\{\mathbf{x}\} \end{bmatrix} + \mathbf{n}'' ,$$

where

$$\mathbf{H}_E^{Q2} = \begin{bmatrix} \Re\{\hat{\mathbf{H}}_E^{Q2}\} & -\Im\{\hat{\mathbf{H}}_E^{Q2}\} \\ \Im\{\hat{\mathbf{H}}_E^{Q2}\} & \Re\{\hat{\mathbf{H}}_E^{Q2}\} \end{bmatrix} .$$

$\hat{\mathbf{H}}_E^{Q2} = [(\hat{\mathbf{H}}_{E,1}^{Q2})^T, (\hat{\mathbf{H}}_{E,2}^{Q2})^T, \ldots, (\hat{\mathbf{H}}_{E,n_R}^{Q2})^T]^T$ and $\hat{\mathbf{H}}_{E,i}^{Q2}$ is given as

$$\hat{\mathbf{H}}_{E,i}^{Q2} = \begin{bmatrix} h_{1i} & h_{2i}^* & h_{3i} & h_{4i}^* \\ -h_{2i} & h_{1i}^* & -h_{4i} & h_{3i}^* \end{bmatrix} ,$$

which is the equivalent channel from the transmitter to the ith receive antenna. This QSTBC is nothing else than a 2-times stacked Alamouti scheme. The more general n-times stacked Alamouti scheme will be analyzed with respect to the mutual information in section 2.4.

LR-aided linear ZF Detection By applying the LLR algorithm, the $m \times n$ equivalent channel $\mathbf{H}_E$ for each transmission scheme can be decomposed as

$$\mathbf{H}_E = \mathbf{QR} , \tag{2.73}$$

where $\mathbf{R}$ is a $n \times n$ matrix with integer entries and $\mathbf{Q}$ is a $m \times n$ matrix, which is better conditioned than $\mathbf{H}_E$, i.e. the columns of $\mathbf{Q}$ are less correlated and shorter. A good indication for the correlation of a matrix is the so called condition number, which is defined as the ratio of the largest to the smallest singular value of the matrix. Using (2.73), the equivalent signal model is then given as

$$\mathbf{y} = \mathbf{H}_E\mathbf{x}_r + \mathbf{n} = \mathbf{QRx}_r + \mathbf{n} = \mathbf{Qz} + \mathbf{n} .$$

Now, by multiplying $\mathbf{Q}^{-1}$ from left to $\mathbf{y}$ we arrive at

$$\tilde{\mathbf{y}} = \mathbf{z} + \mathbf{Q}^{-1}\mathbf{n} ,$$

where the noise enhancement and coloring is relatively low, since $\mathbf{Q}^{-1}$ is also good conditioned. In order to get a estimation for the transmitted symbols, the following operation has to be applied

$$\hat{\mathbf{x}} = C\left(\mathbf{R}^{-1}\mathcal{Q}_{\mathbb{Z}^n}\left[\frac{1}{C}\tilde{\mathbf{y}} - \mathbf{R}\frac{1}{2}\mathbf{1}_n\right] + \frac{1}{2}\mathbf{1}_n\right) ,$$

where $\mathbf{1}_n$ is a $n \times 1$ vector of ones, C is a constant given as $C = \sqrt{\frac{6}{M-1}}$ and $\mathcal{Q}_{\mathbb{Z}^n}[\cdot]$ describes the component-wise quantization with respect to the infinite integer space $\mathbb{Z}$. However, this quantization can only be applied, if the transmit modulation signal set $\mathcal{C}$ is transformed to $\mathbb{Z}$, which is achieved by scaling and shifting $\tilde{\mathbf{y}}$ within the quantization operation. Note that after this quantization some points may lie outside the constellation. A suboptimal solution is to assign these points to the

nearest point within the constellation. For BPSK, the effect of this assignment has a significant effect on the error rate performance, however, this gain diminishes with higher order modulations.

2.3.4.2 Condition number

For illustration, the probability density functions (pdfs) of the natural logarithm

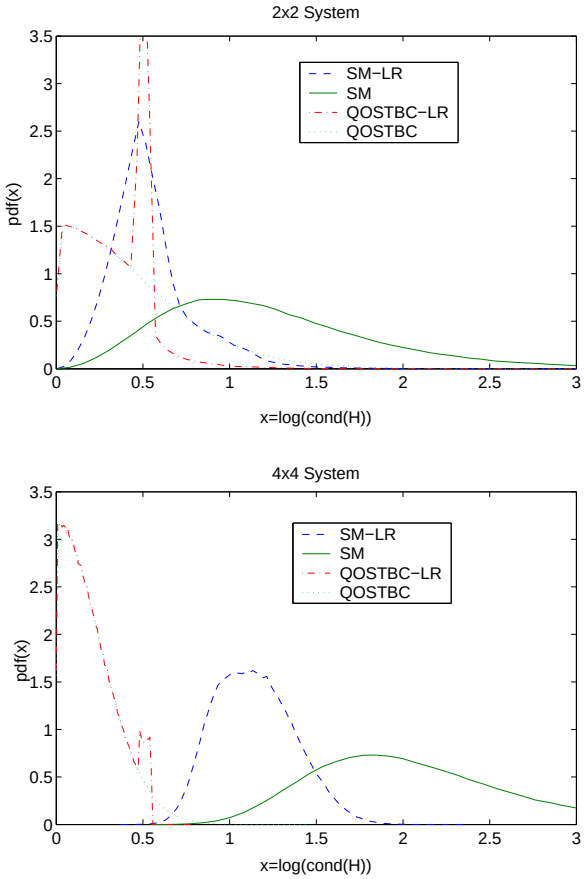

Figure 2.15: Pdfs of channel cond. numbers with SM or code rate one QSTBC with and w/o LR.

of the condition number of channels for SM and for the code rate one QSTBC are depicted in Fig. 2.15. In the upper subplot, the pdfs of a system with $n_T = n_R = 2$ are depicted, first with complex Gaussian distributed entries for SM and secondly with entries as in $\mathbf{H}_E^{Q1}$ (We actually used $n_T = 4, n_R = 1$ here for QSTBC) before and after LR. The same was done for a system with $n_T = n_R = 4$ (which reduces again to the 2×2 channel for QSTBC) in the lower subplot. From both subplots, we observe that the SM-channel is conditioned poorly and that LR has a great impact on the channels. For $n_T = n_R = 2$ the improvement is even better than in the case of $n_T = n_R = 4$. For QSTBC, the impact of LR is not as strong as for SM. Furthermore, for some channels we have no gain with LR, since many samples of the equivalent channel model generated with QSTBC have inherently low condition numbers so that the LR has no effect. Note that for orthogonal channels (e.g. with OSTBC), the pdf is a dirac impulse at position 0. The pdf of the natural logarithm of the condition number for the code rate $\frac{n_T}{2}$ QSTBC is depicted in Fig. 2.16. In

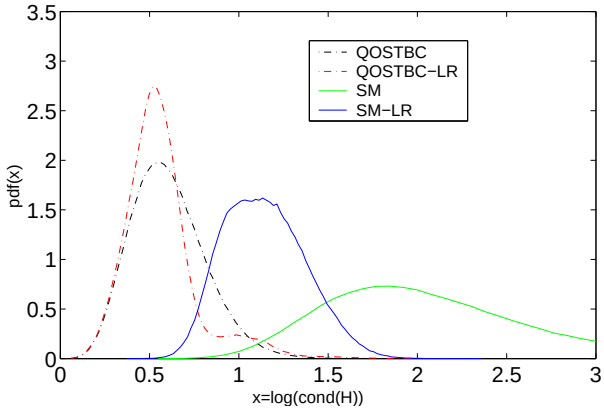

Figure 2.16: Pdfs of channel cond. numbers with SM or code rate $\frac{n_T}{2}$ QSTBC with and w/o LR for a 4×4 system.

order to draw a comparison, the pdf for SM is also plotted. From the Fig., we observe that the impact of LR is not as significant as for SM. However, unlike the rate one QSTBC, a gain is achieved by applying the LR for almost all samples of the equivalent channel model.

2.3.4.3 Numerical simulations

In Fig. 2.17, the BER of SM with BPSK and the code rate one QSTBC with 16-

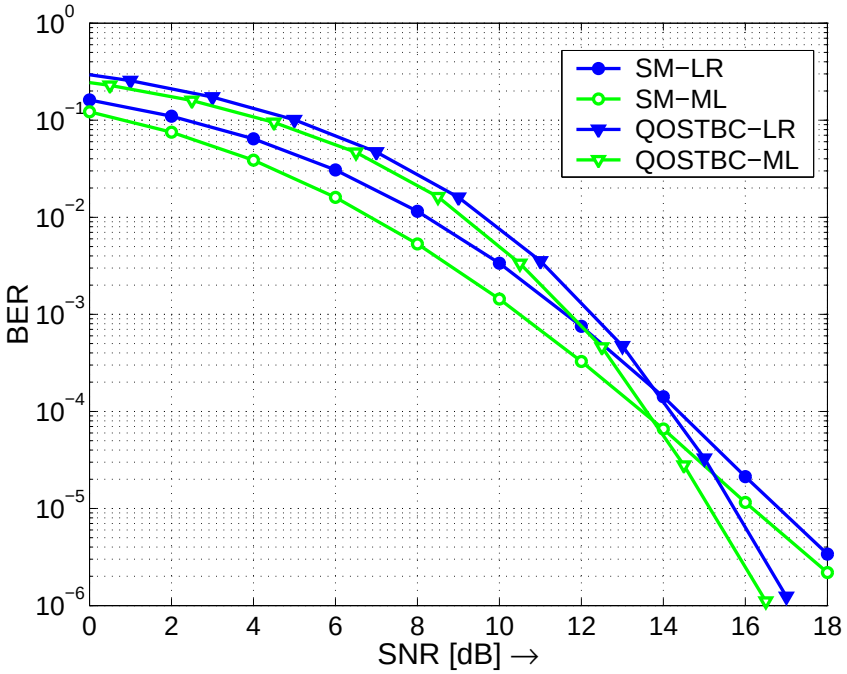

Figure 2.17: BER for SM and QSTBC (code rate one) with ML and LR-ZF, 4 bit/sec/Hz.

QAM is depicted for a system with $n_T = n_R = 4$ and a transmission rate of 4 bits/sec/Hz. The figure shows that the performance of SM with LR-ZF detection

is comparable with the optimal ML detection. In fact, the diversity gain of both detectors is equal and there is only a power penalty of about 1.1dB of LR-ZF to ML. Note that (as aforementioned) this small gap is only due to the BPSK modulation. For higher modulation sizes, this gap is even higher. The gap between ML and LR-ZF detection is even smaller for QSTBC. Here, the power penalty is about 0.6dB. The performance of SM for both ML and LR-ZF detection is better than that of QSTBC for high and moderate BERs. For high SNRs and low BERs of about $1 - 2 \cdot 10^{-4}$, the diversity gain of QSTBC shows its effect and the performance of QSTBC gets better than that of SM. Depending on the application and the operating SNR, one of the schemes is preferable.

The bit error-rate performance of the code rate $\frac{n_T}{2}$ QSTBC with QAM and a transmission rate of 4 bits/sec/Hz is shown in Fig. 2.18. For comparison purposes, we have also plotted the BER of SM for BPSK. Here, we observe that the BER performance with ML-detection of the QSTBC is better than that of SM for all SNR values. With LR-ZF detection, SM performs only better than QSTBC for low SNR of about 2dB. However, the gap in power efficiency between ML and LR-ZF is increased to 1.7dB for the code rate $\frac{n_T}{2}$ QSTBC in comparison to the rate one QSTBC and also to SM with BSPK.

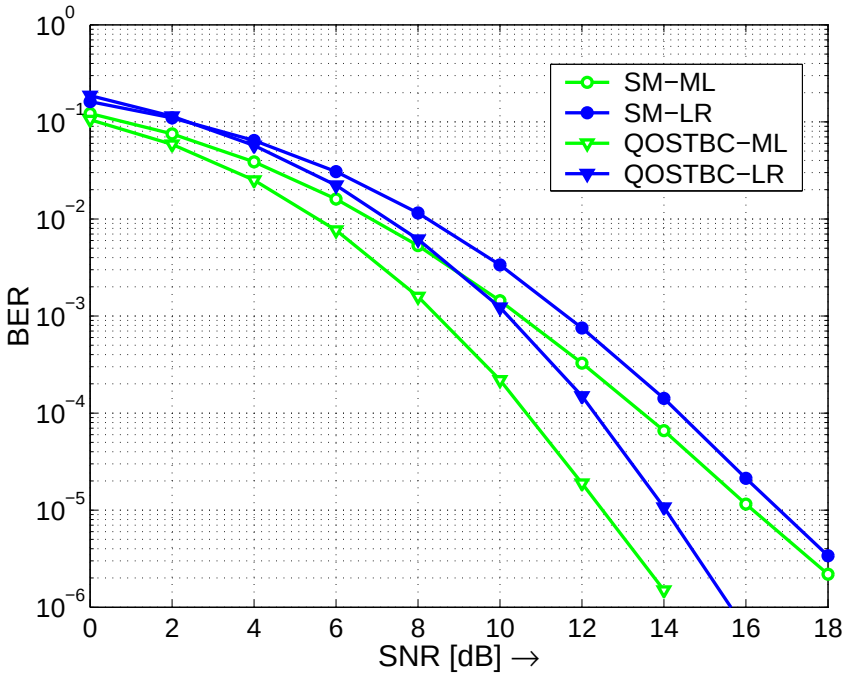

Figure 2.18: BER for SM and QSTBC (code rate $\frac{n_T}{2}$) with ML and LR-ZF, 4 bit/sec/Hz.

By increasing the transmission rate to 8bit/sec/Hz, i.e. QAM for SM and 16QAM for the QSTBC, we observe in Fig. 2.19 that the gap between ML and LR-ZF is dramatically increased with SM to about 6dB. On the other hand, the gap between ML and LR-ZF for the QSTBC and 16QAM is reduced in comparison to the gap achieved with QAM (cf. Fig. 2.18) to about 1.3dB. Although the performance of SM with ML detection is better than that of the QSTBC for low and moderate SNR values, for high SNR values it is the other way around. The performance of the QSTBC with LR-ZF detection is better for the whole SNR range in comparison to SM, which is of higher interest for practical applications, since the computational

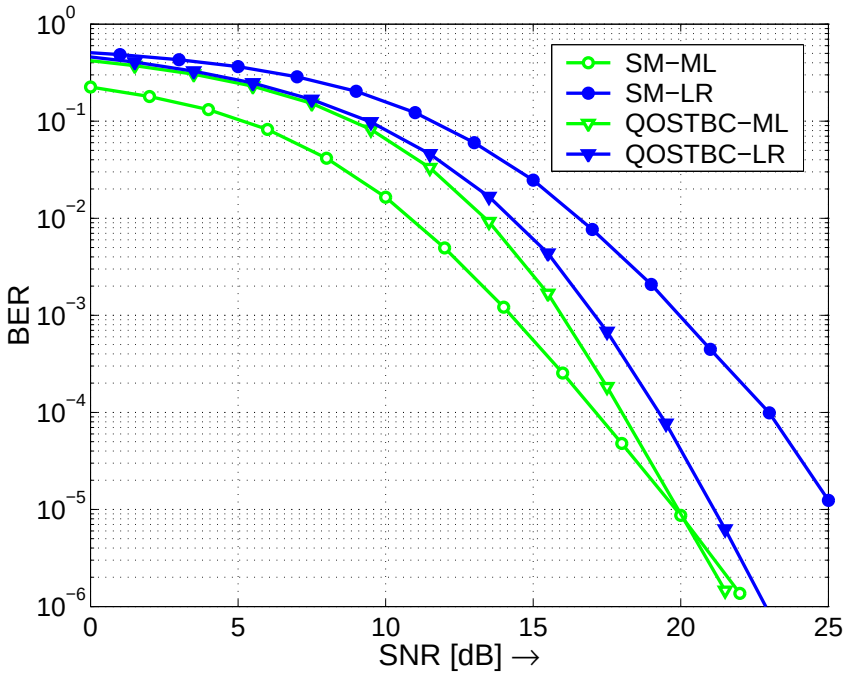

Figure 2.19: BER for SM and QSTBC (code rate $\frac{n_T}{2}$) with ML and LR-ZF, 8bit/sec/Hz.

complexity of the ML detector is exponential in the transmission rate. Another disadvantage of SM is that we need at least as many receive as transmit antennas, i.e. $n_T \leq n_R$, whereas only $\frac{n_T}{2}$ receive antennas are necessary for the code rate $\frac{n_T}{2}$ QSTBC. Multiple receive antennas are only optional for the code rate one QSTBC .

2.4 Capacity Achieving High Rate Quasi-Orthogonal Space-Time Codes

Until now, we focused on achieving full diversity with QSTBC. By relaxing the constraint of full diversity, we arrive at QSTBC capable of achieving high transmission rates. These QSTBC represent n-times stacked version of an OSTBC for two transmit antennas, the Alamouti scheme. The performance of orthogonal space-time codes (OSTBC) [TJC99a] with respect to mutual information has been analyzed (among others) in [NBP04, SP00, BH02] and it was shown that the capacity is achieved only in the case of $n_T = 2$ transmit and $n_R = 1$ receive antennas (the well known Alamouti scheme) due to the rate loss inherent in OSTBC with higher number of transmit antennas.

On the one hand, we have these OSTBC with low complexity and low rates. On the other hand, we have the space-time trellis codes, which achieve higher spectral efficiency in addition to high performance with respect to frame error rates. However, the decoding complexity of space-time trellis codes is increasing exponentially with the number of transmit antennas and the transmission rate. In order to reduce the decoding complexity, a layered space-time architecture was proposed in [TNSC99], where the transmit antennas were partitioned into two-antenna groups and on each group space-time trellis codes were used as component codes. In order to further de-

crease the complexity of this layered space-time architecture, [NSC98, PV01, PV03] used the Alamouti scheme as component code for each group in combination with a suboptimal successive group interference suppression detection strategy. Here, we show that this stacked Alamouti scheme is capable to achieve the capacity in combination with the optimal maximum likelihood detector for the case of $n_T = 2n$ transmit antennas and $n_R = 1$ receive antennas. Furthermore, we show that in the case of more than one receive antenna and if $n_T > n_R$ the stacked Alamouti scheme is capable to achieve a significant portion of the capacity.

2.4.1 Code construction

Starting with the well known (basic) Alamouti scheme [Ala98] (similar to section 2.2.2) for $n_T = 2$ transmit antennas

$$\mathbf{G}_2(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \\ x_2^* & -x_1^* \end{bmatrix} ,$$

the transmit matrix of the rate $n_T/2$ stacked Alamouti scheme with $n_T = 2n$ is constructed in the following way

$$\mathbf{G}_{n_T}\left(\{x_j\}_{j=1}^{n_T}\right) = [\mathbf{G}_2(x_1, x_2), \mathbf{G}_2(x_3, x_4), \ldots, \mathbf{G}_2(x_{n_T-1}, x_{n_T})]$$

Example 2.4.1. *For the case of $n_T = 4$ transmit antennas we have*

$$\mathbf{G}_4(\{x_j\}_{j=1}^4) = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2^* & -x_1^* & x_4^* & -x_3^* \end{bmatrix} .$$

After some manipulations (particularly complex-conjugating) the system model in (2.1) can be rewritten as

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

where $\mathbf{H}' = [(\mathbf{H}'_1)^T, \ldots, (\mathbf{H}'_i)^T, \ldots, (\mathbf{H}'_{n_R})^T]^T$ and $(\mathbf{H}'_i)^T$ is given as

$$\mathbf{H}'_i = \left[\mathbf{H}'_{i,1}, \mathbf{H}'_{i,3}, \ldots, \mathbf{H}'_{i,n_T-1}\right] ,$$

where

$$\mathbf{H}'_{i,j} = \begin{bmatrix} h_{ji} & h_{(j+1)i} \\ -h^*_{(j+1)i} & h^*_{ji} \end{bmatrix} .$$

2.4.2 Mutual Information

In this section we first analyze the case of $n_R = 1$ receive antennas and then generalize the analysis to the case of arbitrary number of receive antennas.

2.4.2.1 Case $n_R = 1$

The capacity of a MIMO system with n_T transmit and n_R receive antennas is given as [Tel99]

$$I = \log_2 \det\left(\mathbf{I}_{n_T} + \frac{\rho}{n_T}\mathbf{H}^H\mathbf{H}\right) .$$

In case of the stacked Alamouti scheme, the achievable portion of the mutual information is

$$I_{sA} = \frac{1}{2}\log_2 \det\left(\mathbf{I}_{n_T} + \frac{\rho}{n_T}(\mathbf{H}_1')^H\mathbf{H}_1'\right) .$$

Using the determinant equality $\det(\mathbf{I}+\mathbf{AB}) = \det(\mathbf{I}+\mathbf{BA})$ and after some manipulations we arrive at

$$I_{sA} = \log_2\left(1 + \frac{\rho}{n_T}\sum_{j=1}^{n_T}|h_{j1}|^2\right) . \tag{2.74}$$

which equals the capacity of a MIMO system with n_T transmit and $n_R = 1$ receive antennas.

2.4.2.2 Case of arbitrary n_R

The available portion of the mutual information achievable with $n_R \geq 1$ for the stacked Alamouti scheme is

$$I_{sA} = \frac{1}{2}\log_2 \det\left(\mathbf{I}_{n_T} + \frac{\rho}{n_T}(\mathbf{H}')^H\mathbf{H}'\right) . \tag{2.75}$$

By applying the trace-determinant inequality $\det(\mathbf{A})^{1/n} \leq \frac{1}{n}\mathrm{tr}(\mathbf{A})$, we arrive at an upper bound on (2.75) given as

$$I_{sA} \leq n_R \log_2\left(1 + \frac{\rho}{n_T n_R}\underbrace{\sum_{j=1}^{n_T}\sum_{i=1}^{n_R}|h_{ji}|^2}_{\lambda}\right) , \tag{2.76}$$

which is equal to n_R times the capacity of a MIMO system with $n_T \times n_R$ transmit antennas and one receive antenna.

2.4.3 Outage Probability P_{out}

The outage probability P_{out} achievable with the stacked Alamouti scheme is defined as the probability that I_{sA} is smaller than a certain rate R, i.e.

$$P_{out}(R, n_T, n_R, \rho) = \Pr[I_{sA} < R] .$$

Using (2.76), the lower bound on the outage probability P_{out} can be written as

$$P_{out}(R, n_T, n_R, \rho) \geq \Pr\left[\lambda < \frac{n_T n_R}{\rho}\left(2^{\frac{R}{n_R}} - 1\right)\right] .$$

Since λ is chi-square distributed random variable with $2n_T n_R$ degrees of freedom, P_{out} is given as [GR83, p.310,3.351(1)]

$$P_{out}(R, n_T, n_R, \rho) \geq 1 - \exp\left(-\frac{n_T n_R}{\rho}\left(2^{\frac{R}{n_R}} - 1\right)\right)\sum_{k=0}^{n_T n_R - 1}\frac{\left(\frac{n_T n_R}{\rho}\left(2^{\frac{R}{n_R}} - 1\right)\right)^k}{k!} . \tag{2.77}$$

2.4.4 Simulations

In Fig. 2.20, the outage mutual information (OMI) of the stacked Alamouti scheme and the outage capacity of a MIMO system (MIMO-OC) with $n_R = 2$ and $n_T = 2, 4$ and $n_T = 8$ is depicted. In case of $n_T = 2$, we have the standard Alamouti scheme. From the Fig., we observe that the difference between the OMI of the stacked Alamouti scheme and the MIMO-OC diminishes significantly by increasing the number of transmit antennas.

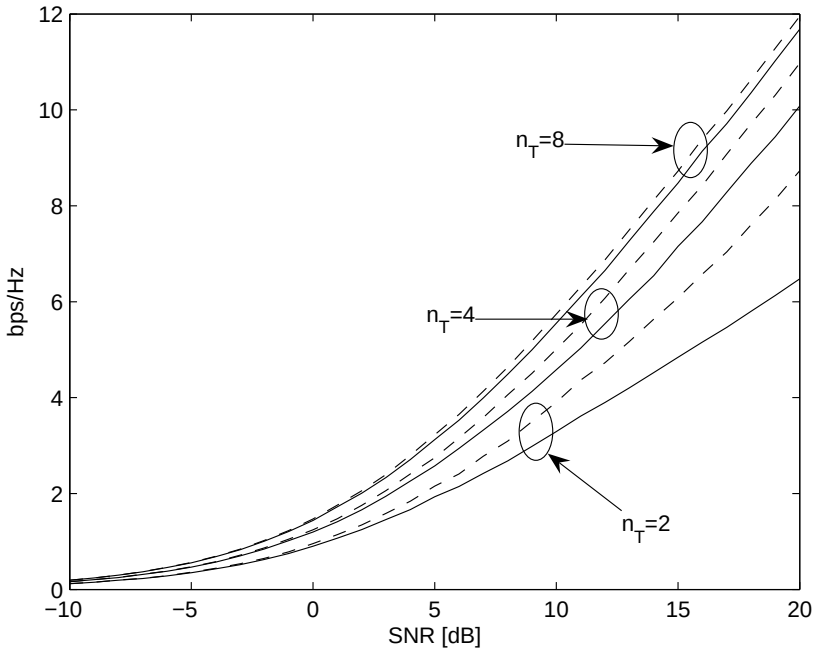

Figure 2.20: 10% Outage capacity of a MIMO system and mutual information (OMI) achievable with the stacked Alamouti scheme with $n_R = 2$ receive and $n_T = 2$,$n_T = 4$ and $n_T = 8$ transmit antennas.

In Fig. 2.21, the outage mutual information (OMI) of the stacked Alamouti scheme and the MIMO-OC with $n_T = 4$ and $n_R = 2, 4$ and $n_R = 8$ is depicted. In contrast to the case of increasing number of transmit antennas, here we observe that the difference between the OMI of the stacked Alamouti scheme and the MIMO-OC increases by increasing the number of receive antennas.

In Fig. 2.22, P_{out} of the stacked Alamouti scheme with $n_T = 4$ transmit and $n_R = 1$ to $n_R = 3$ and $n_R = 6$ receive antennas and the lower bound in (2.77) are depicted. From the Fig. we observe that our lower bound on the performance of the stacked Alamouti scheme with respect to P_{out} is tight for n_R equal to one. This can also be observed by comparing eq. (2.74) with (2.76) and setting $n_R = 1$ in (2.76). For higher number of receive antennas, although not tight, the lower abound appears as an useful and simple approximation of the performance in terms of outage probability for the stacked Alamouti scheme.

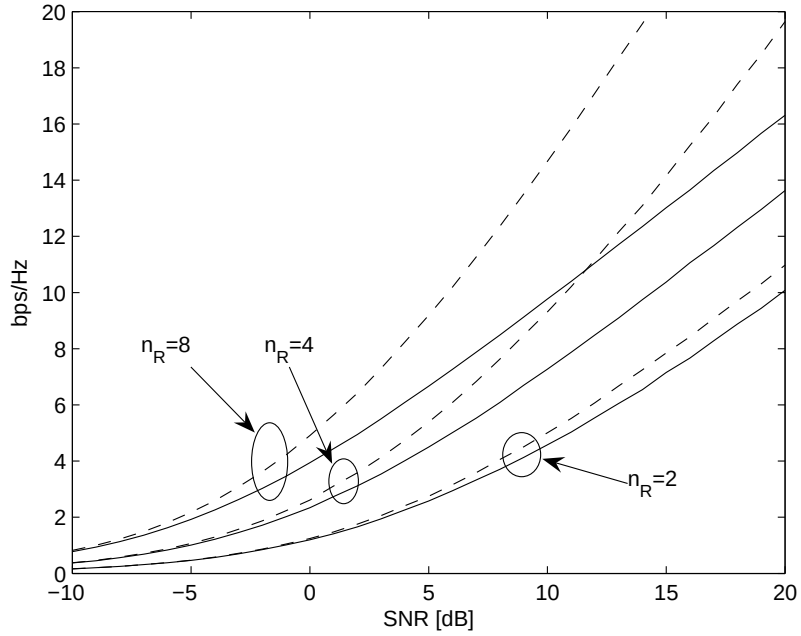

Figure 2.21: 10% MIMO-OC and OMI achievable with the stacked Alamouti scheme with $n_T = 4$ transmit and $n_R = 2$,$n_R = 4$ and $n_R = 8$ receive antennas.

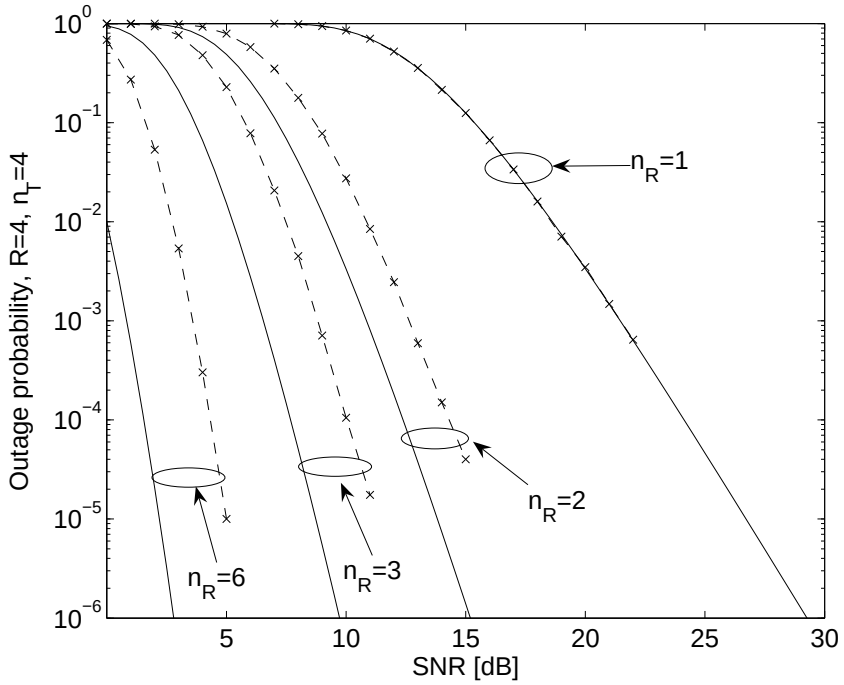

Figure 2.22: Outage probabilities of QSTBC (dashed lines) and lower bound(solid line) for $n_T = 4$ transmit and different numbers of receive antennas n_R, Rate=4.

2.5 Appendix: Proofs

2.5.1 Proof of Lemma 2.2.1

Proof. Proof is given by the principle of induction. We start with the initial case $\mathbf{G}_4$ as given in (2.4) for $n_T = 4$ transmit antennas. It is rather straightforward to realize that

$$\mathbf{G}_4(\tilde{\mathbf{x}}_o)^H \mathbf{G}_4(\tilde{\mathbf{x}}_e) + \mathbf{G}_4(\tilde{\mathbf{x}}_e)^H \mathbf{G}_4(\tilde{\mathbf{x}}_o) = \mathbf{0} \,,$$

where $\tilde{\mathbf{x}}_o = [x_1, 0, x_3, 0]^T$ and $\tilde{\mathbf{x}}_e = [0, x_2, 0, x_4]^T$. Now assume that the following hypothesis holds for $n_T = k/2$, i.e.

$$\mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}_o)^H \mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}_e) + \mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}_e)^H \mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}_o) = \mathbf{0} \,, \tag{2.78}$$

with $\tilde{\mathbf{x}}_o = [x_1, 0, x_3, 0, \ldots, x_{\frac{k}{2}-1}, 0]^T$ and $\tilde{\mathbf{x}}_e = [0, x_2, 0, x_4, 0, \ldots, x_{\frac{k}{2}}]^T$, then the following inductive step is true

$$\begin{aligned}
&\mathbf{G}_k(\tilde{\mathbf{x}}_o)^H \mathbf{G}_k(\tilde{\mathbf{x}}_e) + \mathbf{G}_k(\tilde{\mathbf{x}}_e)^H \mathbf{G}_k(\tilde{\mathbf{x}}_o) \overset{(2.3)}{=} \\
&\begin{bmatrix} \mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}'_o)^H & \boldsymbol{\Theta}\mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}''_o)^H \\ \mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}''_o)^H & -\boldsymbol{\Theta}\mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}'_o)^H \end{bmatrix} \begin{bmatrix} \mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}'_e) & \mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}''_e) \\ \mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}''_e)\boldsymbol{\Theta} & -\mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}'_e)\boldsymbol{\Theta} \end{bmatrix} . \\
&+ \begin{bmatrix} \mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}'_e)^H & \boldsymbol{\Theta}\mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}''_e)^H \\ \mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}''_e)^H & -\boldsymbol{\Theta}\mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}'_e)^H \end{bmatrix} \begin{bmatrix} \mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}'_o) & \mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}''_o) \\ \mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}''_o)\boldsymbol{\Theta} & -\mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}'_o)\boldsymbol{\Theta} \end{bmatrix} \overset{(2.78)}{=} \mathbf{0} \,,
\end{aligned} \tag{2.79}$$

by using the fact that $-\boldsymbol{\Theta}\mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}''_o)^H \mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}'_e)\boldsymbol{\Theta} = \mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}''_o)^H \mathbf{G}_{\frac{k}{2}}(\tilde{\mathbf{x}}'_e)$, where $\tilde{\mathbf{x}}_o = [x_1, 0, x_3, 0, \ldots, x_{k-1}, 0]^T$, $\tilde{\mathbf{x}}_e = [0, x_2, 0, x_4, 0, \ldots, x_k]^T$, $\tilde{\mathbf{x}}'_o = [x_1, 0, x_3, 0, \ldots, x_{\frac{k}{2}-1}, 0]^T$, $\tilde{\mathbf{x}}'_e = [0, x_2, 0, x_4, 0, \ldots, x_{\frac{k}{2}}]^T$, $\tilde{\mathbf{x}}''_o = [x_{\frac{k}{2}+1}, 0, \ldots, \ldots, x_{k-1}, 0]^T$, and $\tilde{\mathbf{x}}''_e = [0, x_{\frac{k}{2}+2}, 0, \ldots, x_k]^T$. Since the initial case of $n_T = 4$ is true and the inductive step is true, the statement in (2.12) is true for all $n_T = 2^n$. That concludes the proof. □

2.5.2 Proof of Lemma 2.2.2

In the following, the arguments of $\mathbf{M}_N, \mathbf{N}_N, \mathbf{S}_N$ and $\mathbf{T}_N$ are omitted occasionally in order to increase the readability of the proof.

Proof. The proof is done by the principle of induction. We start with the initial case $\mathbf{M}_4$. From (2.19), it follows

$$\begin{aligned}
\mathbf{M}_4 &= \begin{bmatrix} \mathbf{M}_2 & \mathbf{N}_2 \\ -\mathbf{N}_2 & \mathbf{M}_2 \end{bmatrix} \overset{(2.21)}{=} \begin{bmatrix} \mathbf{V}_2\mathbf{S}_2\mathbf{V}_2^H & \mathbf{V}_2\mathbf{T}_2\mathbf{V}_2^H \\ -\mathbf{V}_2\mathbf{T}_2\mathbf{V}_2^H & \mathbf{V}_2\mathbf{S}_2\mathbf{V}_2^H \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{V}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 \end{bmatrix} \begin{bmatrix} \mathbf{S}_2 & \mathbf{T}_2 \\ -\mathbf{T}_2 & \mathbf{S}_2 \end{bmatrix} \begin{bmatrix} \mathbf{V}_2^H & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2^H \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
&= \begin{bmatrix} \mathbf{V}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 \end{bmatrix} \boldsymbol{\Pi}_4 \begin{bmatrix} \mu_2^1 & \nu_2^1 & & \\ -\nu_2^1 & \mu_2^1 & & \\ & & \mu_2^2 & \nu_2^2 \\ & & -\nu_2^2 & \mu_2^2 \end{bmatrix} \boldsymbol{\Pi}_4^H \begin{bmatrix} \mathbf{V}_2^H & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2^H \end{bmatrix} \\
&= \underbrace{\begin{bmatrix} \mathbf{V}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 \end{bmatrix} \boldsymbol{\Pi}_4 \begin{bmatrix} \mathbf{V}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 \end{bmatrix}}_{\mathbf{V}_4} \mathbf{S}_4 \underbrace{\begin{bmatrix} \mathbf{V}_2^H & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2^H \end{bmatrix} \boldsymbol{\Pi}_4^H \begin{bmatrix} \mathbf{V}_2^H & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2^H \end{bmatrix}}_{\mathbf{V}_4^H} ,
\end{aligned}$$

with $\mathbf{\Pi}_4$ given as

$$\mathbf{\Pi}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$\mathbf{S}_4 = \mathrm{diag}(\mu_4^1, \mu_4^2, \mu_4^3, \mu_4^4)$, where

$$\mu_4^1 = \mu_2^1 - i\nu_2^1 \ , \ \mu_4^2 = \mu_2^1 + i\nu_2^1 \ , \ \mu_4^3 = \mu_2^2 - i\nu_2^2 \ , \text{ and } \ \mu_4^4 = \mu_2^2 + i\nu_2^2. \tag{2.80}$$

which is equivalent to

$$\mathbf{S}_4(\{\alpha_l\}_{l=1}^4) = \mathbf{\Pi}_4 \begin{bmatrix} \mathbf{S}_2(\alpha_1,\alpha_2) - i\mathbf{T}_2(\alpha_3,\alpha_4) & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_2(\alpha_1,\alpha_2) + i\mathbf{T}_2(\alpha_3,\alpha_4) \end{bmatrix} \mathbf{\Pi}_4^H$$

The same procedure applied to $\mathbf{N}_4$ results in a $\mathbf{T}_4$ given as $\mathbf{T}_4 = \mathrm{diag}(\nu_4^1, \nu_4^2, \nu_4^3, \nu_4^4)$ with

$$\nu_4^1 = \nu_2^3 - i\mu_2^3 \ , \ \nu_4^2 = \nu_2^3 + i\mu_2^3 \ , \ \nu_4^3 = \nu_2^4 - i\mu_2^4 \ , \text{ and } \ \nu_4^4 = \nu_2^4 + i\mu_2^4.$$

which is equivalent to

$$\mathbf{T}_4(\{\alpha_l\}_{l=5}^8) = \mathbf{\Pi}_4 \begin{bmatrix} \mathbf{T}_2(\alpha_7,\alpha_8) - i\mathbf{S}_2(\alpha_5,\alpha_6) & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_2(\alpha_7,\alpha_8) + i\mathbf{S}_2(\alpha_5,\alpha_6) \end{bmatrix} \mathbf{\Pi}_4^H \ .$$

Now assume that the following hypothesis holds for $N = {}^K\!/_2$, i.e.

$$\mathbf{M}_{\frac{K}{2}} = \mathbf{V}_{\frac{K}{2}} \mathbf{S}_{\frac{K}{2}} \mathbf{V}_{\frac{K}{2}}^H \ \text{ and } \ \mathbf{N}_{\frac{K}{2}} = \mathbf{V}_{\frac{K}{2}} \mathbf{T}_{\frac{K}{2}} \mathbf{V}_{\frac{K}{2}}^H \ , \tag{2.81}$$

then the following inductive step is true

$$\begin{aligned} \mathbf{M}_K &= \begin{bmatrix} \mathbf{M}_{\frac{K}{2}} & \mathbf{N}_{\frac{K}{2}} \\ -\mathbf{N}_{\frac{K}{2}} & \mathbf{M}_{\frac{K}{2}} \end{bmatrix} \overset{(2.81)}{=} \begin{bmatrix} \mathbf{V}_{\frac{K}{2}} \mathbf{S}_{\frac{K}{2}} \mathbf{V}_{\frac{K}{2}}^H & \mathbf{V}_{\frac{K}{2}} \mathbf{T}_{\frac{K}{2}} \mathbf{V}_{\frac{K}{2}}^H \\ -\mathbf{V}_{\frac{K}{2}} \mathbf{T}_{\frac{K}{2}} \mathbf{V}_{\frac{K}{2}}^H & \mathbf{V}_{\frac{K}{2}} \mathbf{S}_{\frac{K}{2}} \mathbf{V}_{\frac{K}{2}}^H \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{V}_{\frac{K}{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{\frac{K}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{S}_{\frac{K}{2}} & \mathbf{T}_{\frac{K}{2}} \\ -\mathbf{T}_{\frac{K}{2}} & \mathbf{S}_{\frac{K}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{V}_{\frac{K}{2}}^H & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{\frac{K}{2}}^H \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{V}_{\frac{K}{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{\frac{K}{2}} \end{bmatrix} \mathbf{\Pi}_K \begin{bmatrix} \mathbf{Q}_{\frac{K}{2}}^1 & & \mathbf{0} \\ & \mathbf{Q}_{\frac{K}{2}}^2 & \\ & & \ddots \\ \mathbf{0} & & & \mathbf{Q}_{\frac{K}{2}}^{\frac{K}{2}} \end{bmatrix} \mathbf{\Pi}_K^H \begin{bmatrix} \mathbf{V}_{\frac{K}{2}}^H & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{\frac{K}{2}}^H \end{bmatrix}, \end{aligned}$$

$$= \underbrace{\left(\mathbf{I}_2 \otimes \mathbf{V}_{\frac{K}{2}}\right) \mathbf{\Pi}_K \left(\mathbf{I}_{\frac{K}{2}} \otimes \mathbf{V}_2\right)}_{\mathbf{V}_K} \mathbf{S}_K \underbrace{\left(\mathbf{I}_{\frac{K}{2}} \otimes \mathbf{V}_2^H\right) \mathbf{\Pi}_K \left(\mathbf{I}_2 \otimes \mathbf{V}_{\frac{K}{2}}^H\right)}_{\mathbf{V}_K^H}$$

where

$$\mathbf{Q}_{\frac{K}{2}}^k = \begin{bmatrix} \mu_{\frac{K}{2}}^k & \nu_{\frac{K}{2}}^k \\ -\nu_{\frac{K}{2}}^k & \mu_{\frac{K}{2}}^k \end{bmatrix}$$

and $\mathbf{S}_K = \mathrm{diag}(\mu_K^1, \mu_K^2, \ldots, \mu_K^K)$ with

$$\mu_K^{l-1} = \mu_{\frac{K}{2}}^{\frac{l}{2}} - i\nu_{\frac{K}{2}}^{\frac{l}{2}} \ \text{ and } \ \mu_K^l = \mu_{\frac{K}{2}}^{\frac{l}{2}} + i\nu_{\frac{K}{2}}^{\frac{l}{2}} \qquad \forall l = 2, 4, 6, \ldots, K \tag{2.82}$$

which is equivalent to

$$\mathbf{S}_K = \mathbf{\Pi}_K \begin{bmatrix} \mathbf{S}_{\frac{K}{2}}(\{\alpha_l\}_{l=1}^{\frac{K}{2}}) - i\mathbf{T}_{\frac{K}{2}}(\{\alpha_l\}_{l=\frac{K}{2}+1}^{K}) & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{\frac{K}{2}}(\{\alpha_l\}_{l=1}^{\frac{K}{2}}) + i\mathbf{T}_{\frac{K}{2}}(\{\alpha_l\}_{l=\frac{K}{2}+1}^{K}) \end{bmatrix} \mathbf{\Pi}_K^H.$$

The same procedure applied to $\mathbf{N}_K$ results in the same $\mathbf{V}_K$ and

$$\mathbf{T}_K = \mathrm{diag}(\nu_K^1, \nu_K^2, \ldots, \nu_K^K)$$

with

$$\nu_K^{l-1} = \nu_{\frac{K}{2}}^{\frac{K+l}{2}} - i\mu_{\frac{K}{2}}^{\frac{K+l}{2}} \text{ and } \nu_K^l = \nu_{\frac{K}{2}}^{\frac{K+l}{2}} + i\mu_{\frac{K}{2}}^{\frac{K+l}{2}} \qquad \forall l = 2, 4, 6, \ldots, K \tag{2.83}$$

which is again equivalent to

$$\mathbf{T}_K = \mathbf{\Pi}_K \cdot \begin{bmatrix} \mathbf{T}_{\frac{K}{2}}(\{\alpha_l\}_{l=K+\frac{K}{2}+1}^{2K}) - i\mathbf{S}_{\frac{K}{2}}(\{\alpha_l\}_{l=K+1}^{K+\frac{K}{2}}) & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_{\frac{K}{2}}(\{\alpha_l\}_{l=K+\frac{K}{2}+1}^{2K}) + i\mathbf{S}_{\frac{K}{2}}(\{\alpha_l\}_{l=K+1}^{K+\frac{K}{2}}) \end{bmatrix} \mathbf{\Pi}_K^H.$$

Since the initial case of $N = 4$ is true and the inductive step is true, the statement in (2.22) is true for all $N = 2^n$. That concludes the proof. □

2.5.3 Proof of Lemma 2.2.3

Proof. The proof is done by the principle of induction. The outline of the proof is as follows. For the initial case of $n_T = 4$, we need the eigenvalues $(\tilde{\mu}_2^1)^2$ and $(\tilde{\nu}_2^1)^2$ for $n_T = 2$, i.e. the Alamouti scheme, as indicated in (2.80). The first step is therefore to construct the eigenvalues for $n_T = 4$ with the eigenvalues of $n_T = 2$. Using (2.82) and (2.83), we observe that the eigenvalues for $n_T = 4$ can be also obtained with the eigenvalues for $n_T = 8$, which is the second step revealing an important instruction of constructing eigenvalues μ_K,ν_K from $\mu_{\frac{K}{2}}$,$\nu_{\frac{K}{2}}$. It follows the hypothesis and the inductive step concluding the proof.

Now, we start with the well known Alamouti scheme. By applying the Alamouti scheme ($n_T = 2$), $\frac{2}{n_T}\mathbf{D}_{\frac{n_T}{2}}\mathbf{D}_{\frac{n_T}{2}}$ as well as $\mathbf{x}_\mathrm{o}$ and $\mathbf{x}_\mathrm{e}$ are only scalars. Thus, the only eigenvalue of $\mathbf{D}_1\mathbf{D}_1$ of the decomposed system model for the part with $\mathbf{x}_\mathrm{o}$ (and similar for $\mathbf{x}_\mathrm{e}$) is given as

$$(\tilde{\mu}_2^1)^2 = \sum_{i=1}^{n_R} \alpha_1(h_{1i}, h_{2i}) = \sum_{i=1}^{n_R}\sum_{j=1}^{n_T=2} |h_{ji}|^2 = \sum_{i=1}^{n_R} \mathbf{h}_{1\to2,i}^H \mathbf{A}_2^1 \mathbf{h}_{1\to2,i} \tag{2.84}$$

where $\mathbf{h}_{k\to l,i} = [h_{ki}, \ldots, h_{li}]^T$ and

$$\mathbf{A}_2^1 = \mathbf{A}_2 = \mathbf{I}_2$$

Similarly,

$$(\tilde{\nu}_2^1)^2 = \sum_{i=1}^{n_R} \alpha_2(h_{3i}, h_{4i}) = \sum_{i=1}^{n_R} \mathbf{h}_{3\to4,i}^H \mathbf{A}_2^1 \mathbf{h}_{3\to4,i} \tag{2.85}$$

We are now able to start with the initial case of $n_T = 4$. The first eigenvalue of the QSTBC for the part with $\mathbf{x}_\mathrm{o}$ (and similar for $\mathbf{x}_\mathrm{e}$) is given as in (2.82) (with

$(\tilde{\mu}_{n_T}^j)^2 = \frac{2}{n_T}\mu_{n_T}^j$ and $(\tilde{\nu}_{n_T}^j)^2 = \frac{2}{n_T}\nu_{n_T}^j$)

$$(\tilde{\mu}_4^1)^2 = (\tilde{\mu}_4^1(h_{1i},\ldots,h_{4i}))^2 = \tilde{\mu}_2^1 - i\tilde{\nu}_2^1 \overset{(2.29)}{=} \sum_{i=1}^{n_R}\frac{1}{2}(\alpha_1(h_{1i},\ldots,h_{4i}) + \alpha_2(h_{1i},\ldots,h_{4i}))$$
$$\overset{(2.16)}{=} \sum_{i=1}^{n_R}\frac{1}{2}(\sum_{j=1}^{n_T}|h_{ji}|^2 + 2\mathrm{Im}(h_{1i}^* h_{3i} + h_{4i}^* h_{2i}))$$

$$= \sum_{i=1}^{n_R}\frac{1}{2}\left([\mathbf{h}_{1\to2,i}^H \mathbf{h}_{3\to4,i}^H]\begin{bmatrix}\mathbf{h}_{1\to2,i}\\ \mathbf{h}_{3\to4,i}\end{bmatrix} - i\begin{bmatrix}-h_{3i}^* & h_{4i}^* & h_{1i}^* & -h_{2i}^*\end{bmatrix}\begin{bmatrix}\mathbf{h}_{1\to2,i}\\ \mathbf{h}_{3\to4,i}\end{bmatrix}\right)$$

$$\begin{aligned}&= \sum_{i=1}^{n_R}\mathbf{h}_{1\to4,i}^H\frac{1}{2}\begin{bmatrix}\mathbf{A}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{A}_2\end{bmatrix}\mathbf{h}_{1\to4,i}\\ &\quad - i\begin{bmatrix}h_{1i}^* & -h_{2i}^* & -h_{3i}^* & h_{4i}^*\end{bmatrix}\frac{1}{2}\begin{bmatrix}\mathbf{0} & \mathbf{A}_2\\ \mathbf{A}_2 & \mathbf{0}\end{bmatrix}\mathbf{h}_{1\to4,i}\end{aligned} \tag{2.86}$$

$$= \sum_{i=1}^{n_R}\mathbf{h}_{1\to4,i}^H\frac{1}{2}\begin{bmatrix}\mathbf{A}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{A}_2\end{bmatrix}\mathbf{h}_{1\to4,i} - i\mathbf{h}_{1\to4,i}^H\frac{1}{2}\begin{bmatrix}\mathbf{\Theta}_2 & \mathbf{0}\\ \mathbf{0} & -\mathbf{\Theta}_2\end{bmatrix}\begin{bmatrix}\mathbf{0} & \mathbf{A}_2\\ \mathbf{A}_2 & \mathbf{0}\end{bmatrix}\mathbf{h}_{1\to4,i}$$
$$= \sum_{i=1}^{n_R}\mathbf{h}_{1\to4,i}^H\frac{1}{2}\left(\begin{bmatrix}\mathbf{A}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{A}_2\end{bmatrix} - \begin{bmatrix}\mathbf{0} & i\mathbf{\Theta}_2\mathbf{A}_2\\ -i\mathbf{\Theta}_2\mathbf{A}_2 & \mathbf{0}\end{bmatrix}\right)\mathbf{h}_{1\to4,i}$$
$$= \sum_{i=1}^{n_R}\mathbf{h}_{1\to4,i}^H\mathbf{A}_4^1\mathbf{h}_{1\to4,i}\,, \tag{2.87}$$

where

$$\mathbf{A}_4^1 = \frac{1}{2}\begin{bmatrix}\mathbf{A}_2 & -i\mathbf{\Theta}_2\mathbf{A}_2\\ i\mathbf{\Theta_2}\mathbf{A}_2 & \mathbf{A}_2\end{bmatrix}.$$

In an analogous manner, we get $(\tilde{\nu}_4^1)^2$ given as

$$(\tilde{\nu}_4^1)^2 = \mathbf{h}_{5\to8,i}^H\mathbf{A}_4^1\mathbf{h}_{5\to8,i}$$

On the other hand, with (2.82) we have

$$(\tilde{\mu}_4^1)^2 = (\tilde{\mu}_4^1(h_{1i},\ldots,h_{8i}))^2 = \sum_{i=1}^{n_R}\frac{1}{2}((\tilde{\mu}_8^1)^2 + (\tilde{\mu}_8^2)^2)$$
$$\overset{(2.32)}{=} \sum_{i=1}^{n_R}\frac{1}{2}\frac{\alpha_1(h_{1i},\ldots,h_{8i}) + \alpha_2(h_{1i},\ldots,h_{8i})}{2}$$
$$\overset{(2.18)}{=} \sum_{i=1}^{n_R}\frac{1}{4}\left(\sum_{j=1}^{8}|h_{ji}|^2 + 2\mathrm{Im}(h_{1i}^* h_{3i} + h_{4i}^* h_{2i} + h_{5i}^* h_{7i} + h_{8i}^* h_{6i})\right)$$

$$= \sum_{i=1}^{n_R}\frac{1}{4}\Big(\mathbf{h}_{1\to8,i}^H\mathbf{h}_{1\to8,i}$$

$$
\begin{aligned}
&- i \begin{bmatrix} -h_{3i}^* & h_{4i}^* & h_{1i}^* & -h_{2i}^* & -h_{7i}^* & h_{8i}^* & h_{5i}^* & -h_{6i}^* \end{bmatrix} \mathbf{h}_{1\to 8,i} \Bigg) \\
&= \sum_{i=1}^{n_R} \frac{1}{4} \mathbf{h}_{1\to 8,i}^H \mathbf{h}_{1\to 8,i} \\
&\quad - i \mathbf{h}_{1\to 8,i}^H \frac{1}{4} \begin{bmatrix} \boldsymbol{\Theta}_2 & & 0 & \\ & -\boldsymbol{\Theta}_2 & & \\ 0 & & \boldsymbol{\Theta}_2 & \\ & & & -\boldsymbol{\Theta}_2 \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{A}_2 & 0 & \\ \mathbf{A}_2 & \mathbf{0} & & \\ 0 & & \mathbf{0} & \mathbf{A}_2 \\ & & \mathbf{A}_2 & \mathbf{0} \end{bmatrix} \mathbf{h}_{1\to 8,i} \\
&= \sum_{i=1}^{n_R} \frac{1}{4} \mathbf{h}_{1\to 8,i}^H \\
&\quad \cdot \left(\begin{bmatrix} \mathbf{A}_2 & & 0 & \\ & \mathbf{A}_2 & & \\ 0 & & \mathbf{A}_2 & \\ & & & \mathbf{A}_2 \end{bmatrix} - i \begin{bmatrix} \mathbf{0} & \boldsymbol{\Theta}_2\mathbf{A}_2 & & 0 \\ -\boldsymbol{\Theta}_2\mathbf{A}_2 & \mathbf{0} & & \\ & & \mathbf{0} & \boldsymbol{\Theta}_2\mathbf{A}_2 \\ 0 & & -\boldsymbol{\Theta}_2\mathbf{A}_2 & \mathbf{0} \end{bmatrix} \right) \mathbf{h}_{1\to 8,i} \\
&\quad = \sum_{i=1}^{n_R} \frac{1}{4} \mathbf{h}_{1\to 8,i}^H \left(\begin{bmatrix} \mathbf{A}_2 & -i\boldsymbol{\Theta}_2\mathbf{A}_2 & & 0 \\ i\boldsymbol{\Theta}_2\mathbf{A}_2 & \mathbf{A}_2 & & \\ & & \mathbf{A}_2 & -i\boldsymbol{\Theta}_2\mathbf{A}_2 \\ 0 & & i\boldsymbol{\Theta}_2\mathbf{A}_2 & \mathbf{A}_2 \end{bmatrix} \right) \mathbf{h}_{1\to 8,i} \\
&\quad = \sum_{i=1}^{n_R} \mathbf{h}_{1\to 8,i}^H \left(\frac{1}{2} \begin{bmatrix} \mathbf{A}_4^1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_4^1 \end{bmatrix} \right) \mathbf{h}_{1\to 8,i} = \sum_{i=1}^{n_R} \mathbf{h}_{1\to 8,i}^H \frac{1}{2} \left(\mathbf{I}_2 \otimes \mathbf{A}_4^1 \right) \mathbf{h}_{1\to 8,i}
\end{aligned}
\tag{2.88}
$$

$$
\begin{aligned}
&= \sum_{i=1}^{n_R} \frac{1}{2} [\mathbf{h}_{1\to 4,i}^H \mathbf{h}_{5\to 8,i}^H] \begin{bmatrix} \mathbf{A}_4^1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_4^1 \end{bmatrix} \begin{bmatrix} \mathbf{h}_{1\to 4,i} \\ \mathbf{h}_{5\to 8,i} \end{bmatrix} \\
&= \sum_{i=1}^{n_R} \frac{1}{2} \mathbf{h}_{1\to 4,i}^H \mathbf{A}_4^1 \mathbf{h}_{1\to 4,i} + \frac{1}{2} \mathbf{h}_{5\to 8,i}^H \mathbf{A}_4^1 \mathbf{h}_{5\to 8,i}
\end{aligned}
\tag{2.89}
$$

In an analogous manner, we get $(\tilde{\nu}_4^1)^2$ given as

$$
(\tilde{\nu}_4^1)^2 = \sum_{i=1}^{n_R} \frac{1}{2} [\mathbf{h}_{1\to 4,i}^H \mathbf{h}_{5\to 8,i}^H] \begin{bmatrix} \mathbf{0} & i\boldsymbol{\Theta}_k\mathbf{A}_4^1 \\ -i\boldsymbol{\Theta}_k\mathbf{A}_4^1 & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{h}_{1\to 4,i} \\ \mathbf{h}_{5\to 8,i} \end{bmatrix}
$$

Since the eigenvalues $(\tilde{\mu}_4^2)^2$ and $(\tilde{\nu}_4^2)^2$ can be obtained very easily in a similar way, we omit the derivations here.

Comparing (2.89) with (2.87) shows that in order to get the eigenvalues of $n_T = 8$, the eigenvalues of $n_T = 4$ have to be expanded by using the Kronecker product of $\mathbf{I}_2$ and $\mathbf{A}_4$, divided by $\frac{1}{2}$ in order to incorporate the channel entries $h_{5i}, \ldots, h_{8i}$ as given in (2.88). Actually, this can also be observed in the expansion from $n_T = 2$ to $n_T = 4$ by comparing (2.84) with the first addend in (2.86).

Now assume that the following hypothesis holds

$$(\tilde{\mu}_k^j(h_{1i},\ldots,h_{ki}))^2 = \sum_{i=1}^{n_R} \mathbf{h}_{1\to k,i}^H \mathbf{A}_k^j \mathbf{h}_{1\to k,i} \tag{2.90}$$

$$(\tilde{\mu}_k^j(h_{1i},\ldots,h_{2ki}))^2 = \sum_{i=1}^{n_R} \mathbf{h}_{1\to 2k,i}^H \frac{1}{2}\begin{bmatrix} \mathbf{A}_k^j & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_k^j \end{bmatrix} \mathbf{h}_{1\to 2k,i}\ , \tag{2.91}$$

Similarly

$$(\tilde{\nu}_k^j(h_{1i},\ldots,h_{ki}))^2 = \sum_{i=1}^{n_R} \mathbf{h}_{k+1\to 2k,i}^H \mathbf{A}_k^j \mathbf{h}_{k+1\to 2k,i}$$

$$(\tilde{\nu}_k^j(h_{1i},\ldots,h_{2ki}))^2 = \sum_{i=1}^{n_R} \mathbf{h}_{1\to 2k,i}^H \frac{1}{2}\begin{bmatrix} \mathbf{0} & i\mathbf{\Theta}_k\mathbf{A}_k^j \\ -i\mathbf{\Theta}_k\mathbf{A}_k^j & \mathbf{0} \end{bmatrix} \mathbf{h}_{1\to 2k,i}\ , \tag{2.92}$$

then the following inductive step is true

$$\begin{aligned}
(\tilde{\mu}_{2k}^{j,j+1})^2 &\overset{(2.82)}{=} \sum_{i=1}^{n_R} (\tilde{\mu}_k^{j'}(h_{1i},\ldots,h_{2ki}))^2 \mp (\tilde{\nu}_k^{j'}(h_{1i},\ldots,h_{2ki}))^2 \\
&\overset{(2.91),(2.92)}{=} \sum_{i=1}^{n_R} [\mathbf{h}_{1\to k,i}^H \mathbf{h}_{k+1\to 2k,i}^H] \frac{1}{2}\begin{bmatrix} \mathbf{A}_k^{j'} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_k^{j'} \end{bmatrix} \begin{bmatrix} \mathbf{h}_{1\to k,i} \\ \mathbf{h}_{k+1\to 2k,i} \end{bmatrix} \\
&\quad \mp [\mathbf{h}_{1\to k,i}^H \mathbf{h}_{k+1\to 2k,i}^H] \frac{1}{2}\begin{bmatrix} \mathbf{0} & i\Theta_k\mathbf{A}_k^{j'} \\ -i\Theta_k\mathbf{A}_k^{j'} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{h}_{1\to k,i} \\ \mathbf{h}_{k+1\to 2k,i} \end{bmatrix} \\
&= \sum_{i=1}^{n_R} [\mathbf{h}_{1\to k,i}^H \mathbf{h}_{k+1\to 2k,i}^H] \frac{1}{2}\begin{bmatrix} \mathbf{A}_k^{j'} & \mp i\Theta_k\mathbf{A}_k^{j'} \\ \pm i\Theta_k\mathbf{A}_k^{j'} & \mathbf{A}_k^{j'} \end{bmatrix} \begin{bmatrix} \mathbf{h}_{1\to k,i} \\ \mathbf{h}_{k+1\to 2k,i} \end{bmatrix} \\
&= \sum_{i=1}^{n_R} \mathbf{h}_{1\to 2k,i}^H \mathbf{A}_{2k}^{j,j+1} \mathbf{h}_{1\to 2k,i} \\
&= \sum_{i=1}^{n_R} \mathbf{h}_{1\to 2k,i}^H \frac{1}{2}\begin{bmatrix} \mathbf{A}_k^{j'} & \mp i\Theta_k\mathbf{A}_k^{j'} \\ \pm i\Theta_k\mathbf{A}_k^{j'} & \mathbf{A}_k^{j'} \end{bmatrix} \mathbf{h}_{1\to 2k,i}\ ,
\end{aligned}$$

with $j = 1,3,\ldots,{}^{n_T}/_2 - 1$ and $j' = \frac{j+1}{2}$. That concludes the proof. □

2.5.4 Proof of Lemma 2.2.4

Proof. With 2.26, $(\mathbf{A}_{n_T}^j)^H$ and $(\mathbf{A}_{n_T}^{j+1})^H$, $j = 1,3,\ldots,{}^{n_T}/_2$ and $j' = \frac{j+1}{2}$ are given as

$$(\mathbf{A}_{n_T}^{j,j+1})^H = \frac{1}{2}\begin{bmatrix} (\mathbf{A}_{\frac{n_T}{2}}^{j'})^H & \mp i\mathbf{\Theta}_{n_T}(\mathbf{A}_{n_T}^{j'})^H \\ \pm i\mathbf{\Theta}_{n_T}(\mathbf{A}_{n_T}^{j'})^H & (\mathbf{A}_{\frac{n_T}{2}}^{j'})^H \end{bmatrix}.$$

Thus, $\mathbf{A}_{n_T}^{j,j+1}$ are only Hermitian, if $\mathbf{A}_{\frac{n_T}{2}}^{j'}$ is Hermitian. Since $\mathbf{A}_2$ is Hermitian, it follows that $\mathbf{A}_{n_T}^j$, $n_T = 2^n$,$j = 1,3,\ldots,{}^{n_T}/_2$ are Hermitian. Similarly,

$$(\mathbf{A}_{n_T}^{j,j+1})^H \mathbf{A}_{n_T}^{j,j+1} = \mathbf{A}_{n_T}^{j,j+1}\mathbf{A}_{n_T}^{j,j+1} = \frac{1}{4}\begin{bmatrix} 2\mathbf{A}_{\frac{n_T}{2}}^{j'}\mathbf{A}_{\frac{n_T}{2}}^{j'} & \mp i2\mathbf{\Theta}_{n_T}\mathbf{A}_{n_T}^{j'}\mathbf{A}_{n_T}^{j'} \\ \pm i2\mathbf{\Theta}_{n_T}\mathbf{A}_{n_T}^{j'}\mathbf{A}_{n_T}^{j'} & 2\mathbf{A}_{\frac{n_T}{2}}^{j'}\mathbf{A}_{\frac{n_T}{2}}^{j'} \end{bmatrix}.$$

Thus, $\mathbf{A}_{n_T}^{j,j+1}$ are only idempotent, if $\mathbf{A}_{\frac{n_T}{2}}^{j'}$ is idempotent. Since $\mathbf{A}_2$ is idempotent, it follows that $\mathbf{A}_{n_T}^{j}$, $n_T = 2^n$,$j = 1, 3, \ldots, {}^{n_T}/_2$ are idempotent. That concludes the proof. □

2.5.5 Proof of theorem 2.2.3

Proof. We first prove the independency of the eigenvalues. Since the matrices $\mathbf{A}_{n_T}^{(j)}$ are Hermitian and idempotent, we are able to rewrite (2.25) as

$$(\mu_j^{n_T})^2 = \sum_{i=1}^{n_R} \mathbf{h}_i^H (\mathbf{A}_{n_T}^j)^H \mathbf{A}_{n_T}^j \mathbf{h}_i = \sum_{i=1}^{n_R} \|\mathbf{A}_{n_T}^j \mathbf{h}_i\|^2 \ .$$

Independency between the eigenvalues in the Gaussian case is given if and only if the eigenvalues are uncorrelated, i.e.

$$E[(\mathbf{A}_{n_T}^j \mathbf{h}_i)^H \mathbf{A}_{n_T}^k \mathbf{h}_i] = 0 \ \ \forall j, j \neq k \ ,$$

which is fulfilled if

$$(\mathbf{A}_{n_T}^j)^H \mathbf{A}_{n_T}^k = \mathbf{A}_{n_T}^j \mathbf{A}_{n_T}^k = \mathbf{0} \ \ \forall j, j \neq k \ . \tag{2.93}$$

A more general and formal proof for this can be found in [MP92]. By applying the eigenvalue decomposition to (2.93), one has to distinguish between the case where the eigenvalues are given as

$$(\mathbf{I} - \mathbf{\Theta}_{n_T})\mathbf{A}_{\frac{n_T}{2}}^{j'}(\mathbf{I} + \mathbf{\Theta}_{n_T})\mathbf{A}_{\frac{n_T}{2}}^{k'}, \ (\mathbf{I} + \mathbf{\Theta}_{n_T})\mathbf{A}_{\frac{n_T}{2}}^{j'}(\mathbf{I} - \mathbf{\Theta}_{n_T})\mathbf{A}_{\frac{n_T}{2}}^{k'} \tag{2.94}$$

and

$$(\mathbf{I} - \mathbf{\Theta}_{n_T})\mathbf{A}_{\frac{n_T}{2}}^{j'}(\mathbf{I} - \mathbf{\Theta}_{n_T})\mathbf{A}_{\frac{n_T}{2}}^{k'}, \ (\mathbf{I} + \mathbf{\Theta}_{n_T})\mathbf{A}_{\frac{n_T}{2}}^{j'}(\mathbf{I} + \mathbf{\Theta}_{n_T})\mathbf{A}_{\frac{n_T}{2}}^{k'} \ , \tag{2.95}$$

with $j' = \frac{j+1}{2}$, $k' = \frac{k+1}{2}$ and $j' \neq k'$. Due to the special structure of the matrices $\mathbf{A}$ described in Remark 2.2.5, it follows for (2.94) that

$$(\mathbf{I} \mp \mathbf{\Theta}_{n_T})\mathbf{A}_{\frac{n_T}{2}}^{j'}(\mathbf{I} \pm \mathbf{\Theta}_{n_T})\mathbf{A}_{\frac{n_T}{2}}^{k'} = \underbrace{(\mathbf{I} \mp \mathbf{\Theta}_{n_T})(\mathbf{I} \pm \mathbf{\Theta}_{n_T})}_{\mathbf{0}} \mathbf{A}_{\frac{n_T}{2}}^{j'} \mathbf{A}_{\frac{n_T}{2}}^{k'} = \mathbf{0} \ .$$

Similarly for (2.95), we have

$$(\mathbf{I} \mp \mathbf{\Theta}_{n_T})\mathbf{A}_{\frac{n_T}{2}}^{j'}(\mathbf{I} \mp \mathbf{\Theta}_{n_T})\mathbf{A}_{\frac{n_T}{2}}^{k'} = (\mathbf{I} \mp \mathbf{\Theta}_{n_T})(\mathbf{I} \mp \mathbf{\Theta}_{n_T})\mathbf{A}_{\frac{n_T}{2}}^{j'} \mathbf{A}_{\frac{n_T}{2}}^{k'} \ .$$

Here, it follows that $\mathbf{A}_{n_T}^j \mathbf{A}_{n_T}^k = \mathbf{0}$ only, if $\mathbf{A}_{\frac{n_T}{2}}^{j'} \mathbf{A}_{\frac{n_T}{2}}^{k'} = \mathbf{0}$. Since the basic case $\mathbf{A}_4^{j'} \mathbf{A}_4^{k'} = \mathbf{0}$, it follows that $\mathbf{A}_{n_T}^j \mathbf{A}_{n_T}^k = \mathbf{0}$ $\forall n_T, j, j \neq k$. Thus, the eigenvalues in (2.94) and (2.95) are zero and therefore the eigenvalues in (2.25) are independent.

The probability density function (pdf) of the eigenvalues $p((\mu_j^{n_T})^2)$ can be obtained from (2.25) as follows. The rank (rk(·)) of $\mathbf{A}_2$ is 2. Furthermore,

$$\begin{aligned} \mathrm{rk}(\mathbf{A}_{2^n}^j) &= \mathrm{rk}(\mathbf{U}\mathbf{A}_{2^n}^j\mathbf{U}) = \mathrm{rk}\begin{bmatrix} (\mathbf{I} + \mathbf{\Theta}_{n_T})\mathbf{A}_{2^{n-1}}^{j'} & \mathbf{0} \\ \mathbf{0} & (\mathbf{I} - \mathbf{\Theta}_{n_T})\mathbf{A}_{2^{n-1}}^{j'} \end{bmatrix} \\ &= \mathrm{rk}((\mathbf{I} + \mathbf{\Theta}_{n_T})\mathbf{A}_{2^{n-1}}^{j'}) + \mathrm{rk}((\mathbf{I} - \mathbf{\Theta}_{n_T})\mathbf{A}_{2^{n-1}}^{j'}) = \mathrm{rk}(\mathbf{A}_{2^{n-1}}^{j'}) \ , \end{aligned}$$

where $\mathbf{U}$ contains the eigenvectors of $\mathbf{A}_{2^n}^j$. Since $\mathrm{rk}(\mathbf{A}_2) = 2$, the matrices $\mathbf{A}_{n_T}^j$ have all rank 2, and thus the following holds

$$\mathbf{V}^H(\mathbf{A}_{n_T}^j)\mathbf{V} = \begin{bmatrix} \mathbf{I}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} , \tag{2.96}$$

where $\mathbf{V}$ is a unitary matrix. With (2.96), the pdfs are given as

$$p((\mu_j^{n_T})^2) \quad = \quad p\left(\mathrm{tr}\left[\sum_{i=1}^{n_R} \mathbf{h}_i^H \mathbf{A}_{n_T}^j \mathbf{h}_i\right]\right) = p\left(\mathrm{tr}\left[\sum_{i=1}^{n_R} \bar{\mathbf{h}}_i^H \begin{bmatrix} \mathbf{I}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \bar{\mathbf{h}}_i\right]\right) ,$$

which is the sum of squares of $2n_R$ independent complex normal distributed variables, i.e. a noncentral chi-square distribution with $4n_R$ degrees of freedom. That concludes the proof. □

2.5.6 Proof of corollary 2.3.1

Proof. The inequality (2.36) is tight only, if the ratio of two eigenvalues, i.e. $r = \mu_i^2/\mu_j^2 = 1$, for all $i \neq j$. As a result it has to be shown that the following applies

$$\lim_{n_R \to \infty} \Pr(r = 1) = 1 .$$

Since the eigenvalues are chi-square distributed with each $4n_R$ degrees of freedom, the ratio of the eigenvalues is distributed as follows

$$h(r, n_r) = \frac{\Gamma(4n_R)}{\Gamma(2n_R)^2} \frac{r^{(4n_R-2)/2}}{(1+r)^{4n_R}} ,$$

which is the well-known F distribution [Spr79, p.365, eq.9.9.35]. Therefore, when n_R goes infinity, the F distribution is given as

$$\lim_{n_R \to \infty} (h(r, n_R)) = \delta(r-1) ,$$

where δ is the delta distribution. It follows that the lower bound from (2.35) becomes tight as n_R increases. The lower bound from (2.35) is also tight for low SNR values, which is obvious after rewriting (2.33) as follows

$$I_Q = \frac{2}{n_T} \log_2 \left(\left(1 + \frac{\rho}{n_T}\alpha_1\right)^{\frac{n_T}{2}} - \zeta \right) .$$

Furthermore, $(1+\frac{\rho}{n_T}\alpha_1)^{\frac{n_T}{2}} \gg \zeta$ for small SNR. Therefore, as the SNR gets smaller, the lower bound from (2.35) gets tighter. That concludes the proof. □

2.5.7 Proof of theorem 2.3.2

Proof. Let $x_l = |x_l| \exp(i\theta_{x_l})$, where x_l is the absolute value and θ_{x_l} the angle of x_l. By substituting this in (2.53) it follows that

$$\zeta = \frac{1}{4} \min_{\substack{x_l = c_l - e_l, \\ \mathbf{c} \neq \mathbf{e}}} \left| \left(\sum_{l=1}^{4} |x_l|^2 \right)^2 - 4 \left| |x_4||x_2| \Theta_{x_4}^{x_2} + |x_1||x_3| \Theta_{x_1}^{x_3} \right|^2 \right|^{\frac{1}{4}} , \tag{2.97}$$

where $\Theta_{x_m}^{x_n}$ is given by $\Theta_{x_m}^{x_n} = \sin(\theta_{x_n} - \theta_{x_m})$. The maximum value of $\left|\Theta_{x_m}^{x_n}\right|$ is equal to one. Maximizing $\Theta_{x_m}^{x_n}$ leads to a lower bound for ζ given as

$$\zeta \geq \frac{1}{4} \min_{\substack{x_l = c_l - e_l, \\ \mathbf{c} \neq \mathbf{e}}} \left| \left(\sum_{l=1}^{4} |x_l|^2 \right)^2 - 4 \left| |x_4||x_2| + |x_1||x_3| \right|^2 \right|^{\frac{1}{4}} .$$

After some manipulations we arrive at

$$\zeta \geq$$

$$\frac{1}{4} \min_{\substack{x_l = c_l - e_l, \\ \mathbf{c} \neq \mathbf{e}}} \left| \sum_{l=1}^{2} \left(|x_l|^2 - |x_{l+2}|^2 \right)^2 + 2 \left(|x_1||x_4| - |x_2||x_3| \right)^2 + 2 \left(|x_1||x_2| - |x_4||x_3| \right)^2 \right|^{\frac{1}{4}} .$$

At this point, we have to consider different situations. For the case $|\omega| = 4$, i.e. $x_k \neq 0, \forall k \in \omega, \zeta_4$ is given as $\zeta_4 \geq 0$. For $|\omega| = 3$, e.g., $x_2 = 0$ and $x_k \neq 0, \forall k, k \neq 2$, we have

$$\zeta_3 \geq \frac{1}{4} 5^{\frac{1}{4}} d_{\min}. \tag{2.98}$$

Further on, for $|\omega| = 2$, e.g., $x_2 = x_4 = 0$ and $x_k \neq 0, \forall k, k \neq \{2, 4\}$, we have $\zeta_2 \geq 0$. In the case $|\omega| = 0$, e.g.,$x_l = 0, \forall l, l \neq 4$, we get from (2.97) directly that

$$\zeta_1 = \frac{1}{4} d_{\min}. \tag{2.99}$$

□

2.5.8 Proof of theorem 2.3.3

Proof. Due to the symmetry of M-PSK constellations, rotating the initial constellation around itself by an angle $\phi = 2k\pi/M, k \in \mathbb{Z}$ results in the initial constellation. From this it follows that also the diversity product is rotationally symmetrical with respect to rotations by angles of $\phi = 2k\pi/M$, i.e. $\zeta(\phi) = \zeta(\phi + 2k\pi/M)$. Furthermore, due to symmetry, rotating the constellation by an angle $\phi = k\pi/M + \theta$ or $\phi = k\pi/M - \theta, 0 \leq \theta \leq 2\pi$ yields the same diversity product, i.e. $\zeta(k\pi/M + \theta) = \zeta(k\pi/M - \theta)$. It follows that we have to consider only the range $\frac{2k\pi}{M} \leq \phi < \frac{(2k+1)\pi}{M}$ in the remainder of the proof. For simplicity we choose $k = 0$. Consider (2.97) in order to analyze ζ_4 and ζ_2. After some manipulations we arrive at the following expression for ζ_2

$$\zeta_2 = \frac{1}{4} \min_{\substack{x_l = c_l - e_l, \\ \mathbf{c} \neq \mathbf{e}}} \left| \left(|x_2|^2 - |x_4|^2 \right)^2 + 4|x_2|^2||x_4|^2 \cos^2(\Theta_2) \right|^{\frac{1}{4}} ,$$

where $\Theta_2 = \theta_2 - \theta_4$. Remember from Remark 2.3.1 that $\theta_4 = \theta_4^{orig} - \phi$, where ϕ is the rotation angle. As a result $\Theta_2 = \frac{2\pi}{M} z + \phi, 2 \leq |z| \leq 2(M-1)$. It holds that $\min |\cos(\Theta_2)| = |\cos(\frac{\pi}{2} + \phi)|$ and that $|\cos(\frac{\pi}{2} + \phi)| = |\sin(\phi)|$. With $|x_2|^2 = |x_4|^2 = d_{\min}^2$ it follows that

$$\zeta_2 = \frac{1}{4} d_{\min} \sqrt{2 \left|\sin(\phi)\right|} \; 0 \leq \phi \leq \frac{\pi}{M}. \tag{2.100}$$

For ζ_4 we arrive at (with $|x_1|^2 = |x_3|^2$ and $|x_2|^2 = |x_4|^2$)

$$\zeta_4 = \frac{1}{4} \min_{\substack{x_l = c_l - e_l, \\ \mathbf{c} \neq \mathbf{e}}} \left| 8|x_1|^2|x_2|^2 + 4 \left[\left(|x_1|^2 \cos(\Theta_1) - |x_2|^2 \cos(\Theta_2) \right)^2 + 2|x_1|^2||x_2|^2 \cos(\Theta_1 + \Theta_2) \right] \right|^{\frac{1}{4}} ,$$

where $\Theta_1 = \theta_3 - \theta_1$. The minimum is achieved if $|x_1|^2 \cos(\Theta_1) = |x_2|^2 \cos(\Theta_2)$. Thus, $|x_1|^2 = |x_2|^2 = d_{\min}^2$ and $\cos(\Theta_1) = \cos(\Theta_2)$, i.e. $|\Theta_1| = |\Theta_2|$. After some manipulations we arrive at

$$\zeta_4 = \frac{1}{4} d_{\min} \sqrt{4\,|\sin(\phi)|} \;\; 0 \leq \phi \leq \frac{\pi}{M}. \tag{2.101}$$

By comparing (2.100) and (2.101), it follows that $\zeta_4 \geq \zeta_2$. □

3 Iterative detection and Turbo decoding of Space-Time Codes

In this chapter, we analyze the performance of various Space-Time Codes by deploying an iterative detection and decoding technique at the receiver side of the MIMO system. In case of layered STC, we apply a new iterative signal processing algorithm, which can be seen as an extension of many well-known detection algorithms for multiple-input-multiple-output (MIMO) channels, e.g. zero-forcing (ZF) and successive nulling and cancelation algorithms. As examples, we apply our algorithm to the VBLAST algorithm proposed by [WFGV98] and the ZF algorithm. We show that our approach takes advantage of the maximum available diversity level to improve the initial data estimate obtained by the VBLAST (or ZF) algorithm with minimum additional effort. For the orthogonal and the unitary STC, we employ a "turbo" coding scheme at the transmitter, consisting of the serial concatenation of an outer code and the STC as the inner code. We again apply iterative space-time detection and decoding at the receiver with the difference, that we exchange reliability or soft information about the estimates from the detector to the decoder and vice versa. Similar receiver schemes have been proposed for the layered STC in [SH00]. In the case of orthogonal STC we analyze the impact of different mapping strategies on the information transfer of the soft-input-soft-output (SISO) space-time detector and derive some criteria for the optimum mapping strategy. We show that additional performance gains over Gray mapping can be obtained by different mapping strategies. Furthermore, we use extrinsic information transfer characteristics (EXIT-charts) in order to predict the performance and the behavior of the whole system.

We also study the iterative decoding of Wrapped STC proposed by [CC01, CC03] employing per-survivor-processing at the receiver with the soft-output Viterbi-algorithm (SOVA, [HH89]). With the availability of extrinsic information at the receiver delivered by the SOVA, it is now possible to use a novel receiver scheme performing iterative (turbo) decoding in order to improve the performance of the architecture. According to the ZF or to the MMSE criterion, linear feedforward and -backward filters of the decision-feedback space-time (ST) decoder are derived and the decision metric of the SOVA employing per-survivor-processing is developed.

3.1 Introduction

As distinct from the diversity oriented schemes, the error performance of spatial multiplexing schemes like BLAST is made relatively worse by applying linear detection schemes. In contrast, the optimal (maximum likelihood) detector offers a decent level of performance, however, it is incredibly complex and therefore impractical. However, by applying a suboptimal iterative detection algorithm, as described in chapter 3.2, as an extension to suboptimal detectors the performance can be dramatically improved. Furthermore, in order to better evaluate the performance and compare it with other architectures, the ST scheme is concatenated with a channel encoder. The concept of concatenation of different signal processing constituents and iterative detection and decoding in order to achieve a better performance is

also utilized by turbo coding, a 1993 revolution in coding theory [BG96], achieving near Shannon capacity in additive white Gaussian noise (AWGN) channels. In order to improve the performance of space-time (ST) transmission schemes, the concept of concatenation may also be applied to achieve reliable communication over MIMO channels. Accordingly, a number of different architectures have recently been proposed which utilize conventional concatenated encoder structures in ST coding applications [Bau99, SD99], mostly as replacements or augmentations to the traditional space-time transmission schemes [TSC98, TJC99b, SH00]. In opposition to that, in this chapter, the ST coding part is used as the inner constituent encoder, whereas, at the receiver, the decoder associated with the ST coding component is modified so that it produces soft-output in form of a posteriori probabilities (APPs) of the data bits. Furthermore, decoding is done in an iterative fashion between the ST decoder and the channel decoder until only negligible performance is attained.

3.2 Iterative detection of Layered STC

In order to achieve the capacity of MIMO systems as promised in [Win87, Tel99, FG98], Diagonal Bell Labs Architecture Layered Space-Time (DBLAST)- an architecture which theoretically achieves capacity for such MIMO channels- has been proposed by Foschini in [Fos96]. However, the high complexity of the algorithm implementation is a substantial drawback. Vertical BLAST (VBLAST), a simplified and suboptimal version of the BLAST architecture, using ordered successive nulling and interference cancelation at the receiver, is capable of achieving high capacity with low complexity. A drawback of VBLAST is that the nulling and canceling algorithm works only when there are at least as many receive antennas as transmit antennas. An extension of VBLAST to the case in which the number of receiver antennas is less than the number of transmit antennas is proposed by Hassibi and Hochwald in [HH02] and referred to as high-rate linear dispersion codes. To avoid the repeated computation of the pseudo inverse of the channel matrix in VBLAST, a very efficient detection algorithm based on sorted QR decomposition is used in [WBR$^+$01].

The focus of the BLAST architectures is to achieve high transmission rates in opposition to e.g., STTC [TSC98, GFBK99] and STBC [Ala98, TJC99a]). These approaches exploit multiple antennas at both the transmitter and receiver in order to obtain transmit and receive diversity and therefore increase the reliability of the system. Tse and Zheng [ZT03] have shown that there is a fundamental tradeoff between the amount of diversity (diversity gain) and the number of degrees of freedom (multiplexing gain) in multiple antenna channels. In [ZT03], they focus on the high signal-to-noise-ratio regime and give a simple characterization of the optimal tradeoff curve achievable by any scheme. We will use the methods applied in [ZT03] to underline the performance gain achieved by the algorithm proposed in this section and to develop this new ISIP algorithm.

3.2.1 Transceiver structure

The model of the system described already in chapter 2 still applies, however, for convenience, we rewrite it as

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{N} \,. \tag{3.1}$$

Furthermore, the transmit matrix differs from that of the diversity oriented schemes in the following way. In order to obtain the transmit matrix the data is first demultiplexed in n_T data substreams, so called layers, with equal rates. Each layer

is associated with a transmit antenna. After optionally coding the layers with a forward error correction code (FEC), the layers are then mapped onto a M-Phase-shift-keying (PSK) or M-Quadrature Amplitude Modulation (QAM) signal constellation, where $|M|$ denotes the cardinality of the constellation.

Of course, the optimum way to detect the transmitted data is using a maximum-likelihood-detector (MLD), but since the computational complexity is prohibitively high, it is not feasible. Ref. [WFGV98] therefore proposed a low complexity detection algorithm, which will shortly be reviewed in the following.

Each entry in the receive matrix is a superposition of transmitted symbols from each layer l ($1 \leq l \leq n_T$) for a given time t ($1 \leq t \leq T$) scaled by the multipath fading coefficients and corrupted by AWGN . Conceptually, one symbol of a layer is regarded as the desired signal, and the remainder are considered as interferers. Suppose that the signal of layer l is the desired signal. To suppress the influence of the interferers, a nulling step is performed by linearly weighting the received signal with a so called nulling-vector $\mathbf{w}_m^T$

$$\mathbf{r}_l = (\mathbf{w}_m^T \mathbf{Y})^T,$$

where

$$\mathbf{w}_m^T \mathbf{h}_l = \begin{cases} 0 & m \neq l \\ 1 & m = l \ . \end{cases} \tag{3.2}$$

That means, that the nulling-vector $\mathbf{w}_m^T$ has to be orthogonal to all columns of the channel matrix $\mathbf{H}$, but column l, since column l regulates the influence of the layer l. To obtain the nulling-vector $\mathbf{w}_m$ one way is to compute the Moore-Penrose pseudo-inverse of the channel matrix $\mathbf{H}$

$$\mathbf{G} = \mathbf{H}^+,$$

since $\mathbf{w}_m^T$ is equivalent to the row of G with minimum norm, i.e.

$$\mathbf{w}_m^T = \mathbf{g}_m, \quad \text{s.t.} \quad m = \arg\min_j ||\mathbf{g}_j||^2.$$

In order to obtain the data estimation of the current layer we apply the quantization operation $\mathcal{Q}(\cdot)$ appropriate to the signal constellation in use, i.e.

$$\hat{\mathbf{x}}_l = \mathcal{Q}(\mathbf{r}_l) \ . \tag{3.3}$$

After this step we cancel the interference caused by the layer l on the other layers by subtracting from the receive matrix, i.e.

$$\mathbf{Y}_{i+1} = \mathbf{Y}_i - \mathbf{h}_l[\mathbf{0}, \ldots, \mathbf{0}, \hat{\mathbf{x}}_l, \mathbf{0}, \ldots, \mathbf{0}]^T$$

and zeroing the column l of the channel matrix, which is expressed with the following notation

$$\mathbf{H}_i = \mathbf{H}_{i \setminus \{l\}} \ .$$

3.2.2 Diversity gain of the system

It is well known from [Pro01, vZ00], that the bit error rate (BER) of a system with n_T transmit and n_R receive antennas applying ZF at the receiver can be

approximated for high SNR ρ as

$$P_e \approx \left(\frac{1}{4\rho}\right)^d \binom{2d-1}{d}.$$

Here, we observe that the BER decreases inversely with the d-th power of the signal-to-noise-ratio ρ, where d denotes the diversity gain or diversity order. To derive performance behavior in the high-SNR regime of a VBLAST system we use the technique applied in [ZT03] to compute the diversity gain of a scheme as

$$\lim_{\rho\to\infty} \frac{\log(f(\rho))}{\log(\rho)} = -d, \tag{3.4}$$

where $f(\rho)$ is given as $P_e(\rho) \doteq f(\rho) = \rho^{-d}$ and $\doteq$ denotes exponential equality. For more insight, we refer the interested reader to [ZT03]. For VBLAST without error propagation due to decision errors (genie-aided VBLAST) $f(\rho)$ is given as

$$f(\rho) = \sum_{i=1}^{n_T} \rho^{-d_i}, \tag{3.5}$$

where $d_1 < d_2 < \ldots < d_{n_T}$ and ρ^{-d_i} describes the performance behavior in the high SNR regime of the i-th layer in a VBLAST system. With (3.5), (3.4) can be written as

$$\begin{aligned}
\lim_{\rho\to\infty} \frac{\log(f(\rho))}{\log(\rho)} &= \lim_{\rho\to\infty} \frac{\log\left(\sum_{i=1}^{n_T} \rho^{-d_i}\right)}{\log(\rho)} = \lim_{\rho\to\infty} \frac{\log\left(\rho^{-d_1}\left(1+\sum_{i=2}^{n_T} \rho^{-d_i+d_1}\right)\right)}{\log(\rho)} \\
&= \lim_{\rho\to\infty} \frac{\log\left(\rho^{-d_1}\right) + \log\left(1+\sum_{i=2}^{n_T} \rho^{-d_i+d_1}\right)}{\log(\rho)} \\
&= -d_1 + \underbrace{\lim_{\rho\to\infty} \frac{\log\left(1+\sum_{i=2}^{n_T} \rho^{-d_i+d_1}\right)}{\log(\rho)}}_{=0}.
\end{aligned}$$

Thus, the overall diversity gain of the VBLAST system equals the diversity gain of the first layer to be detected. The performance of the first layer to be detected is therefore important in order to improve the average BER. The proposed ISIP algorithm, which is described in the following section 3.2.3, theoretically (i.e, in the genie-aided case) takes advantage of the maximum available diversity level to detect all layers with the same maximum diversity and therefore improve the BER performance.

3.2.3 Iterative Signal Processing (ISIP) algorithm

As aforementioned the new ISIP algorithm approach can be used as an extension to many detection algorithms for multiple-input-multiple-output channels. Assume now, that an initial estimation of the transmitted symbols is available, which can be obtained by e.g., VBLAST. The aim of the ISIP algorithm is to improve the performance of each layer by using the initial estimates of the other layers. By canceling the interference from the other layers, the desired layer can be detected with the maximum diversity available. Let us now describe the ISIP algorithm in detail: We compute the Maximum-Ratio-Combining (MRC)-vector from the columns of

the channel matrix $\mathbf{H}$. We start at layer l with the worst BER-performance to get the MRC-vector corresponding to this layer with

$$\mathbf{w}_l^T = \frac{\mathbf{h}_l^H}{||\mathbf{h}_l||^2}. \tag{3.6}$$

We get the information of the order of processing the layers from the VBLAST algorithm according to the norm of each row of the matrix $\mathbf{G}$. In case we do not have this order information, e.g. ZF, we start at an arbitrary layer. We have to eliminate the influence of the current layer l by zeroing the column of the actual channel matrix $\mathbf{H}$ corresponding to this layer, which is expressed as

$$\mathbf{H}_l = \mathbf{H}_{\setminus\{l\}} \ .$$

By subtracting the initial data estimation from the receive matrix

$$\mathbf{Y}_l = \mathbf{Y} - \mathbf{H}_l\hat{\mathbf{X}},$$

we now have the maximum diversity level to detect the current layer. At this point, our system is equivalent to a single-input-multiple-output (SIMO) channel with $n_T = 1$ transmit and n_R receive antennas. We have to multiply the remaining receive matrix $\mathbf{Y}_l$ with the MRC-vector

$$\mathbf{r}_l = \mathbf{w}_l^T\mathbf{Y}_l \tag{3.7}$$

and apply the quantization operation like in (3.3)

$$\hat{\mathbf{x}}_l = \mathcal{Q}(\mathbf{r}_l) \ . \tag{3.8}$$

Eq. (3.7) is nothing else than MRC for a system with $n_T = 1$ transmit and n_R receive antennas. The steps from (3.6) to (3.8) are computed iteratively, where the number of iterations is denoted as I_N. For illustration the complete ISIP algorithm is shown in Algorithm 1.

Algorithm 1 : ISIP algorithm

for $i = 1$ to I_N **do**
  for $l = 1$ to n_T **do**

$$1: \quad \mathbf{w}_l^T \Leftarrow \frac{\mathbf{h}_l^H}{||\mathbf{h}_l||^2}$$

$$2: \quad \mathbf{H}_l \Leftarrow \mathbf{H}_{\setminus\{l\}}$$

$$3: \quad \mathbf{Y}_l \Leftarrow \mathbf{Y} - \mathbf{H}_l\hat{\mathbf{X}}$$

$$4: \quad \mathbf{r}_l \Leftarrow \mathbf{w}_l^T\mathbf{Y}_l$$

$$5: \quad \hat{\mathbf{x}}_l \Leftarrow \mathcal{Q}(\mathbf{r}_l)$$

  end for
end for

3.2.4 Numerical simulation

BER versus SNR simulations were carried out in order to compare the regular VBLAST system with and without ISIP in the following subsection 3.2.4.1. Furthermore, we compare a receiver applying regular VBLAST with a receiver applying ZF with and without ISIP in subsection 3.2.4.2.

3.2.4.1 VBLAST vs. VBLAST with ISIP

In Fig. 3.1, we present the bit error rates (BERs) of each layer for a system applying binary-shift-keying (BPSK) modulation with $n_T = n_R = 4$ transmit and receive antennas of genie-aided VBLAST with and without ISIP. In addition, the average BERs of genie-aided VBLAST with and without ISIP are depicted in Fig. 3.1. We assume, that the channel is constant for $T = 100$ symbols per layer. For the genie-aided VBLAST we choose the number of iterations I_N to 1, since, with perfect interference cancelation, there is no need for more iterations.

As expected from the genie-aided VBLAST algorithm the BER performance improves in terms of diversity level from layer to layer. It is worth knowing, that the nulling-vector of the first layer to be detected is the one with the minimum norm and therefore the one with the minimum noise enhancement. Henceforth, we will refer to this layer as layer 1. But since layer 1 is detected first and, therefore, with the minimum diversity level, it provides the worst BER performance and will dominate the error performance of the system (confer Section 3.2.2. After consecutively detecting all layers according to the VBLAST algorithm, we have an initial data estimation of the transmitted symbols. By applying ISIP to the initial data estimations of the individual layers every layer is detected with the highest available diversity level of the system. Therefore layer 1 provides the best BER performance, layer 2 provides the second best BER performance and so on. The BER performance of the last layer of the system, in this case layer 4, is not improved through ISIP in the genie-aided case (vide Fig. 3.1), since it is detected with the maximum available diversity level within the VBLAST algorithm. But as we will see later the performance of layer 4 will significantly improve in the non-genie case of the VBLAST algorithm. It is clearly seen in Fig. 3.1, that VBLAST with ISIP performs significantly better than VBLAST alone.

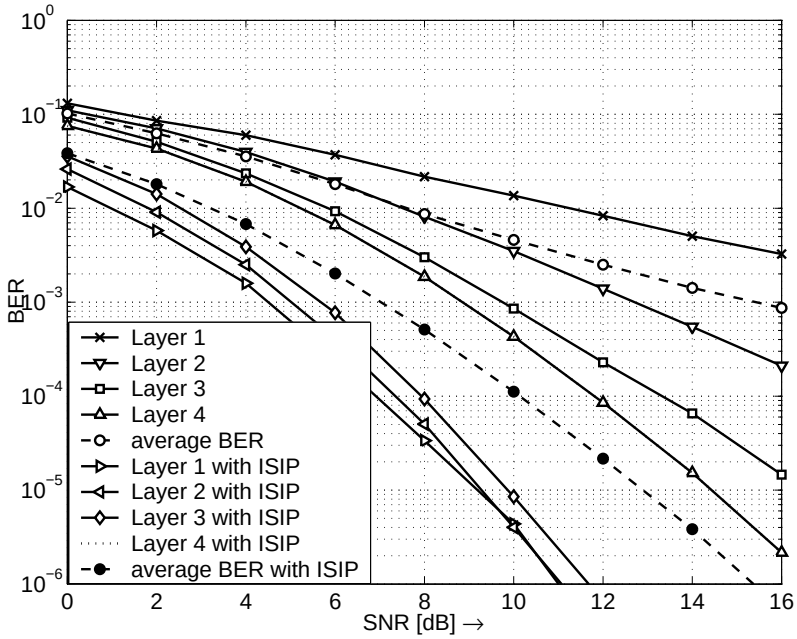

Figure 3.1: BER of the individual layers and average BERs (dashed lines) for the Genie-aided VBLAST with and without ISIP for $n_T = n_R = 4$, uncoded BPSK modulation. BERs of layer 4 with (line with △) and without ISIP (dotted line) are identical.

In Fig. 3.2 the average BERs of the non-genie aided VBLAST, with and without ISIP for a system with $n_T = n_R = 4$ transmit and receive antennas and BPSK modulation, are depicted. Also shown for reference is the simulated BER performance of the maximum-likelihood estimator (MLE). In this non-genie aided case we choose the number of iterations I_N to 4, since after the 4-th iteration the improvement is negligible. While the BER of regular VBLAST decays approximately like $\rho^{-6/5}$, the BER for VBLAST with ISIP decreases with ρ at a faster speed of approximately $\rho^{-3/2}$. Furthermore, there is a shift of the BER-curve of about 5dB at an error rate of $2 \cdot 10^{-3}$. The detrimental effect of error propagation is obvious, when we compare the average BERs in Fig. 3.2 with the average BERs of genie-aided VBLAST in Fig. 3.1. A reduction of this effect can be obtained by applying appropriate coding techniques.

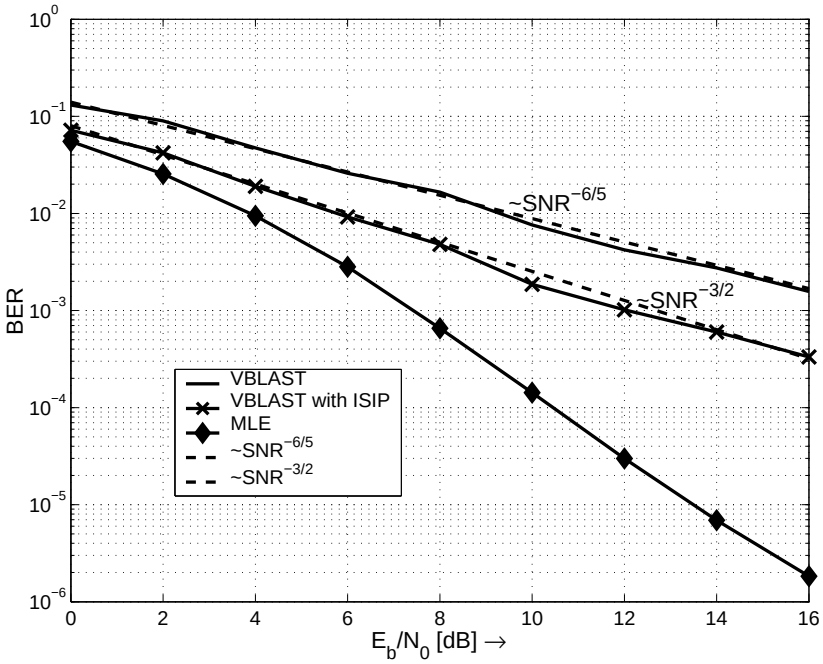

Figure 3.2: Average BER of non-genie VBLAST system ($n_T = n_R = 4$ antennas) with and without ISIP and BER of MLE, uncoded BPSK modulation.

3.2.4.2 VBLAST vs. ZF with ISIP

The average BERs of regular VBLAST, ZF with and without ISIP are shown in Fig. 3.3 for a system with $n_T = n_R = 4$ transmit and receive antennas and BPSK modulation. We choose the number of iterations to be $I_N = 3$. While the performance of ZF alone is relatively worse, the concatenation of ZF with ISIP provides BER performance comparable to VBLAST. In order to achieve a BER of 10^{-4} there is a difference to VBLAST, in terms of required SNR, of about 10dB and about 3dB to ZF and ZF with ISIP, respectively.

To avoid an overload in Fig. 3.3, the average BER after each iteration is depicted in Fig. 3.4. We observe that there is a significant improvement of the BER performance at the first iteration with a saturation of the improvement beyond the second iteration.

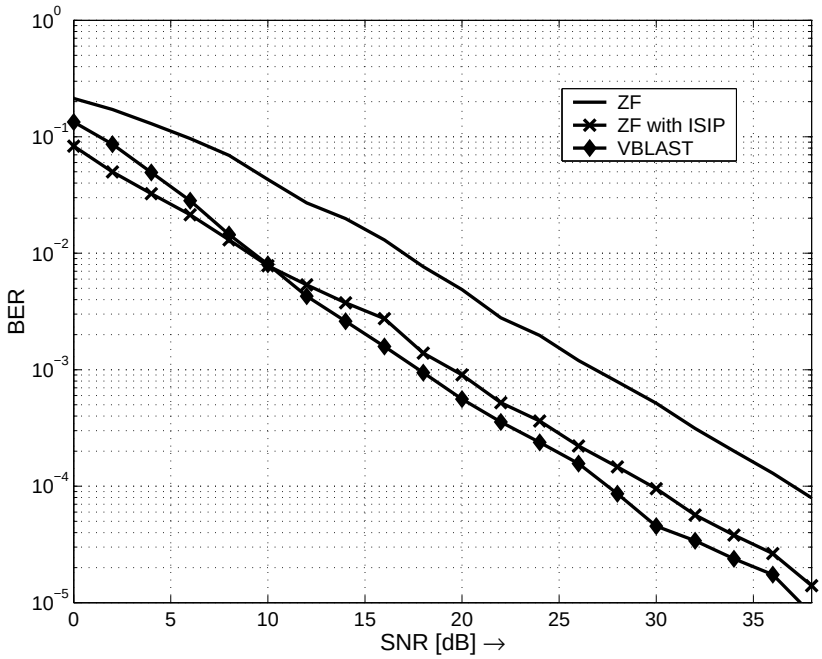

Figure 3.3: Comparing average BERs of a system with $n_T = n_R = 4$ antennas with different receivers like ZF, VBLAST, ZF with ISIP, uncoded BPSK modulation.

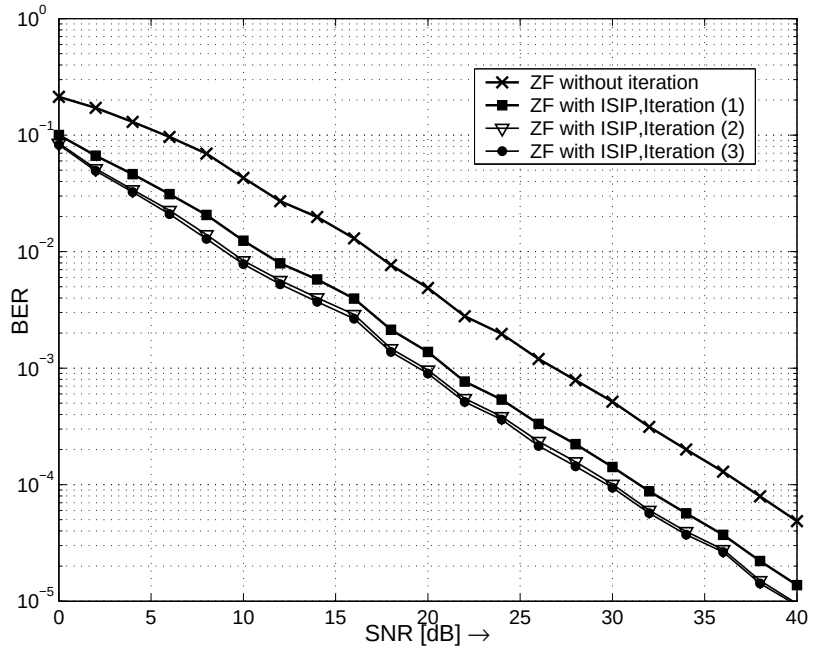

Figure 3.4: Average BERs after each iteration of a system with $n_T = n_R = 4$ antennas applying ZF with ISIP, uncoded BPSK modulation.

3.2.4.3 Performance of ISIP with Channel Coding

In order to improve the performance of the system applying VBLAST or VBLAST with ISIP, each layer is now encoded with a FEC. We apply the convolutional code $CC(7,5)_{oct}$ with code rate $R = 1/2$ to each layer at the transmitter. At the receiver, we apply the Viterbi algorithm after the quantization operation and use the decoded anew encoded data for further processing. In Fig. 3.5, the FER of a coded and uncoded system applying BPSK with $n_T = 4$ transmit and $n_R = 4$ receive antennas is depicted. The frame length is set to $T = 100$ per layer. From the figure, we observe a performance enhancement of the coded system to the uncoded system. Furthermore, the performance gain achieved by applying the ISIP algorithm to VBLAST is significantly improved for the coded system in comparison to the uncoded system.

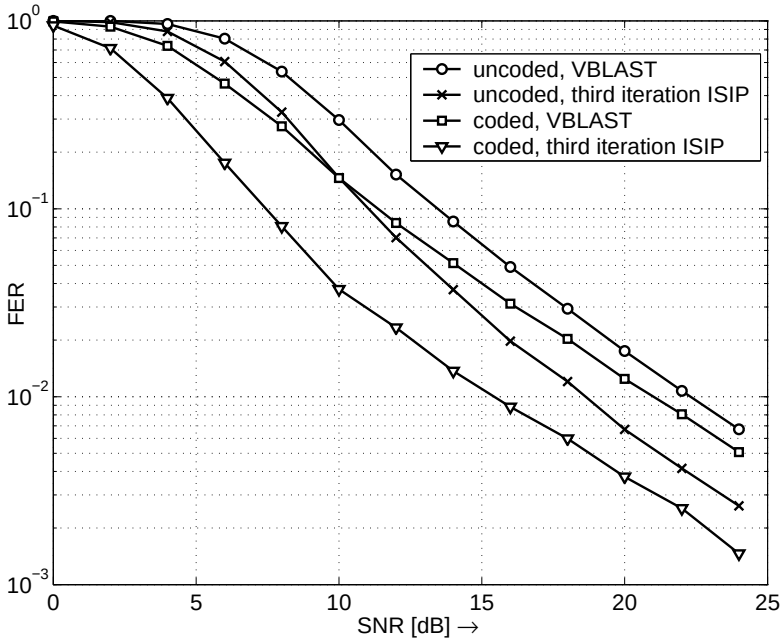

Figure 3.5: FER for VLBAST and VLBAST in concatenation with ISIP of a system with $n_T = n_R = 4$ antennas, uncoded and convolutionary encoded with $CC(7,5)_{oct}$, BPSK modulation.

In order to further improve the performance of the ST scheme, soft-information of the data bits can be produced by the ST detector and fed into the channel decoder. The channel decoder itself produces soft-information about the data bits and this output is fed back to the ST detector and used as additional information during the next iteration. This strategy is applied in the following sections.

3.3 Turbo Decoding of Orthogonal STC

In practice a system designer is interested in the performance of a space-time transmission scheme in concatenation with a channel code, since this is a promising means to improve the performance of a wireless communication system in a fading environment. Accordingly, there is a huge amount of work in the literature

focusing on merging multiple antenna systems with proper channel coding in order to employ both coding and multiple antenna gains. More recently, the authors of [Bau99, LH02, SPS02a] proposed to use a powerful channel code (e.g. turbo codes [BG96]) in concatenation with a space-time unitary matrix differential modulation code or orthogonal space-time block codes (STBC) [Ala98, TJC99a] in order to achieve significant coding gains. STBC from orthogonal design provide the maximum possible diversity gain for a multiple-input-multiple-output (MIMO) channel, but no coding gain.

In this section, we consider the serial concatenation of an outer code with STBC [TJC99a] as the inner code in order to improve the performance of the uncoded system, and in order to approach the capacity promised by the information theoretic results [Win87, Tel99, FG98, BCT01]. Differently from [Bau99, LH02], decoding at the receiver is done in an iterative manner between the soft-input-soft-output (SISO) space-time detector and the SISO channel decoder by exchanging soft-outputs of the data-bits. We consider different mapping strategies and analytically show that it is possible to improve the performance by employing other mapping schemes than Gray mapping. Furthermore, we derive some mapping criteria to obtain better performance for iterative decoding. Very recently, mapping strategies for turbo detection have also been investigated and optimized with respect to different criteria in [BS03]. We analyze the space-time detection and decoding components with extrinsic information transfer (EXIT) charts, which have been proposed in [tB99] as a quasi-analytical tool for predicting the convergence behavior and the performance of concatenated coding systems.

The contribution of this section is the

- analysis of the performance of STBC within an iterative decoding scheme,
- application of both mutual information and EXIT charts to facilitate the choice of an optimal mapping for STBC, and
- derivation of two design criteria for optimal mapping.

3.3.1 Transmitter structure

We combine a STBC with a channel code in serial via a pseudo random interleaver in order to achieve low probability of error for small signal-to-noise-ratios (SNR). As shown in Fig. 3.6, after the encoding and interleaving step the F_L coded bits in the bit sequence $\{c_1, \ldots, c_{F_L}\}$, where F_L denotes the frame length, are mapped onto symbols $s \in \mathcal{C}$ from a given constellation $\mathcal{C}$, e.g. M-PSK.

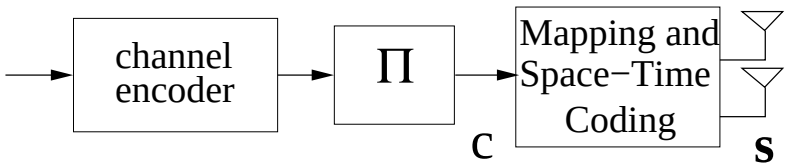

Figure 3.6: Model of the transmitter with channel encoder, interleaver,mapper and Space-Time Block Coder.

The function $s = f(\mathbf{c})$ describes the mapping of $m = \log_2(M)$ consecutive bits contained in the vector $\mathbf{c}$ onto one constellation symbol s. The symbols s are then space-time coded according to the $p \times n_T$ space-time generator matrix $\mathcal{G}_{n_T}$ or $\mathcal{H}_{n_T}$ [TJC99a]. The code rate R is given by $R = q/p$, where q is the number of different symbols and p is the number of time samples.

Our system model described already in chapter 2 still apply, however, we repeat it here briefly for convenience

$$r_t^j = \sum_{i=1}^{n_T} h_{i,j} s_t^i + n_t^j \ , \tag{3.9}$$

where r_t^j is the received signal at time t and receive antenna j, s_t^i is the transmitted signal from a given constellation at time t and transmit antenna i, $h_{i,j}$ is the complex channel path gain from transmit antenna i to receive antenna j and n_t^j is a complex Gaussian random variable at time t and receive antenna j.

3.3.2 Impact of Different Mappings on the Performance

In the following, we analyze the impact of different mappings on the transfer characteristics of the detector for an AWGN channel. The extension to a Rayleigh fading channel is straightforward. The mutual information between transmitted constellation symbol $s = f(\mathbf{c})$ and received AWGN channel output r is given (assuming that all constellation symbols are equiprobable) by

$$I(s;r) = \frac{1}{M} \sum_{n=1}^{M} \int_{-\infty}^{\infty} p(r|s=s_n) \times \log_2 \frac{p(r|s=s_n)}{p(r)} \, dr \tag{3.10}$$

with conditional probability density function (PDF)

$$p(r|s=s_n) = \frac{1}{\pi\sigma^2} \exp\left(-\frac{|r-s_n|^2}{\sigma^2} \right)$$

and

$$p(r) = \frac{1}{M} \sum_{n=1}^{M} p(r|s=s_n) \ ,$$

where $\log_2(\cdot)$ denotes the base 2 logarithm and σ^2 is the noise variance. With the chain rule of mutual information it can be shown that the mutual information can be decomposed into

$$I(s;r) = I(\mathbf{c};r) = I(c_1, \ldots, c_m; r) = \sum_{L=0}^{m-1} I_L^f \ , \tag{3.11}$$

where I_L^f is the average mutual information [tB00], when L bits are already known to the receiver. Note that the I_L^f depend on the mapping function $f(\cdot)$.

Example: For 8PSK (m=3) we have

$$\begin{aligned}
I_0^f &= \frac{1}{m\binom{m-1}{0}} \sum_{i=1}^{m} I^f(c_i;r) \\
I_1^f &= \frac{1}{m\binom{m-1}{1}} \sum_{i=1}^{m} \sum_{\substack{j=1,\\ j\neq i}}^{m} I^f(c_i;r|c_j) \\
I_2^f &= \frac{1}{m\binom{m-1}{2}} \sum_{i=1}^{m} \sum_{\substack{j=1,\\ j\neq i}}^{m} \sum_{\substack{k=j+1,\\ k\neq i}}^{m} I^f(c_i;r|c_j,c_k) \ .
\end{aligned}$$

Note that conditioning (i.e. increasing *a priori* knowledge) increases the mutual information, i.e. $I_L^f \geq I_{L-1}^f$. Further note, that the sum in (3.11) itself is independent of the mapping strategy $s = f(\mathbf{c})$, but that is not the case for the addends. Depending on the constellation $\mathcal{C}$ in use, there are in principal $|\mathcal{C}|!$ different mapping strategies. The question here is how to find a mapping strategy with a good performance in iterative decoding? Before answering the question, let us introduce the subsets $\mathcal{G}_i^+$ and $\mathcal{G}_i^-$, where $\mathcal{G}_i^+ = \{s : c_i = 0\}$ is the set of symbols such that $c_i = 0$ and $\mathcal{G}_i^- = \{s : c_i = 1\}$ is the set of symbols such that $c_i = 1$ for any i with $1 \leq i \leq m$. In order to find some criteria for the optimal mapping strategy with respect to iterative decoding, we are going to analyze the impact of the mapping on I_0^f and I_{m-1}^f from (3.11), which are the most important addends in the sum.

I_0^f: Let us start with I_0^f, which consists of the weighted sum of mutual information $I^f(c_i; r)$

$$I_0^f(c_i; r) = \frac{1}{2}\left(I^f(c_i = 0; r) + I^f(c_i = 1; r)\right) \ , 1 \leq i \leq m$$

and $I^f(c_i = 0; r)$ (and similar for $I^f(c_i = 1; r)$) is given as

$$I^f(c_i = 0; r) = \int_{-\infty}^{\infty} p_i^+(r) \log_2 \left(\frac{p_i^+(r)}{p(r)}\right) \, dr \ , \tag{3.12}$$

where

$$p_i^+(r) = \frac{2}{M} \sum_{\mathcal{G}_i^+} \frac{1}{\pi\sigma^2} e^{-\frac{|r-s|^2}{\sigma^2}} \ .$$

With $\log(x) \approx x - 1$, x around one, and

$$p_i^+(r)^2/p(r) \approx M\pi\sigma^2 p_i^+(r)^2 \ , \tag{3.13}$$

after some manipulations we arrive at

$$I^f(c_i = 0; r) \approx \frac{1}{2\log(2)} \left(1 + \sum_{s_l \in \mathcal{G}_i^+} \sum_{s_k \in \mathcal{G}_i^+ \setminus \{s_l\}} e^{-\frac{1}{2}\frac{|s_l - s_k|^2}{\sigma^2}}\right) \tag{3.14}$$

The approximation in (3.13) is allowed, since (3.12) depends strongly on $p_i^+(r)$.

I_{m-1}^f: Similar to I_0^f, I_{m-1}^f consists of a weighted sum of different conditioned mutual information $I_{m-1}^f(c_i; r|\mathbf{c} \setminus \{c_i\})$, which are given as

$$\begin{aligned} I_{m-1}^f(c_i; r|\mathbf{c} \setminus \{c_i\}) &= \frac{1}{2}\big(I^f(c_i = 0; r|\mathbf{c} \setminus \{c_i\}) + I^f(c_i = 1; r|\mathbf{c} \setminus \{c_i\})\big) \\ &= \frac{1}{2}\int_{-\infty}^{\infty} p_{s^+}(r) \log_2 \left(\frac{p_{s^+}(r)}{p_\pm(r)}\right) \, dr + \frac{1}{2}\int_{-\infty}^{\infty} p_{s^-}(r) \log_2 \left(\frac{p_{s^-}(r)}{p_\pm(r)}\right) \, dr \ , \end{aligned}$$

where $p_{s^+}(r)$ is defined (and similar for $p_{s^-}(r)$) as

$$p_{s^+}(r) = \frac{1}{\pi\sigma^2} e^{-\frac{|r-s^+|^2}{\sigma^2}} \ ,$$

where $s^+ = f(c_1, \ldots, c_i = 0, \ldots, c_m)$ (and similar for s^-) and $p_\pm(r)$ is defined as

$$p_\pm(r) = \frac{1}{2}\left(p_{s^+}(r) + p_{s^-}(r)\right)$$

Again, with $\log(x) \approx x - 1$ and after some manipulations, we arrive at

$$I_{m-1}^{f}(c_i; r|\mathbf{c} \setminus \{c_i\}) \approx -\frac{2\log(\pi\sigma^2)}{\log(2)} - \frac{1}{2}\frac{1 + e^{-\frac{1}{2}\frac{|s^+ - s^-|^2}{\sigma^2}}}{\log(2)\pi\sigma^2} \tag{3.15}$$

From (3.14) and (3.15), we can now derive two important mapping strategy criteria. The first criterion is with respect to I_{m-1}^{f}, say the right most point in the EXIT-charts . In order to achieve low BER with iterative decoding, I_{m-1}^{f} should be as high as possible. Therefore, in order to achieve a high I_{m-1}^{f} (see (3.15)), the first criteria is to get constellation points, which differ in only one symbol bit, as far away as possible from each other.

A highly important fact is that the transfer characteristics of mapping devices are almost straight lines and that the areas under the characteristics are equal. The only difference between different mapping strategies is the slope of the transfer characteristic. Therefore, increasing the right most point of a transfer characteristic results in a decreasing I_0^f, the left most point in the EXIT-chart. The second criterion is with respect to I_0^f. In order to further reduce I_0^f (see (3.14)) and automatically enhance I_{m-1}^{f}, constellation points, which have a symbol bit in common, should again be as far away as possible from each other.

Example: In Table 3.1 we show the mapping function $f(\mathbf{c})$ for five different mappings, which we choose from all possible 8PSK mappings. The constellation positions given in Table 3.1 are depicted in Fig. 3.7 for illustration. Note that the mapping strategies "d23" and anti Gray best fulfil the criteria which we mentioned above. Comparing the information transfer for different mappings in Table 3.2, we

8PSK Mappings					constellation
Gray	natural	"d21"	"d23"	anti Gray	position
000	000	000	000	000	1
001	001	011	011	111	2
011	010	101	101	001	3
010	011	110	110	110	4
110	100	111	001	011	5
111	101	001	010	100	6
101	110	010	100	010	7
100	111	100	111	101	8

Table 3.1: $f(\mathbf{c})$ for 5 different mappings with 8PSK

8PSK mappings	I_L			$\sum I_L = I(s;r)$
	I_0	I_1	I_2	
Gray	0.7805	0.7819	0.7830	2.345
natural	0.6369	0.8265	0.8819	2.345
"d21"	0.6321	0.7736	0.9395	2.345
"d23"	0.5380	0.8182	0.9889	2.345
anti Gray	0.4933	0.8723	0.9796	2.345

Table 3.2: Conditional mutual information I_L for different 8PSK mappings at $E_b/N_0 = 6$ dB

see that for Gray mapping the difference between I_2^f and I_0^f is not as large as in the case of the other mappings. This means that increasing *a priori* knowledge has

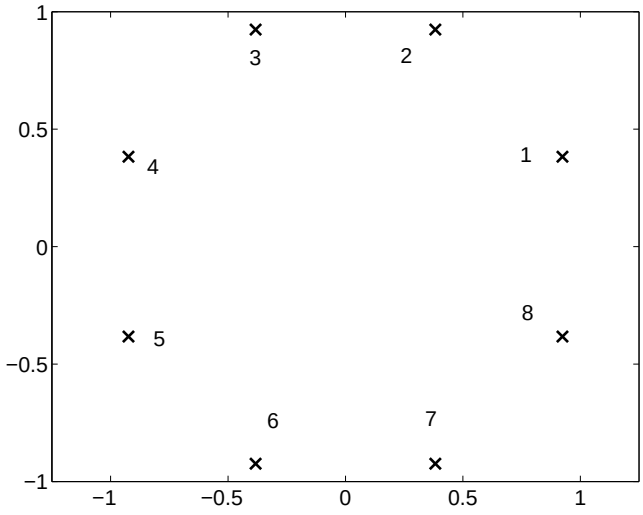

Figure 3.7: 8PSK-constellation

only a small impact on the information transfer of the detector for Gray mapping. Interestingly, this is not true for the other mappings, where we observe an improved information transfer by increasing the *a priori* knowledge. Additionally, I_0^f is larger with Gray mapping in comparison to the other mappings. Therefore, we expect a better performance with Gray mapping in the case, where no *a priori* knowledge is available at the receiver.

3.3.3 Iterative Detection and Decoding

Fig. 3.8 shows the structure of the receiver which consists of two stages: the Space-Time Detector (STD) described in the following subsection 3.3.3.1, and a Maximum A Posteriori (MAP) Decoder [BCJR74]. The two stages are separated by deinterleavers and interleavers. The receiver works as follows: In the first iteration, the switch in Fig. 3.8 is in position 1; assuming equally likely bits, the resulting *a priori* information $\lambda_{A,STD}$ is zero. Now, the Space-Time Detector has to compute the extrinsic information $\lambda_{E,STD}$ only from the observations of the channel output. The extrinsic information $\lambda_{E,STD}$ is deinterleaved and fed into the MAP decoder as *a priori* information $\lambda_{A,MAP}$. Based on this *a priori* information and the trellis representation of the channel code [HOP96], the MAP-Decoder computes the extrinsic information $\lambda_{E,MAP}$. After interleaving, this extrinsic information $\lambda_{E,MAP}$ is fed back to the STD as *a priori* information $\lambda_{A,STD}$ for the following iterations. After the first iteration, the switch in Fig. 3.8 is in position 2.

3.3.3.1 Space-Time Detector with Soft Outputs

In the following, we focus, without loss of generality, on the first p consecutive channel uses. The space-time detector at the receiver computes the log-likelihood ratios of the coded bits $\{c_k\}_{k=1}^{qm}$ corresponding to the transmitted sequence $\mathbf{s} =$

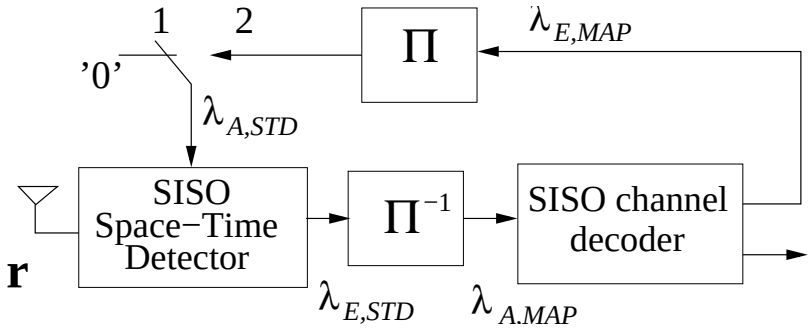

Figure 3.8: Model of the SISO receiver with SISO Space-Time Detector, SISO channel decoder, interleaver and deinterleaver.

$\{s_1 \ldots, s_q\}$ from the received sequence $\{\mathbf{r}_1, \ldots, \mathbf{r}_p\}$, where $\mathbf{r}_t = [r_t^1, \ldots, r_t^{n_R}]$, with

$$\Lambda_{\mathrm{STD}}(c_k) = \log \frac{\sum\limits_{\mathbf{s}\in\mathcal{S}^+} P[\mathbf{r}_1, \ldots, \mathbf{r}_p|\mathbf{s}] \cdot P[\mathbf{s}]}{\sum\limits_{\mathbf{s}\in\mathcal{S}^-} P[\mathbf{r}_1, \ldots, \mathbf{r}_p|\mathbf{s}] \cdot P[\mathbf{s}]} \ . \tag{3.16}$$

In (3.16), $\log(\cdot)$ denotes the natural logarithm, $\mathcal{S}^+ = \{\mathbf{s} : s_i = f(\mathbf{c}_i), c_k = 0\}$ is the set of transmitted sequences such that $c_k = 0$ and $\mathcal{S}^- = \{\mathbf{s} : s_i = f(\mathbf{c}_i), c_k = 1\}$ is the set of transmitted sequences such that $c_k = 1$. After some manipulations we obtain

$$\Lambda_{\mathrm{STD}}(c_k) = \log \frac{\sum\limits_{\mathbf{s}\in\mathcal{S}^+} \exp\left(-\frac{1}{\sigma^2} \sum\limits_{t=1}^{p} \sum\limits_{j=1}^{n_R} \left| r_t^j - \sum\limits_{i=1}^{n_T} h_{i,j} s_t^i \right|^2\right) P[\mathbf{s}]}{\sum\limits_{\mathbf{s}\in\mathcal{S}^-} \exp\left(-\frac{1}{\sigma^2} \sum\limits_{t=1}^{p} \sum\limits_{j=1}^{n_R} \left| r_t^j - \sum\limits_{i=1}^{n_T} h_{i,j} s_t^i \right|^2\right) P[\mathbf{s}]} \ .$$

Since we have a sufficiently long pseudo-random bit interleaver, we may assume that the probabilities $P[c_l]$ of the coded bits c_l are independent. Thus, we have

$$P[\mathbf{s}] = \prod_{l=1}^{qm} P[c_l] \ .$$

Hence,

$$\Lambda_{\mathrm{STD}}(c_k) = \log \frac{\sum\limits_{\mathbf{s}\in\mathcal{S}^+} \exp\left(\frac{-\sum\limits_{t=1}^{p} \sum\limits_{j=1}^{n_R} r_t^j - \sum\limits_{i=1}^{n_T} h_{i,j} s_t^i}{\sigma^2}^2\right) \prod\limits_{l=1}^{qm} P[c_l]}{\sum\limits_{\mathbf{s}\in\mathcal{S}^-} \exp\left(\frac{-\sum\limits_{t=1}^{p} \sum\limits_{j=1}^{n_R} r_t^j - \sum\limits_{i=1}^{n_T} h_{i,j} s_t^i}{\sigma^2}^2\right) \prod\limits_{l=1}^{qm} P[c_l]} \tag{3.17}$$

holds. The probabilities $P[c_k = 0]$ and $P[c_k = 1]$ are present in every term of the summation in the numerator and denominator, respectively. Therefore, we can

write (3.17) as

$$\Lambda_{\mathrm{STD}}(c_k) = \underbrace{\log \frac{\sum\limits_{\mathbf{s}\in\mathbb{S}^+} \exp\left(-\frac{\sum\limits_{t=1}^{p}\sum\limits_{j=1}^{n_R} r_t^j - \sum\limits_{i=1}^{n_T} h_{i,j}s_t^i}{\sigma^2}\right)^2 \prod\limits_{\substack{l=1,\\ l\neq k}}^{qm} P[c_l]}{\sum\limits_{\mathbf{s}\in\mathbb{S}^-} \exp\left(-\frac{\sum\limits_{t=1}^{p}\sum\limits_{j=1}^{n_R} r_t^j - \sum\limits_{i=1}^{n_T} h_{i,j}s_t^i}{\sigma^2}\right)^2 \prod\limits_{\substack{l=1,\\ l\neq k}}^{qm} P[c_l]}}_{\text{extrinsic information } \lambda_{E,STD}} + \underbrace{\log \frac{P[c_k=0]}{P[c_k=1]}}_{a\ priori \text{ information } \lambda_{A,STD}} . \tag{3.18}$$

The log-likelihood ratio Λ_{STD} for c_k is now decomposed into two parts,

$$\Lambda_{\mathrm{STD}} = \lambda_{E,STD} + \lambda_{A,STD} ,$$

where $\lambda_{A,STD}$ is the *a priori* information obtained through the iterative decoding process and $\lambda_{E,STD}$ is obtained from the observations of the channel output.

Remark 3.3.1. *Since the columns of the generator matrix of a STBC are orthogonal to each other, there is no transfer of extrinsic information between symbols (inter-symbol-extrinsic-information) within a STBC.*

Interestingly, using the orthogonality of the columns, we can decompose (3.18) into q parts, where each part is only a function of s_i with $i = 1 \ldots q$, given as

$$\Lambda_{\mathrm{STD}}(c_k) = \log \frac{\sum\limits_{s_i\in\mathbb{S}_i^+} \exp\left(-\frac{|\tilde{r}_i - s_i|^2 + \tilde{h}|s_i|^2}{\sigma^2}\right) \prod\limits_{\substack{l=1,\\ l\neq k}}^{m} P[c_l]}{\sum\limits_{s_i\in\mathbb{S}_i^-} \exp\left(-\frac{|\tilde{r}_i - s_i|^2 + \tilde{h}|s_i|^2}{\sigma^2}\right) \prod\limits_{\substack{l=1,\\ l\neq k}}^{m} P[c_l]} + \log \frac{P[c_k=0]}{P[c_k=1]} \tag{3.19}$$

$$\text{for } (i-1)m < k \leq im ,$$

where $\tilde{h}$ and $\tilde{r}_1, \cdots, \tilde{r}_q$ are given in [TJC99b, pp.458], $\mathbb{S}_i^+ = \{s_i : c_k = 0\}$ is the set of space-time symbols such that $c_k = 0$ and $\mathbb{S}_i^- = \{s_i : c_k = 1\}$ is the set of space-time symbols such that $c_k = 1$ for any k with $(i-1)m < k \leq im$. This dramatically reduces the complexity of the SISO space-time detector.

The code bit probabilities $P[c_k]$ in (3.19) can be expressed in terms of their *a priori* information as

$$P[c_k] = \begin{cases} \frac{1}{1+\exp(+\lambda_{A,STD})} & \text{for } c_k = 1 \\ \frac{1}{1+\exp(-\lambda_{A,STD})} & \text{for } c_k = 0 . \end{cases} \tag{3.20}$$

It is possible to simplify the computational complexity for Λ_{STD} in (3.19), by using the following approximation

$$\Lambda_{\mathrm{STD}}(c_k) \approx \max_{s_i\in\mathbb{S}_i^+} \log P[\mathbf{r}_1, \ldots, \mathbf{r}_p|s_i]P[s_i] - \max_{s_i\in\mathbb{S}_i^-} \log P[\mathbf{r}_1, \ldots, \mathbf{r}_p|s_i]P[s_i] , \tag{3.21}$$

which we use in our simulations in the next section for numerical stability. In order to perform the soft decoding at the next stage of the receiver (i.e. the MAP decoder), we regard these log-likelihoods as the observations from a BPSK modulation over an additive white Gaussian noise channel as described in [Dum98].

3.3.4 EXIT-Chart Analysis

An EXIT chart consists of a pair of curves which represent the mutual information transfer functions or transfer characteristics of the component detectors and decoders in the turbo process. Each of the transfer characteristics is essentially a plot of the a priori mutual information I_A against the extrinsic mutual information I_E for the component decoder of interest. The terms I_A and I_E are related to the probability density functions (pdfs) of the log-likelihood ratios λ_A for the a priori information and λ_E for the extrinsic information, the signal-to-noise ratio E_b/N_0 and the structure of the detector or decoder. The required pdfs can be estimated by generating histograms $p(\lambda_A)$ and $p(\lambda_E)$ of λ_A and λ_E respectively for a particular value of E_b/N_0.

Remark 3.3.2. *Notice that the EXIT-chart analysis assumes that the log-likelihood ratios of the bits are Gaussian distributed. Although that does not hold exactly for the detector considered here, it was shown in [tB01] that the shape of the involved distributions is only of minor importance for the EXIT-chart analysis. This justifies the application of EXIT-charts for our case.*

Fig. 3.9 shows the extrinsic information transfer characteristics of the extended block code BCH$(8, 4)$ and the convolutional code CC$(23, 35)$ with outer rate $R_{out} = 1/2$. Additionally, the curves of the space-time detector for $\mathcal{H}_3$ as the inner decoder

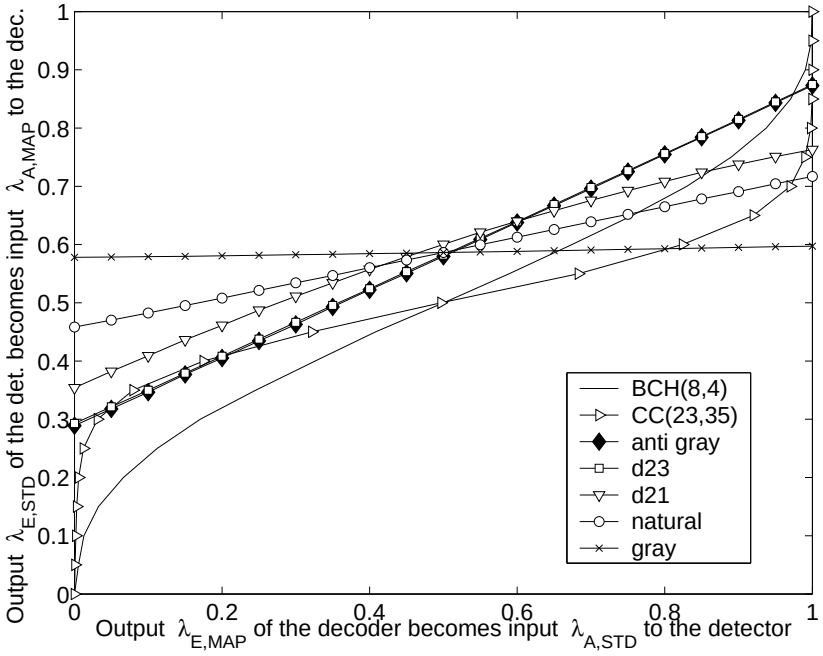

Figure 3.9: Extrinsic information transfer (EXIT) charts of outer rate $R_{out} = 1/2$ decoder and transfer characteristics of inner detector with different mappings.

for different mapping strategies are depicted. The results for other OSTBC such as the Alamouti code [Ala98], $\mathcal{G}_3$, $\mathcal{G}_4$ and $\mathcal{H}_4$ [TJC99a] are similar and therefore omitted. Different mappings result in transfer characteristics of different slopes. It is important to know that for the EXIT chart predictions on code performance we assumed very large interleavers and a fast fading channel, in which the channel is

selected independently for each space-time code matrix (i.e. channel is constant for only p channel uses). However, in practice using large interleavers is not applicable. In our BER simulations later on in this section we use moderate interleaver sizes. Due to both of these assumptions the EXIT charts can be regarded only as asymptotic results. Therefore, the "turbo-cliff"-region does not occur exactly at the predicted SNR values. Note that the detector transfer characteristics are almost straight lines and that at low SNR values increasing the SNR just shifts the curve up. Additionally, note that for high SNR values, the slope of the detector transfer characteristics is also affected (not depicted here). From the figure, we see also that the natural mapping provides good extrinsic output at the beginning but provides diminishing output for higher *a priori* input λ_A. For the anti gray mapping it is the other way around. The detector which uses gray mapping provides almost the same extrinsic output for all *a priori* input, which confirms the results from Table 3.2 in section 3.3.2. Therefore, we expect that the performance of the detector with gray and natural mapping to be good in the low SNR regime and for a few iterations in comparison to the other mappings. But in the high SNR regime and for more iterations, we expect it to be the other way around. Note that the axes are swapped for the outer code: λ_A is on the ordinate, λ_E on the abscissa.

3.3.5 Simulation results

In this section, the simulation results of the proposed schemes and their interpretations are presented. For verification of the observations from the EXIT-Charts,

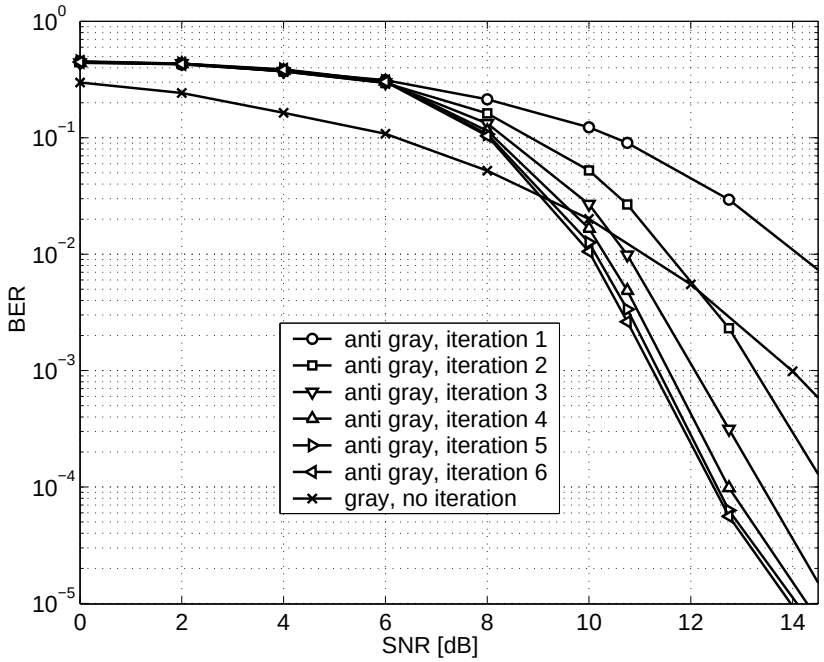

Figure 3.10: Performance of the considered outer code (extended BCH(8,4)) in concatenation with the OSTBC $\mathcal{H}_3$ with different mapping strategies, $n_T = 3$ transmit and $n_R = 1$ receive antennas, 8PSK modulation.

we present the bit error rates (BERs) for gray and anti gray mapping in Fig. 3.10 for a system employing the OSTBC $\mathcal{H}_3$ with $n_T = 3$ transmit and $n_R = 1$ receive antenna. Since the transfer characteristics of all other mapping strategies are

almost always between the transfer characteristics of gray and anti gray mapping, only their BERs are depicted. Coding is performed over multiple block fading channels. The transmitted bits are organized in frames of length $F_L = 432$. We assume that one block fading channel is constant for $\tau = 24$ channel uses. We observe that although the curves are relatively flat at iteration 1, there is a significant improvement with further iterations. However, in the case of gray mapping iteratively decoding yields negligible gain, so only the first iteration is plotted. Furthermore, simulation results show, that there occurs a saturation of the improvement for the other mappings beyond iteration 6. This behavior can be explained by the fact that after each iteration the extrinsic output of the receiver components tends to a Gaussian distribution according to the central limit theorem but the correlation between the extrinsic information and the channel output increases after each iteration [HOP96]. Hence, the improvement of the BER-performance diminishes after each iteration. As predicted in the EXIT-Charts, the BER-performance of the gray mapping is better than the BER-performance of the anti gray mapping in the low SNR regime. In the high SNR regime, however, it is the other way around. Note that this performance difference to gray mapping can be further enhanced with a larger frame length or a smaller τ.

To avoid an overload of Fig. 3.10, the BERs for the uncoded system with gray mapping and coded system with anti gray and gray mapping are depicted in Fig. 3.11 for comparison. Furthermore, the cases where we have assumed that the STD has perfect a priori information are also depicted, to serve as lower bounds. From this figure, we see that there is a significant performance gain compared to the uncoded system. In addition to this, we observe that the performance of the scheme depends strongly on the accuracy of the a priori information in the case of anti Gray mapping, but not for Gray mapping. This also verifies our analysis from the EXIT-charts. The performance of the anti Gray scheme with the assumption of perfect a priori information is about 4 dB better at a BER of 10^{-2} than without this assumption. Note that this difference gets smaller for high SNR, since the accuracy of the a priori information in the non-perfect case gets even better for increasing SNR. By using a more powerful channel code such as the 16-state trellis code $\mathrm{CC}(23, 35)$, the performance with perfect a priori information is even better, as depicted in Fig. 3.12. However, since the transfer characteristics of the $\mathrm{CC}(23, 35)$ do not match as well as the BCH code to the transfer characteristics of the detector, the performance without perfect a priori information is worse in comparison to the BCH code for small SNR, due to the inferior exchange of a priori information between detector and decoder. For higher SNR and better feedback information, the performance of the $\mathrm{CC}(23, 35)$ is better than that of the BCH code. We also observe, that in case of the FER, unlike the case of BER, the performance with anti Gray mapping is always better than the Gray mapping strategy for the whole range of SNR values.

In this section, the serial concatenation of an outer code with different orthogonal space-time block codes has been analyzed. It was assumed, that the transmitter has no channel state information (CSI) whereas the receiver has (CSI). In some cases, this assumption is not valid, e.g., if the channel variations are considerably higher than one symbol duration, it is not possible to apply the OSTBC as inner codes within this "turbo" coded space-time transmission scheme. Then, it is more relevant to employ space-time transmission schemes, which do not have this constraint, as inner codes, which is performed in the following section.

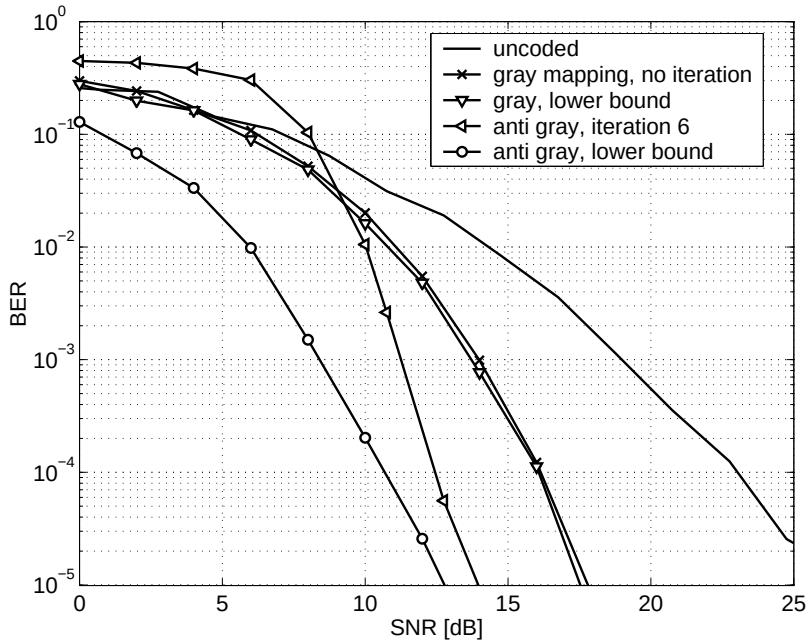

Figure 3.11: Performance comparison of the coded and uncoded system, OSTBC $\mathcal{H}_3$ $n_T = 3$ transmit and $n_R = 1$ receive antennas, 8PSK modulation.

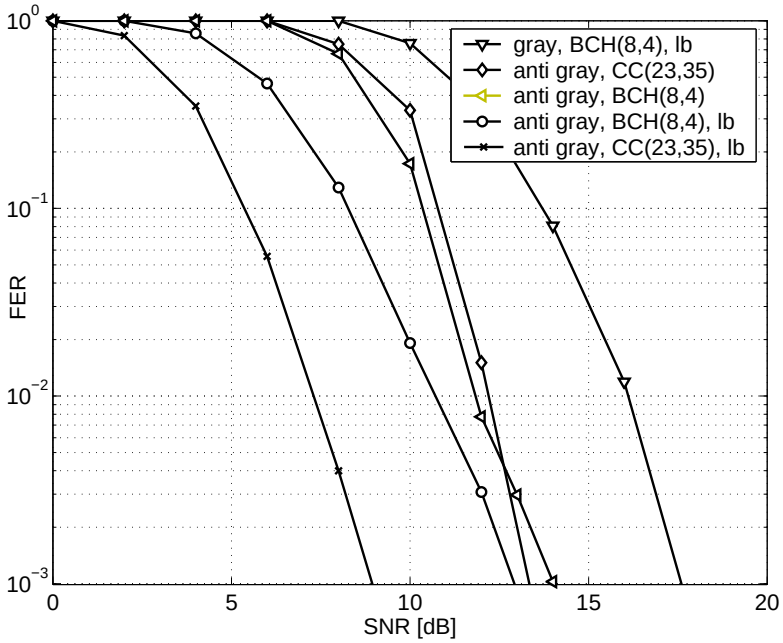

Figure 3.12: Frame error rate of the proposed scheme with different outer codes and mapping strategies, lb (lower bound) stands for assumption of perfect a priori information.

3.4 Turbo Decoding of Unitary STC

Code design for multiple-input-multiple-output (MIMO) channels has mostly concentrated on the case in which the receiver has perfect CSI. However, as argued in [HM00b], [HM00a], in the case when the coherence time is small, estimating the channel coefficients would reduce the spectral efficiency considerably. Furthermore, the assumption of perfect knowledge of the CSI may not be realistic with a large number of transmit antennas.

In this section, we consider the application of channel codes together with a unitary space-time modulation scheme, proposed in [HM00b, HM00a], in order to improve the performance of the uncoded system without CSI at both the receiver and the transmitter, and in order to achieve the capacity promised by the information theoretic results. Note that, there is also much work available for concatenated space-time coded systems with perfect CSI at the receiver, e.g. [LH02, LhHmW02] and references therein.

Interestingly, the serial concatenation of different channel codes (e.g., turbo codes) as outer encoders with a multiple-antenna-differential modulation scheme without CSI at the receiver was proposed very recently in [SPS02b, SG01]. In this section, we use a unitary space-time modulation scheme, instead of the differential modulation scheme in [SPS02b, SG01], as inner code.

In [BD02], the serial concatenation of a turbo code with the unitary space-time modulation scheme was proposed. In contrast to [BD02], here we apply block or convolutional codes as outer encoder instead of a turbo code. In addition to this, we perform an analysis of the detection and decoding components with extrinsic information transfer (EXIT) charts as in section 3.3. We use the major advantage of EXIT-charts; that only simulations of the components are needed to obtain the desired transfer charts in order to analyze the whole system. According to the transfer characteristics of the detector (i.e., the unitary space-time detector) we apply channel codes whose transfer characteristics match relatively well to the detector in order to get good performance results.

3.4.1 Transmitter structure

We combine the unitary space-time coded modulation scheme serially with the following channel block codes,

- an extended BCH(8,4) code,
- a single parity check code SPC(8,7),
- and a convolutional code, namely a $\mathrm{CC}(7,5)_{oct}$,

respectively. This is done in order to achieve low probability of error for small signal-to-noise-ratios. In our system model we used the unitary space-time code with $T = 8$ and $n_T = 2$ [HM00a, Table II], where T denotes the number of channel uses in which the channel is assumed to be constant. It is necessary that the number of transmit antennas is $n_T \leq T$.

After encoding with the channel code, and then interleaving, the F_L coded bits are mapped in sets $\{c_t^s\}_{t=1}^{T}$ $(1 \leq s \leq {}^{F_L}/_{T})$ onto the unitary space-time constellation, which consists of $L = 256$ complex valued $T \times n_T$ matrices $\mathbf{\Phi}_l$ $(1 \leq l \leq L)$. F_L denotes the frame length. The matrices have the property that $\mathbf{\Phi}_l^H \mathbf{\Phi}_l = \mathbf{I}_{n_T}$ and Φ_l is isotropically distributed. It is worth knowing that the orientation of the subspace,

spanned by the n_T columns of $\mathbf{\Phi}_l$, carry the message information. The $(T \times n_T)$ complex transmit matrix $\mathbf{G}_{n_T}$ is given by $\mathbf{G}_{n_T} = \sqrt{T}\mathbf{\Phi}$.

Furthermore, we assume that neither the transmitter nor the receiver has knowledge of the channel. The receiver works in a similar way to the receiver shown in section 3.3, i.e. at the receiver, the unitary space-time detector (STD) computes the log-likelihood ratios of the transmitted coded bits c_t with

$$\Lambda_{\mathrm{STD}}(c_t) = \log \frac{\sum\limits_{\mathbf{G}^+_{n_T}:\mathbf{S}=f(c),c_t=0} P[\mathbf{Y}|\mathbf{G}_{n_T}]P[\mathbf{G}_{n_T}]}{\sum\limits_{\mathbf{G}^-_{n_T}:\mathbf{G}_{n_T}=f(c),c_t=1} P[\mathbf{Y}|\mathbf{G}_{n_T}]P[\mathbf{G}_{n_T}]} \tag{3.22}$$

where $\log(\cdot)$ denotes the natural logarithm, , $\mathbf{Y}$ is the $(T \times n_R)$ complex receive matrix, $\mathbf{G}^+_{n_T}$ is the set of modulation matrices containing the index of code words such that $c_t = 0$ and $\mathbf{G}^-_{n_T}$ is the set of modulation matrices containing the index of code words such that $c_t = 1$. The conditional probability of $\mathbf{Y}$ given that $\mathbf{G}_{n_T}$ is transmitted in (3.22) is given by

$$p(\mathbf{Y}|\mathbf{G}_{n_T}) = \frac{\exp\left(-\mathrm{tr}\left\{\mathbf{\Theta}^{-1}\mathbf{Y}\mathbf{Y}^H\right\}\right)}{\pi^{Tn_R}\det^{n_R}\mathbf{\Theta}}$$

where the $T \times T$ covariance matrix $\mathbf{\Theta}$ of the columns of $\mathbf{Y}$ is defined as $\mathbf{\Theta} = \mathbf{I}_T + (\rho/n_T)\,\mathbf{G}_{n_T}\mathbf{G}^H_{n_T}$. After some manipulations we arrive at

$$\Lambda_{\mathrm{STD}}(c_t) = \log \underbrace{\frac{\sum\limits_{\substack{\mathbf{\Phi}^+:\mathbf{\Phi}=1/\sqrt{T}\mathbf{G}_{n_T},\\ \mathbf{G}_{n_T}=f(c),c_t=0}} \exp\left(\mathrm{tr}\left\{\frac{\mathbf{Y}^H\mathbf{\Phi}\mathbf{\Phi}^H\mathbf{Y}}{1+\frac{n_T}{\rho T}}\right\}\right)\prod\limits_{\substack{l=1,\\ l\neq t}}^{\log_2(L)} P[c_l]}{\sum\limits_{\substack{\mathbf{\Phi}^-:\mathbf{\Phi}=1/\sqrt{T}\mathbf{G}_{n_T},\\ \mathbf{G}_{n_T}=f(c),c_t=1}} \exp\left(\mathrm{tr}\left\{\frac{\mathbf{Y}^H\mathbf{\Phi}\mathbf{\Phi}^H\mathbf{Y}}{1+\frac{n_T}{\rho T}}\right\}\right)\prod\limits_{\substack{l=1,\\ l\neq t}}^{\log_2(L)} P[c_l]}}_{\text{extrinsic information } \lambda_{E,STD}} + \underbrace{\log\frac{P[c_t=0]}{P[c_t=1]}}_{\text{from a priori information } \lambda_{A,STD}}. \tag{3.23}$$

where $\log_2(\cdot)$ denotes the base two logarithm, $\mathbf{\Phi}^+$ and $\mathbf{\Phi}^-$ are defined analogously to $\mathbf{G}^+_{n_T}$ and $\mathbf{G}^-_{n_T}$, respectively. Some simulation results of the proposed scheme and their interpretation are presented in the following section.

3.4.2 Simulation results

Fig. 3.13 shows the transfer characteristics of the block and convolutional code with outer rate $R_{out} = 1/2$ and for the single parity check code with outer rate $R_{out} = 7/8$. For the convolutional code the generator polynomials are given in octal numbers with feedback polynomial $G_r = 7$ and feedforward polynomial $G = 5$. In addition to this, the curves of the detector as the inner decoder are depicted in Fig. 3.13. The detector curves vary for SNRs from 4 dB to 6 dB in steps of 0.5 dB. It can be noted that the detector transfer characteristics are almost straight lines, and increasing the SNR just shifts the curve up. Note that, for high SNR (e.g., 9 dB, as depicted in Fig. 3.13) values, the slope of the detector transfer characteristics is also affected. From the figure, we see also that the extended BCH(8,4) code provides

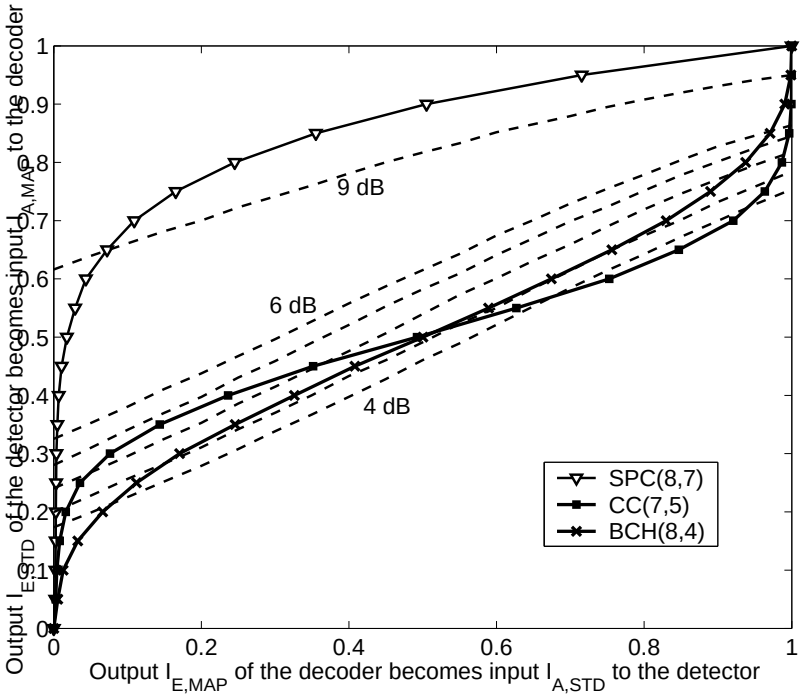

Figure 3.13: Extrinsic information transfer (EXIT) charts of decoders and transfer characteristics of inner detector (dashed lines).

good extrinsic output at the beginning but provides diminishing output for higher a priori input λ_A. For the $\mathrm{CC}(7,5)_{oct}$ code it is the other way around. Furthermore, we expect the "turbo-cliff" to occur near 4.5 dB for the extended BCH(8,4) outer code and near 5.5 dB for the $\mathrm{CC}(7,5)_{oct}$ outer code. For $E_b/N_0 = 4$ dB, the iterative decoding does not improve after the first iteration since the detector and decoder characteristics intersect at low a priori information. At $E_b/N_0 = 5$ dB, the transfer characteristic of the detector has been raised high enough to open a narrow tunnel ("bottleneck") between the detector transfer characteristic and the transfer characteristic of the extended BCH(8,4) code, thereby achieving an improvement of the bit error rates after a few iterations. The opening of a narrow tunnel between the transfer characteristic of the detector and the $\mathrm{CC}(7,5)_{oct}$ is at some higher SNR. Note that the axes are swapped for the outer codes: λ_A is on the ordinate, λ_E on the abscissa.

For verification of the observations from the EXIT-Chart, we present the BER for a system with $n_T = 2$ transmit and $n_R = 1$ receive antenna in Fig. 3.14. We assume that the channel is constant for $T = 8$ channel uses. The transmitted bits are organized in frames of length $F_L = 512$ including tail bits in the case of the convolutional code. We observe, that although for both block and convolutional code the curves at iteration 1 are relatively flat, there is a significant improvement of the curves at iteration 2 with a saturation of the improvement beyond iteration 3. This behavior can be explained by the fact that after each iteration the extrinsic output of the receiver components tends to a Gaussian distribution according to the central limit theorem but the correlation between the extrinsic information and the channel output increases after each iteration. Hence, the improvement of the BER-performance diminishes after each iteration. As predicted in the EXIT-Chart the BER-performance of the block code is better than the BER-performance of the convolutional code in the low SNR regime. In the high SNR regime, however, it is

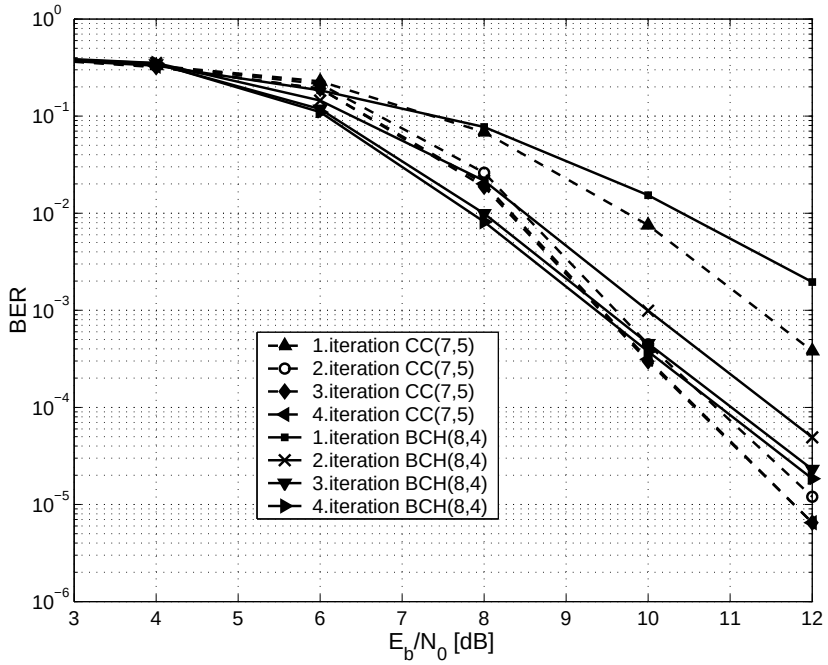

Figure 3.14: Performance of the considered outer codes [$\mathrm{CC}(7,5)_{oct}$ dashed lines; extended BCH(8,4) solid lines] in concatenation with the unitary space-time modulation scheme.

the other way around.

To avoid an overload of Fig. 3.14, the BERs for the uncoded (with rate $R = 1$) and SPC(8,7) coded system (with outer rate $R_{out} = 7/8$) after 4 iterations are depicted in Fig. 3.15 for comparison. From this figure, we see that there is a significant performance gain compared to the uncoded system. Also of interest is the comparison of the performance of our scheme with the turbo-coded unitary space-time modulation scheme in [BD02] with $R_{out} = 7/8$. From the simulations we observe that our scheme performs equally well or better, depending on the interleaver size used in [BD02]. However, note that the decoding complexity is smaller in our scheme. Furthermore, delay requirements of our scheme are lower, because we need only four iterations instead of eight iterations in [BD02].

Until now, in order to iteratively decode the concatenation of the inner space-time code and the outer channel code, we applied the optimum maximum a posteriori (MAP) algorithm. However, for processing purposes it may be necessary for some inner space-time codes to have access to the soft-outputs of the data bits immediately and without any delays, which is not possible due to the availability of the soft-outputs only at the end of the decoding step with the optimum algorithm. One such case is, e.g., the space-time code analyzed in the following section. In contrast to the OSTBC and the USTM scheme in this section, the STC in the following section aims at high transmission rates in order to achieve high gains in capacity that can be achieved through the use of multiple antennas at the transmitter and the receiver.

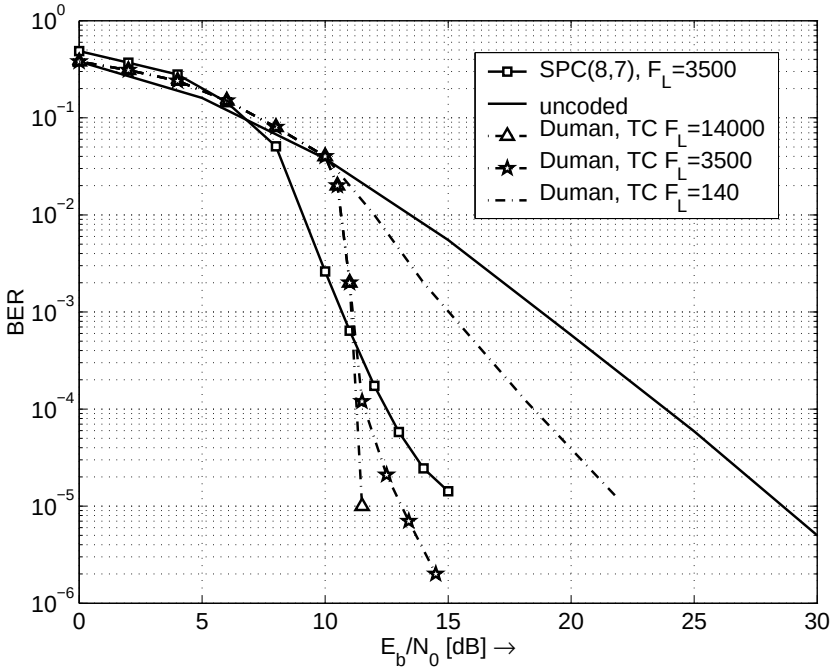

Figure 3.15: Performance comparison of the system with SPC(8,7), TC from [9] and uncoded system.

3.5 Turbo Decoding of Wrapped STC

VBLAST is a bandwidth efficient approach to wireless communication which takes advantage of the spatial dimension by transmitting and detecting several independent co-channel data streams using multiple antennas. However, due to the effect of error propagation caused by wrong estimations of transmitted signals the error rate performance is largely dependent on the substream which is detected first and which has the worst performance. To avoid this drawback, many iterative schemes with high complexity have been proposed in the literature, e.g. [SH00, GAH01].

Another approach to reduce the effect of error propagation and to improve the performance significantly was proposed by Caire et.al in [CC03]. They proposed a low-complexity space-time scheme called Wrapped Space-Time Coding (WSTC) for Rayleigh fading channels to achieve high spectral efficiencies [CC03]. In this scheme, only a single encoder is used. The coded data is diagonally interleaved and transmitted over the n_T transmit antennas. At the receiver, the nulling and cancelation steps are integrated into a Viterbi algorithm employing per-survivor-processing [RPT95]. In this work, we apply the coding and decoding technique from [CC03] as inner coding and decoding components, respectively. However, instead of the receiver used in [CC03], which provides only hard decisions over the information bits, we employ a SOVA providing soft decisions to the outer decoder. Since we need the data estimations at the inner decoder for the interference cancelation with a minimum amount of delay, the application of optimum soft-input-soft-output (SISO) MAP algorithms [BCJR74] is not feasible. Further on, we apply an outer code at the transmitter and couple this with an iterative decoding process in order to improve the performance of the architecture and in order to achieve the capacity promised by the information theoretic results. The performance of our scheme is evaluated by simulations and compared to the scheme proposed in [CC03].

3.5.1 Transmitter structure

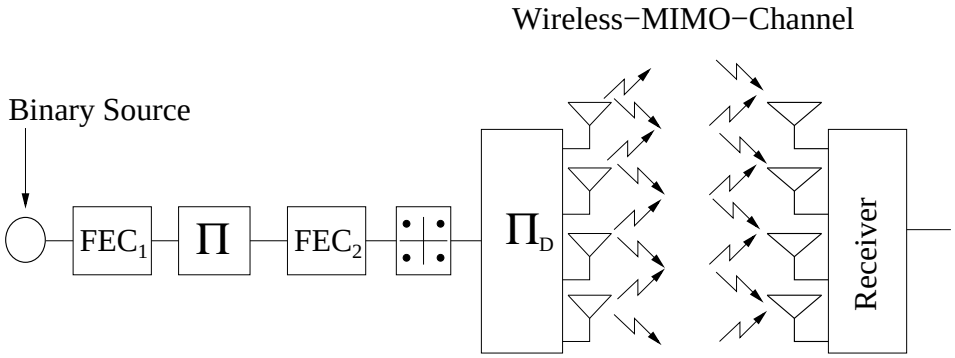

Figure 3.16: System model with binary source, SCCC encoder, Modulation, Diagonal Interleaving (cf. Fig. 3.17), Rayleigh MIMO-Channel and Receiver.

The transmitter consists of two systematic, recursive convolutional codes (CC), denoted as FEC_1 and FEC_2 in Fig. 3.16, concatenated in serial via a pseudo-random interleaver. Codes of such a structure are known as SCCC (Serially concatenated convolutional codes). This interleaver is used in order to uncorrelate the log-likelihoods of adjacent bits and distribute the error events due to a deeply faded block during a transmission. We can obtain different spectral efficiencies by puncturing the parity bits of the component encoders. After encoding the whole information bit sequence with FEC_1 and interleaving, the coded bit sequence is divided into τ blocks. Each block is encoded separately. Let a block of coded bits be $\{c_1, c_2, \ldots, c_{L_B}\}$, where L_B is the block length. These coded bits are then mapped onto symbols from a given constellation, e.g. BPSK, and interleaved via a diagonal (channel) interleaver, which is different from the interleaver used for concatenation of the SCCC code component encoders.

It is further assumed that the transmitter has no CSI and the receiver has perfect CSI. To obtain the transmit matrix $\mathbf{X}$ we use a special interleaver as illustrated in Fig. 3.17.

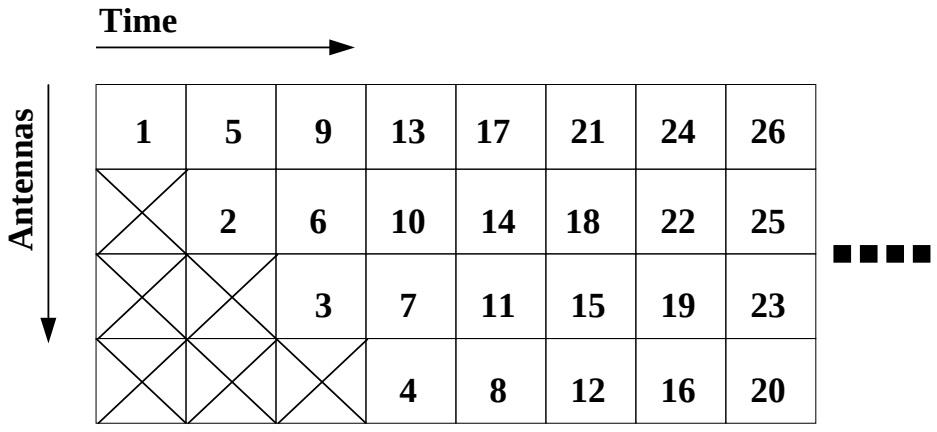

Figure 3.17: Diagonal interleaver for a system with $n_T = 4$ transmit antennas. The entries in the cells indicate the index k of the symbols in the current codeword. A cross in a cell means that at this time the given antenna is not active.

With the index k of the symbols in the current codeword, we get the right cell

position in $\mathbf{X}$ as follows

$$\begin{aligned} r &= k - \left\lfloor \frac{k-1}{n_T} \right\rfloor n_T \\ c &= k - \left\lfloor \frac{k-1}{n_T} \right\rfloor (n_T - 1), \end{aligned}$$

where $[\mathbf{X}]_{r,c}$ is the current cell of $\mathbf{X}$. Herein c corresponds to the column and r to the row of $\mathbf{X}$.

3.5.2 Receiver with Iterative decoding

3.5.2.1 Receiver structure

The structure of the receiver of this scheme is similar to the receiver structures presented in the last sections with the difference that the inner decoder is a space-time SOVA (STS) decoder which is described in the following subsection 3.5.2.2. At the receiver decoding is done by processing the receive matrix according to the diagonal interleaver structure. In Fig. 3.18, the variable $\widehat{\mathbf{X}}$ represents the estimated transmit matrix with entries up to the current decoding step k in the trellis diagram of the SOVA, obtained from the survivor terminating in the code trellis state τ at decoding step k. According to the ZF (Zero-Forcing) or the MMSE (Minimum

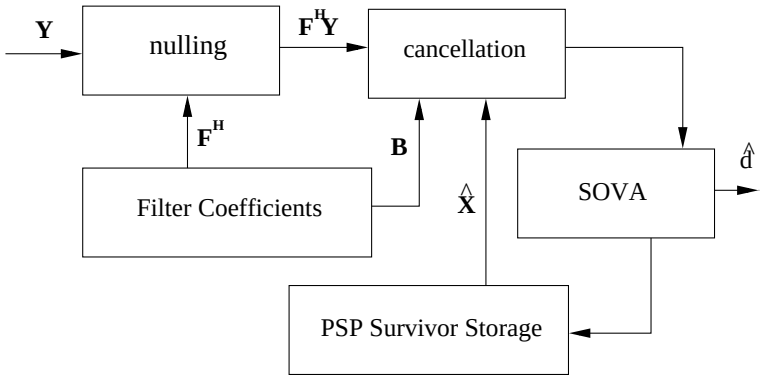

Figure 3.18: Model of the proposed receiver with per-survivor processing (PSP), joint decoding and channel estimating.

Mean-Square Error) criteria for the filter design [BTT02a, CC03], we have for the feedforward (or interference nulling) filter $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_{n_T}]$

$$\mathbf{f}_r = \begin{cases} \mathbf{q}_r & \text{(ZF)} \\ \frac{1}{\sqrt{\mathbf{h}_r^H \mathbf{S}_r^{-1} \mathbf{h}_r}} \mathbf{S}_r^{-1} \mathbf{h}_r & \text{(MMSE)} \end{cases},$$

where

$$\mathbf{S}_r = (\mathbf{H}\boldsymbol{\Gamma})\,(\mathbf{H}\boldsymbol{\Gamma})^H + (n_T/\rho)\mathbf{I}_{n_r} = \sum_{i=1}^{r-1} \mathbf{h}_i \mathbf{h}_i^H + (n_T/\rho)\mathbf{I}_{n_r} \tag{3.24}$$

and

$$\boldsymbol{\Gamma} = \begin{bmatrix} \mathbf{I}_{\mathrm{r}-1} \\ \mathbf{0}_{(\mathrm{n_T}-\mathrm{r}+1)\times(\mathrm{r}-1)} \end{bmatrix}.$$

The vector $\mathbf{q}_r$ is obtained from the **QR**-factorization of the channel matrix $\mathbf{H}$, where the matrix $\mathbf{R}$ is an upper triangular $n_T \times n_T$ matrix, and $\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_{n_T}]$ is an $n_R \times n_T$ unitary matrix with $\mathbf{Q}^H\mathbf{Q} = \mathbf{I}$. The superscript $(\cdot)^H$ denotes matrix conjugate transpose. For the feedback (interference cancelation) filter $\mathbf{B} = [\mathbf{b}_1^T, \mathbf{b}_2^T, \ldots, \mathbf{b}_{n_T}^T]^T$, $\mathbf{b}_r^T = [\mathbf{0}_{1\times r}, \tilde{\mathbf{b}}_r^T]$ we have

$$\tilde{\mathbf{b}}_r^T = \mathbf{f}_r^H \mathbf{H} \begin{bmatrix} \mathbf{0}_{\mathrm{r}\times(\mathrm{n_T}-\mathrm{r})} \\ \mathbf{I}_{\mathrm{n_T}-\mathrm{r}} \end{bmatrix} .$$

Note that we have to compute the filter coefficients only once during a channel realization. At each trellis step and for each trellis state the module called 'PSP survivor storage' gets the hard-decisions of the data sequence corresponding to the survivor at that state from the SOVA. With this input, the estimated transmit matrix $\widehat{\mathbf{X}}$ is constructed for the trellis update leaving each state. This update is done by the cancelation module, which takes $\widehat{\mathbf{X}}$ and computes the interference for the current decoding step in the trellis diagram, for each trellis state. Note that we need soft-decisions for the iterative decoding, but only hard-decisions of the data sequence for the cancelation, since the PSP estimates are based on (hard) hypothesized paths leading to the given trellis state. The nulling module takes the matrix $\mathbf{F}^H$, which is obtained by the filter coefficients module to null out the interference from not yet decoded codeword symbols.

3.5.2.2 ST SOVA Decoder

In order to obtain the path metric for the ST SOVA decoder in the presence of interference from the other layers, we use the matrices $\mathbf{F}$ and $\mathbf{B}$ for nulling out the impact of the upper (not yet detected) layers, and combine this with per-survivor processing for canceling the interference of the lower (already detected) layers. To improve the decoding process we also need a priori information about the transmitted signals in the modified path metric. Let $\mathbf{y}_k$ be the receive vector corresponding to the symbols in $\mathbf{Y}$ and let $\widehat{\mathbf{x}}_k$ be the interference vector corresponding to the symbols in $\widehat{\mathbf{X}}$ at decoding time k, respectively. Furthermore, let the signal-to-interference plus noise ratio at the output of the cancelation module be given by

$$\mu_r = \begin{cases} [\mathbf{R}]_{r,r} & \text{(ZF)} \\ \sqrt{\mathbf{h}_r^H \mathbf{S}_r^{-1} \mathbf{h}_r} & \text{(MMSE)} \end{cases} .$$

where $\mathbf{S}_r$ is given in (3.24) and $[\mathbf{R}]_{r,r}$ is the r-th row and column entry of the upper triangular matrix $\mathbf{R}$. Then, the modified path metric $\mathrm{M}_k(\tau)$ of the path terminating in a state τ in the code trellis at the decoding time k is given by

$$\mathrm{M}_k(\tau) = \min_{\nu \in P(\tau)} \left\{ \mathrm{M}_{k-1}(\nu) + \log p_k(\nu, \tau) + \left| \mathbf{f}_r^H \mathbf{y}_k - \mathbf{b}_r^T \widehat{\mathbf{x}}_k - \mu_r z_{\nu\tau} \right| \right\} ,$$

where $P(\tau)$ denotes the set of parent states of τ, $z_{\nu\tau}$ denotes the modulated symbol on the trellis transition $\nu \to \tau$, $M_{k-1}(\nu)$ is the smallest metric of the path connected to the trellis state ν and $\log p_k(\nu\tau)$ is the logarithm of the a priori probability of the bit c_k corresponding to the trellis transition $\nu \to \tau$. The a priori probability is obtained from the SISO channel decoder. The ST SOVA decoder stepwise decodes the symbols at each stage of the code trellis diagram, storing the survivors terminating in each state of the trellis, and uses these survivors to cancel their impact as interference on following decoding steps. The soft output of the ST SOVA is an approximate log-likelihood ratio of the a posteriori probabilities of the information bits. The soft output can be approximately expressed as the metric difference be-

tween the maximum-likelihood path and its strongest competitor at each decoding step. The strongest competitor of the maximum-likelihood path is the path which has the minimum path metric among a given set of paths. This set is obtained by taking all paths, which have, at the current decoding step, the symbol on their trellis transition complementary to the one on the maximum likelihood path. The ST SOVA decoder provides soft information, which can be expressed as

$$\Lambda(c_k) = \log \frac{P(c_k = 0|\mathbf{Y})}{P(c_k = 1|\mathbf{Y})} = \log \frac{P(z_{\nu\tau} = +1|\mathbf{Y})}{P(z_{\nu\tau} = -1|\mathbf{Y})} = M_k^{-1} - M_k^{+1} \ ,$$

where M_k^{-1} is the minimum path metric corresponding to $z_{\nu\tau} = -1$ and M_k^{+1} is the minimum path metric corresponding to $z_{\nu\tau} = +1$. We can split the soft output of the SOVA into two parts, the extrinsic information $\lambda_{E,STS}(c_k)$ and the intrinsic or a priori information $\lambda_{A,STS}$,

$$\Lambda(c_k) = \lambda_{A,STS}(c_k) + \lambda_{E,STS}(c_k) \ , \tag{3.25}$$

where the a priori information is given as

$$\Lambda_{A,STS}(c_k) = \log \frac{p_k(0)}{p_k(1)} \ . \tag{3.26}$$

Therefore, the extrinsic information, which is fed into the MAP decoder after deinterleaving is obtained from (3.25) and (3.26) as

$$\lambda_{E,STS}(c_k) = \Lambda(c_k) - \lambda_{A,STS}(c_k) \ .$$

Some simulation results of the proposed scheme and their interpretation are presented in the following section.

3.5.3 Numerical simulation

In this section, we illustrate the bit error performance of our proposed scheme, which we call WSTC with iterative decoding (WSTC-ID) in the remainder, and compare it with the performance of the WSTC in [CC01, CC03]. In Fig. 3.19, we present the BER of the WSTC-ID scheme for a system with $n_T = n_R = 4$ transmit and receive antennas and QPSK modulation. The outer coding is performed over multiple block fading channels. After encoding with the outer code and interleaving we divide the whole sequence into $\tau = 8$ blocks. Each block has a length of $L_B = 128$ bits after encoding with the inner code. We assume, that the channel is constant for the transmission for each block and changes independently from block to block. As component codes, we use the binary linear feedback systematic convolutional code $\mathrm{CC}(7, 5)_8$, where the generator polynomials are given in octal numbers with feedback polynomial $G_r = 7$ and feedforward polynomial $G_f = 5$. The overall code rate of our scheme is $R_{WSTC-ID} = 1/4$. Also shown for reference are the simulated BER performance of the WSTC for BPSK modulation. As channel code for WSTC, we used the convolutional code $\mathrm{CC}(7, 5)_8$ with code rate $R_{WSTC} = 1/2$. By comparing the curves of our proposed scheme with the one of WSTC, we observe that although the performance at the first iteration is relatively worse, there is a significant improvement with further iterations, especially for higher SNR values. Furthermore, simulation results show, that there occurs a saturation of the improvement beyond iteration 3. This behavior can be explained by the fact that after each iteration the extrinsic output of the receiver components tends to a Gaussian distribution according to the central limit theorem but the correlation between the extrinsic

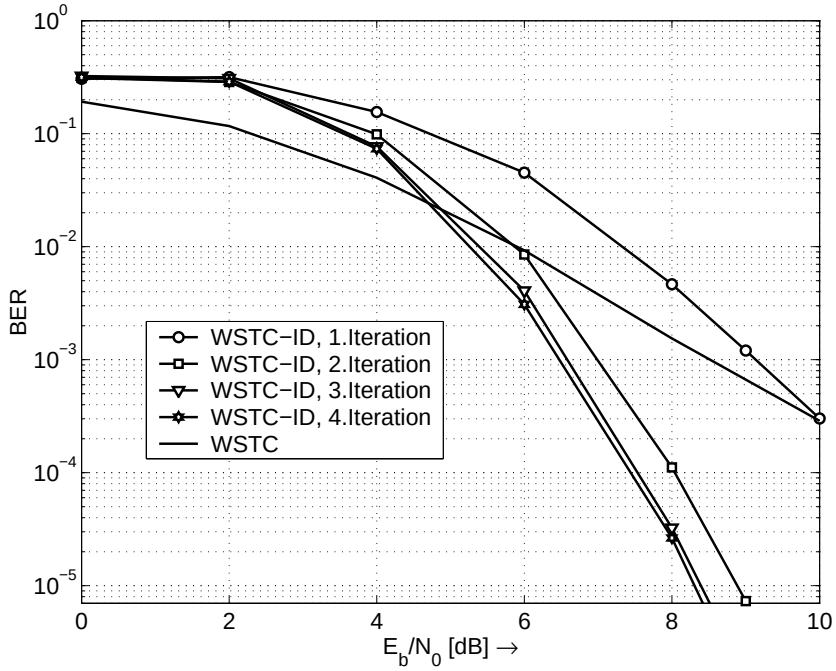

Figure 3.19: Bit error rates, ZF receiver, $n_T = n_R = 4$ antennas, WSTC-ID with coded QPSK modulation with inner and outer code $CC(5,7)_8$, $R_{WSTC-ID} = 1/4$ and WSTC with BPSK and $R_{WSTC} = 1/2$

information and the channel output increases after each iteration [HOP96]. Hence, the improvement of the BER-performance diminishes after each iteration. Even at the first iteration, we observe that the WSTC have a diversity loss in comparison to the new WSTC-ID scheme. The loss is due to the inability of the ZF-or MMSE Decision Feedback Equalization (DFE) within the WSTC to achieve full diversity with given coded modulation over finite alphabets. The diversity gain (or change of slope) of the new WSTC-ID scheme is mainly achieved due to soft-decisions used for the iterative decoding and the distribution of error events (with interleaving) due to a deeply faded block during a transmission. Note that the improvement to WSTC can be further enhanced with a bigger τ and through the enlargement of the whole information bit length.

In addition to the BER for the ZF solution, we present the BER for the MMSE case in Fig 3.20. From the figure, we observe that the new scheme outperforms the WSTC scheme. We also observe, that the improvement in comparison to WSTC is higher than in the ZF case.

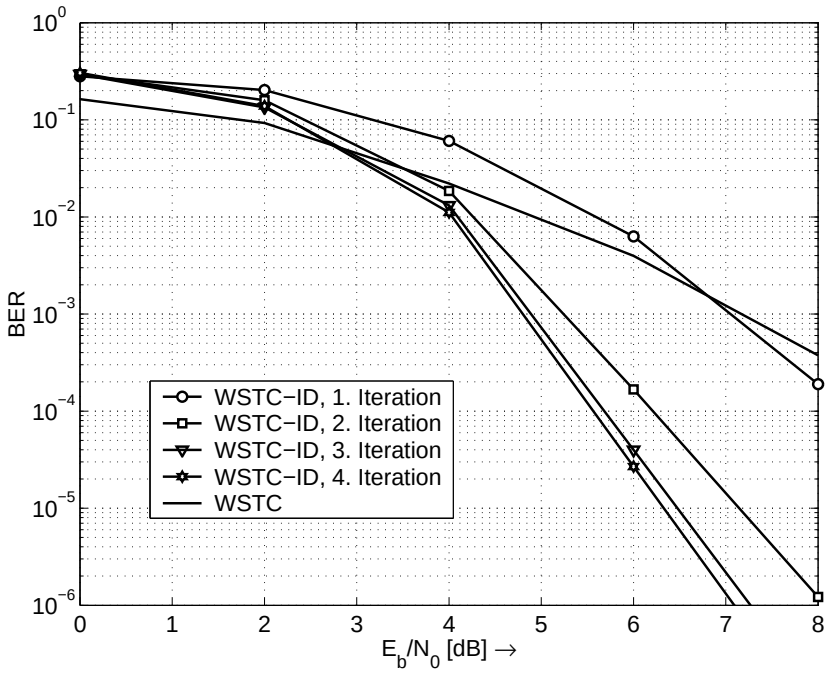

Figure 3.20: Bit error rates, MMSE receiver, $n_T = n_R = 4$ antennas, WSTC-ID with coded QPSK modulation with inner and outer code $CC(5,7)_8$, code rate $R_{WSTC-ID} = 1/4$ and WSTC with BPSK and $R_{WSTC} = 1/2$

4 Conclusions and future research

4.1 Conclusions

In this thesis, the performance of space-time codes for multiple-antenna wireless systems with and without channel coding was studied. In the first part, i.e. without channel coding, we focused on space-time codes from quasi-orthogonal design (QSTBC). There, we completely characterized the equivalent channel representation evolving from the employment of QSTBC at the transmitter in a MIMO system. Based on that, we analyzed the outage probability as well as the average error rate with optimal detection. Suboptimal detection strategies were also considered and their performance analyzed. In the second part of the thesis we considered the combination of channel coding schemes with space-time codes from different classes and iterative detection and decoding at the receiver. Rate oriented as well as diversity oriented space-time codes, schemes assuming perfect CSI at the receiver or no knowledge at all are included in the analysis.

In more detail, the following topics were covered and the following results were derived:

- We have extended the idea of space-time block codes from orthogonal design (OSTBC) to transmission rate one space-time block codes from quasi-orthogonal design (QSTBC) for 2^n transmit and an arbitrary number of receive antennas in order to enhance the achievable rate with OSTBC and to achieve full diversity. We showed, that similar to the OSTBC, the QSTBC transform the MIMO channel into an equivalent channel. This equivalent channel was shown to be still a matrix valued channel, however with very interesting statistics. Most important were the facts that the eigenvectors of the equivalent channel are independent of the channel realizations and that the eigenvalues are i.i.d. non-central chi-square distributed. The rate one QSTBC achieve a high performance by approaching the mutual information in the case of $n_R = 1$ receive antenna. In addition we showed that by increasing the number of receive antennas, the loss in terms of mutual information increases unbounded with the SNR. Furthermore, we developed a new transmit strategy, which allows the application of a linear ML-detector as in the case of OSTBC, so that the symbols from different antennas can be detected separately instead of joint detection, a useful advantage known previously from the OSTBC. The performance of our linear detector in terms of mutual information is equal to the nonlinear ML-detector in [PF03]. We also developed upper and lower bounds on the outage probability with QSTBC. For the ergodic mutual information, we also derived some closed-form expressions.

- From the BER performance point of view, we analyzed two transmit strategies for QSTBC. The objective of the first approach was to reduce the complexity of the detector of the QSTBC, whereas the objective of the second approach was to improve the BER performance. To this end, we analyzed the impact of two performance criteria proposed in earlier works, the minimum Euclidean distance and the diversity product, on the BER performance. We have shown, that the diversity product is the more important performance criterion. Hence, it should be as large as possible. Furthermore, we analyt-

ically derived the BER performance of both approaches and compared the performance with other nonlinear detectors. We observed a tradeoff between receiver-complexity and BER performance.

- In order to further reduce the hardware complexity and costs, we analyzed the impact on antenna selection (AS) based on different selection criteria on both the mutual information and the BER of QSTBC. Furthermore, we derived analytical lower and upper bounds for the outage probability and the average mutual information achieved with QSTBC and AS, respectively. Simulation results show, that with only one additional receive antenna even the ZF-detector with AS outperforms the ML-detector without AS.
- Another approach to reduce the complexity at the receivers is to combine QSTBC with suboptimal detection schemes, such as the ZF detector applying LR aided detection. We showed that the linear QSTBC at the transmitter are highly appropriate for reducing the gap between the maximum-likelihood detector and the LR-ZF detector instead of non-linear signal processing methods at the receiver. We analyzed the effect of LR on the equivalent channel model generated by the QSTBC. From simulation results we observed that the gap between maximum-likelihood and LR-ZF detection is dramatically reduced in comparison to spatial multiplexing (SM) schemes, especially for higher transmission rates.
- By relaxing the constraint of full diversity, we developed high rate QSTBC, which are representing a generalization of the OSTBC. In particular, we showed, that not only the Alamouti-Scheme, a OSTBC for two transmit antennas, achieves the mutual information in the case of one receive antenna, but also its generalized high rate version.
- Another important contribution of this thesis is the combination of channel coding schemes with space-time codes and iterative detection and decoding at the remote unit. To this end, we analyzed a new iterative signal-processing algorithm with very low complexity, which takes advantage of the maximum available diversity. The new algorithm is applicable as an extension to many well-known detection algorithms, such as the ZF or VBLAST detection algorithms. Further on, the proposed algorithm was compared with the regular VBLAST and ZF algorithms, and it was shown by numerical simulations that a high performance gain is achieved in the uncoded as well as in the coded case.
- The concatenation of channel coding and space-time codes was also analyzed for "turbo" decoding, i.e. by allowing the transfer of soft or extrinsic information between the channel decoder and the space-time decoder. We analyzed the impact of different mapping strategies on the information transfer of the SISO detector, derived some criteria for the optimum mapping strategy, and analytically showed that additional performance gain is achieved in comparison to Gray mapping. The performance of this scheme, where OSTBC were chosen as inner space-time code, was investigated with EXIT-charts and BERs and compared with the uncoded scheme in [TJC99b]. The performance of this scheme is discussed and compared to the uncoded system with numerical simulation results.
- The assumption of perfect CSI at the receiver may be unfeasible in different scenarios. Thus, it is not possible to employ OSTBC as inner space-time codes. We therefore analyzed the combination of different channel codes with the noncoherent unitary space-time modulation scheme proposed in [HM00a]. A suboptimal decoding algorithm is derived. Also, as in the coherent case we

studied the performance of our proposed scheme with EXIT-charts, which are used to predict the convergence threshold signal to noise ratios (i.e., "turbo-cliff") for the algorithm.

- A disadvantage of the optimal MAP decoder is the inherent delay of providing soft information at the end of the decoding process. Furthermore, it may not be feasible to combine the MAP decoder with space-time codes, where the soft information is needed immediately in order to perform the decoding process. One such scheme are the WSTC [CC01]. Here, we proposed a novel iterative (turbo) receiver scheme for this low complexity space-time architecture, where the space-time decoding part is carried out by a Space-Time (ST) SOVA decoder. We developed the decision metric for the ST SOVA decoder employing per-survivor processing. Furthermore, we analyzed the performance of our proposed scheme in terms of numerical simulations and compared it with the non-iterative WSTC scheme.

4.2 Future research

There are a number of possibilities for future research. In the following, some research directions are mentioned.

4.2.1 Robustness of space-time codes

In wireless communication scenarios, correlation between the antenna elements or a line-of sight component of the channel (Ricean channels) may have an effect on the diversity gain, or more generally, on the error performance of STC. It is therefore of interest to analyze the robustness of STC in various environments, i.e. channels having other statistics than the Rayleigh fading case like channels with strong line-of-sight components or different channel correlation scenarios and rank-deficient channels (keyholes). In this thesis we have considered one type of fading channel, namely the quasi-static flat fading channel. One of our future research goals is to analyze the robustness and design of space-time coding schemes for general frequency-selective and particularly rapidly time-selective channels, which can be modeled by a Gauss-Markov process. Especially for high data rate mobile applications, the time-invariance assumption, i.e. the channel is constant over a block of data, is violated due to Doppler shifts and therefore temporal channel variations, which become stronger as the carrier frequency increases, have to be considered. High mobility is a huge problem for contemporary wireless systems, which are only able to support low data rates like, e.g. UMTS, or result in complete failure like, e.g. DVB-T. ST coding in a frequency and time-selective environment is very challenging because of the simultaneous presence of time and frequency-selectivity, which usually destroys the structure of the space-time code primarily designed for flat-fading scenarios. However, from a communication theoretic point of view, one can expand well known, or develop new, ST codes for future wireless communication systems enabling the exploitation of time and frequency selectivity in order to achieve higher diversity and coding gains for reliable communications at high speeds.

4.2.2 Code design

The design and performance analysis of Space-Time Transmission schemes has mostly concentrated on two design criteria: the error probability or the ergodic

capacity [TSC98, HH02]. One of the possible future research direction is to design transmission schemes based on a different but highly relevant criterion: The error exponent. Furthermore, the design of these schemes has mostly focused on the case, when there is no channel knowledge at all at the transmitter. However, in practise it is often possible to obtain some kind of channel state information at the transmitter, which has to be taken into account in the design of space-time codes [JSO02]. In addition to this, there has been only little research activity in the field of variable rate space-time codes and space-time codes with adaptive modulation, which is important in order to support services with different QoS and transmission rates. Also, new space-time transmission strategies are needed which show good performance in the low SNR-Regime. In addition to this, the extension of space-time codes to distributed systems like wireless relay networks and also radar needs to be addressed in the future.

4.2.3 Multiuser Multi-cell Multi-Carrier Communications

Another possible extension of this thesis is the deployment of space-time coding schemes in a multi-user multi-cell environment for the analysis of STC in the case of multi-user and inter-cell interference. Most space-time codes have been developed for improving the link quality for noise limited systems in a single-user communication scenario. However, in multi-user networks the space-time coded signals have to cope with interference from other systems users in addition to fading and noise. In fact, using conventional single-user space-time codes in a multi-user environment may potentiality degrade significantly the performance of these codes. It is important to note that practical systems tend to be interference rather than noise limited. Therefore, new concepts are needed to mitigate the impact of interference in multi-user systems. The extension of space-time codes from single-carrier to their application in multi-carrier communication systems like OFDM in order to transform a frequency-selective channel into several frequency-nonselective channels is another interesting research topic, which should be addressed in the future. The additional degrees of freedom in the spectral domain could be exploited in order to improve reliability and to support higher data rates. Furthermore, the space-time codes can be utilized to reduce the complexity at both the base station as well as at the mobile terminals.

In particular, the transformation of the MIMO system, with its complicated expressions for the probability density function of the nonindependent eigenvalues of Wishart matrices into an equivalent channel with independent chi-square distributed eigenvalues due to the application of QSTBC as shown in this work, makes the analysis of different performance measures for such multiuser multi-cell multi-carrier communication systems, like the bit-error rate, the outage probability, the delay limited capacity or the ergodic capacity, more accessible. Furthermore, closed form descriptions of these performance measures may be available by deploying QSTBC at the transmitter side.

Publication List

[JBS04] E.A. Jorswieck, H. Boche, and A. Sezgin. Delay-limited capacity and maximum throughput of spatially correlated multiple antenna systems under average and peak-power constraints. *Proc. of IEEE Info. Theory Workshop 2004, San Antonio, TX, USA*, October 2004.

[JS04] E. A. Jorswieck and A. Sezgin. Impact of spatial correlation on the performance of orthogonal space-time block codes. *IEEE Comm. Letters*, 8(1):21–23, January 2004.

[OS04] T.J. Oechtering and A. Sezgin. A new cooperative transmission scheme using the space-time delay code. *ITG Workshop on Smart Antennas, Munich, Germany*, March 18-19 2004.

[SB03] A. Sezgin and H. Boche. Iterative decoding of low-complexity space-time codes. *PIMRC, Beijing, China*, September 7-10 2003.

[SJ02] A. Sezgin and E.A. Jorswieck. Joint decoding and channel estimation for low-complexity STC. *ASILOMAR CSSC 2002, Pacific Grove, CA USA*, pages 546 –550, November 3-6 2002.

[SJ03a] A. Sezgin and E.A. Jorswieck. Maximum diversity detection for layered space-time codes. *VTC 2003-Spring, Jeju, Korea*, pages 833 –837, April 22-25 2003.

[SJ03b] A. Sezgin and E.A. Jorswieck. On optimal constellations for quasi-orthogonal space-time codes. *ICASSP 2003, Hong Kong, China*, pages 345 –348, 6-10 April 2003.

[SJ04a] A. Sezgin and E.A. Jorswieck. Capacity achieving high rate space-time block codes. *IEEE Comm. Letters*, 9(5):435–437, May 2005.

[SJ04b] A. Sezgin and E.A. Jorswieck. Impact of the mapping strategy on the performance of APP decoded space-time block codes. *to appear in IEEE Transactions on Signal Processing*, 2004.

[SJ05] A. Sezgin and E.A. Jorswieck. On the performance of Partial feedback based Orthogonal Block Coding. *will be presented at VTC 2005-Fall, Dallas, TX USA*, September 25-28 2005.

[SJB03] A. Sezgin, E.A. Jorswieck, and H. Boche. Performance criteria analysis and further performance results for quasi-orthogonal space-time block codes. *ISSPIT 2003, Darmstadt, Germany*, pages 345 –348, December 14-17 2003.

[SJB04] A. Sezgin, E.A. Jorswieck, and H. Boche. On EXIT-chart analysis of coherent and non-coherent space-time codes. *ITG Workshop on Smart Antennas, Munich, Germany*, March 18-19 2004.

[SJC04a] A. Sezgin, E.A. Jorswieck, and E. Costa. Iterative decoding of wrapped space-time codes. *IEEE Transactions on Signal Processing*, 53(5):1937–1941, May 2005.

[SJC04b] A. Sezgin, E.A. Jorswieck, and E. Costa. Lattice-reduction aided detection: Spatial multiplexing versus quasi-orthogonal STBC. *VTC 2004-Fall, Los Angeles, CA USA*, September 26-29 2004.

[SJC04c] A. Sezgin, E.A. Jorswieck, and E. Costa. Optimal transmit strategies in MIMO Ricean channels with MMSE receiver. *VTC 2004-Fall, Los Angeles, CA USA*, September 26-29 2004.

[SJC05] A. Sezgin, E.A. Jorswieck, and E. Costa. Optimal transmit strategies for QSTBC in MIMO Ricean Channels with Linear Detection. *will be presented at PIMRC 2005, Berlin, Germany*, September 11-14 2005.

[SJJ03] A. Sezgin, E.A. Jorswieck, and P. Jung. Analysis of Turbo-Coded MIMO-Systems for noncoherent communication using EXIT-charts. *CISS 2003, Baltimore, MD USA*, March 12 - 14 2003.

[SO04a] A. Sezgin and T. J. Oechtering. Complete characterization of the equivalent MIMO Channel for quasi-orthogonal space-time codes. *submitted to IEEE Transactions on Information Theory. Also available at http://www.user.tu-berlin.de/sezgijdi/*, 2004.

[SO04b] A. Sezgin and T. J. Oechtering. On the outage probability of quasi-orthogonal space-time codes. *Proc. of IEEE Info. Theory Workshop 2004, San Antonio, TX, USA*, October 2004.

[SO04c] A. Sezgin and T.J. Oechtering. Antenna selection with capacity-approaching space-time block codes. *ASILOMAR CSSC 2004, Pacific Grove, CA USA*, November 7-10 2004.

[SO04d] A. Sezgin and T.J. Oechtering. A new resource efficient transmission scheme for cooperative systems. *IEEE Intern. Workshop on SPAWC, Lisbon, Portugal*, July 11-14 2004.

[SWBK03] A. Sezgin, D. Wübben, R. Böhnke, and V. Kühn. On EXIT-charts for space-time block codes. *ISIT 2003, Yokohama, Japan*, June 29 - July 4 2003.

[SWK03] A. Sezgin, D. Wübben, and V. Kühn. Analysis of mapping strategies for turbo-coded space-time block codes. *Proc. of ITW 2003, Paris, France*, pages 103–106, March 30-April 4 2003.

[WJSC05] D. Wang, E.A. Jorswieck, A. Sezgin, and E. Costa. Joint Tomlinson-Harashima Precoding with diversity techniques for multiuser MIMO OFDM systems. *VTC 2005-Spring, Stockholm, Sweden*, May 30- June 1 2005.

References

[Ala98] S.M. Alamouti. A simple transmitter diversity scheme for wireless communications. *IEEE Journal on Selected Areas in Communications*, SAC-16:1451–1458, October 1998.

[Bau99] G. Bauch. Concatenation of space-time block codes and Turbo-TCM. *IEEE Int. Conf. on Communications,Vancouver, BC, Canada*, 2:1202 –1206, June 6-10 1999.

[BBH00] S. Bäro, G. Bauch, and A. Hansmann. Improved codes for space-time trellis-coded modulation. *IEEE Commun. Letters*, 1:20–22, January 2000.

[BCJR74] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv. Optimal decoding of linear codes for minimizing symbol error rate. *IEEE Trans. on Information Theory*, 20:284–287, March 1974.

[BCT01] E. Biglieri, G. Caire, and G. Taricco. Limiting performance of block-fading channels with multiple antennas. *IEEE Trans. on Information Theory*, 47(4):1273–1289, May 2001.

[BD02] I. Bahceci and T.M. Duman. Combined turbo coding and unitary space-time modulation. *IEEE Trans. on Communications*, 50(8):1244–1249, August 2002.

[BG96] C. Berrou and A. Glavieux. Near optimum error correcting coding and decoding:turbo codes. *IEEE Trans. on Communications*, 44(10):1261–1271, October 1996.

[BH02] G. Bauch and J. Hagenauer. Smart versus dumb antennas-capacities and FEC performance. *IEEE Comm. Letters*, 6(2):55–57, February 2002.

[BJ04] H. Boche and E. Jorswieck. Outage probability of multiple antenna systems: Optimal transmission and impact of correlation. *Proc. of IZS*, February 2004.

[BS03] G. Bauch and F. Schreckenbach. How to obtain turbo gains in coherent and non-coherent orthogonal transmit diversity. *IEEE PIMRC, Beijing, China*, pages 1988–1992, September 2003.

[BTT02a] E. Biglieri, G. Taricco, and A. Tulino. Decoding space-time codes with BLAST architectures. *IEE Trans. on Signal Processing*, 50(10):2547–2552, October 2002.

[BTT02b] E. Biglieri, G. Taricco, and A. Tulino. Performance of space-time codes for a large number of antennas. *IEEE Trans. on Information Theory*, 48(7):1794–1803, July 2002.

[BV01] M. Brehler and M.K. Varanasi. Asymptotic error probability analysis of quadratic receivers in rayleigh-fading channels with applications to a unified analysis of coherent and noncoherent space-time

receivers. *IEEE Trans. on Information Theory*, 47(6):2383–2399, September 2001.

[CC01] G. Caire and G. Colavolpe. Wrapped space-time codes for quasi-static multiple-antenna channels. *WPMC*, September 2001.

[CC03] G. Caire and G. Colavolpe. On space-time coding for quasi-static multiple-antenna channels. *IEEE Trans. on Information Theory*, 49(6):1400–1416, June 2003.

[CNC00] W.J. Choi, R. Negi, and J.M. Cioffi. Combined ML and DFE decoding for the V-BLAST system. *ICC*, 3:1243–1248, 2000.

[CVZ03] Z. Chen, J. Vucetic, and Z. Zhou. Performance of Alamouti scheme with transmit antenna selection. *Electronics Letters*, 39(4):379–381, February 2003.

[Dum98] T.M. Duman. *Turbo codes and turbo coded modulation systems: Analysis and performance bounds.* PhD thesis, Northeastern University, Dept. of Electrical and Computer Engineering, Boston, MA, 1998.

[Ede89] A. Edelman. *Eigenvalues and condition numbers of random matrices.* PhD thesis, Department of Mathematicx, Massachusetts Institute of Technology, Cambridge, MA, 1989.

[FG98] G.J. Foschini and M.J. Gans. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communications*, 6(3):311–335, March 1998.

[FK98] M.P. Fitz and J.V. Krogmeier. Further results on space-time codes for Rayleigh fading. *Proc. Allerton Conf. on CCC*, pages 391–400, September 1998.

[Fos96] G.J. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Labs Tech. J.*, 1(2):41–59, Autumn 1996.

[GAH01] H.E. Gamal and Jr. A.R. Hammons. A new approach to layered space-time coding and signal processing. *IEEE Trans. on Information Theory*, 47(6):2321 –2334, September 2001.

[GFBK99] J.C. Guey, M.R. Fitz, M.R. Bell, and W.Y. Kuo. Signal design for transmitter diversity wireless communication systems over rayleigh fading channels. *IEEE Trans. on Communications*, 47(4):527–537, April 1999.

[GGP03a] A. Gorokhov, D.A. Gore, and A.J. Paulraj. Receive antenna selection for MIMO flad-fading channels: Theory and algorithms. *IEEE Trans. on Info. Theory*, 49(10):2687–2696, October 2003.

[GGP03b] A. Gorokhov, D.A. Gore, and A.J. Paulraj. Receive antenna selection for spatial multiplexing systems: Theory and algorithms. *IEEE Trans. on Signal Proc.*, 51(11):2796–2807, November 2003.

[GHP02] D.A. Gore, R.W. Heath, Jr., and A.J. Paulraj. Transmit selection in spatial multiplexing systems. *IEEE Communications Letters*, 6(11):491–493, November 2002.

[GP02] D.A. Gore and A.J. Paulraj. MIMO antenna subset selection with space-time coding. *IEEE Trans. on Signal Proc.*, 50(10):2580–2588, October 2002.

[GR83] I.S. Gradshteyn and I.M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, Inc., 4 edition, 1983.

[GS01] G. Ganeson and P. Stoica. Space-time block codes: A maximum SNR approach. *IEEE Trans. on Info. Theory*, 47(4):1650–1656, May 2001.

[HG00] A.R. Hammons and H.E. Gamal. On the theory of space-time trellis-coded modulation. *IEEE Trans. on Info. Theory*, 46:524–542, March 2000.

[HH89] J. Hagenauer and P. Hoeher. A viterbi algorithm with soft-decision outputs and its applications. *Proc. Globecom 1989, Dallas, USA*, pages 1680–1686, January 1989.

[HH02] B. Hassibi and B.M. Hochwald. High-rate codes that are linear in space and time. *IEEE Trans. on Information Theory*, 48(7):1804 – 1824, July 2002.

[HHSS01] B. Hassibi, B. Hochwald, A. Shokrollahi, and W. Sweldens. Representation theory for high-rate multiple-antenna code design. *IEEE Trans. on Information Theory*, 47(6):2335 –2367, September 2001.

[HJ85] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[HM00a] B.M. Hochwald and T.L. Marzetta. Systematic design of unitary space-time constellations. *IEEE Trans. on Information Theory*, 46(6):1962–1973, September 2000.

[HM00b] B.M. Hochwald and T.L. Marzetta. Unitary space-time modulation for multiple-antenna communciations in rayleigh flat fading. *IEEE Trans. on Information Theory*, 46(2):543–564, March 2000.

[HOP96] J. Hagenauer, E. Offer, and L. Papke. Iterative decoding of binary and convolutional codes. *IEEE Trans. on Information Theory*, 42:429–445, March 1996.

[HS00] B. Hochwald and W. Sweldens. Differential unitary space-time modulation. *IEEE Trans. on Communications*, 48:2041 –2052, December 2000.

[HSP01] R.W. Heath, Jr., S. Sandhu, and A.J. Paulraj. Antenna selection for spatial multiplexing systems with linear receivers. *IEEE Comm. Letters*, 5(4):142–144, April 2001.

[Ion03] D.M. Ionescu. On space-time code design. *IEEE Trans. on Wireless Comm.*, 2(1):20–28, January 2003.

[Jaf01] H. Jafarkhani. A quasi-orthogonal space-time block code. *IEEE Trans. on Comm.*, 49(1):1–4, January 2001.

[JSO02] G. Jöngren, M. Skoglund, and B. Ottersten. Utilizing partial channel information in the design of space-time block codes. *Proc. 5th Intern. Symp. on WPMC*, October 2002.

[LH02] T.H. Liew and L. Hanzo. Space-time codes and concatenated channel codes for wireless communications. *Proceedings of the IEEE*, 90:187 –219, February 2002.

[LhHmW02] Ying Li, Jun hong Hui, and Xin mei Wang. Non-full rank space-time trellis codes for serially concatenated system. *IEEE Comm. Letters*, 6(9):397–399, September 2002.

[Lia03] X.B. Liang. Orthogonal designs with maximal rates. *IEEE Trans. on Info. Theory*, 49(10):2468 – 2503, October 2003.

[LLL82] A.K. Lenstra, H.W. Lenstra, and L. Lovász. Factoring palynomials with rational coefficients. *Math. Ann.*, 261:515–534, 1982.

[LX03] X.B. Liang and X.-G. Xia. Upper bounds of rates of complex orthogonal space-time block codes. *IEEE Trans. on Info. Theory*, 49(10):2788 – 2796, October 2003.

[MP92] A.M. Mathai and S.B. Provost. *Quadratic Forms in random variables, Theory and Applications*, volume 126 of *Statistics: textbooks and monographs*. Marcel Dekker, Inc., 1992.

[MRG04] C.F. Mecklenbräuker, M. Rupp, and G. Gritsch. On mutual information and outage for extended Alamouti space-time block codes. *Proc. IEEE SAM 2004, Barcelona, Spain*, July 2004.

[Mui82] R.J. Muirhead. *Aspects of Multivariate Statistical Theory*. John Wiley & Sons, 1982.

[MW04] A.F. Molisch and M.Z. Win. MIMO systems with antenna selection. *IEEE Microwave Magazine*, 5(1):46–56, March 2004.

[NBP02] R.U. Nabar, H. Bölcskei, and A.J. Paulraj. Outage properties of space-time block codes in correlated rayleigh or rician fading environments. *Proc. of IEEE Intern. Conf. on Acous., Speech, and Sign. Proc., ICASSP'02*, 3:III–2381–III–2384, May 2002.

[NBP04] R.U. Nabar, H. Bölcskei, and A.J. Paulraj. Diversity and outage performance in Ricean MIMO channels. *to appear in IEEE Trans. Wireless Communications*, 2004.

[NSC98] A.F. Naguib, N. Seshadri, and A.R. Calderbank. Applications of space-time block codes and interference suppression for high data rate wireless systems. *Proc. Asilomar Conf. on Signals, Systems and Computers*, pages 1803–1810, 1998.

[PF03] C.B. Papadias and G.J. Foschini. Capacity-approaching space-time codes for systems employing four transmit antennas. *IEEE Trans. on Info. Theory*, 49(3):726 –733, March 2003.

[PK94] A.J. Paulraj and T. Kailath. Increasing capacity in wireless broadcast systems using distributed transmission/directional reception. *U.S. Patent no. 5,345,599*, 1994.

[Pro01] J.G. Proakis. *Digital Communications*. McGrawHill, Inc., 4th edition, 2001.

[PV01] N. Prasad and M.K. Varanasi. Optimum efficiently decodable layered space-time block codes. *Proc. Asilomar Conf. on Signals, Systems, and Computers, Montery, CA*, November 2001.

[PV03] N. Prasad and M.K. Varanasi. Outage analysis and optimization of a stacked orthogonal space-time architecture and near-outage codes. *Proc. Commun. Th. Symp., IEEE GLOBECOM, San Francisco, CA, USA*, December 2003.

[RM02] M. Rupp and C.F. Mecklenbräuker. On extended Alamouti schemes for space-time coding. *Proc. of IEEE WPMC 2002, Honolulu, Hawaii*, October 2002.

[RMG03] M. Rupp, C.F. Mecklenbräuker, and G. Gritsch. High diversity with simple space time block codes and linear receivers. *Proc. IEEE GLOBECOM 2003, San Francisco, USA*, December 2003.

[RPT95] R. Raheli, A. Polydoros, and C.-K. Tzou. Per-survivor processing: A general approach to mlse in uncertain environments. *IEEE Trans. on Commun.*, 43(2):354–364, February 1995.

[SA00] M.K. Simon and M.S. Alouini. *Digital Communication over Fading Channels*. Wiley Series in Telecommunications and Signal Processing. John Wiley & Sons, Inc., 2000.

[SD99] A. Stefanov and T.M. Duman. Turbo coded modulation for wireless communications with antenna diversity. *VTC Fall*, pages 1565–1569, September 1999.

[SFG02] S. Siwamogsatham, M.F. Fitz, and J.H. Grimm. A new view of performance analysis of transmit diversity schemes in correlated rayleigh fading. *IEEE Trans. on Information Theory*, 48(4):950–956, April" 2002.

[SG01] C. Schlegel and A. Grant. Differential turbo space-time coding. *Proc. IEEE Information Theory Workshop 2001, Cairns, Australia*, pages 120–122, 2001.

[SH00] M. Sellathurai and S. Haykin. Turbo-BLAST for high-speed wireless communications. *Wireless Comm. and Netw. Conf. (WCNC)*, 1:1962–1973, 315 -320 2000.

[SJ03] A. Sezgin and E.A. Jorswieck. On optimal constellations for quasi-orthogonal space-time codes. *ICASSP 2003, Hong Kong, China*, pages 345 –348, 6-10 April 2003.

[SJB03] A. Sezgin, E.A. Jorswieck, and H. Boche. Performance criteria analysis and further performance results for quasi-orthogonal space-time block codes. *ISSPIT 2003, Darmstadt, Germany*, pages 345 –348, December 14-17 2003.

[SL03] H. Shin and J.H. Lee. Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole. *IEEE Trans. on Info. Theory*, 49(10):2636–2646, October 2003.

[SP00] S. Sandhu and A.J. Paulraj. Space-time block codes: A capacity perspective. *IEEE Comm. Letters*, 4(12):384 –386, December 2000.

[SP02a] N. Sharma and C.B. Papadias. Improved quasi-orthogonal codes. *IEEE Wireless Comm. and Network Conf., Orlando, FL, USA*, pages 169 –171, 17-21 March 2002.

[SP02b] N. Sharma and C.B. Papadias. Improved quasi-orthogonal codes through constellation rotation. *IEEE Trans. on Comm.*, 51(3):332 –335, March 2002.

[SP04] N. Sharma and C.B. Papadias. Full-rate full-diversity linear quasi-orthogonal space-time codes for any number of transmit antennas. *EURASIP Journal on Applied Sign. Processing*, 9:1246 –1256, March 2004.

[Spr79] M.D. Springer. *The Algebra of Random Variables*. Wiley Series in Probability and Applied Statistics. John Wiley & Sons, 1979.

[SPS02a] A. Steiner, M. Peleg, and S. Shamai (Shitz). Iterative decoding of space-time differentially coded unitary matrix modulation. *IEEE Trans. on Signal Processing*, 50(10):2385–2395, October 2002.

[SPS02b] A. Steiner, M. Peleg, and S. Shamai (Shitz). Iterative decoding of space-time differentially coded unitary matrix modulation. *IEEE Trans. on Sign. Processing*, 50(10):2385–2395, October 2002.

[SX02] W. Su and X.-G. Xia. Quasi-orthogonal space-time block codes with full diversity. *SPIE 2002, Seattle, Washington, USA*, July 2002.

[SX03] W. Su and X.-G. Xia. Two generalized complex orthogonal space-time block codes of rates 7/11 and 3/5 for 5 and 6 transmit antennas. *IEEE Trans. on Information Theory*, 49(1):313–316, January 2003.

[SX04] W. Su and X.-G. Xia. Signal constellations for quasi-orthogonal space-time block codes with full diversity. *IEEE Trans. on Information Theory*, 50(10):2331–2347, October 2004.

[tB99] S. ten Brink. Convergence of iterative decoding. *IEE Electron. Lett.*, 35(10):806–808, May 1999.

[tB00] S. ten Brink. Designing iterative decoding schemes with the extrinsic information transfer chart. *AEUE Intern. Journal of Electr. and Commun.*, 54:187 –219, November 2000.

[tB01] S. ten Brink. Code characteristic matching for iterative decoding of serially concatenated codes. *ANN. DES TÉLÉCOMMUN.*, 56(7-8):394–408, 2001.

[TBH00] O. Tirkkonen, A. Boariu, and A. Hottinen. Minimal non-orthogonality rate 1 space-time block code for 3+ Tx antennas. *IEEE ISSSTA 2000*, pages 429–432, September 2000.

[Tel99] E. Telatar. Capacity of multi-antenna Gaussian channels. *European Trans. on Telecomm. ETT*, 10(6):585–596, November 1999.

[TH02] O. Tirkkonen and A. Hottinen. Square-matrix embeddable space-time block codes for complex signal constellations. *IEEE Trans. on Information Theory*, 48(2):1122–1126, February 2002.

[Tir01] O. Tirkkonen. Optimizing space-time block codes by constellation rotations. *FWCW 2001, Finland*, pages 59–60, October 2001.

[TJ00] V. Tarokh and H. Jafarkhani. A differential detection scheme for transmit diversity. *IEEE J. Selected Areas Commun.*, 3:1043–1047, July 2000.

[TJC99a] V. Tarokh, H. Jafarkhani, and A.R. Calderbank. Space-time block codes from orthogonal designs. *IEEE Trans. on Information Theory*, 45(5):1456–1467, July 1999.

[TJC99b] V. Tarokh, H. Jafarkhani, and A.R. Calderbank. Space-time block coding for wireless communications: Performance results. *IEEE Journal on Sel. Areas in Communications*, 17(3):451 –460, March 1999.

[TNSC99] V. Tarokh, A. Naguib, N. Seshadri, and A.R. Calderbank. Combined array processing and space-time coding. *IEEE Trans. on Info. Theory*, 45(4):1121–1128, May 1999.

[TSC98] V. Tarokh, N. Seshadri, and A.R. Calderbank. Space-time codes for high data rate wireless communication:performance criterion and code

construction. *IEEE Trans. on Information Theory*, 44(2):744–765, March" 1998.

[vZ00] A. van Zelst. Space division multiplexing algorithms. *Electrotechnical Conference, 2000, MELECON 2000.,10th Mediterranean*, 3:1218–1221, May 2000.

[WBKK04] D. Wübben, R. Böhnke, V. Kühn, and K.D. Kammeyer. Near-maximum-likelihood detection of MIMO systems using MMSE-based lattice reduction. *Proc. of IEEE ICC 2004, Paris, France*, June 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, October 2001.

[WF03] C. Windpassinger and R.F.H. Fischer. Low-complexity near-maximum-likelihood detection and precoding for mimo systems using lattice reduction. *Proc. of IEEE ITW 2003, Paris, France*, April 2003.

[WFGV98] P.W. Wolniansky, G.J. Foschini, D.G. Golden, and R.A. Valenzuela. V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel. *PROC. ISSSE*, 1998.

[WG04] Z. Wang and G.B. Giannakis. Outage mutual information of space-time MIMO channels. *IEEE Trans. on Information Theory*, 50(4):657–662, April 2004.

[Win87] J. Winters. On the capactiy of radio communication systems with diversity in a Rayleigh fading environment. *IEEE J. Select. Areas Commun.*, 5:871–878, June 1987.

[Wit91] A. Wittneben. Base Station Modulation Diversity for Digital SIMULCAST. *Proc. of IEEE VTC 1991*, pages 848–853, May 1991.

[WL03] W.H. Wong and E.G. Larsson. Orthogonal space-time block coding with antenna selection and power allocation. *Electronics Letters*, 39(4):379–381, February 2003.

[YB00] Q. Yan and R.S. Blum. Optimum space-time convolutional codes for quasi-static slow fading channels. *Proc. of WCNC*, pages 1351–1355, September 2000.

[YW02] H. Yao and G.W. Wornell. Lattice-reduction-aided detectors for MIMO communication systems. *Proc. of IEEE Globecom 2002, Taipei, Taiwan*, November 2002.

[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):1073–1096, May 2003.