Statistical Analysis and Transmit Prefiltering for MIMO Wireless Systems in Correlated Fading Environments

Von der Fakultät Informatik, Elektrotechnik und Informationstechnik der Universität Stuttgart zur Erlangung der Würde eines Doktor-Ingenieurs (Dr.-Ing.) genehmigte Abhandlung

Vorgelegt von

Mario Kießling

aus Karlsruhe

Hauptberichter: Prof. Dr.-Ing. J. Speidel
Mitberichter: Prof. Dr. Dr. H. Boche

Tag der mündlichen Prüfung: xxx

Institut für Nachrichtenübertragung der Universität Stuttgart

2004
The dissertation at hand evolved from my research activities at the Institute of Telecommunications at the University of Stuttgart and Siemens Information and Communication Mobile in Ulm.

Special thanks go to my professor, Dr.-Ing. Joachim Speidel, for giving me the opportunity to work under his supervision. Numerous profitable discussions and suggestions have sustainably contributed to the success of this work. Moreover, I would like to thank him for his generous support, enabling me to publish major parts of my research results.

Also, I cordially thank Prof. Dr.-Ing. Dr. rer. nat. Holger Boche for taking over the assessment of this thesis.

I am grateful to all colleagues at the Institute of Telecommunications for fruitful discussions and their backing.

Furthermore, I express my sincere gratitude to Dr.-Ing. Hans-Joachim Dreßler and Dr.-Ing. Gottfried Schnabl of the Siemens mobile systems engineering department for their considerate support and for providing me a productive research environment. I am obliged to all my colleagues at Siemens Ulm for their constant support and the valuable discussions. Special thanks go to Dr.-Ing. Markus Reinhardt, who was always open for questions and encouraged my research work.
## Contents

### Acronyms, Notation, and Symbols
- Acronyms: xi
- Notation: xii
- Frequently Used Symbols: xiv

### Abstract
- Kurzfassung: xix

### 1 Introduction

### 2 MIMO Wireless Channel and System Model

#### 2.1 Scalar Wireless Channel
- Multipath and fading: 5

#### 2.2 Matrix Wireless Channel
- General MIMO channel modelling: 9
- Correlation properties of the matrix channel: 10
- Kronecker correlation model: 12
- Statistical channel model with Kronecker covariance assumption: 12
- Note on the flat fading channel assumption: 15
- Determination of the receive and transmit correlation matrices: 16

#### 2.3 System and Signal Model
- Flat fading signal model: 17
- Channel state information: 18
- Equivalent systems: 19
- Note on channel coding: 19
- Two eigenvalue matrix definitions: 20
- Definition of the signal to noise ratio: 21

#### 2.4 Simulation Parameters: 21

### 3 MIMO Mutual Information and Capacity

#### 3.1 MIMO Mutual Information
- General expressions: 25
- Distribution of a complex generalized matrix quadratic form: 27
- Calculation of the MGF of mutual information: 30
- Hypergeometric matrix functions and scalar representations: 32
- MGF for fully correlated channels: 33
- MGF for semi-correlated and uncorrelated channels: 35

#### 3.2 Ergodic Capacity with Uninformed Transmitter
- Fully correlated channels: 38
- Semi-correlated and uncorrelated channels: 39

#### 3.3 Ergodic Capacity Asymptotics: 42
9.2.2 Derivation of Theorem 3.10 ................................. 132
9.2.3 Derivation of Corollary 3.6 ................................. 136
9.2.4 Derivation of Theorem 3.11 ................................. 137
9.3 Ergodic Capacity Bound ................................. 140
  9.3.1 Derivation of Theorem 4.1 ................................. 140

10 Appendix - MIMO Receiver Performance .................. 143
  10.1 MIMO ZF Receiver SER Calculation .......................... 143
    10.1.1 Proof of Theorem 5.1 ................................. 143
    10.1.2 Proof of Theorem 5.2 ................................. 144
    10.1.3 Proof of Lemma 5.1 ................................. 145
    10.1.4 Proof of Theorem 5.7 ................................. 149
  10.2 Statistical Transmit Prefilters for ZF Receivers .......... 150
    10.2.1 Proof of Theorem 5.9 ................................. 150
  10.3 MIMO MMSE Receiver SER Calculation .................. 152
    10.3.1 Proof of Theorem 6.1 ................................. 152
    10.3.2 Proof of Theorem 6.3 ................................. 153
    10.3.3 Proof of Lemma 6.3 ................................. 154
    10.3.4 Proof of Theorem 6.7 ................................. 158
    10.3.5 Proof of Theorem 6.8 ................................. 159
    10.3.6 Proof of Theorem 6.9 ................................. 159
  10.4 MIMO ML Receiver SER Calculation .................. 160
    10.4.1 Proof of Theorem 7.1 ................................. 160
    10.4.2 Proof of Theorem 7.3 ................................. 162
    10.4.3 Proof of Theorem 7.4 ................................. 163

11 Appendix - Mathematical Preliminaries .................. 167
  11.1 Linear Algebra ........................................... 167
    11.1.1 DFT matrix ........................................... 167
    11.1.2 Kronecker product .................................... 167
    11.1.3 Inversion lemmas .................................... 167
    11.1.4 Partitioned matrices ................................ 168
    11.1.5 Determinants ........................................ 168
    11.1.6 Elementary symmetric functions of a matrix ........ 170
  11.2 Multivariate statistics .................................. 171
    11.2.1 Complex normal distributions ......................... 171
  11.3 Complex Integrals ........................................ 172
    11.3.1 Notation ............................................ 172
    11.3.2 Some Gaussian integrals ................................ 173
  11.4 Complex Gaussian Distribution and Associated Integrals .. 173
    11.4.1 Probability density function ......................... 173
    11.4.2 Some integrals ....................................... 174
  11.5 Majorization theory ...................................... 174
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.5.1</td>
<td>Basic definitions</td>
<td>175</td>
</tr>
<tr>
<td>11.5.2</td>
<td>Some lemmas (including results for matrices)</td>
<td>175</td>
</tr>
<tr>
<td>11.5.3</td>
<td>Theorems on Schur-convexity</td>
<td>176</td>
</tr>
<tr>
<td>11.5.4</td>
<td>Theorem on Schur-concavity</td>
<td>176</td>
</tr>
<tr>
<td>11.6</td>
<td>Special functions</td>
<td>177</td>
</tr>
<tr>
<td>11.6.1</td>
<td>Q function</td>
<td>177</td>
</tr>
<tr>
<td>11.6.2</td>
<td>Complementary error function</td>
<td>177</td>
</tr>
<tr>
<td>11.6.3</td>
<td>Exponential integral $E_1(x)$</td>
<td>177</td>
</tr>
<tr>
<td>11.6.4</td>
<td>General hypergeometric function</td>
<td>179</td>
</tr>
<tr>
<td>11.6.5</td>
<td>Hypergeometric function</td>
<td>179</td>
</tr>
<tr>
<td>11.6.6</td>
<td>Kummer $U(a,b,z)$ function</td>
<td>180</td>
</tr>
</tbody>
</table>

**Bibliography** | 183  |
Acronyms, Notation, and Symbols

Acronyms

AS  Angular Spread
AWGN  Additive White Gaussian Noise
BER  Bit Error Rate
CCDF  Complementary Cumulative Distribution Function
CDF  Cumulative Distribution Function
CDIT  Channel Distribution Information at the Transmitter
CDMA  Code Division Multiple Access
CE  Channel Estimation
CF  Characteristic Function
CSI  Channel State Information
dB  Decibel
DOA  Direction Of Arrival
DOD  Direction Of Departure
EC  Ergodic Capacity
ECMM  Exponential Correlation Matrix Model
EM  EigenMode
FCC  Fully Correlated Channel (transmit as well as receive correlation)
FDD  Frequency Division Duplex
i.i.d.  identically independently distributed
LAN  Local Area Network
LOS  Line Of Sight
LT  Long Term
MGF  Moment Generating Function
MIMO  Multiple Input Multiple Output
MISO  Multiple Input Single Output
ML  Maximum Likelihood
MMI  Mean Mutual Information
MMSE  Minimum Mean Squared Error
MRC  Maximum Ratio Combiner; Maximum Ratio Combining
MSE  Mean Squared Error
NLOS  Non Line Of Sight
OC  Optimum Combining
OFDM  Orthogonal Frequency Division Duplex
PDF  Probability Density Function
PA  Power Allocation
PEP  Pairwise Error Probability
QPSK  Quaternary Phase Shift Keying
RCMM  Realistic Correlation Matrix Model
RV  Random Variable
RX  Receiver; receive
SC  Subchannel
SCC Semi Correlated Channel
SER Symbol Error Rate
SIC Successive Interference Cancellation
SISO Single Input Single Output
SIMO Single Input Multiple Output
SINR Signal to Interference plus Noise Ratio
SNR  Signal to Noise Ratio
ST   Short Term
TDD Time Division Duplex
TX  Transmitter; transmit
UCC Uncorrelated Channel
ULA Uniform Linear Array
WF   WaterFilling
w.r.t. with respect to
ZF   Zero Forcer; Zero Forcing

Notation

\( 0 \)  Matrix with all elements of value 0
\( 1 \)  Matrix with all elements of value 1
\( \hat{\alpha}_k \)  Index (sub)set of cardinality \( k \)
\( \langle \hat{\alpha}_k \rangle \)  Cardinality of index subset \( \hat{\alpha}_k \) (here \( \langle \hat{\alpha}_k \rangle = k \))
\( \alpha_m(X), \alpha_m(x) \)  Vandermonde determinant of diagonal matrix \( X \), vector \( x \)
\( X^* \)  Complex conjugate of \( X \)
\( X^\dagger \)  Pseudo-inverse of \( X \)
\( X^T \)  Transpose of matrix \( X \)
\( X^H \)  Hermitian (conjugate transpose) of matrix \( X \)
\( x \sim \) Random variable \( x \) is distributed as
\( x \cong \) Random variable \( x \) is statistically equivalent to
\( x \equiv y \) Expression \( x \) is defined as expression \( y \)
\( X \otimes Y \)  Kronecker product of matrices \( X \) and \( Y \)
\( |x| \)  Absolute value of scalar \( x \)
\( |x_{ij}| \)  Determinant of matrix with \( x_{ij} \) in row \( i \) and column \( j \)
### Acronyms, Notation, and Symbols

<table>
<thead>
<tr>
<th>Expression</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[x_{ij}]$</td>
<td>Matrix with element $x_{ij}$ in row $i$ and column $j$</td>
</tr>
<tr>
<td>$[X]_{ij}$</td>
<td>Element of matrix $X$ in row $i$ and column $j$</td>
</tr>
<tr>
<td>$</td>
<td>X</td>
</tr>
<tr>
<td>$</td>
<td>X</td>
</tr>
<tr>
<td>$</td>
<td>X</td>
</tr>
<tr>
<td>${X}_{\hat{\beta}_k}^{\hat{\alpha}_k}$</td>
<td>Matrix of row subset $\hat{\beta}_k$ and column subset $\hat{\alpha}_k$ of $X$</td>
</tr>
<tr>
<td>$|X|^2$</td>
<td>Frobenius norm of matrix $X$</td>
</tr>
<tr>
<td>$k!$</td>
<td>Factorial of $k$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Expression</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta(t)$</td>
<td>Continuous time Dirac impulse</td>
</tr>
<tr>
<td>$\delta(n)$</td>
<td>Discrete time Dirac impulse</td>
</tr>
<tr>
<td>$\psi(z)$</td>
<td>Digamma function</td>
</tr>
<tr>
<td>$\text{diag}(x_k)$</td>
<td>Diagonal matrix with elements $x_k$</td>
</tr>
<tr>
<td>$\text{diag}(x)$</td>
<td>Diagonal matrix with elements of $x$</td>
</tr>
<tr>
<td>$\text{diag}(X)$</td>
<td>Vector of diagonal elements of $X$</td>
</tr>
<tr>
<td>$\text{blkdiag}(X_1, X_2, \ldots, X_n)$</td>
<td>Block diagonal matrix of matrices $X_1, X_2, \ldots, X_n$</td>
</tr>
<tr>
<td>$E = 0.5772156649\ldots$</td>
<td>Euler’s constant</td>
</tr>
<tr>
<td>$E_1(x)$</td>
<td>Exponential integral</td>
</tr>
<tr>
<td>$E_1[f(x)]$</td>
<td>Expected value of $f(x)$ with respect to $x$</td>
</tr>
<tr>
<td>$\text{eig}(X)$</td>
<td>Diagonal matrix of eigenvalues of matrix $X$</td>
</tr>
<tr>
<td>$\text{eig}_0(X)$</td>
<td>Diagonal matrix of non-zero eigenvalues of matrix $X$</td>
</tr>
<tr>
<td>$\text{erfc}(x)$</td>
<td>Complementary error function</td>
</tr>
<tr>
<td>$_nF_m(a;b;x)$</td>
<td>Scalar hypergeometric function</td>
</tr>
<tr>
<td>$_p\tilde{F}_m(a;b;X)$</td>
<td>Complex hypergeometric function of one matrix argument</td>
</tr>
<tr>
<td>$_n\tilde{F}_m(a;b;X, Y)$</td>
<td>Complex hypergeometric function of 2 matrix arguments</td>
</tr>
<tr>
<td>$\Gamma(x)$</td>
<td>Gamma function</td>
</tr>
<tr>
<td>$\Gamma_m(r)$</td>
<td>Modified complex multivariate Gamma function</td>
</tr>
<tr>
<td>$\tilde{\Gamma}_m(r)$</td>
<td>Complex multivariate Gamma function</td>
</tr>
<tr>
<td>$I_n$</td>
<td>Identity matrix of size $n \times n$</td>
</tr>
</tbody>
</table>
\begin{itemize}
\item $\Im \{z\}$: Imaginary part of complex variable $z$
\item $J_0(x)$: Bessel function of the first kind of order $0$
\item $\ln(x)$: Natural logarithm, base $e$
\item $\log_n(x)$: Logarithm, base $n$
\item $\max(x_1, x_2, \ldots, x_n)$: Maximum of the elements $x_1, x_2, \ldots, x_n$
\item $\max_x f(x)$: Maximum of $f(x)$ with respect to $x$
\item $\min(x_1, x_2, \ldots, x_n)$: Minimum of the elements $x_1, x_2, \ldots, x_n$
\item $\min_x f(x)$: Minimum of $f(x)$ with respect to $x$
\item $N(\mu, \sigma^2)$: Normal distribution with mean $\mu$ and variance $\sigma^2$
\item $\tilde{N}(\mu, \sigma^2)$: Complex normal distribution (mean $\mu$, variance $\sigma^2$)
\item $\tilde{\tilde{N}}(M, \Sigma \otimes \Psi)$: Complex matrix variate normal distribution
\item $\Pr(x < x_0)$: Probability of RV $x$ being smaller than $x_0$
\item $Q(x)$: $Q$ function
\item $\Re \{z\}$: Real part of complex variable $z$
\item $\Res_l(f(z)) \big|_{z = z_l}$: Residuum of complex function $f(z)$ at $l$th pole $z_l$
\item $\sigma(x)$: Unit step function
\item $\sgn(x)$: Signum function
\item $\rk(X)$: Rank of matrix $X$
\item $\mathbb{R}$: Set of real numbers
\item $\tr(X)$: Trace, i.e. sum of diagonal elements of $X$
\item $\tr_k(X)$: $k$th elementary symmetric function of matrix $X$
\item $U(a, b, z)$: Kummer $U$ function
\item $\vec(X)$: Vector with stacked columns of matrix $X$
\item $\tilde{W}_m(n, C)$: Complex $m \times m$ Wishart distr., param. $n$, covariance $C$
\item $x$: Vector $x$
\item $X$: Matrix $X$
\end{itemize}
Frequently Used Symbols

\(\Delta_{RX}\) (Root mean square) angular spread at the receiver

\(\Delta_{TX}\) (Root mean square) angular spread at the transmitter

\(\Delta_{Y_{dB, RX}}\) SNR shift in dB due to receive correlation at high SNR

\(\Delta_{Y_{dB, TX}}\) SNR shift in dB due to transmit correlation at high SNR

\(\varepsilon_a\) Average MSE (averaged over channel statistics)

\(\Phi\) Diagonal power allocation matrix of transmit prefilter

\(\gamma\) Mean SNR per symbol, given by \(\gamma = \frac{E_s}{N_0}\)

\(\gamma_{dB}, \tilde{\gamma}_{dB}\) Normalized mean SNR in dB [see (2.75), (2.77)]

\(\gamma_{SC, k}\) S(I)NR on MIMO subchannel \(k\) (random variable)

\(\Lambda_{TX}\) Diagonal matrix of eigenvalues of TX correlation matrix

\(\mu\) \(\min(R, T)\)

\(\nu\) \(\max(R, T)\)

\(\Omega\) Correlation matrix at side of link with \(\nu\) antennas (3.11)

\(\rho\) Transmit power constraint

\(\Sigma\) Correlation matrix at side of link with \(\mu\) antennas (3.10)

\(A\) Matrix root of RX correlation matrix, i.e. \(A^H A = R_{RX}\)

\(b\) Parameter of conditional SER for QAM constellations

\(B\) Matrix root of TX correlation matrix, i.e. \(B^H B = R_{TX}\)

\(c\) Parameter of conditional SER for QAM constellations

\(C_{\text{erg}}\) Ergodic MIMO capacity with uninformed transmitter

\(C_{\text{erg}}^{\text{CDIT}}\) Ergodic MIMO capacity with CDIT

\(C_{\text{erg}}\) Ergodic MIMO capacity asymptotics in high SNR regime

\(C_{\text{erg}}^{\text{B}}\) Ergodic MIMO capacity asymptotics in low SNR regime

\(C_{\text{erg}}^{\text{B, CDIT}}\) Tight bound on ergodic MIMO capacity

\(C_{\text{erg}}^{\text{B, CDIT}}\) EC with CDIT and PA based on tight bound

\(\tilde{C}_{\text{erg}}^{\text{B}}\) Loose bound on ergodic MIMO capacity
<table>
<thead>
<tr>
<th>Acronym/Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tilde{C}_{erg}^{B, CDIT}$</td>
<td>EC with CDIT and PA based on loose bound</td>
</tr>
<tr>
<td>$D_L$</td>
<td>DFT matrix of size $L$ (11.1)</td>
</tr>
<tr>
<td>$E_s$</td>
<td>Mean transmit symbol energy</td>
</tr>
<tr>
<td>$F$</td>
<td>Linear transmit matrix prefilter</td>
</tr>
<tr>
<td>$G_c$</td>
<td>Coding gain</td>
</tr>
<tr>
<td>$H$</td>
<td>Overall (correlated) MIMO matrix channel</td>
</tr>
<tr>
<td>$I(s, y)$</td>
<td>MIMO mutual information</td>
</tr>
<tr>
<td>$K$</td>
<td>Ricean factor</td>
</tr>
<tr>
<td>$K$</td>
<td>Compound matrix of MIMO channel and transmit prefilter</td>
</tr>
<tr>
<td>$M$</td>
<td>Modulation size, e.g. $M = 4$ for QPSK</td>
</tr>
<tr>
<td>$M_q(s)$</td>
<td>MGF of subchannel $S(I)NR$</td>
</tr>
<tr>
<td>$M^A_q(s)$</td>
<td>Approximation of MGF of subchannel $S(I)NR$</td>
</tr>
<tr>
<td>$\bar{M}_q(s)$</td>
<td>MGF of subchannel $S(I)NR$ in the high SNR regime</td>
</tr>
<tr>
<td>$M_{\Omega}(s)$</td>
<td>MGF of MIMO mutual information, $\Omega \neq I, \Sigma = I$</td>
</tr>
<tr>
<td>$M_{\Omega}(s)$</td>
<td>MGF of MIMO mutual information, $\Sigma \neq I, \Omega = I$</td>
</tr>
<tr>
<td>$M_{\Sigma\Omega}(s)$</td>
<td>MGF of MIMO mutual information, $\Sigma \neq I, \Omega \neq I$</td>
</tr>
<tr>
<td>$N$</td>
<td>Diversity (gain)</td>
</tr>
<tr>
<td>$N_0$</td>
<td>Noise power</td>
</tr>
<tr>
<td>$O$</td>
<td>Overall correlation matrix at receiver side</td>
</tr>
<tr>
<td>$P_{b,k}(\gamma)$</td>
<td>(Average) Bit error rate on subchannel $k$</td>
</tr>
<tr>
<td>$P_{s,c}(\gamma_{SC,k})$</td>
<td>Conditional symbol error rate</td>
</tr>
<tr>
<td>$P_{s,k}(\gamma)$</td>
<td>(Average) Symbol error rate on subchannel $k$</td>
</tr>
<tr>
<td>$\bar{P}_{s,k}(\gamma)$</td>
<td>High SNR asymptotics for SER on subchannel $k$</td>
</tr>
<tr>
<td>$r_{RX}$</td>
<td>Correlation parameter at the receiver for ECMM</td>
</tr>
<tr>
<td>$r_{TX}$</td>
<td>Correlation parameter at the transmitter for ECMM</td>
</tr>
<tr>
<td>$R$</td>
<td>Number of receive antenna elements</td>
</tr>
<tr>
<td>$R_{HH}$</td>
<td>Covariance of MIMO channel matrix $H$</td>
</tr>
<tr>
<td>$R_{nn}$</td>
<td>Covariance of Gaussian noise vector $n$</td>
</tr>
<tr>
<td>$\tilde{R}_{nn}$</td>
<td>Normalized covariance of Gaussian noise vector $n$</td>
</tr>
</tbody>
</table>
$R_{RX}$  Correlation matrix at receiver

$R_{ss}$  Covariance of transmit symbol vector $s$

$\tilde{R}_{ss}$  Normalized covariance of transmit symbol vector $s$

$R_{TX}$  Correlation matrix at transmitter

$S$  Overall correlation matrix at transmitter side

$T$  Number of transmit antenna elements

$V_{TX}$  Eigenvector matrix of transmit correlation matrix
Abstract

In the last five years, multiple-antenna wireless systems have moved into the focus of the research community. They promise to provide a multiple of the throughput of standard single input single output (SISO) systems and a higher link reliability without requiring a larger frequency bandwidth. These properties predestine them for an application in wideband fourth generation systems and have paved their way into first commercial wireless LAN products. However, the performance of multiple input multiple output (MIMO) systems depends heavily on a rich scattering environment with uncorrelated fading between the antenna elements. This condition can not always be fulfilled in practice and the system designer has to carefully take into account fading correlation (e.g. due to narrow antenna element spacing) in the system dimensioning. In this thesis we quantitatively assess the performance of MIMO systems in correlated Rayleigh fading environments via a multivariate statistical analysis. To this end, we study the ergodic capacity and the symbol error rate of linear zero-forcing, minimum mean squared error, and nonlinear maximum likelihood receivers. Specifically, we present novel mathematical approaches, which for the first time permit a unifying analysis with arbitrarily correlated MIMO channels. On this basis, we consider the design of linear transmit prefilters based on statistical channel state information at the transmitter only, which allow for a low complexity implementation in frequency division duplex systems. We demonstrate that the proposed filters can significantly increase ergodic capacity and reduce the symbol error rate (SER) for the various receiver types.

Kurzfassung

1 Introduction

In 1995, the pioneering work on multiple-input multiple-output (MIMO) system capacity by Telatar [186] opened up a novel field of wireless research. His purely theoretical studies showed that the deployment of multiple antennas at transmitter and receiver of the wireless link could dramatically increase the link capacity and thus the throughput of wireless systems without a bandwidth expansion. This was later underpinned by the work of Foschini and Gans [35][36], and high spectral efficiencies of up to 40 bit per second per Hertz of bandwidth were demonstrated in laboratory experiments [202]. At the same time, advanced convolutional space-time coding schemes were invented [140][153][183][184] that exploit both the spatial (antenna) and time domain diversity of the MIMO channel for optimizing transmission reliability and/or data rate. One drawback of these codes, however, is the decoding complexity. Assuming maximum likelihood decoding, the computational requirements increase exponentially with an increasing number of antennas, thus prohibiting an application in terminals with limited processing power. For that reason, low complexity space-time block codes were invented [4][185], which were already adapted for the new Universal Mobile Telecommunications System (UMTS) standard [71]. An advantage of most space-time coding schemes is their independence of channel state information (CSI) at the transmitter, i.e. the coding schemes are independent of the instantaneous channel state or the channel statistics.

On the other hand, it was found in theoretical studies of the ergodic (mean) capacity that accurate instantaneous channel state information at the transmitter can beneficially be exploited, especially at lower signal to noise ratio (SNR). To this end, the MIMO channel is essentially decomposed into a set of independent parallel subchannels by proper transmit and receive processing [34]. The same decomposition can be successfully applied for maximizing throughput and/or transmission reliability in real MIMO systems [98]. Optimal linear transmit processing structures based on short-term channel information also exist for the case of purely linear receivers like minimum mean squared error (MMSE) [156][161] or zero-forcing (ZF) [30] receivers. In MIMO literature, one can find the terms transmit prefiltering, which will be used in this thesis, and transmit precoding for describing the linear transmit processing schemes. Another approach that is based on short-term channel state information at the transmitter, is the nonlinear Tomlinson-Harashima precoding technique [62][188]. While all techniques mentioned in this paragraph achieve a significant improvement in performance (i.e. capacity, throughput, symbol error rate), they all have the disadvantage of requiring accurate instantaneous CSI at the transmitter. In a low-mobility time division duplex (TDD) system this requirement can be fulfilled due to the reciprocity of uplink and downlink. However, with increasing mobility or worse, in a frequency division duplex (FDD) system, duality of uplink and downlink is not fulfilled. In the latter case, one has to resort to a fast feedback link, which is always subject to a delay and requires a significant amount of signalling overhead, thus increasing the complexity of the overall system and in particular the terminal. Moreover, the feedback link is prone to transmission errors, which can noticeably degrade system performance.
Most early theoretical work on MIMO systems including capacity, space-time coding, and receiver algorithms was based on the assumption of uncorrelated fading between the antenna elements, requiring a rich scattering environment at both the transmitter and the receiver. However, it is well known from smart antenna literature [155] that the assumption of uncorrelated fading is violated in most practical propagation scenarios. Especially in cellular mobile wireless systems with exposed base station antenna arrays, scattering in the vicinity of the array is negligible, thus leading to a small angular spread of the impinging wavefronts and strong fading correlation between the antenna elements. This effect is enforced by the fact that for practical reasons, the antenna element spacing is limited (typically 0.5 wavelength, which avoids aliasing in beamforming applications).

Monte-Carlo simulations show that similar to single-input multiple-output (SIMO) smart antenna systems, fading correlation between the antenna elements can seriously impact the performance of MIMO links. Without adequate countermeasures, ergodic capacity can be significantly reduced in the presence of fading correlation, while the symbol error rate (SER) of nonlinear maximum likelihood and particularly linear receivers is drastically increased. In this thesis, we quantitatively study the effects of fading correlation by a mathematical analysis using advanced tools from multivariate statistics. The motivation for carrying out this analysis is twofold. First, the qualitative observations from simulations shall be explained in a solid mathematical framework. Having available closed form expressions for performance measures like SER and ergodic capacity, it is a rather simple exercise to derive rules of thumb for MIMO system design. In this context, we emphasize the significance of asymptotics at high SNR [197], which allow for a simple yet powerful assessment of the system performance by only two parameters, namely coding gain and diversity. Second, we want to design linear transmit prefilters that optimize the MIMO link e.g. in terms of the SER and are based on statistical CSI, i.e. the correlation properties of the channel, only. To this end, we need mathematical expressions of the SER in terms of the correlation properties, which can be optimized via the transmit filter. Basically, the statistical transmit filter can be conceived as a beamformer for each of the independent data streams that is transmitted over the wireless MIMO link. In contrast to the filters based on instantaneous CSI (see above), the statistical prefilters allow for an application in FDD and cellular systems with higher mobility, as they require only the information on the long-term stable correlation properties of the MIMO channel at the transmitter, which will be denoted by channel distribution information at the transmitter (CDIT) in the following. However, long-term CSI is only determined by the propagation environment, i.e. the position of the main scattering clusters and can be provided to the transmitter via a low rate feedback link or via frequency transformation between uplink and downlink [74]. Furthermore, the transmit filters are of low computational complexity, as they have to be updated only on a long-term time scale and thus are suited for simple terminals.

The basis for all following derivations is given in Chapter 2. We introduce the MIMO flat Rayleigh fading channel model with Kronecker product covariance structure, i.e. fading correlation is independently modeled with a separate correlation matrix at the transmitter and the receiver side of the MIMO link. While this channel model is not general, it has been shown in several measurement campaigns (e.g. [18][90]) that it is sufficiently accurate in many practical
situations. Furthermore, its specific structure allows for a concise statistical characterization in terms of complex matrix variate normal distributions. Many results from multivariate statistics can directly be deployed for the MIMO system analysis. In this chapter, we also introduce the notation for the basic parameters of the MIMO system (e.g. the number of transmit and receive antennas, correlation matrices etc.) that will be used consistently throughout this thesis.

In Chapter 3 we analyze mutual information and in particular ergodic capacity of the flat fading MIMO link that was introduced in Chapter 2. Due to the fading nature of the MIMO channel, mutual information is a random variable and we outline a new non-asymptotic analysis via a moment generating function (MGF) approach that is exact for an arbitrary number of transmit and receive antennas. To this end, we make use of advanced methods from multivariate statistics, namely hypergeometric functions of matrix arguments and related integrals. It turns out that the ergodic capacity of MIMO systems in Rayleigh fading environments with channel correlation at transmitter as well as receiver can be expressed in terms of exponentials, powers, and the exponential integral only. Based on these closed-form expressions, we derive asymptotics in the low and high SNR regime that allow for a simple characterization of the effects of fading correlation without the need for lengthy Monte-Carlo simulations. Moreover, using the novel formulas, we introduce a statistical waterfilling (WF) algorithm that achieves ergodic capacity in case of CDIT. The conclusions drawn from the capacity analysis can serve as design guidelines for space-time coded MIMO systems in the presence of fading correlation, as the ergodic capacity can serve as an upper bound on the achievable throughput of coded systems that deploy adaptive modulation and large interleaver sizes that capture a sufficient number of fading periods.

Even though the exact ergodic capacity expressions derived in Chapter 3 can be easily evaluated via mathematical software packages like e.g. Matlab, a further analysis based on these formulas turns out to be difficult because of the mathematical complexity. For example, an interesting question is, whether fading correlation always decreases ergodic capacity and an answer to this question based on the exact formulas appears not to be feasible. Therefore, in Chapter 4 we introduce a tight and simple upper bound on ergodic capacity in terms of elementary symmetric functions of the transmit and receive correlation matrix of the MIMO channel. It is shown that the novel bound can be successfully deployed for fully characterizing the influence of correlation on ergodic capacity via the powerful mathematical tool of majorization theory [130] and we give a rigorous prove of the intuitive perception that stronger fading correlation always leads to lower ergodic capacity. Furthermore, we effectively use the capacity bound for deriving a CDIT based statistical WF scheme that is shown to significantly increase ergodic capacity. We demonstrate via Monte-Carlo simulations that the proposed scheme can achieve almost the same performance as a WF algorithm based on instantaneous CSI in heavily correlated channels. Moreover, we show that this low complexity scheme yields essentially the same ergodic capacity as the CDIT based exact WF scheme introduced in Chapter 3.

The performance of MIMO systems with linear zero-forcing (ZF) receivers is studied in Chapter 5. Linear receivers are attractive because of their low complexity and are thus well suited for simple terminals. However, Monte-Carlo simulations show that their symbol error rate (SER) performance is heavily degraded in the presence of fading correlation at the antenna arrays. In this
work, we present a thorough quantitative assessment of the impact of correlation via a statistical analysis of SER. Exact expressions are given for propagation scenarios with arbitrary fading correlation at the antenna arrays. To this end, we make use of complex Gaussian integrals, a new powerful mathematical approach for analyzing the performance of linear multi-antenna receivers. Due to the mathematical complexity of the resulting formulas, in case of fading correlation at the receive antenna array, we introduce a novel tight SER bound, which yields accurate results in the whole SNR range of practical interest. Exact high SNR asymptotics for arbitrary correlation properties of the channel are calculated, which allow for a concise performance assessment based simply on the two parameters coding gain and diversity [197]. Based on the SER analysis, in the second part of Chapter 5 we design a linear transmit prefilter that minimizes SER and is based on statistical CDIT only. To this end, again majorization theory proves to be a valuable tool for solving the optimization problem that arises in the design process. Interestingly, the prefilter designs for short-term and long-term CSI at the transmitter exhibit an interesting duality, i.e. their basic mathematical structure is the same.

Chapter 6 parallels the structure of Chapter 5, whereas now the focus is on MIMO minimum mean squared error (MMSE) receivers. We demonstrate the close connection between wireless smart antenna systems deploying the so-called optimum combining algorithm and MIMO MMSE processing, such that we can partially reuse of existing results for analyzing the SER performance of MMSE receivers in correlated fading environments. However, it is shown that such an analysis is of considerable mathematical complexity and we again present a novel mathematical approach based on complex Gaussian integrals, which complements and unifies all available partial solutions. Specifically, it allows for the analysis of systems with correlation at the receive antenna array, which was not possible so far with existing approaches. Furthermore, in the second part of Chapter 6 we design a statistical transmit prefilter that minimizes the average MSE and simultaneously the overall SER. The optimization problem is again solved by majorization theory and simulation results show that a considerable performance gain can be achieved by the proposed transmit filter.

Finally, in Chapter 7 we consider the non-linear MIMO maximum likelihood (ML) receiver. Based on a moment generating function approach, we derive exact formulas for the pairwise error probability (PEP) in the presence of fading correlation at transmitter as well as receiver. Using the PEP expressions, we calculate a tight union bound on the SER. Asymptotic results in the high SNR regime again allow for a simple characterization of the influence of fading correlation on the performance of ML receivers. Furthermore, we outline the design of a transmit prefilter based on the SER asymptotics, which is based on the statistical properties (i.e. the correlations) of the channel only and minimizes SER. Monte-Carlo simulations show the tightness of the SER bound and the effectiveness of the transmit filter.

A summary and conclusion of this thesis is provided in Chapter 8.
2 MIMO Wireless Channel and System Model

Starting from the basic theory of scalar, i.e. single input single output (SISO), wireless channels, in this chapter we introduce the flat fading MIMO wireless channel model that will be the foundation for the work in this thesis. Under certain independence assumptions on the direction of departure (DOD) and direction of arrival (DOA) spatial distribution of the electromagnetic wavefronts that are transmitted over the wireless medium, we introduce an abstract statistical MIMO channel model with Kronecker covariance structure (cf. e.g. to [20][41][148][149]). As will be demonstrated below, even though the channel model with Kronecker covariance assumption is not general, it has been shown that it is sufficiently accurate for modelling many practical propagation environments [18][91]. Above all, it is the key for an effective application of results from multivariate statistics. Specifically, the correlated flat Rayleigh fading channel model leads to a complex matrix variate normal distributed MIMO channel matrix [45][59][93]. However, even with the Kronecker covariance assumption, the analysis of MIMO systems in the presence of fading correlation is mathematically challenging and it appears that an analysis of more sophisticated channel models (see e.g. [33] for an overview) is prohibitive, as there are hardly statistical tools available.

Moreover, in this chapter we present the signal and system model with its corresponding notation, which will be used throughout this work. This includes various covariance matrices for noise and signal vectors as well as signal to noise ratio (SNR) definitions, which will be used consistently in the following analysis as well as in Monte-Carlo simulations.

2.1 Scalar Wireless Channel

In this paragraph we shortly summarize the basic development of a scalar channel model with Rayleigh and Ricean fading statistics [77][120][153][174]. We recapitulate the standard multipath propagation model, which is the basis for the evolution of the MIMO matrix channel model utilized in the remainder of the thesis.

2.1.1 Multipath and fading

Let the scalar transmitted (continuous time) bandpass signal of a SISO wireless link be given by

\[ s(t) = Re\{s_0(t) \cdot e^{j2\pi f_c t} \cdot e^{j\phi_c}\}, \]

(2.1)

where \( s_0(t) \) is the complex baseband signal, \( f_c \) is the carrier frequency, \( \phi_c \) is a constant carrier phase, and \( t \) denotes time. Due to multipath fading, the received signal is the sum of multiple delayed copies of the transmitted signal, c.f. Fig. 2.1.
Ignoring additive Gaussian noise at the receiver and a possible additional phase shift due to scattering, the received bandpass signal reads

\[ y(t) = Re \left\{ \sum_{k=1}^{n_p} \alpha_k(t) \cdot s(t - \tau_k) \cdot e^{j2\pi f_{d,\text{max}} \cdot \cos\theta_k \cdot t} \right\}, \]

(2.2)

where \( n_p \) is the number of multipath components, \( \alpha_k \) is the (real) path gain or equivalently path loss of the kth path, \( \tau_k \) is the delay of this path, \( \theta_k \) is the direction of the kth scatterer relative to the receiver, and \( f_{d,\text{max}} \) is the maximum Doppler frequency defined by

\[ f_{d,\text{max}} = \frac{v}{c} f_c, \]

(2.3)

where \( v \) is the velocity of the transmitter and \( c \) is the speed of light. For simplicity, we assume here that scatterers and receiver are stationary, however, an extension is straightforward. Motion introduces a frequency shift and consequently generates a carrier frequency spreading, resulting in a time-varying wireless channel. Considering the multipaths with the highest and lowest delays, the delay spread is given by

\[ \delta_{\tau} = \max_k \tau_k - \min_k \tau_k. \]

(2.4)

**Flat fading channel**

With the help of (2.1) we can rewrite (2.2) as
\[
y(t) = \text{Re} \left\{ \sum_{k=1}^{n_p} \alpha_k(t) \cdot s_0(t - \tau_k) \cdot e^{i\phi_c} \cdot e^{i2\pi f_{d,\max} \cdot \cos \theta_k \cdot t} \cdot e^{-i2\pi f_c \cdot \tau_k} \cdot e^{i2\pi f_c \cdot t} \right\}.
\]  

(2.5)

If the symbol period of the transmit signal is \( T_0 \) much larger than the delay spread, i.e. \( T_0 \gg \delta_\tau \) and all multipath delays are close together \( \tau_k \approx \tau_0 \) for all \( k \), then we can approximate (2.5) by

\[
y(t) \approx \text{Re} \left\{ s_0(t - \tau_0) \cdot \sum_{k=1}^{n_p} \alpha_k(t) \cdot e^{i\phi_c} \cdot e^{i2\pi f_{d,\max} \cdot \cos \theta_k \cdot t} \cdot e^{-i2\pi f_c \cdot \tau_k} \right\} \cdot e^{i2\pi f_c \cdot t}.
\]  

(2.6)

Introducing the phase

\[
\phi_k(t) = 2\pi f_{d,\max} \cdot \cos \theta_k \cdot t - 2\pi f_c \cdot \tau_k + \phi_c
\]  

(2.7)

and defining the channel coefficient \( h(t) \) (which can be extended to include possible transmit and receive filters) at time \( t \)

\[
h(t) = \sum_{k=1}^{n_p} \alpha_k(t) \cdot e^{i\phi_k(t)},
\]  

(2.8)

we can rewrite (2.6) as

\[
y(t) \approx \text{Re} \left\{ s_0(t - \tau_0) \cdot h(t) \cdot e^{i2\pi f_c \cdot t} \right\}.
\]  

(2.9)

Equation (2.9) describes the transmission over a frequency flat fading channel that is invariant across the bandwidth of the transmitted signal. The focus of this thesis is on this important case (see Paragraph 2.2.5 below for comments on this assumption).

Rayleigh fading

Assuming that all scattering contributions are subject to a Doppler shift and a uniform distribution over 0 and \( 2\pi \) of the phases \( \phi_k \) in (2.8), it can be shown that the absolute value of the channel coefficient \( |h(t)| \) is Rayleigh distributed [77]. In the following we assume a normalized channel with \( E[|h(t)|^2] = 1 \). Real and imaginary part \( h_r(t) \) and \( h_i(t) \) of the complex \( h(t) \) are centrally (mean zero) Gaussian (normal) distributed with variance 0.5

\[
h_{r,i}(t) \sim N\left(0, \frac{1}{2}\right),
\]  

(2.10)

whereas the normal probability density function (PDF) is given by [82][147]
where $\mu$ is the mean and $\sigma^2$ is the variance. Equivalently, $h(t)$ has a complex Gaussian distribution of zero mean and variance $1$

$$h(t) \sim \mathcal{N}(0, 1). \quad (2.12)$$

The scalar univariate Gaussian distributions in (2.10) and (2.12) will later be generalized to multivariate vector and matrix variate Gaussian distributions for the MIMO channel model.

Ricean fading

For completeness, we present a simple extension of the Rayleigh fading channel model. In the presence of an additional deterministic non-fading component, we can express the so-called Ricean channel (again assuming a power normalization $E[|h(t)|^2] = 1$) as a superposition

$$h(t) = \sqrt{\frac{K}{1+K}} \cdot h_{\text{fix}}(t) + \sqrt{\frac{1}{1+K}} \cdot h_{\text{Ray}}(t), \quad (2.13)$$

where $K$ is the Ricean factor. The first component $h_{\text{fix}}(t)$ is deterministic of power $1$, while the second component $h_{\text{Ray}}(t)$ is complex normal distributed of power (variance) 1

$$h_{\text{Ray}}(t) \sim \mathcal{N}(0, 1). \quad (2.14)$$

Equivalently, $h(t)$ has a complex normal distribution of non-zero mean

$$h(t) \sim \mathcal{N}\left(\sqrt{\frac{K}{1+K}} \cdot \frac{1}{1+K}\right). \quad (2.15)$$

Note that for $K \to \infty$ we get a totally deterministic channel and for $K = 0$ we arrive at the totally stochastic Rayleigh fading channel model introduced above. At this point, we note that a generalization of (2.13) to the MIMO case is straightforward [34][148]. However, the presence of Ricean fading heavily complicates the statistical analysis of wireless systems.

Discrete channel model

In this thesis, we are interested in a simple mathematical model that captures the statistical properties of the wireless channel. To this end, we introduce a discretized symbol-spaced channel model, which includes all possible transmit and receive filtering (plus upconversion and downconversion etc.). For a flat fading channel, we arrive at the single tap model with sample index $n$

$$h(n) = \delta(n) \cdot h_0, \quad (2.16)$$

where $\delta(n)$ is a discrete unit impulse and based on the results derived above, $h_0$ has a complex Gaussian distribution. The discrete channel model in (2.16) is extended to the MIMO case below.


2.2 Matrix Wireless Channel

The concepts developed for the scalar case can now be applied for the derivation of the matrix MIMO channel model, whereas we directly start with a higher degree of abstraction, i.e. we directly begin with a discrete time flat fading model. As was mentioned above, in this thesis we focus on a channel model with Kronecker product covariance structure. We give a detailed fully discrete derivation of this model, as it appears that there are no comprehensive and concise results available in literature. Furthermore, the insights gained from this derivation prove to be very helpful in later statistical analyses.

2.2.1 General MIMO channel modelling

We focus on a stationary propagation scenario, in the sense that the statistical properties (e.g. fading correlations) of the wireless channel do not change with time. In the following we will therefore drop without loss of generality all time indices. Moreover, as in the scalar channel case above, we make use of the narrowband assumption, implying that runtime differences between the signals received at the antenna elements result in phase shifts only, motivating the use of so-called array response vectors.

Let there be again $n_p$ scatterers leading to a multipath wireless channel. We associate a $T \times 1$ array response vector $a_{T,k}(\theta_{T,k})$ to the array of $T$ transmit antennas and a vector $a_{R,k}(\theta_{R,k})$ to the array of $R$ receive antennas [42][43][117], which depend on the direction of departure (DOD) $\theta_{T,k}$ and direction of arrival (DOA) $\theta_{R,k}$, respectively, to the scatterer $k$ (see Fig. 2.2). The particular characteristic of the array response vector depends on the shape of the array (e.g. uniform linear array ULA, circular array etc.), specifically the element spacing, the carrier frequency, and the radiation pattern of the antenna elements. Obviously, the vector channel from the $T$ transmit antennas to scatterer $k$ is given by

![Fig. 2.2 MIMO channel model](image-url)
\[ h_{T,k} = \alpha_{T,k} \cdot e^{i\phi_{T,k}} \cdot a_{T,k}(\theta_{T,k}), \quad (2.17) \]

where similar to the scalar case the path loss is \( \alpha_{T,k} \) and the phase shift is \( \phi_{T,k} \). On the other hand, the channel from scatterer \( k \) to the \( R \) receive antennas is

\[ h_{R,k} = \alpha_{R,k} \cdot e^{i\phi_{R,k}} \cdot a_{R,k}(\theta_{R,k}). \quad (2.18) \]

Combining the results, the rank one matrix channel constructed by the \( k \)th scatterer reads

\[ H_{TR,k} = \alpha_{TR,k} \cdot e^{i\phi_{TR,k}} \cdot a_{R,k}(\theta_{R,k}) \cdot a^T_{T,k}(\theta_{T,k}), \quad (2.19) \]

with the definitions

\[ \alpha_{TR,k} = \alpha_{R,k} \cdot \alpha_{T,k} \quad (2.20) \]

and

\[ \phi_{TR,k} = \phi_{R,k} + \phi_{T,k}. \quad (2.21) \]

The overall matrix channel \( H \) now results from a superposition of the contributions of all \( np \) scatterers, such that

\[ H = \sum_{k=1}^{np} H_{TR,k} = \sum_{k=1}^{np} \alpha_{TR,k} \cdot e^{i\phi_{TR,k}} \cdot a_{R,k}(\theta_{R,k}) \cdot a^T_{T,k}(\theta_{T,k}). \quad (2.22) \]

After introducing the \( np \times np \) diagonal matrix

\[ D = \text{diag}(\alpha_{TR,k} \cdot e^{i\phi_{TR,k}}) \quad (2.23) \]

and the matrices of array response vectors

\[ A_R = \begin{bmatrix} a_{R,1} & a_{R,2} & \cdots & a_{R,np} \end{bmatrix} \quad A_T = \begin{bmatrix} a_{T,1} & a_{T,2} & \cdots & a_{T,np} \end{bmatrix}, \quad (2.24) \]

we can reformulate (2.22) in matrix notation

\[ H = A_R \cdot D \cdot A^T_T. \quad (2.25) \]

Using the central limit theorem [147], it is clear that the distribution of the elements of the matrix channel in (2.25) tends to a Gaussian distribution, if there is a large number of scatterers.

### 2.2.2 Correlation properties of the matrix channel

In this paragraph we consider the correlation between all elements of the MIMO channel matrix \( H \) in (2.25). We give results both in matrix and a certain sum notation, which will be useful later for deriving the Kronecker product correlation model.
Matrix notation

Consider the covariance matrix of the matrix channel modelled according to (2.25), which shall be defined by (using the vectorization operator that stacks the columns of a matrix)

\[ R_{HH} \equiv E[ \text{vec}(H) \cdot \text{vec}(H)^H ] . \]  

(2.26)

From (2.25) we can directly derive via Kronecker product rule (11.3)

\[ \text{vec}(H) = (A_T \otimes A_R) \cdot \text{vec}(D) . \]  

(2.27)

Plugging this result in (2.26) we find via Kronecker product rule (11.6)

\[ R_{HH} = (A_T \otimes A_R) \cdot R_{DD} \cdot (A_T^H \otimes A_R^H) \]  

(2.28)

with the definition

\[ R_{DD} \equiv E[ \text{vec}(D) \cdot \text{vec}(D)^H ] . \]  

(2.29)

Assuming independent subpath statistics (path attenuations as well as phases), we can calculate

\[ R_{DD} = \text{diag}( \text{vec}(\text{diag}(\alpha_{TR,k}^2))) . \]  

(2.30)

Sum notation

On the other hand, by directly using the sum notation of (2.22) in (2.26), we get

\[ R_{HH} = E \left[ \sum_{k=1}^{n_p} \alpha_{TR,k}^2 \cdot a_T(k) \cdot a_R(k) \cdot \sum_{l=1}^{n_p} \alpha_{TR,l} e^{-j\phi_{TR,k,l}} \cdot a_T^H(l) \cdot a_R^H(l) \right] , \]

(2.31)

where the expected value is with respect to the multipath statistics comprising DOAs \( \theta_{R,k} \), DODs \( \theta_{T,k} \), the phases \( \phi_{TR,k} \), and finally the path gains \( \alpha_{TR,k} \). Using the independence of the path gain and phase statistics, we find from (2.31)

\[ R_{HH} = E \left[ \sum_{k=1}^{n_p} \alpha_{TR,k}^2 \cdot a_T(k) \cdot a_R(k) \cdot a_T^H(k) \cdot a_R^H(k) \right] . \]

(2.32)

Rewriting (2.32) with the help of Kronecker product rule (11.4) results in

\[ R_{HH} = E \left[ \sum_{k=1}^{n_p} \alpha_{TR,k}^2 \cdot \left( a_T(k) \cdot a_T^H(k) \right) \otimes \left( a_R(k) \cdot a_R^H(k) \right) \right] . \]

(2.33)
2.2.3 Kronecker correlation model

Using the assumption of independent DODs and DOAs (see below for the implications of this assumption) in (2.33), we can split the expected values with respect to the DODs and DOAs and find

\[ R_{HH} = \sum_{k=1}^{n_p} \alpha_{TR,k}^2 \cdot E_{\theta_{T,k}}[a_{T,k}(\theta_{T,k}) \cdot a_{T,k}^H(\theta_{T,k})] \otimes E_{\theta_{R,k}}[a_{R,k}(\theta_{R,k}) \cdot a_{R,k}^H(\theta_{R,k})]. \]  

(2.34)

Then introducing the correlation matrices at transmitter and receiver

\[ R_{TX} = E_{\theta_{T,k}}[a_{T,k}(\theta_{T,k}) \cdot a_{T,k}^H(\theta_{T,k})]^T \quad R_{RX} = E_{\theta_{R,k}}[a_{R,k}(\theta_{R,k}) \cdot a_{R,k}^H(\theta_{R,k})] \]  

(2.35)

(note that the expected values are with respect to the distribution of the angular spread at transmitter and receiver, respectively), we find from (2.34)

\[ R_{HH} = \sum_{k=1}^{n_p} \alpha_{TR,k}^2 \cdot R_{TX}^T \otimes R_{RX}. \]  

(2.36)

With normalized multipath gains (power normalization)

\[ \sum_{k=1}^{n_p} \alpha_{TR,k}^2 = 1 \]  

(2.37)

we finally arrive at the special Kronecker product structure of this particular propagation scenario with independent DODs and DOAs

\[ R_{HH} = R_{TX}^T \otimes R_{RX}. \]  

(2.38)

Taking into account the fact that the correlation matrices are Hermitian, we can obviously rewrite (2.38) as

\[ R_{HH} = R_{TX}^* \otimes R_{RX}. \]  

(2.39)

2.2.4 Statistical channel model with Kronecker covariance assumption

In this paragraph, we outline a simple way of modeling a MIMO wireless channel that has the specific correlation structure in (2.39). Moreover, we discuss the statistical properties of this simple model, which will be used extensively throughout this thesis.
Model

We have noted above that the elements of the matrix $H$ tend to a complex Gaussian distribution [202] for a high number of scatterers. Moreover, we have established the Kronecker product structure of the covariance matrix in (2.39). Let us now introduce the matrix decompositions

$$R_{TX} = B^H B \quad R_{RX} = A^H A,$$ (2.40)

which can be established e.g. via a Cholesky factorization. We emphasize that we assume full rank correlation matrices with differing eigenvalues throughout this thesis. This condition is fulfilled in most practical situations and low rank correlation matrices can be taken into account easily. The basic channel model used throughout this thesis is then given by (see e.g. [91][148][149])

$$H = A^H w B,$$ (2.41)

where $H_w$ is a $R \times T$ matrix of i.i.d. complex Gaussian entries. It can readily be checked that the model in (2.41) has exactly the covariance structure in (2.39) using the vectorization operator and Kronecker product rule (11.3). We use the term fully correlated channel (FCC), if both transmit and receive side of the MIMO channel are correlated. If either $A$ or $B$ equals the identity matrix, i.e. only one side is correlated, we name this semi correlated channel (SCC). Finally, without correlation, we talk about an uncorrelated channel (UCC).

In following calculations we make frequent use of the eigenvalue decomposition of the $T \times T$ transmit correlation matrix, where $T$ is the number of transmit antennas, which will be denoted by

$$R_{TX} = V_{TX} \cdot \Lambda_{TX} \cdot V_{TX}^H,$$ (2.42)

with $T \times T$ diagonal matrix of eigenvalues

$$\Lambda_{TX} = \text{diag}(\lambda_{TX,1}, \lambda_{TX,2}, \ldots, \lambda_{TX,T}).$$ (2.43)

Without loss of generality, we assume that the eigenvalues are arranged in decreasing order, i.e. $\lambda_{TX,1}$ is the maximum eigenvalue. Equivalently, we introduce the eigenvalue decomposition of the $R \times R$ receive correlation matrix

$$R_{RX} = V_{RX} \cdot \Lambda_{RX} \cdot V_{RX}^H,$$ (2.44)

where $R$ is the number of receive antennas.

Statistics

A major advantage of the model in (2.41) is the availability of statistical results on the distribution of the $R \times T$ MIMO channel matrix $H$. Namely, in case of a non line of sight (NLOS) environment, $H$ has a zero-mean complex matrix variate Gaussian distribution (see [59],[202], and Appendix 11.2.1), which we denote by
In the remainder of this thesis we make extensive use of these statistics for analyzing MIMO system performance in correlated fading environments. We note the difference between the covariance in (2.39) and (2.45), which results from the special notation for matrix variate distributions.

It is obvious that the statistical MIMO channel model with Kronecker product covariance structure has fewer degrees of freedom than a model with arbitrary covariance. The model implies that the covariance of the receive vector channel is the same, no matter what transmit antenna excites the channel, i.e. if we decompose the channel matrix in column vectors

\[
H = \begin{bmatrix} h_1 & h_2 & \ldots & h_T \end{bmatrix},
\]  

then we find for the vector channels \( h_k \) (SIMO channel from transmit antenna \( k \) to all receive antennas)

\[
E[h_k h_k^H] = R_{RX} \quad \forall k \in \{1, \ldots, T\}
\]  

or equivalently in matrix notation we find via expected value (11.33)

\[
E[HH^H] = \text{tr}(R_{TX}^*) \cdot R_{RX}.
\]  

Similar results hold for the transmit side. With the row-wise decomposition

\[
H = \begin{bmatrix} g_1^H \\ g_2^H \\ \vdots \\ g_k^H \end{bmatrix}
\]

we get for the vector channels \( g_k \) (MISO channel from all transmit antennas to receive antenna \( k \))

\[
E[g_k g_k^H] = R_{TX} \quad \forall k \in \{1, \ldots, R\}
\]  

and in matrix notation

\[
E[H^H H] = \text{tr}(R_{RX}^*) \cdot R_{TX}.
\]

In practice, the Kronecker covariance channel model is accurate if the DOA spectrum at the antenna elements of the receive array is invariant to the choice of transmit antenna that excites the channel (leading to (2.47)). With a large distance between transmitter and receiver, and a small distance between the transmit antenna elements with the same radiation patterns, these conditions are met. Similar considerations lead to (2.50). An example is depicted in Fig. 2.3, where the scatterer clusters are concentrated around the receiver and the transmitter. Assuming a large distance between the two link ends, each transmitter excites the receive antenna array in the same way.
Obviously, the spatial distribution of signal energy at the receiver is independent of the DOD spectrum at the transmitter, i.e. it is invariant to spatial filtering at the transmitter - essentially transmitter and receiver are spatially decoupled [200].

Normalization

In order to allow for fair comparisons between various propagation environments with different channel correlation properties, we have to normalize the channel. Throughout this thesis we set

\[ E[||H||^2] = E[\text{tr}(HH^H)] = T \cdot R. \]  \hspace{1cm} (2.52)

This can be assured by normalizing the correlation matrices at receiver and transmitter such that they carry ones on their main diagonal

\[ \text{diag}(R_{TX}) = 1_{T \times 1}, \quad \text{diag}(R_{RX}) = 1_{R \times 1} \]  \hspace{1cm} (2.53)

with the obvious trace constraint

\[ \text{tr}(R_{TX}) = T, \quad \text{tr}(R_{RX}) = R. \]  \hspace{1cm} (2.54)

Now we get from (2.52) together with (2.41) via expected value (11.33)

\[ E[\text{tr}(A^H H^H w B B^H H^H w A)] = \text{tr}(E[H^H w R_{TX} H^H w] \cdot R_{RX}) = \text{tr}(\text{tr}(R_{TX}) \cdot I \cdot R_{RX}) = T \cdot R \]  \hspace{1cm} (2.55)

as required.

2.2.5  Note on the flat fading channel assumption

The focus of this work is on the investigation of the effects of fading correlation between the antenna elements of wireless MIMO systems. In order to avoid unnecessary distraction from the main problem, we aim for an abstract system and channel model that is as simple as possible, while simultaneously retaining the main properties of real wireless environments. For that reason, we choose the flat fading channel model introduced above, which turns out to be no major con-
constraint. Specifically, it is well known [40][154] that frequency selective MIMO channels can be transformed into a set of parallel and independent flat fading MIMO channels by appropriate transmit and receive processing like in orthogonal frequency division multiplexing (OFDM) systems [143][197]. This transformation is known to preserve capacity and without going into the details, we note that fading correlation between the antenna elements translates into correlated complex Gaussian distributed MIMO channel matrices on the subcarriers. For each of the independent subcarriers, we can therefore apply the analysis presented in this thesis.

2.2.6 Determination of the receive and transmit correlation matrices

In the following, we present two principal methods that are used throughout this thesis for determining the correlation matrices at transmitter and receiver. First, we give a model that explicitly takes into account various propagation parameters like direction of arrival, angular spread etc. This model can be used for the simulation of 'real world' MIMO systems and will be denoted by realistic correlation matrix model (RCMM). On the other hand, we outline an abstract model, the so-called exponential correlation matrix model (ECCM) [122], which can be characterized by a single parameter and is thus more suited for a theoretical analysis.

Propagation based realistic correlation matrix model (RCMM)

The correlation matrices at transmit and receive side of the MIMO link depend on a number of parameters. One one hand, they depend on the specific characteristics of the antenna array (which influence the array response vectors), like array shape (ULA, irregular arrays, circular arrays, 2-D arrays, 3-D arrays), antenna element radiation patterns, and array element spacing. On the other hand, they are determined by the spatial angular spectrum, i.e. the main DOAs and DODs, and the angular power distribution around these main directions, characterized by the distribution type (e.g. uniform, Laplace) and the (root mean square) angular spread (AS). In this thesis, the angular spread at the transmitter is denoted by $\Delta_{TX}$ and at the receiver equivalently by $\Delta_{RX}$. In general, there are no closed-form solutions available for a direct calculation of the correlation matrices from above parameters. However, $R_{TX}$ and $R_{RX}$ can be determined in a straightforward manner via Monte-Carlo integration from the definition in (2.35). We will refer to this specific model as realistic correlation matrix model (RCMM) throughout this thesis.

Abstract exponential correlation matrix model (ECMM)

A major advantage of the exponential correlation matrix model (ECMM) is the fact that it can be characterized by a single parameter (see e.g. [122]). Specifically, we let the correlation matrices at transmitter and receiver be given by

$$[R_{RX/TX}]_{i,j} = r_{RX/TX}^{\frac{|i-j|}{j}} \quad \left| r_{RX/TX} \right| \leq 1 \quad (2.56)$$

with parameter $r_{RX}$ at the receiver side and $r_{TX}$ at the transmitter. The reader can readily check that the correlation matrices in (2.56) have ones on their main diagonal and thus fulfill the standard normalizations (2.53)(2.54) used throughout this work. Moreover, the correlation between
two neighboring antenna elements is given by $r_{\text{RX/TX}}$ and the correlation decays exponentially with the distance between two antenna elements.

### 2.3 System and Signal Model

In order to obtain concise mathematical expressions, all system and signal models in this work are expressed in equivalent baseband notation. Moreover, the focus is on the statistical analysis of uncoded systems with memoryless modulation, such that the memory of the wireless channel has not to be taken into account and the time index in all equations can be dropped.

#### 2.3.1 Flat fading signal model

We consider a flat fading MIMO link modeled by

$$
\mathbf{y} = \mathbf{H}\mathbf{F}\mathbf{s} + \mathbf{n},
$$

(2.57)

where $\mathbf{s}$ is the $L \times 1$ TX symbol vector, i.e. there are $L$ independent subchannels (data streams). $\mathbf{F}$ is a $T \times L$ linear matrix transmit prefilter that maps the $L$ subchannels on the $T$ transmit antennas, whereas $L \leq T$. Similar to a MISO beamforming system, $\mathbf{F}$ can be interpreted as a separate beamformer for each subchannel. Figuratively speaking, the prefilter $\mathbf{F}$ can be used to transmit the various data streams into different spatial directions. By proper choice of $\mathbf{F}$, it is possible to improve transmission quality e.g. in terms of BER/SER or channel capacity. If the prefilter is missing, we just have to set $\mathbf{F} = \mathbf{I}$. $\mathbf{H}$ is the $R \times T$ MIMO channel matrix with correlated Rayleigh fading elements, $\mathbf{n}$ is the $R \times 1$ noise vector, and $\mathbf{y}$ is the $R \times 1$ receive vector (see Fig. 2.4). By $R$ we denote the number of receive antennas. Based on the noisy receive vector $\mathbf{y}$, the receiver can reconstruct the transmit vector $\mathbf{s}$ via linear or nonlinear algorithms like e.g. zero-forcing (ZF), minimum mean squared error (MMSE), successive interference cancellation (SIC) or maximum likelihood (ML).

Moreover, we introduce the $L \times L$ transmit signal covariance matrix

$$
\mathbf{R}_{ss} = E[\mathbf{s}\mathbf{s}^H] = E_s \cdot \mathbf{\tilde{R}}_{ss},
$$

(2.58)

where $E_s$ is the mean energy per transmit symbol and the normalized matrix $\mathbf{\tilde{R}}_{ss}$ with
\[ \text{tr}(\tilde{R}_{ss}) = L. \] (2.59)

For white transmit signals we simply get
\[ R_{ss} = E_s \cdot I_L. \] (2.60)

In presence of a transmit prefilter \( F \), it is straightforward to show that the system can be reduced to an equivalent system without prefilter and transmit signal covariance matrix given by
\[ R_{ss,F} = E_s \cdot F \tilde{R}_{ss} F^H. \] (2.61)

For clarity, without loss of generality, for systems with linear transmit prefilter we set \( \tilde{R}_{ss} = I_L \) and the resulting equivalent transmit signal covariance reads
\[ R_{ss,F} = E_s \cdot FF^H. \] (2.62)

We always normalize the overall transmit power such that
\[ \text{tr}(F \tilde{R}_{ss} F^H) = \text{tr}(FF^H) = \rho = T, \] (2.63)

where \( \rho \) is a transmit power constraint (see below). On the other hand, the noise covariance matrix reads
\[ R_{nn} = E[nn^H] = N_0 \cdot \tilde{R}_{nn} \] (2.64)

with mean noise power per receive antenna \( N_0 \) and normalized matrix \( \tilde{R}_{nn} \) with
\[ \text{tr}(\tilde{R}_{nn}) = R. \] (2.65)

For additive white Gaussian noise (AWGN) we find
\[ R_{nn} = N_0 \cdot I_R. \] (2.66)

Note again that in this thesis we statistically analyze uncoded systems with memoryless modulation. Possible memory effects of the channel due to Doppler fading do not influence the performance analysis in these cases (we mention, that a long interleaver in combination with a fast fading channel (high Doppler spread) is capable of counteracting those effects). Therefore, in order to get a concise mathematical representation, all time indices have been dropped from (2.57).

### 2.3.2 Channel state information

Throughout this thesis, we assume that the instantaneous channel matrix \( H \) is perfectly known at the receiver. This assumption is a simplification for making the analysis of system performance tractable. In practical systems, however, accurate channel estimation (CE) is a critical aspect [130][182]. It is typically supported by cyclic transmission of training sequences in the data stream [8][15], which simultaneously sacrifices spectral efficiency. Therefore, efficient CE schemes are indispensable for a successful operation of MIMO systems. Again, channel correла-
tion also influences CE, however, we note that in contrast to data transmission, fading correlation between the antenna elements can beneficially be exploited for improving the quality of the MIMO channel estimate [106][107].

On the other hand, we differentiate between varying degrees of channel state information (CSI) at the transmitter. For low complexity and standard space time coding applications, it is a common assumption that the transmitter is devoid of any CSI. However, in systems with moderate mobility, it is possible to provide the transmitter with statistical information on the correlation properties of the channel, e.g. via a low rate feedback link or via frequency transformation of the transmit correlation matrix in frequency division duplex (FDD) systems [75]. We refer to this scenario as channel distribution information at the transmitter (CDIT), which will be in the focus of this work. Finally, with the reciprocity in uplink and downlink of low mobility time division duplex (TDD) systems or alternatively via a fast feedback link, it is possible in certain situations to make the full instantaneous CSI (i.e. the channel matrix $H$) available at the transmitter. In the following studies we always explicitly state, which scenario we are referring to.

### 2.3.3 Equivalent systems

In this paragraph, we show that MIMO systems in a Rayleigh fading environment with arbitrary channel fading correlation at receiver and transmitter according to the Kronecker correlation model, arbitrary linear prefilter, and arbitrary noise covariance matrix can be reduced to an equivalent system without prefilter and AWGN. This may be exploited to simplify notation without loss of generality and throughout this thesis we make use of this fact.

Now consider (2.57) together with the channel model (2.41)

$$y = HF s + n = A^H H_\nu B F s + n.$$  \hfill (2.67)

With a noise whitening filter $\tilde{R}_{nn}^{-1/2}$ at the receiver it can be shown that with $\tilde{R}_{ss} = I_L$

$$\tilde{y} = \tilde{R}_{nn}^{-1/2} A^H H_\nu B F s + \tilde{n} = \tilde{A}^H \tilde{H}_\nu \tilde{B} s + \tilde{n}.$$  \hfill (2.68)

with the AWGN vector $\tilde{n}$, $E[\tilde{n}\tilde{n}^H] = N_0 \cdot I_R$, $R \times R$ matrix $\tilde{A}^H = \tilde{R}_{nn}^{-1/2} A^H$, $L \times L$ matrix $\tilde{B} = \Sigma_L \cdot V^H$, the $R \times L$ matrix of complex i.i.d. Gaussian distributed entries $\tilde{H}_w$, and the singular value decomposition (SVD)

$$BF = U \cdot \begin{bmatrix} \Sigma_L \\ 0_{(T-L) \times L} \end{bmatrix} \cdot V^H.$$  \hfill (2.69)

with $L \times L$ diagonal matrix $\Sigma_L$ of singular values, unitary $T \times T$ matrix $U$, and unitary $L \times L$ matrix $V$. It is obvious from (2.68) that we have established an equivalent system to (2.67) with receive correlation modeled by $\tilde{A}^H$, transmit correlation modeled by $\tilde{B}$, and AWGN modeled by $\tilde{n}$. 


2.3.4 Note on channel coding

As was emphasized above, in the analysis of MIMO zero-forcing (ZF), minimum mean squared error (MMSE), and maximum likelihood (ML) receivers, we consider systems without channel coding in this thesis. The presence of channel coding leads to an additional complexity increase of the performance analysis and heavily depends on the deployed coding scheme. For example, coding can be applied on a per subchannel basis (simple time dimension codes) or alternatively across spatial subchannels (space time codes [141][142][184][185][186]). Moreover, the transmitter can apply simple linear block codes (e.g. the Alamouti scheme [4]), convolutional codes [9][195], or coded modulation [11]. At the receiver, the decoding algorithm can deploy hard or soft decisions with possible iterative decoding (so-called turbo decoding [10][61]). While the analysis of the linear Alamouti scheme is straightforward, even for single-input single-output (SISO) systems an assessment of the performance of more advanced coding schemes soon becomes hardly tractable. This is especially true for systems that deploy iterative detection due to the highly nonlinear structure of those systems.

Having in mind above comments, it is not surprising that to the authors’ best knowledge there are no general procedures known in literature that allow for an exact analysis of space-time coded systems. However, in special situations it appears to be possible to derive analytical performance results, which could yield interesting insights that support the design process of coded MIMO systems and on the other hand could serve as bounds for more sophisticated coding schemes. For example, in case of per subchannel coding, the spatial subchannels can be separated at the receiver via low complexity linear interfaces like ZF and MMSE. In this thesis, we provide expressions for the subchannel signal-to-noise ratio (SNR) statistics for these receiver types, which can be used for analyzing the bit error rate (BER) of the coded subchannel data streams. Still, a strong background in (SISO) channel coding theory appears to be necessary for carrying out such an analysis and is therefore beyond the scope of this thesis. Nevertheless, we emphasize that a statistical analysis of space-time coded systems opens up a fruitful field for future research.

2.3.5 Two eigenvalue matrix definitions

In the statistical analysis of capacity and symbol error probability of various MIMO systems it turns out that many results can be expressed in terms of the eigenvalues of certain covariance and correlation matrices. These eigenvalues are explicitly defined in this paragraph to allow for a concise reference in later chapters. First, we define the $R \times R$ diagonal matrix of eigenvalues associated with the receive side

$$O \equiv \text{eig}(\tilde{R}_{nn}^{-1}R_{RX}) = \text{diag}(o_1, \ldots, o_R),$$

(2.70)

which comprises the effects of receive fading correlation and colored additive Gaussian noise. On the transmit side, we define the $L \times L$ matrix of eigenvalues

$$S \equiv \text{eig}(\tilde{R}_{ss}F^H R_{TX}F) = \text{diag}(s_1, \ldots, s_L),$$

(2.71)
which takes into account fading correlation at the transmit antenna array as well as the linear transmit prefilter $F$.

### 2.3.6 Definition of the signal to noise ratio

In this thesis, we use the mean SNR per transmit symbol definition

$$
\gamma = \frac{E_s}{N_0}.
$$

(2.72)

For simulations and asymptotical considerations of ergodic capacity we normalize the noise power with respect to the overall transmit power, such that the SNR in dB is defined by

$$
\gamma_{dB} \equiv 10 \cdot \log_{10} \frac{\rho \cdot E_s}{N_0},
$$

(2.73)

where $\rho$ is the transmit power constraint and throughout the thesis we set

$$
\rho = T.
$$

(2.74)

Plugging (2.72) in (2.73) we get

$$
\gamma_{dB} = 10 \cdot \log_{10} \rho \cdot \gamma,
$$

(2.75)

resulting in the equality

$$
\gamma = \frac{1}{\rho} \cdot 10^{10}. (2.76)
$$

On the other hand, for SER simulations of various receiver types, we use the common alternative SNR in dB definition

$$
\tilde{\gamma}_{dB} \equiv 10 \cdot \log_{10} \frac{T \cdot E_b}{N_0} = 10 \cdot \log_{10} \frac{\rho \cdot E_b}{N_0},
$$

(2.77)

where $E_b$ is the energy per information bit. Obviously, the two SNR scales in (2.73) and (2.77) differ just by a constant shift. Note that in the simulations of this thesis, for simplicity we label diagrams with 'SNR [dB]', thereby implicitly referring to the definitions in (2.75) and (2.77).

### 2.4 Simulation Parameters

If not stated otherwise, without loss of generality according to Paragraph 2.3.3 on equivalent systems, in all simulations of this work we study systems with white input signals of power $E_s$ and additive white Gaussian noise of variance $N_0$, i.e.

$$
R_{ss} = E_s \cdot I \quad R_{nn} = N_0 \cdot I.
$$

(2.78)
Following the channel model proposals for the third generation UMTS system [207], for simulations with propagation based correlation matrices (RCMM) we always assume a single main direction of departure and arrival, respectively, of 20 degrees with respect to the array perpendicular at receiver and transmitter. The array element spacing is 0.5 wavelengths, which is a standard element spacing for smart antenna beamforming arrays. Furthermore, there is a Laplacian power distribution over the angular spread (AS). As was stated above, $R_{RX}$ and $R_{TX}$ can be determined via Monte-Carlo integration according to these assumptions with the help of (2.35).

As will be demonstrated in later chapters, most of the performance measures like SER and ergodic capacity depend exclusively on the eigenvalues of the correlation matrices at transmitter and receiver, whereas a higher spread of the eigenvalues has a negative impact. To give an impression of the influence of the propagation scenario on the eigenvalues of the correlation matrices, with the assumptions above note that for 2°, 10°, and 30° AS in the case of 2 antenna elements we find

$$\text{eig}(R_{2^\circ}, 2 \times 2) = \begin{bmatrix} 1.9947 & 0.0053 \\ 0.0053 & 1.9947 \end{bmatrix}$$

$$\text{eig}(R_{10^\circ}, 2 \times 2) = \begin{bmatrix} 1.8872 & 0.1128 \\ 0.1128 & 1.8872 \end{bmatrix}$$

$$\text{eig}(R_{30^\circ}, 2 \times 2) = \begin{bmatrix} 1.4919 & 0.5081 \\ 0.5081 & 1.4919 \end{bmatrix}$$

(2.79)

With 4 antenna elements we get

$$\text{eig}(R_{2^\circ}, 4 \times 4) = \begin{bmatrix} 3.9488 & 0.0507 & 0.0005 & 0.0000 \\ 0.0507 & 3.9488 & 0.0005 & 0.0000 \\ 0.0005 & 0.0005 & 3.9488 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 3.9488 \end{bmatrix}$$

$$\text{eig}(R_{10^\circ}, 4 \times 4) = \begin{bmatrix} 3.2310 & 0.6444 & 0.1119 & 0.0127 \\ 0.6444 & 3.2310 & 0.1119 & 0.0127 \\ 0.1119 & 0.1119 & 3.2310 & 0.0127 \\ 0.0127 & 0.0127 & 0.0127 & 3.2310 \end{bmatrix}$$

(2.80)

$$\text{eig}(R_{30^\circ}, 4 \times 4) = \begin{bmatrix} 1.9402 & 1.1322 & 0.6362 & 0.2918 \\ 1.1322 & 1.9402 & 0.6362 & 0.2918 \\ 0.6362 & 0.6362 & 1.9402 & 0.2918 \\ 0.2918 & 0.2918 & 0.2918 & 1.9402 \end{bmatrix}$$

and finally with 6 antenna elements we have

$$\text{eig}(R_{2^\circ}, 6 \times 6) = \begin{bmatrix} 5.8282 & 1.673 & 0.0044 & 0.0001 & 0.0000 & 0.0000 \\ 1.673 & 5.8282 & 0.0044 & 0.0001 & 0.0000 & 0.0000 \\ 0.0044 & 0.0044 & 5.8282 & 0.0044 & 0.0001 & 0.0000 \\ 0.0001 & 0.0001 & 0.0044 & 5.8282 & 0.0001 & 0.0000 \\ 0.0000 & 0.0000 & 0.0001 & 0.0001 & 5.8282 & 0.0044 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0044 & 5.8282 \end{bmatrix}$$

$$\text{eig}(R_{10^\circ}, 6 \times 6) = \begin{bmatrix} 4.1460 & 1.3132 & 0.4020 & 0.1097 & 0.0252 & 0.0039 \\ 1.3132 & 4.1460 & 0.4020 & 0.1097 & 0.0252 & 0.0039 \\ 0.4020 & 0.4020 & 4.1460 & 1.3132 & 0.4020 & 0.1097 \\ 0.1097 & 0.1097 & 1.3132 & 4.1460 & 0.4020 & 0.1097 \\ 0.0252 & 0.0252 & 0.4020 & 0.4020 & 4.1460 & 1.3132 \\ 0.0039 & 0.0039 & 0.1097 & 0.1097 & 1.3132 & 4.1460 \end{bmatrix}$$

(2.81)

It becomes obvious from (2.79)-(2.81) that the propagation scenario with only 2 degrees of angular spread results in correlation matrices, where the smallest eigenvalues almost disappear. This leads to a rank reduction of the MIMO channel matrix, thus reducing the number of available spatial dimensions.
3 MIMO Mutual Information and Capacity

Before analyzing practical receiver and transmitter structures in the following chapters, we first put the focus on an information theoretical study of the correlated Rayleigh fading MIMO channel. In particular, we consider the information theoretic limit of achievable data rates with error-free reception over the wireless MIMO channel, which is given by the maximum of mutual information between the input and the output vector of the MIMO channel with respect to the distribution of the random input [24][37][168]. However, while it is well known [187] that the capacity achieving transmit vector distribution is zero-mean Gaussian, adaptation of the input distribution parameters (namely the covariance matrix) depends on the availability of adequate channel state information (CSI) at the transmitter and we have to differentiate between various cases. In the optimal case, with full short-term (ST) CSI the transmitter can adapt to each specific channel fading state by instantaneous waterfilling [34]. On the other hand, if only statistical long-term (LT) CSI is available, which will be denoted by channel distribution information at the transmitter (CDIT) in the following, the optimization of mutual information reduces to finding an optimal covariance matrix of the input signal vector [36][187]. The latter case will be in the focus of this thesis. Finally, without CSI the transmitter equally distributes the available transmit power to the transmit antennas, i.e. the transmit covariance is a scaled identity matrix [187]. We emphasize again that we assume that ideal short-term channel state information is available at the receiver in all cases.

Being a function of the random channel coefficients, MIMO channel mutual information itself is a random variable (RV), i.e. it can fluctuate considerably around its mean value. If the MIMO channel coefficients are in a deep fade or the channel matrix is ill-conditioned, the instantaneous mutual information can even be close to zero. Therefore, we can characterize the RV mutual information by its probability density function (PDF), cumulative distribution function (CDF), complementary cumulative distribution function (CCDF), moments, and moment generating function (MGF). However, due to the considerable mathematical complexity of the problem, up to now there are no results available on the exact distribution of MIMO mutual information, including the simplest i.i.d. Rayleigh fading case. Most of available research results focus on ergodic capacity (EC), which in MIMO literature is often just referred to as capacity. It is the first moment, i.e. the mean value of mutual information averaged over the channel fading statistics, thus resulting in a good characterization of adaptively modulated space-time coded MIMO systems with large interleaver sizes and fast fading, such that the code captures a sufficient number of fading periods to achieve an averaging effect. On the other hand, in a system with short code length and slow fading, outage capacity is a better measure for describing the behavior. In fact, it was shown in [116] that there is a direct relation between outage capacity and the block error rate of coded systems with block fading channels. However, outage capacity is essentially the CCDF of mutual information and the analysis is by far more involved than in case of EC. Therefore, many publications concentrate on Monte-Carlo characterization of outage capacity. A different approach that yields tractable analytical results is the approximation of the distribution of mutual information by standard distributions like Gaussian [177] or Gamma [88] that can be character-
ized by their first and second moments, which in certain cases can be derived analytically. A
detailed overview of existing literature and results on MIMO capacity is provided in [46].

In this thesis, for the first time to the authors’ best knowledge we provide an exact analysis of the
MGF of mutual information for Rayleigh fading MIMO channels with fading correlation at
receiver and transmitter based on sophisticated tools from multivariate statistics, thus generaliz-
ing and unifying existing results in literature. The novel approach is based on a finite antenna ele-
ment number assumption, i.e. in contrast to many existing publications it is non-asymptotic and
exact for an arbitrary number of transmit and receive antennas. Before going into a detailed derive-
tion, we give a short synopsis of available literature and known approaches on MIMO capacity
analysis, in particular the calculation of MIMO channel ergodic and outage capacity.

In his seminal paper [187], Telatar calculated the ergodic capacity of a MIMO link with uncorre-
lated Rayleigh fading in an additive white Gaussian noise (AWGN) environment as a single inte-
gral by integrating over the eigenvalue PDF of certain complex Wishart matrices
[79][80][81][118][126][179], thereby predicting enormous capacity gains by combined spatial
processing at transmitter and receiver, thus initiating immense research activities in this area.
Later, Foschini and Gans presented numerical results and bounds on i.i.d. Rayleigh MIMO outage
capacity in their fundamental work [36]. Bounds on the ergodic capacity of i.i.d. and correlated
Rayleigh fading MIMO channels were given in [52][76][121][170] in a mathematically more
tractable form than the expressions resulting from the exact analysis. Another approach is an
asymptotical analysis, where the number of antenna elements at receiver and transmitter goes to
infinity, however has a fixed ratio. In this case, the empirical eigenvalue distribution of the ran-
don channel matrix can be obtained in a manageable form [171], such that the capacity expres-
sions simplify and it turns out that the results are a good approximation even for practically
relevant systems with a fairly small number of antennas. Asymptotic results can be found for i.i.d.
Rayleigh fading in [155] and for the one-side correlated case (i.e. transmit or receive correlation
exclusively) in [20], where empirical eigenvalue PDFs of certain large dimensional random matri-
ces are used [129][171]. The same techniques were earlier successfully applied in the analysis of
multi user code division multiple access (CDMA) wireless systems [193].

A non-asymtotic yet very powerful approach makes use of the MGF of mutual information. It is
well known that a MGF fully specifies a distribution and thus all its moments, i.e. once the MGF
of Shannon capacity is found, we can e.g. determine EC, outage capacity approximations, or use
numerical MGF inversion techniques to find the PDF and CDF (indeed, the calculation of
moments requires the differentiation of the MGF, which is in principle always possible, but the
resulting expressions can soon become hardly tractable). Specifically, by integrating over the
eigenvalue PDF of an i.i.d. complex Wishart matrix, in [199] the MGF of the mutual information
of an i.i.d. Rayleigh channel is derived. Based on the MGF, numerical Laplace transform inver-
sion techniques are used for calculating outage capacity [73]. A similar MGF approach is taken in
[87], where the authors present results for various propagation scenarios including e.g. i.i.d. and
one-side correlated Rayleigh fading, as well as Ricean fading. Recently, similar results for the
case of a Rayleigh fading channel with fading correlation at one side of the MIMO link were
obtained in [16] and [178]. Again, a mathematically challenging integration over the eigenvalue
PDF of certain (non-central) Wishart matrices [151] is necessary, which requires the use of certain integration results from statistical literature [86][92][95]. However, we emphasize that this eigenvalue based approach also prohibits a general solution for the MGF in Rayleigh fading with both receive and transmit correlation, as there are no compact formulas available for the eigenvalue PDF of certain complex generalized random quadratic forms (see below).

In this thesis, we present a novel approach on the calculation of the MGF of mutual information of MIMO channels with correlated Rayleigh fading. A concise mathematical formulation of the MGF is given in terms of a hypergeometric function of matrix arguments [70][81] and in terms of determinants of scalar hypergeometric functions. In contrast to existing literature [87][199], our approach is not based on eigenvalue PDFs but uses a direct integration technique. Having available a closed-form MGF, it is possible to derive exact, i.e. non-asymptotic moments, including ergodic capacity (which is in the focus of this thesis) for arbitrary array sizes and arbitrary correlation properties at receiver as well as transmitter, thus unifying and completing existing partial solutions for special propagation scenarios [108][109].

Based on the exact analytical expressions for ergodic capacity, we also derive formulas for the asymptotics in the low and high SNR region, which allow for a simple and concise characterization of the influence of correlation [111]. Moreover, we present asymptotical results in the limit of a large number of transmit antennas with a finite number of receive antennas or alternatively in the limit of a large number of receive antennas with a finite number of transmit antennas. Specifically, we study the so-called 'channel-hardening' [71] effect for correlated MIMO channels, which basically states that certain statistics of the fading MIMO channel become deterministic in the limit of a large number of antenna elements at one side of the MIMO link.

We start in this chapter with the capacity analysis of MIMO systems with uninformed transmitters, i.e. for cases, where the transmitter does not have any channel state information. However, we demonstrate that the expressions derived for ergodic capacity of MIMO systems with uninformed transmitters can easily be adapted to take into account CDIT and we present a statistical waterfilling algorithm that achieves capacity by proper adaptation of the covariance matrix of the transmit vector or equivalently the transmit prefilter.

### 3.1 MIMO Mutual Information

In this paragraph we first outline the definition of mutual information of the wireless MIMO channel and introduce a notational framework that describes systems with arbitrary correlation properties at receiver and transmitter with arbitrary (finite) number of transmit and receive antennas. The notation is consistently used for all following derivations, where we calculate the moment generating function (MGF) of mutual information (which is a RV due to the random nature of the wireless channel) in terms of a concise hypergeometric function of matrix argument expression [81]. Based on the matrix variate formulation, we then find an expression for the MGF in terms of well-known scalar hypergeometric functions [1] by exploiting some results given independently in statistical literature [57] and [95].
3.1.1 General expressions

In this paragraph we introduce a unifying notation that greatly simplifies all following derivations and give definitions for MIMO mutual information and capacity. For the given system model, the transmission of the transmit vector \( s \) over the flat Rayleigh fading channel \( H \) with transmit prefilter \( F \) in the presence of additive Gaussian noise can be described by (see Paragraph 2.3.1)

\[
y = HF\hat{s} + n.
\]  

(3.1)

It is well known [34][187] that the mutual information \( I(s, y) \) between input vector \( s \) and output \( y \) of the MIMO channel is given by

\[
I(s, y) = \log_2 \left| I + R_{ss} F^H H^H R_{nn}^{-1} HF \right|.
\]  

(3.2)

Using the normalized noise and signal covariance matrices (2.58)(2.64) for explicitly emphasizing the signal to noise ratio (SNR) dependence of mutual information, we can rewrite (3.2) as

\[
I(s, y) = \log_2 \left| I + \gamma \cdot \tilde{F}\tilde{R}_{ss} F^H H^H \tilde{R}_{nn}^{-1} H \right|
\]  

(3.3)

with the standard mean SNR per transmit symbol definition \( \gamma = \frac{E_s}{N_0} \). Plugging the channel model (2.41) with Kronecker product covariance structure in (3.3), we find

\[
I(s, y) = \log_2 \left| I + \gamma \cdot \tilde{F}\tilde{R}_{ss} F^H B^H \tilde{H}^H \tilde{A} \tilde{R}_{nn}^{-1} A^H H_w B \right|.
\]  

(3.4)

In the following, we reduce (3.4) to a concise equivalent formulation that allows for a unified analysis of correlated MIMO systems. Once again we emphasize that we assume full rank channel correlation and signal and noise covariance matrices in this thesis. An extension is straightforward but would unnecessarily complicate notation, thus detracting from the main problems. By noticing that the distribution of the i.i.d. complex Gaussian matrix \( H_w \) is invariant to left- or right multiplications with unitary matrices \( U \) and \( V \), i.e.

\[
U H_w V \equiv H_w,
\]  

(3.5)

we find with definitions (2.70) for correlation matrix \( S \) at the transmitter and (2.71) for correlation matrix \( O \) at the receiver, the comments on equivalent systems in Paragraph 2.2.3, the \( R \times L \) matrix of i.i.d. complex Gaussian elements \( \tilde{H}_w \), and determinant identity (11.14)

\[
I(s, y) \equiv \log_2 \left| I + \gamma \cdot S H_w^H O \tilde{H}_w \right|,
\]  

(3.6)

where we have used the fact that the non zero eigenvalues of two matrix products agree

\[
\text{eig}_0(XY) = \text{eig}_0(YX)
\]  

(3.7)

for two general matrices \( X \) and \( Y \). Note that with determinant identity (11.14) we can reformulate

\[
I(s, y) \equiv \log_2 \left| I + \gamma \cdot S H_w^H O \tilde{H}_w \right| = \log_2 \left| I + \gamma \cdot O \tilde{H}_w S H_w^H \right|.
\]  

(3.8)
We rewrite (3.8) such that the matrix argument $\mathbf{S} \mathbf{H}_w^H \mathbf{O} \tilde{\mathbf{H}}_w$ or $\mathbf{O} \tilde{\mathbf{H}}_w \mathbf{S} \mathbf{H}_w^H$, respectively, of the determinant is of full rank, thereby simplifying the subsequent statistical analysis. Now we define
\[
\mu \equiv \min(R, L) \quad \nu \equiv \max(R, L)
\]
the $\mu \times \mu$ diagonal matrix
\[
\Sigma = \begin{cases} 
\mathbf{S} & L \leq R \\
\mathbf{O} & L > R 
\end{cases} \quad \Sigma = \text{diag}(\sigma_1, \ldots, \sigma_{\mu})
\]  \(3.10\)
the $\nu \times \nu$ diagonal matrix
\[
\Omega = \begin{cases} 
\mathbf{O} & L \leq R \\
\mathbf{S} & L > R 
\end{cases} \quad \Omega = \text{diag}(\omega_1, \ldots, \omega_{\nu})
\]  \(3.11\)
and the $\nu \times \mu$ matrix of i.i.d. complex Gaussian entries $\mathbf{G}$. With above definitions we can introduce a unifying expression for MIMO mutual information, which will serve as a basis for all following derivations
\[
I(s, y) \equiv \log_2 \left| \mathbf{I} + \gamma \cdot \Sigma \mathbf{H}^H \Omega \mathbf{G} \right|.
\]  \(3.12\)
Furthermore, it turns out in later calculations that it is advantageous to consider MIMO systems with the same number of independent subchannels and receive antennas, such that $\nu = \mu$. To this end, we can artificially set (with $\varepsilon_k$ taking on the value 0 in the limit for all $k$)
\[
\Sigma(\varepsilon) \equiv \text{diag}(\sigma_1, \ldots, \sigma_{\mu}, \varepsilon_{\nu-\mu-1}, \ldots, \varepsilon_0) = \text{diag}(\tilde{\sigma}_1(\varepsilon), \ldots, \tilde{\sigma}_\nu(\varepsilon))
\]  \(3.13\)
Furthermore, we can introduce an enlarged $\nu \times \nu$ matrix $\tilde{\mathbf{G}}$ of complex i.i.d. Gaussian elements. Without loss of generality, we can thus write from (3.12) together with (3.13) and the vector notation $\varepsilon = (\varepsilon_{\nu-\mu-1}, \ldots, \varepsilon_0)^T$
\[
I(s, y) \equiv \lim_{\varepsilon \to 0} \log_2 \left| \mathbf{I} + \gamma \cdot \Sigma(\varepsilon) \tilde{\mathbf{G}}^H \Omega \tilde{\mathbf{G}} \right| = \lim_{\varepsilon \to 0} \log_2 \left| \mathbf{I} + \mathbf{Q} \right|,
\]  \(3.14\)
which means that we can in the general case consider $\nu \times \nu$ MIMO systems with a subsequent limiting process. Obviously, (3.14) is a function of
\[
\mathbf{Q} = \gamma \cdot \Sigma(\varepsilon) \tilde{\mathbf{G}}^H \Omega \tilde{\mathbf{G}},
\]  \(3.15\)
where $\mathbf{Q}$ is a $\nu \times \nu$ random matrix quadratic form of complex normal distributed matrices $\tilde{\mathbf{G}}$. In the next paragraph we study some statistical results on random matrix quadratic forms.

### 3.1.2 Distribution of a complex generalized matrix quadratic form

The PDF of a complex generalized matrix random quadratic form $\mathbf{S}$ of size $m \times m$ and parameter $n \geq m$ has been derived in [94]
\[
p(S) = \frac{e^{-\text{tr}(-q^{-1}M^{-1}S)}}{\Gamma_m(n) \cdot |M|^n \cdot |N|^m} \cdot \mathcal{F}_0^{(n)}(I_n - qN^{-1}; -q^{-1}M^{-1}S) \cdot (dS),
\]

where \( N \) is a deterministic \( n \times n \) covariance matrix, \( M \) is a deterministic \( m \times m \) covariance matrix, and for clarity we have included \( \mathcal{D} \), which is the matrix differential of elements of \( S \) \[59][140], and \( q \) is an arbitrary scalar constant. Using this result, the PDF of \( Q \) in (3.15) reads
\[
p(Q) = \frac{e^{-\text{tr}(-q^{-1}(\gamma \Sigma)^{-1}Q)}}{\Gamma_v(\nu) \cdot |\gamma \Sigma|^{\nu} \cdot |\Omega|^{\nu}} \cdot \mathcal{F}_0^{(n)}(I_v - q\Omega^{-1}; -q^{-1}(\gamma \Sigma)^{-1}Q) \cdot (dQ).
\]

In (3.16) and (3.17) we use the function \( \mathcal{F}_0^{(n)} \), which is a hypergeometric function of 2 matrix arguments \[23][56][58][70][81][132][134], where \( n \times n \) is the maximum dimension of the matrix arguments and in general we have the following series definition with two \( n \times n \) matrices \( X \) and \( Y \) \[81, equation (88)]
\[
p\mathcal{F}_q^{(n)}(a_1, \ldots, a_p; b_1, \ldots, b_q; X, Y) = \sum_{\kappa} \sum_{k=0}^{\infty} \frac{[a_1]_{\kappa} \cdots [a_p]_{\kappa}}{[b_1]_{\kappa} \cdots [b_q]_{\kappa}} \cdot \frac{\tilde{C}_{\kappa}(X) \cdot \tilde{C}_{\kappa}(Y)}{\tilde{C}_{\kappa}(I_n) \cdot k!}.
\]

In (3.18), \( \kappa = (k_1, k_2, \ldots, k_n) \) is a partition of \( k \) into not more then \( n \) parts with \( k_1 \geq k_2 \geq \ldots \geq k_n \geq 0 \) and \( k_1 + k_2 + \ldots + k_n = k \), and the complex multivariate hypergeometric coefficient reads
\[
[a]_{\kappa} = \prod_{i=1}^{n} [a-i+1]_{k_i}
\]
with Pochhammer’s symbol
\[
[a]_{k} = a \cdot (a+1) \cdot \ldots \cdot (a+k-1) \quad [a]_{0} = 1.
\]

The notation \( \tilde{C}_{\kappa}(X) \) denotes a zonal polynomial defined in \[81], which is a symmetric polynomial of degree \( k \) in the eigenvalues of matrix \( X \) and the interested reader is referred e.g. to \[81\] for greater details. It is emphasized that there exist no simple (general) formulas for calculating the coefficients of the zonal polynomials and for higher degrees \( k \) the computational complexity becomes prohibitive ([21][60] and references therein). Furthermore, the complex multivariate Gamma function is defined by (cf. [81, equation (83)])
\[
\Gamma_m(r) = \pi^{\frac{1}{2}m(m-1)} \cdot \prod_{i=1}^{m} \Gamma(r-i+1)
\]
with standard scalar Gamma function \( \Gamma(x) \) (see [1]). Equivalently to (3.18), the hypergeometric function of one \( n \times n \) matrix argument \( X \) is defined by
obviously resulting in the equality

\[
p_{p_F^{(n)}}(a_1, \ldots, a_p; b_1, \ldots, b_q; X) = p_{\tilde{F}_q^{(n)}}(a_1, \ldots, a_p; b_1, \ldots, b_q; X, I_n).
\]  

(3.23)

One can derive with the help of (3.23) that for \( \Omega = I \) the probability density function (PDF) of \( Q \) reduces to the well-known complex Wishart PDF defined by (cf. [81, equation (94)])

\[
p(Q) = \frac{1}{\Gamma_p(v) \cdot |\tilde{\gamma} \Sigma|^v} \cdot e^{-\text{tr}(\tilde{\gamma} \Sigma^{-1} Q)} \cdot (dQ),
\]

(3.24)

whereas the general complex Wishart density of a \( m \times n \) matrix \( W \) with \( m \times m \) covariance \( M \) and parameter \( n \geq m \) reads [81]

\[
p(W) = \frac{1}{\Gamma_m(n) \cdot |M|^n} \cdot e^{-\text{tr}(M^{-1} W) \cdot |W|^{n-m}} \cdot (dW).
\]

(3.25)

We note that eigenvalue PDFs are available for the random matrix \( Q \) with PDF defined in (3.17) as well as (3.24). However, the PDF of the eigenvalues in the general case (3.17) so far can only be given in infinite sums of zonal polynomials, thus complicating its practical use.

For the interested reader, we note that more results on central random quadratic forms (including the vector variate case) can be found in [25][98], hypergeometric function representations of the PDF and CDF in [66], eigenvalue PDFs in [66], in particular the distribution of the maximum eigenvalue in [65], infinite series representations of the PDF in [97][164], arbitrary moments of the determinant in [26], and the distribution of the determinant in [97]. For the study of MIMO systems in Ricean fading channels with both transmit and receive correlation, the centrally (i.e. mean value \( 0 \)) distributed channel matrix \( H \) becomes non-central [23]. In statistical literature a number of results on non-central distributions are available. Real non-central random quadratic forms are the subject of [27]. Many integration results and eigenvalue PDFs are addressed in [67]. Non-central complex random quadratic forms were studied in [22][68], where the PDF, moments of the determinant, and joint eigenvalue PDFs are given. Approximations of the non-central distributions by central distributions are addressed in [180].

In principle, we could carry out the same analysis as in the case of Rayleigh fading for the Ricean channel. However, most of the statistical results given for non-central quadratic forms are represented in terms of infinite sums of certain polynomials, whereas it appears that there are no general formulas available for the calculation of the polynomial coefficients. We note that there has been some effort in statistical literature for extending the theory of zonal polynomials to deal with non-central distribution problems [17]. However, still it appears that there are no efficient algorithms available for a practical evaluation of these extended polynomials. For that reason the use
of these results is limited and we are restricting the analysis in this thesis to the Rayleigh fading (central) case.

### 3.1.3 Calculation of the MGF of mutual information

A MGF uniquely defines a probability distribution [147], i.e. once we can find the MGF of mutual information, we can determine all moments, including the practically important first moment, which is the expected value of the capacity with uninformed transmitter and full channel state information at the receiver, also known as ergodic capacity or simply capacity in MIMO literature. In this paragraph, for the first time we calculate the MGF of mutual information of fully correlated Rayleigh fading MIMO channels, using a novel eigenvalue free approach [113]. To this end, we have to calculate the following expected value with respect to \( \mathbf{Q} \) (see (3.14))

\[
M_{\Sigma\Omega}(s) = E_{\mathbf{Q}}[e^{sI(s,y)}] = E_{\mathbf{Q}}\left[\exp\left(\frac{s}{\ln 2} \cdot \ln|\mathbf{I} + \mathbf{Q}|\right)\right] = E_{\mathbf{Q}}\left[|\mathbf{I} + \mathbf{Q}|^{\frac{s}{\ln 2}}\right],
\]

(3.26)

where \( \hat{M}_{\Sigma\Omega}(s) \) is the MGF of MIMO channel mutual information of the artificially augmented \( \nu \times \nu \) system with transmit as well as receive correlation. The standard approach, e.g. applied in [199] to find the MGF, is to equivalently integrate over the eigenvalue PDF \( p(\Lambda_{\mathbf{Q}}) \) of \( \mathbf{Q} \)

\[
M_{\Sigma\Omega}(s) = E_{\Lambda_{\mathbf{Q}}}\left[|\mathbf{I} + \Lambda_{\mathbf{Q}}|^{\frac{s}{\ln 2}}\right] = \int_{\Lambda_{\mathbf{Q}}} |\mathbf{I} + \Lambda_{\mathbf{Q}}|^{\frac{s}{\ln 2}} \cdot p(\Lambda_{\mathbf{Q}}) d\Lambda_{\mathbf{Q}}
\]

(3.27)

with eigenvalue decomposition (EVD)

\[
\mathbf{Q} = \mathbf{V}_{\mathbf{Q}}\Lambda_{\mathbf{Q}}\mathbf{V}_{\mathbf{Q}}^{H}.
\]

(3.28)

We note that the integral in (3.27) has also been used in [87] for calculating the first and second moment of capacity for Ricean MIMO channels and Rayleigh channels with one-sided fading correlation. As was mentioned already above, the general case with both correlation at receiver and transmitter leads to an infinite sum expression of zonal polynomials for the eigenvalue PDF of \( \mathbf{Q} \) [66], thus limiting its applicability and requiring a different proceeding. In the following paragraphs, we first derive the MGF in terms of a hypergeometric function of matrix arguments. Using these results, a representation of the MGF in terms of scalar hypergeometric functions is calculated.

**Concise matrix notation**

We start with the main theorem.

**Theorem 3.1:** The characteristic function \( M_{\Sigma\Omega} \) of the MIMO channel mutual information according to (3.14) for the artificial \( \nu \times \nu \) system, the definitions of the SNR \( \gamma \), \( \nu \), \( \Sigma \) and \( \Omega \) in Paragraph 3.1.1, and channel model in (2.41) with transmit as well as receive correlation is given by
Proof: We directly focus on solving the integral (see (3.26))

\[
M_{\Sigma \Omega} (s) = 2 \tilde{F}_0^{(v)} \left( -\frac{s}{\ln 2}; v, -\gamma \cdot \Sigma (\epsilon), \Omega \right).
\]

where the PDF \( p(Q) \) is given in (3.17). The key for solving (3.30) is given by the following property of the hypergeometric function \( _1F_0^{(m)} \) of matrix argument (cf. [81, equation (90)]) for the \( m \times m \) matrix \( X \)

\[
_1F_0^{(m)} (a; X) = |I - X|^{-a},
\]

which is the matrix analog to the well-known scalar binomial series [1]

\[
_1F_0 (a; x) = (1 - x)^{-a},
\]

where \( _1F_0 \) is a scalar hypergeometric function (11.81). Application of (3.31) to (3.30) leads to

\[
M_{\Sigma \Omega} (s) = E_Q \left[ |I + Q|^{\frac{s}{\ln 2}} \right] = E_Q \left[ _1F_0^{(v)} \left( -\frac{s}{\ln 2}; -Q \right) \right].
\]

By [94, equation (58)] we have for a Hermitian matrix \( M \) the following expected value

\[
E_Q [\tilde{C}_k (MQ)] = \frac{\tilde{C}_k (M \Sigma (\epsilon)) \cdot \tilde{C}_k (\Omega)}{\tilde{C}_k (I)} \cdot [v]_k.
\]

Then expand (3.33) with the help of the series of the function \( _1F_0^{(v)} \) in (3.23) to find

\[
M_{\Sigma \Omega} (s) = \sum_{k=0}^{\infty} \sum_{\kappa} \left[ -\frac{s}{\ln 2} \right]_k \frac{E_Q [\tilde{C}_k (-Q)]}{k!}.
\]

Application of expected value (3.34) leads to

\[
M_{\Sigma \Omega} (s) = \sum_{k=0}^{\infty} \sum_{\kappa} \left[ -\frac{s}{\ln 2} \right]_k \frac{[v]_k \cdot C_k (-\gamma \Sigma) \cdot C_k (\Omega)}{C_k (I)}.
\]

Now using the series definition of \( 2 \tilde{F}_0^{(v)} \) in (3.18) proves the theorem. QED.

We emphasize again that Theorem 3.1 can be used for the analysis of systems with arbitrary antenna array sizes, arbitrary channel correlation and arbitrary noise and signal covariance matrices. To this end, we need to apply the corresponding limiting processes. Specifically, we consider a scenario with one-sided channel correlation \( (\Omega = I) \) in more detail, where we can make use of
certain properties of the complex Wishart distribution. We note that in this special case we can express the MGF directly in terms of $\Sigma$, avoiding the use of the artificially augmented $\Sigma(\varepsilon)$.

**Corollary 3.1:** The MGF $M_\Sigma(s)$ of the MIMO channel mutual information according to (3.12) with $\Omega = I$ is given by

$$M_\Sigma(s) = 2\tilde{F}_0^{(\mu)}\left(-\frac{s}{\ln 2}, \nu; \gamma \Sigma\right) = 2\tilde{F}_0^{(\mu)}\left(-\frac{s}{\ln 2}, \nu; \gamma \Sigma, I_\mu\right). \quad (3.37)$$

**Proof:** Corollary 3.1 is a direct consequence of Theorem 3.1 and the derivation is left to the reader. However, we present a mathematically interesting alternative derivation of (3.37) by using certain integrals with respect to the complex Wishart PDF in Appendix 9.1.1.

### 3.1.4 Hypergeometric matrix functions and scalar representations

The practical relevance of Theorem 3.1, where the MGF of MIMO mutual information is essentially an infinite sum of zonal polynomials (we emphasize again that there are no general formulas available for their calculation), can be established by the following Lemma 3.1, which was given for the $2 \times 2$ case and conjectured for the general case in [95], later proven in [96], and independently derived in [57]. For completeness, we mention that alternatively there are infinite sum of scalar hypergeometric function expressions available for the $2 \times 2$ case [139]. Laplace approximations for special matrix variate hypergeometric functions can be found in [14].

**Lemma 3.1:** Let $X = \text{diag}(x_1, \ldots, x_m)$ and $Y = \text{diag}(y_1, \ldots, y_m)$ with $x_1 > \ldots > x_m$ and $y_1 > \ldots > y_m$. Furthermore define

$$\Gamma_m(r) = \frac{\Gamma_m(r)}{\pi^{m(m-1)/2}} = \prod_{i=1}^{m} \Gamma(r - i + 1) \quad (3.38)$$

and the Vandermonde determinant $\alpha_m(X)$, which can be expressed as (see also (11.18))

$$\alpha_m(X) = \prod_{i < j} (x_i - x_j). \quad (3.39)$$

Furthermore, define the auxiliary function with vector $b = [b_1, b_2, \ldots, b_q]^T$

$$\psi_q^{(m)}(b) = \prod_{i=1}^{m} \prod_{j=1}^{q} (b_j - i + 1)^{j-1}. \quad (3.40)$$

Then the hypergeometric functions of 2 matrix arguments can be expressed in terms of scalar hypergeometric functions

$$p\tilde{F}_q^{(m)}(a_1, \ldots, a_p; b_1, \ldots, b_q; X, Y) = \frac{\Gamma_m(m) \cdot \psi_q^{(m)}(b)}{\alpha_m(X) \cdot \alpha_m(Y) \cdot \psi_q^{(m)}(a)} \quad (3.41)$$
with \( a = [a_1 \ a_2 \ \ldots \ a_p]^T \) and auxiliary \( m \times m \) matrix \( D \) (i and j run from 1 to m)

\[
D = \begin{bmatrix}
pFq(a_1 - m + 1, \ldots, a_p - m + 1; b_1 - m + 1, \ldots, b_q - m + 1; x_i y_j)
\end{bmatrix},
\]

where the elements of \( D \) are scalar hypergeometric functions.

**Proof**: See [96, Lemma 3].

Using Lemma 3.1, we can directly derive the following specialized Lemma 3.2, which is important for the MIMO mutual information MGF analyzed in this work.

**Lemma 3.2**: The hypergeometric function of 2 matrix arguments \( _2\tilde{F}_0^{(m)}(a_1, a_2; X, Y) \) can be calculated by

\[
_2\tilde{F}_0^{(m)}(a_1, a_2; X, Y) = \Gamma_m(m) \cdot \frac{\text{\( _2F_0(a_1 - m + 1, a_2 - m + 1; : x_i y_j) \)}}{\alpha_m(X) \cdot \alpha_m(Y) \cdot \psi_{20}^{(m)}(a_1, a_2)}
\]

(3.43)

with the scalar hypergeometric function

\[
_2F_0(a_1, a_2; : z) = \sum_{k=0}^{\infty} [a_1]_k [a_2]_k \cdot \frac{z^k}{k!}.
\]

(3.44)

At this point we emphasize again that Lemma 3.2 is the key for establishing the practical relevance of the MGF in Theorem 3.1. Unfortunately, however, the analysis is somewhat complicated by the fact that Lemma 3.1 and Lemma 3.2 are directly valid only for matrices \( X \) and \( Y \) that have distinct eigenvalues (this can be seen by the fact that the Vandermonde determinants in the denominator of (3.41) and (3.43) as well as \( |D| \) in the nominator become 0 for equal eigenvalues), such that we have to calculate certain limits to cover the cases, where several eigenvalues take on the same value. For that reason, in the following paragraphs we have to differentiate between fully correlated channels (FCC) with transmit as well as receive correlation, one-side or semi-correlated channels (SCC), and uncorrelated channels (UCC).

### 3.1.5 MGF for fully correlated channels

In this paragraph, we derive several alternative notations of the MGF of MIMO mutual information in terms of well-known scalar hypergeometric functions. At this point, we emphasize that (3.43) is valid only for two square \( m \times m \) matrices \( X \) and \( Y \), thus (3.43) can only be applied to (3.37) after artificially introducing a \( v \times v \) MIMO system as outlined in Paragraph 3.1.1 with a subsequent limiting process. We summarize the results in

**Theorem 3.2**: The MGF \( M_{\Sigma \Omega}(s) \) of the MIMO channel mutual information according to (3.12) with FCC and arbitrary number of transmit and receive antennas is given by
\[ M_{\Sigma\Omega}(s) = \frac{\Gamma_v(\nu)}{(v \cdot (v-1))^{\frac{1}{2}} \cdot \gamma^{\frac{1}{2}}} \cdot \lim_{\varepsilon \to 0} \frac{\left\{ \frac{2F_0\left[-\frac{s}{\ln 2} - v + 1, 1; -\gamma \tilde{\sigma}_i(\varepsilon) \omega_j\right]}{\alpha_v(\Sigma(\varepsilon)) \cdot \alpha_v(\Omega) \cdot \psi_2^{(v)}\left(-\frac{s}{\ln 2}, v\right)} \right\}}{\nu - (v-1)} \cdot \gamma^{\frac{1}{2}}, \] (3.45)

with the definitions of \( \Sigma(\varepsilon) \) and \( \tilde{\sigma}_i(\varepsilon) \) in (3.13).

**Proof:** After application of (3.43) to (3.29) note that
\[ \alpha_v(-\gamma \Sigma(\varepsilon)) = \alpha_v(\Sigma(\varepsilon)) \cdot (-1)^{\nu} \cdot \gamma^{\frac{1}{2}}. \] (3.46)

We can establish (3.46) by factoring out \((-\gamma)^0, (-\gamma)^1, \ldots, (-\gamma)^{\nu-1}\) from the columns of the Vandermonde determinant \( \alpha_v(-\gamma \Sigma(\varepsilon)) \). **QED.**

After calculating the limit in (3.45), we explicitly find

**Corollary 3.2:** The MGF \( M_{\Sigma\Omega}(s) \) in (3.45) can be written as
\[ M_{\Sigma\Omega}(s) = \frac{(v - \mu - 1)}{(v - \mu) \cdot \gamma^{\frac{1}{2}}} \cdot \frac{|\Sigma|^{\nu - \mu}}{\Gamma_v(\nu)} \cdot \frac{\psi_{\Omega}(\nu - \mu)(s)}{\psi_{\Omega}(\nu - \mu)(s)} \cdot \frac{\psi_{\Omega}(\nu)(s)}{\psi_{\Omega}(\nu)(s)}, \] (3.47)

with the auxiliary \( v \times v \) matrix
\[ \psi_{\Sigma\Omega}(s) = \begin{bmatrix} \psi_{\Sigma\Omega, 1}(s) \\ \psi_{\Sigma\Omega, 2}(s) \end{bmatrix}, \] (3.48)

which is split into the \( \mu \times v \) matrix \((i \) runs from 1 to \( \mu \) and \( j \) from 1 to \( v \))
\[ \psi_{\Sigma\Omega, 1}(s) = \begin{bmatrix} 2F_0\left[-\frac{s}{\ln 2} - v + 1, 1; -\gamma \omega_j \right] \\ U\left(1, \frac{s}{\ln 2} + v + 1, 1, \gamma \omega_j \right) \end{bmatrix}, \] (3.49)

where \( U(a, b, z) \) is Kummer’s \( U \) function (see Appendix 11.6.6 or [1]), and the \((\nu - \mu) \times v \) matrix \((i' \) runs from 1 to \( \nu - \mu \) and \( j' \) from 1 to \( v \))
\[ \psi_{\Sigma\Omega, 2}(s) = \begin{bmatrix} (-\gamma \omega_j)^{\nu - 1} \left[-\frac{s}{\ln 2} - v + 1 \right]_{\nu - 1} \end{bmatrix}. \] (3.50)

**Proof:** See Appendix 9.1.2.

Corollary 3.2 provides an explicit expression for the MGF of mutual information of fully correlated Rayleigh fading MIMO channels in terms of a determinant with elements resulting from evaluation of the Kummer \( U \) function. Via equality (11.89) we note that (3.49) can also be expressed in terms of the incomplete Gamma function. Exploiting certain properties of the \( U \)
function and determinants, it is possible to derive alternative representations of the MGF that are better suited for following calculations, e.g. the calculation of ergodic capacity below.

**Corollary 3.3:** The characteristic function \( M_{\Sigma \Omega}(s) \) in (3.47) can be written as

\[
M_{\Sigma \Omega}(s) = \frac{(-1)^{\frac{\nu(\nu-1)}{2}}} {\gamma } \cdot \frac{(-1)^{\frac{(\nu-\mu)(\nu-\mu-1)}{2}}} {\Gamma(\nu)} \cdot \frac{\Psi_{\Sigma \Omega}(s)} {\alpha_\mu(\Sigma) \cdot \alpha_\nu(\Omega) \cdot \psi_2^{(\nu)}(-\frac{s}{\ln 2}, \nu)},
\]

(3.51)

with the \( v \times v \) matrix

\[
\Psi_{\Sigma \Omega}(s) = \begin{bmatrix}
\Psi_{\Sigma \Omega, 1}(s) \\
\Psi_{\Sigma \Omega, 2}(s)
\end{bmatrix},
\]

(3.52)

where the \( (v-\mu) \times v \) matrix \( \Psi_{\Sigma \Omega, 2}(s) \) is defined in (3.50) and the \( \mu \times v \) matrix (\( i \) runs from 1 to \( \mu \) and \( j \) from 1 to \( v \))

\[
\Psi_{\Sigma \Omega, 1}(s) = \begin{bmatrix}
\sum_{k=\nu-\mu}^{v-2} (-1)^{k} \cdot (\gamma \omega_j)^{-k} \cdot \sigma_i^{k-\nu-\mu} \cdot \left[ -\frac{s}{\ln 2} - \nu + 1 \right]_k + \\
(-1)^{\nu-1} \cdot \sigma_i^{\nu-2} \cdot (\gamma \omega_j)^{\nu-2} \cdot \left[ -\frac{s}{\ln 2} - \nu + 1 \right]_{\nu-1} \cdot U\left(1, \frac{s}{\ln 2} + 2, \frac{1}{\gamma \sigma_i \omega_j}\right)
\end{bmatrix}.
\]

(3.53)

**Proof:** See Appendix 9.1.3.

### 3.1.6 MGF for semi-correlated and uncorrelated channels

In the following, we derive the MGF of MIMO channel mutual information for the case of one-sided correlation or in the absence of correlation. As was noted above, those cases have been derived already in [87] and [199], however, the eigenvalue based approach taken there differs significantly from the novel approach presented in this work. Therefore, for completeness and unification of existing solutions, we demonstrate how those cases can be derived from the general result in Corollary 3.3. Unfortunately, those special cases can be obtained only by tedious yet straightforward calculation of certain limits of the form 0/0 by L’Hospital’s rule. Furthermore, we have to differentiate between various cases, i.e. no channel correlation, channel correlation at that side of the MIMO link with more transmit antennas \( \nu \) (\( \Omega \neq I, \Sigma = I \)) and correlation at the side with less antennas \( \mu \) (\( \Sigma \neq I, \Omega = I \)). The resulting expressions are again given in terms of determinants of scalar Kummer U functions, i.e. single integrals according to (11.88).

**MGF of semi-correlated channel with \( \Sigma \neq I, \Omega = I \)**

If there is no correlation present at that side of the link with the larger number of antenna elements \( \nu \), we have to carry out the formal limit calculation
\[ M_\Sigma(s) = \lim_{\Omega \to I} M_{\Sigma\Omega}(s), \quad (3.54) \]

which requires extensive application of L’Hospital’s rule (see below). We summarize the result of (3.54) in the following theorem.

**Theorem 3.3:** The MGF \( M_\Sigma(s) \) of the semi-correlated MIMO channel mutual information, where that side of the MIMO link with the larger number of antenna elements \( \nu \) is uncorrelated with \( \Sigma \neq I, \Omega = I \) (SCC), is given by

\[
M_\Sigma(s) = \frac{\mu \cdot (\mu - 1)}{\nu - \mu + 1} \cdot \frac{\Psi_\Sigma(s)}{\alpha_\mu(\Sigma)}, \quad (3.55)
\]

with the \( \mu \times \mu \) matrix \( (i, j \text{ run from } 1 \text{ to } \mu) \)

\[
\Psi_\Sigma(s) = \left[ \int_0^\infty e^{-\sigma_i} \cdot t^{v-\mu+j-1} \cdot (1 + \gamma t)^{\frac{s}{2}} \ln dt \right]. \quad (3.56)
\]

**Proof:** See Appendix 9.1.4.

At this point we note that with the Vandermonde determinant identity (11.19) and after rearranging the elements in the determinant of (3.56), Theorem 3.3 can be shown to agree with [87, Theorem 4.1], thus confirming the validity of our analysis.

**MGF of semi-correlated channel with \( \Omega \neq I, \Sigma = I \)**

When the side the MIMO link with less antenna elements is uncorrelated, we have to calculate

\[
M_\Omega(s) = \lim_{\Sigma \to I} M_{\Sigma\Omega}(s) \quad (3.57)
\]

and the results are stated in the following theorem.

**Theorem 3.4:** The MGF \( M_\Omega(s) \) of the correlated MIMO channel mutual information with \( \Omega \neq I, \Sigma = I \) (SCC) is given by

\[
M_\Omega(s) = \frac{\Psi_\Omega, 1(s)}{\Psi_\Omega, 2} \cdot \Psi_\Omega, 1(s) \cdot (v - \mu + 1)^{-2} \cdot (v^2 - \mu^2) \cdot \alpha_\nu(\Omega) \cdot \Gamma_\mu(\mu), \quad (3.58)
\]

with the \( \mu \times \nu \) matrix \( (i \text{ runs from } 1 \text{ to } \mu \text{ and } j \text{ runs from } 1 \text{ to } \nu) \)
\[
\Psi_{\Omega,1}(s) = \left[ \int_0^\infty e^{-\frac{1}{\omega_j} t} \cdot t^{i-1} \cdot (1 + \gamma t)^{\frac{s}{\ln^2 t}} dt \right]
\] (3.59)

and the \((v - \mu) \times v\) matrix \((i^j \text{ runs from 1 to } v - \mu \text{ and } j \text{ from 1 to } v)\)

\[
\Psi_{\Omega,2} = \left( \frac{1}{\omega_{ij}} \right)^{v - \mu - i^j}.
\] (3.60)

**Proof:** See Appendix 9.1.5.

Again, it can be shown that Theorem 3.4 and [87, Theorem 4.2] agree.

**MGF of uncorrelated channel with \(\Sigma = I\) and \(\Omega = I\)**

Obviously, for obtaining the MGF \(M_u(s)\) of the uncorrelated channel, we have to calculate

\[
M_u(s) = \lim_{\Omega \to I} \left( \lim_{\Sigma \to I} M_{\Sigma \Omega}(s) \right) = \lim_{\Omega \to I} M_{\Omega}(s) = \lim_{\Sigma \to I} M_{\Sigma}(s),
\] (3.61)

resulting in

**Theorem 3.5:** The MGF \(M_u(s)\) of the uncorrelated MIMO channel mutual information with \(\Sigma = I\) and \(\Omega = I\) (UCC) is given by

\[
M_u(s) = \frac{\left| \Psi_u(s) \right|}{\Gamma_\mu(\mu) \cdot \prod_{k=1}^{v-\mu} \Gamma(v - \mu + k)}
\] (3.62)

with the \(\mu \times \mu\) matrix

\[
\Psi_u(s) = \left[ \int_0^\infty e^{-t^i} \cdot t^{v - \mu + i + j - 2} \cdot (1 + \gamma t)^{\frac{s}{\ln t}} dt \right],
\] (3.63)

with \(i \text{ and } j \text{ ranging from 1 to } \mu\).

**Proof:** See Appendix 9.1.6.

It can be shown by simple manipulations of (3.63), that Theorem 3.5 and [87, Corollary 2] are the same.
3.2 Ergodic Capacity with Uninformed Transmitter

Having available closed-form expressions for the MGF for various propagation scenarios with
different correlation properties, we can fully statistically characterize MIMO mutual information.
In this thesis, we focus on ergodic capacity $C_{\text{erg}}$. With uninformed transmitter, i.e. without CSI,
transmit power is equally distributed over all transmit antennas [187] and thus ergodic capacity is
just the first moment, i.e. the mean of mutual information with respect to the channel statistics

$$ C_{\text{erg}} = E_G[I(s, y)] = E_G[\log_2 |I + \gamma \cdot \Sigma G^H \Omega G|]. \quad (3.64) $$

$C_{\text{erg}}$ in (3.64) can now be obtained by the well known relation

$$ C_{\text{erg}} = \left. \frac{d}{ds} M_f(s) \right|_{s = 0}, \quad (3.65) $$

where $M_f(s)$ is one of the MGFs of mutual information (depending on the correlation properties
of the channel) according to the previous section. Again, we emphasize that in principle arbitrary
moments of mutual information can be obtained from the MGFs by iterated differentiation

$$ m_{l,k} = \left. \frac{d^k}{ds^k} M_f(s) \right|_{s = 0}, \quad (3.66) $$

where $m_{l,k}$ denotes the $k$th moment of mutual information. These moments can be used for a dis-
tribution fit, e.g. via Gamma or Gaussian approximations [88]. Alternatively, numerical inversion
techniques can be used for finding the PDF [73]. Again, as in case of the MGFs we have to make
a case differentiation between FCCs, SCCs, and UCCs. In the subsequent calculations, we can
make use of the following formula for the derivative of a determinant, which can be established
via the product rule of differentiation

$$ \frac{d}{ds}|X(s)| = \sum_i |X_i(s)|, \quad (3.67) $$

where $|X_i(s)|$ is the determinant of a general matrix $X$, where the $i$th column (or alternatively
row) is differentiated with respect to $s$.

3.2.1 Fully correlated channels

With the prerequisite that all eigenvalues of transmit as well as receive correlation matrix differ,
we obtain the main theorem, which is the basis for all subsequent capacity calculations.

**Theorem 3.6:** The ergodic capacity $C_{\text{erg}}(\Sigma, \Omega)$ of a fully correlated MIMO system with transmit
as well as receive correlation and the definitions of $\Sigma$ and $\Omega$ in Paragraph 3.1.1 is given by
Ergodic Capacity with Uninformed Transmitter

\[ C_{\text{erg}, \Sigma\Omega}(\gamma) = \frac{\Gamma_v(v) \cdot (v-\mu-1)}{\ln 2 \cdot \alpha_\mu(\Sigma) \cdot \alpha_v(\Omega) \cdot \gamma} \cdot \sum_{l=1}^{\mu} \frac{\Xi_{\Sigma\Omega}(l)}{\Psi_{\Sigma\Omega, 2(0)}} \]  

(3.68)

with the \( \mu \times v \) matrix (\( i \) runs from 1 to \( \mu \) and \( j \) from 1 to \( v \))

\[ \Xi_{\Sigma\Omega}(l) = \begin{cases} 
\Gamma(v) \cdot \sigma_i^{\mu-1} \cdot (\gamma \omega_j)^{v-1} \cdot \frac{1}{\gamma \sigma_i, \omega_j} \cdot E_1 \left( \frac{1}{\gamma \sigma_i, \omega_j} \right) & i = l \\
\sum_{k=v-\mu}^{v-1} (-1)^k \cdot (\sigma_i)^{k-(v-\mu)} \cdot (\gamma \omega_j)^k \cdot [1-v]_k & i \neq l 
\end{cases} 
\]  

(3.69)

and \( \Psi_{\Sigma\Omega, 2(s)} \) defined in (3.50). \( E_1(z) \) is the exponential integral (see (11.70)).

**Proof:** See Appendix 9.2.1.

At this point we note that possible numerical problems in the calculation of ergodic capacity according to (3.69) can be overcome simply by expanding the determinants and then factoring out the powers of \( \gamma \). Furthermore, depending on the SNR region of interest, small summands of the expansion can be neglected.

### 3.2.2 Semi-correlated and uncorrelated channels

For completeness, we present the following theorems for systems with SCCs and UCCs. They are given without proof, as they were derived already in [87], however, the interested reader can calculate them via an adequate limit process using L’Hospital’s rule and noting that by Lebesgue’s dominated convergence theorem, we can exchange the sequence of integration and differentiation. Furthermore, we show that our results agree with earlier capacity calculations in SISO and SIMO cases [119].

**Semi-correlated channel with** \( \Omega = I, \Sigma \neq I \)

**Theorem 3.7:** The ergodic capacity \( C_{\text{erg}, \Sigma}(\gamma) \) of a MIMO system with \( \Omega = I, \Sigma \neq I \) (SCC) is given by

\[ C_{\text{erg}, \Sigma}(\gamma) = \frac{(\gamma - \mu - 1)}{\ln 2 \cdot \alpha_\mu(\Sigma) \cdot \sum_{l=1}^{\mu} \Xi_{\Sigma}(l)} \]  

(3.70)

with the \( \mu \times \mu \) matrices (\( i, j \) run from 1 to \( \mu \))
Theorem 3.7 has also been given in [87, Theorem 5.1]. Note that via integration by parts one can derive a closed form of the first integral in (3.71), which in its general form we denote by \( I_1 \), for integer \( n \) and arbitrary constants \( a \) and \( c \) (see e.g. [5, equation (78)])

\[
I_1(c, a, n) = \int_0^\infty e^{-ct} \cdot t^{n-1} \cdot \ln(1 + at) dt = \Gamma(n) \cdot e^{c/a} \cdot \sum_{k=1}^{\infty} \frac{\Gamma\left(-n + k, \frac{c}{a}\right)}{e^k \cdot a^{n-k}}
\]  

(3.72)

with incomplete Gamma function \( \Gamma(z, x) \) [1, Chapter 6]. Moreover, we have from [1, 6.5.19] the special case of the incomplete Gamma function (again with integer \( n \) and arbitrary \( x \))

\[
\Gamma(-n, x) = \frac{(-1)^n}{n!} \cdot \left[ E_1(x) - e^{-x} \cdot \sum_{j=0}^{n-1} \frac{(-x)^j \cdot j!}{x^{j+1}} \right],
\]  

(3.73)

i.e. the only special function occurring in the calculation of \( C_{\text{erg}, \Sigma}(\gamma) \) in (3.70) is the exponential integral \( E_1 \).

Semi-correlated channel with \( \Sigma = I, \Omega \neq I \)

**Theorem 3.8:** The ergodic capacity \( C_{\text{erg}, \Omega}(\gamma) \) of a correlated MIMO system with \( \Sigma = I, \Omega \neq I \) (SCC) is given by

\[
C_{\text{erg}, \Omega}(\gamma) = \frac{\gamma \cdot (\gamma - 1)}{2} \cdot \frac{\ln 2 \cdot \alpha_\gamma(\Omega)}{\Gamma_\mu(\mu)} \cdot \sum_{i = 1}^{\mu} \frac{\Xi_{\Omega, i}}{\Psi_{\Omega, 2}},
\]  

(3.74)

with the \( \mu \times v \) matrix (\( i \) runs from 1 to \( \mu \) and \( j \) runs from 1 to \( v \))

\[
\Xi_{\Omega}(l) = \left[ \begin{array}{c}
\int_0^\infty e^{-\omega_i} \cdot t^{i-1} \cdot \ln(1 + \gamma t) dt = I_1\left(\frac{1}{\omega_i}, \gamma, i\right) \\
\int_0^\infty e^{-\omega_j} \cdot t^{j-1} dt = \Gamma(j) \cdot \omega_j
\end{array} \right],
\]  

(3.75)
the \((v - \mu) \times v\) matrix \(\Psi_{\Omega, 2}\) defined in (3.60), and integral \(I_1(c, a, n)\) defined in (3.72).

Again, we note that Theorem 3.8 has also been derived in [87, Theorem 5.2] by different means.

Uncorrelated channel \((\Sigma = I\, \text{and } \Omega = I)\)

**Theorem 3.9:** The ergodic capacity \(C_{\text{erg,u}}(\gamma)\) of a MIMO system with UCC is given by

\[
C_{\text{erg,u}}(\gamma) = \frac{1}{\mu} \cdot \ln 2 \cdot \Gamma_\mu(\mu) \cdot \prod_{k=1}^{\mu} \Gamma(v - \mu + k) \cdot \sum_{l=1}^{\mu} \left| \Xi_u(l) \right|
\]

(3.76)

with the \(\mu \times \mu\) matrix \((i, j \text{ run from } 1 \text{ to } \mu)\)

\[
\Xi_u(l) = \begin{bmatrix}
\int_0^{\infty} e^{-t} \cdot t^{v-\mu+i+j-2} \cdot \ln(1 + \gamma t) dt = I_1(1, \gamma, v - \mu + i + j - 1) & i = l \\
\int_0^{\infty} e^{-t} \cdot t^{v-\mu+i+j-2} dt = \Gamma(v - \mu + i + j - 1) & i \neq l
\end{bmatrix}
\]

(3.77)

and integral \(I_1\) defined in (3.72).

We note that Theorem 3.9 was given in [87, Corollary 4]. Moreover, it can be shown that the ergodic capacity expression in (3.76) agrees with the results derived in [187], as expected. In case of a single transmit and receive antenna only, we can directly get from Theorem 3.9 the following corollary.

**Corollary 3.4:** For a SISO system, the ergodic capacity with a Rayleigh fading channel is given by

\[
C_{\text{erg,SISO}}(\gamma) = \frac{1}{\ln 2} \cdot I_1(1, \gamma, 1) = \frac{1}{\ln 2} \cdot e^{1/\gamma} \cdot E_1\left(\frac{1}{\gamma}\right).
\]

(3.78)

In order to check the validity of the result in (3.78), we simply calculate the ergodic capacity with a scalar complex Gaussian distributed channel coefficient \(h\)

\[
C_{\text{erg,SISO}}(\gamma) = E_h[\log_2(1 + \gamma \cdot h \cdot h^*)] = \frac{1}{\ln 2} \cdot E_z[\ln(1 + \gamma \cdot z)],
\]

(3.79)

where \(z\) is exponentially distributed. We then get

\[
C_{\text{erg,SISO}}(\gamma) = \frac{1}{\ln 2} \cdot \int_0^{\infty} \ln(1 + \gamma \cdot z) \cdot e^{-z} dz
\]

(3.80)
and with the definition of $I_1$ in (3.72) we find (3.78). We note that (3.78) has been derived in [119]. It is a simple exercise to calculate the corresponding result for SIMO systems, where the RV $z$ in (3.79) has to be replaced by a Gamma distributed variable with $N$ degrees of freedom, where $N$ is the diversity parameter of the system (in this case equal to the number of receive antennas).

### 3.3 Ergodic Capacity Asymptotics

In the last section, exact formulas of ergodic capacity for arbitrary number of transmit and receive antennas and arbitrarily correlated Rayleigh fading MIMO channels were presented. The given expressions can easily be implemented on a computer, using e.g. the Matlab environment and yield extremely accurate results, thus superseding the need for lengthy Monte-Carlo simulations. However, a direct and simple characterization of the influence of the various system parameters like number of transmit and receive antennas and correlation properties of the wireless channel on ergodic capacity is difficult.

Thus, in the following paragraphs, we derive asymptotical formulas for ergodic capacity in the high and low SNR region as well as for a large number of either transmit or receive antennas. In particular, we demonstrate that the capacity is independent of the correlation properties of the channel in the low SNR region. Furthermore, we give an exact straight line approximation of ergodic capacity as a function of the SNR in dB for the high SNR region, which allows for a simple characterization of the influence of the number of transmit and receive antennas and correlation on ergodic capacity. Again we have to differentiate between propagation environments with varying correlation properties. We derive the asymptotics based on our exact ergodic capacity formulas given above. For validating our results, we present alternative derivations for the low and high SNR regime with one-side correlated and uncorrelated channels. On the other hand, the asymptotics for a large number of either transmit or receive antennas can directly be derived without using the exact analytical results. It turns out that in the asymptotic limit of a large number of antennas at one side of the MIMO link, fading correlation at this particular side of the link has no influence on ergodic capacity.

#### 3.3.1 High SNR asymptotics

As was noted above, we first present asymptotic formulas that are based on the exact capacity expressions. In the second part of this paragraph we validate the asymptotics via an alternative derivation.

**Derivation based on exact formulas**

Starting off from the exact expression of ergodic capacity established in Theorem 3.6, it is possible to find the following straight line equation for the high SNR region.
Theorem 3.10: The ergodic capacity asymptotics $C_{\text{erg}, \Sigma \Omega}(\gamma_{dB})$ of a fully correlated channel at high SNR with arbitrary number of transmit and receive antennas are given by

$$C_{\text{erg}, \Sigma \Omega}(\gamma_{dB}) = \frac{1}{\ln 2} \cdot (-\mu \cdot (E + \ln \rho) + \ln |\Sigma| + \zeta_1(\Omega) - \zeta_2(\mu, \nu)) + \frac{\log_{10} 10}{10} \cdot \gamma_{dB},$$  \hfill (3.81)

where $E = 0.5772156649$ is Euler’s constant, $\rho$ is the transmit power constraint, $\gamma_{dB}$ is defined in (2.73), $\zeta_1(\Omega)$ is an auxiliary function depending on $\Omega$ (in the $v \times v$ determinants, $i$ and $j$ run from 1 to $v$)

$$\zeta_1(\Omega) = \frac{1}{\alpha_v(\Omega)} \cdot \sum \frac{\left| \sum_{j=1}^{\mu} (\omega_j^v - i) \cdot \ln \omega_j \right|}{\omega_j^v - i} ,$$  \hfill (3.82)

and the auxiliary function

$$\zeta_2(\mu, \nu) = \sum_{l = 1}^{\nu - 2} a_{E, \nu - l - 1} \cdot \Gamma(\nu - l).$$  \hfill (3.83)

The coefficients $a_{E, k}$ in (3.83) are given by

$$a_{E, k} = \sum_{l = 1}^{k} (-1)^l \cdot \frac{1}{(k - l)! l!} .$$  \hfill (3.84)

Proof: See Appendix 9.2.2.

We emphasize that the slope of the curve in (3.81) is directly proportional to the minimum of the number of transmit and receive antennas $\mu$, independent of the correlation properties of the channel. This agrees with results established via a large-dimensional random matrix analysis in [170]. Moreover, the negative impact of the correlation matrix $\Sigma$ at that side of the MIMO link with a smaller number of antennas is directly obvious in (3.81). Specifically, the determinant of a matrix is a Schur-concave function of the eigenvalues (see [131] and Appendix 11.5), i.e. higher correlation reduces capacity. A special case of Theorem 3.10, which is not directly obvious from (3.81), is given in the following corollary for a correlated $v \times v$ system with the same number of transmit and receive antennas (see also Corollary 3.7 for an alternative formulation).

Corollary 3.5: The ergodic capacity asymptotics of a fully correlated $v \times v$ system are given by

$$C_{\text{erg}, \Sigma \Omega, v \times v}(\gamma_{dB}) = \frac{1}{\ln 2} \cdot (-v \cdot (E + \ln \rho) + \ln (|\Sigma| |\Omega|) - \zeta_2(v, v)) + v \cdot \frac{\log_{10} 10}{10} \cdot \gamma_{dB}$$  \hfill (3.85)

with $\zeta_2(v, v)$ defined in (3.83).

Proof: Corollary 3.5 follows directly from symmetry considerations or by expansion of the determinants in (3.82) via determinant identity (11.13). QED.
We emphasize at this point that Theorem 3.10 is not only valid for fully correlated systems, i.e. with $\Omega \neq I$ and $\Sigma \neq I$, but also for semi-correlated systems with $\Sigma = I$. Another special case of Theorem 3.10 results when $\Omega = I$. Now we have to calculate a limit of the form $\frac{0}{0}$ via L'Hôpital’s rule and we get

**Corollary 3.6:** In the limit $\Omega = I$, the function $\xi_1(\Omega)$ in (3.81) takes on the value (in the $\nu \times \nu$ determinants, $i$ and $j$ run from 1 to $\nu$)

$$
\xi_1(I) = \frac{1}{(-1)^\frac{(\nu-1)}{2}} \cdot \frac{1}{\Gamma(\nu)} \cdot \sum_{i=1}^{\mu} \begin{vmatrix} \alpha_{j-1}(l) & i = l \\ [\nu - i + 2]_{j-1} & i \neq l \end{vmatrix} (3.86)
$$

with the recursive definitions

$$
\alpha_k(i) = (\nu - i - k + 1) \cdot \alpha_{k-1}(i) + \beta_{k-1}(i) \quad \beta_k(i) = [\nu - i - k + 1]_k (3.87)
$$

and

$$
\alpha_0(i) = 0 \quad \beta_0(i) = 1. (3.88)
$$

Proof: See Appendix 9.2.3.

Using a different approach for finding the high SNR asymptotics, in the following sub-paragraph we derive alternative formulations of Corollary 3.5 and Corollary 3.6, which result in somewhat simpler expressions that are not directly obvious.

**Alternative derivations and formulations**

For the special case in Corollary 3.5 it is possible to find a simple alternative derivation, leading to the following corollary.

**Corollary 3.7:** The ergodic capacity asymptotics of a fully correlated $\nu \times \nu$ system can alternatively be written as

$$
\bar{C}_\text{erg}, \Sigma, \nu \times \nu(\gamma_{dB}) = \frac{1}{\ln 2} \left( -\nu \cdot \ln \rho + \ln(|\Sigma|) + \sum_{k=0}^{\nu-1} \psi(\nu - k) \right) + \nu \cdot \frac{\log_2 10}{10} \cdot \gamma_{dB}, (3.89)
$$

where $\psi(z)$ is the Digamma function ([1, Paragraph 6.3] and (11.36)).

Proof: Note from (3.12) that at high SNR the mutual information in the $\nu \times \nu$ case can be approximated by

$$
I(s, y) \approx \log_2 [\gamma \cdot \Sigma G^H \Omega G] = \log_2 |G^H G| + \log_2 |\gamma \cdot \Sigma|. (3.90)
$$

Now realizing that $G^H G$ is complex i.i.d. Wishart $W_{\nu}(\nu, I)$ distributed and using expected value (11.35), we directly get (3.89). QED.
In case of $\Omega = I$, which was studied in Corollary 3.6, again a simple alternative solution can be given, which can be derived analogously to Corollary 3.7, such that we state without proof

**Corollary 3.8:** Alternatively to Corollary 3.6, the ergodic capacity high SNR asymptotics of a MIMO system with semi-correlated channel $\Omega = I$ can be written as

$$C_{\text{erg}, \Sigma, \gamma} = \frac{1}{\ln 2} \left( -\mu \cdot \ln \rho + \ln(\Sigma) + \sum_{k=0}^{\mu-1} \psi(v-k) \right) + \mu \cdot \frac{\log_{10} 10}{10} \cdot \gamma dB,$$

(3.91)

where $\psi(z)$ is the Digamma function [1, Paragraph 6.3].

Obviously, formulation (3.91) avoids recursive calculations, which were necessary in (3.86).

### 3.3.2 Low SNR asymptotics

In the low SNR regime, we can again use a series expansion of the exponential integral in our exact ergodic capacity formulas in order to find the asymptotics. The result presented in Theorem 3.11 can then also be checked by different means outlined below, thus reconfirming the validity of the ergodic capacity analysis.

**Theorem 3.11:** The ergodic capacity $\mathcal{C}_{\text{erg}}(\gamma)$ of an arbitrarily correlated channel in the low SNR region is given by

$$\mathcal{C}_{\text{erg}}(\gamma) = \frac{\text{tr}(\Sigma) \cdot \text{tr}(\Omega)}{\ln 2} \cdot \gamma.$$

(3.92)

With the typical normalization of the correlation matrices (2.54), white input signals, AWGN, and $\rho = T$ we find

$$\mathcal{C}_{\text{erg}}(\gamma) = \frac{R \cdot T}{\ln 2} \cdot \gamma = \frac{R \cdot \rho}{\ln 2} \cdot \gamma = \frac{R \cdot \rho}{\ln 2} \cdot \frac{1}{\rho} \cdot \frac{\gamma dB}{10} = \frac{R \cdot \gamma dB}{\ln 2} \cdot \frac{10}{10} = \frac{R \cdot \gamma dB}{\ln 2} \cdot 10^{\frac{\gamma dB}{10}}.$$

(3.93)

**Proof:** See Appendix 9.2.4 for a proof based on the exact formulas given in Theorem 3.6. Here, we present a concise alternative proof verifying the validity of our analysis. To this end, consider the formula [68]

$$\ln |I - xA| = - \sum_{k=1}^{\infty} \frac{1}{k} \cdot x^k \cdot \text{tr}(A^k)$$

(3.94)

for scalar $x$ and matrix $A$ with $\|xA\| < 1$. For small $x$ we obviously get

$$\ln |I - xA| \approx -x \cdot \text{tr}(A).$$

(3.95)

Then observe the ergodic capacity expression

$$C_{\text{erg}} = E_G[\log_2 |I + \gamma \cdot \Sigma \Omega^H G^H G|].$$

(3.96)
For low SNR we get with (3.95) and the help of expected value (11.33)
\[
C_{\text{erg}} = \frac{1}{\ln 2} \cdot E_G[\gamma \cdot \text{tr}(\Sigma G^H \Omega G)] = \frac{\gamma}{\ln 2} \cdot \text{tr}(\Sigma \cdot E_G[G^H \Omega G]) = \frac{\gamma}{\ln 2} \cdot \text{tr}(\Sigma) \cdot \text{tr}(\Omega). \quad (3.97)
\]
QED.

Interestingly, in Theorem 3.11 it can be seen that capacity increases linearly in SNR in the low SNR regime, whereas in the high SNR regime (Theorem 3.10) the increase is only logarithmic. Moreover, (3.93) shows that ergodic MIMO capacity with uninformed transmitter in the low SNR region is independent of the correlation properties of the channel. This contrasts the behavior in the high SNR region, where fading correlation decreases ergodic capacity (see Chapter 4). On the other hand, with channel state information (CSI) and statistical transmit processing at the transmitter, which adapts the covariance of the transmit vector to the statistics of the MIMO channel, it is well known that ergodic capacity can be increased with strong fading correlation at low SNR (see e.g. [76]). This will be studied in detail in Paragraph 3.4.

### 3.3.3 Large T asymptotics

For a large number of transmit antennas the ergodic capacity expressions simplify significantly. Specifically, we get

**Theorem 3.12:** With a finite number of receive antennas \( R \) and a large number of transmit antennas \( T \), the ergodic capacity asymptotics read

\[
\bar{C}_{\text{erg},T}(\gamma) = R \cdot \frac{\gamma_{dB}}{10} \cdot \log_2 10 + \log_2 |O| + R \cdot \log_2 \left( \frac{\text{tr}(S)}{\rho} \right)
\]

with the definition (2.70) of the diagonal matrix \( O \) at the receiver and the diagonal matrix \( S \) at the transmitter (2.71).

**Proof:** For \( T > R \) the ergodic capacity is given by (with \( T \times R \) matrix \( G \))

\[
C_{\text{erg}} = E_G[\log_2 |I_R + \gamma \cdot O G^H S G|]. \quad (3.99)
\]

By the law of large numbers it can be shown that

\[
\Pr(\{ |G^H S G|_{ij} - E[|G^H S G|_{ij}] > \tau_1 \}) < \tau_2(T), \quad (3.100)
\]

where \( \tau_1 \) is a small positive constant and \( \tau_2(T) \) is a positive constant with \( \tau_2(T) \to 0 \) as \( T \to \infty \), i.e. the random matrix \( G^H S G \) becomes a deterministic quantity with

\[
G^H S G \to E[G^H S G] = \text{tr}(S) \cdot I_R. \quad (3.101)
\]

We note that this phenomenon was aptly termed 'channel hardening’ in [71], where many results on the ergodic capacity asymptotics and the asymptotic normality of EC for various uncorrelated MIMO systems with applications can be found. Using (3.101) in (3.99), we find
C_{\text{erg}}(\gamma) \approx \log_2 | R + \gamma \cdot \text{tr}(S) \cdot O | . \hspace{1cm} (3.102)

A further simplification is possible for large $T$ (also for the high SNR regime), where (3.102) can be approximated by the asymptotics

$$C_{\text{erg},T}(\gamma) = \log_2 | \gamma \cdot \text{tr}(S) \cdot O | = R \cdot \log_2 (\gamma \cdot \text{tr}(S)) + \log_2 | O | . \hspace{1cm} (3.103)$$

Taking into account the Schur-concavity of the determinant, the negative influence of fading correlation at the receiver becomes obvious from (3.103). Moreover, with the SNR in dB definition of (2.76) we find

$$C_{\text{erg},T}(\gamma) = R \cdot \log_2 \left( \frac{\gamma_{\text{db}}}{10} \log_2 10 + \log_2 | O | + R \cdot \log_2 \left( \frac{\text{tr}(S)}{\rho} \right) \right) . \hspace{1cm} (3.104)$$

QED.

Theorem 3.12 essentially states that for a large number of transmit antennas the capacity saturates, i.e. it becomes independent of the number of transmit antennas and is just a function of the SNR, as long as $\text{tr}(S)/\rho = \text{const.}$. However, this condition is fulfilled due to the normalization of the channel correlation matrices. Finally, we note that as expected, again the slope of the capacity asymptotics is proportional to the number of receive antennas $R = \min(R, T)$.

### 3.3.4 Large R asymptotics

Following the same derivation as in case of a large number of transmit antennas, we get

**Theorem 3.13:** In the limit of a large number of receive antennas $R$ the ergodic capacity asymptotics are given by

$$C_{\text{erg},R}(\gamma) = T \cdot \frac{\gamma_{\text{db}}}{10} \log_2 10 + T \cdot \log_2 (\text{tr}(O)) + \log_2 | S | - T \cdot \log_2 (\rho) . \hspace{1cm} (3.105)$$

**Proof:** With a finite number of transmit antennas $T$ and a large number of receive antennas $R$, the ergodic capacity reads (with $R \times T$ matrix $G$)

$$C_{\text{erg}} = E_G [ \log_2 | I_T + \gamma \cdot SGHOG | ] . \hspace{1cm} (3.106)$$

Similar to the case of a large number of transmit antennas, we get

$$C_{\text{erg}}(\gamma) \approx \log_2 | I_T + \gamma \cdot \text{tr}(O) \cdot S | . \hspace{1cm} (3.107)$$

We can further bound (this is also true for the high SNR regime)

$$C_{\text{erg},R}(\gamma) = \log_2 | \gamma \cdot \text{tr}(O) \cdot S | = T \cdot \log_2 (\gamma) + T \cdot \log_2 (\text{tr}(O)) + \log_2 | S | . \hspace{1cm} (3.108)$$
As a function of the SNR in dB we obtain (3.105). QED.

Obviously, ergodic capacity increases logarithmically in the number of receive antennas \( R \) via 
\[ T \cdot \log_2(\text{tr}(\mathbf{O})) \]
and with the Schur-concavity of the determinant, transmit correlation reduces ergodic capacity with uninformed transmitter. We note that the ergodic capacity asymptotics are again linear in the SNR in dB and the number of transmit antennas \( T = \min(R, T) \) determines the slope.

### 3.4 Ergodic capacity with transmit CSI

The ergodic capacity expressions derived above are based on the assumption of an uninformed transmitter, i.e. there is no channel state information (CSI) available at the transmitter and it applies a uniform power allocation on the transmit antennas (the transmit signal covariance matrix is an identity matrix). In this paragraph, we extend these results to cover the case of statistical CSI at the transmitter [110], whereas the transmitter is assumed to be aware of the correlation properties of the channel, which we call channel distribution information at the transmitter (CDIT). It is well known that the transmitter has to adapt the covariance matrix of the transmit signal vector according to the CSI such that it essentially transmits with a proper power allocation (PA) on the eigenmodes of the fading correlation matrix [78][85][194] at the transmitter side. However, in contrast to the case of instantaneous transmit CSI [34], to the authors’ best knowledge, so far the optimum capacity-achieving statistical power allocation strategy is known only for the special case of 2 transmit antennas and uncorrelated receive antennas [173]. Based on the exact mean mutual information (MMI) derived in Paragraph 3.2, in this paper we fill this gap and extend the results of [173] to an arbitrary number of transmit antennas and arbitrary fading correlation at the receive antenna array. To this end, we numerically optimize the transmit signal covariance matrix using the novel exact MMI expressions. Furthermore, in the next Chapter 4 we compare the results with some lower complexity power allocation schemes based on certain bounds on MMI [76][100].

#### 3.4.1 Optimal transmission strategy

Based on the exact MMI expressions derived in the previous paragraphs, it is a straightforward exercise to take into account the capacity-achieving transmit signaling with CDIT. To this end, we note that the optimal transmission strategy is given by [78][85]

\[
F = \tilde{V}_{TX} \cdot \Phi \cdot U \quad \quad R_{ss} = E_s \cdot I,
\]

where \( \tilde{V}_{TX} \) is the \( T \times L \) matrix of \( L \) eigenvectors corresponding to the \( L \) largest eigenvalues of the transmit correlation matrix according to (2.42), diagonal \( L \times L \) power allocation (PA) matrix \( \Phi \) (\( L \) corresponds to the number of independent subchannels in a practical system), and arbitrary unitary \( L \times L \) matrix \( U \). As a consequence, we get from (2.71) with \( \tilde{R}_{ss} = I \)
\[
S(\Phi) = \text{eig}(\tilde{R}_w F^H R_{TX} F) = \text{eig}(\Phi^2 \cdot \Lambda_{TX}),
\]

where \(\Lambda_{TX}\) is the matrix of the \(L\) largest eigenvalues of the transmit correlation matrix. Obviously, mutual information of the MIMO link is now a function of the PA matrix \(\Phi\) and we get from (3.8)

\[
I(s, y, \Phi) \equiv \log_2 \left| I + \gamma \cdot S(\Phi)\tilde{H}_w^H \tilde{O} \tilde{H}_w \right|.
\]

As we have the same type of mathematical problem as in case of uninformed transmitters, in order to obtain the mean value of mutual information averaged over the channel statistics we can directly apply the formulas for MMI with uninformed transmitter derived in Paragraph 3.2, where we just have to replace \(S\) by \(S(\Phi)\). The ergodic capacity with CDIT in a correlated Rayleigh fading environment is then given by

\[
C_{\text{erg}}^{\text{CDIT}}(\gamma) = \max_{\Phi} E[I(s, y, \Phi)] \quad \text{s.t. } \text{tr}(\Phi \Phi^H) = \rho
\]

with transmit power constraint \(\rho\). For calculating the expected value in (3.112) we have again to differentiate between fully correlated channels, semi-correlated channels, and uncorrelated channels. Unfortunately, the solution to the constrained optimization problem for finding the optimum power allocation coefficients can in general not be given in closed form. We therefore have to resort to standard numerical multivariate optimization techniques [19][124]. Results of this optimization process will be presented in the simulations below in section 3.5.

### 3.4.2 Waterfilling in the low and high SNR regime

Due to the complexity of the exact MMI expressions and the numerical optimization process in the statistical WF algorithm, it is difficult to study the analytical behavior of ergodic capacity with CDIT in general. However, with the availability of low and high SNR asymptotics of MMI, we can study the asymptotical behavior, which is at least partially analytically tractable.

**High SNR regime**

We have noticed from the asymptotic formulas that ergodic MIMO capacity is proportional to \(\mu = \min(L, R)\) for high SNR. Therefore, in case of \(T \leq R\) the transmit prefilter should exploit all spatial dimensions at the transmitter side with \(L = T\), such that \(\Phi\) should be chosen to be of full rank \(\text{rk}(\Phi) = T\). With a full rank transmit prefilter, from Theorem 3.10 (and Corollary 3.8) together with the optimal transmission strategy (3.109), it is clear that the ergodic capacity with CDIT in the high SNR regime with \(T \leq R\) and arbitrarily correlated MIMO channels is given by

\[
\overline{C}_{\text{erg}}^{\text{CDIT}}(T \leq R) = \zeta_{C, T \leq R} + \max_{\Phi} \log_2(\left| \Phi^2 \right|) \quad \text{s.t. } \text{tr}(\Phi \Phi^H) = \rho
\]

with a constant \(\zeta_{C, T \leq R}\), which is independent of \(\Phi\). However, the determinant \(\left| \Phi^2 \right|\) is a Schur-concave function of the power allocation coefficients \(\phi_k^2\), such that in the high SNR regime we
have the optimal power allocation strategy $\Phi_{opt, T \leq R} = I_T$, i.e. essentially no waterfilling takes place and therefore the absolute capacity gain with CDIT due to waterfilling is zero

$$\Delta C_{T \leq R}^{WF} = C_{CDIT}^{opt, T \leq R} - C_{erg, T \leq R} = 0,$$  

(3.114)

where $\overline{C}_{erg, T \leq R}$ is the ergodic capacity in the high SNR regime with uninformed transmitter and $T \leq R$. On the other hand, for $T > R$ the slope of the high SNR asymptotics are now determined by $\mu = R$ if we let $L \geq R$, and we have to resort to numerical optimization [19][124] for finding the optimal PA matrix $\Phi$ that achieves capacity. From Theorem 3.10 it can be seen that in the high SNR regime the optimum $\Phi_{opt, T > R}$, which maximizes MMI, is given by

$$\Phi_{opt, T > R} = \arg \max_{\Phi} \left( \zeta_1 (\Lambda_{TX} \cdot \Phi^2) - \zeta_2 (R, rk(\Phi)) \right) \quad \text{s.t.} \quad \tr(\Phi^2) = \rho,$$  

(3.115)

with the definitions of the auxiliary functions $\zeta_1 (\Omega)$ and $\zeta_2 (\mu, \nu)$ in (3.82) and (3.83), and the eigenvalue decomposition of the transmit correlation matrix in (2.42). The absolute gain due to statistical waterfilling with $\Phi_{opt, T > R}$ is then given by

$$\Delta C_{T > R}^{WF} = \frac{1}{\ln 2} \left[ \zeta_1 (\Lambda_{TX} \cdot \Phi_{opt, T > R}^2) - \zeta_2 (R, rk(\Phi_{opt, T > R})) - (\zeta_1 (\Lambda_{TX}) - \zeta_2 (R, T)) \right].$$  

(3.116)

Again, optimization problem (3.115) is not solvable in closed form and we therefore resort to numerical optimization algorithms [19][124]. Results of the optimization will be presented in a comparative study of various statistical WF policies in Chapter 4.

Low SNR regime

With the optimal transmission strategy in (3.109), from Theorem 3.11 the ergodic capacity with channel state information at the transmitter in the low SNR region is given by

$$C_{erg}^{CDIT}(\gamma) = \max_{\Phi} \frac{\tr(\Lambda_{TX} \cdot \Phi^2) \cdot \tr(\Omega)}{\ln 2} \cdot \gamma \quad \text{s.t.} \quad \tr(\Phi^H \Phi) = \rho.$$  

(3.117)

It is straightforward to see that $\tr(\Lambda_{TX} \cdot \Phi^2)$ is Schur-convex in the power allocation coefficients $\phi_k^2$ and therefore optimal transmission in the low SNR regime is beamforming, where all transmit power is put on the strongest long-term eigenmode of the channel via $\Phi_{opt} = \text{diag}(\sqrt{\rho}, 0, ..., 0)$. The relative waterfilling gain $G_C^{WF}$ in EC can now be calculated

$$G_C^{WF} = \frac{C_{erg}^{CDIT}}{C_{erg}} = \frac{\lambda_{TX, 1} \cdot \rho}{\tr(\Lambda_{TX})} = \frac{\lambda_{TX, 1} \cdot T}{\lambda_{TX, 1} \cdot T} = \lambda_{TX, 1},$$  

(3.118)

where $\lambda_{TX, 1}$ is the maximum eigenvalue of the transmit correlation matrix according to (2.43).
3.5 Numerical Results

In the following paragraphs we present ergodic capacity curves that are determined via Monte-Carlo simulations as well as by evaluating the novel closed form capacity expressions for validating the analysis. Furthermore, we demonstrate the application of the asymptotical capacity expressions for the high and low SNR regime as well as for a large number of transmit and receive antennas. Finally, we show the effectiveness of statistical waterfilling with CDIT, which achieves significant capacity increases in highly correlated channels.

3.5.1 Ergodic capacity with uninformed transmitter

Without CSI at the transmitter, power is equally distributed over the transmit antennas, i.e. $R_{ss} = E_s \cdot I$ and $F = I_T$, which implies $L = T$. Simulation results and theoretical curves closely agree in Fig. 3.1 for a 6x6 system with FCC according to the realistic correlation matrix model (RCMM). As expected, the negative impact of channel correlation on ergodic capacity with uninformed transmitter can be observed. Correlation increases (angular spread decreases) from left to right. Note in Fig. 3.1 that the capacity loss due to increased transmit correlation (difference between upper and middle curve) is greater than the loss induced by additional receive correlation (difference between middle and lower curve). However, we emphasize that this observation is valid only for the given SNR range. From (3.85) we know that in the high SNR region both receive and transmit correlation have the same impact on capacity. Moreover, at high SNR all curves achieve the same slope according to the asymptotic analysis in Paragraph 3.3.1, while they converge at low SNR according to the study in Paragraph 3.3.2.

![Fig. 3.1 EC, FCC, RCMM, T=R=6, different AS at transmitter and receiver](image-url)
For a FCC with RCMM and \( \Delta_{RX} = \Delta_{TX} = 10^\circ \) at receiver and transmitter we have depicted the dependence of ergodic capacity on the number of transmit antennas with a fixed number of \( R = 4 \) receive antennas in Fig. 3.2. Obviously, \( T = 1 \) corresponds to the receive diversity case in standard single input multiple output (SIMO) smart antenna systems. As expected, there is a great gain in going from 1 to 2 transmit antennas, while the gain gets smaller if we add more and more transmit antennas. A similar effect can be observed when considering SER curves of diversity systems. Moreover, it is clear from the large \( T \) analysis in Paragraph 3.3.3 that ergodic capacity saturates for a large number of transmit antennas at a fixed value, which depends only on the SNR.

In Fig. 3.3, the lower 3 curves result from a system simulation with \( T = 2 \) transmit antennas and \( R = 2, 4, \) and 6 receive antennas without fading correlation. Similar results are given for a system with \( T = 4 \) and \( R = \{4, 6, 8, 10\} \) in the upper 4 curves. Again, we can observe that the gain of introducing additional receive antennas reduces. However, from the large \( R \) analysis in Paragraph 3.3.4 we know that ergodic capacity grows logarithmically with the number of receive antennas, i.e. there is no fixed saturation point.

The total number of antennas is kept constant \( R + T = 8 \) in Fig. 3.4 with FCC and AS of 10 degrees at transmitter and receiver. Interestingly, the capacity of the 2x6, 3x5, and 4x4 system are not far apart in the considered SNR range. Moreover, in the given SNR range, the asymmetric 3x5 system exhibits the best performance. This contrasts the behavior in the high SNR regime, where the asymptotic slope is proportional to \( \min(R, T) \), see Theorem 3.10. Similar results are depicted in Fig. 3.5 for an uncorrelated system. Compared to the fully correlated case, the 2x6, 3x5, and 4x4 systems exhibit now a different behavior, such that the ergodic capacity curves soon approach their asymptotic behavior which is determined by \( \min(R, T) \). At this point we note that due to a
reduced slope in the low SNR region, ergodic capacity curves for systems with high fading correlation achieve their asymptotic performance only at higher SNR compared to uncorrelated systems. This will also be shown in the next section on asymptotical capacity curves.

Plots of the ergodic capacity for the exponential correlation matrix model (ECMM) are depicted in Fig. 3.6 for various SNR values given as a parameter and uncorrelated fading at the receive antenna array. In this MIMO system with $T = 4$ transmit and $R = 4$ receive antennas, channel correlation at the transmitter has a catastrophic impact on ergodic capacity for correlation coeffi-

Fig. 3.3 EC, UCC, $T = \{2, 4\}$, various $R$

Fig. 3.4 EC, FCC, RCCM, $\Delta_{RX} = \Delta_{TX} = 10^\circ$, $T + R = 8$

Plots of the ergodic capacity for the exponential correlation matrix model (ECMM) are depicted in Fig. 3.6 for various SNR values given as a parameter and uncorrelated fading at the receive antenna array. In this MIMO system with $T = 4$ transmit and $R = 4$ receive antennas, channel correlation at the transmitter has a catastrophic impact on ergodic capacity for correlation coeffi-
MIMO Mutual Information and Capacity

From the asymptotic analysis in Paragraph 3.3.2, we know that in the low SNR regime ergodic capacity essentially becomes independent of the correlation properties of the channel, which can directly be observed in the simulation results. In contrast to that, the influence of correlation on the capacity of SIMO smart antenna systems is negligible (Fig. 3.7). Finally, the influence of fading correlation at the receiver is studied in Fig. 3.8 with uncorrelated fading at the transmitter using the RCMM. Now the angular spread is varied at the receiver, whereas small AS corresponds to high fading correla-

\[ \text{Fig. 3.5 EC, UCC, } T+R=8 \]

\[ \text{Fig. 3.6 EC, SCC, ECMM, } T=R=4, \text{ various } r_{TX}, r_{RX}=0.0 \]
Numerical Results

A dramatic capacity degradation can be observed for low AS, which is especially pronounced at high SNR.

**Fig. 3.7** EC, ECMM, SIMO system, $T=1$, $R=8$

**Fig. 3.8** EC, SCC, RCMM, $T=4$, $R=4$, various $\Delta_{RX}$, TX uncorrelated

dition and vice versa. Again, a dramatic capacity degradation can be observed for low AS, which is especially pronounced at high SNR.
3.5.2 Ergodic capacity asymptotics

High and low SNR asymptotics

Simulation results and theoretical high SNR asymptotics closely agree in Fig. 3.9 for a 4x4 system with RCMM, where the overall correlation of the system increases from left to right (the angular spread, given as a parameter of the curves, decreases). Note that from Corollary 3.5 we can calculate an asymptotic loss of 2.59 bit per channel use for the $\Delta_{RX} = \Delta_{TX} = 30^\circ$ case, 16.81 bit per channel use for the $\Delta_{RX} = \Delta_{TX} = 10^\circ$ case, and 40.00 bit per channel use for the $\Delta_{TX} = 2^\circ, \Delta_{RX} = 10^\circ$ case considered in Fig. 3.9. It is also obvious from Fig. 3.9 that the higher the overall fading correlation in the system, the later the asymptotics and the exact simulated curves converge. Similar results are depicted in Fig. 3.10 for a MIMO system with ECMM correlation model, $T = 4$ transmit antennas, and $R = 6$ receive antennas. The correlation is fixed at the receiver side with correlation coefficient $r_{RX} = 0.7$ and various transmit correlation coefficients $r_{TX} = \{0.3,0.7,0.9,0.95\}$.

On the other hand, ergodic capacity curves for the low SNR regime are given on a logarithmic scale in Fig. 3.11 for a system with ECMM channel model, $T = 3$, and $R = 5$. Again the correlation coefficient at the receiver is fixed $r_{RX} = 0.7$, while the transmit correlation is varying $r_{TX} = \{0.3,0.7,0.9,0.95\}$. Obviously, the theoretical asymptotics agree very well for low SNR values smaller than -15 dB. Between -15 and -5 dB the ergodic capacity curves start to deviate from the asymptotics, and for higher SNR values greater than -5 dB the curves corresponding to various correlation coefficients of the channel clearly diverge.

![Fig. 3.9 EC asymptotics, RCMM, $R=T=4$, various $\Delta_{RX}, \Delta_{TX}$](image_url)
Numerical Results

The dependence of ergodic capacity on the number of receive antennas is plotted in Fig. 3.12 for a fully correlated MIMO system with ECMM channel model with $r_{RX} = r_{TX} = 0.5$. The SNR is fixed at 20 dB and we have depicted curves for $T = \{1, 2, 3, 4\}$. For comparison, we have also included the logarithmic large $R$ asymptotics according to Theorem 3.13. It can be observed that...
they are less tight for a greater number of transmit antennas, i.e. the asymptotics become valid only for an increased number of receive antennas. A similar behavior can be seen for the case of a large number of transmit antennas with a limited number of receive antennas in Fig. 3.13, whereas the SNR is now fixed at 10 dB. Note that as theoretically predicted, the capacity becomes independent of $T$, if $T$ is large.
3.5.3 Statistical waterfilling

Plots of the ergodic capacity for a system with ECMM channel model, $T = 4$, and $R = 6$ are given in Fig. 3.14. The channel exhibits strong fading correlation at the transmitter with $r_{TX} = 0.97$, while the fading at the receiver is only weakly correlated. We compare three different transmission strategies. Without channel state information (CSI) at the transmitter, the optimal transmission strategy is to transmit without a special power allocation, which yields the lowest capacity. With the availability of CDIT, the transmitter can apply long-term (LT) statistical power allocation (PA) based on the exact EC/MMI expressions according to (3.112). One can observe a significant increase in ergodic capacity especially in the lower SNR region. Finally, with the availability of instantaneous CSI at the transmitter we can apply optimum short-term (ST) power allocation [34], which achieves the highest ergodic capacity. However, we note that with this highly correlated MIMO channel the difference between ST PA and LT PA is only minor due to the fact that only a small number (in the limit only one) of eigenmodes can be effectively used for data transmission in the low SNR regime. We note that in the high SNR regime all transmission schemes in general asymptotically yield the same ergodic capacities.

With less fading correlation ($r_{TX} = 0.7$) at the transmitter, the ergodic capacity increase due to adequate power allocation at the transmitter is reduced (Fig. 3.15). Moreover, it can be seen that the SNR range, where CSI at the transmitter is beneficial, is limited. At the same time, there is now a noticeable difference between the case of ST CSI and CDIT. Due to the reduced correlation, it is now obviously more likely that a higher number of channel eigenmodes is used for data transmission and this can be exploited, if the transmitter is aware of the instantaneous channel state.

![Fig. 3.14 EC, ECMM, ST and statistical WF, $T = 4$, $R = 6$, $r_{TX} = 0.97$, $r_{RX} = 0.3$](image-url)
Fig. 3.15 EC, ECMM, ST and statistical WF, $T = 4$, $R = 6$, $r_{TX} = 0.7$, $r_{RX} = 0.3$
4 Bound on Ergodic MIMO Capacity

It has been demonstrated in Chapter 3 that the exact calculation of ergodic MIMO channel capacity with channel correlation leads to expressions that complicate a further analysis due to their mathematical structure. For example, an interesting question is the exact characterization of the influence of correlation on ergodic capacity. The development of low complexity statistical water-filling (WF) schemes with channel distribution information at the transmitter (CDIT) that maximize mean mutual information, or equivalently the calculation of the optimal transmit signal covariance matrix, is another important issue. In this chapter, we therefore derive a closed-form tight upper bound on the ergodic capacity of correlated MIMO channels in terms of elementary symmetric functions of the eigenvalues of transmit and receive correlation matrix. Bounds presented e.g. in [52][76][145][169][170] cover scenarios with fading correlation at one side of the MIMO link exclusively. For systems with correlation at both receiver as well as transmitter, in [121] the authors propose to use a bound based on the dominating correlation matrix of the link. However, for fully correlated channels with strong correlation at both transmitter and receiver, this bound becomes also very loose. In this thesis we find a remedy to this problem and present a novel bound that takes into account both the effects of correlation at the transmitter as well as the receiver and is thus very tight for arbitrary propagation scenarios. Moreover, we give a recursive algorithm for its efficient calculation, which allows for a quick numerical evaluation even for a high number of antenna elements.

It is shown that the novel bound on ergodic capacity is Schur-concave in the eigenvalues of transmit as well receive correlation matrix. Furthermore, by means of the bound we can give a concise proof of the fact that the exact ergodic capacity with uninformed transmitter is Schur-concave and can thus substantiate the observation that higher correlation reduces ergodic capacity, if no channel state information (CSI) is available at the transmitter. We also derive a simple statistical WF scheme, which can be given in closed form for systems with two transmit antennas and arbitrary number of receive antennas. Monte-Carlo simulations demonstrate the tightness of the bound and show that a long-term WF scheme based on the new bound yields essentially the same performance as a scheme based on the exact ergodic capacity analysis introduced in Chapter 3.

4.1 Novel Capacity Bound

In this section we present a novel upper bound on MIMO channel capacity in correlated Rayleigh fading environments based on Jensen’s inequality. We also consider some special cases for various correlation properties of the channel and the low and high SNR region. Furthermore, we give a recursive algorithm for efficiently calculating the novel bound.
4.1.1 General derivation

It was demonstrated in Chapter 3, equation (3.6) that the mutual information of a correlated MIMO channel can in general be expressed as

\[ I(s, y) \equiv \log_2 |I + \gamma \cdot \Sigma G^H \Omega G| \]  

(4.1)

with diagonal deterministic \( \mu \times \mu \) matrix \( \Sigma \), \( \nu \times \nu \) matrix \( \Omega \), and i.i.d. complex Gaussian distributed \( \nu \times \mu \) matrix \( G \) defined in Paragraph 3.1.1. The ergodic capacity with uninformed transmitter is then given by

\[ C_{\text{erg}}(\gamma) = E_G[\log_2 |I + \gamma \cdot \Sigma G^H \Omega G|]. \]  

(4.2)

We directly start with the main theorem of this chapter, which can be directly derived from (4.2).

**Theorem 4.1:** The ergodic capacity of a MIMO link in arbitrarily correlated Rayleigh fading environments with uninformed transmitter can be bounded by

\[ C_{\text{erg}}(\gamma) \leq C_{\text{erg}}^B(\gamma), \]  

(4.3)

with

\[ C_{\text{erg}}^B(\gamma) = \log_2 \left( \sum_{k=0}^{\mu} k! \cdot \gamma^k \cdot \left( \sum_{\hat{\alpha}_k} |\Sigma| \sum_{\hat{\delta}_k} |\Omega| \right) \right), \]  

(4.4)

where \( \hat{\alpha}_k \) and \( \hat{\delta}_k \) denote index subsets of cardinality \( k \) (see also Appendix 11.1.5).

**Proof:** See Appendix 9.3.1.

Furthermore, with the definitions of the elementary symmetric functions in Appendix 11.1.6, we get another interesting representation of the novel capacity bound.

**Corollary 4.1:** In terms of the elementary symmetric functions, the upper bound on ergodic capacity in (4.4) can be rewritten as

\[ C_{\text{erg}}^B(\gamma) = \log_2 \left( \sum_{k=0}^{\mu} k! \cdot \gamma^k \cdot \text{tr}_k(\Sigma) \cdot \text{tr}_k(\Omega) \right). \]  

(4.5)

The importance of Corollary 4.1 becomes clear in later studies of concavity properties of ergodic capacity.

4.1.2 Asymptotics and special cases

In this section we derive asymptotics and some special cases of the capacity upper bound in Theorem 4.1 for different SNR regions, propagation scenarios, and number of antenna elements, which yields new insights into the general properties of the novel capacity bound.
High SNR region

For high SNR we can neglect all smaller powers of $\gamma$ and can approximate (4.4) by

$$C_{\text{erg}}^B(\gamma) \approx \bar{C}_{\text{erg}}^B(\gamma) = \log_2\left( \frac{\mu! \cdot \gamma^\mu \cdot |\Sigma| \cdot \sum_{\delta_\mu} |\Omega|^{\delta_\mu} \delta_\mu!}{\delta_\nu} \right) = \log_2(\mu! \cdot \gamma^\mu \cdot |\Sigma| \cdot \text{tr}_\mu(\Omega)). \quad (4.6)$$

For the special case of a $v \times v$ system (4.6) reduces to

$$\bar{C}_{\text{erg}}^B(\gamma) \approx \log_2(v! \cdot \gamma^v \cdot |\Sigma| \cdot |\Omega|). \quad (4.7)$$

Interestingly, by considering the $v \times v$ system asymptotics in (3.85) resulting from an exact analysis, it becomes clear that the bound asymptotics in (4.7) have an offset that is independent of the correlation properties of the channel. In the general case (4.6), however, the offset is dependent on $\Omega$.

Low SNR region

On the other hand, for low SNR we neglect higher powers of $\gamma$ and find

$$C_{\text{erg}}^B(\gamma) \approx \log_2(1 + \gamma \cdot \text{tr}(\Sigma) \cdot \text{tr}(\Omega)). \quad (4.8)$$

Then using the series expansion of the ln function ([1])

$$\ln(1 + z) = z - \frac{1}{2}z^2 + \frac{1}{3}z^3 - \ldots \quad (4.9)$$

we get

$$C_{\text{erg}}^B(\gamma) = \bar{C}_{\text{erg}}^B(\gamma) = \frac{1}{\ln 2} \cdot \gamma \cdot \text{tr}(\Sigma) \cdot \text{tr}(\Omega). \quad (4.10)$$

Obviously, by comparing (3.92) and (4.10), it can be seen that the novel bound is asymptotically tight in the low SNR regime.

Uncorrelated fading, AWGN

When considering a MIMO system without power allocation with uncorrelated fading, white transmit signals, and white Gaussian noise, i.e. $R_{ss} = E_s \cdot I$, $R_{nn} = N_0 \cdot I$ and $R_{TX} = I$, $R_{RX} = I$, we have

$$\Sigma = I_\mu \quad \Omega = I_\nu. \quad (4.11)$$

Plugging (4.11) in (4.4) results in
\[ C_{\text{erg},u}^B(\gamma) = \log_2 \left( \sum_{k=0}^{\mu} k! \cdot \gamma^k \cdot \sum_{\hat{\alpha}_k} I_{\mu} \cdot \sum_{\hat{\delta}_k} I_{V} \right). \] (4.12)

Now there are exactly \( \binom{\mu}{k} \) possibilities to choose \( k \) elements out of \( \mu \) and exactly \( \binom{V}{k} \) to choose \( k \) elements out of \( V \). Using these relations in (4.12) results in

\[ C_{\text{erg},u}^B(\gamma) = \log_2 \left( \sum_{k=0}^{\mu} k! \cdot \gamma^k \cdot \binom{\mu}{k} \cdot \binom{V}{k} \right). \] (4.13)

We note that a similar result for the uncorrelated case has been derived in [52]. Now consider the case of a large number of receive antennas \( R \), whereas (4.13) reads

\[ C_{\text{erg},u}^B(\gamma) = \log_2 \left( \sum_{k=0}^{T} k! \cdot \gamma^k \cdot \binom{T}{k} \cdot \binom{R}{k} \right) = \log_2 \left( T! \cdot \gamma^T \cdot \binom{R}{T} \right) = \log_2 \left( \gamma^T \cdot \frac{R!}{(R-T)!} \right). \] (4.14)

Using Stirling’s approximation formula [1]

\[ n! \approx n^n \cdot e^{-n} \cdot \sqrt{2\pi n} \] (4.15)

we find from (4.14)

\[ C_{\text{erg},u}^B(\gamma) \approx \log_2 \left( \gamma^T \cdot \frac{R!}{(R-T)!} \cdot e^{-R} \cdot \sqrt{\frac{2\pi R}{R-T}} \right) = \log_2 \left( \gamma^T \cdot \frac{\sqrt{\frac{R}{R-T}} \cdot e^{-R} \cdot \sqrt{\frac{2\pi (R-T)}{R}}} \right). \] (4.16)

In the limit of a large number of receive antennas we then get with \( \lim_{n \to \infty} \left( 1 - \frac{x}{n} \right)^n = e^{-x} \)

\[ C_{\text{erg},u}^B(\gamma) \approx \log_2 (\gamma^T \cdot R^T). \] (4.17)

The result in (4.17) as expected perfectly agrees with expression (3.108) for uncorrelated fading, which has been established by different means. We note that a derivation of the asymptotics based on the novel bound for the general case of fading correlation appears to be mathematically challenging and we do not further pursue this problem in this thesis.

Arbitrarily correlated 2×R case

In case of only 2 transmit and \( R \geq 2 \) receive antennas we directly find from Corollary 4.1

\[ C_{\text{erg, 2×R}}^B(\gamma) = \log_2 \left( 1 + \gamma \cdot \text{tr}(S) \cdot \text{tr}(O) + 2 \cdot \gamma^2 \cdot |S| \cdot \text{tr}_2(O) \right). \] (4.18)

This results can be used to find a closed form statistical WF scheme (see below) with CDIT.
4.1.3 Recursive capacity calculation

Especially for a higher number of transmit and receive antennas the calculation of the partial sums in (4.4) can be computationally challenging. Therefore, we propose a recursive calculation of the capacity bound. To this end, we focus on the elementary symmetric functions, which we denote by $a_k$ for brevity

$$a_k = \sum_{\delta_i} |\Omega| \hat{\delta_i}. \quad (4.19)$$

These coefficients can be shown to be the coefficients of the polynomial in $z$

$$P(z) = \prod_{k=1}^{\nu} (1 + \omega_k \cdot z) = \sum_{k=0}^{\nu} a_k \cdot z^k \quad (4.20)$$

with $\Omega = \text{diag}(\omega)$. In [38], a recursive algorithm was given for the calculation of the $a_k$, which is stated here for completeness. The polynomial coefficients can be calculated via

$$a_k = \frac{E_k}{k!} \quad k = 0...\nu, \quad (4.21)$$

where the $E_k$ can be determined recursively

$$E_0 = 1 \quad E_k = \sum_{i=1}^{k} (-1)^{i-1} \cdot P_{i-1}^{k-1} \cdot S_i \cdot E_{k-i} \quad k \geq 1. \quad (4.22)$$

The auxiliary variables in (4.22) can be obtained via

$$S_i = \sum_{n=1}^{\nu} \omega_n^{2i} \quad i \geq 1 \quad (4.23)$$

$$P_i^n = n \cdot (n-1) \cdot \ldots \cdot (n-i+1)$$

Having the same structure as in (4.19), the other coefficients in (4.4), namely

$$b_k = \sum_{\hat{\alpha}_k} |\Sigma| \hat{\alpha}_k, \quad (4.24)$$

can obviously be calculated in the same way.
4.2 Concavity Properties of Ergodic Capacity

A powerful tool for characterizing the influence of correlation on ergodic capacity is majorization theory (see Appendix 11.5 and [131]). We can say that a channel is more correlated if one vector of eigenvalues of a correlation matrix majorizes another. Loosely speaking, in this case the eigenvalues are more spread out (see also the examples in Paragraph 2.4). We first use majorization theory to characterize the influence of correlation on the novel bound on ergodic capacity. Later, we extend the results to the exact ergodic capacity, which was calculated in Chapter 3. Via the bound representation in (4.5) we can directly derive the following intuitive theorem.

Theorem 4.2: The upper bound on correlated channel capacity in (4.5) is a Schur-concave function of the eigenvalues of $\Sigma$ and $\Omega$, i.e. higher correlation at either side of the MIMO link reduces the capacity bound with uninformed transmitter, independent of the SNR.

Proof: By Theorem 11.3, the argument of the logarithm in (4.5) is a sum of Schur-concave elementary symmetric functions and thus Schur-concave. The logarithm is a monotonic (concave) function and thus preserves Schur-concavity. QED.

At this point we note that the full characterization of exact ergodic capacity for arbitrarily correlated MIMO systems is an open problem. First steps in this direction can be found in [85] for the case of two receive antennas. For MISO systems on the other hand, the impact of correlation at the transmitter is well understood [12]. We now make use of some auxiliary results that were derived in the process of calculating the novel ergodic capacity bound to characterize exact ergodic MIMO capacity via majorization theory. To this end, we introduce the so-called Schur’s condition.

Theorem 4.3: [131, theorem 3.A.4] Let $I \subset \mathbb{R}$ be an open interval and let $\Phi: I^n \to \mathbb{R}$ be continuously differentiable. Necessary and sufficient conditions for $\Phi$ to be Schur-convex on $I^n$ are

$$\Phi$$ is symmetric on $I^n$ \hspace{1cm} (4.25)

and

$$\frac{\partial}{\partial z_i} \Phi(z_1, \ldots, z_n)$$ is decreasing in $i = 1 \ldots n$ for all $(z_1, \ldots, z_n) \in I^n$. \hspace{1cm} (4.26)

Alternatively, $\Phi$ is Schur-convex on $I^n$ if and only if (4.25) and for all $i \neq j$

$$(z_i - z_j) \cdot \left[ \frac{\partial}{\partial z_i} \Phi(z_1, \ldots, z_n) - \frac{\partial}{\partial z_j} \Phi(z_1, \ldots, z_n) \right] \geq 0$$ \hspace{0.5cm} for all $(z_1, \ldots, z_n) \in I^n$. \hspace{1cm} (4.27)

For Schur-concave functions "decreasing" is replaced by "increasing" in (4.26) and inequality (4.27) is reversed.

We apply Theorem 4.3 for deriving the following
**Theorem 4.4:** The exact ergodic MIMO channel capacity in (4.2) is a Schur-concave function of the eigenvalues of $\Sigma$ and $\Omega$, i.e. higher correlation reduces ergodic capacity with uninformed transmitter, independent of the SNR.

**Proof:** We first prove the Schur-concavity of ergodic capacity with respect to the eigenvalues of $\Sigma$. Via symmetry considerations it is then straightforward to extend the result to $\Omega$. As the expected value operator is linear, after application of (4.27) to (4.2) we have to show that

\[
(\sigma_i - \sigma_j) \cdot E_G \left[ \left( \frac{\partial}{\partial \sigma_i} - \frac{\partial}{\partial \sigma_j} \right) \left( \log_2 |I + \gamma \cdot \Sigma G^H \Omega G| \right) \right] \leq 0,
\]

(4.28)

with $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_\mu)$. After differentiating the logarithm, we find

\[
(\sigma_i - \sigma_j) \cdot E_G \left[ \frac{1}{|I + \gamma \cdot \Sigma G^H \Omega G|} \left( \left( \frac{\partial}{\partial \sigma_i} - \frac{\partial}{\partial \sigma_j} \right) |I + \gamma \cdot \Sigma G^H \Omega G| \right) \right] \leq 0.
\]

(4.29)

Then note that (with probability 1) $\gamma \cdot \Sigma G^H \Omega G$ is a positive definite matrix and thus with probability 1 we have $|I + \gamma \cdot \Sigma G^H \Omega G| > 1$. Therefore, we can simplify (4.29) and we have to show

\[
(\sigma_i - \sigma_j) \cdot E_G \left[ \left( \frac{\partial}{\partial \sigma_i} - \frac{\partial}{\partial \sigma_j} \right) |I + \gamma \cdot \Sigma G^H \Omega G| \right] \leq 0.
\]

(4.30)

Again exchanging the sequence of differentiation and expected value operator and applying the expected value that was derived in (4.5) we find

\[
(\sigma_i - \sigma_j) \cdot \frac{\partial}{\partial \sigma_i} - \frac{\partial}{\partial \sigma_j} \left( \sum_{k=0}^{\mu} k! \cdot \gamma^k \cdot \text{tr}_k(\Sigma) \cdot \text{tr}_k(\Omega) \right) \leq 0.
\]

(4.31)

However, it can be seen that (4.31) is exactly Schur’s condition for a sum of elementary symmetric functions. By Theorem 11.3 we know that the elementary symmetric functions are Schur-concave, thus proving Theorem 4.4. QED.

### 4.3 Long-Term Statistical Waterfilling

In Paragraph 3.4 we have presented results on the exact ergodic capacity of correlated MIMO systems with statistical channel distribution information at the transmitter (CDIT). To this end, we have shown that the optimum capacity-achieving strategy is transmission on the long-term eigen-modes of the transmit correlation matrix with proper power allocation (PA). For the determination of the power allocation coefficients we have formulated a constrained optimization problem, which could be solved only numerically [19][124] due to the mathematical complexity of the exact mean mutual information expressions. In order to overcome the problem of lengthy numerical optimization processes, in this paragraph we study the design of a statistical transmit prefilter $F$ (or equivalently the transmit vector covariance) that maximizes a bound instead of exact
expressions of mean MIMO mutual information. It can be shown that the basic prefilter structure with long-term eigenmode transmission and appropriate power allocation is still valid in this case and the problem boils down to finding a statistical power allocation algorithm that optimizes the bound. Closed-form formulas for the statistical power allocation are given for $2 \times R$ MIMO systems, while we have to resort to numerical optimization techniques for finding the prefilter coefficients in the general case. However, even though again numerical optimization is required, the bound can be evaluated much easier and thus a faster and more stable optimization process results.

Finally, we shortly introduce an alternative power allocation scheme based on a loose bound on mean mutual information (MMI), which takes into account only fading correlation at the transmit antenna array [76]. We use this low complexity scheme as a benchmark for the statistical water-filling schemes based on the exact ergodic capacity analysis and the novel tight bound on MMI.

### 4.3.1 Statistical transmit prefilter structure

The ergodic capacity of a MIMO link with CDIT and transmit prefilter was given in Paragraph 3.4 as the solution to the constrained optimization problem (with transmit power constraint $\rho$)

$$
C_{\text{erg}}^{\text{CDIT}}(\gamma) = \max_{\Phi} E[\log_2|I + \gamma \cdot S(\Phi)\tilde{H}_w^H\tilde{O}_w^H|] \quad \text{s.t. } \text{tr}(\Phi\Phi^H) = \rho
$$

with optimal prefilter structure and transmit signal covariance matrix

$$
F = \tilde{V}_{TX} \cdot \Phi \cdot U \quad R_{ss} = E_s \cdot I_L
$$

and the prefilter-dependent matrix (with the EVD of the transmit correlation matrix in (2.42))

$$
S(\Phi) = \text{eig}(\Phi^2 \cdot \tilde{A}_{TX}).
$$

It is a straightforward exercise to show that the transmission strategy given in (4.33) is also optimal in the sense that it maximizes the upper bound on ergodic capacity with statistical CSI at the transmitter

$$
C_{\text{erg}}^{B,\text{CDIT}}(\gamma) = \max_{\Phi} \log_2E[|I + \gamma \cdot S(\Phi)\tilde{H}_w^H\tilde{O}_w^H|] \quad \text{s.t. } \text{tr}(\Phi\Phi^H) = \rho.
$$

Again, it remains to find the optimum power allocation matrix $\Phi$ in the sense of (4.35). In the next paragraph it is demonstrated that a closed form solution can be found for $T = 2$ transmit antennas, while for the general case one has to resort to numerical techniques.
4.3.2 Determination of the power allocation matrix $\Phi$

With the specific transmission strategy in (4.33), we are faced with the problem of determining the optimum power allocation (PA) matrix. The problem of optimizing the bound on mean mutual information with respect to $\Phi$ reads (analogously to (4.35))

$$\Phi^{B}_{\text{opt}} = \arg \max_{\Phi} \log_2 \left| \left[ I + \gamma \cdot \Phi^2 \Lambda_{TX} H_{w} H_{OH} \right] \right| \quad \text{s.t. } \text{tr}(\Phi^2) = \rho \,. \quad (4.36)$$

It can easily be shown that based on (4.4), this problem can be rewritten as

$$\Phi^{B}_{\text{opt}} = \arg \max_{\Phi} \log_2 \left[ \sum_{k = 0}^{\mu} k! \cdot \gamma^k \cdot \sum_{\hat{\alpha}_k} \Phi^{2\hat{\alpha}_k} A_{TX} \sum_{\hat{\delta}_k} O^{\hat{\delta}_k} \right] \quad \text{s.t. } \text{tr}(\Phi^2) = \rho \,. \quad (4.37)$$

Special $2 \times R$ antenna case

We first shall restrict ourselves to the $2 \times R$ antenna case (see also (4.18)) to allow for a concise analytical solution. Omitting details, the diagonal elements of $\Phi^2$ obtained from a Lagrange optimization process are

$$\phi_1^2 = \frac{\gamma^2 \cdot |R_{TX}| \cdot \text{tr}_2(O) \cdot \rho + \frac{\gamma}{2} \cdot (\lambda_{TX,1} - \lambda_{TX,2}) \cdot \text{tr}(O)}{2 \cdot \gamma^2 \cdot |R_{TX}| \cdot \text{tr}_2(O)}$$
$$\phi_2^2 = \frac{\gamma^2 \cdot |R_{TX}| \cdot \text{tr}_2(O) \cdot \rho + \frac{\gamma}{2} \cdot (\lambda_{TX,2} - \lambda_{TX,1}) \cdot \text{tr}(O)}{2 \cdot \gamma^2 \cdot |R_{TX}| \cdot \text{tr}_2(O)} \quad , \quad (4.38)$$

where we have to assure that $\phi_{1/2}^2 \geq 0$, otherwise the coefficient $\phi_2^2$ is set to 0 and all available transmit power is given to $\phi_1^2 = \rho \,. \,$ Note that in agreement with results stated e.g. in [169], the total transmit power is equally distributed on both subchannels, i.e. $\phi_1^2 = \phi_2^2 = \rho/2 \,$, if no transmit correlation is present, i.e. $R_{TX} = I \,$, no matter what receive correlation is prevailing. This result also holds for the general case with an arbitrary number of transmit antennas.

Receive correlated case

Note that for the case of receive correlation only we find from (4.37) the optimization problem in terms of the elementary symmetric functions

$$\Phi^{B,RX}_{\text{opt}} = \arg \max_{\Phi} \log_2 \left[ \sum_{k = 0}^{\mu} k! \cdot \gamma^k \cdot \text{tr}_k(\Phi^2) \cdot \text{tr}_k(O) \right] \quad \text{s.t. } \text{tr}(\Phi^2) = \rho \,. \quad (4.39)$$

The objective function in (4.39) is (according to the proof of Theorem 4.2) Schur-concave in the $\phi_k^2$ with $\Phi = \text{diag}(\phi_1, \ldots, \phi_L) \,$. By Lemma 11.2, the optimal power allocation strategy is thus
uniform. This is in accordance with known results in literature [187] and intuition: if there is no transmit correlation present, there are no prominent directions and the transmitter equally distributes power.

**General case**

As was mentioned already above, for the general case of arbitrary array sizes and arbitrary transmit and receive correlation, a closed form solution of the problem in (4.37) can not be given. Numerical methods have to be applied to find the optimal PA coefficients. The results of such numerical optimization processes [19][124] will be shown in the simulations below.

### 4.3.3 Alternative power allocation scheme

In this paragraph, we study another bound [76] on mean mutual information that can also be derived via Jensen’s inequality and the concavity of the log det function [24]. A power allocation scheme based on this loose bound is used as a benchmark for the power allocation schemes based on the exact EC/MMI analysis and the novel tight bound presented above. Furthermore, we point out an interesting structural equivalence of the waterfilling algorithms in case of short-term and long-term CSI at the transmitter.

In its first version, the loose bound takes into account only the transmit correlation matrix $R_{TX}$ and reads [76]

$$\bar{C}_{erg}^{B, CDIT}(\gamma) = \max_F \log_2 \left| I + \gamma \cdot F^H \cdot E[H^H \cdot \tilde{R}^{-1}_{nn} \cdot H] \cdot F \right| \quad \text{s.t.} \quad \text{tr}(FF^H) = \rho, \quad (4.40)$$

which together with expected value formula (11.33) can be written as

$$\bar{C}_{erg}^{B, CDIT}(\gamma) = \max_F \log_2 \left| I + \gamma \cdot \text{tr}(\tilde{R}^{-1}_{nn} R_{RX}) \cdot F^H \cdot R_{TX} \cdot F \right| \quad \text{s.t.} \quad \text{tr}(FF^H) = \rho. \quad (4.41)$$

It is straightforward to show that the optimal prefilter structure and transmit signal covariance matrix in the sense of (4.41) have the same structure as in case of the tight bound and are thus again given by

$$F = \tilde{V}_{TX} \cdot \Phi \cdot U \quad R_{ss} = E_s \cdot I_L \quad (4.42)$$

with arbitrary unitary $L \times L$ matrix $U$. Using the optimal filter structure, we again arrive at a constrained optimization problem for the power allocation matrix

$$\bar{B}^{\Phi}_{opt} = \max_{\Phi} \log_2 \left| I + \gamma \cdot \text{tr}(\tilde{R}^{-1}_{nn} R_{RX}) \cdot \Phi^2 \Delta_{TX} \right| \quad \text{s.t.} \quad \text{tr}(\Phi^2) = \rho. \quad (4.43)$$

In contrast to the statistical power allocation schemes presented above, problem (4.43) can be solved in closed form by the standard iterative waterfilling algorithm (e.g. [34]) for an arbitrary number of transmit and receive antenna elements. Specifically, there is an interesting short-term
long-term duality, which shall be emphasized in the following. In case of the availability of short-
term CSI at the transmitter, the optimum transmit strategy is given by

\[ F_{ST} = \tilde{V}_{ST} \cdot \Phi_{ST} \cdot U \quad R_{ss} = E_s \cdot I_L \]  

(4.44)

with arbitrary unitary \( L \times L \) matrix \( U \) and power allocation matrix

\[ \Phi_{ST} = \max_{\Phi} \log_2 \left| I + \gamma \cdot \Phi^2 \tilde{\Lambda}_{ST} \right| \quad \text{s.t.} \ \text{tr}(\Phi^2) = \rho \]  

(4.45)

and eigenvalue decomposition of the (instantaneous) matrix

\[ H^H R_{nn}^{-1} H = \begin{bmatrix} \tilde{V}_{ST} & \Lambda_{ST} \end{bmatrix} \begin{bmatrix} \Lambda_{ST} & \tilde{V}_{ST} \end{bmatrix}^H \]  

(4.46)

where \( \Lambda_{ST} \) is the matrix of the \( L \) largest eigenvalues.

The analogy between (4.44) and (4.42) and between (4.45) and (4.43) is obvious and we will see
that similar long-term short-term dualities can in certain cases be established for the prefilter
design for linear zero-forcing and minimum mean squared error receivers.

Finally, note that without channel state information at the transmitter, we just set \( F = I_F \) and find
the loose bound on ergodic capacity

\[ \tilde{C}^{B}_{\text{erg}}(\gamma) = \log_2 \left| I + \gamma \cdot \text{tr}(\tilde{R}_{nn}^{-1} R_{RX}) \cdot R_{TX} \right| . \]  

(4.47)

We emphasize that the effects of the receive correlation matrix \( R_{RX} \) are basically not captured by
this bound and numerical results underpin this observation. However, we know from
Paragraph 3.3.4 that (4.47) is an approximation for a large number of antenna elements \( R \) due to
the channel hardening effect. One can therefore expect that the bound in (4.47) becomes tight for
large \( R \) and numerical results below will confirm this expectation.

On the other hand, we can establish another second version of the loose bound paralleling (4.40)

\[ \hat{C}^{B,CDIT}_{\text{erg}}(\gamma) = \max_F \log_2 \left| I + \gamma \cdot \text{tr}(\tilde{R}_{nn}^{-1} R_{RX}) \cdot R_{TX} \right| \quad \text{s.t.} \ \text{tr}(FF^H) = \rho , \]  

(4.48)

which can be calculated equivalently to (4.41)

\[ \hat{C}^{B,CDIT}_{\text{erg}}(\gamma) = \max_F \log_2 \left| I + \gamma \cdot \text{tr}(F^H \cdot R_{TX} \cdot F) \cdot \tilde{R}_{nn}^{-1} R_{RX} \right| \quad \text{s.t.} \ \text{tr}(FF^H) = \rho . \]  

(4.49)

We note that (4.49) exactly agrees with the large \( T \) asymptotics (3.102) established for uninformed transmitters in Paragraph 3.3.3 with \( F = I \)

\[ C_{\text{erg}}(\gamma) = \log_2 \left| I_R + \gamma \cdot \text{tr}(S) \cdot O \right| . \]  

(4.50)
It is a straightforward exercise to show that the optimal power allocation strategy in the sense of (4.49) is given by

\[ \Phi_{opt}^B = \text{diag}(\sqrt{\rho}, 0, \ldots, 0). \]  

(4.51)

i.e. all transmit power is assigned to the strongest eigenmode of the channel, independent of the SNR.

### 4.4 Numerical Results

In this section we present plots verifying the tightness of the novel capacity bound. Furthermore, we study the properties of the various statistical power allocation strategies introduced above in more detail. To this end, the behavior of the power allocation coefficients is investigated and the effects of statistical waterfilling with CDIT on ergodic MIMO capacity are demonstrated. We note that in the following figures we use the label 'EC bound' for the novel tight bound on ergodic capacity (EC) in Theorem 4.1. On the other hand, we use 'Loose EC bound' for the bound in (4.47). Finally, 'Exact EC' refers to the exact capacity analysis presented in Chapter 3. The terms 'Waterfilling (WF)' and 'Power Allocation (PA)' are used interchangeably.

#### 4.4.1 Ergodic capacity with uninformed transmitter

First, we consider systems with uninformed transmitter, such that \( F = I \). The tightness of the novel bound is shown in Fig. 4.1, where we compare the ergodic capacity curves obtained from Monte-Carlo simulations and the EC bound according to Theorem 4.1 for an exponential correla-
Numerical Results

The effect of statistical waterfilling on ergodic MIMO capacity with correlated channels is studied in Fig. 4.3 for a system with $R = T = 4$. On the transmitter side, we assume strong fading correlation with $r_{TX} = 0.9$ and the receiver side is correlated with $r_{RX} = 0.7$. For validating our theoretical results, we have plotted Monte-Carlo simulation results and theoretical curves according to the analysis presented in Chapter 3. Again, there is a perfect agreement between theory and simulation. First, we can observe a large gain due to waterfilling, which is especially pronounced in the low SNR region. As expected, the PA schema based on short-term (ST) CSI performs best (see (4.45)), while the performance of the long-term (LT) PA schemes with CDIT comes close to
the optimum at low SNR. Furthermore, we can observe that LT PA based on the exact EC/MMI analysis according to (3.112) is superior than a scheme based on the loose EC/MMI bound according to (4.41) in the medium SNR range. We emphasize that the EC curve for the PA scheme based on the novel tight bound in (4.32) cannot be distinguished from the scheme based on exact PA and we have therefore removed it from Fig. 4.3 for clarity. More insight can be obtained by looking at the power allocation coefficients of the matrix $\Phi = \text{diag}(\phi_1, \ldots, \phi_r)$, which are depicted in Fig. 4.4. Obviously, the PA coefficients (4.43) resulting from the PA allocation algo-
Numerical Results

A more detailed view on the relative gain of statistical waterflling (compared to a system without CSI at the transmitter) is given in Fig. 4.6 for a $T = R = 4$ system with correlation at the receiver $r_{RX} = 0.7$ and varying correlation at the transmitter $r_{TX} = \{0.5, 0.7, 0.9, 0.97\}$. A close agreement between the WF scheme based on the exact EC/MMI analysis and the WF scheme based on the novel EC/MMI bound according to (4.36) can be observed. Furthermore, in the low SNR regime the theoretical result (3.118) is confirmed, i.e. statistical waterfilling achieves a relative gain equal to the largest eigenvalue of the transmit correlation matrix $R_{TX}$. With increasing SNR, the WF scheme based on the loose bound deviate significantly from the other two schemes, which are very close together. Specifically, additional eigenmodes are activated already at lower SNR, thus leading to a degradation of mean mutual information. At high SNR, however, all three schemes lead to a uniform power allocation, such that all eigenmodes are used for transmission and thus the statistical waterfilling gain vanishes at higher SNR. Equivalently, in the sense of maximizing EC, CDIT is useless in the high SNR regime for this particular system, where $T = R = 4$. However, in contrast to that, we note that below we present results for a system with $T > R$, where a capacity gain can be achieved even in the high SNR region (see Paragraph 3.4.2).

In case of only minor fading correlation at the transmit antenna array, as expected, the ergodic capacity increase due to statistical CSI at the transmitter is only small (cf. to Fig. 4.5). With a correlation coefficient of $r_{TX} = 0.3$, a capacity gain due to statistical LT PA can only be observed in the low SNR regime, while with the availability of ST CSI at the transmitter still a considerable WF gain can be achieved.

4.4.3 Asymptotical study of statistical waterfilling

High and low SNR regime

A more detailed view on the relative gain of statistical waterfilling (compared to a system without CSI at the transmitter) is given in Fig. 4.6 for a $T = R = 4$ system with correlation at the receiver $r_{RX} = 0.7$ and varying correlation at the transmitter $r_{TX} = \{0.5, 0.7, 0.9, 0.97\}$. A close agreement between the WF scheme based on the exact EC/MMI analysis and the WF scheme based on the novel EC/MMI bound according to (4.36) can be observed. Furthermore, in the low SNR regime the theoretical result (3.118) is confirmed, i.e. statistical waterfilling achieves a relative gain equal to the largest eigenvalue of the transmit correlation matrix $R_{TX}$. With increasing SNR, the WF scheme based on the loose bound deviate significantly from the other two schemes, which are very close together. Specifically, additional eigenmodes are activated already at lower SNR, thus leading to a degradation of mean mutual information. At high SNR, however, all three schemes lead to a uniform power allocation, such that all eigenmodes are used for transmission and thus the statistical waterfilling gain vanishes at higher SNR. Equivalently, in the sense of maximizing EC, CDIT is useless in the high SNR regime for this particular system, where $T = R = 4$. However, in contrast to that, we note that below we present results for a system with $T > R$, where a capacity gain can be achieved even in the high SNR region (see Paragraph 3.4.2).

In case of only minor fading correlation at the transmit antenna array, as expected, the ergodic capacity increase due to statistical CSI at the transmitter is only small (cf. to Fig. 4.5). With a correlation coefficient of $r_{TX} = 0.3$, a capacity gain due to statistical LT PA can only be observed in the low SNR regime, while with the availability of ST CSI at the transmitter still a considerable WF gain can be achieved.

4.4.3 Asymptotical study of statistical waterfilling

High and low SNR regime
transmit correlation (increasing $r_{TX}$) the largest eigenvalue of the transmit correlation matrix increases and thus the relative WF gain increases. The maximum eigenvalues of the transmit correlation matrix for the different correlation coefficients are $\lambda_{TX,1} = \{2.08, 2.72, 3.53, 3.85\}$. At high SNR we note that the relative capacity gain converges to 1. A closer look on the performance of a WF scheme based on the loose capacity bound in (4.47) is taken in Fig. 4.7 for the same system parameters as in Fig. 4.6. While the asymptotical performance in the low and high SNR regime is the same as for the WF scheme based on the exact EC analysis, it is obviously inferior...
Numerical Results

in the medium SNR region. The loss in capacity is in the range of 3 to 4 percent for the given scenario. We note that this loss is higher for strong fading correlation at the receiver, as the loose capacity bound does not take into account correlation at the receive antenna array.

The absolute gain in bit per channel use due to statistical waterfilling based on the exact EC analysis in the high SNR regime (see also (3.116)) is depicted in Fig. 4.8 for a system with $R = 2$ receive antennas and $r_{RX} = 0.7$. As expected, the gain increases with a higher number of transmit antennas and with stronger correlation at the transmit side. Basically, in the presence of CDIT, the transmitter can make use of the beamforming capabilities of the transmit antenna array for improving the capacity of the link. In Fig. 4.9, the squared PA coefficients are depicted for a system with $T = 8$ transmit and $R = 2$ receive antenna elements for $r_{RX} = 0.7$ in the high SNR regime. While for vanishing transmit correlation ($r_{TX} \approx 0$) power is equally distributed on all 8 transmit eigenmodes of the MIMO channel, with strong fading correlation at the transmitter ($r_{TX} \rightarrow 1$) the overall transmit power $\rho = 8$ is equally split between the two strongest transmit eigenmodes.

Large number of receive antennas

Due to the channel hardening effect it was shown that ergodic capacity essentially becomes independent of fading correlation at the receive antenna array for a large number of receive antennas $R \gg T$. However, this means that also the power allocation coefficients become independent of correlation at the receiver. This is confirmed by the numerical results in Fig. 4.10, where we depict the power allocation coefficient $\phi_i^2$ for a system with $T = 2$ transmit antennas and a varying number of receive antennas resulting from a waterfilling strategy based on the exact EC/MMI analysis. We note that the second coefficient is implicitly given by $\phi_2^2 = \rho - \phi_1^2 = 2 - \phi_1^2$ due to
the power constraint. While there is still a significant discrepancy of the PA coefficient $\phi_1^2$ for up to $R = 12$ receive antennas for a receive correlation coefficient of $r_{RX} = 0.7$ and $r_{RX} = 0.9$, the statistical WF algorithm essentially yields the same PA coefficients for $R = 20$ receive antenna elements independent of the receive fading correlation. Moreover, it can be seen that for a large number of receive antennas, a (almost) uniform PA is optimal for a wide range of transmit correlation values $r_{TX}$, i.e. CDIT has basically no effect on ergodic capacity.

**Fig. 4.9** PA coefficients, high SNR, FCC, ECMM. $T=8$, $R=2$, $r_{RX}=0.7$, variable $r_{TX}$

**Fig. 4.10** Squared PA coefficients, FCC, ECMM. $T=2$, SNR=0 dB, $r_{RX}=$\{0.7,0.9\}
5 Zero-Forcing Receivers

Monte-Carlo simulation studies show that fading correlation between the antenna elements of a wireless system can seriously affect the symbol error rate (SER) of MIMO receivers, especially in case of linear processing. However, in contrast to ergodic MIMO capacity, relatively little is known about the analytical performance of practical receiver types like e.g. linear zero-forcing (ZF) and minimum mean squared error (MMSE) in the presence of correlation. This holds also for the more specialized single-input multiple-output (SIMO) and multiple-input single-output (MISO), respectively, multi-user cellular beamforming applications. While the ZF receiver performance is well analyzed in uncorrelated Rayleigh fading [201], only few results are available on correlated scenarios. To the authors’ best knowledge, analytical results in literature comprise only the case of transmit correlation [47]. Again, as was already observed in the capacity analysis, compared to the uncorrelated case, the mathematical analysis is greatly complicated in the presence of fading correlation. This holds especially for scenarios with fading correlation at the receive antenna array.

Motivated by the need to better characterize and quantify the influence of channel correlation on practical systems, in this chapter we study the performance of a MIMO ZF receiver in a correlated Rayleigh fading environment with arbitrary receive and transmit correlation. To this end, we use results from the theory of matrix variate normal distributions [59] and certain properties of complex Gaussian integrals [137]. In this context, we present an alternative and more flexible derivation for the subchannel SNR statistics in case of arbitrary transmit correlation and unrestricted array sizes given in [49]. Based on the subchannel SNR statistics, closed-form expressions are derived for the symbol error rate. Furthermore, for the first time we present general analytical results for systems with spatial fading correlation at the receiver. Using the eigenvalue distribution of a special random matrix, an exact analysis was presented by the authors in [104]. Due to mathematical difficulties with this approach, the analysis was restricted to the case of $L = 2$ independent subchannels and $R = 2$ receive antennas. In this thesis, we generalize these results to an arbitrary number of independent subchannels and receive antennas. Specifically, we show that the subchannel SNR moment generating function (MGF) can be expressed as the expected value of a ratio of random determinants [115]. We establish closed form formulas for the MGF by exploiting complex Gaussian integrals, such that the SER expressions can be given as a single integral [6][172] by deploying a special representation of the $Q$ function. Furthermore, a novel concise and asymptotically tight bound (at high SNR) on the SER is given for arbitrary system parameters [114]. High SNR asymptotics are derived that allow for a compact and insightful comparison of systems with transmit and receive correlation, respectively. We show via Monte-Carlo simulations the close agreement of theoretical and simulation results. The novel bound is demonstrated to yield surprisingly exact results in the SNR range of interest that should be accurate enough for all practical purposes. An interesting application is the analysis of multi-user smart antenna cellular systems with linear signal processing at the receiver for spatial separation of the various users. Via the novel bound, lengthy simulations can be avoided and the influence of fading correlation can be studied theoretically.
On the other hand, with the availability of adequate short-term (instantaneous) channel state information (CSI) at the transmitter, it is well known that linear transmit prefiltering can significantly improve the performance of MIMO systems with ZF receivers [30]. However, it is difficult if not impossible in many practical systems (e.g. frequency division duplex (FDD) systems) to provide accurate short-term CSI to the transmitter. For that reason, we use the statistical SER analysis derived in this work to design SER minimizing transmit prefilters that are based on the correlation (statistical) properties of MIMO channels and require channel distribution information at the transmitter (CDIT) only. To this end, we make extensive use of majorization theory for solving the optimization problem that occurs in the prefilter design [112]. Interestingly, it is demonstrated that the statistical prefilter designed in this thesis has the same basic structure as its short-term CSI based equivalent. Simulation results indicate a considerable performance gain of the proposed statistical prefilter scheme in strongly correlated MIMO channels.

5.1 Performance Analysis in Rayleigh Fading

In this paragraph, we first derive expressions for the SNR on each of the L subchannels after receive processing, which are the basis for all subsequent derivations. Later, the subchannel SNR statistics are analyzed and the MGF is calculated. Based on this analysis, we present SER expressions and corresponding high SNR asymptotics for various propagation environments that are used for the statistical transmit prefilter design in the second part of this chapter. All derivations are based on the assumption of the Rayleigh fading channel model with Kronecker product covariance structure (2.41).

5.1.1 Zero-forcing receiver basics

Consider the transmission over a correlated flat fading MIMO channel with spatially colored additive Gaussian noise (Fig. 5.1)

$$y = HF s + n.$$  \hspace{1cm} (5.1)

The receiver zero-forcing matrix filter $G$ is given by the combination of a noise whitening filter and the pseudo-inverse of the compound channel consisting of prefilter $F$, channel matrix $H$, and noise whitening filter $R_{nn}^{-1/2}$

$$G = (R_{nn}^{-1/2}HF)^{\dagger} \cdot R_{nn}^{-1/2} = (F^HH^{H}R_{nn}^{-1}HF)^{-1} \cdot F^HH^{H}R_{nn}^{-1},$$  \hspace{1cm} (5.2)

with $GHF = I$. Then denote the signal at the output of the receive filter by $z$ with

![Fig. 5.1 System model with prefilter and ZF receiver](image)
which is split into a signal part
\[ z_s = G H F_s , \]  
and a noise part
\[ z_n = G n . \]

It is then straightforward by the zero-forcing property to show that there is no cross-interference between the subchannels, i.e. we have to take into account only an additive noise term, such that the covariance of the signal part of the received vector \( z_s \) after receive filtering is given by
\[ Z_s = E [ z_s z_s^H ] = R_{ss} = E_s \cdot I \]  
and the covariance of the noise part \( z_n \) reads after simplifying (see also (2.64))
\[ Z_n = E [ z_n z_n^H ] = (F^H H^H R_{nn}^{-1} H F)^{-1} = N_0 \cdot (F^H H^H \tilde{R}_{nn}^{-1} H F)^{-1} . \]  

Without loss of generality, in (5.6) and in the following derivations we assume white transmit signal vectors with \( R_{ss} = E_s \cdot I \) (other choices of the signal covariance can easily be realized by proper choice of the prefilter) and on the other hand we consider an AWGN scenario, i.e. \( R_{nn} = N_0 \cdot I \) (colored noise can be taken into account in the channel receive correlation matrix) to simplify notation (see also Paragraph 2.3.3 on equivalent systems). From above results we can directly derive that the SNR on subchannel \( k \) after the receive filter \( G \) is given by
\[ \gamma_{SC,k} = \frac{\gamma}{\left[ (H F)^\dagger (H F)^\dagger H \right]_{kk}} = \frac{\gamma}{\left[ (H F)^\dagger H F \right]^{-1}} . \]  

Equation (5.8) is the basis for the following paragraph.

### 5.1.2 Subchannel SNR statistics

In this paragraph, we derive compact and flexible expressions for the subchannel SNR statistics. Particularly, it is demonstrated that for transmit correlation only the subchannel SNR is simply Gamma distributed. On the other hand, in the presence of receive fading correlation, the subchannel SNR distribution can not explicitly be given. However, as was mentioned above, in the following paragraphs we study an alternative, MGF based approach that can successfully be applied for the SER analysis in the presence of fading correlation at the receive antenna array.

We start with an expression of the SNR statistics in terms of a random quadratic form in i.i.d. complex normal vectors.

**Theorem 5.1:** The SNR statistics on subchannel \( k \) for a MIMO system with ZF receiver, \( L \) independent subchannels, \( R \) receive antennas, and fully correlated MIMO channel with transmit as well as receive correlation can be expressed as
\( \gamma_{SC,k} \equiv \gamma \beta_k \cdot u^H Qu \equiv \gamma \beta_k \cdot \sum_{l=1}^{L} \lambda_{Q,l} \cdot |u_l|^2, \)  
(5.9)

where \( u \) is a \( R \times 1 \) column vector of unit variance i.i.d. complex Gaussian elements, \( |u_l|^2 \) are exponentially distributed random variables (RV), the constant

\[ \beta_k = \frac{1}{[C^{-1}]_{kk}}, \]  
(5.10)

the auxiliary definitions of the \( L \times L \) matrix

\[ C = F^H \cdot R_{TX} \cdot F, \]  
(5.11)

and the \( R \times R \) random matrix \( Q \)

\[ Q = R_{RX} - R_{RX} \tilde{H}_w (\tilde{H}_w^H R_{RX} \tilde{H}_w)^{-1} \tilde{H}_w^H R_{RX}, \]  
(5.12)

where \( \tilde{H}_w \) is a \( R \times (L-1) \) matrix of complex Gaussian i.i.d. elements, and with eigenvalue decomposition

\[ Q = V_Q \cdot \Lambda_Q \cdot V_Q^H = V_Q \cdot \text{diag}(\lambda_{Q,1}, \lambda_{Q,2}, \ldots, \lambda_{Q,R}) \cdot V_Q^H. \]  
(5.13)

**Proof:** See Appendix 10.1.1.

The determination of the PDF of the subchannel SNR \( \gamma_{SC,k} \) is a difficult problem for the general case with receive correlation. To this end, note in (5.9) that \( \gamma_{SC,k} \) is conditioned on the eigenvalues of \( Q \) - a sum of weighted exponentially distributed variables. However, in case of a transmit correlated channel we get the following simple result, which has been obtained independently in [47][49], where the authors make use of certain marginal statistics of complex Wishart matrices.

**Corollary 5.1:** The SNR \( \gamma_{SC,k} \) on subchannel \( k \) for a ZF receiver in the presence of transmit correlation only, i.e. with uncorrelated fading at the receiver, has a gamma distribution with PDF

\[ f(\gamma_{SC,k}) = \frac{\exp\left(-\frac{\gamma_{SC,k}}{\gamma \beta_k}\right) \left(\frac{\gamma_{SC,k}}{\gamma \beta_k}\right)^{N-1}}{\gamma \beta_k \cdot \Gamma(N)} , \]  
(5.14)

where \( N = R - L + 1 \) is the number of degrees of freedom and \( \beta_k \) is defined in (5.10).

**Proof:** Without receive correlation, we get from Theorem 5.1 using \( R_{RX} = A = I \)

\[ \gamma_{SC,k} \equiv \gamma \beta_k \cdot u^H (I_R - \tilde{H}_w \tilde{H}_w^H \tilde{H}_w^{-1}) u. \]  
(5.15)

However, it can easily be seen that \( \tilde{H}_w \tilde{H}_w^H \tilde{H}_w^{-1} \) is an idempotent matrix with \( \text{rk}(\tilde{H}_w) = L - 1 \) eigenvalues of value 1 and all other eigenvalues of value 0. Thus, the number of eigenvalues of value one of \( I_R - \tilde{H}_w \tilde{H}_w^H \tilde{H}_w^{-1} \) is \( N = R - L + 1 \). Obviously, \( \gamma_{SC,k} \) in (5.15)
is the sum of $N$ independent exponentially distributed random variables, i.e. it has a Gamma distribution \[82\] with $N$ degrees of freedom. \textit{QED.}

5.1.3 \ SER calculation in the presence of transmit correlation only

In this section, the explicit SNR distribution given in Corollary 5.1 is used for calculating the SER performance of the ZF receiver by using a well known approximate conditional SER expression. Moreover, we give simple asymptotic formulas of the SER for the high SNR range. Finally, using the Chernoff bound on the conditional SER, we calculate a simple upper bound on the SER, which will be used for the statistical prefilter design in the second part of this chapter.

\textbf{Exact analysis}

In the following derivations of the ZF receiver performance and the MMSE receiver in the next chapter, we use the following conditional SER (conditioned on the subchannel SNR $\gamma_{SC,k}$) approximation for square $M$-QAM constellations \[153\]

$$P_{s,c}(\gamma_{SC,k}) \approx b \cdot \text{erfc}(\sqrt{c} \gamma_{SC,k}) = 2b \cdot Q(\sqrt{2c} \gamma_{SC,k})$$ \hspace{1cm} (5.16)

with the modulation-dependent parameters

$$b = 2\left(1 - \frac{1}{\sqrt{M}}\right) \quad c = \frac{3}{2(M - 1)}$$ \hspace{1cm} (5.17)

and constellation size $M$. Note that with the Chernoff upper bound on the complementary error function

$$\text{erfc}(\sqrt{ax}) \leq e^{-ax}$$ \hspace{1cm} (5.18)

we can bound the conditional SER via

$$P_{s,c}(\gamma_{SC,k}) \leq b \cdot e^{-c\gamma_{SC,k}}$$ \hspace{1cm} (5.19)

which will be used for the statistical prefilter design in the second part of this chapter.

Having available an exact expression of the subchannel SNR PDF in Corollary 5.1 for the case of transmit correlation only, we can calculate expressions of the SER by direct integration.

\textbf{Theorem 5.2:} The subchannel SER $P_{s,k}(\gamma)$ with ZF receiver and transmit correlated MIMO channel only, i.e. uncorrelated fading at the receiver, can be accurately calculated by

$$P_{s,k}(\gamma) = b \cdot \left[1 - \sqrt{\frac{c \beta_k \gamma}{1 + c \beta_k \gamma}} \cdot \sum_{k=0}^{N-1} \binom{2k}{k} \left(\frac{1}{4(1 + c \beta_k \gamma)}\right)^k\right]$$ \hspace{1cm} (5.20)

with the auxiliary parameter
\[ \beta_k = \frac{1}{[C^{-1}]_{kk}}, \]  

(5.21)

where the \( L \times L \) matrix \( C \) is defined in (5.11), conditional SER approximation (5.19), and diversity parameter \( N = R - L + 1 \).

**Proof:** See Appendix 10.1.2.

### High SNR asymptotics

**Theorem 5.3:** The high SNR asymptotics \( P_{s,k}(\gamma) \) of the SER on subchannel \( k \) derived in Theorem 5.2 with transmit correlation only and ZF receiver are given by

\[ P_{s,k}(\gamma) = b \cdot \frac{\Gamma\left(N + \frac{1}{2}\right)}{\frac{1}{\Gamma\left(\frac{1}{2}\right)} \cdot \Gamma(N + 1)} \cdot (c \beta_k \gamma)^{-N} = b \cdot \frac{1 \cdot 3 \cdot 5 \ldots (2N-1)}{2^N \cdot \Gamma(N+1)} \cdot (c \beta_k \gamma)^{-N}, \]

(5.22)

where \( \beta_k \) is given in (5.21), \( N = R - L + 1 \) is the diversity of the system, and the modulation-dependent parameters \( b \) and \( c \) are defined in (5.17).

**Proof:** Using the hypergeometric function formulation in (10.20) for the SER expression in (5.20), then applying the approximation of the hypergeometric function in (11.84) and simplifying the result using

\[ \Gamma\left(N + \frac{1}{2}\right) = \frac{1 \cdot 3 \cdot 5 \ldots (2N-1)}{2^N \cdot \Gamma(N+1)} \cdot \frac{1}{\Gamma\left(\frac{1}{2}\right)} \]

(5.23)

yields (5.22). **QED.**

### Simple upper bounds on the SER

For a simple assessment of the ZF receiver performance and more importantly, for a closed-form statistical transmit prefilter design in the second part of this chapter, in this section we derive concise upper bounds on the SER.

**Theorem 5.4:** The subchannel SER \( P_{s,k}(\gamma) \) with ZF receiver and transmit correlated MIMO channel (no correlation at the receiver) can be upper bounded by

\[ P_{s,k}(\gamma) \leq b \cdot (1 + c \beta_k \gamma)^{-N}, \]

(5.24)

where \( \beta_k \) is given in (5.21), \( N = R - L + 1 \) is the diversity of the system, and the modulation-dependent parameters \( b \) and \( c \) are defined in (5.17).

**Proof:** Using Chernoff bound approximation (5.19) on the conditional symbol error rate and averaging over the subchannel SNR statistics using
\[ \int_{0}^{\infty} x^a \cdot e^{-bx} \, dx = b^{-(a+1)} \cdot \Gamma(a+1) \quad b > 0 \]  
(5.25)

directly yields (5.24). QED.

From (5.24) it follows directly

**Corollary 5.2:** The subchannel SER with ZF receiver and transmit correlated MIMO channel can be upper bounded by

\[ P_{s,k}(\gamma) \leq b \cdot (c \beta_k \gamma)^{-\frac{N}{L}}. \]  
(5.26)

At this point, we emphasize that the upper bound (5.26) and the exact result in (5.20) differ just by a constant in the high SNR region. This becomes clear by considering (5.22).

### 5.1.4 SER calculation with transmit and receive correlation

In the general case of a fully correlated channel, the SNR in (5.9) is conditioned on the eigenvalues of the random \( R \times R \) matrix \( Q \) - a weighted sum of exponentially distributed RVs. One possible approach to the problem of calculating SER expressions is an integration over the eigenvalue PDF of the random matrix \( Q \). However, for arbitrary correlations and antenna array sizes, it appears that there are no general expressions available in literature on the eigenvalue PDF. Therefore, the eigenvalue PDF based analysis so far is restricted to the case of subchannels with receive antennas \([104]\).

However, in this thesis we present a novel approach to the SER analysis based on the subchannel SNR moment generating function (MGF) and demonstrate that the problem of calculating the SER in the presence of receive correlation can be reduced to finding the expected value of a special ratio of random determinants of matrix quadratic forms, which is an extension of a standard ratio of random quadratic forms \([133]\). To this end, we derive closed form expressions for the expected value by exploiting certain complex Gaussian integrals. We emphasize that this powerful mathematical tool is deployed for the first time in the context of MIMO receiver performance analysis and we expect that it can be successfully applied for solving other problems in communication theory. Based on the subchannel SNR MGF, the subchannel SER can be expressed as a single integral with finite integration limits \([172]\) by deploying a special representation of the \( Q \) function. Furthermore, we propose an asymptotically tight approximation that yields very accurate results for the analytical SER in the SNR range of interest (see the numerical results at the end of this chapter). Again, simple and exact formulas for the high SNR asymptotics are presented for the case of fading correlation at the receive antenna array, allowing for a simple and concise characterization of the influence of fading correlation on the performance of ZF receivers.

Using the integral representation of the complementary error function in (11.69) together with the conditional SER approximation in (5.16) and (5.17), it is well known \([172]\) that the SER \( P_{s,k}(\gamma) \) on subchannel \( k \) can be calculated via the subchannel SNR MGF.
and the single integral

$$P_{s,k}(\gamma) = \frac{2b}{\pi} \int_0^{\pi/2} M_{\gamma,k}(s) \frac{c}{\sin^2 \theta} d\theta.$$  \hspace{1cm} (5.28)

Furthermore, with Gray encoding and the assumption that a wrong symbol decision causes just one single bit error, which is fulfilled for higher SNR, we can approximate the bit error rate (BER) on subchannel $k$ for a modulation size $M$ by

$$P_{b,k}(\gamma) \approx \frac{1}{\log_2 M} \cdot P_{s,k}(\gamma).$$ \hspace{1cm} (5.29)

First, we derive an explicit expression of the MGF of the subchannel SNR.

**Theorem 5.5:** The MGF $M_{\gamma,k}(s)$ of the subchannel SNR in a correlated Rayleigh fading environment with MIMO ZF receiver and AWGN is given by

$$M_{\gamma,k}(s) = E_{\gamma_{\text{SC},k}}\left[ \exp(-s\gamma_{\text{SC},k}) \right],$$ \hspace{1cm} (5.30)

where $\tilde{H}_w$ is a $R \times (L-1)$ matrix of i.i.d. complex Gaussian elements, $\beta_k$ is defined in (5.10), and the $R \times R$ matrix $O = \text{eig}(R_{RX})$ given in (2.70).

**Proof:** Using the subchannel SNR expression (5.9) and $\alpha_k = \gamma \beta_k$ we find

$$M_{\gamma,k}(s) = E_{\gamma_{\text{SC},k}}\left[ \exp(-s\alpha_k \cdot u^H Q u) \right].$$ \hspace{1cm} (5.31)

Then applying the expected value formula (11.37) for complex normal vectors we get

$$M_{\gamma,k}(s) = E_{\gamma_{\text{SC},k}}\left[ \exp(-s\alpha_k \cdot u^H Q u) \right] = E_Q \left[ \frac{1}{|I + s\alpha_k Q|} \right].$$ \hspace{1cm} (5.32)

Together with the expression of $Q$ in (5.12) it can be shown that

$$M_{\gamma,k}(s) = E_{\tilde{H}_w}\left[ \frac{1}{|I + s\alpha_k(O - \tilde{O} \tilde{H}_w (\tilde{H}_w^H O \tilde{H}_w)^{-1} \tilde{H}_w^H O)|} \right].$$ \hspace{1cm} (5.33)

Now factoring out $|I + c \alpha_k O|$ in the denominator, exploiting determinant identity (11.14) and multiplying nominator and denominator by $|\tilde{H}_w^H O \tilde{H}_w|$ yields the result (5.30). \textit{QED}.

It appears that there are no results available in literature on calculating the expected value of the ratio of determinants in (5.30) in the general case of arbitrary number of antenna elements and subchannels. In this thesis, for the first time we determine exact closed form expressions by exploiting certain complex integrals. To this end, we first derive the following lemma.
Lemma 5.1: Let $G$ be a $m \times n$ ($m \geq n$) matrix of i.i.d. complex Gaussian elements, whereas $\tilde{A} = \text{diag}(a_1, a_2, \ldots, a_m)$ and $\tilde{B} = \text{diag}(b_1, b_2, \ldots, b_m)$ are positive definite diagonal $m \times m$ matrices, whereas the diagonal elements of $\tilde{B}$ differ. Then

$$r(\tilde{A}, \tilde{B}, m, n) = E_G \left[ \left| G^H \tilde{A} G \right| / \left| G^H \tilde{B} G \right| \right] = \sum_{\tilde{\alpha}_n} \left| \tilde{A}^{\tilde{\alpha}_n} \right| \cdot \sum_{k = 1}^{n} \int_0^\infty \frac{t^{n-1}}{(1 + \left\{ \left\{ \tilde{B} \right\}^{\tilde{\alpha}_n} \right\}_{kk} \cdot t) |I + t\tilde{B}|} dt, \quad (5.34)$$

where $\left| \tilde{A}^{\tilde{\alpha}_n} \right|$ is the determinant of the matrix that results from selecting the row and column subset $\tilde{\alpha}_n$ from matrix $\tilde{A}$, $\left\{ \tilde{B} \right\}^{\tilde{\alpha}_n}$ is the $n \times n$ matrix that results from selecting the row and column subset $\tilde{\alpha}_n$ from matrix $\tilde{B}$, and $\left\{ \left\{ \tilde{B} \right\}^{\tilde{\alpha}_n} \right\}_{kk}$ is its $k$th diagonal element. The integrals in (5.34) can be calculated in closed form and we obtain for the prototype integral with $b_j = \left\{ \left\{ \tilde{B} \right\}^{\tilde{\alpha}_n} \right\}_{kk}$

$$I_{rat}(\tilde{B}, n, j) = \int_0^\infty \frac{t^{n-1}}{(1 + b_j \cdot t) |I + t\tilde{B}|} dt = K_1(\tilde{B}, n, j) + K_2(\tilde{B}, n, j), \quad (5.35)$$

where

$$K_1(\tilde{B}, n, j) = (-1)^n \cdot \sum_{k = 1 \atop k \neq j}^{m} \frac{\ln(b_k) \cdot b_k^{m-n}}{(b_k - b_j) \prod_{l = 1 \atop l \neq k}^{m} (b_k - b_l)} \quad (5.36)$$

and

$$K_2(\tilde{B}, n, j) = \frac{b_j^{m-n-1} \cdot \left[ 1 - \ln(b_j) \cdot \left( \sum_{l = 1 \atop l \neq j}^{m} \frac{b_l}{b_j - b_l} + n - 1 \right) \right]}{\prod_{l = 1 \atop l \neq j}^{m} (b_j - b_l)} \quad (5.37)$$

Proof: See Appendix 10.1.3.

We note that the expected value in (5.34) reduces to a standard ratio of random quadratic forms for $n = 1$, whereas $G$ becomes a $m \times 1$ column vector. The interested reader is referred to [133] for a detailed overview of results on random quadratic forms. Results on ratios of random quadratic forms are given in [84][123][133][136][175][176][191].

With the help of Lemma 5.1 we can directly get from Theorem 5.5 the following closed form expression for the subchannel SNR MGF.

Theorem 5.6: The MGF $M_{\gamma,k}(s)$ of the subchannel SNR in a correlated Rayleigh fading environment with MIMO ZF receiver and AWGN is given by
where \( r(\mathbf{A}, \mathbf{B}, m, n) \) is given in Lemma 5.1, \( \beta_k \) is defined in (5.10), and the \( R \times R \) matrix \( \mathbf{O} = \text{eig}(\mathbf{R}_{RX}) \) given in (2.70).

With the closed form MGF expression in (5.38) and formula (5.28), the SER calculation can be reduced to a simple integral with finite integration limits that can be evaluated in a straightforward manner with standard math software like e.g. Matlab. Results will be presented below in the numerical results section at the end of this chapter. Due to the complexity of the MGF expression (this is especially the case for a greater number of independent subchannels), however, we present a different approximate approach to the problem of determining the expected value in (5.30) and (5.34), respectively. We propose the following tight approximation on the MGF.

**Theorem 5.7:** The MGF \( M_{y,k}(s) \) of the SNR on subchannel \( k \) according to Theorem 5.5 can be approximated by

\[
M_{y,k}(s) \approx M_{y,k}^A(s) = \frac{1}{\left| I + s\gamma \beta_k \mathbf{O} \right|} \cdot \frac{\text{tr}_{L-1}(\mathbf{O})}{\text{tr}_{L-1}(\mathbf{O}(I + s\gamma \beta_k \mathbf{O})^{-1})},
\]

where \( \text{tr}_n(\mathbf{X}) \) is the elementary symmetric function of order \( n \) of the eigenvalues of matrix \( \mathbf{X} \). The approximation is exact in the high SNR regime and reads

\[
M_{y,k}(s) = \frac{\text{tr}_{L-1}(\mathbf{O})}{|\mathbf{O}| \cdot \binom{R}{L-1}} \cdot (s\beta_k)^{-N} \cdot \gamma^{-N},
\]

with diversity parameter \( N = R - L + 1 \), and \( \beta_k \) defined in (5.10).

**Proof:** See Appendix 10.1.4.

Again, with the closed form MGF approximation in (5.40) and formula (5.28), the (approximate) SER calculation can be reduced to a simple integral that can be solved numerically. It turns out that the approximation is very tight and results are presented at the end of this chapter. On the other hand, in the high SNR region we can solve the integral in (5.28) in closed form to derive exact SER formulas.

**Theorem 5.8:** The SER asymptotics on subchannel \( k \) \( \overline{P}_{s,k}(\gamma) \) in the high SNR region with fully correlated (transmit as well as receive correlation) Rayleigh fading MIMO channel with AWGN and ZF receiver is given by

\[
\overline{P}_{s,k}(\gamma) = b \cdot \frac{\text{tr}_{L-1}(\mathbf{O})}{|\mathbf{O}| \cdot \binom{R}{L-1}} \cdot \frac{1 \cdot 3 \cdot 5 \ldots (2N-1)}{2^N \cdot \Gamma(N+1)} \cdot (c\beta_k)^{-N} \cdot \gamma^{-N},
\]

where \( b \) is given in (5.38), \( R(x) \) is defined in (5.10), and \( \mathbf{O} = \text{eig}(\mathbf{R}_{RX}) \) given in (2.70).
where $b$ and $c$ are modulation-dependent parameters according to (5.17), $N = R - L + 1$ is the diversity of the system, $\beta_k$ is defined in (5.10), and $O = \text{eig}(R_{RX})$. Again, $\text{tr}_n(X)$ is the elementary symmetric function of order $n$ of the eigenvalues of matrix $X$ according to Appendix 11.1.6.

**Proof:** Using the exact asymptotic MGF formula (5.40) in (5.28), we directly arrive at (5.41) with the integration result [50]

$$
\frac{\pi}{2} \int_{0}^{\pi} \sin^{2N}(\theta) d\theta = \frac{\sqrt{\pi} \cdot \Gamma\left(\frac{1}{2} + N\right)}{2 \cdot \Gamma(1 + N)} = \frac{1 \cdot 3 \cdot 5 \ldots (2N - 1)}{2^N \cdot \Gamma(N + 1)} \cdot \frac{\pi}{2}.
$$

(5.42)

**QED.**

Interestingly, the exact asymptotical SER expression in Theorem 5.8 is a formal proof the fact that the diversity $N$ of a MIMO system with ZF receiver is not dependent on the correlation properties of the channel. We note that as expected, (5.41) reduces in the absence of receive correlation ($O = I$) to the results already derived for the transmit correlated case in (5.22). Another special case is given for a single transmit data stream only. In this case, the ZF receiver reduces to the well-known maximum ratio combiner (MRC) and we find from Theorem 5.8 with $N = R$ and $\beta_1 = 1$

**Corollary 5.3:** With $L = 1$ subchannel, the ZF receiver reduces to a MRC and the SER in the high SNR region with receive correlated Rayleigh fading SIMO channel is given by

$$
P_{s}^{\text{MRC}}(\gamma) = \frac{b}{|O|} \cdot \frac{1 \cdot 3 \cdot 5 \ldots (2R - 1)}{2^R \cdot c^R \cdot \Gamma(R + 1)} \cdot \gamma^R.
$$

(5.43)

It is a simple exercise to directly obtain this result from a MGF approach for analyzing the MRC and the calculation is left to the reader.

**5.1.5 Coding and diversity gain**

It has been observed in different contexts that SER expressions $\overline{P}_s(\gamma)$ of uncoded and coded systems at high SNR can in certain cases be approximated by [153][184][198]

$$
P_s(\gamma) = (G_c \cdot \gamma)^{-N},
$$

(5.44)

where $G_c$ is the so-called coding gain and $N$ is the diversity of the system. From the high SNR asymptotics of the SER performance of ZF receivers given in (5.41) we obtain the subchannel specific coding gain $G_{c,k}$ and the diversity $N$

$$
G_{c,k} = \left(b \cdot \frac{\text{tr}_{L-1}(O)}{|O|} \cdot \frac{1 \cdot 3 \cdot 5 \ldots (2N - 1)}{2^N \cdot \Gamma(N + 1)} \right)^{-1/N} \cdot c\beta_k \quad N = R - L + 1.
$$

(5.45)
5.1.6 Asymptotic SNR shift due to fading correlation

The coding gain in (5.45) incorporates the influence of transmit as well as receive correlation. Both effects shall be separated in the following.

**Corollary 5.4:** The asymptotic SNR shift in the high SNR regime of the SER curve in dB $\Delta \gamma_{dB, RX}$ due to receive fading correlation is given by

$$\Delta \gamma_{dB, RX} = \frac{10}{R - L + 1} \cdot \log_{10} \left( \frac{R}{L - 1} \cdot |R_{RX}| \right) \quad [dB] \quad (5.46)$$

and is independent of the transmit correlation and the subchannel index $k$.

On the other hand we find

**Corollary 5.5:** The subchannel specific asymptotic shift in the high SNR regime of the SER curve $\Delta \gamma_{dB, TX}(k)$ due to transmit fading correlation for subchannel $k$ is given by

$$\Delta \gamma_{dB, TX}(k) = 10 \cdot \log_{10} (|C^{-1}_{kk}|) \quad [dB], \quad (5.47)$$

where the $L \times L$ matrix $C$ is defined in (5.11).

As in case of the ergodic capacity analysis, again the asymptotic performance penalties due to transmit and receive correlation are independent of each other and thus can be separated. This is a consequence of the channel model with Kronecker product covariance structure.

5.2 Prefilter Design for ZF Receivers

The closed-form results on the SER derived above in the first part of this chapter can effectively be exploited for the derivation of a linear matrix transmit prefilter that minimizes SER. Adaptation of the prefilter requires only CDIT and thus allows for an implementation in systems, where no instantaneous channel state information is available. Moreover, estimation of the statistical properties of the channel can take place in a wide time window and the estimates can therefore be assumed to be perfect, i.e. the prefiltering scheme becomes very robust. Last but not least, the interval between two updates of the prefilter is long, therefore minimizing the computational requirements. We note that the matrix transmit prefilter approach presented in this thesis can be considered as a generalization of transmit antenna selection schemes (e.g. [48][69]). Basically, the transmit prefilter has more degrees of freedom and is thus more flexible.

Again, majorization theory [131], which has been successfully applied for the characterization of the influence of fading correlation on ergodic capacity, proves to be a powerful tool for solving the SER optimization problem.
5.2.1 Statistical prefilter design based on majorization theory

For finding the optimum transmit prefilter in the sense that it minimizes the resulting SER, in principle we could use the exact SER analysis for arbitrary correlated MIMO channels or the approximate MGF based approach in Theorem 5.7. However, we would have to resort to lengthy numerical optimization techniques. The insights gained from numerical optimizations are limited and the complexity of such techniques obviates an efficient implementation. Therefore, in the following we focus on the exact SER formulas derived for the high SNR regime in Theorem 5.8. This has major advantages. First of all, it is demonstrated that a closed-form prefilter design can be derived. Moreover, the effects of fading correlation at transmit and receive antenna array can be separated completely (see Theorem 5.8), thus leading to a transmit prefilter that is independent of receive fading correlation. Besides, the prefilter becomes independent of the SNR, thus avoiding the need for costly SNR measurements and further simplifying the prefilter. Finally, simulation results demonstrate that the proposed prefilter design is effective in the SER range of interest, thus again confirming its practical importance.

In order to find the optimum prefilter $F_{opt}$ that minimizes the cumulative error probability of all subchannels in the high SNR regime, it can be shown from (5.41) that we have to solve the equivalent constrained optimization problem

$$F_{opt} = \arg \min_F \sum_{k=1}^{L} \{(F^H R_{TX} F)^{-1})_{kk}\}^N \quad \text{s.t.} \quad \text{tr}(FF^H) = \rho,$$

(5.48)

where the transmit power is restricted to $\rho$. The solution to problem (5.48) is given in

**Theorem 5.9:** The optimum SER minimizing prefilter based on CDIT for a MIMO system with ZF receiver in the sense of (5.48) is given by

$$F_{opt} = \tilde{V}_{TX} \cdot \Phi_{ZF} \cdot D_L,$$

(5.49)

where $D_L$ is a $L \times L$ discrete Fourier transform (DFT) matrix (see Appendix 11.1.1) and $\tilde{V}_{TX}$ is a matrix of the eigenvectors corresponding to the $L$ strongest eigenvalues of the eigenvalue decomposition (EVD) in

$$R_{TX} = \begin{bmatrix} V_{TX} & \tilde{V}_{TX} \end{bmatrix} \begin{bmatrix} \Lambda_{TX} & \end{bmatrix} \begin{bmatrix} V_{TX} & \tilde{V}_{TX} \end{bmatrix}^H,$$

(5.50)

where the matrix $\tilde{\Lambda}_{TX}$ contains the $L$ largest eigenvalues in increasing order. The diagonal power allocation matrix $\Phi_{ZF}$ reads

$$\Phi_{ZF} = \left( \frac{\rho}{\text{tr}\left(\tilde{\Lambda}_{TX}\right)} \right)^{1/2} \cdot \tilde{\Lambda}_{TX}^{1/4}.$$

(5.51)

**Proof:** See Appendix 10.2.1.
5.2.2 Prefilter design based on short-term CSI

An interesting property of the statistical prefilter design presented in Theorem 5.9 is its resemblance of the short-term CSI based prefilter for ZF receivers given in [30]. For completeness, without proof we outline the design of the ST prefilter.

**Theorem 5.10:** A possible choice for a SER minimizing prefilter based on short-term channel state information has the structure

\[
F_{opt}^{ST} = \tilde{V}_H \cdot \Phi_{ZF}^{ST} \cdot D_L, \tag{5.52}
\]

where \(D_L\) is a DFT matrix of size \(L\), \(\Phi_{ZF}^{ST}\) is a diagonal power allocation matrix with

\[
\Phi_{ZF}^{ST} = \left( \rho / \text{tr} \left( \tilde{\Lambda}_H^{-1/2} \right) \right)^{1/2} \cdot \tilde{\Lambda}_H^{-1/4}, \tag{5.53}
\]

and the eigenvalue decomposition

\[
H^HH = \begin{bmatrix} V_H & \tilde{V}_H \end{bmatrix} \begin{bmatrix} \Lambda_H & 0 \\ 0 & \Lambda_H \end{bmatrix} \begin{bmatrix} V_H & \tilde{V}_H \end{bmatrix}^H, \tag{5.54}
\]

where the \(L \times L\) diagonal matrix \(\tilde{\Lambda}_H\) contains the \(L\) largest eigenvalues.

Obviously, the basic structure of the ST CSI and the CDIT based prefilters agree. Furthermore, while the ST prefilter is based on the instantaneous channel state via (5.54), the statistical prefilter is based on its expected value via (5.50), as we have for the correlated Rayleigh fading MIMO channel with (2.51)

\[
E[H^HH] = \text{tr} \left( R_{RX} \right) \cdot R_{TX}. \tag{5.55}
\]

At this point, we mention that the same analogy can be established for a transmit prefilter for MIMO minimum mean squared error (MMSE) receivers (see Chapter 6). Furthermore, it was shown that the statistical waterfilling based on the loose ergodic capacity bound presented in Paragraph 4.3.3 also resembles the optimum algorithm based on ST CSI.

5.3 Numerical Results

In this section we study the performance of MIMO systems in correlated Rayleigh fading environments deploying a ZF algorithm at the receiver. Specifically, we demonstrate the validity of our theoretical studies and the tightness of the analytical SER expressions. Furthermore, we show the effectiveness of the proposed statistical transmit prefiltering algorithms that can achieve a significant SER reduction in MIMO systems with non-vanishing transmit fading correlation.
5.3.1 BER performance

Exact BER analysis

In Fig. 5.2 we have depicted BER curves for a MIMO system with $L = 4$ independent subchannels, $T = 4$ transmit antennas (i.e. each subchannel is mapped on one transmit antenna), and a varying number of receive antennas $R = \{4, 6, 8, 10\}$. We consider a semi correlated channel (SCC) with correlation at the receiver only ($r_{RX} = 0.7$) according to the exponential correlation matrix model (ECMM). A perfect agreement between the analytical curves according to (5.28)-(5.29) with the MGF from Theorem 5.6 and the simulated curves can be observed. Furthermore, we have included the asymptotics according to Theorem 5.8, which clearly indicate the increasing diversity with a greater number of receive antennas.

Tight bound on BER

The negative impact of fading correlation at the receive antenna array on a MIMO system with ZF receiver and $T = 4$ transmit antennas, $L = 4$ independent subchannels, and $R = 6$ receive antennas can be observed in Fig. 5.3. At the transmitter, the fading correlation between the antenna elements is kept constant with $r_{TX} = 0.7$, while the correlation at the receiver is varied. In contrast to Fig. 5.2, now the theoretical curves stem from the bound expression derived above in Theorem 5.7 in combination with (5.28)-(5.29). However, it can be seen that the bound is very accurate even for lower SNR values. We note that for the given system parameters, a diversity of $N = 3$ results.
Zero-Forcing Receivers

In Chapter 6 it is demonstrated that the performance analysis of MIMO MMSE receivers exhibits a significantly higher complexity than the analysis of ZF receivers. However, it is well known that the performance of MMSE and ZF receivers is very similar in the high SNR regime. Therefore, the low-complexity tight BER bound from Theorem 5.7 with (5.28)-(5.29), which was originally derived for MIMO ZF receivers, can effectively serve as an upper bound on the performance of MMSE receivers. This is demonstrated for a system with \( T = L = 4 \) and \( R = 6 \) in Fig. 5.4. Obviously, the approximation is tight with less than 1 dB difference in SNR in the high SNR regime. As was pointed out above, the novel bound on ZF BER performance can therefore serve as a valuable means for analyzing multi-user smart antenna cellular systems in correlated Rayleigh fading environments, where the base stations deploy linear optimum combining algorithms, which are just a special form of MIMO MMSE receive algorithms. The novel bound allows for a quick assessment of the negative impact of fading correlation, thereby deepening the insight in the overall system and helping to avoid the need for lengthy Monte-Carlo simulations.

### 5.3.2 Statistical transmit prefiltering

An example of a \( T = R = 4 \) system, where only a subset \( L = 2 \) of the available subchannels is used, is given in Fig. 5.5. We consider a system with uncorrelated receiver and fading correlation according to the realistic correlation matrix model (RCMM) at the transmitter, whereas the angular spread is \( \Delta_{TX} = 2^\circ \) and \( \Delta_{TX} = 10^\circ \), respectively. Curves are plotted for a system without transmit prefilter (in this case, only the outer two transmit antennas are used), with statistical prefilter based on channel distribution information at the transmitter (CDIT), with prefilter based on...
ST CSI, and finally for reference we have given the performance of a system with an uncorrelated channel (UCC). For the case of less fading correlation, it can be observed that with statistical prefiltering we can outperform the system with uncorrelated fading. Essentially, the statistical prefilter can beneficially exploit the beamforming-like capabilities of the transmit antenna array. However, ST CSI at the transmitter with corresponding ST transmit prefiltering can achieve a significant additional performance gain. On the other hand, in case of strong fading correlation $\Delta_{TX} = 2^\circ$, CDIT with adequate statistical prefiltering becomes more effective with a greater per-
performance improvement compared to the case without prefilter. Moreover, the gap between ST and LT prefiltering is only small.

Similar observations are possible in Fig. 5.6 with $T = L = 4$ and $R = 6$. Now we have plotted curves for a system with FCC, whereas the receive correlation is fixed $r_{RX} = 0.5$ according to the ECMM, while the transmit correlation is varied. There is an increasing gain due to statistical prefiltering with increasing fading correlation at the transmit antenna array. However, for the given transmit correlation parameter range, there is a significant gap between prefiltering based on CDIT and ST CSI.
6 Minimum Mean Squared Error Receivers

In the first part of this chapter, corresponding to the proceeding in case of the ZF receiver, we consider the symbol error rate (SER) performance of a MIMO link with linear prefilter at the transmitter and minimum mean squared error (MMSE) receiver. It turns out that there is a close connection to the analysis of the optimum combining (OC) algorithm in the context of smart antenna signal processing, such that the considered MIMO system with MMSE receiver is a generalization of smart antenna cellular systems with multiple users and optimum combining algorithm at the base station. The subchannel signal to interference plus noise ratio (SINR) statistics after OC processing are very well analyzed in case of vanishing fading correlation, e.g. in [39], where the authors base their analysis on a special random quadratic form that was studied in the statistical literature [94] to derive the exact distribution of the subchannel SINR. Recently, an exact analysis of the SER with OC for the uncorrelated case based on the eigenvalue distribution of a complex Wishart matrix [81] was presented in [128]. Other approximate results on the performance of OC in uncorrelated Rayleigh fading environments can be found e.g. in [3][62][165][166].

However, only little is known about the exact performance of OC in correlated fading environments, i.e. in presence of fading correlation at the receive antenna array, and it appears that known mathematical approaches are not suited for finding a solution. Available results in literature are semi-analytic or use approximations [28][150]. On the other hand, exact non-asymptotic analytical results on the performance of MIMO MMSE receivers appear not to exist at all in literature. In order to fill this gap, in this thesis we derive for the first time analytical expressions of the SER of MIMO MMSE receivers for arbitrarily correlated Rayleigh fading MIMO channels. We emphasize that for the first time this also includes the case of fading correlation at the receiver, which generalizes a result for systems with \(L = 2\) subchannels and \(R = 2\) receive antennas that was presented by the authors in [105]. To this end, as in case of the ZF receiver in Chapter 5, we make use of a moment generating function (MGF) approach, which again leads to the problem of calculating the expected value of a ratio of random determinants. Using complex Gaussian integrals, we present closed-form expressions for these expected values, which appear to be not available in statistical literature. Based on that, exact non-asymptotic (i.e. finite number of antenna elements) SER expressions are given in terms of single integral formulas [115]. Moreover, we derive SER asymptotics for the high SNR regime, which allow for a simple and concise assessment of the influence of the various system parameters. Specifically, we demonstrate the importance of certain elementary symmetric functions of the eigenvalues of the correlation matrices on system performance.

In the second part of this chapter, we outline the design of a linear transmit prefilter that minimizes the overall MSE at the output of a MIMO MMSE receiver, while simultaneously minimizing SER [101][102][112]. Again, the prefilter is based on statistical distribution information at the transmitter (CDIT) only, thus allowing for a low-complexity robust implementation e.g. in frequency division duplex (FDD) systems. The prefilter optimization problem is solved by majorization theory and we present a closed-form concise matrix formulation of the prefilter. Interestingly,
we find again a short-term long-term duality, i.e. the basic structure of transmit prefilters based on short-term channel state information (CSI) and CDIT is the same.

6.1 Subchannel SINR Analysis

Similar to the proceeding in case of the ZF receiver, in this section we first derive general expressions of the subchannel SINR after MMSE processing at the receiver with arbitrarily correlated MIMO channels. Subsequently, we analyze the SINR statistics, which exhibit an increased complexity compared to the ZF case. However, it is demonstrated that the SINR statistics simplify significantly under the assumption of long-term (LT) eigenmode (EM) transmission.

6.1.1 MMSE receiver basics

Without loss of generality according to the comments on equivalent systems in Paragraph 2.3.3, in order to simplify notation we consider the transmission over a correlated flat fading MIMO channel in the presence of AWGN (Fig. 6.1)

\[ y = HF_s + n \]  

with signal and noise covariance matrices

\[ R_{ss} = E_s \cdot I \quad R_{nn} = N_0 \cdot I. \]  \hspace{1cm} (6.2)

As a basis for all following SER calculations, we give a general expression for the subchannel SINR after the MMSE receive filter \( G \) in the following theorem. It appears that only special cases of this theorem can be found in literature (e.g. [138][192]) in different contexts and therefore we present a more detailed derivation.

**Theorem 6.1:** The SINR on subchannel \( k \) of a MIMO system with transmit prefilter \( F \), MMSE receiver, and covariances according to (6.2) is given by

\[ \gamma_{SC,k} = \gamma \cdot k_k^H Q_k^{-1} k_k = \frac{1}{1 + \gamma \cdot k_k^H Q_k^{-1} k_k} - 1 = \frac{1}{[\gamma \cdot KK^H + I_R]^{-1}_{kk}} - 1 \]  

with the auxiliary (subchannel dependent) \( R \times R \) matrix

\[ Q_k = \gamma \cdot \tilde{K}_k \tilde{K}_k^H + I_R, \]  \hspace{1cm} (6.4)
the compound $R \times L$ matrix $K$ consisting of matrix channel and prefilter matrix, and its decomposition with $k$th $R \times 1$ column vector $k_k$ and auxiliary $R \times (L - 1)$ matrix $\tilde{K}_k$

$$K = HF \quad K = \begin{bmatrix} K_{1,k} & k_k \\ K_{2,k} \end{bmatrix} \quad \tilde{K}_k = \begin{bmatrix} K_{1,k} & K_{2,k} \end{bmatrix}. \quad (6.5)$$

The MSE $\varepsilon_k$ on subchannel $k$ reads

$$\varepsilon_k = \frac{1}{1 + \gamma \cdot k_k^H Q_k^{-1} k_k}. \quad (6.6)$$

Furthermore, there is a close relation between the SINR and the mean squared error on subchannel $k$

$$\gamma_{SC,k} = \gamma \cdot k_k^H (\gamma \cdot \tilde{K}_k \tilde{K}_k^H + I)^{-1} k_k = \frac{1}{\varepsilon_k} - 1. \quad (6.7)$$

**Proof:** See Appendix 10.3.1.

### 6.1.2 Subchannel SINR statistics

In this paragraph, we derive statistically equivalent expressions for (6.7) that facilitate a further analysis. To this end, we first consider the general case with arbitrary transmit correlation and transmit prefilter. Due to the complexity of the resulting formulas and the space limitation, we are then restricting ourselves in this thesis to the practically important case of long-term eigenmode (LT EM) transmission, which is shown to significantly simplify the subchannel SINR expressions.

**General case**

For notational simplicity, again we first focus on the SINR $\gamma_{SC,1}$ on subchannel 1 and generalize the results later to an arbitrary subchannel. From Theorem 6.1 we have

$$\gamma_{SC,1} = \gamma \cdot k_1^H (\gamma \cdot \tilde{K}_1 \tilde{K}_1^H + I)^{-1} k_1 = k_1^H (\tilde{K}_1 \tilde{K}_1^H + \frac{1}{\gamma} I_R)^{-1} k_1. \quad (6.8)$$

The marginal distributions of $k_1$ and $\tilde{K}_1$ were derived in the context of the ZF receiver and with $C = F^H \cdot R_{TX} \cdot F$ and the partitioning according to (11.10) it was shown that [see (10.11)]

$$k_1 \equiv x + \tilde{K}_1 a \quad a = (C_{22}^*)^{-1} C_{21}^* \quad (6.9)$$

with random vector and matrix variables (see also Appendix 11.2.1)

$$x \sim N_{R,1}(0, C_{11}^* \cdot R_{RX}) \quad \tilde{K}_1 \sim N_{R,L-1}(0, R_{RX} \otimes C_{22}^*). \quad (6.10)$$

Plugging (6.9) into (6.8) results in
\[ \gamma_{SC, 1} = (x + \tilde{K}_1 a)^H (\tilde{K}_1 \tilde{K}_1^H + \frac{1}{\gamma} I_R)^{-1} (x + \tilde{K}_1 a). \] (6.11)

Unfortunately, unlike the ZF case, the expression in (6.11) cannot be further simplified. The SINR expression in (6.11) is in fact a random quadratic form in the complex normally distributed vector \( x + \tilde{K}_1 a \), however, conditioned on \( \tilde{K}_1 a \), the term \( x + \tilde{K}_1 a \) is non-central, i.e. it has non-zero mean, thus making (6.11) a non-central random quadratic form. In its general form (6.11), the subchannel SINR statistics in the presence of correlation at the transmitter complicate the analysis and thus in this thesis we put the focus on long-term eigenmode transmission.

**Long-term eigenmode transmission**

Long-term (LT) eigenmode (EM) transmission is an important and practically relevant case. First, it can be shown that LT EM transmission corresponds to the case of multiple unequal power users in a multi-user smart antenna cellular environment with optimum combining algorithm at the base station. Second, we will show below in the second part of this chapter that LT EM transmission with proper power allocation is the average MSE minimizing strategy. In this section, it is demonstrated that the subchannel SINR expression in (6.11) greatly simplifies for LT EM transmission, thus allowing for reduced complexity SER analysis. Specifically, we again arrive at a certain random quadratic form in central (mean value free) complex normal vectors.

**Theorem 6.2:** With LT EM transmission, the subchannel SINR on subchannel \( k \) after MMSE processing at the receiver can be expressed as

\[ \gamma_{SC, k} \equiv c_k \cdot y^H \left( \bar{Y}_k \bar{Y}_k^H + \frac{1}{\gamma} I_R \right)^{-1} y \] (6.12)

with the auxiliary random \( R \times (L - 1) \) matrix

\[ \bar{Y}_k \sim \mathcal{N}_{R, L - 1}(0, R_{RX} \otimes \tilde{C}_k), \] (6.13)

the auxiliary complex normal \( R \times 1 \) random vector \( y \)

\[ y \sim \mathcal{N}_{R, 1}(0, R_{RX}), \] (6.14)

the auxiliary \( L \times L \) matrix \( C \), defined in (5.11), which for LT EM transmission with \( F = \tilde{V}_{TX} \cdot \Phi \) simplifies to

\[ C = F^H \cdot R_{TX} \cdot F = \tilde{A}_{TX} \cdot \Phi^2 = \text{diag}(c_1, c_2, \ldots, c_L), \] (6.15)

the partitioning, where \( c_k \) is the \( k \)th diagonal element

\[ C = \text{blkdiag}(C_{1, k}, c_k, C_{2, k}) \quad \tilde{C}_k = \text{blkdiag}(C_{1, k}, C_{2, k}), \] (6.16)
the EVD of $R_{TX}$ in (2.42), such that $\Lambda_{TX} = \text{diag}(\lambda_{TX,1}, \ldots, \lambda_{TX,L})$ is a $L \times L$ matrix of a subset of eigenvalues with the associated matrix $\tilde{V}_{TX}$ of a subset of $L$ eigenvectors, and diagonal power allocation (PA) matrix $\Phi = \text{diag}(\phi_1, \ldots, \phi_L)$.

**Proof:** In case of LT EM transmission, the transmit prefilter takes on the form $F = \tilde{V}_{TX} \cdot \Phi$, resulting with the notation of Appendix 11.2.1 in a diagonal matrix $C$ according to (6.15) with $C_{21} = 0$ and from (6.9) $a = 0$. Furthermore, we find $C_{11.2} = \lambda_{TX,1} \cdot \phi_1^2 \equiv c_1$. Applying these results to (6.11), we get with (6.10) and (6.14) the simplified SINR expression for subchannel 1

$$\gamma_{SC,1} \equiv c_1 \cdot y^H (\tilde{Y}_1 \tilde{Y}_1^H + \frac{1}{\gamma} I_R)^{-1} y. \quad (6.17)$$

After generalizing the results to an arbitrary subchannel, we find Theorem 6.2. QED.

### 6.2 Subchannel SINR Moment Generating Functions

Similar to the approach that was used for analyzing the SER performance of ZF receivers in the presence of receive fading correlation, again we present a MGF based method for analyzing the MMSE receiver. As was already noted, however, the resulting expressions exhibit a significantly increased mathematical complexity. Moreover, as in case of the ergodic capacity analysis, we have to differentiate between various propagation scenarios with varying fading correlation at transmitter and receiver. Again, we focus on the case of LT EM transmission. However, we emphasize that a more general derivation is possible at the cost of a further increase in complexity.

#### 6.2.1 MGFs in terms of ratios of random determinants

In this paragraph, we present expressions for the subchannel SINR MGFs in terms of the expected value of ratios of certain random determinants, which are the basis for all following derivations.

**Theorem 6.3:** With eigenmode transmission and AWGN, the MGF of the subchannel SINR after MMSE receive processing for subchannel $k$ is given by

$$M_{\gamma}(s,k) = \frac{1}{[1 + s c_k \gamma O]} \cdot E_{\tilde{H}_w} \left[ \frac{\gamma \tilde{H}_w^H O \tilde{H}_w \tilde{C}_k + I_{L-1}}{\left(1 + s c_k \cdot I \right)^{-1} \tilde{H}_w \tilde{C}_k + I_{L-1}} \right], \quad (6.18)$$

where $\tilde{H}_w$ is a $R \times (L - 1)$ matrix of i.i.d. complex Gaussian entries, with $O = \text{eig}(R_{RX})$ from (2.70), and the definitions of $c_k$ and $\tilde{C}_k$ in Theorem 6.2.

**Proof:** See Appendix 10.3.2.

It can be seen from (6.18) that the performance analysis for the MMSE receiver becomes more involved than for the ZF case, whereas we note that the corresponding MGF for the ZF case is
Minimum Mean Squared Error Receivers

given in Theorem 5.5. Unlike the ZF case, we do not have a determinant of a generalized quadratic form in nominator and denominator of (6.18), but there is an additional offset term, namely the identity matrix $I_{L-1}$, present. Again, it appears that there are no general formulas available in literature for calculating the expected value of this ratio of random determinants. As in case of the ZF receiver performance analysis, below we make use of certain complex (Gaussian) integrals.

We have already pointed out the importance of SER asymptotics for a concise assessment of the influence of the correlation properties of the MIMO channel on the system performance. Therefore, we explicitly state the following asymptotic MGF expression.

**Corollary 6.1:** With eigenmode transmission, the MGF $M_y(s, k)$ of the SINR on subchannel $k$ after MMSE receive processing given in Theorem 6.3 can be approximated in the high SNR region by

$$M_y(s, k) \approx \tilde{M}(s, k) = \frac{C_k}{sC_kO} \cdot \gamma^N$$

with the definitions of Theorem 6.2, diversity parameter $N = R - L + 1$, and $O = \text{eig}(R_{RX})$.

We note that as expected the diversity of the systems with ZF and MMSE receiver (given as the negative exponent of $\gamma$ in (6.19)) agrees.

### 6.2.2 Analytical MGF expressions in the high SNR regime

Based on the formulation in (6.19), in this paragraph we determine exact analytical formulas for the subchannel SNR MGF for various correlation properties of the channel in the high SNR regime. It turns out that we have to differentiate between MIMO channels with fading correlation at the transmitter and uncorrelated fading at the transmitter.

**No correlation at the transmitter ($C = I$)**

For determining an analytical expression for $M_y(s, k)$ in (6.19) we need the following lemma, which appears to be not available in literature. It can be obtained similar to Lemma 5.1 in case of the ZF receiver and we thus state without proof due to the space limitation

**Lemma 6.1:** Let $G$ be an i.i.d. $m \times n$ matrix ($m \geq n$) of complex Gaussian elements. Then

$$r_1(a, M) = E_G\left[\frac{G^HGM}{I + a \cdot G^H G}\right] = \text{tr}_n(M) \cdot n \cdot \int_0^{\infty} t^{n-1} \cdot e^{-t} \frac{1}{(1 + at)^{m+1}} dt$$

for the $m \times m$ deterministic diagonal matrix $M$ and constant $a$. Again, $\text{tr}_n(M)$ denotes the $n$th elementary symmetric function according to Appendix 11.1.6. We note that the integral in (6.20) can be written in terms of incomplete Gamma functions by applying integration by parts.
whereas the incomplete Gamma functions can be expressed in terms of the exponential integral via identity (3.73).

Using (6.20) in (6.19), we can directly find

**Theorem 6.4:** The MGF $M_\gamma(s)$ of the SINR after MMSE receive processing is for a channel without correlation at the transmitter ($C = I$) independent of the subchannel $k$ and can be approximated in the high SNR region by

\[ M_\gamma(s, k) = \frac{(L - 1) \cdot \text{tr}_{L - 1}(O)}{|O| \cdot s^R} \cdot \int_0^\infty \frac{t^{L - 2} \cdot e^{-t}}{1 + \frac{1}{s} t + R + 1} dt \cdot \gamma^N. \]  

(6.22)

**Correlation at the transmitter ($C \neq I$)**

We start again with a lemma, which can be derived similar to Lemma 5.1. For brevity we state without proof

**Lemma 6.2:** Let $G$ be an i.i.d. $m \times n$ matrix ($m \geq n$) of complex Gaussian elements. Then

\[ r_2(a, M, N) = E_G \left[ \frac{|G^H M G|}{|I + a \cdot N G^H G|} \right] = \frac{\text{tr}_n(M) \cdot n!}{\Gamma(m + 1)} \cdot \sum_{k=1}^{n} \prod_{i=1 \atop i \neq k}^{n} \frac{\tilde{n}_k}{\tilde{n}_k - \tilde{n}_i} \cdot \int_0^\infty \frac{t^{m} e^{-t}}{1 + a \tilde{n}_k t} dt \]  

(6.23)

for $m \times m$ deterministic diagonal matrix $M$, $n \times n$ deterministic diagonal matrix $N = \text{diag}(\tilde{n}_1, \ldots, \tilde{n}_n)$ with different diagonal elements, and constant $a$. Again, $\text{tr}_n(M)$ denotes the $n\text{th}$ elementary symmetric function according to Appendix 11.1.6.

We note that the integral in (6.23) can be expressed with (6.21) in terms of an incomplete Gamma function. Now making use of (6.23) in (6.19) yields

**Theorem 6.5:** The MGF $M_\gamma(s, k)$ of the SINR on subchannel $k$ after MMSE receive processing can be approximated in the high SNR region for a channel with correlation at the transmitter ($C \neq I$) by

\[ M_\gamma(s, k) = \frac{\tilde{C}_k}{c_k^R} \cdot \frac{\text{tr}_{L - 1}(O) \cdot (L - 1)!}{\Gamma(R + 1) \cdot |O|} \cdot \sum_{k=1}^{L - 1} \prod_{i=1 \atop i \neq k}^{L - 1} \frac{\tilde{c}_k}{\tilde{c}_k - \tilde{c}_i} \cdot \int_0^\infty \frac{t^{R} e^{-t}}{1 + \frac{1}{s} t + c_k^R} dt \cdot \gamma^N \]  

(6.24)

with the definition of $\tilde{C}_k = \text{diag}(\tilde{c}_1, \ldots, \tilde{c}_{L - 1})$ and $c_k$ in (6.16).
6.2.3 General analytical MGF expressions

In this paragraph we present exact expressions for the subchannel SNR MGF, whereas the derivation is again based on complex Gaussian integrals. In principle, we have to differentiate between various propagation scenarios with varying correlation properties. Due to the space limitation, we exemplarily focus on the case of receive correlation only. Other cases can equivalently be derived with the given mathematical tools. As for the asymptotical analysis, we first need a lemma for calculating the expected value of a special ratio of random determinants.

**Lemma 6.3:** Let $G$ be a $m \times n$ ($m \geq n$) matrix of i.i.d. complex Gaussian elements and

$$
\tilde{A} = \text{diag}(a_1, a_2, \ldots, a_m) \quad \tilde{B} = \text{diag}(b_1, b_2, \ldots, b_m)
$$

(6.25)

are positive definite deterministic diagonal $m \times m$ matrices, whereas the diagonal elements of $\tilde{B}$ are different. Then

$$
\hat{r}_1(\tilde{A}, \tilde{B}, m, n) = E_G \left[ \frac{I_n + G^H \tilde{A} G}{I_n + G^H \tilde{B} G} \right] = \sum_{k=0}^{n} \binom{n}{k} \sum_{\tilde{\alpha}_k} |\tilde{A}|^{\tilde{\alpha}_k} \cdot \hat{r}_{\alpha, k},
$$

(6.26)

with

$$
\hat{r}_{\alpha, k} = \frac{\Gamma(k+1)}{\Gamma(n+1)} \left[ \left( n-k \right) \int_0^{\infty} \int_0^{\infty} e^{-t} \frac{\rho^{n-1} e^{-t}}{|I + t\tilde{B}|} dt \cdot \left( \sum_{l=1}^{k} \int_0^{\infty} \frac{\rho^{n-1} e^{-t}}{1 + t \cdot [\{\tilde{B}\}_{\tilde{\alpha}_l} \{\tilde{B}\}_{\tilde{\alpha}_l}]} |I + t\tilde{B}| dt \right) \right].
$$

(6.27)

where $|\tilde{A}|_{\tilde{\alpha}_k}$ is the determinant of the matrix that results from selecting the row and column subset $\tilde{\alpha}_k$ from matrix $\tilde{A}$, $[\{\tilde{B}\}_{\tilde{\alpha}_l}]_{ll}$ is the $k \times k$ matrix that results from selecting the row and column subset $\tilde{\alpha}_k$ from matrix $\tilde{B}$, and $[\{\tilde{B}\}_{\tilde{\alpha}_l}]_{ll}$ is its $l$th diagonal element.

**Proof:** See Appendix 10.3.3.

We note that the integral in (6.27) can be calculated in terms of the exponential integral by proper decomposition in partial fractions, however, the calculations are lengthy and thus beyond the scope of this thesis. With Lemma 6.3 we get

**Theorem 6.6:** The MGF $M_{\gamma, k}(s)$ of the subchannel SINR in a correlated Rayleigh fading environment with receive correlation only, $C = I$, AWGN, and MIMO MMSE receiver is explicitly given by

$$
M_{\gamma, k}(s) = \frac{1}{|I + s\gamma O|} \cdot \hat{r}_1(\gamma O, \left( \frac{1}{\gamma} O^{-1} + s \cdot I \right)^{-1}, R, L - 1),
$$

(6.28)

where $\hat{r}_1(\tilde{A}, \tilde{B}, m, n)$ is given in Lemma 6.3 and the $R \times R$ matrix $O = \text{eig}(R_{RX})$. 

6.3 Symbol Error Rate Analysis

In this section, we make use of the subchannel SINR MGFs derived above for calculating SER expressions. Again we differentiate between high SNR asymptotics and an exact analysis.

6.3.1 High SNR asymptotics

No correlation at the transmitter \( (C = I) \)

Using relation (5.28) between the MGF and the SER, we can derive from Theorem 6.4

**Theorem 6.7:** The SER \( \bar{P}_{s,k} \) on subchannel \( k \) after MMSE receive processing can be approximated in the high SNR region for a channel without correlation at the transmitter \( (C = I) \) by

\[
\bar{P}_{s,k} = \frac{b}{cN} \cdot \frac{\Gamma(L)}{|O| \cdot 4^R} \cdot \left( \frac{2R}{R} \right) \cdot U\left(L - 1, -N + \frac{1}{2}, c\right) \cdot \gamma^{-N}
\]

(6.29)

with modulation-dependent constants \( b \) and \( c \) according to (5.17), Kummer’s \( U(a, b, z) \) function according to Appendix 11.6.6, \( N = R - L + 1 \), and \( O = \text{eig}(R_{RX}) \).

**Proof:** See Appendix 10.3.4.

Correlation at the transmitter \( (C \neq I) \)

Using relation (5.28) between the MGF and the SER, we can derive from Theorem 6.5

**Theorem 6.8:** The SER \( \bar{P}_{s,k} \) on subchannel \( k \) after MMSE receive processing can be approximated in the high SNR region for a channel with correlation at the transmitter \( (C \neq I) \) by

\[
\bar{P}_{s,k} = b \cdot \frac{\Gamma(L)}{\Gamma(R + 1)} \cdot \frac{\text{tr}_{L-1}(O)}{|O|} \cdot \chi_k \cdot \gamma^{-N}
\]

(6.30)

with the auxiliary definition

\[
\chi_k = (-1)^R \sum_{l=1}^{L-1} \left[ \frac{1}{cI} \prod_{i=1}^{l-1} \frac{\tilde{c}_i}{\tilde{c}_i - \tilde{c}_l} \cdot \frac{c_{\mathcal{C}k}}{e^c I^l} \cdot \text{erf}\left( \frac{c_{\mathcal{C}k}}{\tilde{c}_l} \right) - \sum_{j=0}^{R-1} (-1)^j \cdot \frac{c_k}{2^j} \cdot \frac{1}{L^j} \cdot \Gamma(j + 1) \right],
\]

(6.31)

the definition of \( \tilde{C}_k = \text{diag}(\tilde{c}_1, \ldots, \tilde{c}_{L-1}) \) and \( c_k \) in (6.16), \( N = R - L + 1 \), and the modulation-dependent constants \( b \) and \( c \) according to (5.17).

**Proof:** See Appendix 10.3.5.

6.3.2 Exact analysis

In this paragraph, we present expressions for the exact SER in various propagation environments. The resulting SER formulas are basically given in terms of a single integral. For systems with
uncorrelated fading and systems with fading correlation at the transmitter only, we present a derivation based on exact subchannel SINR distributions that were given in the context of optimum combining performance analysis in [38]. On the other hand, however, for the case of correlation at the receiver, we make use of the novel subchannel SINR MGF given above in Theorem 6.6. We emphasize again that this approach can in principle be used for the performance analysis of arbitrary propagation environments.

Transmit correlation

We directly start with

**Theorem 6.9:** The resulting SER on subchannel \( k \) for a MIMO system with transmit fading correlation and long-term eigenmode transmission and AWGN reads for the conditional SER approximation in (5.16) for \( M \)-QAM constellations and the definitions of Theorem 6.2

\[
P_{s,k}(\gamma) = b \cdot \left[ 1 - \frac{c}{\sqrt{\pi}} \sum_{j=1}^{R} \frac{\hat{B}_{j-1}}{\tilde{C}} (j-\frac{1}{2}) \left( \sum_{n=1}^{L-1} w_n \frac{\theta_n^2 - j}{\theta_n \Gamma(\frac{3}{2} - j, \frac{\hat{\delta}}{\theta_n})} \right) \right], \tag{6.32}
\]

where \( \hat{\delta} = c + \frac{1}{c_k \gamma} \), constants \( w_n \) defined by

\[
w_n = \prod_{j=1, j \neq n}^{L-1} \frac{\theta_n}{\theta_n - \theta_j}, \tag{6.33}
\]

and the auxiliary diagonal \((L-1) \times (L-1)\) matrix

\[
\Theta = \text{diag}(\theta_1, \ldots, \theta_{L-1}) = \tilde{C}_k / c_k, \tag{6.34}
\]

where \( \tilde{C}_k \) and \( c_k \) are defined in (6.16). Furthermore, the \( \hat{B}_i \) are the coefficients of \( z^i \) in the series expansion of (we note that there is a recursive formula for the calculation [39])

\[
\exp(z/(c_k \gamma)) \cdot \prod_{j=1}^{L-1} (1 + \theta_j z).
\]

**Proof:** See Appendix 10.3.6.

Uncorrelated fading and AWGN

With uncorrelated fading, we get just a special case of Theorem 6.9 and the SER can be obtained again by direct integration

**Theorem 6.10:** With uncorrelated fading at transmitter and receiver of the MIMO link with AWGN, the SER \( P_s(\gamma) \) is the same on all subchannels and for the conditional SER approximation in (5.16) it is given by
\[ P_s(\gamma) = b \left[ 1 - \frac{1}{\sqrt{\pi}} \sum_{j=1}^{R} \bar{\beta}_0 \Gamma \left( j - \frac{1}{2} \right) U \left( j - \frac{1}{2}, j - L + \frac{3}{2}, \delta \right) \right] \] (6.36)

with \( \delta = c + 1/\gamma \) and the constants \( \bar{\beta}_i \) are the coefficients of \( z^i \) in the series expansion of \( \exp(\gamma z) \cdot (1+z)^{L-1} \), which can be explicitly calculated

\[ \bar{\beta}_i = \sum_{k = \max(0, i-(L-1))}^{i} \frac{1}{k!} \cdot \gamma^k \cdot \binom{L-1}{i-k} \] (6.37)

\( U(a, b, z) \) is Kummer’s \( U \) function [1, Chapter 13].

**Proof:** Following the proceeding in the previous section in the derivation of Theorem 6.9, for uncorrelated fading without prefiltering (\( \Theta = I \)), we can use the integral identity [50, 3.383.5] with constants \( a, c, \) and \( N \)

\[ \int_{0}^{\infty} \frac{a - \frac{3}{2}}{(1+z)^N} \cdot e^{-cz} \, dz = \Gamma \left( a - \frac{1}{2} \right) \cdot U \left( a - \frac{1}{2}, a - N + \frac{1}{2}, c \right) \] (6.38)

and obtain directly the SER expression given in (6.36). Alternatively, (6.38) directly follows from the integral identity for Kummer’s \( U \) function [1, equation 13.2.5]. QED.

**Receive correlation**

For the case of receive correlation only, we can directly use the subchannel SINR MGF expression derived in Theorem 6.6 together with the integral representation of the SER in (5.28). In principle, an exact SER expression can be given in terms of a single integral again. However, the derivation is lengthy and tedious. Due to the space limitation, in this thesis we present only numerical results below at the end of this chapter, whereas the integral in (5.28) is calculated numerically.

**6.4 Prefiltering for MMSE Receivers**

As in case of the ZF receiver, in this section we derive statistical transmit prefilters for the case of channel distribution information at the transmitter (CDIT). To this end, again we need closed-form expressions of the performance measure, e.g. the SER, that shall be optimized via the prefilter. However, it was demonstrated above that the SER analysis of the MMSE receiver in correlated Rayleigh fading environments results in lengthy expressions that are not well suited for a further analysis. For that reason, we present a statistical prefilter design that is based on a simple bound on the average MSE, which is the natural choice of performance measure for the MMSE receiver.
Using majorization theory, we first derive the general structure of the prefilter that minimizes average MSE. It turns out again that the optimal transmission strategy is to transmit on the long-term eigenmodes (EM) of the channel with proper power allocation (PA) for the subchannels. In a second step, we derive a low-complexity power allocation scheme based on an upper bound of the average MSE. Again without loss of generality, for brevity we assume white transmit signal vectors with \( R_{ss} = E_s \cdot I \) and AWGN with \( R_{nn} = N_0 \cdot I \) in all following derivations (see Paragraph 2.3.3 for comments on equivalent systems).

### 6.4.1 Statistical prefilter structure

The resulting average (averaged over the channel statistics) mean squared error summed over all subchannels \( \varepsilon_a(F) \), which is a function of the transmit prefilter \( F \), reads from (10.74) (‘a’ stands for average)

\[
\varepsilon_a(F) = E_{H_w} [E_s \cdot \text{tr}((\gamma F^H B^H H_w A A^H H_w B F + I)^{-1})].
\]  

(6.39)

Now we arrive at the optimization problem for the prefilter (minimum average MSE)

\[
F_{opt} = \arg \min_F \varepsilon_a(F) \quad \text{s.t.} \quad \text{tr}(FF^H) = \rho.
\]  

(6.40)

The solution is given in

**Theorem 6.11**: The optimum transmit prefilter for a MIMO system with MMSE receiver in the sense that it minimizes the average MSE and forces equal SER on each subchannel and thus minimizes overall SER in a correlated Rayleigh fading environment, is given by

\[
F_{opt} = \tilde{V}_{TX} \cdot \Phi_{\varepsilon} \cdot D_L,
\]  

(6.41)

with diagonal power allocation (PA) matrix \( \Phi_{\varepsilon} \), \( L \times L \) DFT matrix \( D_L \) according to (11.1) and the EVD

\[
R_{TX} = \begin{bmatrix} V_{TX} & \tilde{V}_{TX} \end{bmatrix} \begin{bmatrix} \Lambda_{TX} & \Lambda_{TX} \end{bmatrix} \begin{bmatrix} V_{TX} & \tilde{V}_{TX} \end{bmatrix}^H,
\]  

(6.42)

where the matrix \( \tilde{\Lambda}_{TX} \) contains the \( L \) largest eigenvalues in increasing order.

**Proof**: Introducing the singular value decomposition (SVD) of the compound \( R \times L \) matrix \( BF \) with unitary matrices \( Y \) and \( Z \)

\[
BF = Y D_{BF} Z^H,
\]  

(6.43)

where \( Y \) contains the left singular vectors, \( Z \) the right singular vectors and \( D_{BF} \) the singular values, we get for the average MSE as a function of \( D_{BF} \) from (6.39)
with a $R \times L$ matrix of complex Gaussian i.i.d. entries $\tilde{H}_w$. Obviously, the average MSE is a function of the singular values in $D_{BF}$ only. Now noting that left multiplication of $F$ with unitary $Z$ does not change the objective function and the constraint in (6.40), we can introduce

$$\tilde{F} = FZ \iff \tilde{F}Z^H = F$$  \hspace{1cm} (6.45)

and solve problem (6.40) with $F$ being replaced by $\tilde{F}$. We can arbitrarily chose the unitary matrix $Z$. Here, we let $Z^H = D_L$ be a DFT matrix for forcing equal SER on all subchannels and thus minimizing the overall SER [188]. We find from (6.43) and (6.45)

$$\tilde{F}^H B^H B \tilde{F} = \tilde{F}^H R_{TX} \tilde{F} = D_{BF}^2.$$  \hspace{1cm} (6.46)

Finally, we can apply Lemma 11.5 and arrive at (6.41). QED.

At this point we emphasize the parallels of the statistical prefilter structure for ergodic capacity maximization, SER minimization in case of the ZF receiver, and finally MSE minimization in case of the MMSE receiver. All prefilters transmit on the long-term eigenmodes of the channel with proper power allocation.

6.4.2 Statistical low complexity power allocation

Theorem 6.11 describes only the general structure of the prefilter and we still have to define the power allocation (PA) matrix $\Phi_\varepsilon$. To this end, we introduce the following low-complexity power allocation scheme.

**Theorem 6.12:** A statistical power allocation scheme that reduces the average MSE can be calculated via

$$\Phi_\varepsilon = (\mu^{-1/2} \Lambda_h^{-1/2} - \Lambda_h^{-1})^{1/2} = \text{diag}(\phi_1, \ldots, \phi_L),$$  \hspace{1cm} (6.47)

where the constant $\mu$ is chosen according to the power constraint $\text{tr}(\Phi_\varepsilon^2) = \rho$, resulting in

$$\mu^{-1/2} = \frac{\text{tr}(\Lambda_h^{-1}) + \rho}{\text{tr}(\Lambda_h^{-1/2})},$$  \hspace{1cm} (6.48)

and $\Lambda_h$ are the largest $L$ eigenvalues from an eigenvalue decomposition

$$E[\gamma \cdot H^H H] = \gamma \cdot \text{tr}(R_{RX}) \cdot R_{TX} = [V_h \tilde{V}_h][\Lambda_h - \Lambda_h][V_h \tilde{V}_h]^H.$$  \hspace{1cm} (6.49)
We have to assure $\phi_l > 0$ for all $l$, which is indicated by the plus sign in (6.47). This can be guaranteed in an iterative procedure and means that in certain situations the weakest LT eigenmodes are not used for transmission by setting the corresponding PA coefficient to 0.

Derivation: From (6.39) we have to minimize the average MSE

$$\varepsilon_a(F) = E\left[E_s \cdot \text{tr}\left((\gamma F^H H^H F + I)^{-1}\right)\right]. \tag{6.50}$$

Again, an exact calculation of the expected value in (6.50) is difficult. Therefore, in this work, we consider a simple lower bound $\varepsilon_a(F)$ of the average MSE, namely

$$\varepsilon_a(F) \geq \varepsilon_a(F) = E_s \cdot \text{tr}\left((I + \gamma F^H \cdot E[H^H H] \cdot F)^{-1}\right). \tag{6.51}$$

which can be optimized more easily. In order to derive (6.51), we have applied Jensen’s inequality [24][131] for a convex function (cf. [127], where it is shown that the inverse is a convex function of a positive definite matrix argument). Now the simplified optimization problem reads

$$F_{opt} = \arg \min_F \varepsilon_a(F) \text{ s.t. } \text{tr}(FF^H) = \rho. \tag{6.52}$$

However, after calculation of the expected value in (6.51) via (11.33), the basic structure of the optimization problem with the simplified bound is equivalent to the case, where short-term channel state information is available at the transmitter [157][158][159][160][161][162]. This problem can easily be solved via constrained Lagrange optimization and the solution can be found e.g. in [160], which is given in (6.46) and (6.47). $QED$.

Finally, we note that in (6.49) obviously $\begin{bmatrix} V_h & \tilde{V}_h \end{bmatrix} = \begin{bmatrix} V_{TX} & \tilde{V}_{TX} \end{bmatrix}$ defined in (6.42), as the expected value in (6.49) is just a scaled version of the transmit correlation matrix $R_{TX}$.

### 6.4.3 Prefilter based on short-term CSI

With the availability of instantaneous channel state information (CSI) at the transmitter, a prefilter can be designed that optimally adapts to the prevailing channel state such that the MSE is minimized. The designs are well known (e.g. [160]), and we present the short-term CSI based prefilters here without proof.

**Theorem 6.13:** With short-term CSI at the transmitter, the linear prefilter that minimizes the MSE and simultaneously minimizes the overall symbol error rate can be calculated via

$$F_{opt}^{ST} = \tilde{V}_{HH} \cdot \Phi_{\varepsilon}^{ST} \cdot D_L, \tag{6.53}$$

with the $L \times L$ DFT matrix $D_L$ according to Appendix 11.1.1, the power allocation matrix

$$\Phi_{\varepsilon}^{ST} = (\mu^{-1/2} \Lambda_{HH} - \Lambda_{HH})_+^{1/2} \cdot \tag{6.54}$$
where the constant $\mu$ is chosen according to the power constraint, resulting in

$$\mu^{-1/2} = \frac{\text{tr}(\tilde{\Lambda}_{HH}) + \rho}{\text{tr}(\Lambda_{HH})^2},$$  \hspace{1cm} (6.55)

and where $\tilde{\Lambda}_{HH}$ and $\tilde{V}_{HH}$ are the matrix of the strongest $L$ eigenvalues and the corresponding matrix of eigenvectors from an eigenvalue decomposition

$$\gamma \cdot H^H H = \begin{bmatrix} V_{HH} & \tilde{V}_{HH} \end{bmatrix} \begin{bmatrix} \Lambda_{HH} & \tilde{\Lambda}_{HH} \\ \tilde{\Lambda}_{HH} & \tilde{V}_{HH} \end{bmatrix} H.$$ \hspace{1cm} (6.56)

Again, we have to assure $\phi_l > 0$ for all $l$, which is indicated by the plus sign in (6.54).

Proof: See e.g. [160].

Similar to the ZF case, there is again an interesting duality between the prefilter design based on statistical CDIT and instantaneous CSI. Obviously, we just have to replace the instantaneous channel eigenvectors and eigenvalues by their LT counterparts and by doing so, we can directly derive the prefilter structure and the power allocation matrix for the LT case.

### 6.5 Numerical Results

In this paragraph we validate the SER performance analysis and show the effectiveness of the proposed CDIT based prefilter design.

#### 6.5.1 BER performance

In Fig. 6.2 we have plotted various BER curves for a system with $T = L = 4$ and uncorrelated fading between the antenna elements, which we refer to as uncorrelated channel (UCC). There is a perfect agreement between simulation results and analytical curves according to Theorem 6.10 together with BER approximation (5.29). We emphasize the high accuracy even at lower SNR for BERs higher than $10^{-1}$. The asymptotics from Theorem 6.7 clearly indicate the increasing diversity with a higher number of receive antennas. From right to left the diversity is $N = \{1, 3, 5\}$.

Simulated BER curves, BER curves resulting from the exact analysis according to (5.28) together with Theorem 6.6, and finally the asymptotics according to Theorem 6.10 are depicted in Fig. 6.3 for a system with semi-correlated channel (SCC) and fading correlation at the receiver according to the exponential correlation matrix model (ECMM). The number of transmit antennas $T = 4$, number of receive antennas $R = 6$, and number of independent subchannels $L = 4$ is fixed. As expected, one can observe the negative impact of increasing fading correlation between the receive antenna elements from left to right, whereas the correlation coefficient increases from
The subchannel BER curves for a system with transmit correlation and long-term (LT) eigenmode (EM) transmission exhibit an interesting behavior in Fig. 6.4. At the transmitter, the correlation matrix results from the realistic correlation matrix model (RCMM) with an angular spread $\Delta_{TX} = 10^\circ$, while the receiver is assumed to be uncorrelated. One can clearly observe the different strengths of the 4 eigenmodes, whereas the 3 strongest eigenmodes approach their asymptotic

$$r_{RX} = 0.3 \text{ to } r_{RX} = 0.9$$

We note that between these two curves there is a significant shift of almost 8 dB in SNR.

$\Delta_{TX} = 10^\circ$
performance from above, while the weakest eigenmode approaches its asymptotic performance from below. This clearly contrasts the behavior of the examples in Fig. 6.2 and Fig. 6.3.

### 6.5.2 Statistical transmit prefiltering

A study of the CDIT based statistical prefilter is presented in Fig. 6.5 for a SCC with varying correlation at the transmitter according to the ECMM with $T = L = 4$ and $R = 6$ (we emphasize
that all available 4 eigenmodes of the system are used for transmission). Clearly, the prefilter becomes more effective with increased fading correlation, as expected, and the gain is about 1.5 dB in the high SNR regime for $r_{TX} = 0.9$. On the other hand, it can be observed that the gain vanishes at lower SNR. Basically, the performance of the MMSE receiver is dominated by additive noise, such that the transmit prefilter cannot improve performance. Moreover, we have depicted performance results for a system with a ST CSI based prefilter according to Theorem 6.13. Obviously, ST CSI can beneficially be exploited for achieving a significant additional performance gain in the high SNR region. However, due to the noise limitation there is again no gain in the low SNR regime.

The picture changes for a system, where the number of eigenmodes is greater than the number of independent subchannels. In Fig. 6.6 results are presented for a system with $L = 3$ and $T = R = 5$. In order to allow for a fair comparison, the system without prefilter is assumed to transmit on the outer 3 antenna elements for minimizing the correlation between them. One can observe a gain due to CDIT based statistical prefiltering even for negligible transmit correlation of $r_{TX} = 0.3$ and low SNR. Basically, with the availability of CDIT, the transmitter can now beneficially exploit the beamforming-like capabilities of the transmit antenna array that leads to a SNR gain.
7 Maximum Likelihood Receivers

In this chapter we analyze the performance of a MIMO system with maximum likelihood (ML) receiver, which has reasonable complexity for a small number of transmit antennas and low-order constellations [204][205]. We extend the results presented in [53][54][206] for uncorrelated fading and white Gaussian noise such that we take into account transmit matrix prefiltering, fading correlation at both transmitter and receiver, as well as colored Gaussian interference [104].

Using a union bound approach on the pairwise error probability (PEP), closed form symbol error rate (SER) approximations are given for arbitrary modulation alphabets. Moreover, we calculate SER asymptotics for the high signal to noise ratio (SNR) regime. Our analysis reveals that similar to the observations made in case of ergodic capacity and linear receivers, the performance penalty at high SNR due to transmit and receive correlation may be quantified independently of each other for MIMO ML receivers. Moreover, it turns out that the performance degradation due to receive correlation is a function of the determinant of the receive correlation matrix, independent of the deployed modulation scheme. This behavior is well known for maximum ratio combining performance (Corollary 5.3), pointing out that the ML receiver can achieve full diversity for each single subchannel.

On the other hand, we show that transmit correlation causes a degradation that depends on the deployed modulation alphabet. However, the modulation alphabet can be directly shaped by an adequate transmit matrix filter. We study the design of such a filter that is shown to be capable of alleviating the performance degradation due to transmit correlation for certain propagation scenarios and modulation alphabets.

7.1 Maximum Likelihood Detector Performance

In this section, we outline the derivation of an asymptotically tight union bound on the SER of a ML detector in a correlated Rayleigh fading propagation environment. To this end, we derive suitable expressions for the detection metric that allow for a concise statistical description. Using a moment generating function (MGF) approach, we calculate the probability density function (PDF) of the detection metric and derive closed form expressions for the pairwise error probability, which are used for the calculation of the union bound on the SER. We also derive high SNR asymptotics of the SER. These asymptotics are used for designing a SER minimizing statistical prefilter that is based on the transmit correlation matrix only.

7.1.1 Detection metric

With transmit matrix prefiltering and noise whitening at the receiver (Fig. 7.1), the detection metric $\eta(\nu|\mu)$ associated with transmit vector hypothesis $s_\mu$ (under the assumption that actually $s_\mu$
was transmitted) is given by (see e.g. [90][135][153][174][192] for an introduction to ML detection)

\[ \eta(\mathbf{v}|\mu) = \| \tilde{R}_{nn}^{-1/2} \mathbf{y} - \tilde{R}_{nn}^{-1/2} \mathbf{H} \mathbf{f} \|_2^2. \]  
\[ (7.1) \]

The ML receiver searches for the minimum of the detection metrics over all transmit vector hypotheses

\[ \hat{s} = \arg \min_{\tilde{s}_v} \eta(\mathbf{v}|\mu). \]  
\[ (7.2) \]

We reformulate the detection metric in (7.1) to bring it into a form that appeared already in [206] so as to simplify the following statistical analysis. To this end, we introduce the singular value decompositions (SVD) at the transmitter side (see 2.42)

\[ B = U_{TX} \begin{bmatrix} \tilde{\Lambda}_{TX}^{1/2} & \tilde{V}_{TX}^H \\ \Lambda_{TX}^{1/2} & V_{TX}^H \end{bmatrix}, \]  
\[ (7.3) \]

where \( \tilde{\Lambda}_{TX}^{1/2} \) contains the \( L \) largest singular values, and at the receiver side (see 2.70)

\[ \tilde{R}_{nn}^{-1/2} A = U_o O^{1/2} V_o^H. \]  
\[ (7.4) \]

Furthermore, without loss of generality, the \( R \times L \) matrix prefilter \( \mathbf{F} \) can be decomposed as

\[ \mathbf{F} = \tilde{V}_{TX} \cdot \mathbf{L} \]  
\[ (7.5) \]

with general \( L \times L \) matrix \( \mathbf{L} \). We get from (7.1) after left-multiplication with \( U_o \) (this leaves the norm unaffected, as \( U_o \) is a unitary rotation matrix)

\[ \eta(\mathbf{v}|\mu) \equiv \| \tilde{\mathbf{y}} - O^{1/2} \tilde{\mathbf{H}} \tilde{\mathbf{A}}_{TX}^{1/2} \mathbf{L} s_v \|_2^2, \]  
\[ (7.6) \]

which with the help of the decomposition of the \( R \times L \) random i.i.d. complex Gaussian matrix \( \tilde{\mathbf{H}}_w \) in \( 1 \times L \) row vectors \( \mathbf{g}_r^H \) similar to the decomposition given in (2.49) can be written as

\[ \eta(\mathbf{v}|\mu) \equiv \sum_{r=1}^{R} \| \tilde{\mathbf{y}}_r - O^{1/2} \mathbf{g}_r^H \tilde{\mathbf{A}}_{TX}^{1/2} \mathbf{L} s_v \|_2^2, \]  
\[ (7.7) \]

where we have introduced the auxiliary vector
\[
\tilde{y} = U_o^H \tilde{R}^{-1/2} y = (\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_R)^T
\]
and the singular values \( o_i^{1/2} \) on the diagonal of the matrix \( O^{1/2} \). We emphasize again that we consider full rank matrices \( R_{RX} \) and \( R_{TX} \) in this work. Furthermore, we have used the property of i.i.d. complex Gaussian matrices that the distribution does not change by left- or right-multiplications with a unitary matrix.

### 7.1.2 Union bound on symbol error probability

For the analysis of the symbol and bit error probability we introduce the following definitions. Let \( M \) be the constellation size (e.g. \( M = 4 \) for QPSK), whereas we assume in this work that the modulation on each of the \( L \) parallel subchannels is the same for notational simplicity (however, the generalization for different modulation schemes is straightforward). Let \( s_q \) be a symbol out of the signal constellation \( q \in \{0, 1, \ldots, M - 1\} \), \( \{s\} \) the set of all \( M^L \) possible transmit vectors, \( \{s_i\} \) the set of transmit vectors with \( s_q \) as their \( k \)th element ( \( k \) is the subchannel index), \( \{s_j\} \) the set of transmit symbol vectors with their \( k \)th element differing from \( s_q \), \( \eta_k(i|\{i\}, q) \) the metric associated with transmit symbol vector hypothesis \( s_i \) (assuming that in fact \( \{s_i\} \) was transmitted), \( \eta_k(j|\{i\}, q) \) the metric associated with hypothesis \( s_j \) (again assuming that actually \( s_i \) was transmitted). Furthermore, we denote the difference between the metrics associated with the correct and incorrect hypotheses (i.e. correct and incorrect decision on subchannel \( k \)) as

\[
D_k(i, j, q) = \eta_k(j|\{i\}, q) - \eta_k(i|\{i\}, q).
\]

A pairwise error with incorrect symbol decision on subchannel \( k \) occurs if

\[
P_k(i, j, q) = Pr(D_k(i, j, q) < 0),
\]

i.e. the probability, that actually \( s_i \) was transmitted with \( s_q \) as its \( k \)th element and the receiver erroneously decides in favor of the incorrect hypothesis \( s_j \). It can be calculated from (with shorthand notation \( D_k \) for \( D_k(i, j, q) \))

\[
P_k(i, j, q) = \int_{-\infty}^{0} p(D_k) dD_k,
\]

i.e. we have to find the PDF of \( D_k(i, j, q) \), namely \( p(D_k) \), and carry out the integration. The union bound of the SER on subchannel \( k \) in the general case with transmit prefiltering reads

\[
P_{s, k}(\gamma) \leq \frac{1}{M^L} \sum_q \sum_i \sum_j P_k(i, j, q).
\]
It should be mentioned that simplifications are possible by exploiting possible symmetries of the constellation diagram (cf. [206]) such that the number of sum terms in (7.13) can be reduced. Furthermore, an approximate calculation of the bit error probability \( P_{b,k}(\gamma) \) is possible via

\[
P_{b,k}(\gamma) \approx \frac{1}{M} P_{s,k}(\gamma),
\]

by assuming that each symbol error causes just exactly one bit error, which is true for Gray encoding in the high SNR region.

### 7.1.3 Calculation of pairwise error probability

For determining the PEP in (7.12), we have to find the PDF of \( D_k(i, j, q) \). To this end, we first determine its MGF and then use an inverse Laplace transform to find the PDF.

**MGF of the detection metric**

**Theorem 7.1**: The MGF of the metric difference \( D_k(i, j, q) \), namely \( M_{D_k}(s) \), is given by

\[
M_{D_k}(s) = \prod_{r=1}^{R} \frac{1}{(1 + \lambda_{r,1}s)(1 + \lambda_{r,2}s)}
\]

with the auxiliary variables

\[
\lambda_{r,1/2} = E_s \cdot \frac{\tilde{d}_r \pm \sqrt{\tilde{d}_r^2 + 4\gamma}}{2},
\]

the modified transmit vector differences

\[
\tilde{d}_r = o_r \cdot \| \Lambda_{\tilde{X}_r} L(\tilde{s}_i - \tilde{s}_j) \|^2,
\]

and the normalized transmit vectors (\( E_s \) is the mean symbol energy per subchannel)

\[
s_i = \sqrt{E_s} \cdot \tilde{s}_i, \quad s_j = \sqrt{E_s} \cdot \tilde{s}_j.
\]

**Proof**: See Appendix 10.4.1.

It remains now to determine the PDF of \( D_k(i, j, q) \) and perform the integration in (7.12) for finding the PEP.

**Expressions for the pairwise error probability**

**Case I**: Same singular values \( o_r = 1 \) for all \( r \)
This case occurs for spatially white Gaussian noise and vanishing receive correlation. It has partially been studied in [53][54][206] for the case of vanishing fading correlation at the transmitter and we summarize the results for completeness.

**Theorem 7.2:** The pairwise error probability on subchannel $k$ for a MIMO ML receiver without fading correlation at the receive antenna array is given by

$$ P_k(i, j, q) = \frac{1}{(1+r_{ij})^{2R-1}} \cdot \sum_{r=0}^{R-1} \binom{2R-1}{r} r_{ij}^r. $$

(7.19)

with the auxiliary definition

$$ r_{ij} = \frac{\tilde{d}_I \cdot \gamma}{2} + \sqrt{\left(\frac{\tilde{d}_I \cdot \gamma}{2}\right)^2 + \tilde{d}_I \cdot \gamma + 1} $$

(7.20)

and together with the definition of the normalized transmit vectors in (7.18)

$$ \tilde{d}_I = \|\Lambda_{\tilde{r}X}^\frac{1}{2} L (\tilde{s}_i - \tilde{s}_j)\|^2. $$

(7.21)

**Proof:** See [206], whereas it is a simple exercise to take into account transmit correlation and the prefilter via the term $\Lambda_{\tilde{r}X}^\frac{1}{2} L$ in a modified transmit symbol vector (cf. (7.21)).

**Case II:** Different singular values $\sigma_r$

This case occurs for colored Gaussian noise and/or non-vanishing receive fading correlation (cf. to the definition in (7.4)).

**Theorem 7.3:** In the presence of correlation at the receive antenna array, the PEP with an error on subchannel $k$ is given by

$$ P_k(i, j, q) = 1 - \sum_{r=1}^{R} \prod_{\substack{i=1 \atop i \neq r}}^{2R} \lambda_r \frac{\lambda_r}{\lambda_r - \lambda_i}, $$

(7.22)

whereas we define the $2R$-dimensional vector

$$ \begin{bmatrix} \lambda_1 & \lambda_2 & \ldots & \lambda_{2R} \end{bmatrix}^T = \begin{bmatrix} \lambda_{1,1} & \lambda_{2,1} & \ldots & \lambda_{R,1} & \lambda_{1,2} & \lambda_{2,2} & \ldots & \lambda_{R,2} \end{bmatrix}^T $$

(7.23)

with the eigenvalues $\lambda_{r,\frac{1}{2}}$ defined in (7.16), such that obviously the positive eigenvalues $\lambda_{r,1}$ are arranged first in (7.23).

**Proof:** See Appendix 10.4.2.
7.1.4 Pairwise error probability asymptotics

We have demonstrated in earlier chapters that the asymptotics in the high SNR regime prove to be a powerful means to characterize the influence of fading correlation on various performance measures. In this section we show that the effects of fading correlation at transmitter and receiver can again be separated for high SNR in case of the ML receiver. Moreover, below we will use the asymptotics to derive SNR independent statistical prefilters that minimize SER. First, we state a theorem on the high SNR asymptotics of the PEP.

**Theorem 7.4:** At high SNR, the pairwise error probability of a ML receiver in a Rayleigh fading environment with transmit as well as receive correlation and transmit prefilter (7.5) is given by

\[
\bar{P}_k(i, j, q) = \frac{1}{|O|} \left| \tilde{\Lambda}_{TX}^{1/2} L(\tilde{s}_i - \tilde{s}_j) \right|^{2R} \cdot \left( \frac{2R - 1}{R - 1} \right) \cdot \gamma^{-R}
\]

(7.24)

with the transmit vector normalization

\[
s_i = \sqrt{\overline{E}_s} \cdot \tilde{s}_i \quad s_j = \sqrt{\overline{E}_s} \cdot \tilde{s}_j.
\]

(7.25)

**Proof:** See Appendix 10.4.3.

Equation (7.24) explicitly shows a diversity level of \(R\) (independent of fading correlation and the number of input subchannels \(L\)) that can be achieved by ML detection. It is also obvious that the effects of fading correlation at transmitter and receiver can be separated completely. Furthermore, the SNR penalty induced by fading correlation at the receive array is a function of the determinant of the receive correlation matrix (see below). Note that (7.24) reduces to the results presented in [53][54][206] in case of vanishing receive correlation and white Gaussian noise.

7.1.5 Impact of fading correlation at high SNR

The SNR penalty due to fading correlation at the receive antenna array in dB compared to a system with vanishing receive correlation directly follows from Theorem 7.4 and is given by

\[
\Delta \gamma_{dB, RX} = \frac{10}{R} \cdot \log_{10} \frac{1}{|O|}.
\]

(7.26)

Moreover, the degradation is independent of the modulation scheme. As was noted already above, with the underlying channel model, the effects of receive and transmit correlation can be separated. To this end, note that the SNR penalty in dB due to transmit correlation (again compared to an uncorrelated system) is (from (7.13) and (7.24))

\[
\Delta \gamma_{dB, TX} = 10 \cdot \log_{10} \left( \frac{\sum \sum \sum \mid \tilde{\Lambda}_{TX}^{1/2} L(s_i - s_j) \mid^{-2R} \cdot \gamma^{-R}}{\sum \sum \sum \mid (s_i - s_j) \mid^{-2R} \cdot \gamma^{-R}} \right)^{1/R}.
\]

(7.27)
Now the SNR penalty is a direct (yet complicated) function of the deployed modulation scheme and the transmit matrix filter.

### 7.1.6 Statistical transmit filter design

The SNR penalty in (7.27) is obviously a function of the transmit prefilter \( F \) in (7.5) via the general \( L \times L \) matrix \( L \). It is thus possible to improve the overall performance of the MIMO system by adequate transmit filtering. We can minimize the performance degradation via

\[
L_{opt} = \arg \min_L \sum_{q} \sum_{i} \sum_{j} \left\| \tilde{A}_{TX}^{1/2} L(s_i - s_j) \right\|^{-2R} \quad \text{s.t. } \text{tr}(LL^H) = \rho, \tag{7.28}
\]

where \( \rho \) is the transmit power constraint. Numerical optimization algorithms can be used to find the optimum filter of the constrained optimization problem in (7.28). However, unfortunately the problem in (7.28) is not convex and thus of considerable complexity. Efficient optimization procedures would have to be found for a practical implementation, especially in case of a large number of independent subchannels. In this thesis, we resort to simple standard optimization algorithms [19][124] and the effectiveness of the proposed prefilter design is demonstrated below in Monte-Carlo simulations.

### 7.2 Numerical Results

#### 7.2.1 Validation of the SER bound

Symbol error rate curves are depicted in Fig. 7.2 for a system with semi-correlated channel (SCC) according to the exponential correlation matrix model (ECMM) with \( T = L = R = 4 \). The correlation coefficient at the transmitter varies from \( r_{TX} = 0.3 \) to \( r_{TX} = 0.9 \), while the receiver is assumed to be uncorrelated. As expected, one can observe that the full diversity of \( N = R = 4 \) can be achieved by the ML receiver, independent of the number of subchannels that are transmitted. Obviously, the SER union bound is tight at high SNR, while it significantly diverges from the simulated curves at SERs higher than \( 10^{-2} \).

Similar results are presented for a system with uncorrelated channel (UCC) and \( T = L = 2 \) in Fig. 7.3. Now the number of receive antennas \( R \) is varied and the increasing diversity of the system can be seen from the high SNR asymptotics. Moreover, the performance gain resulting from increasing the number of antenna elements \( R \) from 2 to 4 is large, while it decreases rapidly for higher number of antennas in the SNR range of practical interest.

#### 7.2.2 Linear prefiltering based on the union bound

Simulation results of the BER performance of a system with prefiltering are depicted in Fig. 7.4 for a SCC with transmit correlation according to the realistic correlation matrix model (RCMM).
A standard gradient-based numerical optimization algorithm has been used to design transmit pre-filters for a MIMO system with $T = 2$ transmit antennas, $R = 2$ receive antennas, and QPSK modulation via (7.28). It is difficult to differentiate between the results for the uncorrelated reference case and the case with transmit prefiltering and transmit correlation corresponding to $\Delta_{TX} = 10^\circ$ AS. Obviously, for this particular scenario we can completely eliminate the SNR penalty due to transmit correlation by statistical long-term based transmit prefiltering and even slightly outperform the uncorrelated system at lower SNR. This is an observation that has already
Numerical Results

been noted for ergodic MIMO capacity. For comparison, we have also plotted the curve with transmit correlation but without transmit prefiltering that is subject to an asymptotic SNR penalty according to (7.27) of 2.5 dB.

Further insight may be obtained by looking at the virtual signal constellation Fig. 7.5 that results from prefiltering and TX channel correlation \((\Lambda_{TX} L s_j)\) for all \(j\). We have plotted the first element of the constellation vector \(\Lambda_{TX} L s_j\) for all \(j\) (a superposition of two scaled QPSK constellations) and an AS of \(\Delta_{TX} = 10^\circ\) at the transmitter including transmit prefiltering (left) and without prefiltering (right). It is clear that the optimum transmit filter arranges the virtual signal constellation more regularly, as otherwise the symbol pair with the minimum distance would dominate the performance penalty at high SNR. Thus the prefilter automatically leads to an increase of the minimum constellation distance, a design criteria that was used in [101].
Interestingly, even for strong fading correlation with an angular spread of $\Delta_{TX} = 2^\circ$ at the transmitter, statistical transmit prefiltering is very effective. This can be seen in Fig. 7.6, where the asymptotic performance penalty of 4.6 dB in SNR according to (7.27) can be eliminated totally.

**Fig. 7.6** BER with prefiltering, RCMM, $T=L=R=2$, $\Delta_{TX} = 2^\circ$, RX unc.
8 Conclusion

Based on multivariate statistical theory, in this thesis we have presented a performance analysis of MIMO wireless systems in correlated Rayleigh flat fading environments with arbitrary fading correlation at both the transmitter as well as the receiver. We could quantitatively assess the influence of fading correlation on system performance, in particular the ergodic capacity and the symbol error rate for various receiver types.

Specifically, by using advanced tools from multivariate statistics, namely hypergeometric functions of matrix arguments, a closed form formula of the moment generating function (MGF) of MIMO mutual information was derived in Chapter 3. For the first time, an eigenvalue free approach was used that permits a unifying analysis for arbitrary propagation environments. Using the MGF, we could calculate the exact ergodic capacity in terms of simple functions for finite array sizes and find high and low SNR asymptotics that allow for a concise characterization of the influence of fading correlation. Moreover, a tight bound on ergodic capacity was established in Chapter 4, which was used to rigorously prove the intuitive perception that increasing fading correlation reduces ergodic capacity. In addition, waterfilling schemes based on the exact ergodic capacity analysis and the novel bound were presented, which are based on statistical channel state information (CSI) only. It was demonstrated that a considerable capacity gain can be achieved by these latter schemes.

With the aim of quantitatively assessing the performance in the presence of correlated Rayleigh fading and designing symbol error rate (SER) minimizing linear transmit prefilters that use long-term CSI only, we have analyzed the SER of MIMO systems deploying linear zero-forcing (ZF) receivers in Chapter 5. It turned out that an exact SER analysis in general is possible for arbitrary correlation properties of the channel by applying a novel mathematical approach based on complex Gaussian integrals. Specifically, we have shown via a MGF approach that SER calculation in arbitrarily correlated MIMO channels boils down to determining the expected value of a ratio of determinants of generalized random matrix quadratic forms. To the authors’ best knowledge, these expected values are not available in statistical literature and have been calculated in this thesis for the first time. We could calculate exact SER expressions for arbitrary system parameters and established a very tight low-complexity bound, which is asymptotically exact in the high SNR regime. With the availability of exact asymptotic SER formulas, we could establish a simple way of characterizing the impact of correlation on ZF receivers. Furthermore, we have designed a low-complexity statistical transmit prefilter based on these asymptotics, which was shown to effectively reduce SER in correlated MIMO channels and allows for an application in simple terminals. For the design of this prefilter, majorization theory proved to be a powerful mathematical tool. Interestingly, the proposed statistical prefilter design exhibits the same structure as its counterpart that makes use of instantaneous channel state information.

A similar analysis as in case of the ZF receiver was presented for minimum mean squared error (MMSE) MIMO receivers in Chapter 6. However, it was demonstrated that the complexity of the analysis increases considerably due to the noise-dependent structure of the MMSE receive filter.
Via a new MGF approach, we could show that the key for calculating exact SER expressions in case of the MMSE receiver is again the determination of certain expected values of ratios of determinants of random matrices. As in case of the ZF receiver, these expected values could be calculated for the first time using complex Gaussian integrals. Based on the novel eigenvalue free approach, exact unifying SER expressions for arbitrary system parameters could be given in terms of a single integral. Furthermore, a simple and concise SER performance assessment was possible via asymptotical SER formulas. Due to considerable complexity of the exact SER expressions, we have proposed a simple statistical transmit prefilter design that is based on a minimum average mean squared error (MSE) criterion. Again, we could observe a duality between short-term and long-term CSI based filters.

Chapter 7 contains a thorough analysis of the SER of MIMO maximum likelihood (ML) receivers. To this end, exact formulas for the pairwise error probability (PEP) were derived based on a MGF approach for arbitrary correlated Rayleigh fading MIMO channels. Based on that, we gave an asymptotically tight union bound on the SER. Via a series expansion, we could derive high SNR asymptotics of the SER, which explicitly demonstrate the dependence on the correlation properties of the channel. A transmit prefilter design that effectively minimizes the SER was developed based on these SER asymptotics.


9 Appendix - MIMO Capacity

**Note:** In all derivations we use the abbreviation \( \tilde{s} = \frac{s}{\ln 2} \) for a concise notation.

9.1 Moment Generating Functions

9.1.1 Derivation of Corollary 3.1

We now have to calculate similar to (3.33)

\[
M_{\Sigma}(s) = \mathbb{E}_G \left[ \tilde{F}_0(\mu) \left( -\frac{s}{\ln 2}; \mathcal{Q} \right) \right] = \mathbb{E}_W \left[ \tilde{F}_0(\mu) \left( -\frac{s}{\ln 2}; -\mathcal{W} \right) \right].
\]

(9.1)

where for clarity with \( \Omega = \mathbf{I} \) the matrix quadratic form \( \mathcal{Q} \) is replaced by a complex Wishart distributed \( \nu \times \mu \) matrix \( \mathcal{W} \). The PDF of \( \mathcal{W} \) for the given problem reads (see (3.25))

\[
p(\mathcal{W}) = \frac{1}{\Gamma(\nu) \cdot |\gamma\Sigma|^{\nu}} \cdot e^{-tr(\gamma\Sigma^{-1}\mathcal{W} \cdot |\mathcal{W}|^{\nu} - \mu \cdot (d\mathcal{W})}.
\]

(9.2)

Now, analogously to (3.34), we want to calculate

\[
E_W [ \tilde{C}_\kappa(M\mathcal{W}) ] = \int_{\mathcal{W}} \tilde{C}_\kappa(M\mathcal{W}) \cdot p(\mathcal{W}) d\mathcal{W},
\]

(9.3)

which reads with (9.2)

\[
E_W [ \tilde{C}_\kappa(M\mathcal{W}) ] = \frac{1}{\Gamma(\nu) \cdot |\gamma\Sigma|^{\nu}} \cdot \int_{\mathcal{W}} \tilde{C}_\kappa(M\mathcal{W}) \cdot e^{-tr(\gamma\Sigma^{-1}\mathcal{W} \cdot |\mathcal{W}|^{\nu} - \mu \cdot (d\mathcal{W})}.
\]

(9.4)

For solving (9.4) we can make use of [94, equation (53)] with \( m \times m \) matrices \( \tilde{A}, \tilde{B} \), and scalar \( a \)

\[
\int_{\tilde{A}} \tilde{C}_\kappa(\tilde{A}\tilde{B}) \cdot e^{-tr(\tilde{A}) \cdot |\tilde{A}|^{\nu-m} \cdot \tilde{A}} = \tilde{\Gamma}_m(a, \kappa) \cdot \tilde{C}_\kappa(\tilde{B}) = [a]_\kappa \cdot \tilde{\Gamma}_m(a) \cdot \tilde{C}_\kappa(\tilde{B}),
\]

(9.5)

where we directly have deployed [94, equation (50)]

\[
[a]_\kappa = \frac{\tilde{\Gamma}_m(a, \kappa)}{\tilde{\Gamma}_m(a)}.
\]

(9.6)

First, we make the transformation \( \mathcal{W} \rightarrow \gamma\Sigma \cdot \tilde{A} \) in (9.4) with Jacobian [29][59][144] \( J(\mathcal{W} \rightarrow \gamma\Sigma \cdot \tilde{A}) = |\gamma\Sigma|^{\mu} \) and get

\[
E_W [ \tilde{C}_\kappa(M\mathcal{W}) ] = \frac{1}{\Gamma(\nu) \cdot |\gamma\Sigma|^{\nu}} \cdot \int_{\tilde{A}} \tilde{C}_\kappa(M\gamma\Sigma \tilde{A}) \cdot e^{-tr(\tilde{A}) \cdot |\gamma\Sigma\tilde{A}|^{\nu} - \mu \cdot |\gamma\Sigma|^{\mu} \cdot d\tilde{A}}.
\]

(9.7)
resulting in
\[ E_W[\tilde{C}_\kappa(M\Sigma)] = \frac{1}{\Gamma_\mu(\nu)} \cdot \int \tilde{C}_\kappa(M\gamma\Sigma\tilde{\alpha}) \cdot e^{-\pi(\tilde{\alpha})} \cdot |\tilde{\alpha}|^{\nu-\mu} \cdot d\tilde{\alpha}. \] (9.8)

Direct application of (9.5) to (9.8) results in the reproductive property of the zonal polynomial with respect to the Wishart distribution
\[ E_W[\tilde{C}_\kappa(M\Sigma)] = [\nu]_\kappa \cdot \tilde{C}_\kappa(M\gamma\Sigma). \] (9.9)

Then following the derivation in (3.35) and (3.36) we get (9.1). QED.

9.1.2 Derivation of Corollary 3.2

Using L’Hospital’s rule, we have to differentiate the determinant in the nominator and the Vandermonde determinant in the denominator in (3.45) times with respect to \( \varepsilon_k \) (this shall be indicated by the symbolic notation \( \frac{\partial}{\partial \varepsilon} \)) and then set \( \varepsilon = 0 \). To this end, note that
\[ \frac{\partial^k}{\partial w^k} \left( \sum_{a_1} \ldots \sum_{a_k} \left( -c^{a_1} \right) \right) = \left[ a_1 \right] \left[ a_2 \right] \ldots \left[ a_k \right] (-c)^k \cdot \sum_{a_1} \ldots \sum_{a_k} \left( -c^{a_1} \right) \] (9.10)

and after the differentiation and limiting process above the Vandermonde determinant reads
\[ \left. \frac{\partial}{\partial \varepsilon} \alpha_\nu(\Sigma(\varepsilon)) \right|_{\varepsilon = 0} = \alpha_\mu(\Sigma) \cdot |\Sigma|^{\nu-\mu} \cdot \Gamma_{\nu-\mu}(\nu-\mu) \cdot (-1)^{\frac{(\nu-\mu)(\nu-\mu-1)}{2}}. \] (9.11)

Using (9.10) and (9.11) in (3.45) we get (3.47) with the help of (11.87). QED.

9.1.3 Derivation of Corollary 3.3

We find after iteratively applying Kummer U function recurrence relations (11.90) and (11.91) for the entries of the matrix \( \Psi_{\Sigma,\Omega,\Omega} (s) \) in (3.49)
\[ \frac{1}{\gamma \sigma_i \omega_j} \cdot U\left( 1, \tilde{s} + v + 1, \frac{1}{\gamma \sigma_i \omega_j} \right) \]
\[ = \sum_{k=0}^{v-2} (-1)^k \cdot (\gamma \sigma_i \omega_j)^k \cdot [-\tilde{s} - v + 1] + \]
\[ (-1)^{v-1} \cdot (\gamma \sigma_i \omega_j)^v \cdot [-\tilde{s} - v + 1]_{v-1} \cdot U\left( 1, \tilde{s} + 2, \frac{1}{\gamma \sigma_i \omega_j} \right) \] (9.12)

and the remaining Kummer U function has from (11.88) the integral representation
Using (9.13) and (9.12) in (3.47), then subtracting properly scaled multiples of the rows of $\Psi_{\Sigma \Omega, 2}$ from the rows of $\Psi_{\Sigma \Omega, 1}$ (this does not alter the determinant $|\Psi_{\Sigma \Omega}|$), and finally factoring out $\sigma_i^{-\mu}$ from the resulting rows of $\Psi_{\Sigma \Omega, 1}$ yields (3.51) with (3.52) and (3.53). QED.

9.1.4 Derivation of Theorem 3.3

We have to calculate the limit $\Omega \to I$ in (3.47). To this end, we can apply L’Hospital’s rule, i.e. we differentiate nominator and denominator $(k - 1)$ times with respect to $\omega_k$ ($k \in \{1, \ldots, v\}$), which we symbolically describe by $\frac{\partial}{\partial \omega}$, and then set $\omega = 1$. We find ($i$ runs from 1 to $\mu$ and $j$ from 1 to $v$)

$$\left. \frac{\partial \Psi_{\Sigma \Omega, 1}}{\partial \omega} \right|_{\omega = 1} = \left[ (-\bar{s} - v + 1)_{j - 1} \cdot [1]_{j - 1} \cdot (-\gamma \sigma_i)^{j - 1} \cdot {}_2F_0(-\bar{s} - v + j, j; -\gamma \sigma_i) \right].$$

(9.14)

which can be rewritten in terms of the Kummer $U$ function via (11.87)

$$\left. \frac{\partial \Psi_{\Sigma \Omega, 1}}{\partial \omega} \right|_{\omega = 1} = \left[ (-\bar{s} - v + 1)_{j - 1} \cdot [1]_{j - 1} \cdot (-1)^{j - 1} \cdot \frac{1}{\gamma \sigma_i} \cdot U\left( j, \bar{s} + v + 1, 1 \right) \right].$$

(9.15)

On the other hand we get ($i'$ runs from 1 to $v - \mu$ and $j'$ from 1 to $v$)

$$\left. \frac{\partial \Psi_{\Sigma \Omega, 2}}{\partial \omega} \right|_{\omega = 1} = \left[ (-1)^{i' - 1} \cdot \gamma^{i' - 1} \cdot (-\bar{s} - v + 1)_{i' - 1} \cdot (-1)^{i' - 1} \cdot [1 - i']_{i' - 1} \right].$$

(9.16)

This is a lower triangular matrix (note that $[-a]_b = 0$ for $b > a > 0$). The resulting expression $|\Psi_{\Sigma \Omega}|$ according to (3.48) can further be simplified by developing the determinant along the diagonal of the triangular matrix and a $\mu \times \mu$ matrix remains. Without going too much into the details, we can row and column wise extract common factors from (9.15), which cancel with $\psi^{(v)}(-\bar{s}, v)$ in the denominator of (3.47), such that only the Kummer $U$ function remains in the elements of the determinant. Then using the recurrence relation (11.92), we can reduce the determinant entries of $|\Psi_{\Sigma \Omega}|$ to the following form ($i, j$ run from 1 to $v$)

$$U\left( j + v - \mu, \bar{s} + v - \mu + j + 1, \frac{1}{\gamma \sigma_i} \right) = \frac{1}{\Gamma(v - \mu + j)} \int_0^\infty \frac{1}{(v - \mu + j)} e^{-\gamma \sigma_i t} \cdot t^{v - \mu + j - 1} \cdot (1 + t)^{\bar{s}} dt$$

(9.17)

by iteratively subtracting scaled versions of the $(j + 1)$ th column from the $j$ th column. Then note that the integral in (9.17) can be rewritten as
Finally, we can find by straightforward application of L’Hospital’s rule
\[
\left[ \frac{\partial}{\partial \omega} \alpha_{v}(\Omega) \right]_{\omega = 1} = (-1)^{\frac{v(v-1)}{2}} \cdot \Gamma(v) .
\]  
(9.19)

Plugging all results in (3.47) and simplifying yields Theorem 3.3. QED.

### 9.1.5 Derivation of Theorem 3.4

Again we calculate a limit \( \Sigma \to I \) of (3.47) via L’Hospital’s rule, yielding in (3.49)
\[
\left[ \frac{\partial^\Psi}{\partial \sigma} \Omega^1 \right]_{\sigma = 1} = \left[ (-1)^{i-1} \cdot \left[ -\tilde{s} - v + 1 \right]_{i-1} \cdot \Gamma(i) \cdot \frac{1}{\gamma \omega_j} \cdot U \left( i, \tilde{s} + v + 1, \frac{1}{\gamma \omega_j} \right) \right].
\]  
(9.20)

Then consider \( [\Psi_{\Sigma} \Omega] \) together with the result in (9.20). We can obviously factor out \( (-1)^{i-1} \cdot [ -\tilde{s} - v + 1 ]_{i-1} \cdot \Gamma(i) \) in the first \( \mu \) rows. With the recurrence relations of the Kummer \( U \) function (11.90) and (11.91) we can iteratively subtract a scaled version of ith row from \( (i+1) \)th row (this does not change the determinant) and arrive at a determinant, where the first \( \mu \) rows have elements
\[
(-1)^{v-\mu} \cdot \left[ -\tilde{s} - v + 1 \right]_{v-\mu} \cdot (\gamma \omega_j)^{v-\mu-1} \cdot U \left( i, \tilde{s} + v + 1, \frac{1}{\gamma \omega_j} \right).
\]  
(9.21)

Now iteratively add a scaled version of \((i+1)\)th row to the ith row and apply (11.92), such that we finally find the elements of the first \( \mu \) rows
\[
(-1)^{v-\mu} \cdot \left[ -\tilde{s} - v + 1 \right]_{v-\mu} \cdot (\gamma \omega_j)^{v-\mu-1} \cdot U \left( i, \tilde{s} + i + 1, \frac{1}{\gamma \omega_j} \right).
\]  
(9.22)

Then note that the Kummer \( U \) function can explicitly be written as an integral
\[
U \left( i, \tilde{s} + i + 1, \frac{1}{\gamma \omega_j} \right) = \frac{1}{\Gamma(i)} \cdot \int_{0}^{\infty} e^{-\gamma \omega_j t} \cdot t^{i-1} \cdot (1 + t)^{\tilde{s}} dt.
\]  
(9.23)

Plugging the results in (3.47) and simplifying yields (3.58)-(3.60). QED.

### 9.1.6 Derivation of Theorem 3.5

We can start the derivation with the results of (3.55) and (3.56) for a system with \( \Sigma \neq I \) and \( \Omega = I \), where we have to find the limit \( \Sigma \to I \), which can be obtained by L’Hospital’s rule along the same lines as in the derivation of Theorem 3.4 and Theorem 3.5. By Lebesgue’s dominated
convergence theorem we can exchange the sequence of integration and differentiation in the process of differentiating \[|\Psi_{\Sigma}(s)|\], where we can use the following derivative

\[
\frac{\partial^k}{\partial e^k} e^{-t/c} = \left( \sum_{l=1}^{k-1} \frac{a_l}{c^{l+k}} \cdot t^l + \frac{e^k}{c^{2k}} \right) \cdot e^{-t/c} \tag{9.24}
\]

with constants \(a_l\). Adding a properly scaled multiple of the 1st, 2nd, …, \((i-1)\)th row to the \(i\)th row in the resulting \[|\Psi_{\Sigma}(s)|\] and using (9.19) for differentiating the Vandermonde determinant in the denominator yields finally (3.62) and (3.63). QED.

### 9.2 Ergodic Capacity

#### 9.2.1 Derivation of Theorem 3.6

By the quotient rule of differentiation we get from (3.65) with (3.51)

\[
C_{\text{erg}, \Sigma} = \zeta \cdot \left. \frac{\partial}{\partial s} \left[ \Psi_{\Sigma}(s) \big| \Psi_{\Sigma}^{(v)}(\hat{s}, v) - \Psi_{\Sigma}(s) \big| \cdot \frac{\partial}{\partial s} \psi_2^{(v)}(\hat{s}, v) \right] \right|_{s=0} \tag{9.25}
\]

with

\[
\zeta = \frac{(-1)^{\frac{v}{2}} \cdot (v-1)^{-\nu}}{\Gamma(v)} \cdot \frac{1}{\alpha_{\mu}(\Sigma) \cdot \alpha_v(\Omega)} \cdot \frac{(1-v) \cdot (\nu-\mu) \cdot (\nu-\mu-1)}{v \cdot (v-1)} \cdot \frac{(v \cdot (v-1))}{2} \cdot \frac{(v \cdot (v-1))}{2} \cdot \Gamma(v) \tag{9.26}
\]

Omitting the details, we find

\[
\psi_2^{(v)}(\hat{s}, v) \big|_{s=0} = (-1)^{\frac{v}{2}} \cdot (v-1)^{-\nu} \cdot \prod_{k=1}^{v-1} k^v \tag{9.27}
\]

and

\[
\frac{\partial}{\partial s} \psi_2^{(v)}(\hat{s}, v) \big|_{s=0} = \frac{v-1}{\ln 2} \cdot \psi_2^{(v)}(\hat{s}, v) \big|_{s=0} \tag{9.28}
\]

Moreover, we introduce the auxiliary function

\[
\xi_k(s) = [-\tilde{s} - v + 1]_k \prod_{l=1}^{k} (-\tilde{s} - v + l) \quad \xi_0(s) = 1 \tag{9.29}
\]

with
and derivative

\[ \frac{\partial \xi_k(s)}{\partial s} \bigg|_{s=0} = \frac{1}{\ln 2} \xi_k(0) \cdot \sum_{l=1}^{k} \frac{1}{-v + l}. \]  

(9.31)

For finding the derivative \( \frac{\partial }{\partial s} \left[ \Psi_{\Sigma \Omega}(s) \right] \), we can apply formula (3.67) for differentiation of a determinant. Moreover, we can use (see also (11.88))

\[ U(1, \tilde{s} + 2, \frac{1}{\gamma \sigma_i \omega_j}) \bigg|_{s=0} = \int_0^\infty e^{-\gamma \sigma_i \omega_j} dt = \gamma \sigma_i \omega_j \]  

(9.32)

and by exchanging the sequence of integration and differentiation by Lebesgue’s dominated convergence theorem, it is possible to derive

\[ \frac{\partial }{\partial s} U(1, \tilde{s} + 2, \frac{1}{\gamma \sigma_i \omega_j}) \bigg|_{s=0} = \int_0^\infty e^{-\gamma \sigma_i \omega_j} \cdot \ln(1 + t) dt = \gamma \sigma_i \omega_j \cdot E_1 \left( \frac{1}{\gamma \sigma_i \omega_j} \right), \]  

(9.33)

where \( E_1(x) \) is the exponential integral (see appendix 11.6.3). Without going into the details, by application of (9.27)-(9.33) for the calculation of (9.25), combination and simplification of the resulting determinants, we find (9.34). QED.

9.2.2 Derivation of Theorem 3.10

From Theorem 3.6 we can derive the alternative notation

\[ C_{\text{erg}, \Sigma \Omega} = \frac{\Gamma_{\gamma}(v)}{\ln 2 \cdot \alpha_\mu(\Sigma) \cdot \alpha_\gamma(\Omega) \cdot \gamma^{v-1}} \cdot \sum_{l=1}^{\mu} \left| \xi_{\Sigma \Omega}(l) \right| \]  

(9.34)

with the \( \mu \times v \) matrix (\( i \) runs from 1 to \( \mu \) and \( j \) from 1 to \( v \))

\[ \xi_{\Sigma \Omega}(l) = \begin{cases} \Gamma(v) \cdot \sigma_i^{v-1} \cdot (\gamma \omega_j)^{v-1} \cdot e^{-\gamma \sigma_i \omega_j} \cdot E_1 \left( \frac{1}{\gamma \sigma_i \omega_j} \right) & i = l \\ \sum_{k=\nu-\mu}^{\nu-1} \sigma_i^{k-(v-\mu)} \cdot (\gamma \omega_j)^k \cdot \frac{\Gamma(v)}{\Gamma(v-k)} & i \neq l \end{cases} \]  

(9.35)

and the \( (v-\mu) \times v \) matrix
\[
\Psi_{\Sigma \Omega, 2}(0) = \left[ (\gamma \omega_j)^{v-\mu-i} \cdot \frac{\Gamma(v)}{\Gamma(\mu + i)} \right]. \tag{9.36}
\]

For the derivation of (9.34)-(9.36) we have used
\[
(-1)^k [1 - n]_k = \frac{\Gamma(n)}{\Gamma(n - k)} \tag{9.37}
\]
and exchanged the rows of \( \Psi_{\Sigma \Omega, 2}(0) \) given in (3.50). With the help of the series expansion in (11.77) and determinant formula (11.13) we can split
\[
\begin{vmatrix}
\Xi_{\Sigma \Omega}(l) \\
\Psi_{\Sigma \Omega, 2}(0)
\end{vmatrix} =
\begin{vmatrix}
U_1(l) \\
\Psi_{\Sigma \Omega, 2}(0)
\end{vmatrix} +
\begin{vmatrix}
U_2(l) \\
\Psi_{\Sigma \Omega, 2}(0)
\end{vmatrix} -
\begin{vmatrix}
U_3(l) \\
\Psi_{\Sigma \Omega, 2}(0)
\end{vmatrix}, \tag{9.38}
\]
with the \( \mu \times \nu \) auxiliary matrices (\( i \) runs from 1 to \( \mu \) and \( j \) from 1 to \( \nu \))
\[
U_1(l) = \begin{bmatrix}
\Gamma(v) \cdot \sigma_i^{\mu-1} \cdot (\gamma \omega_j)^{v-1} \cdot (-E + \ln \gamma \sigma_i) \cdot \sum_{n=0}^{\infty} \frac{(\gamma \sigma_i \omega_j)^{-n}}{n!} & i = l \\
\sum_{k=v-\mu}^{v-1} \sigma_i^{k-(v-\mu)} \cdot (\gamma \omega_j)^{k} \cdot \frac{\Gamma(v)}{\Gamma(v-k)} & i \neq l
\end{bmatrix}, \tag{9.39}
\]
\[
U_2(l) = \begin{bmatrix}
\Gamma(v) \cdot \sigma_j^{\mu-1} \cdot (\gamma \omega_j)^{v-1} \cdot \ln \omega_j \cdot \sum_{n=0}^{\infty} \frac{(\gamma \sigma_i \omega_j)^{-n}}{n!} & i = l \\
\sum_{k=v-\mu}^{v-1} \sigma_j^{k-(v-\mu)} \cdot (\gamma \omega_j)^{k} \cdot \frac{\Gamma(v)}{\Gamma(v-k)} & i \neq l
\end{bmatrix}, \tag{9.40}
\]
and finally
\[
U_3(l) = \begin{bmatrix}
\Gamma(v) \cdot \sigma_j^{\mu-1} \cdot (\gamma \omega_j)^{v-1} \cdot \sum_{k=1}^{\infty} a_{E,k} \cdot (\gamma \sigma_i \omega_j)^{-k} & i = l \\
\sum_{k=v-\mu}^{v-1} \sigma_j^{k-(v-\mu)} \cdot (\gamma \omega_j)^{k} \cdot \frac{\Gamma(v)}{\Gamma(v-k)} & i \neq l
\end{bmatrix}. \tag{9.41}
\]

Combining the partial asymptotical results of the following sub-paragraphs 9.2.2.1-9.2.2.3 yields Theorem 3.10. \textit{QED.}
9.2.2.1 Asymptotics of \( \frac{U_1(l)}{\Psi \Sigma \Omega, 2(0)} \)

By elementary row operations and neglecting negative powers of \( \gamma \) at high SNR, we can derive from (9.39) (\( i \) runs from 1 to \( \mu \) and \( j \) from 1 to \( \nu \))

\[
\begin{vmatrix}
U_1(l) \\
- \Psi \Sigma \Omega, 2(0)
\end{vmatrix} \approx \begin{vmatrix}
\frac{1}{\Psi \Sigma \Omega, 2(0)} \\
\frac{1}{\Psi \Sigma \Omega, 2(0)} \\
\vdots \\
\frac{1}{\Psi \Sigma \Omega, 2(0)}
\end{vmatrix}
\]

\[
\left( - E + \ln \gamma \sigma_i \right) \cdot \sum_{k = \nu - \mu}^{\nu - 1} \sigma_i^{k-(\nu - \mu)} \cdot (\gamma \sigma_j)^k \cdot \frac{G(v)}{G(v-k)} \quad i = l
\]

\[
\sum_{k = \nu - \mu}^{\nu - 1} \sigma_i^{k-(\nu - \mu)} \cdot (\gamma \sigma_j)^k \cdot \frac{G(v)}{G(v-k)} \quad i \neq l
\]

After expanding the determinant via determinant formula (11.13), with

\[
\prod_{k=0}^{\nu-1} (-1)^k[1-v]_k = \prod_{k=0}^{\nu-1} \frac{G(v)}{G(v-k)} = \frac{[G(v)]^v}{G_v(v)}
\]

and

\[
\prod_{k=0}^{\nu-1} \gamma = \gamma^\frac{v \cdot (v-1)}{2}
\]

we can row-wise extract common factors and find (\( i \) runs from 1 to \( \mu \) and \( j \) from 1 to \( \nu \))

\[
\begin{vmatrix}
U_1(l) \\
- \Psi \Sigma \Omega, 2(0)
\end{vmatrix} \approx \gamma \frac{v \cdot (v-1)}{2} \cdot \frac{[G(v)]^v}{G_v(v)} \cdot (- E + \ln \gamma \sigma_i) \cdot \sum_{k = \nu - \mu}^{\nu - 1} \sigma_i^{k-(\nu - \mu)} \cdot \omega_j^k
\]

with the \((\nu - \mu)\times\nu\) matrix (\( i' \) runs from 1 to \( \nu - \mu \) and \( j' \) from 1 to \( \nu \))

\[
V = [\omega_j^{\nu - \mu - i'}].
\]

With the Vandermonde product determinant identity (11.24) it can then be shown that

\[
\begin{vmatrix}
U_1(l) \\
- \Psi \Sigma \Omega, 2(0)
\end{vmatrix} \approx \gamma \frac{v \cdot (v-1)}{2} \cdot \frac{[G(v)]^v}{G_v(v)} \cdot (- E + \ln \gamma \sigma_i) \cdot \alpha_{\mu}(\Sigma) \cdot \alpha_{v}(\Omega).
\]
9.2.2.2 Asymptotics of \( \frac{U_2(l)}{\Psi_{\Sigma, \Omega}(0)} \)

Similar to the derivation that lead to (9.45), we can find

\[
\frac{U_2(l)}{\Psi_{\Sigma, \Omega}(0)} \approx \gamma \frac{v \cdot (v-1)}{2 \cdot \Gamma(v)} \cdot \frac{[\Gamma(v)]^y}{\Gamma_v(v)} \cdot \sum_{k = v - \mu}^{v - 1} \sigma_j^{k-(v - \mu)} \cdot \omega_j^k \cdot \ln \omega_j \cdot \alpha^y_i \cdot \ln \omega_j \cdot \omega_j^{y-i}
\]

where matrix \( V \) is defined in (9.46). With a matrix factorization similar to (11.24) it is straightforward to show that \((i, j \text{ run from 1 to } v)\)

\[
\frac{U_2(l)}{\Psi_{\Sigma, \Omega}(0)} \approx \gamma \frac{v \cdot (v-1)}{2 \cdot \Gamma(v)} \cdot \frac{[\Gamma(v)]^y}{\Gamma_v(v)} \cdot \alpha^y_i \cdot \sum_{k = v - \mu}^{v - 1} \sigma_j^{k-(v - \mu)} \cdot \omega_j^k \cdot \ln \omega_j \cdot \omega_j^{y-i} \quad i = l
\]

9.2.2.3 Asymptotics of \( \frac{U_3(l)}{\Psi_{\Sigma, \Omega}(0)} \)

Discarding sum terms in (9.41) that are negligible at high SNR and re-indexing the summation, it can be shown that \((i \text{ runs from 1 to } \mu \text{ and } j \text{ from 1 to } v)\)

\[
\frac{U_3(l)}{\Psi_{\Sigma, \Omega}(0)} \approx \gamma \frac{v \cdot (v-1)}{2 \cdot \Gamma(v)} \cdot \frac{[\Gamma(v)]^y}{\Gamma_v(v)} \cdot \sum_{l = v - \mu}^{v - 1} \frac{\Gamma(v)}{v - k} \cdot \sigma_i^{k-(v - \mu)} \cdot \frac{\Gamma(v)}{v - k} \cdot \omega_j^{y-i} \quad i \neq l
\]

Then consider \((i \text{ runs from 1 to } \mu \text{ and } j \text{ from 1 to } v)\)
By expanding the determinants via formula (11.13) it can now be shown that

\[
\sum_{l=1}^{\mu} \left| \frac{U_3(l)}{\Psi \Sigma \Omega, 2(0)} \right| = \Gamma(v) \cdot \sum_{l=1}^{\mu} \left| \begin{array}{c}
\nu - 2 \\
\sum_{l=1}^{\mu} a_{E, v-1-l} \cdot (\gamma \omega_j)^l \cdot \sigma_i^{-(\nu - \mu)} i = l \\
\sum_{l=1}^{\mu} a_{E, v-1-l} \cdot (\gamma \omega_j)^l \cdot \Gamma(v) \Gamma(v-l) i \neq l \\
\sum_{k=\nu-1}^{\mu} \sigma_i^{k-(\nu - \mu)} (\gamma \omega_j)^k \Gamma(v) \Gamma(v-k)
\end{array} \right|.
\]  

(9.51)

This can be rewritten as

\[
\sum_{l=1}^{\mu} \left| \frac{U_3(l)}{\Psi \Sigma \Omega, 2(0)} \right| = \frac{[\Gamma(v)]^v \cdot \alpha_\mu(\Sigma) \cdot \alpha_\nu(\Omega)}{\Gamma(v)} \cdot \sum_{l=1}^{\mu} a_{E, v-1-l} \cdot \prod_{k=0}^{v-1} \frac{\Gamma(v)}{\Gamma(v-k)}.
\]  

(9.52)

9.2.3 Derivation of Corollary 3.6

Consider the limit \((i, j)\) run from 1 to \(v\) of \(\zeta_1(\Omega)\) according to (3.82)

\[
\lim_{\Omega \to I} \zeta_1(\Omega) = \sum_{l=1}^{\mu} \lim_{\omega \to I} \frac{\omega_j^{v-i} \cdot \ln \omega_j}{\alpha_\nu(\Omega)}.
\]  

(9.54)

which can be calculated via L’Hospital’s rule, where we take the \((j-1)\)th derivative with respect to \(\omega_j\) in nominator and denominator. To this end, note that it can be shown that

\[
\frac{\partial^k}{\partial x^k}(x^n \cdot \ln x) = x^{n-k} \cdot \left( b_k \cdot \ln x + a_k \right)
\]  

(9.55)

with the recursive definitions

\[
a_k = (n-k+1) \cdot a_{k-1} + b_{k-1} \quad b_k = [n-k+1]_k
\]  

(9.56)

and

\[
a_0 = 0 \quad b_0 = 1.
\]  

(9.57)

From (9.55) and (9.56) we directly get
with the recursive definitions

\[ a_k(i) = (v - i - k + 1) \cdot a_{k-1}(i) + b_{k-1}(i) \quad b_k(i) = [v - i - k + 1]_k \]  

and

\[ a_0(i) = 0 \quad b_0(i) = 1. \]  

In the limit we have

\[ \frac{\partial^k}{\partial \omega_j x^v} (x^{v-i} \cdot \ln x) \bigg|_{x = 1} = a_k(i). \]  

On the other hand, we get

\[ \frac{\partial^k}{\partial \omega_j x^v} (x^{v-i} \cdot \ln x) = [v - i - k + 1]_k \cdot \omega_j^{v-i-k} \]  

with the limit

\[ \frac{\partial^k}{\partial \omega_j x^v} \bigg|_{x = 1} = [v - i - k + 1]_k. \]  

Similarly, after the limiting process we find the resulting expression of the Vandermonde determinant in the denominator of (9.54)

\[ \frac{\partial}{\partial \omega} \alpha_{\nu}(\Omega) \bigg|_{\omega = 1} = (-1)^{v-1} \frac{v \cdot (v-1)}{2} \cdot \Gamma(v). \]  

Summarizing the results, we have Corollary 3.6. QED.

### 9.2.4 Derivation of Theorem 3.11

We start the derivation with the ergodic capacity expression (9.34). Using approximation (11.80) in (9.35) we arrive at (i runs from 1 to \( \mu \) and j runs from 1 to \( v \))

\[ \Xi_{\Sigma \Omega} (l) = \begin{cases} \Gamma(v) \cdot \sigma_\mu^l \cdot (\gamma \omega_j)^v & i = l \\ \sum_{k = v - \mu}^{v-1} \sigma_{i-k}^{l-(v-\mu)} \cdot (\gamma \omega_j)^{i-k} \cdot \frac{\Gamma(v)}{\Gamma(v-k)} & i \neq l \end{cases}. \]  

Then consider the determinant
At very low SNR we can neglect higher powers of $\gamma$. From the expansion of the determinant according to (11.13) we know that only those determinants of the expansion remain (do not vanish to 0), where terms with different powers of $\gamma$ stand in the rows (otherwise we get linearly dependent rows). We can then neglect those determinants with power $\gamma^{\nu-1}$ in one of the rows, such that we find ($i$ runs from 1 to $\mu$ and $j$ runs from 1 to $\nu$)

$$
\begin{vmatrix}
\Xi_{\Sigma\Omega}(l) \\
\Psi_{\Sigma\Omega, 2}(0)
\end{vmatrix}
= \begin{cases}
\Gamma(v) \cdot \sigma_i^\mu \cdot (\gamma \omega_j)^v & i = l \\
\sum_{k = \nu - \mu}^{\nu - 2} \sigma_i^{k-(\nu - \mu)} \cdot (\gamma \omega_j)^k \cdot \frac{\Gamma(v)}{\Gamma(v - k)} & i \neq l
\end{cases}
\frac{1}{\Psi_{\Sigma, 2}(0)}.
$$

(9.67)

Now factor out the terms

$$
\prod_{k = 0}^{\nu - 1} \frac{\Gamma(v)}{\Gamma(v - k)} = \frac{[\Gamma(v)]^\nu}{\Gamma_\nu(v)}
$$

(9.68)

and

$$
\gamma^\nu \cdot \prod_{k = 0}^{\nu - 2} \gamma^k = \gamma^\nu \cdot \gamma^{(v-2) \cdot (v-1)} = \gamma^{\frac{v^2 - v + 2}{2}}
$$

(9.69)

from (9.67) such that we find

$$
\begin{vmatrix}
\Xi_{\Sigma\Omega}(l) \\
\Psi_{\Sigma\Omega, 2}(0)
\end{vmatrix}
= \gamma \frac{[\Gamma(v)]^\nu}{\Gamma_\nu(v)} \cdot \begin{cases}
\sigma_i^\mu \cdot \omega_j^\nu & i = l \\
\sum_{k = \nu - \mu}^{\nu - 2} \sigma_i^{k-(\nu - \mu)} \cdot \omega_j^k & i \neq l
\end{cases}
\begin{vmatrix}
V
\end{vmatrix}
$$

(9.70)

with the $(\nu - \mu) \times \nu$ matrix $V$ defined in (9.46). It can be shown by an expansion of the determinants in (9.70) according to (11.13) that ($i$ runs from 1 to $\mu$ and $j$ from 1 to $\nu$)
\[
\sum_{l=1}^{\mu} \frac{\Xi_{\Sigma \Omega}(l)}{\Psi_{\Sigma \Omega, 2}(0)} = \gamma \left( \frac{v^2 - v + 2}{2} \right) \cdot \frac{[\Gamma(v)]^\nu}{\Gamma_v(v)} \cdot \sigma_i^\mu \sigma_i^{\mu-2} \sigma_i^{\mu-3} \cdot \sigma_i 1 \cdot \omega_j^\nu \end{array} \right|_{\nu = 0} = \left| \begin{array}{c} \omega_j^\nu \\
\omega_j^{\nu-2} \\
\omega_j^{\nu-3} \\
\vdots \\
1 \\
\end{array} \right| \right. \quad (9.71)
\]

Without loss of generality, first consider the determinant \((i \text{ runs from } 1 \text{ to } \mu)\)
\[
d \equiv \left| \begin{array}{ccc}
\sigma_i^\mu & \sigma_i^{\mu-2} & \sigma_i^{\mu-3} & \sigma_i 1 \\
\end{array} \right| \quad (9.72)
\]

For finding an explicit expression of the determinant \(d\) we can follow the derivation of the Vandermonde determinant. First of all, \(d\) is a polynomial in the \(\sigma_i\). Furthermore, \(d\) equals to 0, if \(\sigma_i = \sigma_i\) for any \(i, i'\) (two rows in the determinant are equal). By the factor theorem for polynomials, it is clear that any \(\sigma_i - \sigma_i\) divides the polynomial. This, however, implies that we can express the determinant by
\[
d = \prod_{i > i'} (\sigma_i - \sigma_i) \cdot \phi(\sigma_1, \sigma_2, \ldots, \sigma_\mu) = \alpha_\mu(\Sigma) \cdot \phi(\sigma_1, \sigma_2, \ldots, \sigma_\mu), \quad (9.73)
\]

where \(\phi(\sigma_1, \sigma_2, \ldots, \sigma_\mu)\) is a polynomial in the \(\sigma_i\). Then note that the Vandermonde determinant is a sum of polynomials of maximum degree \(\mu \cdot (\mu - 1)/2\). On the other hand, the determinant \(d\) is a sum of determinants of degree \(\mu \cdot (\mu - 1)/2 + 1\), i.e. \(\phi\) can be of maximum degree 1, i.e. it has the structure
\[
\phi(\sigma_1, \sigma_2, \ldots, \sigma_\mu) = \sum_{l=1}^{\mu} a_l \cdot \sigma_l, \quad (9.74)
\]

where the \(a_l\) are constant coefficients. Now consider the product of the elements of the main diagonal of \(d\) and the product of elements on the main diagonal of the Vandermonde determinant \(\alpha_\mu(\Sigma)\). One can directly derive \(a_1 = 1\) and due to the symmetry of the problem it easily follows that \(a_l = 1\) for all \(l\). Summarizing the results, we have found
\[
d = \alpha_\mu(\Sigma) \cdot \sum_{l=1}^{\mu} \sigma_l = \alpha_\mu(\Sigma) \cdot \tr(\Sigma). \quad (9.75)
\]

For the determinant in (9.71) we find (after a similar derivation for the determinant of the \(\omega_j\))
\[
\sum_{l=1}^{\mu} \frac{\Xi_{\Sigma \Omega}(l)}{\Psi_{\Sigma \Omega, 2}(0)} = \gamma \left( \frac{v^2 - v + 2}{2} \right) \cdot \frac{[\Gamma(v)]^\nu}{\Gamma_v(v)} \cdot \alpha_\mu(\Sigma) \cdot \tr(\Sigma) \cdot \alpha_\mu(\Omega) \cdot \tr(\Omega). \quad (9.76)
\]

Plugging (9.76) in (9.34) and simplifying yields Theorem 3.11. \(QED.\)
9.3 Ergodic Capacity Bound

9.3.1 Derivation of Theorem 4.1

Due to the concavity of the log function we get via Jensen’s inequality [24] the upper bound

\[ C_{\text{erg}, \Sigma \Omega}(\gamma) \leq C_{\text{erg}}^B(\gamma) = \log_2 E[|I + \gamma \Sigma G^H \Omega G|]. \tag{9.77} \]

The expected value in (9.77) can be calculated in closed form via determinant expansion and the theory of complex Wishart matrices. To this end, we are applying determinant expansion formula (11.15) and get

\[ C_{\text{erg}}^B(\gamma) = \log_2 E \left[ \sum_{k=0}^{\mu} \sum_{\alpha_k} |\gamma \Sigma G^H \Omega G|_{\alpha_k}^{\hat{\alpha}_k} \right] = \log_2 E \left[ \sum_{k=0}^{\mu} \sum_{\alpha_k} |\gamma \Sigma^{1/2} G^H \Omega G \Sigma^{1/2}|_{\alpha_k}^{\hat{\alpha}_k} \right]. \tag{9.78} \]

Obviously, the expression within the expected value in (9.78) is a sum of quadratic forms. In principle, we could apply results on the expected value of quadratic forms, which is given for the real case in [167], for solving (9.78). However, in this thesis we take a different, independent approach. The subdeterminants of size $k \times k$ in (9.78) can be expanded via the determinant formula (11.17)

\[ |\gamma \Sigma G^H \Omega G|_{\alpha_k}^{\hat{\alpha}_k} = \sum_{\beta_k} \sum_{\delta_k} |\gamma \Sigma|_{\beta_k}^{\hat{\beta}_k} \cdot |G^H|_{\delta_k}^{\hat{\delta}_k} \cdot |\Omega|_{\delta_k}^{\hat{\delta}_k} \cdot |G|_{\alpha_k}^{\hat{\alpha}_k}. \tag{9.79} \]

In a next step, the diagonal structure of $\Omega$ can be exploited by noticing that

\[ |\Omega|_{\delta_k}^{\hat{\delta}_k} = 0 \quad \hat{\delta}_k \neq \sigma_k, \tag{9.80} \]

as the submatrix that results from selecting the row and column subsets indexed by $\hat{\delta}_k$ and $\sigma_k$ has at least one row that contains only zeros. Now we get from (9.79)

\[ |\gamma \Sigma G^H \Omega G|_{\alpha_k}^{\hat{\alpha}_k} = \sum_{\beta_k} \sum_{\delta_k} |\gamma \Sigma|_{\beta_k}^{\hat{\beta}_k} \cdot |G^H|_{\delta_k}^{\hat{\delta}_k} \cdot |\Omega|_{\delta_k}^{\hat{\delta}_k} \cdot |G|_{\alpha_k}^{\hat{\alpha}_k}. \tag{9.81} \]

Obviously, by the same arguments we also have

\[ |\gamma \Sigma|_{\beta_k}^{\hat{\beta}_k} = 0 \quad \hat{\beta}_k \neq \beta_k, \tag{9.82} \]

such that we can further simplify (9.81) and arrive at

\[ |\Sigma G^H \Omega G|_{\alpha_k}^{\hat{\alpha}_k} = \sum_{\delta_k} |\gamma \Sigma|_{\delta_k}^{\hat{\delta}_k} \cdot |G^H|_{\delta_k}^{\hat{\delta}_k} \cdot |\Omega|_{\delta_k}^{\hat{\delta}_k} \cdot |G|_{\alpha_k}^{\hat{\alpha}_k}. \tag{9.83} \]

By inspecting (9.83), it can be seen that
\[ |G_H^{\hat{\alpha}_k} \hat{\delta}_k \cdot |G|_{\alpha_k}^{\hat{\delta}_k} = |W|, \]  

(9.84)

where the \( k \times k \) matrix \( W \) has a complex Wishart distribution with identity covariance matrix, i.e. \( W \sim \tilde{W}_k(k, I_k) \). It is well known (e.g. [44]) that

\[ E[|W|] = k!, \]  

(9.85)

such that we get from (9.83) with (9.84) and (9.85)

\[ E[|\gamma \Sigma G^H \Omega G|_{\alpha_k}^{\hat{\alpha}_k}] = \sum_{\hat{\delta}_k} |\Sigma|_{\alpha_k}^{\hat{\alpha}_k} \cdot |\Omega|_{\delta_k}^{\hat{\delta}_k} \cdot k! \cdot \gamma^k. \]  

(9.86)

Plugging (9.86) in (9.78) we finally arrive at

\[ C_{erg}^B(\gamma) = \log_2 \left( \sum_{\hat{\alpha}_k, \delta_k} \sum_{\alpha_k} |\Sigma|_{\alpha_k}^{\hat{\alpha}_k} \cdot |\Omega|_{\delta_k}^{\hat{\delta}_k} \cdot k! \cdot \gamma^k \right), \]  

(9.87)

leading to (4.3) with (4.4). \( QED. \)
10 Appendix - MIMO Receiver Performance

10.1 MIMO ZF Receiver SER Calculation

10.1.1 Proof of Theorem 5.1

For brevity, we introduce the compound channel $R \times L$ matrix $K$ with the rank $L$ matrix prefilter $F$

$$K = HF.$$ \hspace{1cm} (10.1)

Exploiting the properties of the vec and Kronecker operators and using results on complex matrix multivariate normal distributions (see appendix 11.2.1), it can be seen that the statistics of $K$ for the channel model in (2.41) are given by (with $C$ defined in (5.11))

$$K \sim \tilde{N}_{\text{rx,L}}(0, R_{RX} \otimes (F^H R_{TX} F)^+) = \tilde{N}_{\text{rx,L}}(0, R_{RX} \otimes C^+),$$ \hspace{1cm} (10.2)

which has zero mean, as we assume a vanishing Ricean (deterministic) component. The receiver zero-forcing matrix filter (for AWGN) is given by the pseudo-inverse [from (5.2)]

$$G = K^\dagger = (K^H K)^{-1} K^H,$$ \hspace{1cm} (10.3)

with $GK = I$. Under the AWGN assumption it is then straightforward to show that the signal-to-noise ratio on subchannel 1 can be expressed as [see (5.8)]

$$\gamma_{SC,1} = \frac{\gamma}{[K^\dagger (K^\dagger)^H]_{11}} = \frac{\gamma}{[(K^H K)^{-1}]_{11}}.$$ \hspace{1cm} (10.4)

We note that we present a derivation for the SNR expression on subchannel 1 in order to simplify the notation. An extension to an arbitrary subchannel is then a simple task by symmetry considerations and left to the reader. Partitioning the $R \times L$ matrix $K$ as

$$K = \begin{bmatrix} k_1 & \tilde{K} \end{bmatrix}$$ \hspace{1cm} (10.5)

with $R \times 1$ column vector $k_1$, $R \times (L - 1)$ matrix $\tilde{K}$, and using the inversion identity for partitioned matrices (11.11) in the appendix, we can rewrite (5.7) as

$$\gamma_{SC,1} = \gamma \cdot k_1^H (I - \tilde{K} (\tilde{K}^H \tilde{K})^{-1} \tilde{K}^H) k_1.$$ \hspace{1cm} (10.6)

However, from (11.31) we know that (using the same notation as in (11.10) for partitioning $C$)

$$\tilde{K} \sim \tilde{N}_{R,L-1}(0, R_{RX} \otimes C_{22}^*),$$ \hspace{1cm} (10.7)

and equivalently (with scalar $C_{11,2}^*$) we obtain the conditional statistics
\[ k_1 \mid \tilde{K} - \tilde{N}_R(\tilde{K}(C_{22}^{-1})C_{21}^*, C_{11} \cdot C_{21} \cdot R_{RX}). \] (10.8)

We emphasize that the covariance matrix of \( k_1 \) is just a scaled version of the receive correlation matrix and with equality (11.11) we find that

\[ C_{11} = \frac{1}{1} = \frac{1}{[C^{-1}]_{kk}}. \] (10.9)

For brevity, let \( a = (C_{22}^*)^{-1}C_{21}^* \) and introduce a new random vector

\[ x \sim \tilde{N}_{R, 1}(0, C_{11} \cdot C_{21} \cdot R_{RX}). \] (10.10)

Using the definitions above we find from (10.6) that

\[ \gamma_{SC, 1} \equiv \gamma \cdot (x + \tilde{K}a)^H(I - \tilde{K}(\tilde{K}^H\tilde{K})^{-1}\tilde{K}^H)(x + \tilde{K}a). \] (10.11)

However, noting that

\[ a^H\tilde{K}^H(I - \tilde{K}(\tilde{K}^H\tilde{K})^{-1}\tilde{K}^H) = 0 \]
\[ (I - \tilde{K}(\tilde{K}^H\tilde{K})^{-1}\tilde{K}^H)a = 0, \] (10.12)

and inserting this result in (10.11) we get

\[ \gamma_{SC, 1} \equiv \gamma \cdot x^H(I - \tilde{K}(\tilde{K}^H\tilde{K})^{-1}\tilde{K}^H)x. \] (10.13)

We now explicitly model \( \tilde{K} \) and \( x \) as

\[ \tilde{K} \equiv A^H\tilde{H}_w C_{22}^{1/2} \quad x \equiv \frac{1}{\sqrt{[C^{-1}]_{kk}}}A^H u, \] (10.14)

where the \( R \times (L - 1) \) matrix \( \tilde{H}_w \) and \( R \times 1 \) vector \( u \) have unit variance i.i.d. complex Gaussian elements. Then we arrive at (note that \( C_{22}^{1/2} \) is invertible) the quadratic form of random variables

\[ \gamma_{SC, 1} \equiv \alpha_1 \cdot u^H A(I - A^H\tilde{H}_w(\tilde{H}_w^H AA^H\tilde{H}_w)^{-1} \tilde{H}_w^HA)A^H u = \alpha_1 \cdot u^H Qu, \] (10.15)

where we have introduced a scaling factor \( \alpha_1 = \gamma / ([C^{-1}]_{11}) \) and the random matrix \( Q \)

\[ Q = A(I - A^H\tilde{H}_w(\tilde{H}_w^H AA^H\tilde{H}_w)^{-1} \tilde{H}_w^HA)A^H \] (10.16)

as short-hand notation. Generalization to an arbitrary subchannel then yields Theorem 5.1. QED.

**10.1.2 Proof of Theorem 5.2**

We have to average the conditional SER expression in (5.16) over the subchannel SNR statistics to find the unconditional average SER
Plugging (5.14) and (5.16) in (10.17) we get with

\[ P_{s,k}(\gamma) = 2b \cdot \frac{1}{\alpha_k^N \cdot \Gamma(N)} \cdot \int_0^\infty Q(\sqrt{2c} \gamma_k) \cdot \exp\left(-\frac{\gamma_k}{\alpha_k}\right) \gamma_k^{N-1} d\gamma_k. \]  

Then note that in [32] the following integration result was established

\[ P(a, b, c) = \frac{a^b}{\Gamma(N)} \cdot \int_0^\infty Q(\sqrt{ct}) \cdot \exp(-at) \cdot t^{b-1} dt. \]  

The integration result reads for arbitrary \( b \)

\[ P(a, b, c) = \sqrt{\frac{\kappa}{1 + \kappa}} \cdot \frac{(1 + \kappa)^{-b} \cdot \Gamma\left(\frac{b + 1}{2}\right)}{2\sqrt{\pi} \cdot \Gamma(b + 1)} \cdot \, _2F_1\left(1, b + \frac{1}{2}; b + 1; (1 + \kappa)^{-1}\right) \]  

with hypergeometric function \(_2F_1(a_1, a_2; b_1; z)\) (see appendix 11.6.5) and parameter

\[ \kappa = \frac{c}{2a}. \]

For integer \( b \), the hypergeometric series in (10.20) terminates and it can be shown [32] that

\[ P(a, b, c) = \frac{1}{2} \left[ 1 - \sqrt{\frac{\kappa}{1 + \kappa}} \cdot \sum_{k=0}^{b-1} \binom{2k}{k} \left(\frac{1}{4(1 + \kappa)}\right)^k \right]. \]

Noting in (10.19) that

\[ P_{s,k}(\gamma) = 2b \cdot P\left[\frac{1}{\alpha_k}, N, 2c\right] \]  

and simplifying the result proves the theorem. \( QED. \)

### 10.1.3 Proof of Lemma 5.1

#### 10.1.3.1 Single integral representation

For brevity, with \( m \times m \) deterministic matrices \( \tilde{A}, \tilde{B} \), and \( m \times n \) i.i.d. complex Gaussian matrix \( G \), let
\[ r = E_G \left[ \frac{G^H \tilde{A} G}{G^H B G} \right]. \]  

(10.24)

We use integral identity (11.43) and get with \( n \times 1 \) complex vector \( x \)

\[ r = E_G \left[ |G^H \tilde{A} G| \cdot \int e^{-x^H (G^H \tilde{B} G) x} D_e x \right]. \]  

(10.25)

Using (11.42) for separating \( G \) and \( G^H \) (this is necessary for integrating out \( G \) later), we can express in (10.25)

\[ e^{-x^H G^H \tilde{B} G x} = (-1)^m e^{-(y^H y + x^H \tilde{B}^{1/2} y + y^H \tilde{B}^{1/2} G x)} D_e y \]  

(10.26)

with \( m \times 1 \) complex vector \( y \) . This can be reformulated

\[ e^{-x^H G^H \tilde{B} G x} = (-1)^m \int e^{y^H y} e^{-\text{tr}(x^H \tilde{B}^{1/2} G + G^H \tilde{B}^{1/2} y y^H)} D_e y. \]  

(10.27)

Introducing the auxiliary matrix

\[ C = \tilde{B}^{1/2} y x^H \]  

(10.28)

we get

\[ e^{-x^H G^H \tilde{B} G x} = (-1)^m \int e^{y^H y} e^{-\text{tr}(C^H G + G^H C)} D_e y. \]  

(10.29)

Plugging (10.29) in (10.25), we find

\[ r = (-1)^m \int \left[ E_G \left[ |G^H \tilde{A} G| \cdot e^{-\text{tr}(C^H G + G^H C)} \right] \cdot e^{y^H y} D_e y D_e x \right]. \]  

(10.30)

Now we focus on the expected value with respect to the Gaussian distribution of \( G \) in (10.30), such that we have to calculate

\[ g = E_G \left[ |G^H \tilde{A} G| \cdot e^{-\text{tr}(C^H G + G^H C)} \right]. \]  

(10.31)

With the PDF of \( G \) in (11.45) this can be written in integral form

\[ g = \int |G^H \tilde{A} G| \cdot e^{-\text{tr}(G^H G + C^H G + G^H C)} D_e G. \]  

(10.32)

Completing the square in the exponential, we find

\[ g = \int |G^H \tilde{A} G| \cdot e^{-\text{tr}(G + C)^H (G + C)} \cdot e^{\text{tr}(C^H C)} D_e G. \]  

(10.33)

The determinant in (10.33) can be split into a sum of determinants via determinant identity (11.17)
\[ |G^H \tilde{A} G| = \sum_{\tilde{\alpha}_n} |A|_{\tilde{\alpha}_n} |G^H|_{\{1, \ldots, n\}} |G|_{\{1, \ldots, n\}}. \]  

(10.34)

Now we define a complementary index subset

\[ \hat{\beta}_{m-n} = \{1, \ldots, m\} \backslash \tilde{\alpha}_n \]  

(10.35)

and for brevity

\[ G_{1, \alpha} = \{G\}_{\{1, \ldots, n\}} \quad G_{2, \alpha} = \{G\}_{\{1, \ldots, n\}}. \]  

(10.36)

Equivalently, we let

\[ C_{1, \alpha} = \{C\}_{\{1, \ldots, n\}} \quad C_{2, \alpha} = \{C\}_{\{1, \ldots, n\}}. \]  

(10.37)

Now (10.33) can be written with the help of (10.34) and (10.36) as

\[ g = \int \sum_{\tilde{\alpha}_n} |A|_{\tilde{\alpha}_n} |G^H_{1, \alpha} G_{1, \alpha}| \cdot e^{-\text{tr}(G + C)^{\mu}(G + C)} \cdot e^{\text{tr}(C^{\mu}C)D_c G}, \]  

(10.38)

and

\[ g = \sum_{\tilde{\alpha}_n} (|A|_{\tilde{\alpha}_n} \cdot I_{\tilde{\alpha}_n}) \cdot e^{\text{tr}(C^{\mu}C)} \]  

(10.39)

with the prototype integral

\[ I_{\tilde{\alpha}_n} = \int |G^H_{1, \alpha} G_{1, \alpha}| \cdot e^{-\text{tr}(G + C)^{\mu}(G + C)D_c G}. \]  

(10.40)

It is possible to rewrite (10.40) as

\[ I_{\tilde{\alpha}_n} = \int |G^H_{1, \alpha} G_{1, \alpha}| \cdot e^{-\text{tr}(G_{1, \alpha} + C_{1, \alpha})^{\mu}(G_{1, \alpha} + C_{1, \alpha})D_c G_{1, \alpha}} \cdots \int e^{-\text{tr}(G_{2, \alpha} + C_{2, \alpha})^{\mu}(G_{2, \alpha} + C_{2, \alpha})D_c G_{2, \alpha}}. \]  

(10.41)

Now we can use integral identity (11.52)

\[ \int |G^H_{1, \alpha} G_{1, \alpha}| \cdot e^{-\text{tr}(G_{1, \alpha} + C_{1, \alpha})^{\mu}(G_{1, \alpha} + C_{1, \alpha})D_c G_{1, \alpha}} = \left[ \Gamma(n + 1) + \Gamma(n) \cdot \text{tr}(C_{1, \alpha} C_{1, \alpha}^H) \right]. \]  

(10.42)

and on the other hand it follows

\[ \int e^{-\text{tr}(G_{2, \alpha} + C_{2, \alpha})^{\mu}(G_{2, \alpha} + C_{2, \alpha})D_c G_{2, \alpha}} = 1. \]  

(10.43)

Summarizing, we have

\[ I_{\tilde{\alpha}_n} = \left[ \Gamma(n + 1) + \Gamma(n) \cdot \text{tr}(C_{1, \alpha} C_{1, \alpha}^H) \right]. \]  

(10.44)
After partitioning the matrix $B$

$$
B_{1, \alpha} = \{\tilde{B}\}_{\alpha_n} \hspace{1cm} B_{2, \alpha} = \{\tilde{B}\}_{\beta_m - \alpha_n} \hspace{1cm} (10.45)
$$

and the vector $y$

$$
y_{1, \alpha} = \{y\}_{1}^{\hat{\alpha}_n} \hspace{1cm} y_{2, \alpha} = \{y\}_{1}^{\hat{\beta}_m - \alpha_n} \hspace{1cm} (10.46)
$$

we get from (10.28)

$$
C_{1, \alpha} = B_{1, \alpha}^{1/2} y_{1, \alpha} x^H \hspace{1cm} C_{2, \alpha} = B_{2, \alpha}^{1/2} y_{2, \alpha} x^H. \hspace{1cm} (10.47)
$$

From (10.30) together with (10.39) and (10.44) we get

$$
r = (-1)^m \sum_{\hat{\alpha}_n} (\hat{\alpha}_n \cdot r_{\alpha}) \hspace{1cm} (10.48)
$$

with

$$
r_{\alpha} = \left[ \left( n \cdot \text{tr}(y_{1, \alpha}^H B_{1, \alpha} y_{1, \alpha} x^H x) \right) - \beta \text{tr}(y_{1, \alpha}^H B_{1, \alpha} y_{1, \alpha} x^H x) \right] \cdot \text{tr}(y_{2, \alpha}^H B_{2, \alpha} y_{2, \alpha} x^H x) \cdot \text{tr}(y_{2, \alpha}^H B_{2, \alpha} y_{2, \alpha} x^H x) \cdot \text{tr}(y_{2, \alpha}^H D x D y) \hspace{1cm} (10.49)
$$

This can be written as

$$
r_{\alpha} = \Gamma(n) \left[ \int [ n \cdot \text{tr}(y_{1, \alpha}^H B_{1, \alpha} y_{1, \alpha} x^H x) \cdot e^{i\pi y_{1, \alpha}^H B_{1, \alpha} y_{1, \alpha} x^H x} \cdot e^{i\pi y_{1, \alpha}^H B_{1, \alpha} y_{1, \alpha} x^H x} \cdot e^{i\pi y_{2, \alpha}^H D x D y} ] \right] \hspace{1cm} (10.50)
$$

Now integrate with respect to the $n \times 1$ vector $y_{1, \alpha}$ with the help of (11.41) (using the abbreviation $b_{j(k)} = [B_{1, \alpha}]_{kk}$)

$$
r_{\alpha} = \Gamma(n) \left[ \int \frac{n}{1 + x^H x B_{1, \alpha}} \left[ \sum_{k=1}^{n} \frac{x^H x \cdot b_{j(k)}}{1 + x^H x \cdot b_{j(k)} + I} \right] e^{i\pi y_{1, \alpha}^H (x^H x B_{1, \alpha} + I) y_{1, \alpha}} e^{i\pi x^H D x D y} \right] \hspace{1cm} (10.51)
$$

Then integrating with respect to the $(m-n) \times 1$ vector $y_{2, \alpha}$ we find

$$
r_{\alpha} = \Gamma(n) \cdot (-1)^m \left[ \int \frac{n}{1 + x^H x B_{1, \alpha}} - \sum_{k=1}^{n} \frac{x^H x \cdot b_{j(k)}}{1 + x^H x \cdot b_{j(k)} + I} \right] \frac{1}{1 + x^H x D_{1, \alpha}} \hspace{1cm} (10.52)
$$

and after simplifying

$$
r_{\alpha} = \Gamma(n) \cdot (-1)^m \left[ \int \frac{n}{1 + x^H x B_{1, \alpha}} - \sum_{k=1}^{n} \frac{x^H x \cdot b_{j}}{1 + x^H x \cdot b_{j(k)} + I} \right] \hspace{1cm} (10.53)
$$

We can transform the integral in (10.53) into a single integral by using identity (11.44)
and after combining (10.48) and (10.54), the first part of Lemma 5.1 is established. QED.

### 10.1.3.2 Closed form representation

Consider the prototype integral

\[
I_{rat}(\tilde{B}, n, j) = \int_0^\infty \frac{t^{n-1}}{(1 + b_j \cdot t)|I + i\tilde{B}|} dt,
\]

(10.55)

with

\[
\tilde{B} = \text{diag}(b_1, b_2, \ldots, b_m).
\]

(10.56)

By proper contour integration and the residue theorem, it can be shown that the integral in (10.55) is given by

\[
I_{rat}(\tilde{B}, n, j) = -\sum_l \text{Res}_l \left( \frac{\ln(|z|) \cdot z^{n-1}}{(1 + z b_j)|I + z\tilde{B}|} \right)_{z = z_l},
\]

(10.57)

where the sum is over all residues at the poles \(z_l\) of the indicated function. To this end, note that there are single poles at \(z_k = -1/b_k\) for all \(k = \{1, \ldots, m\}\) and a double pole at \(z_j = -1/b_j\). At the single poles we get

\[
\text{Res}_k \left( \frac{\ln(|z|) \cdot z^{n-1}}{(1 + z b_j)|I + z\tilde{B}|} \right)_{z = -1/b_k} = \frac{\ln(|z|) \cdot z^{n-1} \cdot \left(z - \frac{1}{b_k}\right)}{(1 + z b_j)|I + z\tilde{B}|} |_{z = -1/b_k}.
\]

(10.58)

After some straightforward manipulations we find (5.36). On the other hand, at the double pole \(z_j = -1/b_j\) we find

\[
\text{Res}_j \left( \frac{\ln(|z|) \cdot z^{n-1}}{(1 + z b_j)|I + z\tilde{B}|} \right)_{z = -1/b_j} = \frac{d}{dz} \left( \frac{\ln(|z|) \cdot z^{n-1} \cdot \left(z - \frac{1}{b_j}\right)^2}{(1 + z b_j)|I + z\tilde{B}|} \right) |_{z = -1/b_j}.
\]

(10.59)

Now carrying out the differentiation and simplifying the results, we can establish (5.37). QED.
10.1.4 Proof of Theorem 5.7

We approximate the expected value in (5.30) by

\[ E_{\hat{H}_w} \left[ \frac{\hat{H}_w^H O \hat{H}_w}{\hat{H}_w^H O (I + s\alpha_k O)^{-1} \hat{H}_w} \right] \approx \frac{E_{\hat{H}_w} \left[ \hat{H}_w^H O \hat{H}_w \right]}{E_{\hat{H}_w} \left[ \hat{H}_w^H O (I + s\alpha_k O)^{-1} \hat{H}_w \right]} \]  

(10.60)

Now using the expected value in (11.34) yields (5.39). The separation of the expected value in (10.60) is motivated by the so-called Laplace approximation presented in [123] for the vector-valued case (with i.i.d. complex Gaussian vector \( u \))

\[ E_u \left[ \frac{u^H \tilde{A} u}{u^H \tilde{B} u} \right] = \frac{E_u [u^H \tilde{A} u]}{E_u [u^H \tilde{B} u]} \]  

(10.61)

and the fact that the approximation becomes exact in the high SNR regime (see below). For high SNR we have \( \alpha_k \to \infty \) and (5.30) reduces to

\[ \bar{M}_{T,k}(s) = \frac{1}{|O|} \cdot E_{\hat{H}_w} \left[ \frac{\hat{H}_w^H O \hat{H}_w}{|\hat{H}_w^H \hat{H}_w|} \right] \cdot (s\alpha_k)^{-N}. \]  

(10.62)

Based on a result derived in [152] for chi-squared distributed random variables, in [84] it was demonstrated for the vector valued case that

\[ E_u \left[ \frac{u^H O u}{u^H u} \right] = \frac{E_u [u^H O u]}{E_u [u^H u]} \]  

(10.63)

We conjecture that the expected value separation in (10.63) also holds for the matrix valued case

\[ E_{\hat{H}_w} \left[ \frac{\hat{H}_w^H O \hat{H}_w}{|\hat{H}_w^H \hat{H}_w|} \right] = \frac{E_{\hat{H}_w} \left[ \hat{H}_w^H O \hat{H}_w \right]}{E_{\hat{H}_w} \left[ \hat{H}_w^H \hat{H}_w \right]} \]  

(10.64)

The conjecture can be verified by Monte-Carlo simulations. Moreover, in [44] it is shown that the distribution of a determinant of a complex Wishart matrix equals to the distribution of the product of independent chi-squared distributed variables. We therefore have an obvious analogy to the vector-valued case in (10.63). An exact proof of the conjecture is possible via a Gaussian integral approach also used in Lemma 5.1, which is omitted here for brevity. Finally, using expected value (11.34) and noting that

\[ E_{\hat{H}_w} \left[ \hat{H}_w^H \hat{H}_w \right] = \left( \frac{R}{L-1} \right) \cdot \Gamma(L) \]  

(10.65)

yields (5.40). QED.
10.2 Statistical Transmit Prefilters for ZF Receivers

10.2.1 Proof of Theorem 5.9

We use majorization theory for solving problem (5.48). To this end, we first introduce the auxiliary \( L \times L \) matrix

\[
X = (F^H R_{TX} F)^{-1}
\]  
(10.66)

with the vector of diagonal elements (that are functions of the prefilter \( F \))

\[
x = \text{diag}(X) = \begin{bmatrix} x_1(F) & x_2(F) & \cdots & x_L(F) \end{bmatrix}.
\]  
(10.67)

Without loss of generality, assume that the elements of \( x \) are arranged in decreasing order. Now we can reformulate the problem in (5.48) as

\[
F_{opt} = \arg \min_{F} \sum_{k=1}^{L} [x_k(F)]^N \text{ s.t. } \text{tr}(FF^H) = \rho .
\]  
(10.68)

Note that \( F^H R_{TX} F \) is a positive definite matrix, i.e. all minors are positive definite and thus the diagonal elements of the inverse are positive [74]. Now consider the objective function in (10.68). Obviously, it is symmetric in its arguments. Furthermore, it is convex (it is convex in each of its arguments). It follows from Theorem 11.2 that it is a Schur-convex function. Using Lemma 11.2, the objective function is minimized for equal elements in \( x \). From Lemma 11.3, we can find a real symmetric matrix \( Q \) such that \( Q^H X Q \) has identical diagonal elements, i.e. we use prefilter matrix

\[
F = \tilde{F} Q .
\]  
(10.69)

Note that this does not change the transmit power. Another straightforward choice for the matrix \( Q \) is a DFT matrix of size \( L \times L \). Using the transmit filter structure in (10.69), we find

\[
[(F^H R_{TX} F)^{-1}]_{kk} = \frac{1}{L} \text{tr}((F^H R_{TX} F)^{-1}).
\]  
(10.70)

Obviously, with (10.70) the optimization problem in (10.68) can be reduced to (minimum trace minimizes each diagonal element)

\[
\tilde{F}_{opt} = \arg \min_{\tilde{F}} \text{tr}((\tilde{F}^H R_{TX} \tilde{F})^{-1}) \text{ s.t. } \text{tr}(\tilde{F} \tilde{F}^H) = \rho .
\]  
(10.71)

Now (10.71) is a Schur-concave function of the diagonal elements of \( (\tilde{F}^H R_{TX} \tilde{F})^{-1} \). Applying Lemma 11.1 we can choose \( \tilde{F} \) by proper application of a rotation matrix such that \( (\tilde{F}^H R_{TX} \tilde{F})^{-1} \) is diagonal with elements in decreasing order or equivalently such that the diagonal elements of \( \tilde{F}^H R_{TX} \tilde{F} \) are in increasing order. The diagonalizing rotation matrix consists of \( L \) eigenvectors of \( R_{TX} \). It can then be shown by Lemma 11.5 that we have to chose the eigenvectors corresponding
to the largest $L$ eigenvalues of $R_{TX}$, such that the optimum $F$ has the structure in (5.49). Constrained Lagrange optimization of the problem in (10.71) then leads to the power allocation matrix in (5.51). QED.

10.3 MIMO MMSE Receiver SER Calculation

10.3.1 Proof of Theorem 6.1

The receiver MMSE matrix filter $G$ (Fig. 6.1) is given by the well known expression [89] (with (6.5))

$$G = R_{ss}K^H(KR_{ss}K^H + R_{nn})^{-1}.$$  \hfill (10.72)

Via the matrix inversion lemma (11.9) it is straightforward to show that (10.72) is equivalent to

$$G = (R_{ss}^{-1} + K^HR_{nn}^{-1}K)^{-1}K^HR_{nn}^{-1}.$$  \hfill (10.73)

On the other hand, for the given MMSE receive filter $G$, it can be shown that the error covariance matrix reads [89]

$$R_{e e} = E[(s - Gy) \cdot (s - Gy)^H] = (R_{ss}^{-1} + K^HR_{nn}^{-1}K)^{-1}.$$  \hfill (10.74)

The diagonal elements of the covariance in (10.74) determine the mean squared error (MSE) on the subchannels, which is the mean power of cross-interference plus noise. Then denote the signal at the output of the receive filter by $z$ (Fig. 6.1) with

$$z = Gy.$$  \hfill (10.75)

For the following analysis, we split the vector $z$ in signal part $z_s$, interference (from other subchannels) part $z_i$, and noise part $z_n$

$$z = z_s + z_i + z_n.$$  \hfill (10.76)

We emphasize at this point the difference compared to the ZF receiver, where the interference term is completely suppressed. Then we consider subchannel 1, whereas $K$ is decomposed according to (6.5) $K = [k_1 \bar{K}]$ and the receive matrix filter is split with row vector $g_1^H$ into

$$G = \begin{bmatrix} g_1^H \\ \bar{G}^H \end{bmatrix}.$$  \hfill (10.77)

Later we generalize the results for an arbitrary subchannel. The signal part for subchannel 1 after receive filter $G$ is given by

$$z_{s, 1} = g_1^Hk_1s_1,$$  \hfill (10.78)
where \( s_1 \) is the transmit symbol on subchannel 1 and the transmit symbol vector can be partitioned into \( s = [s_1 \; \tilde{s}^T]^T \). Cross-interference due to other subchannels is given by

\[
\begin{align*}
    z_{i, 1} &= g_1^H \tilde{K}_1 \tilde{s}, \quad (10.79)
\end{align*}
\]

and finally the noise part reads

\[
\begin{align*}
    z_{n, 1} &= g_1^H n. \quad (10.80)
\end{align*}
\]

The subchannel signal to interference plus noise ratio (SINR) on subchannel 1 is again denoted by \( \gamma_{SC, 1} \) and we find for the MMSE receiver

\[
\begin{align*}
    \gamma_{SC, 1} &= \frac{E[|z_{s, 1}|^2]}{E[|z_{n, 1}|^2 + |z_{i, 1}|^2]} \quad (10.81)
\end{align*}
\]

For deriving an explicit expression of the SINR in (10.81), we first calculate a concise expression for \( g_1^H \) from (10.72) with the white covariance assumption in (6.2)

\[
\begin{align*}
    g_1^H &= E_s \cdot k_1^H (E_s \cdot KK^H + N_0 \cdot I)^{-1} = \gamma \cdot k_1^H (\gamma \cdot KK^H + I)^{-1}. \quad (10.82)
\end{align*}
\]

We can reformulate (10.82) by decomposing

\[
\begin{align*}
    g_1^H &= \gamma \cdot k_1^H (\gamma \cdot \tilde{K}_1 \tilde{k}_1^H + \gamma \cdot k_1^H + I)^{-1} \quad (10.83)
\end{align*}
\]

and with the help of matrix inversion lemma (11.8) we get after defining \( Q_1 \) according to (6.4) the result

\[
\begin{align*}
    g_1^H &= \gamma \cdot k_1^H Q_1^{-1} \frac{\gamma^2 \cdot k_1^H Q_1^{-1} k_1^H Q_1^{-1} \cdot 1 + \gamma k_1^H Q_1^{-1} k_1}{1 + \gamma k_1^H Q_1^{-1} k_1} \quad (10.84)
\end{align*}
\]

This simplifies to

\[
\begin{align*}
    g_1^H &= \gamma \cdot k_1^H Q_1^{-1} \cdot \frac{1}{1 + \gamma k_1^H Q_1^{-1} k_1} \quad (10.85)
\end{align*}
\]

Using (10.85), we can directly find the nominator of (10.81) with (10.78)

\[
\begin{align*}
    E[|z_{s, 1}|^2] &= E_s \cdot \frac{\gamma^2 \cdot (k_1^H Q_1^{-1} k_1)^2}{(1 + \gamma k_1^H Q_1^{-1} k_1)^2} \quad (10.86)
\end{align*}
\]

Equivalently, we have for the denominator of (10.81) with (10.79)(10.80)

\[
\begin{align*}
    E[|z_{n, 1}|^2 + |z_{i, 1}|^2] &= g_1^H Q_1 \cdot g_1 = N_0 \cdot \frac{\gamma^2}{(1 + \gamma k_1^H Q_1^{-1} k_1)^2} \cdot k_1^H Q_1^{-1} k_1 \quad (10.87)
\end{align*}
\]

Summarizing the results, we get (6.3) from (10.81) with (10.86)(10.87) after generalization to an arbitrary subchannel. On the other hand, with a similar proceeding as above we can derive from
the expression for the MSE on subchannel \( k \) in (6.6). The relation between the SINR and the MSE on subchannel \( k \) in (6.7) directly follows from (6.3) and (6.6). \( QED. \)

### 10.3.2 Proof of Theorem 6.3

We still assume eigenmode transmission, such that we have from (6.12) the MGF expression for the SINR on subchannel \( k \)

\[
M_{\gamma}(s, k) = E_{\tilde{y}, \tilde{Y}} \left[ \exp \left( -s c_k \cdot y^H (\tilde{Y}_k \tilde{Y}_k^H + \frac{1}{\gamma} I_R)^{-1} y \right) \right].
\]

With the help of the expected value formula for complex normal vectors (11.37) and the distribution of \( y \) and \( \tilde{Y} \) in (6.13) and (6.14), respectively, we find

\[
M_{\gamma}(s, k) = E_{\tilde{y}} \left[ \frac{1}{I_R + s c_k \cdot \left( \tilde{Y}_k \tilde{Y}_k^H + \frac{1}{\gamma} I_R \right)^{-1}} \right] = E_{\tilde{h}} \left[ \frac{1}{I_R + s c_k \cdot \left( \tilde{h}_w \tilde{c}_k \tilde{h}_w^H + \frac{1}{\gamma} O^{-1} \right)^{-1}} \right]
\]

with \( R \times (L - 1) \) matrix of i.i.d. complex Gaussian entries \( \tilde{h}_w \). Starting from (10.89), we find after obvious multiplication of nominator and denominator

\[
M_{\gamma}(s, k) = E_{\tilde{h}} \left[ \frac{\tilde{h}_w \tilde{c}_k \tilde{h}_w^H + \frac{1}{\gamma} O^{-1}}{\tilde{h}_w \tilde{c}_k \tilde{h}_w^H + \frac{1}{\gamma} O^{-1} + s c_k \cdot I_R} \right]
\]

Factoring out from nominator and denominator we find

\[
M_{\gamma}(s, k) = E_{\tilde{h}} \left[ \frac{\frac{1}{\gamma} O^{-1}}{\tilde{h}_w \tilde{c}_k \tilde{h}_w^H + s c_k \cdot I_R} \right]
\]

and with the help of the determinant identity (11.14) we get after simplifying the expression (6.18). \( QED. \)

### 10.3.3 Proof of Lemma 6.3

We use integral identity (11.43) and get with \( n \times 1 \) complex vector \( x \) from (6.26)

\[
\tilde{r}_1 = E_{x} \left[ I + G^H \tilde{A} G \right] \cdot \left[ e^{-x^H \tilde{G}_B \tilde{B} \tilde{G}_x} e^{-x^H D \tilde{G}_x} \right]
\]

Using (11.42) for separating \( G \) and \( G^H \) (this is necessary for integrating out \( G \) later), we can express in (10.92)
\[ e^{-x^H G^H \tilde{B} G x} = (-1)^n \int e^{y^H y} e^{-\text{tr}(C^H G + G^H C)} D_y D_x. \] (10.93)

with the auxiliary matrix
\[ C = \tilde{B}^{1/2} y x^H. \] (10.94)

Plugging (10.93) in (10.92), we find
\[ \tilde{r}_1 = (-1)^n \int E_y \left[ |I + \tilde{G}^H \tilde{A} \tilde{G}| \cdot e^{-\text{tr}(C^H G + G^H C)} \right] e^{y^H y} e^{-x^H D_x y D_x} D_x. \] (10.95)

Now we focus on the expected value with respect to the Gaussian distribution of \( G \) in (10.95), such that we have to calculate
\[ g = E_y \left[ |I + \tilde{G}^H \tilde{A} \tilde{G}| \cdot e^{-\text{tr}(C^H G + G^H C)} \right] = \int |I + \tilde{G}^H \tilde{A} \tilde{G}| \cdot e^{-\text{tr}(G^H G + C^H G + G^H C)} D_x G. \] (10.96)

Completing the square in the exponential, we find
\[ g = \int |I + \tilde{G}^H \tilde{A} \tilde{G}| \cdot e^{-\text{tr}(G + C)^H (G + C)} \cdot e^{\text{tr}(C^H C)} D_x G. \] (10.97)

The determinant in (10.97) can be split into a sum of determinants via determinant identity (11.17)
\[ |I + \tilde{G}^H \tilde{A} \tilde{G}| = \sum_{k=0}^{n} \tilde{G}^H \tilde{A} \tilde{G} |_{\hat{\alpha}_k \hat{\beta}_k}. \] (10.98)

Splitting the determinant yields with \( \hat{\alpha}_k \subseteq \{1, \ldots, m\} \) and \( \hat{\beta}_k \subseteq \{1, \ldots, n\} \)
\[ |I + \tilde{G}^H \tilde{A} \tilde{G}| = \sum_{k=0}^{n} \sum_{\hat{\alpha}_k} \tilde{G}^H \tilde{A} |_{\hat{\alpha}_k \hat{\beta}_k} \cdot |G|_{\hat{\alpha}_k \hat{\beta}_k}. \] (10.99)

Now we define a complementary index subset
\[ \hat{\alpha}_{m-k} = \{1, \ldots, m\} \backslash \hat{\alpha}_k \] (10.100)
as well as
\[ \hat{\beta}_{n-k} = \{1, \ldots, n\} \backslash \hat{\beta}_k. \] (10.101)

Then partition \( G \) into 4 blocks with short-hand notation
and equivalently $C$

$\begin{pmatrix} \{G\}_{\beta_k}^{\tilde{\alpha}_k} & \{G\}_{\beta_{k-1}}^{\tilde{\alpha}_{k-1}} \\ \{C\}_{\beta_k}^{\tilde{\alpha}_k} & \{C\}_{\beta_{k-1}}^{\tilde{\alpha}_{k-1}} \end{pmatrix} = \begin{pmatrix} G_{1,1} & G_{1,2} \\ G_{2,1} & G_{2,2} \end{pmatrix}$ \hspace{1cm} (10.102)

Now (10.99) can be written as

$g = \int \sum_{k=0}^{n} \sum_{\tilde{\alpha}_k} \sum_{\tilde{\beta}_k} |\hat{A}|_{\tilde{\alpha}_k}^{\tilde{\beta}_k} \cdot \left| G^H_{1,1} G_{1,1} \right| \cdot e^{-\text{tr}(G+C)\gamma(G+C)} \cdot e^{\text{tr}(C^H C)} D_C G$ \hspace{1cm} (10.104)

and we get

$g = \sum_{k=0}^{n} \sum_{\tilde{\alpha}_k} \sum_{\tilde{\beta}_k} (|\hat{A}|_{\tilde{\alpha}_k}^{\tilde{\beta}_k}, I_{\alpha, \beta, k}) \cdot e^{\text{tr}(C^H C)}$ \hspace{1cm} (10.105)

with the prototype integral

$I_{\alpha, \beta, k} = \int |G^H_{1,1} G_{1,1}| \cdot e^{-\text{tr}(G+C)\gamma(G+C)} D_C G.$ \hspace{1cm} (10.106)

By straightforward manipulations of the exponent we can rewrite (10.106) as

$I_{\alpha, \beta, k} = \left\{ \begin{array}{c} \int |G^H_{1,1} G_{1,1}| \cdot e^{-\text{tr}(G_{1,1}+C_{1,1})\gamma(G_{1,1}+C_{1,1})} D_C G_{1,1} \cdot \ldots \\ \int e^{-\text{tr}(G_{2,2}+C_{2,2})\gamma(G_{2,2}+C_{2,2})} D_C G_{2,2} \cdot \ldots \\ \int e^{-\text{tr}(G_{1,2}+C_{1,2})\gamma(G_{1,2}+C_{1,2})} D_C G_{1,2} \cdot \ldots \\ \int e^{-\text{tr}(G_{2,1}+C_{2,1})\gamma(G_{2,1}+C_{2,1})} D_C G_{2,1} \end{array} \right.$ \hspace{1cm} (10.107)

Now we can use integral identity (11.52) and find

$\int |G^H_{1,1} G_{1,1}| \cdot e^{-\text{tr}(G_{1,1}+C_{1,1})\gamma(G_{1,1}+C_{1,1})} D_C G_{1,1} = [\Gamma(k + 1) + \Gamma(k) \cdot \text{tr}(C_{1,1}^H C_{1,1})].$ \hspace{1cm} (10.108)

On the other hand it follows for $u \in \{1, 2\}$ and $v \in \{1, 2\}$

$\int e^{-\text{tr}(G_{u,v}+C_{u,v})\gamma(G_{u,v}+C_{u,v})} D_C G_{u,v} = 1.$ \hspace{1cm} (10.109)
Summarizing, we have

\[ I_{\alpha, \beta, k} = [\Gamma(k + 1) + \Gamma(k) \cdot \text{tr}(C_{1, 1} C_{1, 1}^H)]. \tag{10.110} \]

After partitioning the diagonal \( m \times m \) matrix \( \tilde{B} \) with short-hand notation

\[
\begin{bmatrix}
\{ \tilde{B} \}_{\hat{G}_k} & 0 \\
0 & \{ \tilde{B} \}_{\hat{G}_{m-k}}
\end{bmatrix}
= \begin{bmatrix}
\tilde{B}_1 & 0 \\
0 & \tilde{B}_2
\end{bmatrix}
\tag{10.111}
\]

and the \( m \times 1 \) vector \( y \)

\[
\begin{bmatrix}
\{ y \}_{\hat{G}_k} \\
\{ y \}_{\hat{G}_{m-k}}
\end{bmatrix}
= \begin{bmatrix}
y_1 \\
y_2
\end{bmatrix}
\tag{10.112}
\]

and the \( n \times 1 \) vector \( x \)

\[
\begin{bmatrix}
\{ x \}_{\hat{G}_k} \\
\{ x \}_{\hat{G}_{m-k}}
\end{bmatrix}
= \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
\tag{10.113}
\]

we get from (10.94)

\[
\begin{align*}
C_{1, 1} &= \tilde{B}_1^{1/2} y_1 x_1^H \\
C_{1, 2} &= \tilde{B}_1^{1/2} y_1 x_2^H \\
C_{2, 1} &= \tilde{B}_2^{1/2} y_2 x_1^H \\
C_{2, 2} &= \tilde{B}_2^{1/2} y_2 x_2^H
\end{align*}
\tag{10.114}
\]

We find with (10.114) and (10.110) in (10.105)

\[ \tilde{r}_1 = (-1)^m \sum_{k=0}^{n} \sum_{\tilde{G}_k} \sum_{\tilde{B}_k} \langle \hat{G}_k, \tilde{r}_{\alpha, \beta, k} \rangle, \tag{10.115} \]

where

\[ \tilde{r}_{\alpha, \beta, k} = \Gamma(k) \int \left[ k + \text{tr}(y^H (x^H \tilde{B}_1 y_1 x_1^H)) \cdot e^{i \text{tr}(y^H (x^H \tilde{B}_1 y_1 x_1^H))} \cdot e^{i \text{tr}(y^H (x^H \tilde{B}_1 y_1 x_1^H))} \cdot e^{-x^H D_1 x} D_1 y \right]. \tag{10.116} \]

Now integrate with respect to the vector \( y_1 \) with the help of (11.41) (using the abbreviation \( b_{j(l)} = [\tilde{B}_1]_{ll} = \{ [\tilde{B}]_{\hat{G}_k} \}_{ll} \) )

\[ \hat{r}_{\alpha, \beta, k} = \frac{\Gamma(k)}{(-1)^k} \int k \frac{x^H x_1 \cdot b_{j(l)}}{I + x^H x \tilde{B}_1} \sum_{l=1}^{k} \frac{x^H x_1 \cdot b_{j(l)}}{1 + x^H x \cdot b_{j(l)}} \cdot e^{i \text{tr}(y^H (x^H \tilde{B}_1 y_1 x_1^H))} \cdot e^{-x^H D_1 x} D_1 y_2. \tag{10.117} \]

Then integrating with respect to \( y_2 \) we find after simplifying
\[ \hat{\alpha}_{k} = \Gamma(k) \cdot (-1)^m \cdot \int \left[ \frac{k}{|I + x^H x|} - \sum_{l=1}^{k} \frac{x_i^H x_j \cdot b_{j(l)}}{(1 + x^H x \cdot b_{j(l)})|I + x^H x|} \right] e^{-x_m^H x} \, dx. \] (10.118)

This can be written with the help of (11.44) as

\[ \hat{\alpha}_{k} = \Gamma(k + 1) \cdot (-1)^m \cdot \int \left[ \frac{1}{\Gamma(n)} \int_0^{\infty} e^{-t} \frac{t^{n-1}}{|I + t B|} \right] dt - \frac{1}{\Gamma(n + 1)} \int_0^{\infty} \left[ \sum_{l=1}^{k} \frac{b_{j(l)}}{(1 + t \cdot b_{j(l)})|I + t B|} \right] t^n e^{-t} \, dt \] (10.119)

After rewriting (10.119) as

\[ \hat{\alpha}_{k} = \frac{\Gamma(k + 1)}{\Gamma(n + 1)} \cdot (-1)^m \cdot \left[ (n-k) \int_0^{\infty} \frac{t^{n-1} e^{-t}}{|I + t B|} \, dt + \left( \sum_{l=1}^{k} \int_0^{\infty} \frac{t^{n-1} e^{-t}}{|I + t B|} \, dt \right) - \right] \]

\[ \left[ \int_0^{\infty} \sum_{l=1}^{k} \frac{b_{j(l)}}{(1 + t \cdot b_{j(l)})|I + t B|} t^n e^{-t} \, dt \right] \] (10.120)

we get with some simplifications

\[ \hat{\alpha}_{k} = \frac{\Gamma(k + 1)}{\Gamma(n + 1)} \cdot (-1)^m \cdot \left[ (n-k) \int_0^{\infty} \frac{t^{n-1} e^{-t}}{|I + t B|} \, dt + \left( \sum_{l=1}^{k} \int_0^{\infty} \frac{t^{n-1} e^{-t}}{|I + t B|} \, dt \right) - \right] \]

\[ \left[ \int_0^{\infty} \sum_{l=1}^{k} \frac{b_{j(l)}}{(1 + t \cdot b_{j(l)})|I + t B|} t^n e^{-t} \, dt \right] \] (10.121)

However, \( \hat{\alpha}_{k} \) is invariant to \( \hat{\beta}_k \), and after noting that there are \( \binom{n}{k} \) different subsets \( \hat{\beta}_k \) we arrive at (6.26). \textit{QED.}

\subsection*{10.3.4 Proof of Theorem 6.7}

After exchanging the sequence of integrals, from (5.28) and (6.22) we get

\[ \bar{P}_{s, k} = \frac{2b}{\pi} \cdot \frac{(L-1) \cdot \text{tr}_{L-1}(O)}{|O|} \cdot \left[ \int_0^{\pi/2} \int_0^{\infty} \frac{1}{\left( \frac{c}{\sin^2 \theta} \right)^R + 1} \, d\theta \right] \cdot \int_0^{\infty} t^{L-2} e^{-t} \, dt \cdot \gamma^N. \] (10.122)

By exploiting the integral [172]

\[ \int_0^{\pi/2} \left( \frac{\sin^2 \theta}{\kappa + \sin^2 \theta} \right)^R \, d\theta = \frac{\pi}{2} \left[ 1 - \frac{\kappa}{1 + \kappa} \cdot \sum_{k=0}^{R-1} \binom{2k}{k} \left( \frac{1}{4(1 + \kappa)} \right)^k \right] \] (10.123)

it can be shown that
Using (10.124) in (10.122) yields the theorem after making use of the integral representation of the Kummer $U$ function (11.88) and simplifying the result. $QED$.

### 10.3.5 Proof of Theorem 6.8

Plugging (6.24) in (5.28) yields after reformulation

$$
\frac{\pi}{2} \int_0^\infty \left( \frac{1}{c} \right)^R \left( \frac{1}{1 + \frac{\sin^2 \theta}{c}} \right)^{R+1} d\theta = \frac{\pi \sqrt{\pi}}{2 \cdot 4^R} \left( \frac{\gamma_{MC}}{R} \right) \cdot \left( t + c \right)^{\left( R + \frac{1}{2} \right)}.
$$

(10.124)

Now we can make use of [172]

$$
\frac{\pi}{2} \int \frac{\sin^2 \theta}{c + \sin^2 \theta} d\theta = (-1)^m \cdot \frac{c^{m-1}}{2} \left[ \sqrt{\frac{1 + c}{c}} - \sum_{i=0}^{m-1} (-1)^i \cdot \frac{1}{2^{2i} \cdot c^i} \right].
$$

(10.126)

and after simplifying we can establish the theorem from (10.125) with (10.126) and the integration results

$$
\int_0^\infty \frac{1}{\sqrt{c + t}} e^{-t} dt = e^c \cdot \sqrt{\pi} \cdot \text{erfc}(\sqrt{c})
$$

(10.127)

and

$$
\int_0^\infty t^{n-1} e^{-t} dt = \Gamma(n).
$$

(10.128)

$QED$.

### 10.3.6 Proof of Theorem 6.9

Without receive correlation, the SINR expression in (6.12) reduces to

$$
\gamma_{SC,k} = \mathbf{u}^H \left( \mathbf{H}_w \mathbf{C} \mathbf{H}_w^H + \frac{1}{\gamma} \mathbf{I} \right)^{-1} \mathbf{u},
$$

(10.129)
where \( \mathbf{u} \) is a \( R \times 1 \) vector and \( \tilde{H}_w \) is a \( R \times (L - 1) \) matrix of i.i.d. complex Gaussian random variables. We now bring the SINR expression into a form that was analyzed in [39]. To this end, we simply rewrite

\[
\gamma_{SC,k} \equiv \mathbf{u}^H (\tilde{H}_w (\frac{1}{c_k} \mathbf{c}) \tilde{H}_w^H + \frac{1}{c_{k}^{\mathcal{Y}}} \mathbf{I})^{-1} \mathbf{u} = \mathbf{u}^H (\tilde{H}_w \Theta \tilde{H}_w^H + \frac{1}{c_{k}^{\mathcal{Y}}} \mathbf{I})^{-1} \mathbf{u}
\]  

(10.130)

with \( \Theta \) defined in (6.34). The complementary cumulative distribution function CCDF (reliability function) \( R(\gamma_{SC,k}) \) of the random variable \( \gamma_{SC,k} \) was given in [39] (corresponding to the case of multiple uncorrelated users in a smart antenna system with different transmit powers \( \theta_i \))

\[
R(\gamma_{SC,k}) = e^{-\frac{\gamma_{SC,k}}{c_{k}^{\mathcal{Y}}}} \left( \sum_{i=1}^{R} \beta_i^{-1} \cdot \gamma_{SC,k}^{-i-1} \right) \left( \prod_{j=1}^{L-1} (1 + \theta_j \gamma_{SC,k}) \right),
\]  

(10.131)

where the \( \beta_i \) are the coefficients of \( z^i \) in the series expansion of (6.35). Note that for a random variable \( u \) the CCDF (reliability function) is defined by \( R(u) = Pr(u > u_0) \), and for \( \gamma_{SC,k} \) obviously \( R(0) = 1 \) and \( R(\infty) = 0 \). Applying integration by parts and noting for the conditional symbol error probability in (5.16) that \( P_{s,c}(\infty) = 0 \) and \( P_{s,c}(0) = b \), one can derive in general for the average SER on subchannel \( k \)

\[
P_{s,k}(\gamma) = b + \int_{0}^{\infty} \left( \frac{d}{dy_{SC,k}} P_{s,c}(\gamma_{SC,k}) \right) R(\gamma_{SC,k}) dy_{SC,k}.
\]  

(10.132)

For the conditional symbol error rate formula for square QAM constellations (5.16) we find the derivative

\[
\frac{d}{dy_{SC,k}} P_{s,c}(\gamma_{SC,k}) = -b \sqrt{\frac{\gamma_{SC,k}^{1/2}}{\pi}} \cdot e^{-c \gamma_{SC,k}}.
\]  

(10.133)

For evaluating the integral (10.132) we make use of the expansion into partial fractions

\[
\prod_{n=1}^{N} \frac{1}{1 + \theta_n z} = \sum_{n=1}^{N} \left( \prod_{j=1}^{N} \frac{\theta_n}{\theta_n - \theta_j} \right) \frac{1}{1 + \theta_n z}
\]  

(10.134)

and the integration result [50, 3.383.10], yielding for constants \( a, b, \) and \( c \)

\[
\int_{0}^{\infty} \frac{z^{a-3/2}}{1 + b z} \cdot e^{-c z} \, dz = \Gamma\left(a - \frac{1}{2}\right) \cdot \left(\frac{1}{2} - a\right) \cdot e^{\mathcal{Y} b} \cdot \Gamma\left(\frac{3}{2} - a, \frac{c}{b}\right).
\]  

(10.135)

\( \Gamma(z) \) and \( \Gamma(a, z) \) denote the Gamma and incomplete Gamma function (cf. [1]), respectively. QED.
10.4 MIMO ML Receiver SER Calculation

10.4.1 Proof of Theorem 7.1

We are using a MGF approach similar to the proceeding in [206] for the evaluation of the PEP. First, we rewrite (7.7) as

\[ \eta(v | \mu) \equiv \sum_{r=1}^{R} |\hat{y}_r - \tilde{g}_r^H x_v|^2, \]  

(10.136)

with the auxiliary vectors

\[ \tilde{g}_r = \sigma_r^{1/2} g_r, \quad x_v = \Lambda_{1/2}^{-1} \Phi s_v. \]  

(10.137)

The basic structure of (10.136) and [206, equation (3)] agrees and we can derive the MGF along analogue lines as in case of uncorrelated fading. The key for finding the MGF is to rewrite (10.136) as a sum of random quadratic forms in normal variables. It can be shown by straightforward expansion that for scalar \( a \), vectors \( b \) and \( c \), we get

\[ |a - b^H c|^2 = z^H Q z, \]  

(10.138)

with the auxiliary definitions

\[ z = \begin{bmatrix} a \\ b^* \end{bmatrix}, \quad d = \begin{bmatrix} 1 \\ -c \end{bmatrix}, \quad Q = d^* d^T = \begin{bmatrix} 1 & -c^T \\ -c^* & c^* c^T \end{bmatrix}. \]  

(10.139)

Application of (10.138)(10.139) to (10.136) leads to

\[ \eta(v | \mu) \equiv \sum_{r=1}^{R} z_r^H \tilde{Q} z_r, \]  

(10.140)

with complex Gaussian random \((L + 1) \times 1\) column vectors \( z_r \) and \((L + 1) \times (L + 1)\) deterministic matrix \( \tilde{Q} \)

\[ z_r = \begin{bmatrix} \hat{y}_r \\ \tilde{g}_r^* \end{bmatrix}, \quad \tilde{Q} = \begin{bmatrix} 1 & -x_v^T \\ -x_v^* x_v^* x_v^T \end{bmatrix}. \]  

(10.141)

The difference of the distance metrics (7.9) is then given by a sum of independent random quadratic forms

\[ D_k(i, j, q) = \eta_k(j | i, q) - \eta_k(i | i, q) = \sum_{r=1}^{R} z_r^H P z_r \]  

(10.142)
with auxiliary matrix \( x_i = \tilde{\Lambda}_{TX}^{1/2} \Phi s_i \) and \( x_j = \tilde{\Lambda}_{TX}^{1/2} \Phi s_j \) from (10.137)

\[
P = \begin{bmatrix}
0 & -(x_j^T - x_i^T) \\
-(x_j^* - x_i^*) & x_j^* x_j^T - x_i^* x_i^T
\end{bmatrix}. \tag{10.143}
\]

Omitting details, the covariance matrices \( R_r \) of the complex normal random vectors \( z_r \) are given by

\[
R_r = E[z_r z_r^H] = o_r \cdot \begin{bmatrix}
\|x_i\|^2 + \frac{N_0}{o_r} x_i^T \\
x_i^* & I
\end{bmatrix}. \tag{10.144}
\]

Due to the independence of the quadratic forms in the expression for \( D_k(i, j, q) \) in (10.142) the moment generating function can be calculated via (11.37) in product form, and we obtain

\[
M_{D_k}(s) = E[e^{-sD_k}] = \prod_{r=1}^{R} \frac{1}{1 + sR_r P} = \prod_{r=1}^{R} \frac{1}{(1 + \lambda_{r,1} s)(1 + \lambda_{r,2} s)}. \tag{10.145}
\]

According to (10.143), \( P \) is a matrix of rank 2 and consequently the matrix \( R_r P \) has only 2 non-zero eigenvalues \( \lambda_{r,1/2} \), which are used in (10.145). It can be shown that the eigenvalues are given by (basically we have to solve a quadratic equation in the eigenvalues)

\[
\lambda_{r,1/2} = \frac{d_r \pm \sqrt{d_r^2 + 4N_0 d_r}}{2}, \tag{10.146}
\]

where we have introduced the scaled transmit symbol vector distance that reads with (10.137)

\[
d_r = o_r \cdot \|x_i - x_j\|^2 = o_r \cdot \|\tilde{\Lambda}_{TX}^{1/2} \Phi (s_i - s_j)\|^2. \tag{10.147}
\]

Introducing the normalized transmit vectors (7.18) and

\[
d_r = E_s \cdot o_r \cdot \|\tilde{\Lambda}_{TX}^{1/2} \Phi (\tilde{s}_i - \tilde{s}_j)\|^2 = E_s \cdot \tilde{d}_r, \tag{10.148}
\]

we can rewrite (10.146) as is given in (7.16) and (7.17). \textit{QED.}

### 10.4.2 Proof of Theorem 7.3

In order to be able to perform the integration in (7.12) for finding the pairwise error probability, we first have to find the PDF of the metric difference \( D_k(i, j, q) \). To this end, note that with (10.145)
\[ M_{D_k}(s) = E[e^{-sD_k}] = \int_{-\infty}^{\infty} e^{-sD_k} \cdot p(D_k)dD_k = \prod_{r=1}^{R} \frac{1}{(1 + \lambda_{r,1}s)(1 + \lambda_{r,2}s)}, \quad (10.149) \]

where \( p(D_k) \) is the PDF of \( D_k \). However, (10.149) defines the Laplace transform of the PDF \( p(D_k) \) and we can thus obtain \( p(D_k) \) via an inverse Laplace transform. To this end, we first rewrite (10.149) using definition (7.23) as

\[ M_{D_k}(s) = \prod_{k=1}^{2R} \frac{1}{1 + \lambda_k s}. \quad (10.150) \]

Note again that in (7.23) the positive eigenvalues are arranged first. Now we apply a partial fractional expansion and write (10.150) as

\[ M_{D_k}(s) = \prod_{k=1}^{2R} \frac{1}{1 + \lambda_k s} = \sum_{i=1}^{2R} \left( \prod_{i \neq k}^{\lambda_k} \frac{\lambda_k}{\lambda_k - \lambda_i} \right) \cdot \frac{1}{1 + \lambda_k s}. \quad (10.151) \]

This is the MGF of a sum of weighted independently exponentially distributed variables. It is straightforward to derive the double sided Laplace transform pairs

\[ \frac{1}{b} \cdot e^{-x/b} \cdot \sigma(x) \rightarrow \frac{1}{1 + sx}, \quad b > 0 \land x \geq 0 \]

\[ \frac{1}{|b|} \cdot e^{-x/|b|} \cdot \sigma(-x) \rightarrow \frac{1}{1 + sx}, \quad b < 0 \land x \leq 0 \quad (10.152) \]

where \( \sigma(x) \) is the unit step function. Introducing the short-hand notation

\[ \xi_k = \prod_{i=1}^{2R} \frac{\lambda_k}{\lambda_k - \lambda_i} \quad (10.153) \]

we can find the PDF \( p(D_k) \) of \( D_k(i,j,q) \) from (10.151) and (10.152)

\[ p(D_k) = \sum_{k=1}^{2R} \xi_k \cdot e^{-\lambda_k \cdot \lambda_k} \cdot \sigma(-\lambda_k \cdot D_k) \quad (10.154) \]

with signum function \( \sigma(x) \). Now we can directly calculate (note that the following sum in (10.155) is only over the positive eigenvalues \( \lambda_k \) with \( k = 1 \cdots R \))

\[ P_k(i,j,q) = \int_{-\infty}^{\infty} p(D_k)dD_k = 1 - \int_{0}^{\infty} p(D_k)dD_k = 1 - \int_{0}^{\infty} \sum_{k=1}^{R} \xi_k \cdot \frac{D_k}{\lambda_k} \cdot dD_k. \quad (10.155) \]
After carrying out the straightforward integration in (10.155), we finally arrive at (7.22). \textit{QED.}

### 10.4.3 Proof of Theorem 7.4

Application of a series expansion of the square root leads to the high SNR approximation of the eigenvalues in (7.16)

\[
\tilde{\lambda}_{r,1/2} = E_s \left\{ \begin{array}{l} \frac{-d_r + \frac{1}{\gamma}}{1/\gamma} \\
\frac{-1}{\gamma}
\end{array} \right. 
\]  

(10.156)

We find from PEP expression (7.22) and eigenvalue approximation (10.156) after straightforward manipulations

\[
P_k(i,j,q) \approx 1 - \sum_{r=1}^{R} \left[ \prod_{i=1}^{R} \frac{1}{d_r - d_i} \left( \frac{1}{d_r + 2/\gamma} \right)^R (d_r + 1/\gamma)^{2R-1} \right].
\]  

(10.157)

For obtaining the asymptotics of (10.157) for \(1/\gamma \approx 0\), we introduce the auxiliary function

\[
f(1/\gamma) = \sum_{r=1}^{R} \left[ \prod_{i=1}^{R} \frac{1}{d_r - d_i} \left( \frac{1}{d_r + 2/\gamma} \right)^R (d_r + 1/\gamma)^{2R-1} \right] (1/\gamma)^{R-1}.
\]  

(10.158)

with

\[
f(0) = \sum_{r=1}^{R} \left[ \prod_{i=1}^{R} \frac{1}{d_r - d_i} \right] (d_r)^{R-1}.
\]  

(10.159)

Now (10.157) can be expressed via a Taylor series expansion of \(f(1/\gamma)\) as

\[
P_k(i,j,q) \approx 1 - \sum_{l=0}^{\infty} \frac{1}{l!} \left. \frac{d^l}{d\left(\frac{1}{\gamma}\right)^l} \right|_{\frac{1}{\gamma} = 0} f\left(\frac{1}{\gamma}\right)^l.
\]  

(10.160)

For the Taylor series expansion of \(f(1/\gamma)\), first consider the auxiliary function

\[
g(x) = \left( \frac{1}{a + 2x} \right)^R (a + x)^{2R-1}.
\]  

(10.161)
By direct differentiation or series expansion for small $x$ it can be shown after some tedious math that

$$
\left. \frac{d^k g(x)}{dx^k} \right|_{x = 0} = K \cdot a^{R-k-1}, \quad (10.162)
$$

where $K$ is a constant independent of $a$. Specifically, in the special case $k = R$ we obtain

$$
\frac{d^R}{dx^R} g(x) = (-1)^R \cdot \frac{\Gamma(2R)}{\Gamma(R)} \cdot \frac{1}{a} = (-1)^R \cdot \frac{(2R-1)!}{(R-1)!} \cdot \frac{1}{a}. \quad (10.163)
$$

On the other hand, it is possible to find

$$
\sum_{r=1}^{R} \prod_{\substack{i=1 \atop i \neq r}}^{R} \frac{1}{d_r - d_i} = \begin{cases} 1 & k = 0 \\ 0 & 0 < k < R \\ (-1)^{R+1} \frac{R!}{\prod_{r=1}^{R} d_r} & k = R \end{cases} \quad (10.164)
$$

Applying the auxiliary results (10.162) and (10.163) for finding the Taylor series expansion of $f(1/\gamma)$ in (10.160) and taking into account summation result (10.164), it can be shown that the high SNR approximation of (10.157) is

$$
\tilde{P}_k(i, j, q) = \frac{1}{\prod_{r=1}^{R} \tilde{d}_r} \cdot \frac{(2R-1)!}{R! \cdot (R-1)!} \cdot \gamma^{-R} = \frac{1}{\prod_{r=1}^{R} d_r} \cdot \left( \frac{2R-1}{R-1} \right) \cdot \gamma^{-R}, \quad (10.165)
$$

which together with (7.16) results in (7.24). QED.
11 Appendix - Mathematical Preliminaries

In this appendix we present results on linear algebra, multivariate statistics, complex integrals, majorization theory, and special functions that are frequently used in this thesis. Most of the results are stated without proof, however, if they are not available in literature, we present a more detailed derivation.

11.1 Linear Algebra

A powerful source of results on linear matrix algebra is [125] and if not stated otherwise, all following formulas can be found in this reference.

11.1.1 DFT matrix

The normalized \( L \times L \) discrete Fourier transform (DFT) matrix \( D_L \) is defined by (\( i \) and \( k \) run from 1 to \( L \))

\[
D_L = \left[ \frac{1}{\sqrt{L}} e^{-\frac{2\pi i k}{L}} \right]
\]

with the property

\[
D_L D_L^H = D_L^H D_L = I_L.
\]

11.1.2 Kronecker product

For more details on the Kronecker product see [51].

\[
\text{vec}(ABC) = (C^T \otimes A) \cdot \text{vec}(B)
\]

\[
(A \otimes B) \cdot (C \otimes D) = AC \otimes BD
\]

\[
(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}
\]

\[
(A \otimes B)^H = A^H \otimes B^H
\]

11.1.3 Inversion lemmas

\[
(A + UV^H)^{-1} = A^{-1} - A^{-1}U(I + V^HA^{-1}U)^{-1}V^HA^{-1}
\]

\[
(P^{-1} + MHQ^{-1}M)^{-1} = P - PM^H(Q + MPM^H)^{-1}MP
\]

\[
(P^{-1} + MHQ^{-1}M)^{-1}M^HQ = PM^H(Q + MPM^H)^{-1}
\]
11.1.4 Partitioned matrices

Partition a quadratic matrix $A$ and its inverse $A^{-1}$ as
\[
A = \begin{bmatrix}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{bmatrix}, \quad A^{-1} = \begin{bmatrix}
A_{11}^{-1} & A_{12}^{-1} \\
A_{21}^{-1} & A_{22}^{-1}
\end{bmatrix}.
\tag{11.10}
\]

It is then well-known that
\[
A_{11}^{-1} = (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} \equiv (A_{11} \cdot 2)^{-1}.
\tag{11.11}
\]

11.1.5 Determinants

General

Let $a$ and $b$ be two $m \times 1$ column vectors and $C_1$ and $C_2$ be matrices of compatible size. Then it is possible to expand the determinant
\[
\begin{vmatrix}
C_1 
\begin{array}{c}
a + b \\
C_2
\end{array}
\end{vmatrix}
= \begin{vmatrix}
C_1 
\begin{array}{c}
a \\
C_2
\end{array}
\end{vmatrix} + \begin{vmatrix}
C_1 
\begin{array}{c}
b \\
C_2
\end{array}
\end{vmatrix}
\tag{11.12}
\]

with the obvious equivalent
\[
\begin{vmatrix}
C_1^T 
\begin{array}{c}
a^T + b^T \\
C_2^T
\end{array}
\end{vmatrix}
= \begin{vmatrix}
C_1^T 
\begin{array}{c}
a^T \\
C_2^T
\end{array}
\end{vmatrix} + \begin{vmatrix}
C_1^T 
\begin{array}{c}
b^T \\
C_2^T
\end{array}
\end{vmatrix}.
\tag{11.13}
\]

For arbitrary compatible matrices $A$ and $B$ it can be shown that
\[
|I + AB| = |I + BA|.
\tag{11.14}
\]

Consider the determinant of the $n \times n$ matrix $I_n + X$. It can be shown that the determinant can be expanded into subdeterminants [2]
\[
|I_n + X| = \sum_{k=0}^{n} \sum_{\hat{\alpha}_k} |X|^{\hat{\alpha}_k},
\tag{11.15}
\]

with the index set $\hat{\alpha}_k = \{\alpha_1, \alpha_2, \ldots, \alpha_k\}$, where $\alpha_k \subseteq \{1, 2, \ldots, n\}$ with cardinality $|\hat{\alpha}_k| = k$. For clarity, the second sum in (11.15) is over all index sets of cardinality $k$.

Let $T$ be a $k \times k$ matrix resulting from the matrix product (where $A, B, \ldots, R, S$ are of compatible sizes)
\[
T = A \cdot B \ldots \cdot R \cdot S.
\tag{11.16}
\]
Then we can expand the determinant of $T$ in a multiple sum expression \[2\] \[13\] of subdeterminants
\[ (11.17) \]

Vandermonde determinant

The Vandermonde determinant of a diagonal $m \times m$ matrix $X = \text{diag}(x_1, x_2, \ldots, x_m)$, or equivalently a vector $x = (x_1, x_2, \ldots, x_m)^T$, reads in general
\[ \alpha_m(X) = \alpha_m(x) = \begin{vmatrix} x_1^{m-1} & x_1^{m-2} & \cdots & x_1 & 1 \\ x_2^{m-1} & x_2^{m-2} & \cdots & x_2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_m^{m-1} & x_m^{m-2} & \cdots & x_m & 1 \end{vmatrix}. \] \[ (11.18) \]

By column-wise factoring out $x_i^{m-1}$ in (11.18), it can be shown that
\[ \alpha_m(X) = |X|^{m-1} \cdot \alpha_m(X^{-1}) \cdot (-1)^{\frac{m \cdot (m-1)}{2}}. \] \[ (11.19) \]

Now consider the product of two Vandermonde determinants of size $n$
\[ \alpha_n(\Sigma) \cdot \alpha_n(\Omega) \] \[ (11.20) \]
with $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n)$ and $\Omega = \text{diag}(\omega_1, \ldots, \omega_n)$. We can use the properties of determinants
\[ |X| = |X^T| \quad |X||Y| = |XY| \] \[ (11.21) \]
for square matrices $X, Y$, and find
\[ \alpha_n(\Sigma) \cdot \alpha_n(\Omega) = \begin{vmatrix} n-1 & n-1 & n-1 & n-1 & n-1 \\ \sum_{k=0}^{n-1} (\sigma_1 \omega_1)^k & \sum_{k=0}^{n-1} (\sigma_1 \omega_2)^k & \cdots & \sum_{k=0}^{n-1} (\sigma_1 \omega_{n-1})^k & \sum_{k=0}^{n-1} (\sigma_1 \omega_n)^k \\ \sum_{k=0}^{n-1} (\sigma_2 \omega_1)^k & \sum_{k=0}^{n-1} (\sigma_2 \omega_2)^k & \cdots & \sum_{k=0}^{n-1} (\sigma_2 \omega_{n-1})^k & \sum_{k=0}^{n-1} (\sigma_2 \omega_n)^k \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sum_{k=0}^{n-1} (\sigma_n \omega_1)^k & \sum_{k=0}^{n-1} (\sigma_n \omega_2)^k & \cdots & \sum_{k=0}^{n-1} (\sigma_n \omega_{n-1})^k & \sum_{k=0}^{n-1} (\sigma_n \omega_n)^k \end{vmatrix}. \] \[ (11.22) \]
Then consider the product of two Vandermonde determinants of size \( T \) and \( R \) (\( R \geq T \)) \( \alpha_T(\Sigma) \cdot \alpha_R(\Omega) \) with \( \Sigma = \text{diag}(\sigma_1, ..., \sigma_T) \) and \( \Omega = \text{diag}(\omega_1, ..., \omega_R) \). Artificially, we can increase the size of the first determinant to \( R \times R \) by introducing a size \( R-T \) identity matrix, which does not alter the determinant

\[
\alpha_T(\Sigma) = \begin{vmatrix}
\sigma_1^{T-1} & \sigma_1^{T-2} & \cdots & \sigma_1 & 1 \\
\sigma_2^{T-1} & \sigma_2^{T-2} & \cdots & \sigma_2 & 1 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\sigma_T^{T-1} & \sigma_T^{T-2} & \cdots & \sigma_T & 1 \\
1 & 1 & \cdots & 1 & I_{R-T}
\end{vmatrix}. \quad (11.23)
\]

It is then straightforward via direct matrix multiplication to show that

\[
\alpha_T(\Sigma) \cdot \alpha_R(\Omega) = \begin{vmatrix}
\sum_{k=0}^{T-1} \sigma_1^k \omega_1^{R-k} + \sum_{k=0}^{T-1} \sigma_2^k \omega_2^{R-k} + \cdots + \sum_{k=0}^{T-1} \sigma_T^k \omega_T^{R-k} \\
\sum_{k=0}^{T-1} \sigma_1^k \omega_1^{R-k} + \sum_{k=0}^{T-1} \sigma_2^k \omega_2^{R-k} + \cdots + \sum_{k=0}^{T-1} \sigma_T^k \omega_T^{R-k} \\
\vdots \\
\sum_{k=0}^{T-1} \sigma_1^k \omega_1^{R-k} + \sum_{k=0}^{T-1} \sigma_2^k \omega_2^{R-k} + \cdots + \sum_{k=0}^{T-1} \sigma_T^k \omega_T^{R-k} \\
\omega_1^{R-1} & \omega_2^{R-1} & \cdots & \omega_{R-1} & \omega_R \\
1 & 1 & \cdots & 1 & 1
\end{vmatrix}. \quad (11.24)
\]

### 11.1.6 Elementary symmetric functions of a matrix

In statistical literature we often find the definition of the \( k \) th elementary symmetric function in the eigenvalues of a \( n \times n \) matrix \( X \), which can be expressed as (see e.g. [59])

\[
\text{tr}_k(X) \equiv \sum_{\delta_i} |X|^\delta_i = \sum_{\delta_i} |\text{eig}(X)|^\delta_i. \quad (11.25)
\]

where \( |X|^\delta_i \) is a so-called principal minor of matrix \( X \) [125]. The standard trace operator is then given by

\[
\text{tr}(X) = \sum_{\delta_1} |X|^\delta_1 = \text{tr}_1(X) \quad (11.26)
\]

and the determinant of the \( n \times n \) matrix \( X \) reads
\[ |X| \equiv \text{tr}_n(X). \tag{11.27} \]

11.2 Multivariate statistics

Classical introductory textbooks on multivariate statistics are [7][140]. A lot of results and references can be found in [83] and a very detailed overview on matrix variate distributions is given in [59]. If not stated otherwise, most of the results can be found in this reference.

11.2.1 Complex normal distributions

Definition

Let the vector-valued complex normal distribution with \( m \) elements, covariance matrix \( \Sigma \), and mean \( \mu \) be denoted by \( \tilde{N}_m(\mu, \Sigma) \). A \( m \times n \) matrix \( X \) is said to have a matrix variate complex normal distribution [59] with covariance matrices \( \Sigma, \Psi \), and mean \( M \), denoted by \( \tilde{N}_{m,n}(M, \Sigma \otimes \Psi) \), if

\[ \text{vec}(X^T) \sim \tilde{N}_{mn}(\text{vec}(M^T), \Sigma \otimes \Psi). \tag{11.28} \]

Marginal distributions

Now introduce the following partitions

\[ X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}, \quad M = \begin{bmatrix} M_1 & M_2 \end{bmatrix}, \quad \Psi = \begin{bmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{bmatrix}. \tag{11.29} \]

with \( X_1 \) of size \( m \times r \), \( X_2 \) of size \( m \times (n-r) \), equivalent partitions for \( M \), \( \Psi_{11} \) of size \( r \times r \), \( \Psi_{22} \) of size \( (n-r) \times (n-r) \), \( \Psi_{12} \) of size \( r \times (n-r) \), \( \Psi_{21} \) of size \( (n-r) \times r \), and the definition

\[ \Psi_{11 \cdot 2} = \Psi_{11} - \Psi_{12} \Psi_{22}^{-1} \Psi_{21}. \tag{11.30} \]

It can then be shown that

\[ X_2 \sim \tilde{N}_{m,n-r}(M_2, \Sigma \otimes \Psi_{22}), \tag{11.31} \]

and conditioned on \( X_2 \), the distribution of \( X_1 \) is given by

\[ X_1 | X_2 \sim \tilde{N}_{m-r}(M_1 + (X_2 - M_2) \Psi_{22}^{-1} \Psi_{21}, \Sigma \otimes \Psi_{11 \cdot 2}). \tag{11.32} \]

Expected values

For the interested reader we note that references [181][183] deal with expected values and moments with respect to multivariate complex normal distributions.
Let $H_w$ be a matrix with zero-mean i.i.d. complex Gaussian elements of size $R \times T$. For a general deterministic matrix $X$ of size $T \times T$ one can derive
\[
E_{H_w}[H_wXH_w^H] = \text{tr}(X) \cdot I_R. \tag{11.33}
\]

The following expected value seems not to explicitly appear in statistical literature. Again let $H_w$ be a matrix with zero-mean i.i.d. complex Gaussian elements of size $R \times T$ ($R \geq T$), then we have (see Paragraph 9.3.1 for hints on the derivation)
\[
E_{H_w}[H_w^HXH_w] = \text{tr}_F(X) \cdot \Gamma(T + 1). \tag{11.34}
\]

As the determinant of a complex Wishart distributed matrix has the same distribution as the product of independently distributed chi-squared distributed RVs [44], it is possible by straightforward integration to show that for $W \sim \tilde{W}_m(n, I)$ we have (see e.g. [52, lemma A.2])
\[
E_W[\ln|W|] = \sum_{k=0}^{m-1} \psi(n-k). \tag{11.35}
\]
where $\psi(z)$ is the so-called Psi or Digamma function [1, Paragraph 6.3], which reads for integer arguments $n$ and Euler’s constant $E=0.5772156649$ (see e.g. [1, equation 6.1.3])
\[
\psi(1) = -E \quad \psi(n) = -E + \sum_{k=1}^{n-1} \frac{1}{k}. \tag{11.36}
\]

An important result, which is given e.g. in [163] and indirectly in [190], is the following expected value. For a complex Gaussian distributed random vector $z \sim \tilde{N}(\mu, \Sigma)$ with expected value $\mu = E[z]$, covariance matrix $\Sigma = E[zz^H] - \mu\mu^H$, and a Hermitian matrix $A$ it follows
\[
E_z[\exp(-z^HAz)] = \frac{1}{|I + \Sigma A|} \cdot \exp(-\mu^HA(I + \Sigma A)^{-1}\mu). \tag{11.37}
\]

### 11.3 Complex Integrals

#### 11.3.1 Notation

In this thesis, we use the following notation for the complex matrix integral measure for complex $M \times N$ matrix $X = [x_{mn}]$ ($m = 1 \ldots M, n = 1 \ldots N$)
\[
D_{c}X = \prod_{m=1}^{M} \prod_{n=1}^{N} \frac{d\text{Re}\{x_{mn}\} \cdot d\text{Im}\{x_{mn}\}}{\pi}. \tag{11.38}
\]
Furthermore, we simply write for a function $f(X)$
11.3.2 Some Gaussian integrals

The following results are given e.g. in [31][137], whereas the basics can be found in [55][64].

For complex column vectors $x, a, b$ it can be shown that

$$
\int e^{-\frac{1}{2} x^H A x + a^H x + x^H b} D_x X = \frac{1}{|A|} \cdot e^{a^H A^{-1} b}.
$$

(11.40)

A more general case of the integral in (11.40) can be derived from [31] and reads

$$
\int \frac{1}{|B|} \cdot [\text{tr}(A B^{-1}) - \text{tr}(B^{-1} A) - \text{tr}(A B^{-1}) A B^{-1} B + 2 \text{Re}[a_0] \cdot \exp(b^H B^{-1} b - 2 \text{Re}[b_0])]
$$

(11.41)

The result in (11.40) can be extended and for complex $m \times n$ matrices $X, A, B$, $m \times m$ matrix $M$, and $n \times n$ matrix $N$ ($n \leq m$) we get

$$
\int e^{-\frac{1}{2} x^H M x + a^H x + x^H b} D_x X = \frac{1}{|N \otimes M|} \cdot e^{\text{tr}(N^{-1} A^{-1} M^{-1} B)}.
$$

(11.42)

An important special case of (11.40)(11.42) is

$$
\int e^{-\frac{1}{2} x^H M x} D_x X = \frac{1}{|M|}.
$$

(11.43)

One can establish the following integral identity with scalar function $f(x^H x)$ and $m \times 1$ complex vector $x$

$$
\int f(x^H x) \cdot e^{-\frac{1}{2} x^H x} D_x X = \frac{1}{\Gamma(m)} \cdot \int_0^\infty f(t) \cdot t^{m-1} \cdot e^{-t} dt.
$$

(11.44)

The identity follows from the fact that (11.44) is the expected value of $f(x^H x)$ with respect to a i.i.d. complex Gaussian $m \times 1$ vector $x$, i.e. $t = x^H x$ has a Gamma distribution with $m$ degrees of freedom.

11.4 Complex Gaussian Distribution and Associated Integrals

11.4.1 Probability density function

Let $G$ be an i.i.d. $m \times n$ matrix ($m \geq n$) of zero mean complex Gaussian elements. Then the joint probability density function (PDF) is
\[ p(G) = e^{-\text{tr}(G^H G)} D_G G. \]  
(11.45)

If \( \bar{G} \) is non-central with \( E[\bar{G}] = M \), then the PDF is given by
\[ p(\bar{G}) = e^{-\text{tr}((\bar{G} - M)(\bar{G} - M)^H)} D_G \bar{G}. \]  
(11.46)

In case of a \( m \times n \) matrix \( \tilde{G} \) with central complex Gaussian distribution and covariance matrix \( \Sigma \) we get the PDF
\[ p(\tilde{G}) = \frac{1}{|\Sigma|^n} e^{-\text{tr}(\Sigma^{-1}\tilde{G}^H \tilde{G})} D_G \tilde{G}. \]  
(11.47)

### 11.4.2 Some integrals

In all following integrals \( G \) is a \( m \times n \) matrix with complex elements. Using the expected value of the Wishart determinant in (11.34) and the PDF in (11.45), it can be seen that
\[ \int |G^H G| e^{-\text{tr}(G^H G)} D_G G = \binom{m}{n} \cdot n!. \]  
(11.48)

Using the PDF with non-identity covariance \( \Sigma \), we equivalently get
\[ \int |G^H G| e^{-\text{tr}(\Sigma^{-1}G^H G)} D_G G = \binom{m}{n} \cdot n! \cdot |\Sigma|^{n+1}. \]  
(11.49)

If \( \bar{G} \) has a non-central complex Gaussian distribution according to (11.46), then it was shown in [167][196] that
\[ E_{\bar{G}}[|\bar{G}^H \bar{G}|] = \frac{\Gamma(m + 1)}{\Gamma(m - n + 1)} \sum_{i=0}^{n-1} \frac{\Gamma(m - i)}{\Gamma(m - n + 1)} \cdot \text{tr}_{i+1}(Q), \]  
(11.50)

where \( Q = MM^H \). In case of \( Q \) being of rank 1 this obviously reduces to
\[ E_{\bar{G}}[|\bar{G}^H \bar{G}|] = \frac{1}{\Gamma(m - n + 1)} [\Gamma(m + 1) + \Gamma(m) \cdot \text{tr}(Q)]. \]  
(11.51)

Therefore, we have for this special case the integration result
\[ \int |G^H G| e^{-\text{tr}((G - M)(G - M)^H)} D_G G = \frac{1}{\Gamma(m - n + 1)} [\Gamma(m + 1) + \Gamma(m) \cdot \text{tr}(MM^H)] . \]  
(11.52)

### 11.5 Majorization theory

A powerful source of results on majorization theory is [131]. Most of the following definitions, lemmas, and theorems stem from this reference.
11.5.1 Basic definitions

**Definition 11.1:** For any \( x \in \mathbb{R}^n \), let

\[ x_{[1]} \geq \ldots \geq x_{[n]} \]  

(11.53)

denote the elements of \( x \) in decreasing order (also termed order statistics).

**Definition 11.2:** [131, 1.A.1] Let \( x, y \in \mathbb{R}^n \). Vector \( x \) is majorized by vector \( y \) (\( y \) majorizes \( x \)) if

\[
\sum_{i=1}^k x_{[i]} \leq \sum_{i=1}^k y_{[i]} \quad 1 \leq k \leq n - 1
\]

\[
\sum_{i=1}^n x_{[i]} = \sum_{i=1}^n y_{[i]}
\]

(11.54)

and is represented by \( x \prec y \).

**Definition 11.3:** [131, 3.A.1] A real-valued function \( \phi \) defined on a set \( A \subseteq \mathbb{R}^n \) is said to be Schur-convex if

\[
x \prec y \text{ on } A \Rightarrow \phi(x) \leq \phi(y).
\]

(11.55)

Similarly, \( \phi \) is said to be Schur-concave on \( A \) if

\[
x \prec y \text{ on } A \Rightarrow \phi((x) \geq \phi(y)).
\]

(11.56)

As a consequence, if \( \phi \) is Schur-convex on \( A \), then \(-\phi\) is Schur-concave on \( A \) and vice-versa.

11.5.2 Some lemmas (including results for matrices)

**Lemma 11.1:** [131, 9.B.1] Let \( R \) be an \( n \times n \) Hermitian matrix with diagonal elements denoted by the vector \( d \) and eigenvalues denoted by the vector \( \lambda \), then

\[ d \prec \lambda. \]  

(11.57)

**Lemma 11.2:** [131, page 7] Let \( x \in \mathbb{R}^n \) and let \( 1_{n \times 1} \) be a \( n \times 1 \) vector of all ones. Then with

\[
\kappa = \left( \sum_{j=1}^n x_j \right) / n
\]

(11.58)

it follows

\[ \kappa \cdot 1_{n \times 1} \prec x. \]  

(11.59)
Lemma 11.3: [131, 9.B.2] If $h \prec \lambda$ on $\mathbb{R}^n$, then there exists a real symmetric matrix $H$ with diagonal elements $h$ and characteristic roots $\lambda$.

Lemma 11.4: [131, 9.H.1.h] If $U$ and $V$ are $n \times n$ positive semidefinite Hermitian matrices, then

$$\text{tr}(UV) \geq \sum_{i=1}^{n} \lambda_i(U) \cdot \lambda_{n-i+1}(V),$$

where the eigenvalues are sorted in decreasing order.

Lemma 11.5: [146, lemma 12 in appendix] Let $R$ be a full rank positive semidefinite Hermitian matrix with eigenvalue decomposition

$$R = \begin{bmatrix} \tilde{L} & \tilde{U} \end{bmatrix} \begin{bmatrix} L & U \end{bmatrix}^H,$$

where the matrix $\tilde{L}$ contains the $L$ largest eigenvalues in increasing order and $\tilde{U}$ the corresponding eigenvectors. Given a $R \times L$ ($R \geq L$) matrix $B$ such that $B^HRB$ is diagonal with diagonal elements in increasing order, it is always possible to find a matrix $\tilde{B}$ of the form $\tilde{B} = \tilde{U} \Sigma$ with diagonal matrix $\Sigma$, such that it satisfies $\tilde{B}^H \tilde{B} = B^HRB$ with $\text{tr}(\tilde{B}^H) \leq \text{tr}(BB^H)$.

11.5.3 Theorems on Schur-convexity

Theorem 11.1: [131, 3.A.3] Let $\mathcal{D} = \{x : x_1 \geq \ldots \geq x_n\}$ and define

$$\phi_{(k)}(z) = \frac{\partial}{\partial z_k} \phi(z),$$

where $\phi$ is a real valued function, defined and continuous on $\mathcal{D}$ and continuously differentiable on the interior of $\mathcal{D}$. Then

$$x \prec y \text{ on } \mathcal{D} \text{ implies } \phi(x) \leq \phi(y)$$

if and only if

$$\phi_{(k)}(z) \text{ is decreasing in } k = 1 \ldots n.$$  (11.64)

Note that there are lots of results on compositions involving Schur-convex functions in [131, 3.B].

Theorem 11.2: [131, 3.C.2] If $\phi$ is symmetric (i.e. symmetric in its arguments) and convex, then $\phi$ is Schur-convex.

11.5.4 Theorem on Schur-concavity

Theorem 11.3: [131, 3.F.1] The elementary symmetric function $\text{tr}_k(X)$ of the positive semi-definite matrix $X$ is Schur-concave on $\mathbb{R}_n^+$. 
11.6 Special functions

If not stated otherwise, most of the results given in this paragraph can be found in [1] and [50].

11.6.1 Q function

\[ Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-\frac{t^2}{2}} dt. \]  \hspace{1cm} (11.65)

11.6.2 Complementary error function

\[ \text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_{x}^{\infty} e^{-t^2} dt. \]  \hspace{1cm} (11.66)
\[ \text{erfc}(\sqrt{ax}) = 2Q(\sqrt{2ax}). \]  \hspace{1cm} (11.67)

The so-called Chernoff bound can be found e.g. in [153]

\[ \text{erfc}(\sqrt{ax}) \leq e^{-ax}. \]  \hspace{1cm} (11.68)

An integral representation with finite integration limits [172] is given by

\[ \text{erfc}(\sqrt{x}) = \frac{2}{\pi} \int_{0}^{\frac{\pi}{2}} \exp\left(-\frac{x}{\sin^2 \theta}\right) d\theta. \]  \hspace{1cm} (11.69)

11.6.3 Exponential integral \( E_1(x) \)

The exponential integral is defined by [1, 5.1.1]

\[ E_1(x) = \int_{x}^{\infty} \frac{e^{-t}}{t} dt. \]  \hspace{1cm} (11.70)

Series expansion of \( e^{1/x} \cdot E_1\left(\frac{1}{x}\right) \)

From [1, 5.1.11], we have

\[ E_1\left(\frac{1}{z}\right) = -E + \ln z - \sum_{n=1}^{\infty} \frac{(-1)^n \cdot z^{-n}}{n \cdot n!} = -E + \ln z - \left(\frac{1}{z} + \frac{1}{4z^2} - \frac{1}{18z^3} + \ldots\right) \]  \hspace{1cm} (11.71)

with Euler’s constant \( E=0.5772156649 \). On the other hand
\[ e^{1/z} = \sum_{n=0}^{\infty} \frac{z^{-n}}{n!}. \tag{11.72} \]

Combining the results of (11.71) and (11.72) we find

\[ e^{1/z} \cdot E_1\left(\frac{1}{z^2}\right) = (-E + \ln z) \cdot \sum_{n=0}^{\infty} \frac{z^{-n}}{n!} - \left( \sum_{n=1}^{\infty} \frac{(-1)^n \cdot z^{-n}}{n \cdot n!} \right) \cdot \sum_{n=0}^{\infty} \frac{z^{-n}}{n!}. \tag{11.73} \]

Then consider the series expansion with coefficients \( a_{E, k} \)

\[ \left( \sum_{n=1}^{\infty} \frac{(-1)^n \cdot z^{-n}}{n \cdot n!} \right) \cdot \sum_{n=0}^{\infty} \frac{z^{-n}}{n!} = \sum_{k=1}^{\infty} a_{E, k} \cdot z^{-k} = \frac{a_{E, 1}}{z} + \frac{a_{E, 2}}{z^2} + \ldots + \frac{a_{E, k}}{z^k} + \ldots, \tag{11.74} \]

where ‘E’ in \( a_{E, k} \) is reminiscent of the exponential integral. By sorting the terms of the product of sums in (11.74), we can find the series coefficients

\[ a_{E, k} = \sum_{l=1}^{k} \frac{(-1)^l}{l \cdot l!} \cdot \frac{1}{(k-l)!} \tag{11.75} \]

with the first 3 coefficients

\[ a_{E, 1} = -1 \]
\[ a_{E, 2} = -\frac{3}{4} \]
\[ a_{E, 3} = -\frac{11}{36} \tag{11.76} \]

Obviously, using above results we can rewrite (11.73) as

\[ e^{1/z} \cdot E_1\left(\frac{1}{z^2}\right) = (-E + \ln z) \cdot \sum_{n=0}^{\infty} \frac{z^{-n}}{n!} - \sum_{k=1}^{\infty} a_{E, k} \cdot z^{-k}. \tag{11.77} \]

**Approximation for \( e^{1/x} \cdot E_1\left(\frac{1}{x}\right) \), \( x \) small**

From [1.5.1.19], we have

\[ \frac{1}{z+1} \leq e^z \cdot E_1(z) \leq \frac{1}{z}. \tag{11.78} \]

For large \( z \) we obviously find
\[
e^z \cdot E_1(z) \approx \frac{1}{z}, \quad (11.79)
\]

Equivalently, we have for small \(x\)
\[
e^{\frac{1}{x}} \cdot E_1\left(\frac{1}{x}\right) \approx x. \quad (11.80)
\]

### 11.6.4 General hypergeometric function \(pF_q(a_1, a_2, \ldots, a_p; b_1, b_2, \ldots, b_q; z)\)

\[
pF_q(a_1, a_2, \ldots, a_p; b_1, b_2, \ldots, b_q; z) = \sum_{k=0}^{\infty} \frac{[a_1]_k [a_2]_k \cdots [a_p]_k}{[b_1]_k [b_2]_k \cdots [b_q]_k} \frac{z^k}{k!}. \quad (11.81)
\]

### 11.6.5 Hypergeometric function \(2F_1(a_1, a_2; b_1; z)\)

See e.g. [1, Chapter 15] for the series definition
\[
\sum_{k=0}^{\infty} \frac{[a_1]_k [a_2]_k}{[b_1]_k} \frac{z^k}{k!}. \quad (11.82)
\]

**Large \(z\) asymptotics**

First consider the linear transformation formula [1, 15.3.7]
\[
\Gamma(b_1) \cdot \Gamma(a_2 - a_1) \Gamma(a_2) \cdot \Gamma(b_1 - a_1) \cdot (-z)^{a_1} \cdot \frac{2F_1(a_1, 1 - b_1 + a_1; 1 - a_2 + a_1; \frac{1}{z})}{\Gamma(a_2) \cdot \Gamma(b_1 - a_1) \cdot \Gamma(a_1) \cdot \Gamma(b_1 - a_2)} \cdot (-z)^{a_2} \cdot \frac{2F_1(a_2, 1 - b_1 + a_2; 1 - a_1 + a_2; \frac{1}{z})}{\Gamma(b_1) \cdot \Gamma(a_1 - a_2) \cdot \Gamma(b_1 - a_2)}.
\]

Using the series definition (11.82) in (11.83) and neglecting higher negative powers of \(z\), we can approximate
\[
2F_1(a_1, a_2; b_1; z) \approx \frac{\Gamma(b_1) \cdot \Gamma(a_2 - a_1)}{\Gamma(a_2) \cdot \Gamma(b_1 - a_1)} \cdot (-z)^{a_1} \left(1 + \frac{a_1 \cdot (1 - b_1 + a_1)}{1 - a_2 + a_1} \cdot \frac{1}{z}\right) + \frac{\Gamma(b_1) \cdot \Gamma(a_1 - a_2)}{\Gamma(a_1) \cdot \Gamma(b_1 - a_2)} \cdot (-z)^{a_2} \left(1 + \frac{a_2 \cdot (1 - b_1 + a_2)}{1 - a_1 + a_2} \cdot \frac{1}{z}\right). \quad (11.84)
\]
11.6.6 Kummer U(a,b,z) function

Relation to hypergeometric function

We give a relation between the scalar hypergeometric function $\binom{a}{1 + a - b; -\frac{1}{z}}$ and the Kummer $U(a, b, z)$ function. To this end we have from [1, Paragraph 13.1], the alternative notations for the Kummer $U$ function

$$U(a, b, z) = \Psi(a; b; z) = z^{-a} \binom{a}{1 + a - b; -\frac{1}{z}}.$$  \hfill (11.85)

Furthermore, the Kummer transformation reads (cf. [1, equation 13.1.29])

$$U(a, b, z) = z^{1-b} \cdot U(1 + a - b, 2 - b, z).$$  \hfill (11.86)

From (11.85), together with (11.86) we find

$$\binom{a}{1 + a - b; -\frac{1}{z}} = z^{1+a-b} \cdot U(1 + a - b, 2 - b, z).$$ \hfill (11.87)

Integral representation

Equation (11.87) is important, as there are integral representations (in contrast to the infinite sum representation of the hypergeometric function (see [1] or (11.81)), which could possibly exhibit convergency problems) available for the Kummer $U$ function, whereas [1, equation 13.2.5] reads

$$U(a, b, z) = \frac{1}{\Gamma(a)} \cdot \int_0^\infty e^{-zt} \cdot t^{a-1} \cdot (1 + t)^{b-a-1} dt.$$ \hfill (11.88)

A special case

Using the integral representation (11.88) we find

$$U(1, b, z) = \int_0^\infty e^{-zt} \cdot (1 + t)^{b-2} dt = \frac{e^z \cdot \Gamma(b - 1, z)}{z^{b-1}},$$ \hfill (11.89)

where $\Gamma(a, z)$ is the incomplete Gamma function [1].

Recurrence relations

We further state the following recurrence relation for the Kummer $U$ function

$$U(1, b, z) = \frac{1}{z} + \frac{1}{z^2} (b - 2) U(1, b - 1, z),$$ \hfill (11.90)

which can be derived from (11.88) via integration by parts and in the general case [1, 13.4.18].
\[ U(a, b, z) = \frac{1}{z} \cdot [(b - a - 1) \cdot U(a, b - 1, z) + U(a - 1, b - 1, z)]. \quad (11.91) \]

Moreover, from [1, 13.4.17] we have
\[ U(a, b - 1, z) = U(a, b, z) - a \cdot U(a + 1, b, z). \quad (11.92) \]
Bibliography


[85] Jorswieck E., Boche H., “Channel capacity and capacity-range of beamforming in MIMO wireless systems under correlated fading with covariance feedback,” *IEEE Journal on Selected Areas in Communications*, 2003, submitted


[120] Lee W. C. Y., Mobile communications design fundamentals, John Wiley&Sons, 1993


[125] Lütkepohl H., Handbook of matrices, John Wiley&Sons, 1996


[133] Mathai A. M., Provost S. B., Quadratic forms in random variables, Marcel Dekker, 1992
Bibliography 189


[207] 3GPP, “A standardized set of MIMO radio propagation channels,“ *TSG#23 RI-01-1179*, Jeju, Korea, Nov. 2001