# Evaluatie van de gemiddelde bitfoutprobabiliteit voor orthogonale spatio-temporele blokcodes in aanwezigheid van kanaalschattingsfouten <br> Bit Error Rate Evaluation <br> for Orthogonal Space-Time Block Codes in the Presence of Channel Estimation Errors <br> Lennert Jacobs 

Promotor: prof. dr. ir. M. Moeneclaey
Proefschrift ingediend tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen: Elektrotechniek
Vakgroep Telecommunicatie en Informatieverwerking
Voorzitter: prof. dr. ir. H. Bruneel
IIIIII
Faculteit Ingenieurswetenschappen en Architectuur
Academiejaar 2011-2012

## Dankwoord

Vermits deze doctoraatsthesis de resultaten bevat van het wetenschappelijk onderzoek waaraan ik mij de voorbije vijf jaar gewijd heb, is het niet meer dan vanzelfsprekend dat ik de vele mensen wil bedanken zonder wie dit wapenfeit onmogelijk of toch op zijn minst minder aangenaam zou geweest zijn. En dat zijn er heel wat.

De eerste naam in dit lijstje is, uiteraard, mijn promotor Marc Moeneclaey, die de wonderlijke gave bezit om zijn doctoraatsstudenten zowel voldoende vrijheid en verantwoordelijkheid te geven, als verder te helpen met een to the point en relevant antwoord wanneer eender welk probleem opduikt. Ik kan hem niet genoeg bedanken voor de nagenoeg perfecte begeleiding die ik de voorbije jaren heb mogen ervaren.

De twee collega's die mijn professionele beslommeringen de voorbije jaren het meest opgeleukt hebben, en daarvoor zonder enige twijfel een eigen paragraaf in dit dankwoord verdienen, zijn Dieter Van Welden en Jabran Bhatti, tevens mijn vaste kantoor- en lunchmakkers. Niet enkel brachten zij mij de principes van KSP-OFDM en de discrete cosine transform bij, zonder hen zou mijn interesse en appreciatie voor de betere Italo-disco, underground techno en funky beats uit Pakistan en ver daarbuiten nooit zo een hoge vlucht
genomen hebben. Onvergetelijk waren de vrije momenten op conferenties in Barcelona, Tokyo of San Francisco, waarin de toeristische trekpleisters het moesten afleggen tegen onze honger naar stoffige kratten vol vinyl. Krautrock, psych, Japanese dark noise, Taiwanese Pop of stapels seventies funk, overal vielen wel pareltjes te rapen.

Ook mijn andere huidige en voormalige collega's bij DIGCOM verdienen mijn welgemeende dank: Heidi Steendam, Nele Noels, Julie Neckebroeck, Dieter Duyck, Nico Aerts, Iancu Avram, Samuel Van de Velde, Jeroen Van Hecke, Frederik Simoens en Frederik Van Haverbeke. Absoluut niet te vergeten in dit uitgebreide lijstje zijn mijn kantoorgenoten Bart Feyaerts en Sebastian Grünwedel en hun sympathieke voorgangers Laurence Hoflack en Ewout Vansteenkiste. Daarnaast dank ik Philippe Serbruyns en Davy Moreels voor de deskundige IT-support, en Patrick Schaillée, Annette Nevejans, Alice Verheylesonne en Sylvia Moeneclaey voor hun prima werk op het secretariaat.

Ook buitenlandse collega's hebben een rol gespeeld in de totstandkoming van dit doctoraat. Graag had ik mijn Griekse vriend Georgios ‘George' Alexandropoulos en diens aimabele promotor prof. Takis Mathiopoulos bedankt voor de uitstekende samenwerking. Verder dank ik ook prof. Andrea Conti voor zijn gastvrijheid. Mijn korte maar boeiende verblijf aan de universiteit van Ferrara in Italië leverde een vruchtbare samenwerking op.

Uiteraard dank ik mijn aangename ouders, broer en familie voor de onvoorwaardelijke steun en het vertrouwen gedurende de voorbije 28 jaar.

Een uiterst belangrijk woord van dank gaat uit naar mijn talloze muzikale vrienden, met wie ik niet enkel graag samen speel omwille van hun muzikaal talent, maar ook omdat ze zo sympathiek, intelligent, tegendraads en/of plezant zijn. In de eerste plaats dank ik uiteraard mijn olijke vrienden van The Germans: Vincent, Jakob en mijn broer Timothy (al voor de tweede keer). Het plezier waarmee wij nu al bijna tien jaar lang commerciële zelfmoord plegen is onbetaalbaar. Daarnaast dank ik ook mijn geweldige vrienden Boris, Sarah en Geoffrey van Hong Kong Dong, en Jonathan, David en Wesley (tevens een ideale huisgenoot!) van de notoire Rape Blossoms.

Last but allerminst least zou ik, in willekeurige volgorde, de volgende uitermate sympathieke vrienden en vriendinnen willen bedanken, met wie het steevast plezierig pilsjes consumeren is: Charlotte, Elise, Thomas, Jason, Benne, Dijf, Bert H, Jeroen S, Liesbeth, Jan, Koen, Pieter-Paul, Pieter, Jonas, Daan, Jeroen V, Bert M, Hein, en de onafscheidelijke duo's Fran en Ellen, Cat Power en Inge, en Nana en Marjolein. De mensen die ik hier schromelijk over het hoofd gezien heb, zijn natuurlijk ook bedankt. Of net niet, om met een vrolijke noot te eindigen.

## Contents

List of Abbreviations ..... V
List of Notations ..... vii
Nederlandstalige Samenvatting ..... ix
English Summary ..... xiii
1 Introduction ..... 1
1.1 Background and Motivation ..... 1
1.2 Outline ..... 4
2 Estimation and Detection Theory ..... 7
2.1 Definitions ..... 8
2.2 Estimation ..... 9
2.3 Detection ..... 11
2.4 Chapter Summary ..... 11
3 Conventional Digital Communication ..... 13
3.1 Band-Pass Signals and Systems ..... 14
3.2 System Description ..... 15
3.3 Bit Error Rate ..... 28
3.4 Chapter Summary ..... 33
4 MIMO Communication ..... 35
4.1 Diversity ..... 35
4.2 Orthogonal Space-Time Block Codes ..... 41
4.3 MIMO Channel Model ..... 48
4.4 Chapter Summary ..... 59
5 Channel Estimation ..... 61
5.1 Insertion of Pilot Symbols ..... 61
5.2 ML Channel Estimation ..... 63
5.3 LMMSE Channel Estimation ..... 64
5.4 Chapter Summary ..... 67
6 BER Analysis of OSTBCs in Rayleigh Fading ..... 69
6.1 Rule of Thumb ..... 70
6.2 BER Analysis for i.i.d. Rayleigh Fading ..... 74
6.3 BER Analysis for Correlated Rayleigh Fading ..... 83
6.4 Chapter Summary ..... 95
6.A Appendix ..... 96
7 BER Analysis of OSTBCs in Arbitrary Fading ..... 103
7.1 Approximate BER Analysis ..... 104
7.2 Exact BER Analysis ..... 105
7.3 Evaluation of the BER ..... 108
7.4 Numerical Results ..... 113
7.5 Chapter Summary ..... 121
7.A Appendix ..... 123
8 Food for Future Thought ..... 127
8.1 Channel Coding ..... 127
8.2 Adaptive Modulation and Coding ..... 129
8.3 MIMO-OFDM ..... 139
8.4 Chapter Summary ..... 142
9 Summary and Conclusions ..... 145
Publications ..... 149
Bibliography ..... 151

## Abbreviations

| AMC | Adaptive Modulation and Coding |
| :--- | :--- |
| ASE | Achieved Spectral Efficiency |
| AWGN | Additive White Gaussian Noise |
| BEO | Bit Error Outage |
| BER | Bit Error Rate |
| CSCG | Circularly Symmetric Complex Gaussian |
| CSI | Channel State Information |
| D-BLAST | Diagonal Bell Labs Space-Time |
| EGC | Equal Gain Combining |
| FAM | Fast Adaptive Modulation |
| FFT | Fast Fourier Transform |
| IBI | Inter-Block Interference |
| ICI | Inter-Carrier Interference |
| ICSI | Imperfect Channel State Information |
| i.i.d. | independent and identically distributed |
| ISI | Inter-Symbol Interference |
| LMMSE | Linear Minimum Mean Square Error |
| LOS | Line-Of-Sight |


| MAP | Maximum A Posteriori |
| :--- | :--- |
| MGF | Moment Generating Function |
| MIMO | Multiple-Input Multiple-Output |
| ML | Maximum Likelihood |
| MMSE | Minimum Mean Square Error |
| MRC | Maximal-Ratio Combining |
| MSE | Mean Square Error |
| OFDM | Orthogonal Frequency Division Multiplexing |
| OSFBC | Orthogonal Space-Frequency Block Code |
| OSTBC | Orthogonal Space-Time Block Code |
| PAM | Pulse Amplitude Modulation |
| PCSI | Perfect Channel State Information |
| PDF | Probability Density Function |
| PEP | Pairwise Error Probability |
| PSK | Phase Shift Keying |
| QAM | Quadrature Amplitude Modulation |
| RF | Radio Frequency |
| RV | Random Variable |
| SAM | Slow Adaptive Modulation |
| SC | Selection Combining |
| SE | Spectral Efficiency |
| SER | Symbol Error Rate |
| SIMO | Single-Input Multiple-Output |
| SISO | Single-Input Single-Output |
| SNR | Signal-to-Noise Ratio |
| STBC | Space-Time Block Code |
| STTC | Space-Time Trellis Code |
| TCM | Trellis-Coded Modulation |
| V-BLAST | Vertical Bell Labs Space-Time |
| WLAN | Wireless Local Area Network |
| ZM | Zero-Mean |
|  |  |

## Notations

| $x$ | variable |
| :--- | :--- |
| $\mathbf{x}$ | column vector |
| $\mathbf{X}$ | matrix |
| $x^{*}$ | complex conjugate of $x$ |
| $\Re[x], \Im[x]$ | real and imaginary part of $x$ |
| $\|x\|$ | absolute value of $x$ |
| $\operatorname{var}(x)$ | variance of a random variable $x$ |
| $\operatorname{cov}(x, y)$ | covariance of two random variables $x$ and $y$ |
| $\exp (x)$ | exponential function of $x$ |
| $\log (x)$ | (natural) logarithm of $x$ |
| $\\|\mathbf{x}\\|$ | norm of $\mathbf{x}$ |
| $(\mathbf{X})_{k, \ell}$ or $x_{k, \ell}$ | $(k, \ell)$-th element of $\mathbf{X}$ |
| $\operatorname{tr}(\mathbf{X})$ | trace of $\mathbf{X}$ |
| $\operatorname{det}(\mathbf{X})$ | determinant of $\mathbf{X}$ |
| $\operatorname{vec}(\mathbf{X})$ | vec operator (stacks columns of $\mathbf{X}$ in a large column vector) |
| $\mathbf{X}^{T}$ | transpose of $\mathbf{X}$ |
| $\mathbf{X}^{H}$ | Hermitian transpose of $\mathbf{X}$ |
| $\mathbf{X}^{-1}$ | inverse of $\mathbf{X}$ |

$\|\mathbf{X}\|_{\mathrm{F}} \quad$ Frobenius norm of $\mathbf{X}\left(\|\mathbf{X}\|_{\mathrm{F}}=\sqrt{\operatorname{tr}\left[\mathbf{X}^{H} \mathbf{X}\right]}\right)$
$\mathbf{X} \otimes \mathbf{Y} \quad$ Kronecker product of $\mathbf{X}$ and $\mathbf{Y}$
$\operatorname{diag}[\mathbf{x}]$ diagonal matrix with elements of vector $\mathbf{x}$ on diagonal
$\mathbf{I}_{N} \quad$ identity matrix of size $N \times N$
$\mathbb{E}[\cdot] \quad$ expectation operator
$\delta_{k} \quad$ discrete Dirac function $(k \in \mathbb{Z})$

## Samenvatting

Dankzij het comfort, de flexibiliteit en de mobiliteit die digitale draadloze communicatiesystemen, zoals laptops, tablet-pc's en smartphones, de gebruiker bieden, boomt de draadloze markt als nooit tevoren. Dit succes brengt echter met zich mee dat een steeds groeiend aantal draadloze toepassingen gebruik dient te maken van een eindig en beperkt spectrum. Om verdere verzadiging van het beschikbare spectrum te vermijden, worden voortdurend nieuwe en spectraal efficiënte technieken voorgesteld en onderzocht, met als doel het debiet en de betrouwbaarheid van draadloze systemen te verhogen.

Het is bekend dat de betrouwbaarheid van draadloze communicatiesystemen sterk wordt beperkt door multipath fading. Dit fenomeen wordt veroorzaakt door reflecties en verstrooiingen van het uitgezonden signaal op (bewegende) objecten in de omgeving, en resulteert in een tijdsafhankelijke attenuatie van het signaal tijdens de propagatie over het draadloze kanaal. Aan het reeds verzwakte signaal wordt bovendien thermische ruis toegevoegd door de ontvanger. Het spreekt voor zich dat wanneer het nuttig signaal sterk geattenueerd wordt door fading, de ontvanger het moeilijk zal hebben om de verstuurde informatie correct te reconstrueren uit het ontvangen signaal. Meestal wordt de kwaliteit van het signaal uitgedrukt aan de hand van de
signaal-ruisverhouding (signal-to-noise ratio of SNR), die wordt gedefinieerd als de verhouding van het vermogen van het nuttige signaal tot het vermogen van de achtergrondruis. Aangezien een lage SNR rechtstreeks aanleiding geeft tot een slechte performantie, wordt de betrouwbaarheid van draadloze transmissie vaak verhoogd door gebruik te maken van diversiteit. Op die manier wordt de ontvanger voorzien van meerdere replica's van hetzelfde signaal via verschillende onafhankelijke paden, zodat de kans dat alle replica's tegelijk sterk geattenueerd worden, zo klein mogelijk wordt. De verschillende diversiteitspaden kunnen op verschillende manieren worden gegenereerd, bijvoorbeeld in frequentie, tijd of ruimte. In dit proefschrift richten we ons op multiple-input multiple-output (MIMO) systemen die de ruimtelijke diversiteit benutten met behulp van spatio-temporele codering. In het bijzonder beschouwen we orthogonale spatio-temporele blokcodes (OSTBCs), die in staat zijn om een maximale diversiteit van $L=L_{t} L_{r}$ te behalen, waarbij $L_{t}$ en $L_{r}$ respectievelijk het aantal zend- en ontvangstantennes voorstellen. Bovendien herleidt optimale detectie van OSTBCs zich tot eenvoudige symbool-per-symbool detectie, enkel gebaseerd op lineaire signaalverwerking aan de ontvanger.

De belangrijkste maat om de performantie van een digitaal communicatiesysteem te evalueren is de gemiddelde bitfoutprobabiliteit (bit error rate of BER), die wordt gedefinieerd als de verhouding van het aantal fout gedetecteerde bits tot het totaal aantal verzonden bits. Hoewel de BER kan worden bepaald door het aantal bitfouten te tellen dat optreedt bij een stochastische computersimulatie van het systeem, geniet deze aanpak niet altijd de voorkeur. In geval van lage gemiddelde BER, bijvoorbeeld, zijn bijzonder veel systeemrealisaties nodig om een voldoende aantal bitfouten te observeren en een bepaalde nauwkeurigheid te garanderen. Omdat dit erg lange simulatietijden tot gevolg heeft, is een elegant en efficiënt alternatief voor computersimulaties nodig. Aangezien orthogonale spatio-temporele blokcodering aanleiding geeft tot zeer lage BERs, zelfs voor matige SNR, voeren we in dit proefschrift een nauwkeurige analytische BER studie uit voor OSTBCs op fading kanalen met een vlakke frequentiekarakteristiek. In de wetenschappelijke literatuur is de performantie van OSTBCs reeds uitgebreid onderzocht in de veronderstelling dat de toestand van het draadloze kanaal perfect gekend is aan de ontvanger. Vermits deze veronderstelling niet opgaat in een realistische draadloze omgeving, beschouwen we een OSTBC ontvanger die het kanaal schat aan de hand van gekende pilootsymbolen. Bovendien laten we ruimtelijke correlatie toe tussen de verschillende coëfficiënten van het MIMO kanaal. Op die manier geven de berekende BER uitdrukkingen systeemontwerpers de kans om nauwkeurig het effect te onderzoeken dat kanaalschattingsfouten en correlatie tussen de verschillende MIMO paden hebben op de performantie van OSTBCs. Hoewel de uitdrukkingen zijn verkregen voor de spectraal efficiënte quadrature amplitude modulation (QAM) techniek, zijn ze ook eenvoudig toepasbaar op pulse amplitude modulation (PAM). Om een nauwkeu-
rige modellering van verschillende draadloze omgevingen mogelijk te maken, omvat onze BER analyse meerdere fading distributies.

In geval van ongecorreleerde Rayleigh fading kanalen en linear minimum mean square error (LMMSE) kanaalschatting, presenteren we gesloten uitdrukkingen voor de BER. Deze uitdrukkingen zijn exact voor vierkante OSTBCs en blijken zeer nauwkeurig te zijn voor niet-vierkante OSTBCs. Ook de BER degradatie veroorzaakt door imperfecte kanaalschatting is gegeven in gesloten vorm. Voor gecorreleerde Rayleigh fading kanalen, geven we zeer nauwkeurige benaderingen van de BER in gesloten vorm, die asymptotisch exact zijn voor vierkante OSTBCs. Daarnaast leiden we een eenvoudige vuistregel af die nuttig is als indicator van de BER degradatie veroorzaakt door imperfecte kanaalschatting, en het exacte resultaat oplevert in geval van hoge SNR, vierkant OSTBCs, ongecorreleerde Rayleigh fading en constellaties met constante energie.

Wanneer we willekeurig verdeelde MIMO kanalen met maximum likelihood (ML) kanaalschatting veronderstellen, slagen we erin om de exacte BER te reduceren tot een verwachtingswaarde over niet meer dan $N_{\mathrm{S}}$ discrete toevalsgrootheden en $2 L+3$ reële continue toevalsgrootheden, waarbij $N_{s}$ het aantal informatiesymbolen in de beschouwde OSTBC matrix voorstelt. Voor vierkante OSTBCs kan het aantal reële continue toevalsgrootheden gereduceerd worden tot 4 , ongeacht het aantal antennes. Ook tonen we hoe de exacte BER uitdrukkingen efficiënt en nauwkeurig geëvalueerd kunnen worden door middel van numerieke integratietechnieken, zoals de kwadratuurregel en Monte-Carlo integratie met importance sampling, of een combinatie van beide. We gaan dieper in op de numerieke evaluatie van de exacte BER uitdrukkingen voor het specifieke geval van gecorreleerde Nakagami$m$ fading. De Nakagami- $m$ distributie wordt beschouwd als een veelzijdige distributiefunctie die niet enkel de Rayleigh distributie omvat voor $m=1$, maar ook verschillende andere fading omgevingen kan modelleren mits een geschikte keuze van de fading parameter $m$. Naast de exacte uitdrukkingen voor de BER, presenteren we ook een computationeel eenvoudige benadering van de BER, gebaseerd op het beschouwen van de symboolinterferentie te wijten aan imperfecte kanaalschatting als witte Gaussiaanse ruis. Hoewel de resulterende uitdrukking in het algemeen niet asymptotisch exact is, leidt ze tot zeer nauwkeurige BER resultaten wanneer de fading distributie niet al te erg afwijkt van de Rayleigh distributie en wanneer voldoende pilootsymbolen worden gebruikt.

Tot slot suggereren we hoe de gepresenteerde technieken en uitdrukkingen kunnen worden uitgebreid naar een aantal interessante MIMO OSTBC toepassingen die gebruik maken van andere of extra transmissie- of modulatietechnieken, zoals kanaalcodering, adaptieve modulatie en codering, of orthogonal frequency division multiplexing (OFDM).

## Summary

Due to the comfort, flexibility, and mobility offered by digital wireless communication systems, such as laptops, tablets, and smartphones, the wireless market is booming and an exponentially growing amount of wireless applications is competing for finite bandwidth resources. In order to avoid further saturation of the available spectrum, new and spectrally efficient techniques which increase the data rate or improve the reliability of wireless systems are continuously being proposed and investigated.

It is well-known that the reliability of wireless communication systems is strongly limited by multipath fading. This phenomenon is caused by reflections and scatterings of the transmitted signal on (moving) objects in the environment, and results in a time-varying attenuation of the information bearing signal. Moreover, when captured by the receiver, the already attenuated signal is further corrupted by thermal noise. It is readily verified that deep channel fades impede the receiver's ability to correctly retrieve the information from the received signal. Usually, the signal quality is expressed in terms of the signal-to-noise ratio (SNR), which is defined as the ratio of the power of the useful signal to the power of the background noise. Since a low SNR gives rise to bad system performance, the reliability of wireless transmission is of-
ten increased by using a proper diversity scheme. In this way, the receiver is provided with multiple replicas of the same signal through different independent paths, such that the probability that all replicas simultaneously suffer from severe fading is minimized. The different diversity paths can be generated in several ways, e.g., in frequency, time, or space. In this dissertation, we focus on multiple-input multiple-output (MIMO) systems exploiting spatial diversity by using space-time coding. More specifically, we consider orthogonal space-time block codes (OSTBCs), which are able to achieve a full diversity order of $L=L_{\mathrm{t}} L_{\mathrm{r}}$, with $L_{\mathrm{t}}$ and $L_{\mathrm{r}}$ denoting the number of transmit and receive antennas, respectively. Moreover, because of the orthogonality of the OSTBC matrices, the optimal detection algorithm reduces to symbol-by-symbol detection, based on linear signal processing at the receiver.

In digital communications, the principal figure of merit to evaluate the system performance is the bit error rate (BER), which is defined as the ratio of the number of erroneously received bits to the total number of bits. Although the BER can be obtained using a stochastic computer simulation of the system involving bit error counting, this approach is not always to be preferred. In case of low average BER, for instance, many simulation runs are required to observe a sufficient number of bit errors and ensure a given accuracy. Since this results in undesirably long simulation times, an elegant and efficient alternative for straightforward computer simulations is required. As, owing to their high diversity order, OSTBCs achieve very low BERs, even for moderate SNR, we provide in this dissertation an accurate analytical BER analysis for OSTBCs on flat-fading channels. In the literature, the performance of OSTBCs has been investigated extensively under the assumption that the channel state information (CSI) is perfectly known by the receiver. Since in a realistic wireless environment, the channel is a priori unknown and has to be estimated, we consider an OSTBC receiver that estimates the channel by means of known pilot symbols sent among the data. Moreover, we allow spatial correlation between the different coefficients of the MIMO channel. In this way, the presented BER expressions allow system designers to accurately examine the effect of channel estimation errors and fading correlation on the performance of OSTBCs. Although our expressions have been obtained for the spectrally efficient quadrature amplitude modulation (QAM) scheme, they are also easily applicable to pulse amplitude modulation (PAM) schemes. In order to enable an accurate modeling of different wireless environments, our BER analysis incorporates several fading distributions.

In case of uncorrelated Rayleigh fading channels and linear minimum mean square error (LMMSE) channel estimation, we present closed-form expressions for the BER, which are shown to be exact for square OSTBCs and very accurate for non-square OSTBCs. Also, a closed-form expression for the BER degradation due to imperfect channel estimation is provided. For arbitrarily correlated Rayleigh fading channels, we derive very accurate closed-form BER
approximations, which are asymptotically exact for square OSTBCs. In addition, we provide a simple rule of thumb that serves as an indicator for the BER degradation caused by imperfect channel estimation and yields the exact result in case of high SNR, square OSTBCs, uncorrelated Rayleigh fading channels, and equal-energy constellations.

Under the assumption of arbitrarily distributed MIMO channels with maximum likelihood (ML) channel estimation, we manage to reduce the exact BER to an expectation over $N_{\mathrm{S}}$ discrete random variables (RVs) and $2 L+3$ real-valued continuous RVs, with $N_{S}$ denoting the number of information symbols in the considered OSTBC matrix. For square OSTBCs, the number of real-valued continuous RVs can be reduced to 4 , regardless of the number of antennas. It is also shown how the presented exact BER expressions can be efficiently and accurately evaluated by means of numerical integration techniques, such as the quadrature rule and Monte-Carlo integration with importance sampling, or a combination thereof. We elaborate further on the numerical evaluation of the exact BER expressions for the specific case of correlated Nakagami- $m$ fading channels. The Nakagami- $m$ distribution is considered as a versatile statistical distribution that not only includes the Rayleigh distribution but is also able to accurately model a vast variety of fading environments by selecting a proper value for the fading parameter $m$. In addition to the exact BER expressions, we provide a computationally simple approximation of the BER, based on treating the symbol interference due to imperfect channel estimation as white Gaussian noise. Although the resulting expression is in general not asymptotically exact, it yields very accurate BER results when the fading distribution is similar to Rayleigh and when a sufficient number of pilot symbols is used.

Finally, we suggest how the presented techniques and performance results can be extended to a number of interesting MIMO OSTBC applications using different or additional transmission or modulation techniques, such as channel coding, adaptive modulation and coding, or orthogonal frequency division multiplexing (OFDM).

## 1

## Introduction

In this doctoral thesis, we investigate the impact of imperfect knowledge of the wireless channel on the performance of orthogonal space-time block codes (OSTBCs). We derive closed-form bit error rate (BER) expressions for OSTBCs operating over Rayleigh fading channels, whereas accurate and easy-toevaluate analytical BER expressions are obtained under generalized fading conditions. In section 1.1, we provide some interesting background information and explain why the presented research results are useful. An outline of this dissertation is provided in section 1.2.

### 1.1 Background and Motivation

In general, the purpose of any communication system is to transfer information from one point to another over a physical medium or link. In case of digital communication, the information to be sent is represented by a sequence of binary digits (bits) that take on the values ' 0 ' or ' 1 '. Hence, digital communication comprises any type of information that can be digitized, be it
text, data, voice or video. In communication theory, the medium over which the transmission takes place is called the channel. Examples of wired communication channels include copper wires and optical fibres. In case of wireless communication, however, the channel consists of the (air) space between the transmitter and receiver antennas. As opposed to wired communication, wireless technologies and applications provide the user with a great sense of comfort and mobility. For that reason, many important mobile technologies, such as mobile telephony and wireless local area networks (WLANs), have become indispensable and ubiquitous in our daily life. Moreover, due to the recent success of smartphones and tablets, more and more wireless applications requiring high-rate data transfer are entering our information sharing society. In order to satisfy the ever growing demand for bandwidth and at the same time not to overload the available spectrum, new and spectrally efficient wireless communication techniques are continuously being proposed and investigated.

Despite the appealing properties of digital wireless communication, system designers have to cope with one major challenge which is called multipath fading. This phenomenon is caused by reflections and scatterings of the transmitted signal on (moving) objects in the environment, and results in a time-varying attenuation of the information bearing signal. In other words, the wireless channel affecting the transmitted signal may vary considerably in time and its status cannot be known a priori. When captured by the receiver, the already attenuated signal is further corrupted by thermal noise. Usually, the signal quality is expressed in terms of the signal-to-noise ratio (SNR), which is defined as the ratio of the power of the useful signal to the power of the background noise. It is readily verified that multipath fading has a detrimental effect on the SNR of the received signal and, hence, on the performance of wireless communication systems. For example, when the channel is in a deep fade, the useful signal will be severely attenuated, and the presence of noise will make it impossible for the receiver to correctly retrieve the information from the received signal. Often, the reliability of wireless transmission is increased by using a proper diversity scheme, which provides the receiver with multiple replicas of the same signal through different independent paths. In this way, the probability that all replicas simultaneously suffer from severe fading is minimized.

One way to provide the receiver with multiple copies of the same signal consists of introducing multiple spatial paths between the transmitter and the receiver. To this end, both sides of the transmission link are equipped with multiple antennas. Although the first ideas on multiple-input multiple-output (MIMO) communication date back to the seventies [1-3], exciting results revealing the potential spectral efficiencies of MIMO systems were reported by Winters [4], Foschini [5], and Telatar [6]. A schematic example of a MIMO system with 2 transmit and 2 receive antennas is shown in Fig. 1.1. It is


Figure 1.1: A MIMO system with 2 transmit and 2 receive antennas.
important to note that the multiple virtual communication channels between the transmitter and receiver can be used both to increase the throughput by transmitting different data streams in parallel, a technique which is usually referred to as spatial multiplexing, and to improve the reliability of the transmission by exploiting spatial diversity; the latter technique requires the use of proper space-time coding. However, the throughput and the diversity order cannot be maximized simultaneously, as there exists a fundamental trade-off between diversity and multiplexing gain [7]. In this dissertation, we focus on the transmit diversity technique of orthogonal space-time block coding, which was introduced in 1999 by Tarokh et al. [8] as a generalization of Alamouti's remarkable transmit diversity scheme [9]. By coding the information symbols across different transmit antennas (space) and subsequent time slots (time), OSTBCs manage to achieve a full diversity order of $L=L_{t} L_{r}$, with $L_{t}$ and $L_{r}$ denoting the number of transmit and receive antennas, respectively, while their optimal detection algorithm reduces to symbol-by-symbol detection. Because of these beneficial properties, OSTBCs are considered a particularly interesting diversity technique.

Before technical advances as MIMO and OSTBCs are adopted in wireless communication standards that pave the way for future technologies, their performance needs to be carefully examined and compared with existing techniques. In digital communications, the principal figure of merit to evaluate and compare the system performance is the bit error rate, which is defined as the ratio of the number of erroneously received bits to the total number of bits. The BER of a system can be calculated either in an analytical fashion or using stochastic computer simulations. The latter approach requires that the input parameters of the system, such as the information symbols, the channel, and the noise samples are generated according to their corresponding statistical distributions. Consequently, the BER is obtained by applying the transmitter and receiver operations to the generated parameters and counting the number of erroneously received bits. As a result, in case of low average BER, extremely long simulation times are typically required to obtain a sufficient number of bit errors and achieve a given accuracy. Hence, for reliable commu-
nication systems achieving low BERs, efficient and easy-to-evaluate analytical BER results offer a pleasing and necessary alternative for brute-force computer simulations. However, in order to enable analytical performance analysis, a relatively simple channel model and data source model must be assumed.

Because of their high diversity order, OSTBCs achieve very low BERs, even for moderate SNR. Therefore, BER analysis through straightforward computer simulations is usually not appropriate and analytical BER results are to be preferred. In the literature, the performance of OSTBC s has been investigated extensively. Most results, however, were obtained under the assumption that the channel state information (CSI) is perfectly known by the receiver. In realistic receivers, this is clearly not the case and system designers have to consider the impact of imperfect channel estimation on the system performance. In this dissertation, we derive accurate BER expressions for a wireless receiver that estimates the channel by means of known pilot symbols sent among the data. We also allow spatial correlation between the MIMO subchannels. Our BER results allow to accurately examine the impact of both imperfect channel estimation and fading correlation on the BER performance of OSTBCs. Not only do our analytical BER expressions offer an elegant, accurate, and fast alternative for brute-force computer simulations, and enable system designers to select the optimal transmission parameters, they can also be used to verify the accuracy of the BER approximations that are currently available in the literature.

### 1.2 Outline

This dissertation is organized as follows:
Chapter 2 provides an overview of the fundamental principles of estimation and detection theory. This branch of statistics and signal processing is particularly important in communication theory, since the receiver operations required to extract the transmitted digital information and other unknown transmission parameters from the received signal, rely on it.

Chapter 3 presents a single-input single-output (SISO) digital wireless communication system and introduces the relevant terminology and notations. In this chapter, the different blocks of a conventional communication system are explained in more detail. We show how the digital information is modulated onto a continuous-time carrier wave and transmitted over the channel. Finally, it is illustrated how the transmitted information is recovered from the signal captured by the receiver and how the associated BER can be obtained analytically.

Chapter 4 shows how the exploitation of spatial diversity enables to mitigate multipath fading and improve the performance of wireless systems. This chapter presents the MIMO channel model that will be used throughout this thesis and introduces the concept of orthogonal space-time block coding.

Chapter 5 discusses two widespread pilot aided channel estimation methods. It also explains how pilot symbols are inserted in the data stream and why there is a trade-off between resources dedicated to pilot and data symbols.

Chapter 6 presents exact and approximate closed-form BER results for a receiver that estimates the CSI by means of pilot aided channel estimation. We also derive a simple rule of thumb that serves as an indicator for the BER degradation caused by imperfect channel estimation and yields the exact result under certain conditions.

Chapter 7 provides an exact BER analysis for square and non-square OSTBCs, under the assumption of arbitrarily distributed flat-fading channels with imperfect CSI. We show how the exact BER expressions can be efficiently and accurately evaluated using numerical integration techniques, and elaborate further on the case of correlated Nakagami- $m$ fading channels.

Chapter 8 illustrates how the expressions and techniques derived in this dissertation can be applied to assess the performance of MIMO OSTBC systems using different or additional transmission or modulation techniques. It shows both preliminary performance results and interesting ideas for further research.

Chapter 9 summarizes this dissertation and sums up the main conclusions. This chapter is followed by a complete list of our publications.

## 2

## Estimation and Detection Theory

In digital communications, the aim of the receiver is to recover the transmitted digital information from the received signal, which can be regarded as a probabilistic function of the transmitted signal. Moreover, before the receiver can retrieve this information, it needs to obtain the value of several unknown parameters, such as timing and channel parameters. Since the receiver operations for extracting both these signal parameters and the transmitted digital information rely on estimation and decision theory, we present in this chapter some fundamental definitions and principles of this particular branch of statistics and signal processing.

In section 2.1, we introduce the relevant definitions and statistical distributions. Using these distributions, we assess the estimation of continuous parameters and the detection of discrete parameters in sections 2.2 and 2.3, respectively. A chapter summary is given in section 2.4.

### 2.1 Definitions

Let us assume that we have an observation $\mathbf{y} \in \mathcal{Y}$, which depends on an unknown parameter $\mathbf{x} \in \mathcal{X}$ in a probabilistic manner. The purpose of estimation and detection theory is to derive an estimate $\hat{\mathbf{x}}(\mathbf{y})$ of the unknown $\mathbf{x}$ as a function of the observation $\mathbf{y}$. To this end, a statistical model is formulated for $\mathbf{x}$ and $\mathbf{y}$. When the parameter $\mathbf{x}$ is a discrete variable, i.e., it takes values from a finite set $\mathcal{X}$, estimation is called detection and the resulting estimate is called a decision.

In this dissertation, two important applications of estimation and detection theory are encountered, i.e., channel estimation and data detection. Chapter 5 is devoted to channel estimation and provides more details about the derivation and performance of different types of channel estimators. In this case, the unknown parameter $\mathbf{x}$ represents the channel state, and the observation $y$ consists of the signals captured by the receiver during the transmission of known pilot symbols. In the case of data detection, which is presented in section 3.2.3.3, the receiver detects the unknown transmitted information symbols based on the corresponding signals captured by the receiver. Here, $\mathbf{x}$ represents the (discrete) information symbol vector, whereas $\mathbf{y}$ consists of the corresponding received signals.

### 2.1.1 Likelihood Function

Since the unknown parameter $\mathbf{x}$ does not fully determine $\mathbf{y}$, we consider the probability density function (PDF)

$$
\begin{equation*}
p(\mathbf{y} \mid \mathbf{x}) \tag{2.1}
\end{equation*}
$$

of $\mathbf{y}$ conditioned on $\mathbf{x}$. This PDF specifies how likely the parameter value $\mathbf{x}$ gives rise to a given observation $\mathbf{y}$. When considered as a function of $\mathbf{x}$, this PDF is called the likelihood function of $\mathbf{x}$. It represents the knowledge we have about $\mathbf{x}$ as a result of the observation $\mathbf{y}$. The $\operatorname{logarithm} \log (p(\mathbf{y} \mid \mathbf{x}))$ of the likelihood function is called the log-likelihood function of $\mathbf{x}$.

### 2.1.2 Prior Distribution

Essentially, there are two ways to assess an estimation or detection problem. The standard or non-Bayesian approach considers the unknown parameter $\mathbf{x}$ to be deterministic and does not make any further a priori assumptions about it. The Bayesian approach, on the other hand, assumes $\mathbf{x}$ to be random and associates a prior distribution $p(\mathbf{x})$ to it, which represents the fundamental statistical knowledge about $\mathbf{x}$ before observing $\mathbf{y}$. Since the non-Bayesian approach can easily be incorporated in the Bayesian framework by associating a uniform prior distribution to $\mathbf{x}$, we will use the more general Bayesian approach throughout this chapter.

In principle, there is only one correct prior distribution $p(\mathbf{x})$ that characterizes the random parameter $\mathbf{x}$, yet it may be hard to obtain. Because of this argument the suitability of the Bayesian approach is sometimes questioned. However, a useful and accurate approximation of $p(\mathbf{x})$ can often be derived from either the physical constraints of the situation or from empirical measurements. In many situations it is preferred to represent the prior distribution by a standard distribution. This distribution should be chosen such that it both reduces the mathematical complexity of the problem and still sufficiently resembles the actual (empirical) distribution.

### 2.1.3 Posterior Distribution

The information about $\mathbf{x}$ provided by the prior distribution $p(\mathbf{x})$ and the likelihood function $p(\mathbf{y} \mid \mathbf{x})$ can be combined through the application of Bayes' rule

$$
\begin{equation*}
p(\mathbf{x} \mid \mathbf{y})=\frac{p(\mathbf{y} \mid \mathbf{x}) p(\mathbf{x})}{p(\mathbf{y})} \tag{2.2}
\end{equation*}
$$

where the posterior distribution $p(\mathbf{x} \mid \mathbf{y})$ represents all available information about $\mathbf{x}$. Note that the normalization factor $p(\mathbf{y})$ in (2.2) is given by

$$
\begin{equation*}
p(\mathbf{y})=\int_{\mathbf{x}} p(\mathbf{y} \mid \mathbf{x}) p(\mathbf{x}) \mathrm{d} \mathbf{x} \tag{2.3}
\end{equation*}
$$

and ensures that $\int_{\mathbf{x}} p(\mathbf{x} \mid \mathbf{y}) \mathrm{d} \mathbf{x}=1$.

### 2.2 Estimation

Let us consider in this section a continuous parameter $\mathbf{x}$. Based on the statistical distributions considered in section 2.1, various estimators $\hat{\mathbf{x}}(\mathbf{y})$ of the unknown parameter $\mathbf{x}$ can be derived. A common measure to evaluate and compare the performance of different estimators is the mean square error (MSE), which is defined as the expectation of the squared norm of the difference between the the estimate $\hat{\mathbf{x}}(\mathbf{y})$ and the actual parameter $\mathbf{x}$

$$
\begin{equation*}
\operatorname{MSE}=\mathbb{E}\left[\|\mathbf{x}-\hat{\mathbf{x}}(\mathbf{y})\|^{2}\right]=\int_{\mathbf{x}, \mathbf{y}}\|\mathbf{x}-\hat{\mathbf{x}}(\mathbf{y})\|^{2} p(\mathbf{x}, \mathbf{y}) \mathrm{d} \mathbf{x} \mathrm{~d} \mathbf{y} \tag{2.4}
\end{equation*}
$$

The MSE indicates how much $\hat{\mathbf{x}}(\mathbf{y})$ and $\mathbf{x}$ differ on average. Although numerous estimators are available from the literature, we give in this section an overview of the most important ones.

Minimum Mean Square Error Estimation The estimator that minimizes the MSE is called the minimum mean square error (MMSE) estimator. It is readily
verified that the MMSE estimate is obtained by minimizing, for every possible observation $\mathbf{y}$, the posterior mean square error

$$
\begin{equation*}
\hat{\mathbf{x}}_{\mathrm{MMSE}}(\mathbf{y})=\arg \min _{\tilde{\mathbf{x}} \in \mathcal{X}} \int_{\mathbf{x}}\|\mathbf{x}-\tilde{\mathbf{x}}\|^{2} p(\mathbf{x} \mid \mathbf{y}) \mathrm{d} \mathbf{x} \tag{2.5}
\end{equation*}
$$

Moreover, (2.5) can be easily shown to reduce to

$$
\begin{equation*}
\hat{\mathbf{x}}_{\mathrm{MMSE}}(\mathbf{y})=\mathbb{E}[\mathbf{x} \mid \mathbf{y}]=\int_{\mathbf{x}} \mathbf{x} p(\mathbf{x} \mid \mathbf{y}) \mathrm{d} \mathbf{x} \tag{2.6}
\end{equation*}
$$

such that the MMSE estimate is given by the posterior mean of the unknown parameter $\mathbf{x}$.

In many cases, a closed-form expression for the MMSE estimator is hard to obtain. A possible solution for this problem is to consider only the class of linear estimators, i.e., estimators of the form $\hat{\mathbf{x}}=\mathbf{M} \mathbf{y}^{1}$. The linear minimum mean square error (LMMSE) estimator is the linear estimator that minimizes the MSE. It follows from (2.4) that the LMMSE estimate is given by

$$
\begin{equation*}
\hat{\mathbf{x}}_{\text {LMMSE }}(\mathbf{y})=\mathbf{R}_{x y}\left(\mathbf{R}_{y y}\right)^{-1} \mathbf{y} \tag{2.7}
\end{equation*}
$$

where the cross-covariance matrix $\mathbf{R}_{x y}$ and the covariance matrix $\mathbf{R}_{y y}$ are defined as

$$
\begin{equation*}
\mathbf{R}_{x y}=\mathbb{E}\left[\mathbf{x} \mathbf{y}^{H}\right] \tag{2.8}
\end{equation*}
$$

where the superscript $H$ denotes the Hermitian transpose, and

$$
\begin{equation*}
\mathbf{R}_{y y}=\mathbb{E}\left[\mathbf{y} \mathbf{y}^{H}\right] . \tag{2.9}
\end{equation*}
$$

It is known that the MMSE estimator is linear if $\mathbf{x}$ and $\mathbf{y}$ are jointly Gaussian. Hence, in this case, the LMMSE and MMSE estimators coincide.

Maximum A Posteriori Estimation Another possibility to avoid the computational complexity associated with the MMSE estimate, is to use the maximum a posteriori (MAP) estimator. The MAP estimate selects the value of the unknown parameter $\mathbf{x}$ that maximizes its posterior distribution

$$
\begin{align*}
\hat{\mathbf{x}}_{\mathrm{MAP}}(\mathbf{y}) & =\arg \max _{\mathbf{x} \in \mathcal{X}} p(\mathbf{x} \mid \mathbf{y}) \\
& =\arg \max _{\mathbf{x} \in \mathcal{X}} p(\mathbf{y} \mid \mathbf{x}) p(\mathbf{x}) . \tag{2.10}
\end{align*}
$$

The second equation follows from (2.2) and the fact that $p(\mathbf{y})$ is independent of $\mathbf{x}$. It is easily understood that when the posterior distribution $p(\mathbf{x} \mid \mathbf{y})$ is symmetric and unimodal, the MAP and MMSE estimators are identical.

[^0]Maximum Likelihood Estimation A third widespread estimator is the maximum likelihood (ML) estimator, which selects the value of the unknown parameter $\mathbf{x}$ that maximizes its likelihood function

$$
\begin{equation*}
\hat{\mathbf{x}}_{\mathrm{ML}}(\mathbf{y})=\arg \max _{\mathbf{x} \in \mathcal{X}} p(\mathbf{y} \mid \mathbf{x}) . \tag{2.11}
\end{equation*}
$$

In contrast to the MMSE and MAP estimators, the ML estimator is strictly speaking not a Bayesian estimator, since it does not use any prior information about $\mathbf{x}$. However, the ML estimation rule (2.11) can be derived from the MAP estimation rule (2.10) by associating a uniform prior distribution to $\mathbf{x}$.

### 2.3 Detection

In detection problems, the parameter $\mathbf{x}$ belongs to a discrete set $\mathcal{X}$, whereas the observation y may be discrete or continuous. A typical application in digital communications is the detection of the transmitted bits from the continuous signal observed at the receiver.

Since a decision is either right or wrong, the MSE is no longer an appropriate means to compare the performance of different detectors. A better performance measure is the average probability of an erroneous decision

$$
\begin{align*}
P_{\text {err }} & =\operatorname{Pr}[\hat{\mathbf{x}}(\mathbf{y}) \neq \mathbf{x}] \\
& =1-\sum_{\mathbf{x}} \int_{\mathbf{y}} \mathbb{I}[\hat{\mathbf{x}}(\mathbf{y})=\mathbf{x}] p(\mathbf{x}, \mathbf{y}) \mathrm{d} \mathbf{y}, \tag{2.12}
\end{align*}
$$

where $\mathbb{I}[\cdot]$ denotes the indicator function, whose value is 1 if its argument is true and 0 otherwise.

Maximum A Posteriori Detection It is known that the error probability (2.12) is minimized by the MAP detector, which returns the value of $\mathbf{x}$ with the highest posterior probability

$$
\begin{equation*}
\hat{\mathbf{x}}_{\mathrm{MAP}}(\mathbf{y})=\arg \max _{\mathbf{x} \in \mathcal{X}} p(\mathbf{x} \mid \mathbf{y}) . \tag{2.13}
\end{equation*}
$$

Maximum Likelihood Detection The ML detector returns the value of $\mathbf{x}$ that maximizes the likelihood function

$$
\begin{equation*}
\hat{\mathbf{x}}_{\mathrm{ML}}(\mathbf{y})=\arg \max _{\mathbf{x} \in \mathcal{X}} p(\mathbf{y} \mid \mathbf{x}) . \tag{2.14}
\end{equation*}
$$

When $\mathbf{x}$ is uniformly distributed, MAP and ML detection are equivalent.

### 2.4 Chapter Summary

In this chapter, we introduced the fundamental principles of estimation and detection theory. Various widespread estimation and detection methods were
presented. In communication theory, these methods are applied to recover the transmitted digital information and other unknown transmission parameters from the received signal.

# Conventional Digital Communication 


#### Abstract

In this chapter, we present a conventional SISO digital wireless communication system and introduce the relevant terminology and notations. The more advanced MIMO system model, which will be used throughout this thesis, is presented in chapter 4. The purpose of any digital communication system is to transfer digital information from one point to another over a physical medium or channel. To this end, the digital information is modulated onto a radio frequency ( RF ) carrier wave, which is transmitted over the channel. As the transmitted electromagnetic or RF waves are band-pass signals, we devote section 3.1 to the mathematical representation of band-pass signals and systems. The different blocks of the communication system are explained in more detail in section 3.2. Section 3.3 illustrates how the BER can be calculated both analytically and through computer simulations, whereas section 3.4 summarizes the chapter.


### 3.1 Band-Pass Signals and Systems

### 3.1.1 Representation of Band-Pass Signals

A real-valued signal $s_{\mathrm{BP}}(t)$ is considered a band-pass signal if its Fourier transform or frequency response $S_{\mathrm{BP}}(f)$ satisfies $S_{\mathrm{BP}}(f)=0$ for $\left||f|-f_{\mathrm{C}}\right|>B$, with $f_{\mathrm{c}}>B$. In other words, $s_{\mathrm{BP}}(t)$ occupies an interval of length $2 B$ on both the positive and negative frequency axis. $f_{c}$ and $2 B$ are called the center frequency and the RF bandwidth of $s_{\mathrm{BP}}(t)$; usually $f_{\mathrm{c}} \gg B$. Since $s_{\mathrm{BP}}(t)$ is real-valued, it follows that $S_{\mathrm{BP}}(f)=S_{\mathrm{BP}}^{*}(-f)$.

The band-pass signal $s_{\mathrm{BP}}(t)$ can be related to an equivalent complex-valued baseband or low-pass signal $s_{\mathrm{LP}}(t)$, with $S_{\mathrm{LP}}(f)=0$ for $|f|>B$

$$
\begin{align*}
& s_{\mathrm{BP}}(t)=\sqrt{2} \Re\left[s_{\mathrm{LP}}(t) e^{j 2 \pi f_{\mathrm{c}} t}\right]  \tag{3.1a}\\
& s_{\mathrm{LP}}(t)=\left\{\sqrt{2} s_{\mathrm{BP}}(t) e^{-j 2 \pi f_{\mathrm{c}} t}\right\}_{\mathrm{LP}} \tag{3.1b}
\end{align*}
$$

where $\Re[\cdot]$ denotes the real part and $\{s(t)\}_{\text {LP }}$ denotes the components of $s(t)$ that are in the frequency band $[-B, B]$. Because of the normalization factor $\sqrt{2}$ in (3.1), the baseband signal $s_{\mathrm{LP}}(t)$ and the band-pass signal $s_{\mathrm{BP}}(t)$ have the same power. Note that multiplying a time domain signal with $e^{j 2 \pi f_{c} t}$ corresponds to a shift of $f_{\mathrm{c}} \mathrm{Hz}$ in the frequency domain. Hence, for the frequency responses of $s_{\mathrm{BP}}(t)$ and $s_{\mathrm{LP}}(t)$, we have

$$
\begin{align*}
& S_{\mathrm{BP}}(t)=\frac{\sqrt{2}}{2} S_{\mathrm{LP}}\left(f-f_{\mathrm{c}}\right)+\frac{\sqrt{2}}{2} S_{\mathrm{LP}}^{*}\left(-f-f_{\mathrm{c}}\right)  \tag{3.2a}\\
& S_{\mathrm{LP}}(t)=\sqrt{2} S_{\mathrm{BP}}\left(f+f_{\mathrm{c}}\right) \Pi_{\mathrm{LP}}(f) \tag{3.2b}
\end{align*}
$$

where $\Pi_{\mathrm{LP}}(f)$ is the frequency response of an ideal unit-amplitude low-pass filter with bandwidth $B$

$$
\Pi_{\mathrm{LP}}(f) \triangleq \begin{cases}1 & \text { if }|f| \leq B  \tag{3.3}\\ 0 & \text { otherwise }\end{cases}
$$

### 3.1.2 Representation of Linear Band-Pass Systems

Each linear band-pass system is characterized by a linear real-valued bandpass filter with impulse response $h_{\mathrm{BP}}(t)$ and frequency response $H_{\mathrm{BP}}(f)=$ 0 for $\left||f|-f_{c}\right|>B$. Let us now define a low-pass filter whose frequency response $H_{\mathrm{LP}}(f)=0$ is given by

$$
\begin{equation*}
H_{\mathrm{LP}}(t)=H_{\mathrm{BP}}\left(f+f_{\mathrm{c}}\right) \Pi_{\mathrm{LP}}(f) \tag{3.4}
\end{equation*}
$$

In this way, the impulse response $h_{\mathrm{LP}}(t)$ of the equivalent low-pass filter is related to $h_{\mathrm{BP}}(t)$ as

$$
\begin{align*}
& h_{\mathrm{BP}}(t)=2 \Re\left[h_{\mathrm{LP}}(t) e^{j 2 \pi f_{\mathrm{c}} t}\right]  \tag{3.5a}\\
& h_{\mathrm{LP}}(t)=\left\{h_{\mathrm{BP}}(t) e^{-j 2 \pi f_{\mathrm{c}} t}\right\}_{\mathrm{LP}} \tag{3.5b}
\end{align*}
$$

Note that (3.5) is similar to (3.1) except for the normalization factor and that the impulse response $h_{\mathrm{LP}}(t)$ is in general complex-valued.

### 3.1.3 Response of a Linear Band-Pass System to a Band-Pass Signal

If we apply the band-pass signal $s_{\mathrm{BP}}(t)$ to a linear band-pass filter with impulse response $h_{\mathrm{BP}}(t)$, the signal $r_{\mathrm{BP}}(t)$ at the output of the filter is also a band-pass signal and is given by the convolution of $s_{\mathrm{BP}}(t)$ and $h_{\mathrm{BP}}(t)$

$$
\begin{equation*}
r_{\mathrm{BP}}(t)=\int_{u} h_{\mathrm{BP}}(u) s_{\mathrm{BP}}(t-u) \mathrm{d} u \tag{3.6}
\end{equation*}
$$

In the frequency domain, we have

$$
\begin{equation*}
R_{\mathrm{BP}}(f)=H_{\mathrm{BP}}(f) S_{\mathrm{BP}}(f) \tag{3.7}
\end{equation*}
$$

such that $R_{\mathrm{BP}}(f)=0$ for $\left||f|-f_{\mathrm{c}}\right|>B$. Similar to the transmitted signal $s_{\mathrm{BP}}(t)$, the received signal $r_{\mathrm{BP}}(t)$ can be represented by an equivalent baseband signal $r_{\mathrm{LP}}(t)$. Using (3.5), it is readily verified that (3.6) and (3.7) each have their low-pass counterpart

$$
\begin{gather*}
r_{\mathrm{LP}}(t)=\int_{u} h_{\mathrm{LP}}(u) s_{\mathrm{LP}}(t-u) \mathrm{d} u  \tag{3.8}\\
R_{\mathrm{LP}}(f)=H_{\mathrm{LP}}(f) S_{\mathrm{LP}}(f) \tag{3.9}
\end{gather*}
$$

In section 3.2, it is shown how electromagnetic waves transfer digital information over a band-pass channel. However, we will derive a convenient baseband system model which makes abstraction of the RF signals and uses the baseband representations of the transmitted and received band-pass signals and the channel impulse response.

### 3.2 System Description

A conventional uncoded digital wireless communication system is depicted in Fig. 3.1. In general, we recognize three main blocks:

- the transmitter
- the channel


Figure 3.1: A conventional uncoded digital wireless communication system.

- the receiver

These three blocks and their subsystems are described in more detail in the following sections.

### 3.2.1 Transmitter

In digital communications, the information to be transferred over the channel is usually represented by a sequence of bits. We assume that the bits are independent and uniformly distributed, i.e., they take on the values 0 and 1 with equal probability. This assumption is valid when an efficient compression or source coding algorithm is applied to the data, and no channel coding is used. The purpose of the transmitter is to convert the given bit sequence into a continuous waveform, which can be transmitted onto the channel. This process is called digital modulation and consists of three stages: symbol mapping, pulse shaping, and upconversion.

### 3.2.1.1 Symbol Mapping

A symbol mapper translates a bit sequence $\left\{b_{k}\right\}$ into a sequence of complexvalued symbols $\left\{s_{k}\right\}$. To this end, the bit sequence is split into blocks of $m_{\mathrm{b}}$ bits, which are mapped to symbols belonging to a constellation $\Psi$ consisting of $M \triangleq 2^{m_{\mathrm{b}}}$ constellation points. The parameter $M$ is called the size or order of the constellation. Denoting by the superscript $T$ the transpose of a vector or matrix, the symbol sequence $\boldsymbol{s}=\left[s_{1}, s_{2}, \ldots, s_{N_{s}}\right]^{T}$ of $N_{\mathrm{s}}$ symbols represents a bit sequence $\mathbf{b}=\left[b_{1}, b_{2}, \ldots, b_{N_{\mathrm{b}}}\right]^{T}$ of $N_{\mathrm{b}}$ bits, with $N_{\mathrm{b}}=\log _{2}(M) N_{\mathrm{s}}$. When the mapping of a block of $m_{\mathrm{b}}$ bits to a symbol $s_{k}$ does not depend on previously transmitted symbols, the mapping is said to be memoryless.

Hence, each symbol $s_{k}$ can be generated through a bijective mapping function $\mathcal{M}:\{0,1\}^{m_{\mathrm{b}}} \rightarrow \Psi:$

$$
\begin{equation*}
s_{k}=\mathcal{M}\left(b_{(k-1) m_{\mathrm{b}}+1}, \ldots, b_{k m_{\mathrm{b}}}\right), \quad k=1, \ldots, N_{\mathrm{s}} \tag{3.10}
\end{equation*}
$$

Moreover, we assume that the symbol constellation $\Psi=\left\{\psi_{1}, \psi_{1}, \ldots, \psi_{M}\right\}$ is normalized, i.e.,

$$
\begin{equation*}
\mathbb{E}\left[\left|\psi_{i}\right|^{2}\right]=\frac{1}{M} \sum_{i=1}^{M}\left|\psi_{i}\right|^{2}=1 \tag{3.11}
\end{equation*}
$$

In practice, the following memoryless mapping strategies are often used:

## - M-ary Pulse Amplitude Modulation (M-PAM)

In case of $M$-PAM, the symbol constellation $\Psi$ is real-valued and given by

$$
\begin{equation*}
\Psi=\left\{(2 i-1-M) d_{\mathrm{PAM}}: i=1,2, \ldots, M\right\} \tag{3.12}
\end{equation*}
$$

where $d_{\text {PAM }}$ denotes half the distance between adjacent constellation points and is given by

$$
\begin{equation*}
d_{\mathrm{PAM}}=\sqrt{\frac{3}{M^{2}-1}} \tag{3.13}
\end{equation*}
$$

- M-ary Quadrature Amplitude Modulation (M-QAM)

In case of $M$-QAM, the symbol constellation $\Psi$ is complex-valued. We consider square $M$-QAM constellations such that the real and imaginary parts of the constellation points take values out of the set $\Psi^{\prime}$, which is given by

$$
\begin{equation*}
\Psi^{\prime}=\left\{(2 i-1-\sqrt{M}) d_{\mathrm{QAM}}: i=1,2, \ldots, \sqrt{M}\right\} \tag{3.14}
\end{equation*}
$$

where $d_{\mathrm{QAM}}$ is given by

$$
\begin{equation*}
d_{\mathrm{QAM}}=\sqrt{\frac{3}{2(M-1)}} \tag{3.15}
\end{equation*}
$$

Hence, the $M-Q A M$ constellation $\Psi$ is given by

$$
\begin{equation*}
\Psi=\left\{\psi: \Re[\psi], \Im[\psi] \in \Psi^{\prime}\right\} \tag{3.16}
\end{equation*}
$$

where $\Im[\cdot]$ denotes the imaginary part. Due to its high spectral efficiency (SE), M-QAM has been adopted in various standards, e.g., DVB [10], WLAN [11] and LTE [12].


Figure 3.2: 16-QAM constellation with Gray mapping.

- M-ary Phase Shift Keying (M-PSK)

In case of $M$-PSK, the symbol constellation $\Psi$ is complex-valued and given by

$$
\begin{equation*}
\Psi=\left\{\exp \left(j \frac{2 \pi(i-1)}{M}\right): i=1,2, \ldots, M\right\} \tag{3.17}
\end{equation*}
$$

Usually, 2-PSK and 4-PSK are called BPSK (binary PSK) and QPSK (quadrature PSK), respectively. Moreover, the BPSK and QPSK symbol constellations are identical to the constellations for 2-PAM and 4-QAM (rotated by $45^{\circ}$ ), respectively.

The assignment of $m_{\mathrm{b}}=\log _{2}(M)$ bits to $M$ constellation points can be done through many possible mapping functions. Because the mapping function affects the overall system performance, however, the selection of a proper mapping function is of particular importance. For uncoded transmission, it can be shown that the bit error rate is minimized for a mapping function which is such that constellation points at minimum Euclidean distance differ in exactly one bit. This mapping is called Gray mapping. Since an erroneous symbol detection due to noise will most likely result in the selection of an adjacent constellation point, only one bit in the detected $m_{\mathrm{b}}$-bit sequence will be incorrect in this case. An example of a 16-QAM constellation with Gray mapping is shown in Fig. 3.2. Note that the two least significant bits specify the real part of the constellation point, whereas the imaginary part depends on the two most significant bits only. The least and most significant bits are also called in-phase and quadrature bits, respectively.


Figure 3.3: Pulse shaping.

### 3.2.1.2 Pulse Shaping

We consider linear digital modulation, such that the transmitted continuoustime signal is a linear function of the discrete symbol symbol sequence $\left\{s_{k}\right\}$. To this end, a sequence of Dirac impulses at rate $R_{\mathrm{s}}=1 / T$ with weights $\left\{\sqrt{E_{\mathrm{S}}} s_{k}\right\}$ is fed to a baseband transmit filter, as shown in Fig. 3.3; $T$ is called the symbol period. The transmit filter is characterized by an impulse response $p(t)$ and a frequency response $P(f)$, with $P(f)=0$ for $|f|>B$. The parameter $B$ is called the baseband bandwidth of the filter. At its output the transmit filter produces a complex-valued baseband signal $s_{\mathrm{LP}}(t)$, which is given by

$$
\begin{equation*}
s_{\mathrm{LP}}(t)=\sqrt{E_{\mathrm{S}}} \sum_{k} s_{k} p(t-k T) \tag{3.18}
\end{equation*}
$$

where $E_{\mathrm{S}}$ is the energy per symbol, provided that $p(t)$ has unit energy, i.e.,

$$
\begin{equation*}
\int|p(t)|^{2} \mathrm{~d} t=\int|P(f)|^{2} \mathrm{~d} f=1 \tag{3.19}
\end{equation*}
$$

How the transmit pulse $p(t)$ should be selected is explained in section 3.2.3.2.

### 3.2.1.3 Upconversion

Since each wireless communication system is allowed to transmit within a limited frequency band only, the complex-valued baseband signal (3.18) has to be up-converted into a real-valued continuous-time band-pass signal. As shown in Fig. 3.4, the baseband signal $s_{\mathrm{LP}}(t)$ is modulated onto a sinusoidal carrier wave with frequency $f_{\mathrm{c}}>B$ according to the equivalence (3.1a), which yields the following band-pass signal

$$
\begin{equation*}
s_{\mathrm{BP}}(t)=\sqrt{2} \Re\left[s_{\mathrm{LP}}(t) e^{j 2 \pi f_{\mathrm{c}} t}\right] . \tag{3.20}
\end{equation*}
$$

This RF signal is fed to the transmit antenna and transmitted onto the channel.


Figure 3.4: Upconversion.

### 3.2.2 Channel

In wireless communications, the channel forms the physical medium over which the electromagnetic waves propagate. The way how the transmitted signal is affected along its way from the transmitter to the receiver is described mathematically by the channel impulse response. Because of reflections and scatterings on objects in the propagation environment, it is often assumed that the transmitted signal reaches the receiver through $N$ resolvable propagation paths [13], which are each characterized by a real-valued attenuation factor $\gamma_{n}$ and a time delay $\tau_{n}$, with $n=1,2, \ldots, N$. In this way, the received band-pass signal $r_{\mathrm{BP}}(t)$ is given by ${ }^{1}$

$$
\begin{equation*}
r_{\mathrm{BP}}(t)=\sum_{n=1}^{N} \gamma_{n} s_{\mathrm{BP}}\left(t-\tau_{n}\right)+w_{\mathrm{BP}}(t), \tag{3.21}
\end{equation*}
$$

where the thermal noise term $w_{\mathrm{BP}}(t)$ is assumed to be band-pass white noise [13] with power spectral density $\left(N_{0} / 2\right)\left|\Pi_{\mathrm{BP}}(f)\right|^{2}$, with

$$
\begin{equation*}
\Pi_{\mathrm{BP}}(f)=\Pi_{\mathrm{LP}}\left(f-f_{\mathrm{c}}\right)+\Pi_{\mathrm{LP}}\left(f+f_{\mathrm{c}}\right), \tag{3.22}
\end{equation*}
$$

denoting the frequency response of an ideal unit-amplitude band-pass filter centered around the carrier frequency $f_{\mathrm{c}}$ with passband bandwidth $2 B$. Note that in real communication systems, the bandwidth of the band-pass noise will be larger than $2 B$. Nevertheless, we neglect all noise components outside the considered frequency interval around $f_{\mathrm{c}}$, since they are rejected by the receiver's low-pass filter $\Pi_{\mathrm{LP}}(f)$ (see section 3.2.3.1). A typical example of a wireless multipath channel between a base station and a mobile user is shown in Fig. 3.5. Clearly, the transmitted signal reaches the receiver through different paths which experience different time delays and attenuations.

According to (3.1b), the received band-pass signal $r_{\mathrm{BP}}(t)$ can be represented by its equivalent complex-valued baseband signal $r_{\text {LP }}(t)$

$$
\begin{equation*}
r_{\mathrm{LP}}(t)=\sqrt{2}\left\{r_{\mathrm{BP}}(t) e^{-j 2 \pi f_{\mathrm{c}} t}\right\}_{\mathrm{LP}} . \tag{3.23}
\end{equation*}
$$

[^1]

Figure 3.5: Multipath channel between base station and mobile user.

Moreover, since it is easily seen from (3.21) that the channel acts as a linear filter, the equivalences in section 3.1 allow us to derive a baseband system model which makes abstraction of the band-pass nature of the electromagnetic waves and describes the relation between the received and transmitted baseband representations

$$
\begin{equation*}
r_{\mathrm{LP}}(t)=\int_{u} h(u) s_{\mathrm{LP}}(t-u) \mathrm{d} u+w_{\mathrm{LP}}(t) \tag{3.24}
\end{equation*}
$$

where $w_{\mathrm{LP}}(t)=\sqrt{2}\left\{w_{\mathrm{BP}}(t) e^{-j 2 \pi f_{\mathrm{c}} t}\right\}_{\mathrm{LP}}$ denotes low-pass filtered white noise with power spectral density $N_{0}\left|\Pi_{\mathrm{LP}}(f)\right|^{2}$ and $h(t)$ denotes the channel impulse response. From (3.21) and (3.23), it follows that the impulse response of the multipath channel can be written as

$$
\begin{equation*}
h(t)=\sum_{n=1}^{N} \alpha_{n} \delta\left(t-\tau_{n}\right) \tag{3.25}
\end{equation*}
$$

where $\alpha_{n} \triangleq \gamma_{n} e^{-j 2 \pi f_{c} \tau_{n}}$. Hence, each resolvable path is characterized by a complex attenuation $\alpha_{n}$ and a time delay $\tau_{n}$. Note that the frequency response $H(f)$ of the channel (3.25) is not limited to the interval $[-B, B]$. However, since the frequency response of $s_{\mathrm{LP}}(t)$ is zero outside the interval $[-B, B]$, only the frequency components of $H(f)$ within the interval $[-B, B]$ are relevant and it does not matter whether we use the channel (3.25) or a low-pass filtered version with frequency response $H_{\mathrm{LP}}(f)=H(f) \Pi_{\mathrm{LP}}(f)$ and corresponding impulse response

$$
\begin{equation*}
h_{\mathrm{LP}}(t)=(2 B) \sum_{n=1}^{N} \alpha_{n} \operatorname{sinc}\left(2 B\left(t-\tau_{n}\right)\right) \tag{3.26}
\end{equation*}
$$

where $\operatorname{sinc}(\cdot)$ denotes the sinc function, which is defined as

$$
\begin{equation*}
\operatorname{sinc}(x)=\frac{\sin (\pi x)}{\pi x} \tag{3.27}
\end{equation*}
$$

### 3.2.2.1 Frequency-Flat versus Frequency-Selective Channels

When the frequency response of a channel changes considerably over the frequency interval occupied by the transmitted signal, the channel is said to be frequency-selective. Frequency-flat or flat-fading channels, on the other hand, remain more or less constant over the considered interval.

The frequency response of the multipath channel (3.25) is given by

$$
\begin{equation*}
H(f)=\sum_{n=1}^{N} \alpha_{n} e^{-j 2 \pi f \tau_{n}} \tag{3.28}
\end{equation*}
$$

which is clearly a function of the frequency because of the non-zero time delays $\tau_{n}$. Let us assume that the path with index $n=1$ is the shortest path,
such that the associated time delay $\tau_{1}<\tau_{n}$, with $n=2, \ldots, N$. In this way, $\tau_{1}$ represents the propagation delay between transmitter and receiver and the channel (3.28) can be regarded as the cascade of two filters; the first filter is a unit-amplitude filter which accounts for the propagation delay and has frequency response $e^{-j 2 \pi f \tau_{1}}$, whereas the second filter's frequency response is given by

$$
\begin{equation*}
\bar{H}(f)=\sum_{n=1}^{N} \alpha_{n} e^{-j 2 \pi f \Delta \tau_{n}} \tag{3.29}
\end{equation*}
$$

where $\Delta \tau_{n} \triangleq \tau_{n}-\tau_{1}$ denotes the relative propagation delay of the $n$-th path with respect to the shortest path. Note that $\Delta \tau_{1}=0$ and $\Delta \tau_{n}>0$ for $n=2, \ldots, N$. The channel (3.29) can be considered to be frequency-flat when $\bar{H}\left(f_{1}\right) \approx \bar{H}\left(f_{2}\right)$, for any two frequencies $f_{1}$ and $f_{2}$ belonging to the frequency interval $[-B, B]$ occupied by the transmitted signal (3.18). Obviously, this is the case when $e^{-j 2 \pi\left|f_{1}-f_{2}\right| \Delta \tau_{n}} \approx 1$, i.e., when $\left|f_{1}-f_{2}\right| \ll 1 / \Delta \tau_{\text {max }}$, where

$$
\begin{equation*}
\Delta \tau_{\max }=\max _{n} \Delta \tau_{n} \tag{3.30}
\end{equation*}
$$

is called the delay spread and denotes the time delay difference between the longest and the shortest path. Usually, the coherence bandwidth $B_{\text {coh }}$ of the channel is defined as the inverse of the delay spread

$$
\begin{equation*}
B_{\mathrm{coh}} \triangleq \frac{1}{\Delta \tau_{\max }} \tag{3.31}
\end{equation*}
$$

Hence, a channel is said to be frequency-flat over the frequency interval $[-B, B]$ if $2 B \ll B_{\text {coh }}$. Frequency-flat channels can be characterized by a frequency response

$$
\begin{equation*}
H_{\text {flat }}(f)=\alpha e^{-j 2 \pi f \tau} \tag{3.32}
\end{equation*}
$$

which has a flat amplitude response $|H(f)|=|\alpha|$ and a phase response which is a linear function of the frequency. Note that the impulse response corresponding to (3.32) corresponds to a multipath channel with one path only

$$
\begin{equation*}
h_{\mathrm{flat}}(t)=\alpha \delta(t-\tau), \tag{3.33}
\end{equation*}
$$

### 3.2.2.2 Slow versus Fast Fading

In (3.25), the attenuations and time delays of the paths are assumed to be fixed, such that the channel is time-invariant. In reality, however, the channel parameters continuously vary over time, because of motion of the transmitter, motion of the receiver or motion of objects in the environment. We incorporate the time-variant nature of the channel by denoting its impulse response by $h(u ; t)$ and its frequency response by $H(f ; t)$. It is well-known that a timevarying channel affects the frequency content of the transmitted signal; this phenomenon is known as the Doppler effect. The larger the mobility of the

## CHAPTER 3. CONVENTIONAL DIGITAL COMMUNICATION

channel, the larger the Doppler frequency shift. For instance, if we assume a fixed transmitter and propagation environment, and a receiver which is moving at speed $v$, the maximum Doppler shift can be shown to be $f_{\mathrm{D}}=f_{\mathrm{c}}(v / c)$, where $f_{\mathrm{c}}$ and $c \approx 3 \cdot 10^{8} \mathrm{~m} / \mathrm{s}$ denote the center frequency of the transmitted signal and the speed at which electromagnetic waves travel through the air.

Taking into consideration that the maximum Doppler shift is related to the mobility of the channel, it can be shown that $H(f ; t) \approx H(f ; t+\Delta t)$ when $\Delta t \ll 1 / f_{\mathrm{D}}$. Let us define the coherence time $T_{\text {coh }}$ of the channel as the inverse of the maximum Doppler shift

$$
\begin{equation*}
T_{\mathrm{coh}} \triangleq \frac{1}{f_{\mathrm{D}}} \tag{3.34}
\end{equation*}
$$

In this way, a fading channel is considered to be time-invariant over a time interval of length $\Delta t \ll T_{\text {coh }}$. When the symbol period $T$ is much smaller than the coherence time $T_{\text {coh }}$, the channel is called a slowly fading channel and can be modelled by a time-invariant impulse response. Note that many definitions of $T_{\text {coh }}$ are available in the literature, yet the coherence time is always inversely proportional to the maximum Doppler shift.

### 3.2.3 Receiver

The purpose of the receiver is to recover the transmitted bits from the received band-pass signal $r_{\mathrm{BP}}(t)$. To this end, the receiver performs several steps which are depicted in Fig. 3.1 and are derived in the following sections under the assumption of a frequency-flat slowly fading channel with impulse response

$$
\begin{equation*}
h(t)=h \delta(t) \tag{3.35}
\end{equation*}
$$

### 3.2.3.1 Downconversion and Low-Pass Filtering

As already mentioned in section 3.2.2, the receiver first converts the received band-pass signal $r_{\mathrm{BP}}(t)$ into its equivalent complex-valued baseband representation (3.23). To this end, $r_{\mathrm{BP}}(t)$ is multiplied with $\sqrt{2} e^{-j 2 \pi f_{\mathrm{c}} t}$ and applied to an ideal low-pass filter with baseband bandwidth $B$, as shown in Fig. 3.6. The received baseband signal $r_{\mathrm{LP}}(t)$ is related to the transmitted baseband signal $s_{\mathrm{LP}}(t)$ through (3.24).

### 3.2.3.2 Matched Filtering and Sampling

As illustrated in Fig. 3.7, the received baseband signal $r_{\mathrm{LP}}(t)$ is fed to a receiving filter and then sampled at the symbol rate $R_{\mathrm{s}}=1 / T$. We use a receiving filter which is matched to the transmit pulse $p(t)$, i.e., the impulse and frequency responses of the filter are given by $p^{*}(-t)$ and $P^{*}(f)$, respectively. Since by definition the matched filter and the ideal low-pass filter from Fig. 3.6


Figure 3.6: Downconversion and low-pass filtering.


Figure 3.7: Matched filtering and sampling.
have the same bandwidth $B$, it follows that the frequency response of the cascade of both filters yields $\Pi_{\mathrm{LP}}(f) P^{*}(f)=P^{*}(f)$, such that the low-pass filter can actually be omitted in a practical receiver.

From (3.18), (3.24), and (3.35), it follows that the signal at the output of the matched filter is given by

$$
\begin{equation*}
r_{\mathrm{MF}}(t)=\sqrt{E_{\mathrm{s}}} h \sum_{k} s_{k} g(t-k T)+w_{\mathrm{MF}}(t) \tag{3.36}
\end{equation*}
$$

where $w_{\mathrm{MF}}(t)=\int w_{\mathrm{LP}}(u) p^{*}(u-t) \mathrm{d} u$, and the pulse $g(t)$ is given by

$$
\begin{equation*}
g(t)=\int p(u) p^{*}(u-t) \mathrm{d} u \tag{3.37}
\end{equation*}
$$

Note that taking the Fourier transform of (3.37) yields $G(f)=|P(f)|^{2}$, such that $G(f)=0$ for $|f|>B$. After sampling $r_{\mathrm{MF}}(t)$ at instants $t=k T$, we obtain the received symbol vector $\left\{r_{k}\right\}$

$$
\begin{equation*}
r_{k} \triangleq r_{\mathrm{MF}}(k T)=\sqrt{E_{\mathrm{s}}} h \sum_{n} s_{n} g(k T-n T)+w_{\mathrm{MF}}(k T) \tag{3.38}
\end{equation*}
$$

Since it follows from (3.19) that $g(0)=1$, we can rewrite $r_{k}$ as a function of the transmitted symbol $s_{k}$

$$
\begin{equation*}
r_{k}=\sqrt{E_{\mathrm{s}}} h s_{k}+\mathrm{ISI}_{k}+w_{k} \tag{3.39}
\end{equation*}
$$

where the disturbance terms $w_{k} \triangleq w_{\mathrm{MF}}(k T)$ and $\mathrm{ISI}_{k}$ represent Gaussian noise and inter-symbol interference (ISI), respectively, with

$$
\begin{equation*}
\mathrm{ISI}_{k}=\sqrt{E_{\mathrm{S}}} h \sum_{n \neq k} s_{n} g(k T-n T) \tag{3.40}
\end{equation*}
$$

The ISI (3.40) is eliminated when we choose a transmit pulse $p(t)$ which yields

$$
g(n T)= \begin{cases}1 & \text { if } n=0  \tag{3.41}\\ 0 & \text { otherwise }\end{cases}
$$

It is shown in [13, Sect. 9.2.1] that the condition (3.41) is equivalent to

$$
\begin{equation*}
\sum_{m=-\infty}^{\infty} G(f+m / T)=T \tag{3.42}
\end{equation*}
$$

which is known as the Nyquist pulse-shaping criterion or the Nyquist condition for zero ISI. According to (3.42), the sum of replicas of $G(f)$, separated by $1 / T$, should be a constant function. It is readily verified that this is only possible when $B \geq 1 /(2 T)$. A pulse $g(t)$ which satisfies (3.42) is called a Nyquist pulse.

When $B=1 /(2 T)$, the Nyquist criterion (3.42) is satisfied by only one $G(f)$, namely

$$
G(f)= \begin{cases}T & \text { if }|f| \leq B  \tag{3.43}\\ 0 & \text { otherwise }\end{cases}
$$

which corresponds to the scaled frequency response of an ideal low-pass filter with bandwidth $B$. The frequency response of the transmit pulse $p(t)$ yielding (3.43) is given by $P(f)=\sqrt{|G(f)|}$. Hence, $p(t)$ reduces to a (scaled) sinc pulse

$$
\begin{equation*}
p(t)=\frac{1}{\sqrt{T}} \operatorname{sinc}(t / T) \tag{3.44}
\end{equation*}
$$

Although the sinc pulse achieves the theoretical maximum bandwidth efficiency, it is not suited for practical filter design, because of its infinite length and infinitely steep frequency response $P(f)$.

Therefore, in reality, $B>1 /(2 T)$ and numerous candidates for $G(f)$ that satisfy the Nyquist criterion can be found. A particular spectrum that has been widely used because of its beneficial time-decay properties and smooth frequency characteristics is the raised cosine spectrum [13, Eq. (9.2-26)]. The transmit pulse $p(t)$ corresponding to this spectrum is called the square-root raised cosine pulse.

Using a proper Nyquist pulse $g(t)$ and corresponding transmit pulse $p(t)$, the received samples (3.39) reduce to

$$
\begin{equation*}
r_{k}=\sqrt{E_{\mathrm{s}}} h s_{k}+w_{k} \tag{3.45}
\end{equation*}
$$

where the noise samples $w_{k}$ can be shown to be zero-mean (ZM) circularly symmetric complex Gaussian (CSCG) random variables (RVs), the real and imaginary parts of which have variance $N_{0} / 2$. Note that (3.45) can be seen as a discrete-time system model for a frequency-flat slowly varying fading channel.

### 3.2.3.3 Detection

According to (3.45), each received sample $r_{k}$ is a function of the corresponding transmitted symbol $s_{k}$, the channel $h$, and a white Gaussian noise term $w_{k}$. Since the symbols $s_{k}$ to be estimated are independent and no ISI is present at the receiver, their detection can be performed symbol-by-symbol. Moreover, because all constellation points have equal prior probability, ML detection is optimal in the sense that it minimizes the error probability (2.12), which in this case reduces to the symbol error rate (SER), i.e., the ratio of the number of incorrectly detected symbols to the total number of transmitted symbols. From (3.45), it follows that the likelihood function of $s_{k}$ is given by

$$
\begin{equation*}
p\left(r_{k} \mid s_{k}\right)=\frac{1}{\pi N_{0}} \exp \left(-\frac{\left|r_{k}-\sqrt{E_{\mathrm{s}}} h s_{k}\right|^{2}}{N_{0}}\right) . \tag{3.46}
\end{equation*}
$$

Maximizing (3.46) yields the following ML decision rule

$$
\begin{equation*}
\hat{s}_{k}=\arg \min _{\tilde{s} \in \Psi}\left|u_{k}-\tilde{s}\right|^{2}, \tag{3.47}
\end{equation*}
$$

where the minimization is over all symbols belonging to the considered constellation $\Psi$ and the decision variable $u_{k}$ is given by

$$
\begin{equation*}
u_{k}=\frac{r_{k}}{\sqrt{E_{\mathrm{s}}} h} \tag{3.48}
\end{equation*}
$$

Taking (3.45) into account, the decision variable (3.48) can be written as

$$
\begin{equation*}
u_{k}=s_{k}+n_{k} \tag{3.49}
\end{equation*}
$$

where the noise term

$$
\begin{equation*}
n_{k}=\frac{w_{k}}{\sqrt{E_{\mathrm{s}}} h^{\prime}} \tag{3.50}
\end{equation*}
$$

is a ZM CSCG RV with variance $N_{0} /\left(E_{\mathrm{S}}|h|^{2}\right)$. Hence, the decision variable $u_{k}$ is the sum of the transmitted symbol $s_{k}$ and a Gaussian noise term $n_{k}$. In order to obtain $\hat{s}_{k}$, the receiver selects the constellation point at minimum Euclidean distance from the decision variable $u_{k}$. This allows us to associate a decision area $D_{m}$ to each of the constellation points $\psi_{m}, m=1, \ldots, M$

$$
\begin{equation*}
D_{m}=\left\{u:\left|u-\psi_{m}\right|^{2} \leq\left|u-\psi_{n}\right|^{2} \text { for } n \neq m\right\} . \tag{3.51}
\end{equation*}
$$

In this way, the constellation point $\psi_{m}$ is selected when the decision variable is located inside the corresponding decision area $D_{m}$. The decision areas of a 16-QAM constellation are depicted in Fig. 3.8. Because of the noise term $n_{k}$ in (3.49), the decision variable $u_{k}$ may be located in the decision area of a constellation point different from the transmitted symbol $s_{k}$, which results in a symbol detection error.


Figure 3.8: Decision areas for 16-QAM.

### 3.2.3.4 Demapping

Finally, the detected symbol sequence $\left\{\hat{s}_{k}\right\}$ is converted into a bit sequence $\left\{\hat{b}_{k}\right\}$ by the inverse of the mapping function (3.10)

$$
\begin{equation*}
\left[\hat{b}_{(k-1) m_{\mathrm{b}}+1}, \ldots, \hat{b}_{k m_{\mathrm{b}}}\right]=\mathcal{M}^{-1}\left(\hat{s}_{k}\right), \quad k=1, \ldots, N_{\mathrm{s}} . \tag{3.52}
\end{equation*}
$$

In case of errorless transmission, the bit sequence (3.52) is identical to the transmitted bit sequence. When a data symbol is detected erroneously, however, one or more of the associated bits will be incorrect.

### 3.3 Bit Error Rate

Since each data symbol represents a block of $\log _{2}(M)$ bits, it would be convenient to have a performance measure for the number of erroneously received bits, rather than for the incorrect symbol decisions. To this end, we introduce the bit error rate, which we define as the ratio of the average number of erroneously received bits per symbol to the number of bits per symbol.

### 3.3.1 Analytical BER Calculation

Using the ML decision rule (3.47), the BER is given by

$$
\begin{equation*}
P_{\mathrm{b}}=\frac{1}{\log _{2}(M)} \sum_{\substack{m, n=1 \\ m \neq n}}^{M} d_{\mathrm{H}}\left(\psi_{m}, \psi_{n}\right) \operatorname{Pr}\left[\hat{s}_{k}=\psi_{m} \mid s_{k}=\psi_{n}\right] \operatorname{Pr}\left[s_{k}=\psi_{n}\right] \tag{3.53}
\end{equation*}
$$

where $d_{\mathrm{H}}\left(\psi_{m}, \psi_{n}\right)$ denotes the Hamming distance between the bits associated to the constellation points $\psi_{m}$ and $\psi_{n}, \operatorname{Pr}\left[s_{k}=\psi_{n}\right]$ is the a priori probability of $\psi_{n}$, and $\operatorname{Pr}\left[\hat{s}_{k}=\psi_{m} \mid s_{k}=\psi_{n}\right]$ is the probability that the symbol $\psi_{m}$ is detected when $\psi_{n}$ is transmitted. Note that (3.53) does not depend on the value of the time index $k$, such that $k$ can be omitted for notational convenience in the remainder of this section. Moreover, taking into account that all symbols occur with equal probability, (3.53) reduces to

$$
\begin{equation*}
P_{\mathrm{b}}=\frac{1}{M \log _{2}(M)} \sum_{\substack{m, n=1 \\ m \neq n}}^{M} d_{\mathrm{H}}\left(\psi_{m}, \psi_{n}\right) \operatorname{Pr}\left[\hat{s}=\psi_{m} \mid s=\psi_{n}\right] . \tag{3.54}
\end{equation*}
$$

In order to obtain (3.54) analytically, the probabilities $\operatorname{Pr}\left[\hat{s}=\psi_{m} \mid s=\psi_{n}\right]$ have to be calculated given a specific channel model.

### 3.3.2 Monte-Carlo Simulations

When an analytical BER calculation is not mathematically feasible, the BER can also be obtained by straightforward Monte-Carlo simulations. In error analysis, these simulations are also often used to confirm the analytical BER curves. In order to obtain the BER for the system model (3.45) by Monte-Carlo simulations, the input RVs of the system, i.e., the channel $h$, the information symbols $s_{k}$, and the noise samples $w_{k}$, need to be generated repeatedly according to their corresponding distributions. For each set of input RVs, the receiver detects the information symbols from the received signal according to (3.47). Finally, the BER is obtained as the ratio of the number of bit errors to the total number of bits transmitted. Clearly, the smaller the number of observed bit errors, the less accurate the BER result will be. Hence, the accuracy of the BER can be improved by increasing the number of symbols transmitted. Note that in case of low average BER, the number of transmitted symbols required to obtain a certain accuracy may become very high, resulting in very long simulation times.

### 3.3.3 Example

As an example, we explicitly derive an analytical expression for the BER in case of $M$-QAM transmission. As will be shown in chapter 6, a similar approach can be used to obtain the BER in case of imperfect channel estimation. Let us introduce the real and imaginary parts of $\hat{s}$ and $u$ as $\hat{s}_{R}=\Re[\hat{s}], \hat{s}_{I}=\Im[\hat{s}]$, $u_{R}=\Re[u]$, and $u_{\mathrm{I}}=\Im[u]$. In this way, it follows directly from (3.14) and (3.47) that $\hat{s}_{\mathrm{R}}$ and $\hat{s}_{\mathrm{I}}$ can be obtained separately as

$$
\begin{align*}
& \hat{s}_{\mathrm{R}}=\arg \min _{\tilde{s} \in \Psi^{\prime}}\left|u_{\mathrm{R}}-\tilde{s}\right|^{2}  \tag{3.55a}\\
& \hat{s}_{\mathrm{I}}=\arg \min _{\tilde{s} \in \Psi^{\prime}}\left|u_{\mathrm{I}}-\tilde{s}\right|^{2}, \tag{3.55b}
\end{align*}
$$

where $\Psi^{\prime}$ is defined in (3.14) as the set consisting of the real (and imaginary) parts of $\Psi$. Since $\hat{s}_{R}$ and $\hat{s}_{I}$ determine the in-phase and quadrature bits associated to $\hat{s}$, respectively, the BER for M-QAM can be obtained as the average of the BERs of the in-phase and quadrature bits. Moreover, owing to the rotational symmetry of the QAM constellation and the circular symmetry of the noise term in (3.49), the BERs of the in-phase and quadrature bits are identical. Let us consider a QAM symbol $b=b_{\mathrm{R}}+j b_{\mathrm{I}}$, with $b_{\mathrm{R}}$ and $b_{\mathrm{I}}$ denoting the real and imaginary parts of $b$, respectively; we refer to the projections of the decision area of $b$ on the real and imaginary axis as the decision regions of $b_{R}$ and $b_{\mathrm{I}}$, respectively. When a QAM symbol $s$ is transmitted, a detection error occurs when $u_{\mathrm{q}}$, with $\mathrm{q}=\mathrm{R}$ or $\mathrm{q}=\mathrm{I}$, is located inside the decision area of $b_{\mathrm{q}} \neq s_{\mathrm{q}}$. Taking into account that the BER for M-QAM can be written as the BER of the in-phase bits, we have

$$
\begin{equation*}
P_{\mathrm{b}}=\frac{1}{\sqrt{M} \log _{2}(\sqrt{M})} \sum_{\substack{s_{\mathrm{R}}, b_{\mathrm{R}} \in \Psi^{\prime} \\ s_{\mathrm{R}} \neq b_{\mathrm{R}}}} d_{\mathrm{H}}\left(s_{\mathrm{R}}, b_{\mathrm{R}}\right) \operatorname{Pr}\left[\hat{s}_{\mathrm{R}}=b_{\mathrm{R}} \mid s_{\mathrm{R}}\right] \tag{3.56}
\end{equation*}
$$

where $d_{\mathrm{H}}\left(s_{\mathrm{R}}, b_{\mathrm{R}}\right)$ is the Hamming distance between the (in-phase) bits associated to the real part $s_{\mathrm{R}}$ of the transmitted symbol $s$ and the real part $b_{\mathrm{R}}$ of the detected symbol $b$.

We introduce $d_{1}\left(b_{\mathrm{q}}\right)$ and $d_{2}\left(b_{\mathrm{q}}\right)$ as the boundaries of the decision area of $b_{\mathrm{q}}$, with $d_{1}\left(b_{\mathrm{q}}\right)<d_{2}\left(b_{\mathrm{q}}\right)$. For outer constellation points, we set $d_{1}\left(b_{\mathrm{q}}\right) \rightarrow-\infty$ or $d_{2}\left(b_{\mathrm{q}}\right) \rightarrow \infty$. Using these boundaries, the conditional probability in (3.56) is easily shown to reduce to

$$
\begin{align*}
\operatorname{Pr}\left[\hat{s}_{\mathrm{R}}=b_{\mathrm{R}} \mid s_{\mathrm{R}}\right] & =\operatorname{Pr}\left[d_{1}\left(b_{\mathrm{R}}\right) \leq u_{\mathrm{R}} \leq d_{2}\left(b_{\mathrm{R}}\right) \mid s_{\mathrm{R}}\right] \\
& =Q\left(\frac{d_{1}\left(b_{\mathrm{R}}\right)-s_{\mathrm{R}}}{\sigma}\right)-Q\left(\frac{d_{2}\left(b_{\mathrm{R}}\right)-s_{\mathrm{R}}}{\sigma}\right) \tag{3.57}
\end{align*}
$$

where the second equality relies on (3.49), $\sigma^{2}=N_{0} /\left(2 E_{\mathrm{S}}|h|^{2}\right)$ is the variance of the real part of the noise term (3.50), and $Q(x)$ is the Gaussian $Q$-function

$$
\begin{equation*}
Q(x)=\int_{x}^{\infty} \frac{1}{\sqrt{2 \pi}} \exp \left(-\frac{y^{2}}{2}\right) \mathrm{d} y . \tag{3.58}
\end{equation*}
$$

Note that the $Q$-functions in (3.57) can have a negative argument, which is undesirable in some situations. Therefore, we introduce the following equivalent expression for (3.57)

$$
\begin{equation*}
\operatorname{Pr}\left[\hat{s}_{\mathrm{R}}=b_{\mathrm{R}} \mid s_{\mathrm{R}}\right]=Q\left(\frac{D_{1}\left(s_{\mathrm{R}}, b_{\mathrm{R}}\right)}{\sigma}\right)-Q\left(\frac{D_{2}\left(s_{\mathrm{R}}, b_{\mathrm{R}}\right)}{\sigma}\right) \tag{3.59}
\end{equation*}
$$

where $D_{1}\left(s_{\mathrm{R}}, b_{\mathrm{R}}\right)>0$ and $D_{2}\left(s_{\mathrm{R}}, b_{\mathrm{R}}\right)>0$ are defined as

$$
\begin{align*}
& D_{1}\left(s_{\mathrm{q}}, b_{\mathrm{q}}\right) \triangleq \begin{cases}s_{\mathrm{q}}-d_{2}\left(b_{\mathrm{q}}\right) & \text { if } s_{\mathrm{q}}>b_{\mathrm{q}} \\
d_{1}\left(b_{\mathrm{q}}\right)-s_{\mathrm{q}} & \text { otherwise }\end{cases}  \tag{3.60a}\\
& D_{2}\left(s_{\mathrm{q}}, b_{\mathrm{q}}\right) \triangleq \begin{cases}s_{\mathrm{q}}-d_{1}\left(b_{\mathrm{q}}\right) & \text { if } s_{\mathrm{q}}>b_{\mathrm{q}} \\
d_{2}\left(b_{\mathrm{q}}\right)-s_{\mathrm{q}} & \text { otherwise }\end{cases} \tag{3.60b}
\end{align*}
$$

Note that $D_{1}\left(s_{\mathrm{q}}, b_{\mathrm{q}}\right)$ and $D_{2}\left(s_{\mathrm{q}}, b_{\mathrm{q}}\right)$ denote the distances between $s_{\mathrm{q}}$ and the boundaries of the decision area of $b_{\mathrm{q}}$, with $D_{1}\left(s_{\mathrm{q}}, b_{\mathrm{q}}\right)<D_{2}\left(s_{\mathrm{q}}, b_{\mathrm{q}}\right)$; if $b_{\mathrm{q}}$ is an outer constellation point, we have $D_{2}\left(s_{\mathrm{q}}, b_{\mathrm{q}}\right) \rightarrow \infty$. Both (3.57) and (3.59) depend on the channel $h$ through $\sigma$. Hence, substituting (3.57) or (3.59) in (3.56) yields an expression for the conditional BER, conditioned on the channel. A more compact expression for the conditional BER can be straightforwardly obtained from [14, Eq. (8.14)] or [15, Eq. (14)]. Usually, the BER is plotted versus the energy per information bit $E_{\mathrm{b}}$, which is given by $E_{\mathrm{b}}=E_{\mathrm{s}} / \log _{2}(M)$. In case of 4-QAM, the conditional BER is shown to reduce to

$$
\begin{equation*}
P_{\mathrm{b}, 4-\mathrm{QAM}}(h)=Q\left(\sqrt{2 \frac{E_{\mathrm{b}}}{N_{0}}|h|^{2}}\right) . \tag{3.61}
\end{equation*}
$$

It is easily understood that the BER will be large when $|h|^{2}$ is very small, i.e., when the channel is in a deep fade. When $|h|^{2}$ approaches zero, the BER equals $1 / 2$; hence, the observation gives no useful information about the transmitted bits.

In order to obtain the average BER, the conditional BER needs to be averaged over the channel statistics. In section 4.3.2, the statistical modeling of channel coefficients is assessed and several widespread fading distributions are provided. In this example, we use the common assumption that the channel norm $|h|$ is distributed according to the Rayleigh distribution, with $E\left[|h|^{2}\right]=1$

$$
\begin{equation*}
p_{|h|}(x)=2 x \exp \left(-x^{2}\right) \tag{3.62}
\end{equation*}
$$

such that the squared norm $|h|^{2}$ has an exponential distribution

$$
\begin{equation*}
p_{|h|^{2}}(x)=\exp (-x) \tag{3.63}
\end{equation*}
$$

A closed-form solution for the BER for M-QAM under Rayleigh fading can be found by using the following identity to average the $Q$-functions in (3.57) over $|h|^{2}$

$$
\begin{equation*}
\int_{0}^{\infty} Q(\sqrt{\beta x}) \exp (-x) \mathrm{d} x=\frac{1}{2}\left(1-\sqrt{\frac{\beta}{2+\beta}}\right) \tag{3.64}
\end{equation*}
$$



Figure 3.9: BER for $M$-QAM under Rayleigh fading, with $M \in\{4,16,64,256\}$.
where the parameter $\beta>0$, and by noting that $Q(-x)=1-Q(x)$. In case of 4-QAM, the average BER reduces to

$$
\begin{equation*}
P_{\mathrm{b}, 4-\mathrm{QAM}}=\frac{1}{2}\left(1-\sqrt{\frac{\frac{E_{\mathrm{b}}}{N_{0}}}{1+\frac{E_{\mathrm{b}}}{N_{0}}}}\right) . \tag{3.65}
\end{equation*}
$$

For high $E_{\mathrm{b}} / N_{0}$, a series expansion of (3.65) yields $P_{\mathrm{b}, 4-\mathrm{QAM}} \approx N_{0} /\left(4 E_{\mathrm{b}}\right)$, such that the asymptotic behavior of the BER is inversely proportional to $E_{\mathrm{b}} / N_{0}$. Fig. 3.9 shows the analytical BER curves for $M$-QAM, with $M \in$ $\{4,16,64,256\}$. Also shown in the figure are the results from Monte-Carlo simulation. Note that all BER curves are asymptotically inversely proportional to $E_{\mathrm{b}} / N_{0}$. Exact and accurate approximate BER expression for M-QAM over Rayleigh fading channels can be found in, e.g., [14, sec. 8.2.1.2] and [16,17].

It is worth mentioning that in order to simplify the analytical averaging of the conditional BER over the channel statistics, the Gaussian $Q$-functions are often replaced by exponential functions according to the so-called Chernoff bound [14]:

$$
\begin{equation*}
Q(x) \leq \frac{1}{2} \exp \left(-\frac{1}{2} x^{2}\right) \tag{3.66}
\end{equation*}
$$

For example, by applying the Chernoff bound to (3.61), an upper bound on (3.65) is readily obtained as

$$
\begin{equation*}
P_{\mathrm{b}, 4-\mathrm{QAM}} \leq \frac{1}{2\left(1+\frac{E_{\mathrm{b}}}{N_{0}}\right)} \tag{3.67}
\end{equation*}
$$

which for high $E_{\mathrm{b}} / N_{0}$ differs from the exact result by a factor of 2 . However, as we aim to derive an accurate set of BER expressions, we maintain the Gaussian $Q$-functions throughout our analysis.

### 3.4 Chapter Summary

In this chapter, we have presented the fundamentals of uncoded digital wireless communication. In general, a communication systems consists of three main blocks: the transmitter, the channel, and the receiver.

- The transmitter maps the bits representing the digital information to symbols, modulates the symbols onto a continuous-time carrier wave and transmits the resulting RF signal on the channel.
- The channel is the physical transmission medium over which the RF waves propagate. Because of scattering and reflections in the environment, the channel transforms the transmitted signal on its way to the receiver. The resulting variations of the received signal strength are called fading. We distinguish between frequency-flat and frequency-selective channels, and between slow and fast fading channels.
- The aim of the receiver is to extract the transmitted information from the signal captured by the receive antenna. This process is called detection and consists of several steps: downconversion, low-pass filtering, matched filtering, sampling, symbol detection and demapping.

Throughout this dissertation, we assume slow frequency-flat fading, in which case we can make abstraction of many of the transmitter and receiver operations by using the straightforward discrete-time channel model (3.45). To conclude this chapter, we introduced the bit error rate, which serves as an important performance measure and can be obtained analytically or through Monte-Carlo simulations.

## 4

## MIMO Communication

In chapter 3, we introduced a conventional uncoded SISO communication system and assessed its error performance. It was shown in section 3.3.3 that the probability of detection errors to occur is high when the channel is in a deep fade. In section 4.1, we illustrate how the exploitation of spatial diversity allows to tackle this problem and thus improves the performance of wireless systems. Furthermore, we introduce the appealing concept of orthogonal space-time block coding in section 4.2. In section 4.3, we present the MIMO channel model that will be used throughout this thesis, and discuss its statistical properties. Section 4.4 wraps up this chapter.

### 4.1 Diversity

In wireless communications, the detrimental effect of deep channel fades on the system performance can be tackled by using a proper diversity scheme [14,18], which provides the receiver with multiple replicas of the same signal through different propagation paths. Ideally, these paths are affected by
independent fading such that the probability of the received replicas to simultaneously suffer from deep fading is minimized. At the receiver, the multiple replicas are combined using a proper diversity combining scheme in order to minimize the resulting error probability. In practice, there are several ways to provide the receiver with $L$ independent replicas of the same signal. The most commonly used diversity techniques are the following:

- The $L$ signal replicas are transmitted on $L$ carriers which are separated by at least the coherence bandwidth $B_{\text {coh }}$ of the channel. This technique is called frequency diversity. Note that by applying frequency diversity the occupied bandwidth increases by a factor $L$.
- A second method consists of employing time diversity, where the same signal is transmitted in $L$ different time slots which are separated by at least the coherence time $T_{\text {coh }}$ of the channel. Consequently, the spectral efficiency of the system decreases by a factor $L$.
- By using multiple antennas at the transmitter and/or receiver, spatial diversity can be exploited. It is important to note that the multiple antennas should be sufficiently separated in order for the different channels between the transmit and receive antennas to undergo independent fading; usually, a separation in the order of one wavelength is sufficient.
- In case of frequency-selective fading, i.e., when the bandwidth of the transmitted signal exceeds the coherence bandwidth $B_{\text {coh }}$, multipath diversity can be exploited by resolving the multipath components arriving with different delays. The optimum multipath receiver was invented by Price and Green in 1958 and is called the RAKE demodulator [19].

In section 3.3.3, we have derived that in case of Rayleigh fading a conventional uncoded digital communication system without diversity achieves a BER which is asymptotically inversely proportional to $E_{\mathrm{b}} / N_{0}$. The exploitation of diversity, however, allows to increase the rate of descent of the BER curve. Therefore, in the context of error probability, the term diversity gain or diversity order is often used to denote the increase in the slope of the error probability curve. That is, when the BER of a reference system without diversity behaves asymptotically as $P_{\mathrm{b}} \propto\left(E_{\mathrm{b}} / N_{0}\right)^{-\zeta}$ and exploiting diversity gives rise to $P_{\mathrm{b}} \propto\left(E_{\mathrm{b}} / N_{0}\right)^{-\zeta G_{\mathrm{d}}}$, the exponent $G_{\mathrm{d}}$ represents the diversity gain. For a conventional uncoded communication system without diversity, the diversity gain $G_{d}=1$, whereas the use of one or more diversity techniques gives rise to a diversity gain $G_{d}>1$. Moreover, we have illustrated in section 3.3.3 that $\zeta=1$ in case of Rayleigh fading.

In this dissertation, we focus on the use of spatial diversity to enhance the performance of wireless systems. Depending on whether the signal replicas are produced by the receiver or the transmitter, we distinguish between re-
ceive and transmit diversity. Eventually, receive and transmit diversity can be combined to achieve maximal spatial diversity.

### 4.1.1 Receive Diversity

In chapter 3 , we derived the discrete-time baseband system model (3.45) for an uncoded SISO system without diversity. When extending this model to a receiver equipped with $L_{r}$ receive antennas, the received signal samples on the $n$-th receive antenna are given by

$$
\begin{equation*}
r_{n}=\sqrt{E_{\mathrm{s}}} h_{n} s+w_{n} \tag{4.1}
\end{equation*}
$$

where $n=1, \ldots, L_{\mathrm{r}}$. Note that we have omitted the time index $k$ for notational convenience. In (4.1), the complex-valued channel coefficient $h_{n}$ characterizing the link between the transmit antenna and the $n$-th receive antenna is assumed to remain constant during a frame of $K_{\text {fr }}$ symbols. The ZM CSCG noise samples $w_{n}$ are assumed to be independent and identically distributed (i.i.d.), the real and imaginary parts of which have variance $N_{0} / 2$. Let us gather the received samples, the channel coefficients and the noise samples in the $L_{r}$-dimensional column vectors

$$
\begin{align*}
\mathbf{r} & =\left[r_{1}, \ldots, r_{L_{\mathrm{r}}}\right]^{T},  \tag{4.2}\\
\mathbf{h} & =\left[h_{1}, \ldots, h_{L_{\mathrm{r}}}\right]^{T},  \tag{4.3}\\
\mathbf{w} & =\left[w_{1}, \ldots, w_{L_{\mathrm{r}}}\right]^{T} . \tag{4.4}
\end{align*}
$$

In this way, the single-input multiple-ouput (SIMO) system model can be written as

$$
\begin{equation*}
\mathbf{r}=\sqrt{E_{\mathrm{s}}} \mathbf{h} s+\mathbf{w} \tag{4.5}
\end{equation*}
$$

where the SIMO channel $\mathbf{h}$ is assumed to be normalized

$$
\begin{equation*}
\mathbb{E}\left[\|\mathbf{h}\|^{2}\right]=L_{\mathrm{r}} \tag{4.6}
\end{equation*}
$$

When the $L_{\mathrm{r}}$ received signal samples are combined in a proper way, the SIMO receiver can take advantage of the spatial diversity and a better error performance can be obtained as compared to SISO reception. When applying ML detection to the received signal (4.5), the following optimal decision rule is derived

$$
\begin{equation*}
\hat{s}=\arg \min _{\tilde{s} \in \Psi}|u-\tilde{s}|^{2} \tag{4.7}
\end{equation*}
$$

where the minimization is over all symbols belonging to the considered constellation $\Psi$ and the decision variable $u$ is given by

$$
\begin{equation*}
u=\frac{\mathbf{h}^{H} \mathbf{r}}{\sqrt{E_{\mathrm{s}}}\|\mathbf{h}\|^{2}} \tag{4.8}
\end{equation*}
$$



Figure 4.1: Discrete-time SIMO system model.

Since (4.8) stems from the ML decision rule, it represents the optimal diversity combining scheme, which is known as maximal-ratio combining (MRC) [14]. Note that also other combining schemes, such as equal gain combining (EGC) [20] or selection combining (SC) [21] could be used to combine the multiple received signal samples, yet the resulting performance would be suboptimal. In Fig. 4.1, an uncoded SIMO system with diversity combining is depicted. In case of MRC, the decision variable (4.8) reduces to

$$
\begin{equation*}
u=s+\frac{\mathbf{h}^{H} \mathbf{w}}{\sqrt{E_{\mathbf{s}}}\|\mathbf{h}\|^{2}} \tag{4.9}
\end{equation*}
$$

where the right term in the sum can be shown to be a ZM CSCG noise term with variance $N_{0} /\left(E_{\mathrm{S}}\|\mathbf{h}\|^{2}\right)$. Since the only difference between (4.9) and (3.49) is the additive noise term in the decision variable, the variance of which is given by $N_{0} /\left(E_{\mathrm{S}}\|\mathbf{h}\|^{2}\right)$ in (4.9) and by $N_{0} /\left(E_{\mathrm{S}}|h|^{2}\right)$ in (3.49), it is readily verified that the BER for a SIMO receiver with MRC can be calculated in a similar way as in the case of SISO communication. Hence, for QAM constellations, the conditional BER is again obtained from (3.56) and (3.57), except that $\sigma^{2}$ in (3.57) is now given by $N_{0} /\left(2 E_{s}\|\mathbf{h}\|^{2}\right)$ and the resulting BER expression needs to be averaged over the distribution of the squared channel norm $\|\mathbf{h}\|^{2}$.

In order to get more insight in the effect that receive diversity has on the system's error performance, we calculate the average BER for an MRC system operating over $L_{r}$ i.i.d. Rayleigh fading channels, with $\mathbb{E}\left[\left|h_{n}\right|^{2}\right]=1$ for $n=$ $1, \ldots, L_{\mathrm{r}}$. In this case, the squared channel norm $\|\mathbf{h}\|^{2}$ is distributed according
to the $\chi^{2}$-distribution with $2 L_{r}$ degrees of freedom [13, eq. (14.4-13)]

$$
\begin{equation*}
p_{\|\mathbf{h}\|^{2}}(x)=\frac{1}{\left(L_{\mathrm{r}}-1\right)!} x^{L_{\mathrm{r}}-1} \exp (-x) \tag{4.10}
\end{equation*}
$$

A closed-form solution for the BER for M-QAM under Rayleigh fading can be found by using the following identity to average the $Q$-functions in (3.57) over $\|\mathbf{h}\|^{2}$ [13, eq. (14.4-15)]

$$
\begin{align*}
& \frac{1}{\left(L_{\mathrm{r}}-1\right)!} \int_{0}^{\infty} Q(\sqrt{\beta x}) x^{L_{\mathrm{r}}-1} \exp (-x) \mathrm{d} x \\
& \quad=\left[\frac{1}{2}\left(1-\sqrt{\frac{\beta}{2+\beta}}\right)\right]^{L_{\mathrm{r}}} \sum_{k=0}^{L_{\mathrm{r}}-1}\binom{L_{\mathrm{r}}-1+k}{k}\left[\frac{1}{2}\left(1+\sqrt{\frac{\beta}{2+\beta}}\right)\right]^{k} \tag{4.11}
\end{align*}
$$

where $\beta>0$. In case of 4-QAM, the average BER for an MRC receiver with $L_{r}$ receive antennas reduces to

$$
\begin{align*}
P_{\mathrm{b}, 4-\mathrm{QAM}}=\left[\frac{1}{2}(1-\right. & \left.\left.\sqrt{\frac{\frac{E_{\mathrm{b}}}{N_{0}}}{1+\frac{E_{\mathrm{b}}}{N_{0}}}}\right)\right]^{L_{\mathrm{r}}} \\
& \times \sum_{k=0}^{L_{\mathrm{r}}-1}\binom{L_{\mathrm{r}}-1+k}{k}\left[\frac{1}{2}\left(1+\sqrt{\frac{\frac{E_{\mathrm{b}}}{N_{0}}}{1+\frac{E_{\mathrm{b}}}{N_{0}}}}\right)\right]^{k} \tag{4.12}
\end{align*}
$$

which for high $E_{\mathrm{b}} / N_{0}$ reduces to [13, Eq. 14.4-18]

$$
\begin{equation*}
P_{\mathrm{b}, 4-\mathrm{QAM}}^{(\mathrm{as})} \approx\left(\frac{N_{0}}{4 E_{\mathrm{b}}}\right)^{L_{\mathrm{r}}}\binom{2 L_{\mathrm{r}}-1}{L_{\mathrm{r}}} \tag{4.13}
\end{equation*}
$$

such that the asymptotic behaviour of the BER is proportional to $\left(E_{\mathrm{b}} / N_{0}\right)^{-L_{\mathrm{r}}}$. Since the diversity gain $L_{\mathrm{r}}$ determines the slope of the error probability curve, it is clear that exploiting receive diversity allows to improve the system performance significantly. Moreover, by using multiple receive antennas, more energy is captured by the receiver even though the transmit energy remains unchanged. This so-called array gain further improves the performance of a SIMO receiver. Whereas the diversity gain determines the slope of the BER curve, the array gain determines the horizontal shift of the BER. In case of an array of length $L_{r}$, the array gain equals $L_{r}$ since on average $L_{r}$ times more energy is captured as compared to the SISO case. Hence, the BER curve is shifted to the left over an amount of $10 \log _{10}\left(L_{r}\right)$ dB. Fig. 4.2 shows the BER for an MRC receiver with $L_{\mathrm{r}}=1, \ldots, 5$ and 4-QAM over i.i.d. Rayleigh fading channels.


Figure 4.2: BER for MRC with $L_{\mathrm{r}}=1, \ldots, 5$ and 4-QAM over i.i.d. Rayleigh fading channels.

### 4.1.2 Transmit Diversity versus Spatial Multiplexing

In the previous section, we have shown that using multiple antennas at the receiver side enables to achieve spatial diversity and, thus, improves the reliability without expanding the required transmission bandwidth or reducing the spectral efficiency of the system. Employing multiple antennas at the transmitter side in conjunction with a proper transmit diversity scheme, however, can also exploit spatial diversity. For example, using Alamouti's code [9], it is possible to achieve dual diversity with two transmit antennas and one receive antenna. In general, by introducing redundancy in the spatial and the temporal domain, a MIMO system with $L_{t}$ transmit and $L_{r}$ receive antennas can achieve a maximum diversity gain of $G_{d}=L_{t} L_{r}$ [7], which corresponds to the number of independently faded paths between the different transmit and receive antennas. Note that unlike the case of receive diversity, the redundancy resulting from space-time coding may have a negative impact on the system's spectral efficiency.

Except for providing diversity, multiple transmit antennas can also be applied to increase the system's data rate. To this end, multiple independent information streams are transmitted in parallel through different spatial chan-
nels created by the multiple-antenna transmitter; this is often referred to as spatial multiplexing. With $L_{t}$ transmit and $L_{\mathrm{r}}$ receive antennas, e.g., it is possible to achieve an $L_{\mathrm{t}}$-fold increase in the data rate along with a receive diversity gain of $G_{d}=L_{\mathrm{r}}$. Examples of practical spatial multiplexing schemes are the diagonal Bell Labs space-time (D-BLAST) [5] and vertical Bell Labs space-time (V-BLAST) [22] architectures. It is important to note that these schemes mainly focus on maximizing the data rate and do not achieve full spatial diversity. In general, it turns out to be impossible to maximize the data rate and the diversity simultaneously. Intuitively, the reason for this is that transmit diversity techniques require a certain redundancy in the space-time domain, whereas high data rates can only be achieved when this redundancy is absent. The fundamental trade-off between diversity and multiplexing gain is well documented by Zheng and Tse in their landmark paper [7].

In this dissertation, we are mainly interested in transmit diversity schemes that minimize the error probability by taking advantage of the spatial diversity. Early examples of transmit diversity techniques achieving full diversity are the delay diversity schemes proposed by Wittneben [23,24], and Seshadri and Winters [25], which date back to the early nineties. A few years later, Tarokh et al. generalized these schemes to the so-called space-time trellis codes (STTCs) [26], which combine trellis-coded modulation (TCM) with transmit diversity. Although STTCs provide full diversity and perform very well in terms of data rate and coding gain, their decoding complexity, which is measured in number of trellis states in the decoder, can be relatively high. Therefore, space-time block codes (STBCs) [8,9,27], which apply block coding across multiple transmit antennas, may be a good alternative for STTCs. Although STBCs suffer from a small loss in performance (coding gain) compared to STTCs, they also achieve full spatial diversity and are endowed with a remarkably simple decoding scheme.

### 4.2 Orthogonal Space-Time Block Codes

In this section, we assess the construction and properties of orthogonal spacetime block codes, which have gained a lot of attention recently because of their straightforward symbol-by-symbol decoding scheme and their ability to provide full spatial diversity. We start this section with Alamouti's transmit diversity technique, which laid the foundation for the generalized OSTBCs that followed soon in the literature.

### 4.2.1 Alamouti's Code

In 1998, Alamouti [9] reported a simple transmit diversity technique, which was designed for a system with two transmit antennas and transforms two
independent information symbols $s_{1}$ and $s_{2}$ into a $2 \times 2$ coded symbol matrix

$$
\mathbf{C}=\left[\begin{array}{cc}
s_{1} & -s_{2}^{*}  \tag{4.14}\\
s_{2} & s_{1}^{*}
\end{array}\right],
$$

where the rows and columns of $\mathbf{C}$ are related to the spatial and temporal dimension, respectively. Hence, the transmitter applies the $(\ell, k)$-th element of the code matrix $\mathbf{C}$, which we denote by $c_{\ell, k}$, to the $\ell$-th transmit antenna during the $k$-th time slot. Since on average one information symbol is sent per time slot, Alamouti's code is said to achieve full rate. Note that the rows of (4.14) are orthogonal, yielding

$$
\begin{equation*}
\mathbf{C C}^{H}=\left(\left|s_{1}\right|^{2}+\left|s_{2}\right|^{2}\right) \mathbf{I}_{2} \tag{4.15}
\end{equation*}
$$

where $\mathbf{I}_{2}$ denotes the $2 \times 2$ identity matrix. Extending the discrete-time SIMO model (4.5) to the MIMO case, the received signals at the $L_{\mathrm{r}}$ receive antennas can be represented by the $L_{r} \times 2$ matrix

$$
\begin{equation*}
\mathbf{R}=\sqrt{E_{\mathrm{S}}} \mathbf{H C}+\mathbf{W} \tag{4.16}
\end{equation*}
$$

where the $L_{r} \times 2$ matrix $\mathbf{W}$ represents additive spatially and temporally white noise and consists of i.i.d. ZM CSCG RVs with variance $N_{0}$, and the $L_{r} \times 2$ channel matrix $\mathbf{H}$ comprises the channel coefficients $h_{n, \ell}$ characterizing the channel link between the $\ell$-th transmit and $n$-th receive antenna:

$$
\mathbf{H}=\left[\begin{array}{cc}
h_{1,1} & h_{1,2}  \tag{4.17}\\
\vdots & \vdots \\
h_{L_{\mathrm{r}}, 1} & h_{L_{\mathrm{r}}, 2}
\end{array}\right] .
$$

Furthermore, we assume that the channel is normalized:

$$
\begin{equation*}
\mathbb{E}\left[\|\mathbf{H}\|_{\mathrm{F}}^{2}\right]=2 L_{\mathrm{r}} \tag{4.18}
\end{equation*}
$$

where $\|\cdot\|_{F}$ denotes the Frobenius norm. According to (4.16), ML detection of the information symbols $\left[s_{1}, s_{2}\right]$ yields the following decision rule

$$
\begin{equation*}
\left[\hat{s}_{1}, \hat{s}_{2}\right]=\arg \min _{\tilde{s}_{1}, \tilde{s}_{2}}\left\|\mathbf{R}-\sqrt{E_{\mathrm{S}}} \mathbf{H} \tilde{\mathbf{C}}\right\|_{\mathrm{F}}^{2} \tag{4.19}
\end{equation*}
$$

where $\tilde{\mathbf{C}}$ is the Alamouti matrix consisting of the symbols $\tilde{s}_{1}$ and $\tilde{s}_{2}$. Expanding the Frobenius norm in (4.19) yields

$$
\begin{equation*}
\left\|\mathbf{R}-\sqrt{E_{\mathrm{S}}} \mathbf{H} \tilde{\mathbf{C}}\right\|_{\mathrm{F}}^{2}=\|\mathbf{R}\|_{\mathrm{F}}^{2}-2 \Re\left[\operatorname{tr}\left[\mathbf{C}^{H} \mathbf{H}^{H} \mathbf{R}\right]\right]+\operatorname{tr}\left[\left(\mathbf{C C}^{H}\right)\left(\mathbf{H}^{H} \mathbf{H}\right)\right] \tag{4.20}
\end{equation*}
$$

where $\operatorname{tr}[\cdot]$ denotes the trace. Due to the orthogonality condition (4.15), the Frobenius norm (4.20) simplifies to a function of $\tilde{s}_{1}$ and $\tilde{s}_{2}$ without cross terms involving both $\tilde{s}_{1}$ and $\tilde{s}_{2}$. In this way, the minimization in (4.19) reduces to two


Figure 4.3: Discrete-time model of a MIMO system employing Alamouti's code.
minimizations over $s_{1}$ and $s_{2}$ separately and ML detection of the information symbols in the Alamouti matrix boils down to symbol-by-symbol decision for both $s_{1}$ and $s_{2}$

$$
\begin{equation*}
\hat{s}_{i}=\arg \min _{\tilde{s} \in \Psi}\left|u_{i}-\tilde{s}\right|^{2}, \quad i=1,2 \tag{4.21}
\end{equation*}
$$

where the minimization is over all symbols belonging to the considered constellation $\Psi$. With $\mathbf{H}=\left[\mathbf{h}_{1}, \mathbf{h}_{2}\right]$ and $\mathbf{R}=\left[\mathbf{r}_{1}, \mathbf{r}_{2}\right]$, the decision variables $u_{1}$ and $u_{2}$ are given by

$$
\begin{align*}
& u_{1}=\frac{\mathbf{h}_{1}^{H} \mathbf{r}_{1}+\mathbf{h}_{2}^{T} \mathbf{r}_{2}^{*}}{\sqrt{E_{\mathrm{s}}}\|\mathbf{H}\|_{\mathrm{F}}^{2}},  \tag{4.22}\\
& u_{2}=\frac{\mathbf{h}_{2}^{H} \mathbf{r}_{1}-\mathbf{h}_{1}^{T} \mathbf{r}_{2}^{*}}{\sqrt{E_{\mathrm{s}}}\|\mathbf{H}\|_{\mathrm{F}}^{2}} . \tag{4.23}
\end{align*}
$$

Note that the receiver has to buffer the samples received on all antennas during the two timeslots wherein the code matrix is transmitted, since the computation of the decision variables requires the knowledge of $\mathbf{r}_{1}$ and $\mathbf{r}_{2}$. In Fig. 4.3, a MIMO system employing Alamouti's code is displayed. It is readily verified that the decision variables $u_{i}$, with $i=1,2$, can be written as the sum of the transmitted information symbol $s_{i}$ and a ZM CSCG noise term $w_{i}$ with variance $N_{0} /\left(E_{\mathrm{s}}\|\mathbf{H}\|_{\mathrm{F}}^{2}\right)$

$$
\begin{equation*}
u_{i}=s_{i}+w_{i} \tag{4.24}
\end{equation*}
$$

Due to the similarity between (4.24) and (4.9), the BER for Alamouti's code can be easily calculated along the lines provided for a SIMO MRC receiver
in section 4.1.1. In case of i.i.d. Rayleigh fading, the squared Frobenius norm $\|\mathbf{H}\|_{\mathrm{F}}^{2}$ is distributed according to the $\chi^{2}$-distribution with $2 L$ degrees of freedom, with $L=2 L_{\mathrm{r}}$ being the product of the number of transmit and receive antennas. It is also important to note that because of coding in the space-time domain, the total average energy to transmit two information symbols is now given by $E_{\mathrm{s}} \mathbb{E}\left[\|\mathbf{C}\|_{\mathrm{F}}^{2}\right]$, such that the relation between $E_{\mathrm{S}}$ and $E_{\mathrm{b}}$ becomes

$$
\begin{equation*}
E_{\mathrm{s}}=\frac{1}{2} \log _{2}(M) E_{\mathrm{b}} . \tag{4.25}
\end{equation*}
$$

For example, in case of 4-QAM, the average BER for Alamouti's code with ML symbol-by-symbol detection and $L_{\mathrm{r}}$ receive antennas reduces to

$$
\begin{align*}
P_{\mathrm{b}, 4-\mathrm{QAM}}= & {\left[\frac{1}{2}\left(1-\sqrt{\frac{\frac{E_{\mathrm{b}}}{N_{0}}}{2+\frac{E_{\mathrm{b}}}{N_{0}}}}\right)\right]^{L} } \\
& \times \sum_{k=0}^{L-1}\binom{L-1+k}{k}\left[\frac{1}{2}\left(1+\sqrt{\frac{\frac{E_{\mathrm{b}}}{N_{0}}}{2+\frac{E_{\mathrm{b}}}{N_{0}}}}\right)\right]^{k} \tag{4.26}
\end{align*}
$$

which for high $E_{\mathrm{b}} / N_{0}$ reduces to [13, Eq. 14.4-18]

$$
\begin{equation*}
P_{\mathrm{b}, 4-\mathrm{QAM}}^{(\mathrm{as})} \approx\left(\frac{N_{0}}{2 E_{\mathrm{b}}}\right)^{L}\binom{2 L-1}{L} \tag{4.27}
\end{equation*}
$$

such that the asymptotic behaviour of the BER is proportional to $\left(E_{\mathrm{b}} / N_{0}\right)^{-L}$ and the maximal diversity gain of $L=2 L_{\mathrm{r}}$ is indeed achieved. Although an Alamouti scheme with $N$ receive antennas achieves the same diversity gain as a SIMO MRC system with $2 N$ receive antennas, the BERs realized by both systems are not identical, as can be observed from (4.12) and (4.26). To explain intuitively the reason for this difference, we consider a SIMO scheme with dual-antenna MRC reception, and an Alamouti scheme with 1 receive antenna. Furthermore, we assume that the same total amount of transmit power is available in both scenarios. In the MRC scheme, one symbol is transmitted per time slot and the total available energy is allocated to it. Hence, both diversity channels benefit from full power. In the Alamouti transmit diversity scheme, however, on average also one symbol is transmitted per time slot, yet the available transmit power is split equally between both transmit antennas, such that only half of the available power is used on both diversity channels. This results in a 3 dB loss of power efficiency, which is illustrated in Fig. 4.4. From the figure, we observe that the BER curves for the Alamouti scheme are parallel to those of the equivalent MRC scheme and shifted 3 dB to the right. In other words, the MRC and Alamouti schemes achieve the same diversity gain, yet the MRC receive diversity scheme benefits from an additional array gain of 3 dB as compared to the Alamouti transmit diversity scheme, as explained in section 4.1.1.


Figure 4.4: Alamouti transmit diversity scheme versus receive diversity.

### 4.2.2 Generalization of Alamouti's Code to Orthogonal SpaceTime Block Codes

As an extension of the theory of orthogonal designs studied by Radon and Hurwitz [28], Tarokh et al. introduced the notion of generalized complex orthogonal designs, which allows to generalize Alamouti's transmit diversity scheme and construct orthogonal space-time block codes achieving full diversity for any number of transmit antennas and any signal constellation [8].

Let us consider a MIMO wireless communication system with $L_{t}$ transmit and $L_{r}$ receive antennas. In this dissertation, we consider OSTBCs from complex orthogonal designs which transform $N_{\mathrm{s}}$ information symbols $s_{i}$, with $1 \leq i \leq N_{\mathrm{s}}$, into an $L_{\mathrm{t}} \times K_{\mathrm{c}}$ coded symbol matrix $\mathbf{C}$, the entries of which are linear combinations of $s_{i}$ and $s_{i}^{*}$

$$
\begin{equation*}
\mathbf{C}=\sum_{i=1}^{N_{\mathrm{s}}}\left(\mathbf{C}_{i} s_{i}+\mathbf{C}_{i}^{\prime} s_{i}^{*}\right) \tag{4.28}
\end{equation*}
$$

with $K_{c}$ being the number of time slots required to transmit one OSTBC matrix, and $\mathbf{C}_{i}$ and $\mathbf{C}_{i}^{\prime}$ denoting $L_{\mathrm{t}} \times K_{\mathrm{c}}$ matrices consisting of the coefficients of the information symbols $s_{i}$ and $s_{i}^{*}$, respectively. For example, the coefficient
matrices for Alamouti's code are given by

$$
\mathbf{C}_{1}=\left[\begin{array}{ll}
1 & 0  \tag{4.29}\\
0 & 0
\end{array}\right], \mathbf{C}_{2}=\left[\begin{array}{ll}
0 & 0 \\
1 & 0
\end{array}\right], \mathbf{C}_{1}^{\prime}=\left[\begin{array}{ll}
0 & 0 \\
0 & 1
\end{array}\right], \mathbf{C}_{2}^{\prime}=\left[\begin{array}{cc}
0 & -1 \\
0 & 0
\end{array}\right]
$$

OSTBCs from complex orthogonal designs satisfy the following important orthogonality condition

$$
\begin{equation*}
\mathbf{C C}^{H}=C\|\mathbf{s}\|^{2} \mathbf{I}_{L_{\mathrm{t}}} \tag{4.30}
\end{equation*}
$$

where $C$ is a strictly positive constant and $\mathbf{s}=\left[s_{1}, s_{2}, \ldots, s_{N_{s}}\right]^{T}$ is the data symbol vector. Since scaling of the OSTBC matrices does not affect their orthogonality, we assume without loss of generality that the OSTBC matrices are normalized in such way that they satisfy

$$
\begin{equation*}
\mathbf{C C}^{H}=\lambda\|\mathbf{s}\|^{2} \mathbf{I}_{L_{\mathrm{t}}} \tag{4.31}
\end{equation*}
$$

where $\lambda \triangleq K_{\mathrm{c}} / N_{\mathrm{s}}$. In this way, using a normalized signal constellation $\Psi$, such that $\mathbb{E}\left[\left|s_{i}\right|^{2}\right]=1$, (4.31) yields

$$
\begin{equation*}
\frac{1}{L_{\mathrm{t}} K_{\mathrm{c}}} \mathbb{E}\left[\|\mathbf{C}\|_{\mathrm{F}}^{2}\right]=1 . \tag{4.32}
\end{equation*}
$$

Various examples of OSTBCs can be found in, e.g., [8,9,27,29]. It is shown in [8] that OSTBCs achieving full diversity and a code rate $1 / 2$ can be constructed for any number of transmit antennas and any signal constellation. For 3 transmit antennas, e.g., we have

$$
\mathbf{C}_{3 \times 8}=\left[\begin{array}{cccccccc}
s_{1} & -s_{2} & -s_{3} & -s_{4} & s_{1}^{*} & -s_{2}^{*} & -s_{3}^{*} & -s_{4}^{*}  \tag{4.33}\\
s_{2} & s_{1} & s_{4} & -s_{3} & s_{2}^{*} & s_{1}^{*} & s_{4}^{*} & -s_{3}^{*} \\
s_{3} & -s_{4} & s_{1} & s_{2} & s_{3}^{*} & -s_{4}^{*} & s_{1}^{*} & s_{2}^{*}
\end{array}\right],
$$

which transmits 4 information symbols within 8 time slots and thus achieves an effective rate of $1 / 2$. It is also possible to construct OSTBCs with higher rates than $1 / 2$. For $L_{t}=2$, Alamouti's scheme achieves full rate, whereas for $L_{t}=3$ and 4 , OSTBCs with rate $3 / 4$ have been reported [8, Eqs. (39) and (40)]. As an example, we provide the rate $3 / 4$ OSTBC for $L_{t}=3$, which transmits 3 information symbols within 4 time slots

$$
\mathrm{C}_{3 \times 4}=\frac{2}{\sqrt{3}}\left[\begin{array}{cccc}
s_{1} & -s_{2}^{*} & \frac{s_{3}^{*}}{\sqrt{2}} & \frac{s_{3}^{*}}{\sqrt{2}}  \tag{4.34}\\
s_{2} & s_{1}^{*} & \frac{s_{3}^{*}}{\sqrt{2}} & -\frac{s_{3}^{*}}{\sqrt{2}} \\
\frac{s_{3}}{\sqrt{2}} & \frac{s_{3}}{\sqrt{2}} & \frac{-s_{1}-s_{1}^{1}+s_{2}-s_{2}^{*}}{2} & \frac{s_{2}+s_{2}^{*}+s_{1}-s_{1}^{*}}{2}
\end{array}\right] .
$$

Note that we apply the scaling factor $2 / \sqrt{3}$ in (4.34) in order that $\mathbf{C}_{3 \times 4}$ satisfies (4.31).

From (4.28) and (4.31), the following properties can be derived for the coefficient matrices $\mathbf{C}_{i}$ and $\mathbf{C}_{i}^{\prime}$

$$
\begin{gather*}
\mathbf{C}_{i} \mathbf{C}_{n}^{H}+\mathbf{C}_{n}^{\prime} \mathbf{C}_{i}^{\prime H}=\lambda \delta_{i-n} \mathbf{I}_{L_{\mathrm{t}}}  \tag{4.35a}\\
\mathbf{C}_{i} \mathbf{C}_{n}^{\prime H}+\mathbf{C}_{n} \mathbf{C}_{i}^{\prime H}=\mathbf{0}_{L_{\mathrm{t}^{\prime}}} \tag{4.35b}
\end{gather*}
$$

where $1 \leq i, n \leq N_{\mathrm{s}}$ and $\delta_{k}$ denotes the discrete Dirac function. Moreover, for square OSTBCs, i.e., when $L_{t}=K_{c}$, it is readily verified that

$$
\begin{equation*}
\mathbf{C}^{H} \mathbf{C}=\mathbf{C C}^{H} \tag{4.36}
\end{equation*}
$$

such that for square OSTBCs the coefficient matrices also satisfy

$$
\begin{gather*}
\mathbf{C}_{i}^{H} \mathbf{C}_{n}+\mathbf{C}_{n}^{\prime}{ }^{H} \mathbf{C}_{i}^{\prime}=\lambda \delta_{i-n} \mathbf{I}_{L_{\mathrm{t}}}  \tag{4.37a}\\
\mathbf{C}_{i}^{H} \mathbf{C}_{n}^{\prime}+\mathbf{C}_{n}^{H} \mathbf{C}_{i}^{\prime}=\mathbf{0}_{L_{\mathrm{t}}} . \tag{4.37b}
\end{gather*}
$$

Examples of square OSTBCs are Alamouti's code (4.14) and the $4 \times 4$ codes given in [8, Eq. (40)], [27, Eq. (62)], and [30, Eq. (41)]. Uncoded SIMO systems can be treated as a special case of square OSTBCs, with $L_{t}=N_{s}=1$, and coefficient matrices $\mathbf{C}_{1}=1$ and $\mathbf{C}_{1}^{\prime}=0$.

Similar to the signal model (4.16) for Alamouti's code, the received signals corresponding to the transmitted OSTBC C can be represented by an $L_{r} \times K_{\mathrm{C}}$ matrix $\mathbf{R}$

$$
\begin{equation*}
\mathbf{R}=\sqrt{E_{\mathrm{s}}} \mathbf{H C}+\mathbf{W} \tag{4.38}
\end{equation*}
$$

where $\mathbf{W}$ consists of i.i.d. ZM CSCG RVs, the real and imaginary parts of which have variance $N_{0} / 2$, and $\mathbf{H}$ denotes the $L_{r} \times L_{\mathrm{t}}$ MIMO channel matrix

$$
\mathbf{H}=\left[\begin{array}{ccc}
h_{1,1} & \ldots & h_{1, L_{\mathrm{t}}}  \tag{4.39}\\
\vdots & \ddots & \vdots \\
h_{L_{\mathrm{r}}, 1} & \ldots & h_{L_{\mathrm{r}}, L_{\mathrm{t}}}
\end{array}\right]
$$

Given the fact that $N_{\mathrm{s}}$ information symbols are sent within one OSTBC matrix C and taking (4.31) into account, $E_{\mathrm{S}}$ in (4.38) is given by

$$
\begin{equation*}
E_{\mathrm{s}}=\rho \log _{2}(M) E_{\mathrm{b}} \tag{4.40}
\end{equation*}
$$

where $\rho \triangleq N_{\mathrm{s}} /\left(L_{\mathrm{t}} K_{\mathrm{c}}\right)$.
Using (4.28) and (4.31), it can be shown that ML detection of the information symbols $s_{i}$ in the OSTBC matrix reduces to symbol-by-symbol detection

$$
\begin{equation*}
\hat{s}_{i}=\arg \min _{\tilde{s} \in \Psi}\left|u_{i}-\tilde{s}\right|, \quad i=1, \ldots, N_{\mathrm{S}} \tag{4.41}
\end{equation*}
$$

where the minimization is over the symbols $\tilde{s}$ belonging to the considered constellation $\Psi$ and the decision variables $u_{i}$ are given by

$$
\begin{equation*}
u_{i}=\frac{\operatorname{tr}\left(\mathbf{C}_{i}^{H} \mathbf{H}^{H} \mathbf{R}+\mathbf{R}^{H} \mathbf{H C} C_{i}^{\prime}\right)}{\lambda \sqrt{E_{\mathrm{s}}}\|\mathbf{H}\|_{\mathrm{F}}^{2}} \tag{4.42}
\end{equation*}
$$

By substituting (4.38) in (4.42), it follows from (4.28) and (4.35) that the decision variables $u_{i}$ can be written as the sum of the transmitted information symbol $s_{i}$ and a ZM CSCG noise term $w_{i}$ with variance $N_{0} /\left(\lambda E_{\mathrm{S}}\|\mathbf{H}\|_{\mathrm{F}}^{2}\right)$

$$
\begin{equation*}
u_{i}=s_{i}+w_{i} \tag{4.43}
\end{equation*}
$$

such that the BER can be calculated in the same way as for SIMO and Alamouti's transmit scheme. From (4.40) and the variance of $w_{i}$ in (4.43), we can derive the important property that the BER curve plotted versus $E_{\mathrm{b}} / N_{0}$ only depends on the number of transmit and receive antennas when a specific constellation is considered. Therefore, different OSTBCs constructed for the same number of transmit antennas, e.g., (4.33) and (4.34) achieve the same BER, even though the corresponding code rates are not identical. For 4-QAM, e.g., the conditional BER of any OSTBC is given by

$$
\begin{equation*}
P_{\mathrm{b}, 4-\mathrm{QAM}}\left(\|\mathbf{H}\|_{\mathrm{F}}^{2}\right)=Q\left(\sqrt{\frac{2}{L_{\mathrm{t}}} \frac{E_{\mathrm{b}}}{N_{0}}\|\mathbf{H}\|_{\mathrm{F}}^{2}}\right), \tag{4.44}
\end{equation*}
$$

such that in case of i.i.d. Rayleigh fading channels, the average BER reduces to

$$
\begin{align*}
P_{\mathrm{b}, 4-\mathrm{QAM}}= & {\left[\frac{1}{2}\left(1-\sqrt{\frac{\frac{E_{\mathrm{b}}}{N_{0}}}{L_{\mathrm{t}}+\frac{E_{\mathrm{b}}}{N_{0}}}}\right)\right]^{L} } \\
& \times \sum_{k=0}^{L-1}\binom{L-1+k}{k}\left[\frac{1}{2}\left(1+\sqrt{\frac{\frac{E_{\mathrm{b}}}{N_{0}}}{L_{\mathrm{t}}+\frac{E_{\mathrm{b}}}{N_{0}}}}\right)\right]^{k} . \tag{4.45}
\end{align*}
$$

where $L=L_{\mathrm{t}} L_{\mathrm{r}}$. For high $E_{\mathrm{b}} / N_{0}$, (4.45) is shown to reduce to [13, Eq. 14.4-18]

$$
\begin{equation*}
P_{\mathrm{b}, 4-\mathrm{QAM}}^{(\mathrm{as})} \approx\left(\frac{L_{\mathrm{t}}}{4} \frac{N_{0}}{E_{\mathrm{b}}}\right)^{L}\binom{2 L-1}{L} \tag{4.46}
\end{equation*}
$$

such that the asymptotic behaviour of the BER is proportional to $\left(E_{\mathrm{b}} / N_{0}\right)^{-L}$ and the maximal diversity gain of $L_{\mathrm{t}} L_{\mathrm{r}}$ is indeed achieved.

### 4.3 MIMO Channel Model

In the previous section, we have introduced the slow flat-fading MIMO channel model (4.38) as an extension of the discrete-time SISO channel model
(3.45). We have also shown that in order to obtain the average BER for OSTBCs, the conditional BER has to be averaged over $\|\mathbf{H}\|_{\mathrm{F}}^{2}$. So far, we have assumed that the channel coefficients are i.i.d. ZM CSCG RVs, i.e., the norms of the channel coefficients are Rayleigh distributed, whereas the phases are uniformly distributed. This common assumption not only allows to simply obtain a closed-form solution for the BER, it also represents an empirically good fit for real-world MIMO channels as long as the environment consists of a large number of reflectors and scatterers and sufficient antenna spacing is assured. When the antennas are too closely spaced, however, fading correlation may appear between adjacent antennas, such that the assumption of i.i.d. fading is no longer valid. Moreover, when a line-of-sight (LOS) path is present or when an insufficiently rich scattering environment is considered, the Rayleigh distribution may not be accurate enough to characterize the fading and other fading distributions may be more appropriate.

### 4.3.1 Spatial Correlation

In general, diversity techniques rely on the assumption that all replicas of the transmitted signal undergo independent fading, which minimizes the probability that all diversity branches are in a deep fade simultaneously. Hence, in case of MIMO systems exploiting spatial diversity, fading correlation between channel coefficients is highly undesirable as it deteriorates the diversity technique's effectiveness. In practice, the correlation between channel coefficients depends both on antenna spacing and angle-of-arrival spread [31,32]. In a rich scattering environment, where the angular spread is large, a widely used rule of thumb is that an antenna separation of a half wavelength is sufficient to obtain independent fading channels. However, when the angular spread becomes small, e.g., at a base station on a high structure with few local scatterers, the antennas should be separated by several wavelengths to obtain decorrelation. When these physical requirements are not met, the channel coefficients in the MIMO channel matrix cannot be treated as i.i.d. RVs. Therefore, we establish in this section a model that describes the spatial correlation between the channel coefficients.

Let us introduce the vector equivalent

$$
\begin{equation*}
\mathbf{h} \triangleq \operatorname{vec}(\mathbf{H}) \tag{4.47}
\end{equation*}
$$

of the channel matrix $\mathbf{H}$, where $\operatorname{vec}(\mathbf{X})$ denotes the vec-operator, which stacks the columns of the $M \times N$ matrix $\mathbf{X}$ in one column vector of dimension $M N$. Hence, $\mathbf{h}=\left[h_{1}, \ldots, h_{L}\right]^{T}$ is an $L$-dimensional column vector, with $L=L_{\mathrm{t}} L_{\mathrm{r}}$. Note that $h_{k+(\ell-1) L_{\mathrm{r}}}=h_{k, \ell}$. The channel covariance matrix $\mathcal{R}$ of the MIMO channel is a positive semi-definite Hermitian matrix that captures the correlation between the different spatial paths between transmitter and receiver. For a discrete-time frequency-flat slow fading channel with ZM channel coef-
ficients, the $L \times L$ matrix $\mathcal{R}$ is defined as

$$
\begin{equation*}
\mathcal{R} \triangleq \mathbb{E}\left[\mathbf{h} \mathbf{h}^{H}\right] . \tag{4.48}
\end{equation*}
$$

In this dissertation, we will regularly use the Kronecker model, which assumes that the effect of fading correlation induced at the transmitter and receiver side can be decoupled. In this way, the channel covariance matrix is given by

$$
\begin{equation*}
\boldsymbol{\mathcal { R }}=\boldsymbol{\mathcal { R }}_{\mathrm{t}} \otimes \boldsymbol{\mathcal { R }}_{\mathrm{r}} \tag{4.49}
\end{equation*}
$$

which is the Kronecker product of the $L_{t} \times L_{t}$ transmit covariance matrix $\boldsymbol{\mathcal { R }}_{\mathrm{t}}$ and the $L_{\mathrm{r}} \times L_{\mathrm{r}}$ receive covariance matrix $\boldsymbol{\mathcal { R }}_{\mathrm{r}}$. The Kronecker product of an $M \times N$ matrix $\mathbf{X}$ and a $P \times Q$ matrix $\mathbf{Y}$ is the $M P \times N Q$ block matrix $\mathbf{Z}$, given by

$$
\mathbf{Z}=\mathbf{X} \otimes \mathbf{Y} \triangleq\left[\begin{array}{ccc}
x_{1,1} \mathbf{Y} & \cdots & x_{1, N} \mathbf{Y}  \tag{4.50}\\
\vdots & \ddots & \vdots \\
x_{M, 1} \mathbf{Y} & \cdots & x_{M, N} \mathbf{Y}
\end{array}\right]
$$

The entries of $\boldsymbol{\mathcal { R }}_{\mathrm{t}}$ and $\boldsymbol{\mathcal { R }}_{\mathrm{r}}$ in (4.49) are defined as

$$
\begin{array}{ll}
\left(\boldsymbol{\mathcal { R }}_{\mathrm{t}}\right)_{i, n} \triangleq \mathbb{E}\left[h_{k, i}, h_{k, n}^{*}\right], & 1 \leq k \leq L_{\mathrm{r}} \\
\left(\boldsymbol{\mathcal { R }}_{\mathrm{r}}\right)_{i, n} \triangleq \mathbb{E}\left[h_{i, k}, h_{n, k}^{*}\right], & 1 \leq k \leq L_{\mathrm{t}} \tag{4.51b}
\end{array}
$$

Hence, $\boldsymbol{\mathcal { R }}_{\mathrm{t}}$ describes the correlation between the fading of the different transmit antennas, which is assumed to be independent of the considered receive antenna. Likewise, $\boldsymbol{\mathcal { R }}_{\mathrm{r}}$ describes the receive correlation, irrespective of the particular transmit antenna. In this way, the correlation between two arbitrary channel coefficients is given by the product of the corresponding transmit and receive correlation coefficients

$$
\begin{equation*}
\mathbb{E}\left[h_{k, \ell} h_{i, n}^{*}\right]=\left(\boldsymbol{\mathcal { R }}_{\mathrm{t}}\right)_{\ell, n}\left(\boldsymbol{\mathcal { R }}_{\mathrm{r}}\right)_{k, i} \tag{4.52}
\end{equation*}
$$

Note that if the entries of $\mathbf{H}_{\mathrm{w}}$ are independent ZM RVs with unit variance, the covariance matrix of the channel

$$
\begin{equation*}
\mathbf{H}=\boldsymbol{\mathcal { R }}_{\mathrm{r}}^{1 / 2} \mathbf{H}_{\mathrm{w}}\left(\boldsymbol{\mathcal { R }}_{\mathrm{t}}^{1 / 2}\right)^{T} \tag{4.53}
\end{equation*}
$$

where $(\cdot)^{1 / 2}$ denotes any matrix square root satisfying $\mathbf{X}^{1 / 2}\left(\mathbf{X}^{1 / 2}\right)^{H}=\mathbf{X}$, will be given by (4.49). Although the Kronecker model has become popular owing to its simple analytical treatment, it is not always valid, in particular when large antenna arrays are considered [33,34]. In these cases, more elaborate channel models may yield more accurate results [35].

In addition to the covariance matrix $\boldsymbol{\mathcal { R }}$, we also define the $L \times L$ power correlation matrix $\Sigma$. With $\alpha_{\ell}=\left|h_{\ell}\right|$ denoting the fading envelope of the channel coefficient $h_{\ell}$, the entries of $\Sigma$ are defined as [14, Eq. (9.195)]

$$
\begin{equation*}
(\boldsymbol{\Sigma})_{i, n} \triangleq \frac{\operatorname{cov}\left(\alpha_{i}^{2}, \alpha_{n}^{2}\right)}{\sqrt{\operatorname{var}\left(\alpha_{i}^{2}\right) \operatorname{var}\left(\alpha_{n}^{2}\right)}} \tag{4.54}
\end{equation*}
$$

where $\operatorname{var}(x)$ and $\operatorname{cov}(x, y)$ denote the variance of $x$ and the covariance of $x$ and $y$, respectively, and $i, n=1,2, \ldots, L$. Using the MIMO Kronecker model proposed in [36], $\Sigma$ is decomposed as

$$
\begin{equation*}
\boldsymbol{\Sigma}=\boldsymbol{\Sigma}_{\mathrm{t}} \otimes \boldsymbol{\Sigma}_{\mathrm{r}} \tag{4.55}
\end{equation*}
$$

where $\Sigma_{t}$ and $\Sigma_{r}$ are the $L_{t} \times L_{t}$ transmit and $L_{r} \times L_{r}$ receive power correlation matrices, respectively. With $\alpha_{k, \ell}=\left|h_{k, \ell}\right|$ denoting the envelope of the channel coefficient $h_{k, \ell}$, the elements of $\Sigma_{\mathrm{t}}$ and $\Sigma_{\mathrm{r}}$ are defined as

$$
\begin{equation*}
\left(\Sigma_{\mathrm{t}}\right)_{i, n} \triangleq \frac{\operatorname{cov}\left(\alpha_{k, i}^{2}, \alpha_{k, n}^{2}\right)}{\sqrt{\operatorname{var}\left(\alpha_{k, i}^{2}\right) \operatorname{var}\left(\alpha_{k, n}^{2}\right)}} \quad 1 \leq k \leq L_{\mathrm{r}} \tag{4.56a}
\end{equation*}
$$

with $i, n=1,2, \ldots, L_{\mathrm{t}}$, and

$$
\begin{equation*}
\left(\Sigma_{\mathrm{r}}\right)_{i, n} \triangleq \frac{\operatorname{cov}\left(\alpha_{i, k}^{2} \alpha_{n, k}^{2}\right)}{\sqrt{\operatorname{var}\left(\alpha_{i, k}^{2}\right) \operatorname{var}\left(\alpha_{n, k}^{2}\right)}} \quad 1 \leq k \leq L_{\mathrm{t}} \tag{4.56b}
\end{equation*}
$$

with $i, n=1,2, \ldots, L_{\mathrm{r}}$. According to (4.56), $\Sigma_{\mathrm{t}}$ and $\Sigma_{\mathrm{r}}$ are independent of the index of the considered receive antenna and transmit antenna, respectively. From the Kronecker model (4.55), it follows that the normalized covariance between the fading powers $\alpha_{k, \ell}^{2}$ and $\alpha_{i, n}^{2}$ equals $\left(\Sigma_{\mathrm{t}}\right)_{\ell, n}\left(\Sigma_{\mathrm{r}}\right)_{k, i}$.

### 4.3.2 Fading Distributions

According to the MIMO channel model (4.38), the received power is determined by the fading envelopes, which depend on the positions of the transmitter and the receiver, and on the presence of absorbing, reflecting, and scattering objects in the propagation environment. In general, we make abstraction of path loss and shadowing effects [31], and assume a normalized channel matrix affected by small-scale fading only. This fading phenomenon stems from the fact that in reality each channel coefficient represents a cluster of multipath components with almost identical time delays but different amplitudes and phases. Small time variations in the phases may cause the different signals to add either constructively or destructively, causing large amplitude changes in the received signal. Since the statistical behavior of the envelope of the fading coefficients depends on the nature of the propagation environment,
an appropriate statistical model has to be selected in order to ensure a satisfactory empirical fit. In this section, we provide an overview of a number of widespread (marginal) fading distributions that model the statistical behavior of the individual channel coefficients.

### 4.3.2.1 Rayleigh Fading

In a dense scattering environment with no LOS components, each channel coefficient comprises a large number of independent multipath components within a very short time window. In this way, it follows from the central limit theorem that the channel coefficients can be modeled as a ZM CSCG RVs. When for a ZM CSCG channel coefficient $h, \mathbb{E}\left[|h|^{2}\right]=\Omega$, we denote $h \sim N_{\mathrm{C}}(0, \Omega)$ and it can be shown that the channel fading amplitude $\alpha=|h|$ is distributed according to the Rayleigh distribution

$$
\begin{equation*}
p_{\alpha}(x)=\frac{2 x}{\Omega} \exp \left(-\frac{x^{2}}{\Omega}\right), \quad x \geq 0 . \tag{4.57}
\end{equation*}
$$

Consequently, the squared fading amplitude $\alpha^{2}$ is distributed according to the exponential distribution

$$
\begin{equation*}
p_{\alpha^{2}}(x)=\frac{1}{\Omega} \exp \left(-\frac{x}{\Omega}\right), \quad x \geq 0 \tag{4.58}
\end{equation*}
$$

Note that $\Omega=1$ in case of a normalized channel.

### 4.3.2.2 Rice Fading

When the channel consists of both a LOS path not affected by fading and a cluster of multipath components subjected to Rayleigh fading, the channel coefficients can be written as

$$
\begin{equation*}
h=\sqrt{\frac{\kappa}{\kappa+1}} \sqrt{\Omega} \exp (-j \theta)+\sqrt{\frac{1}{\kappa+1}} y \tag{4.59}
\end{equation*}
$$

where $y \sim \mathrm{~N}_{\mathrm{c}}(0, \Omega)$ and $\kappa$ denotes the Rician $K$-factor, which ranges from 0 to $\infty$. It follows from (4.59) that $\mathbb{E}\left[|h|^{2}\right]=\Omega$ and that $\kappa$ indicates the ratio of the power of the LOS component to the power of the Rayleigh multipath component. When $\kappa=0$, the LOS component can be neglected and the Rice fading channel becomes a Rayleigh fading channel. When $\kappa \rightarrow \infty$, the Rayleigh multipath component can be neglected and the Rice fading channel becomes an additive white Gaussian noise (AWGN) channel.

In case of Rice fading channel, the fading amplitude $\alpha=|h|$ is distributed according to the Rice distribution, also known as the Nakagami- $n$ distribution

$$
\begin{equation*}
p_{\alpha}(x)=\frac{2(1+\kappa) \exp (-\kappa) x}{\Omega} \exp \left(-\frac{(1+\kappa) x^{2}}{\Omega}\right) I_{0}\left(2 x \sqrt{\frac{\kappa(1+\kappa)}{\Omega}}\right), \quad x \geq 0 \tag{4.60}
\end{equation*}
$$

where $I_{0}(\cdot)$ is the zeroth-order modified Bessel function of the first kind.

### 4.3.2.3 Nakagami- $m$ Fading

The Nakagami-m distribution [37] is considered as a versatile statistical distribution that accurately models a variety of fading environments by selecting a proper value for the fading parameter $m \geq 1 / 2$. It includes the Rayleigh ( $m=1$ ) and the one-sided Gaussian ( $m=1 / 2$ ) distributions as special cases. Moreover, for $m \rightarrow \infty$, the fading channel converges to an AWGN channel. Denoting by $\alpha$ the magnitude of a complex-valued channel coefficient, its PDF in case of Nakagami- $m$ fading is given by

$$
\begin{equation*}
p_{\alpha}(x)=\frac{2}{\Gamma(m)}\left(\frac{m}{\Omega}\right)^{m} x^{2 m-1} \exp \left(-\frac{m}{\Omega} x^{2}\right), \quad x \geq 0 \tag{4.61}
\end{equation*}
$$

with $\Omega=\mathbb{E}\left[\alpha^{2}\right]$ being the average fading power and $\Gamma(\cdot)$ being the Gamma function

$$
\begin{equation*}
\Gamma(x)=\int_{0}^{\infty} t^{x-1} \exp (-t) \mathrm{d} t \tag{4.62}
\end{equation*}
$$

Equivalently, the fading power $\alpha^{2}$ follows a Gamma distribution with shape parameter $m$ and mean $\Omega$

$$
\begin{equation*}
p_{\alpha^{2}}(x)=\frac{1}{\Gamma(m)}\left(\frac{m}{\Omega}\right)^{m} x^{m-1} \exp \left(-\frac{m}{\Omega} x\right), \quad x \geq 0 \tag{4.63}
\end{equation*}
$$

In Fig. 4.5, the Nakagami- $m$ distribution is depicted for $\Omega=1$ and $m=$ $0.5,1,2,5$. It is shown in [14] that for $m>1$, a one-to-one mapping can be obtained between the $m$ parameter of the Nakagami- $m$ distribution and the Rician K-factor. In this way, the Rice distribution can be closely approximated by the mathematically less complicated Nakagami- $m$ distribution. The mapping function is given by

$$
\begin{align*}
m & =\frac{(1+\kappa)^{2}}{1+2 \kappa}, \quad \kappa \geq 0  \tag{4.64a}\\
\kappa & =\frac{\sqrt{m^{2}-m}}{m-\sqrt{m^{2}-m}}, \quad m \geq 1 \tag{4.64b}
\end{align*}
$$

In Fig. 4.6, both the Rice distribution and the corresponding Nakagami-m distribution are depicted for several values of the Rician K-factor $\mathcal{\kappa}$.

### 4.3.2.4 Other

Other distributions that are sometimes used to describe small-scaling fading are the Hoyt distribution, also known as the Nakagami- $q$ distribution, and the Weibull distribution. It is shown in [14] that both distributions can be closely approximated by the Nakagami- $m$ distribution.


Figure 4.5: Nakagami- $m$ distribution with $\Omega=1$ and $m=0.5,1,2,5$.

### 4.3.3 Generating Correlated MIMO Channels

When the BER is obtained through Monte-Carlo simulation, numerous realisations of the MIMO channel have to be generated according to a particular fading distribution. In this section, we explain how spatially correlated Rayleigh and Nakagami-m channels can be obtained from easy-to-generate i.i.d. Gaussian RVs.

### 4.3.3.1 Rayleigh Fading

A channel vector $\mathbf{h}$ with Rayleigh fading envelopes $\alpha_{\ell}=\left|h_{\ell}\right|$ and covariance matrix $\mathcal{R}$ can simply be obtained as

$$
\begin{equation*}
\mathbf{h}=\mathbf{P} \mathbf{x} \tag{4.65}
\end{equation*}
$$

where the $L$-dimensional column vector $\mathbf{x}$ consists of i.i.d. Gaussian RVs with ZM and unit variance and the $L \times L$ matrix $\mathbf{P}$ results from the Cholesky decomposition of the covariance matrix $\mathcal{R}$, i.e., $\mathcal{R}=\mathbf{P} \mathbf{P}^{H}$.


Figure 4.6: Rice distribution versus approximate Nakagami- $m$ distribution.

### 4.3.3.2 Nakagami- $m$ Fading

Usually, arbitrarily correlated Nakagami-m RVs are generated from either Gamma RVs [38-40] or Gaussian RVs [41-45]. An efficient method for generating bivariate Nakagami- $m$ samples based on the rejection method is given in [46], for arbitrary values of $m \geq 0.8$. For integer and identical fading parameters, i.e., $m_{\ell}=m, \forall \ell$, it is shown in [44] that $L$ correlated Nakagami- $m$ RVs $\alpha_{\ell}$, can be obtained from $2 m$ i.i.d. real-valued ZM Gaussian random vectors $\mathbf{y}_{k}=\left[y_{k, 1}, y_{k, 2}, \ldots, y_{k, L}\right]^{T}$, with $k=1,2, \ldots, 2 m$. In particular, by defining

$$
\begin{equation*}
\alpha_{\ell}^{2} \triangleq \sum_{k=1}^{2 m} y_{k, \ell}^{2} \tag{4.66}
\end{equation*}
$$

it is readily verified that $\alpha_{\ell}$ 's are correlated Nakagami- $m$ RVs with $\mathbb{E}\left[\alpha_{\ell}^{2}\right]=\Omega_{\ell}$ and power correlation matrix $\Sigma$, if the covariance matrix $\mathbf{Q}=\mathbb{E}\left[\mathbf{y}_{k} \mathbf{y}_{k}^{T}\right]$ of the column vectors $\mathbf{y}_{k}$ is given by

$$
\begin{equation*}
\mathbf{Q}=\frac{1}{2 m} \sqrt{\boldsymbol{\Omega}} \Sigma_{\mathbf{G}} \sqrt{\boldsymbol{\Omega}} \tag{4.67}
\end{equation*}
$$

where the $L \times L$ diagonal matrix $\Omega$ is given by $\Omega=\operatorname{diag}\left\{\Omega_{1}, \Omega_{2}, \ldots, \Omega_{L}\right\}$ and $\Sigma_{\mathbf{G}}=\sqrt{\Sigma}$, with $\sqrt{\mathbf{X}}$ denoting the element-wise square root of a matrix $\mathbf{X}$.

### 4.3.4 PDF of the Squared Channel Norm

Since it follows from (4.43) that the analytical computation of the BER for OSTBCs requires averaging of the conditional BER over $\|\mathbf{H}\|_{\mathrm{F}}^{2}$, we present in this section the PDF of $\|\mathbf{H}\|_{\mathrm{F}}^{2}$ for i.i.d. and correlated Rayleigh and Nakagami$m$ channels. Note that $\|\mathbf{h}\|^{2}=\|\mathbf{H}\|_{\mathrm{F}}^{2}$.

### 4.3.4.1 Rayleigh Fading

i.i.d. Rayleigh Fading Let us consider a MIMO channel $\mathbf{H}$ with $L=L_{t} L_{\mathrm{r}}$ i.i.d. channel coefficients $h_{k, \ell}$, the envelopes of which are Rayleigh distributed with $\mathbb{E}\left[\left|h_{k, \ell}\right|^{2}\right]=\sigma^{2}$. Assuming uniformly distributed phases for all channel coefficients, the channel vector $\mathbf{h}$ is a vector consisting of $L$ i.i.d. ZM CSCG RVs $h_{\ell}$, with $\mathbb{E}\left[\left|h_{\ell}\right|^{2}\right]=\sigma^{2}$ for $\ell=1, \ldots, L$. It can easily be derived that $\|\mathbf{h}\|^{2}=\sum_{\ell=1}^{L}\left|h_{\ell}\right|^{2}$ is distributed according to a scaled $\chi^{2}$-distribution with $2 L$ degrees of freedom [13, Eq. 2.1-110]:

$$
\begin{equation*}
p_{\|\mathbf{h}\|^{2}}(x)=\frac{1}{\left(\sigma^{2}\right)^{L}(L-1)!} x^{L-1} \exp \left(-\frac{x}{\sigma^{2}}\right), \quad x \geq 0 . \tag{4.68}
\end{equation*}
$$

Similarly, $\|\mathbf{h}\|$ follows a scaled $\chi$-distribution with $2 L$ degrees of freedom. Note that the $\chi$ and $\chi^{2}$-distributions are special cases of the Nakagami- $m$ and Gamma distributions, respectively. Exact BER expressions for OSTBCs on i.i.d. Rayleigh fading channels can be found in [47] for M-PSK and in [48] for M-QAM constellations.

Arbitrarily Correlated Rayleigh Fading We consider a Rayleigh MIMO channel with an arbitrary positive semi-definite covariance matrix $\mathcal{R}$. Using a moment generating function (MGF) approach, it is readily verified that the PDF of $\|\mathbf{h}\|^{2}$ is given by [49]

$$
\begin{equation*}
p_{\|\mathbf{h}\|^{2}}(x)=\sum_{i=1}^{\kappa} \sum_{q=1}^{c_{i}} \frac{D_{i, q}}{\left(\lambda_{i}\right)^{q}(q-1)!} x^{q-1} \exp \left(-\frac{x}{\lambda_{i}}\right), \quad x \geq 0 \tag{4.69}
\end{equation*}
$$

where $\lambda_{i}, i=1,2, \ldots, \kappa$, are the distinct eigenvalues of the channel covariance matrix $\mathcal{R}$, with corresponding algebraic multiplicities $c_{i}$. Note that (4.69) is a finite weighted sum of $\chi^{2}$-distributions with $2 q$ degrees of freedom; the coefficients $D_{i, q}$ are given by

$$
\begin{equation*}
D_{i, q}=\left.\frac{\left(\lambda_{i}\right)^{q-c_{i}}}{\left(c_{i}-q\right)!}\left[\frac{\mathrm{d}^{c_{i}-q}}{\mathrm{~d} s^{c_{i}-q}} F(s)\left(1+\lambda_{i} s\right)^{c_{i}}\right]\right|_{s=-\frac{1}{\lambda_{i}}} \tag{4.70}
\end{equation*}
$$

where

$$
\begin{equation*}
F(s)=\prod_{n=1}^{\kappa} \frac{1}{\left(1+\lambda_{n} s\right)^{c_{n}}} \tag{4.71}
\end{equation*}
$$

Exact BER expressions for OSTBCs in correlated Rayleigh channels are provided in [50]. In case of i.i.d. Rayleigh fading with $\mathcal{R}=\mathbf{I}_{L}$, there is only one distinct eigenvalue $\lambda_{1}=1$ with multiplicity $c_{1}=L$, such that the coefficients (4.70) are given by

$$
D_{1, q} \triangleq \begin{cases}\frac{1}{(L-1)!} & \text { if } q=L  \tag{4.72}\\ 0 & \text { otherwise }\end{cases}
$$

and the PDF (4.69) reduces to (4.68).

### 4.3.4.2 Nakagami- $m$ Fading

i.i.d. Nakagami- $m$ Fading Let us consider an i.i.d. MIMO channel $\mathbf{H}$ with corresponding channel vector $\mathbf{h}$. When the fading envelopes $\alpha_{\ell}=\left|h_{\ell}\right|$ are i.i.d. Nakagami- $m$ distributed RVs with arbitrary and identical $m_{\ell}=m$ and $\Omega_{\ell}=\Omega, \forall \ell$, the PDF of $\|\mathbf{h}\|$ is shown to be distributed according to the Nakagami- $m$ distribution with parameters $L m$ and $L \Omega$ [51]. Accordingly, $\|\mathbf{h}\|^{2}$ follows a Gamma distribution with parameters $L m$ and $L \Omega$. Exact analytical expressions for the SER of OSTBCs in i.i.d. Nakagami- $m$ fading channel are given in [52]

Arbitrarily Correlated Nakagami- $m$ Fading In case the Nakagami- $m$ MIMO channel is correlated, the fading envelopes $\alpha_{\ell}$ are distributed according to (4.61) with parameters $m_{\ell}$ and $\Omega_{\ell}$ possibly depending on the index $\ell$.

In the past, different approaches have been presented for deriving analytical expressions for the distribution of $\|\mathbf{h}\|^{2}$ in the case of arbitrarily correlated Nakagami- $m$ fading channels, e.g., see [38-44] and references therein. Analytical expressions for the MGF of $\|\mathbf{h}\|^{2}$ have been derived for integer $m_{\ell}=m$, $\forall \ell,[41,42,44]$, integer $m_{\ell}$ [43] and arbitrary $m_{\ell}$ [40]. Although the obtained expressions in $[41,42,44]$ can be straightforwardly used for the derivation of $p_{\|\mathbf{h}\|^{2}}(x)$, this seems complicated using the MGF expression presented in [43] and rather difficult with that in [40]. On the other hand, the PDF-based approach has been used for deriving the distribution of $\|\mathbf{h}\|^{2}$ for arbitrary $m_{\ell}=m, \forall \ell,[38]$ and for integer $m_{\ell}$ with the restriction that $\Omega_{\ell} / m_{\ell} \neq \Omega_{k} / m_{k}$ if $k \neq \ell$ [39].

For integer and identical $m_{\ell}=m, \forall \ell$, and $\mathbb{E}\left[\alpha_{\ell}^{2}\right]=\Omega_{\ell}$, the PDF of $\|\mathbf{h}\|^{2}$ is given by [41,42]

$$
\begin{equation*}
p_{\|\mathbf{h}\|^{2}}(x)=\sum_{i=1}^{\kappa} \sum_{q=1}^{c_{i} m} \frac{D_{i, q}}{\left(2 \lambda_{i}\right)^{q}(q-1)!} x^{q-1} \exp \left(-\frac{x}{2 \lambda_{i}}\right), \quad x \geq 0 \tag{4.73}
\end{equation*}
$$

where $\lambda_{i}{ }^{\prime}$ s, $i=1,2, \ldots, \kappa$, are the distinct eigenvalues of $\mathbf{Q}$ given by (4.67), with corresponding algebraic multiplicities $c_{i}$. In (4.73), the parameters $D_{i, q}$


Figure 4.7: PDF of squared channel norm $\|\mathbf{h}\|^{2}$ in case of correlated Nakagami- $m$ channels.
are given by

$$
\begin{equation*}
D_{i, q}=\left.\frac{\left(2 \lambda_{i}\right)^{q-c_{i} m}}{\left(c_{i} m-q\right)!}\left[\frac{\mathrm{d}^{c_{i} m-q}}{\mathrm{~d} s^{c_{i} m-q}} F(s)\left(1+2 \lambda_{i} s\right)^{c_{i} m}\right]\right|_{s=-\frac{1}{2 \lambda_{i}}} \tag{4.74}
\end{equation*}
$$

where

$$
\begin{equation*}
F(s)=\prod_{n=1}^{\kappa} \frac{1}{\left(1+2 \lambda_{n} s\right)^{c_{n} m}} . \tag{4.75}
\end{equation*}
$$

Alternatively, by applying a tridiagonal decomposition to $\left(\Sigma_{\mathbf{G}}\right)^{-1}$ for integer $m_{\ell}=m$ and $\Omega_{\ell}=\Omega, \forall \ell, p_{\|\mathbf{h}\|^{2}}(x)$ can be obtained from [44] as fast convergent infinite summations. For arbitrary and identical $m_{\ell}=m, \forall \ell$, the PDF of $\|\mathbf{h}\|^{2}$ can be easily obtained from [38, Eq. (5)] as an infinite summation. For arbitrary and non-identical $m_{\ell}$, the PDF of $\|\mathbf{h}\|^{2}$ is given in [53] as an infinite summation and a good truncation of the PDF's infinite summation is proposed.

In Fig. 4.7, we show the PDF of the squared channel norm $\|\mathbf{h}\|^{2}$ for a $2 \times 2$ MIMO channel with transmit and receive power correlation matrices given by

$$
\Sigma_{\mathrm{t}}=\boldsymbol{\Sigma}_{\mathrm{r}}=\left[\begin{array}{ll}
1 & \rho  \tag{4.76}\\
\rho & 1
\end{array}\right]
$$

where the correlation coefficient $\rho$ takes on the values $0,0.2$, and 0.5 . The PDFs are shown for $m=1$ (Rayleigh fading) and $m=5$. Note that because of fading correlation small values of $\|\mathbf{h}\|^{2}$ are more likely to occur.

Exact analytical expressions for the SER of OSTBCs operating over correlated Nakagami- $m$ fading can be found in [54] for PSK and QAM constellations and in [55] for PAM/PSK/QAM modulation. Analytical BER expressions are provided in [56].

### 4.4 Chapter Summary

In this chapter, we have shown how using multiple receive antennas enables to exploit spatial diversity and mitigate the effect of multipath fading without extending the required bandwidth or decreasing the spectral efficiency. When the signals received at each of the $L_{r}$ antennas are combined by MRC, both a diversity gain of $L_{\mathrm{r}}$ and an array gain of $L_{\mathrm{r}}$ can be achieved. Multiple transmit antennas, on the other hand, can be used not only to exploit spatial diversity by applying proper space-time coding, but also to increase the system's data rate by transmitting different data streams in parallel; the latter transmit technique is usually referred to as spatial multiplexing. Nevertheless, it is impossible to maximize the diversity and the data rate simultaneously, as there exists a fundamental trade-off between diversity and multiplexing gain. In this dissertation, we focus on the appealing transmit diversity technique of orthogonal space-time block coding. When combined with receive diversity, orthogonal space-time block codes achieve a full spatial diversity gain of $L=L_{t} L_{r}$, with $L_{t}$ denoting the number of transmit antennas, whereas ML data detection reduce to symbol-by-symbol detection based only on signal processing at the receiver side. Furthermore, we introduced a slow flatfading MIMO channel model and discussed its statistical properties. Various widespread fading distributions were presented and we illustrated how correlation between fading coefficients can be taken into account.

## Channel Estimation

In chapters 3 and 4, we have assumed that the receiver detects the data symbols under the assumption that perfect channel state information (PCSI) is available at the receiver. In practical wireless scenarios, however, the channel is not a priori known because of its random nature so the assumption of PCSI is not valid. Typically, the receiver estimates the channel with the aid of known pilot symbols sent among the data [57], although blind or semi-blind joint channel estimation and detection techniques can also be applied [58,59]. This chapter discusses two widespread pilot aided channel estimation methods. In section 5.1, we show how pilot symbols are inserted in the data stream, whereas sections 5.2 and 5.3 deal with ML and LMMSE channel estimation, respectively. Section 5.4 concludes the chapter.

### 5.1 Insertion of Pilot Symbols

To facilitate pilot aided channel estimation, we organize the data transmission in frames consisting of $K_{\mathrm{fr}}=K+K_{\mathrm{p}}$ time slots; $K_{\mathrm{p}}$ time slots are associated to


Figure 5.1: A data frame with a $2 \times 4$ pilot matrix and five $2 \times 2$ code matrices.
pilot symbols, whereas the other $K$ time slots are reserved for the transmission of OSTBC matrices. Hereby, we assume that $K$ is a multiple of $K_{c}$, such that $K / K_{c}$ code matrices $\mathbf{C}(k)$, with $k$ denoting the matrix index, are sent within one frame. In Fig. 5.1, we show a data frame for a dual-antenna transmitter employing Alamouti's code, consisting of a $2 \times 4$ pilot matrix $C_{p}\left(K_{p}=4\right)$ and a sequence of five $2 \times 2$ code matrices $\mathbf{C}(k)(K=10)$. Furthermore, we assume orthogonal and normalized pilot sequences, i.e., $\mathbf{C}_{\mathrm{p}} \mathbf{C}_{\mathrm{p}}^{H}=K_{\mathrm{p}} \mathbf{I}_{L_{\mathrm{t}}}$. In this way, the average energy of the entries of $\mathbf{C}_{\mathrm{p}}$ is given by

$$
\begin{equation*}
\frac{1}{L_{\mathrm{t}} K_{\mathrm{p}}} \mathbb{E}\left[\left\|\mathbf{C}_{\mathrm{p}}\right\|_{\mathrm{F}}^{2}\right]=1 \tag{5.1}
\end{equation*}
$$

Assuming that the length of one frame of $K_{f r}$ symbols does not exceed the channel coherence time, the channel remains constant during the transmission of one frame (block fading), and the receiver separately observes the $L_{\mathrm{r}} \times K_{\mathrm{c}}$ matrices

$$
\begin{equation*}
\mathbf{R}(k)=\sqrt{E_{\mathrm{s}}} \mathbf{H C}(k)+\mathbf{W}(k), \tag{5.2}
\end{equation*}
$$

with $k=1, \ldots, K / K_{\mathrm{c}}$, and the $L_{\mathrm{r}} \times K_{\mathrm{p}}$ matrix

$$
\begin{equation*}
\mathbf{R}_{\mathrm{p}}=\sqrt{E_{\mathrm{p}}} \mathbf{H} \mathbf{C}_{\mathrm{p}}+\mathbf{W}_{\mathrm{p}} \tag{5.3}
\end{equation*}
$$

where the noise matrices $\mathbf{W}(k)$ and $\mathbf{W}_{\mathrm{p}}$ affecting the transmission of data and pilot symbols, respectively, consist of i.i.d. ZM CSCG RVs with variance $N_{0}$. Because of (4.32) and (5.1), $E_{\mathrm{s}}$ and $E_{\mathrm{p}}$ in (5.2) and (5.3) can be considered as the average data and pilot energy, respectively. In the remainder of this chapter, we will omit the matrix index $k$ for notational convenience.

When spatially correlated channels are considered, we often use the channel vector notation (4.47). It is possible, however, to derive an equivalent vector model for (5.2) and (5.3), which uses the channel vector $\mathbf{h}$ instead of $\mathbf{H}$. To this end, we introduce the $L$-dimensional column vectors $\mathbf{r} \triangleq \operatorname{vec}(\mathbf{R})$, $\mathbf{w} \triangleq \operatorname{vec}(\mathbf{W}), \mathbf{r}_{\mathrm{p}} \triangleq \operatorname{vec}\left(\mathbf{R}_{\mathrm{p}}\right)$, and $\mathbf{w}_{\mathrm{p}} \triangleq \operatorname{vec}\left(\mathbf{W}_{\mathrm{p}}\right)$. Using these vectors, the received signal matrices (5.2) and (5.3) are shown in [60] to be equivalent to

$$
\begin{equation*}
\mathbf{r}=\sqrt{E_{\mathrm{s}}} \mathbf{B h}+\mathbf{w} \tag{5.4}
\end{equation*}
$$

$$
\begin{equation*}
\mathbf{r}_{\mathrm{p}}=\sqrt{E_{\mathrm{p}}} \mathbf{B}_{\mathrm{p}} \mathbf{h}+\mathbf{w}_{\mathrm{p}} \tag{5.5}
\end{equation*}
$$

where $\mathbf{B} \triangleq \mathbf{C}^{T} \otimes \mathbf{I}_{L_{\mathrm{r}}}$ and $\mathbf{B}_{\mathrm{p}} \triangleq \mathbf{C}_{\mathrm{p}}^{T} \otimes \mathbf{I}_{L_{\mathrm{r}}}$. It follows from (4.31) that for OSTBCs the following criterion is met

$$
\begin{equation*}
\mathbf{B}^{H} \mathbf{B}=\lambda\|\mathbf{s}\|^{2} \mathbf{I}_{L}, \tag{5.6}
\end{equation*}
$$

whereas for square OSTBCs, (4.36) yields

$$
\begin{equation*}
\mathbf{B B}^{H}=\mathbf{B}^{H} \mathbf{B}=\lambda\|\mathbf{s}\|^{2} \mathbf{I}_{L} \tag{5.7}
\end{equation*}
$$

The equivalent pilot matrix $\mathbf{B}_{\mathrm{p}}$ in (5.5) satisfies $\mathbf{B}_{\mathrm{p}}^{H} \mathbf{B}_{\mathrm{p}}=K_{\mathrm{p}} \mathbf{I}_{L}$.
Given a total amount of energy per data frame, increasing the energy $K_{p} E_{p}$ allocated to pilot symbols allows to improve the channel estimate, but also reduces the symbol energy $E_{\mathrm{s}}$ available for data transmission. Hence, the optimal number of pilot symbols minimizing the BER will be a trade-off between accurate channel estimation on the one hand, and sufficient energy for data transmission on the other hand. With $E_{\mathrm{b}}$ denoting the average energy per information bit, so that the total energy per frame is constrained to $E_{\text {tot }}=\frac{K}{K_{\mathrm{c}}} N_{\mathrm{s}} \log _{2}(M) E_{\mathrm{b}}$, we have [61]

$$
\begin{equation*}
E_{\mathrm{s}}=\frac{K}{K+\eta K_{\mathrm{p}}} \rho \log _{2}(M) E_{\mathrm{b}} \tag{5.8}
\end{equation*}
$$

where $\eta \triangleq E_{\mathrm{p}} / E_{\mathrm{s}}$ denotes the ratio of $E_{\mathrm{p}}$ to $E_{\mathrm{s}}, M$ is the constellation size, and $\rho=N_{\mathrm{s}} /\left(L_{\mathrm{t}} K_{\mathrm{c}}\right)$. Clearly, $E_{\mathrm{s}}$ is a decreasing function of $K_{\mathrm{p}}$. Note that increasing $K_{\mathrm{p}}$ also decreases the information bit rate. Denoting by $R_{s}$ the symbol rate per transmit antenna, the resulting information bit rate is given by

$$
\begin{equation*}
R_{b}=\frac{K}{K+K_{\mathrm{p}}} \frac{N_{\mathrm{s}}}{K_{\mathrm{c}}} \log _{2}(M) R_{s} \tag{5.9}
\end{equation*}
$$

which indicates that the addition of pilot symbols reduces the bandwidth efficiency.

Although numerous definitions of the SNR are used in the literature when MIMO communication is considered, we use the following simple definition

$$
\begin{equation*}
\mathrm{SNR} \triangleq \frac{E_{\mathrm{S}}}{N_{0}} \tag{5.10}
\end{equation*}
$$

Throughout this dissertation, the SNR refers to the above definition, unless otherwise mentioned.

### 5.2 ML Channel Estimation

Channel estimation consists of deriving an estimate $\hat{\mathbf{H}}$ from the observed received signal matrix (5.3). As shown in chapter 2, various estimators are avail-
able from the literature. A popular MIMO channel estimator is the maximumlikelihood estimator (2.11), which does not use any a priori information about the channel. From (5.3) the likelihood function of the matrix $\mathbf{H}$ is obtained as

$$
\begin{equation*}
p\left(\mathbf{R}_{\mathrm{p}} \mid \mathbf{H}\right)=\frac{1}{\left(\pi N_{0}\right)^{L}} \exp \left(-\frac{\left\|\mathbf{R}_{\mathrm{p}}-\sqrt{E_{\mathrm{p}}} \mathbf{H} \mathbf{C}_{\mathrm{p}}\right\|_{\mathrm{F}}^{2}}{N_{0}}\right) \tag{5.11}
\end{equation*}
$$

where $L=L_{\mathrm{t}} L_{\mathrm{r}}$. Maximizing (5.11) results in the $L_{\mathrm{r}} \times L_{\mathrm{t}}$ ML channel estimate $\hat{\mathbf{H}}_{\mathrm{ML}}$, which is given by [61]

$$
\begin{equation*}
\hat{\mathbf{H}}_{\mathrm{ML}}=\frac{1}{K_{\mathrm{p}} \sqrt{E_{\mathrm{p}}}} \mathbf{R}_{\mathrm{p}} \mathbf{C}_{\mathrm{p}}^{H} \tag{5.12}
\end{equation*}
$$

such that $\hat{\mathbf{H}}_{\mathrm{ML}}$ can be decomposed into the following sum of two statistically independent contributions

$$
\begin{equation*}
\hat{\mathbf{H}}_{\mathrm{ML}}=\mathbf{H}+\mathbf{N} \tag{5.13}
\end{equation*}
$$

The entries of the $L_{\mathrm{r}} \times L_{\mathrm{t}}$ estimation noise matrix $\mathbf{N}=\left[1 /\left(K_{\mathrm{p}} \sqrt{E_{\mathrm{p}}}\right)\right] \mathbf{W}_{\mathrm{p}} \mathbf{C}_{\mathrm{p}}^{H}$ are i.i.d. ZM CSCG RVs, the real and imaginary parts of which have variance [61]

$$
\begin{equation*}
\sigma_{\mathrm{N}}^{2}=N_{0} /\left(2 K_{\mathrm{p}} E_{\mathrm{p}}\right) \tag{5.14}
\end{equation*}
$$

Hence, when conditioned on $\mathbf{H}$, the estimated channel coefficients $\left(\hat{\mathbf{H}}_{\mathrm{ML}}\right)_{\ell, k}$ are CSCG RVs with mean $h_{\ell, k}$ and variance $2 \sigma_{\mathrm{N}}^{2}$.

Note that, using (5.5), the vector equivalent to the ML channel estimate (5.12) is given by

$$
\begin{equation*}
\hat{\mathbf{h}}_{\mathrm{ML}}=\frac{\mathbf{B}_{\mathrm{p}}^{H} \mathbf{r}_{\mathrm{p}}}{K_{\mathrm{p}} \sqrt{E_{\mathrm{p}}}} \tag{5.15}
\end{equation*}
$$

Obviously, similar conclusions can be drawn for the components of $\hat{\mathbf{h}}_{\mathrm{ML}}$ as for the entries of (5.12).

### 5.3 LMMSE Channel Estimation

Under the assumption that the channel coefficients are ZM RVs, it follows from (5.5) that the LMMSE channel estimate (2.7) is given by [60]

$$
\begin{equation*}
\hat{\mathbf{h}}_{\mathrm{LMMSE}}=\frac{\sqrt{E_{\mathrm{p}}}}{N_{0}}\left(\mathbf{I}_{L}+\frac{K_{\mathrm{p}} E_{\mathrm{p}}}{N_{0}} \boldsymbol{\mathcal { R }}\right)^{-1} \mathcal{R} \mathbf{B}_{\mathrm{p}}^{H} \mathbf{r}_{\mathrm{p}} \tag{5.16}
\end{equation*}
$$

which is a function of the channel covariance matrix $\mathcal{R}=\mathbb{E}\left[\mathbf{h} \mathbf{h}^{H}\right]$. It is readily verified that for high SNR, the LMMSE and ML estimates coincide. Moreover, in case of Rayleigh fading channels, the LMMSE, MMSE, and MAP channel
estimators are identical. For i.i.d. channels with $\mathcal{R}=\mathbf{I}_{L}$, the LMMSE channel estimate (5.16) is easily shown to reduce to

$$
\begin{equation*}
\hat{\mathbf{h}}_{\mathrm{LMMSE}}=\frac{\sqrt{E_{\mathrm{p}}}}{N_{0}+K_{\mathrm{p}} E_{\mathrm{p}}} \mathbf{B}_{\mathrm{p}}^{H} \mathbf{r}_{\mathrm{p}}=\frac{K_{\mathrm{p}} E_{\mathrm{p}}}{N_{0}+K_{\mathrm{p}} E_{\mathrm{p}}} \hat{\mathbf{h}}_{\mathrm{ML}} \tag{5.17}
\end{equation*}
$$

where $\hat{\mathbf{h}}_{\mathrm{ML}}$ is defined by (5.15). Hence, the LMMSE estimate reduces to a scaled version of the ML estimate.

Let us now define the channel estimation error as $\varepsilon \triangleq \mathbf{h}-\hat{\mathbf{h}}$. In case of LMMSE channel estimation, it can be shown that the estimation error $\varepsilon$ and the channel estimate $\hat{\mathbf{h}}$ are uncorrelated. Moreover, the following properties can be derived for $\varepsilon$ and $\hat{\mathbf{h}}$ :

- The covariance matrix of $\hat{\mathbf{h}}$ is given by

$$
\begin{equation*}
\mathcal{R}_{\hat{\mathrm{h}}} \triangleq \mathbb{E}\left[\hat{\mathbf{h}} \hat{\mathbf{h}}^{H}\right]=\frac{K_{\mathrm{p}} E_{\mathrm{p}}}{N_{0}} \mathcal{R}\left(\mathbf{I}_{L}+\frac{K_{\mathrm{p}} E_{\mathrm{p}}}{N_{0}} \mathcal{R}\right)^{-1} \mathcal{R} \tag{5.18}
\end{equation*}
$$

In case of i.i.d. fading with $\mathcal{R}=\mathbf{I}_{L}$, the estimated channel coefficients are i.i.d. ZM RVs with variance $K_{\mathrm{p}} E_{\mathrm{p}} /\left(K_{\mathrm{p}} E_{\mathrm{p}}+N_{0}\right)$, since (5.18) reduces to

$$
\begin{equation*}
\mathcal{R}_{\hat{\mathrm{h}}}=\frac{K_{\mathrm{p}} E_{\mathrm{p}}}{K_{\mathrm{p}} E_{\mathrm{p}}+N_{0}} \mathbf{I}_{L} \tag{5.19}
\end{equation*}
$$

- The covariance matrix of $\varepsilon$ is given by

$$
\begin{equation*}
\mathcal{R}_{\varepsilon} \triangleq \mathbb{E}\left[\varepsilon \varepsilon^{H}\right]=\left(\mathbf{I}_{L}+\frac{K_{\mathrm{p}} E_{\mathrm{p}}}{N_{0}} \boldsymbol{\mathcal { R }}\right)^{-1} \mathcal{R} \tag{5.20}
\end{equation*}
$$

In case of i.i.d. fading with $\mathcal{R}=\mathbf{I}_{L}$, the elements of $\varepsilon$ are i.i.d. ZM RVs with variance $N_{0} /\left(K_{\mathrm{p}} E_{\mathrm{p}}+N_{0}\right)$, since (5.20) reduces to

$$
\begin{equation*}
\mathcal{R}_{\varepsilon}=\frac{N_{0}}{K_{\mathrm{p}} E_{\mathrm{p}}+N_{0}} \mathbf{I}_{L} \tag{5.21}
\end{equation*}
$$

Furthermore, when the components of $\mathbf{h}$ are ZM CSCG RVs, which is the case for Rayleigh fading, $\hat{\mathbf{h}}$ and $\varepsilon$ are statistically independent and their components are ZM CSCG RVs. Moreover, for high SNR, the elements of the channel noise vector $\varepsilon$ can be considered as i.i.d. ZM CSCG RVs with variance $N_{0} /\left(K_{\mathrm{p}} E_{\mathrm{p}}\right)$, irrespective of the channel covariance matrix $\mathcal{R}$, since $\boldsymbol{\mathcal { R }}_{\varepsilon}$ becomes

$$
\begin{equation*}
\mathcal{R}_{\varepsilon} \approx \frac{N_{0}}{K_{\mathrm{p}} E_{\mathrm{p}}} \mathbf{I}_{L}, \quad \frac{K_{\mathrm{p}} E_{\mathrm{p}}}{N_{0}} \gg 1 \tag{5.22}
\end{equation*}
$$

This result is in correspondence with the fact that the LMMSE and ML channel estimates coincide for high SNR.


Figure 5.2: MSE of ML and LMMSE channel estimators.

To compare the performance of the ML and LMMSE channel estimators, we compute the MSE of both estimators, yielding

$$
\begin{equation*}
\mathrm{MSE}_{\mathrm{ML}}=L \frac{N_{0}}{K_{\mathrm{p}} E_{\mathrm{p}}}, \tag{5.23}
\end{equation*}
$$

where $L=L_{\mathrm{t}} L_{\mathrm{r}}$, and

$$
\begin{equation*}
\mathrm{MSE}_{\mathrm{LMMSE}}=\operatorname{tr}\left(\boldsymbol{\mathcal { R }}_{\varepsilon}\right), \tag{5.24}
\end{equation*}
$$

where $\boldsymbol{\mathcal { R }}_{\varepsilon}$ is given by (5.20). Note that the MSE of the LMMSE estimator depends on the channel covariance matrix $\mathcal{R}$, whereas the MSE of the ML estimator does not depend on any a priori information about the channel. In case of i.i.d. fading with $\mathcal{R}=\mathbf{I}_{L}$, it follows from (5.21) that (5.24) reduces to

$$
\begin{equation*}
\mathrm{MSE}_{\mathrm{LMMSE}}=L \frac{N_{0}}{N_{0}+K_{\mathrm{p}} E_{\mathrm{p}}} . \tag{5.25}
\end{equation*}
$$

In Fig. 5.2, we show the MSE curves for both the ML and LMMSE channel estimators in case of a $2 \times 2$ MIMO channel with transmit and receive covariance matrices

$$
\boldsymbol{\mathcal { R }}_{\mathrm{t}}=\boldsymbol{\mathcal { R }}_{\mathrm{r}}=\left[\begin{array}{ll}
1 & \rho  \tag{5.26}\\
\rho & 1
\end{array}\right] .
$$

The results are shown for $\rho \in\{0,0.8\}$ and for $K_{p} \in\{2,20\}$. As expected, the MSE curves converge for high SNR, whereas for low SNR, the LMMSE channel estimator outperforms the ML channel estimator in terms of MSE.

### 5.4 Chapter Summary

In this chapter, we have presented two widespread pilot based channel estimation techniques, i.e., ML and LMMSE channel estimation. Both estimators allow the receiver to obtain an estimate of the MIMO channel, which is a priori unknown due to its random character. Also, we established a data frame model that incorporates both pilot and data symbols.

## 6

## BER Analysis of OSTBCs in Rayleigh Fading

In chapter 4, we showed that ML detection of the different information symbols in an OSTBC matrix reduces to symbol-by-symbol detection, based on linear signal processing at the receiver. More specifically, the decision variables corresponding to the transmitted information symbols are given by (4.42). Because the results from chapter 4 were obtained for a system with PCSI, they do not represent the realistic situation where the channel has to be estimated by the receiver, as illustrated in chapter 5 .

Under the assumption of Rayleigh fading, the impact of imperfect channel state information (ICSI) on the performance of OSTBCs has been investigated extensively in the literature. In [62], the effect of channel estimation errors on the BER was demonstrated by means of Monte-Carlo simulations. As indicated in section 3.3.2, however, this approach requires that all input RVs of the system, i.e., the $L$ complex-valued entries of the channel matrix $\mathbf{H}$, the information symbols in the data frame, and the AWGN channel noise matrices
are generated repeatedly according to their corresponding distributions. For each set of input RVs, the channel is estimated, the data symbols are recovered from the received signal, and the number of bit errors is counted. Finally, the BER is approximated as the ratio of the total number of bit errors to the total number of bits transmitted. Since the full diversity gain achieved by OSTBCs results in very low BERs, very long simulation times are usually necessary to obtain accurate BER results, even at moderate SNR. Therefore, Monte-Carlo simulations are, in general, inappropriate for accurate and efficient BER computations for OSTBCs, and analytical performance results provide a powerful alternative. Analytical SER expressions for OSTBCs in absence of PCSI were presented in [63] and [64]. In the case of M-PSK constellations, exact closedform BER expressions as well as tight upper bounds were given in [65] for pilot-based MMSE channel estimation. In [66], the effect of ML channel estimation on the performance of Alamouti's scheme was examined for QPSK modulation in rapid i.i.d. Rayleigh fading channels. High-SNR expressions for the pairwise error probability (PEP) were derived under quite general conditions in [67], using an eigenvalue approach. In [68], an exact closed-form expression for the PEP of both orthogonal and non-orthogonal space-time codes in the case of least-squares channel estimation was obtained by means of characteristic functions. In [69], this result was extended to the case of correlated Rayleigh fading with receive correlation only. However, from the PEP one can compute only an upper bound on the BER, which in a fading environment does not converge to the true BER at high SNR.

In order to investigate the impact of ICSI on the BER of OSTBCs, we consider in this chapter a receiver that uses the estimated channel in the same way as an ML receiver would apply the true channel; this type of receiver is usually called a mismatched ML receiver. In section 6.1, we present a simple rule of thumb that serves as an indicator for the BER degradation caused by imperfect channel estimation and yields the exact result for high SNR under certain conditions. Assuming LMMSE channel estimation and i.i.d. Rayleigh fading, we derive in section 6.2 an exact closed-form BER expression for square OSTBCs and an accurate approximation of the BER for non-square OSTBCs. In section 6.3, we present accurate BER approximations for both square and non-square OSTBCs under the assumption of arbitrarily correlated Rayleigh fading. The chapter is summarized in section 6.4. Although all BER expressions in this chapter are derived for QAM constellations, they can be easily modified to the case of PAM constellations.

### 6.1 Rule of Thumb

When the mismatched receiver and the PCSI receiver achieve some target BER at $E_{\mathrm{b}} / N_{0}=\left(E_{\mathrm{b}} / N_{0}\right)_{\text {ICSI }}$ and $E_{\mathrm{b}} / N_{0}=\left(E_{\mathrm{b}} / N_{0}\right)_{\text {PCSI }}$, respectively, the BER degradation of the mismatched receiver as compared to the PCSI receiver is
expressed as

$$
\begin{equation*}
\Delta_{\mathrm{BER}}=\frac{\left(E_{\mathrm{b}} / N_{0}\right)_{\mathrm{ICSI}}}{\left(E_{\mathrm{b}} / N_{0}\right)_{\mathrm{PCSI}}} \tag{6.1}
\end{equation*}
$$

Usually, the BER degradation is expressed in dB: $\Delta_{\mathrm{BER}, \mathrm{dB}}=10 \log _{10}\left(\Delta_{\mathrm{BER}}\right)$.
In case of PCSI, the information symbols included in the OSTBC matrix $\mathbf{C}$ are detected from the received signal

$$
\begin{equation*}
\mathbf{R}=\sqrt{E_{\mathrm{s}}} \mathbf{H C}+\mathbf{W} \tag{6.2}
\end{equation*}
$$

using the decision variables (4.42). Defining the channel estimation error matrix $\mathbf{E}$ as $\mathbf{E} \triangleq \mathbf{H}-\hat{\mathbf{H}}$, we can rewrite (6.2) for a mismatched receiver that assumes $\hat{\mathbf{H}}$ to be the correct channel matrix as

$$
\begin{equation*}
\mathbf{R}=\sqrt{E_{s}} \hat{\mathbf{H}} \mathbf{C}+\sqrt{E_{s}} \mathbf{E C}+\mathbf{W} \tag{6.3}
\end{equation*}
$$

where $\sqrt{E_{s}} \hat{H} \mathbf{C}$ is the useful component and $\sqrt{E_{S}} \mathbf{E C}$ is an interference term caused by imperfect channel estimation. If the useful term $\sqrt{E_{s}} \hat{H} C$ and the disturbance term $\sqrt{E_{s}} \mathbf{E C}+\mathbf{W}$ in (6.3) are uncorrelated and have similar statistics as the useful term $\sqrt{E_{s}} \mathbf{H C}$ and the noise term $\boldsymbol{W}$ in (6.2), respectively, the mismatched and PCSI receiver will achieve the same BER when operating at the same SNR, where the SNR is defined as the ratio of the energy of the useful term to the energy of the disturbance term. Since the noise term $\mathbf{W}$ consists of i.i.d. ZM CSCG RVs with variance $N_{0}$, we can adjust the SNR in (6.2) and (6.3) by selecting a proper symbol energy $E_{\mathrm{s}}$. Hence, for both receivers to achieve the same $\operatorname{SNR}, E_{\mathrm{S}}$ should be selected such that

$$
\begin{equation*}
\frac{\mathbb{E}\left[\left\|\sqrt{\left(E_{\mathrm{s}}\right)_{\mathrm{ICSI}}} \hat{\mathbf{H}} \mathbf{C}\right\|_{\mathrm{F}}^{2}\right]}{\mathbb{E}\left[\left\|\sqrt{\left(E_{\mathrm{s}}\right)_{\mathrm{ICSI}}} \mathbf{E C}+\mathbf{W}\right\|_{\mathrm{F}}^{2}\right]}=\frac{\mathbb{E}\left[\left\|\sqrt{\left(E_{\mathrm{s}}\right)_{\mathrm{PCSI}}} \mathbf{H C}\right\|_{\mathrm{F}}^{2}\right]}{\mathbb{E}\left[\|\mathbf{W}\|_{\mathrm{F}}^{2}\right]} \tag{6.4}
\end{equation*}
$$

Since the BER degradation is given by (6.1) and $N_{0}$ is kept fixed for both receivers, it follows from (4.40), (5.8), and (6.4) that

$$
\begin{align*}
\Delta_{\mathrm{BER}}=\frac{\left(E_{\mathrm{b}}\right)_{\mathrm{ICSI}}}{\left(E_{\mathrm{b}}\right)_{\mathrm{PCSI}}} & =\left(1+\frac{\eta K_{\mathrm{p}}}{K}\right) \frac{\left(E_{\mathrm{S}}\right)_{\mathrm{ICSI}}}{\left(E_{\mathrm{s}}\right)_{\mathrm{PCSI}}} \\
& =\left(1+\frac{\eta K_{\mathrm{p}}}{K}\right) \frac{\mathbb{E}\left[\|\mathbf{H C}\|_{\mathrm{F}}^{2}\right]}{\mathbb{E}\left[\|\hat{\mathbf{H} C}\|_{\mathrm{F}}^{2}\right]} \frac{\mathbb{E}\left[\left\|\sqrt{\left(E_{\mathrm{S}}\right)_{\mathrm{ICSI}}} \mathbf{E C}+\mathbf{W}\right\|_{\mathrm{F}}^{2}\right]}{\mathbb{E}\left[\|\mathbf{W}\|_{\mathrm{F}}^{2}\right]}, \tag{6.5}
\end{align*}
$$

which is a function of $\sqrt{\left(E_{\mathrm{S}}\right)_{\text {ICSI }}}$. Before we further simplify the expression (6.5), we take a look at the conditions for which (6.5) is exact.

- The useful term $\sqrt{E_{s}} \hat{H} \mathbf{C}$ and the disturbance term $\sqrt{E_{s}} \mathbf{E C}+\mathbf{W}$ in (6.3) must be uncorrelated. Clearly, this is the case when we use LMMSE channel estimation, for which the channel estimate $\hat{\mathbf{H}}$ and the estimation error $\mathbf{E}$ are known to be uncorrelated.
- The useful terms in (6.2) and (6.3) must have similar statistical properties. Since $\hat{\mathbf{H}}$ is a linear function of both the channel $\mathbf{H}$ and the noise matrix $\mathbf{W}_{\mathrm{p}}$, this condition will be satisfied when the channel coefficients are Rayleigh distributed, such that both $\mathbf{H}$ and $\hat{\mathbf{H}}$ consist of ZM CSCG RVs. Moreover, it follows from (5.18) that $\boldsymbol{\mathcal { R }}_{\hat{\mathrm{h}}}$ is proportional to $\boldsymbol{\mathcal { R }}$ only when $\mathcal{R}=\mathbf{I}_{L}$. Hence, (6.5) is exact for i.i.d. Rayleigh fading only.
- The interference term $\sqrt{E_{s}} \mathbf{E C}$ must consist of i.i.d. ZM CSCG RVs, with a variance that does not depend on the values of the information symbols. Since it follows from (5.21) that the estimation error E consist of i.i.d. ZM CSCG RVs with variance $N_{0} /\left(K_{\mathrm{p}} E_{\mathrm{p}}+N_{0}\right)$ in case of i.i.d. Rayleigh fading, the entries of the interference term are ZM CSCG RVs. The correlation between the entries of the interference term is given by

$$
\begin{equation*}
\mathbb{E}\left[\left(\sqrt{E_{\mathrm{s}}} \mathbf{E C}\right)_{m, k}\left(\sqrt{E_{\mathrm{s}}} \mathbf{E C}\right)_{m^{\prime}, k^{\prime}}^{*}\right]=\frac{E_{\mathrm{s}} N_{0}}{K_{\mathrm{p}} E_{\mathrm{p}}+N_{0}} \delta_{m-m^{\prime}}\left(\mathbf{C}^{H} \mathbf{C}\right)_{k^{\prime}, k} \tag{6.6}
\end{equation*}
$$

Hence, the components of $\sqrt{E_{S}} \mathrm{EC}$ are spatially uncorrelated but a temporal correlation might exist, unless $\mathbf{C}^{H} \mathbf{C}$ is a diagonal matrix, which is the case for square OSTBCs. Because of (4.36), however, the variance of the interference term is still a function of the symbol vector $\mathbf{s}$, unless the symbols belong to a PSK constellation. Hence, for square OSTBCs and PSK constellations, we have

$$
\begin{equation*}
\mathbb{E}\left[\left(\sqrt{E_{\mathrm{s}}} \mathbf{E C}\right)_{m, k}\left(\sqrt{E_{\mathrm{s}}} \mathbf{E C}\right)_{m^{\prime}, k^{\prime}}^{*}\right]=\frac{E_{\mathrm{s}} N_{0} L_{\mathrm{t}}}{K_{\mathrm{p}} E_{\mathrm{p}}+N_{0}} \delta_{m-m^{\prime}} \delta_{k-k^{\prime}} \tag{6.7}
\end{equation*}
$$

Assuming LMMSE channel estimation, i.i.d. Rayleigh fading, and PSK constellations, the different factors in (6.5) are easily obtained:

$$
\begin{gather*}
\mathbb{E}\left[\|\mathbf{H C}\|_{\mathrm{F}}^{2}\right]=\mathbb{E}\left[\operatorname{tr}\left(\mathbf{H C C}{ }^{H} \mathbf{H}^{H}\right)\right]=L_{\mathrm{t}} \mathbb{E}\left[\|\mathbf{H}\|_{\mathrm{F}}^{2}\right]=K_{\mathrm{c}} L_{\mathrm{t}} L_{\mathrm{r}},  \tag{6.8}\\
\mathbb{E}\left[\|\hat{\mathbf{H} C}\|_{\mathrm{F}}^{2}\right]=K_{\mathrm{c}} L_{\mathrm{t}} L_{\mathrm{r}} \frac{K_{\mathrm{p}} E_{\mathrm{p}}}{K_{\mathrm{p}} E_{\mathrm{p}}+N_{0}},  \tag{6.9}\\
\mathbb{E}\left[\left\|\sqrt{\left(E_{\mathrm{S}}\right)_{\mathrm{ICSI}}} \mathbf{E C}+\mathbf{W}\right\|_{\mathrm{F}}^{2}\right]=K_{\mathrm{c}} L_{\mathrm{t}} L_{\mathrm{r}} \frac{\left(E_{\mathrm{s}}\right)_{\mathrm{ICSI}} N_{0}}{K_{\mathrm{p}} E_{\mathrm{p}}+N_{0}}+K_{\mathrm{c}} L_{\mathrm{r}} N_{0},  \tag{6.10}\\
\mathbb{E}\left[\|\mathbf{W}\|_{\mathrm{F}}^{2}\right]=K_{\mathrm{c}} L_{\mathrm{r}} N_{0} . \tag{6.11}
\end{gather*}
$$

Using (6.8)-(6.11), the BER degradation (6.5) reduces to

$$
\begin{equation*}
\Delta_{\mathrm{BER}}=\left(1+\frac{\eta K_{\mathrm{p}}}{K}\right)\left(1+\frac{L_{\mathrm{t}}}{\eta K_{\mathrm{p}}}+\frac{N_{0}}{K_{\mathrm{p}} E_{\mathrm{p}}}\right) \tag{6.12}
\end{equation*}
$$

which is a function of $E_{\mathrm{p}} / N_{0}$. Note that (6.12) is only exact for square OSTBCs, yet it can be used as an approximation for other OSTBCs as well. Since it
follows from (5.17) that for i.i.d. fading, LMMSE and ML channel estimation differ only by a scale factor, which does not affect the detection in case of PSK symbols, the BER degradation (6.12) can also be used for ML channel estimation. Moreover, in case of spatial multiplexing MIMO systems transmitting $L_{t}$ independent uncoded symbol streams over i.i.d. Rayleigh fading channels using a PSK constellation, $\mathbf{C}$ reduces to an $L_{t} \times 1$ vector and $\mathbf{C}^{H} \mathbf{C}=L_{t}$, such that (6.12) also holds for spatial multiplexing and uncoded SIMO communication $\left(L_{t}=1\right)$. For high SNR, (6.12) reduces to

$$
\begin{equation*}
\Delta_{\mathrm{BER}}=\left(1+\frac{\eta K_{\mathrm{p}}}{K}\right)\left(1+\frac{L_{\mathrm{t}}}{\eta K_{\mathrm{p}}}\right) \tag{6.13}
\end{equation*}
$$

which is a function of the structure of the data frames $\left(L_{t}, K, K_{\mathrm{p}}\right)$ and the ratio of pilot to symbol energy $(\eta)$ only. Let us recall that the simple yet important expression (6.13) yields the asymptotic BER degradation due to channel estimation errors, for square OSTBCs with PSK symbols under i.i.d. Rayleigh fading. For non-square OSTBCs, PAM or QAM constellations, or arbitrary fading distributions, (6.13) is useful to approximate the actual BER degradation and to estimate the optimal number of pilot symbols. Note that the first factor in (6.13) grows with $K_{p}$, since adding pilot symbols decreases the energy available for the transmission of data symbols, whereas the second term decreases with growing $K_{\mathrm{p}}$, since using more pilot symbols improves channel estimation. Because of this trade-off, it is possible, for given $K$, to find an optimal value for $K_{p}$ that minimizes the BER degradation of the mismatched receiver with respect to the PCSI receiver. It is easily derived that the optimal value for $K_{p}$ is given by

$$
\begin{equation*}
K_{\mathrm{p}, \mathrm{opt}}=\frac{\sqrt{L_{\mathrm{t}} K}}{\eta} \tag{6.14}
\end{equation*}
$$

such that the minimal BER degradation becomes

$$
\begin{equation*}
\Delta_{\mathrm{BER}, \min }=\left(1+\sqrt{\frac{L_{\mathrm{t}}}{K}}\right)^{2} \tag{6.15}
\end{equation*}
$$

which for large $K$ asymptotically approaches 1 (or 0 dB ). For best performance we should take $K$ as high as possible, taking into account that the frame length must not exceed the coherence time of the channel. When $\eta=1$, i.e., when the pilot energy $E_{\mathrm{p}}$ equals the symbol energy $E_{\mathrm{s}}$, the optimal number of pilot symbols is given by $\sqrt{L_{\mathrm{t}} K}$. If we allow $\eta$ to be higher than 1 , however, this degree of freedom can be used to reduce the number of pilot symbols $K_{p}$ below $\sqrt{L_{\mathrm{t}} K}$ in order to increase the information bit rate without increasing the BER degradation. However, the lower $K_{p}$, the larger $E_{p}$ to maintain optimal performance, such that higher peak transmit powers are needed at the transmitter side. Also, in order that the rows of $\mathbf{A}_{p}$ be orthogonal, $K_{p}$ should not be less than $L_{t}$.

### 6.2 BER Analysis for i.i.d. Rayleigh Fading

Since the mismatched receiver uses the estimated channel $\hat{\mathbf{H}}$ in the same way as an ML receiver would apply $\mathbf{H}$, its detection algorithm is given by

$$
\begin{equation*}
\hat{\mathbf{C}}=\arg \min _{\widetilde{\mathrm{C}}}\left\|\mathbf{R}-\sqrt{E_{\mathrm{S}}} \hat{\mathbf{H}} \widetilde{\mathbf{C}}\right\|_{\mathrm{F}^{\prime}}^{2} \tag{6.16}
\end{equation*}
$$

where the minimization is over the valid code matrices $\widetilde{\mathbf{C}}$ satisfying (4.28). In this way, the detection algorithm still reduces to symbol-by-symbol detection

$$
\begin{equation*}
\hat{s}_{i}=\arg \min _{\tilde{s} \in \Psi}\left|u_{i}-\tilde{s}\right|, \quad 1 \leq i \leq N_{\mathrm{S}} \tag{6.17}
\end{equation*}
$$

where the minimization is over the symbols $\tilde{s}$ belonging to the considered constellation $\Psi$ and the decision variables $u_{i}$ are obtained by replacing the channel $\mathbf{H}$ in (4.42) by $\hat{\mathbf{H}}$

$$
\begin{equation*}
u_{i}=\frac{\operatorname{tr}\left(\mathbf{C}_{i}^{H} \hat{\mathbf{H}}^{H} \mathbf{R}+\mathbf{R}^{H} \hat{\mathbf{H}} \mathbf{C}_{i}^{\prime}\right)}{\lambda \sqrt{E_{\mathrm{s}}}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}} \tag{6.18}
\end{equation*}
$$

According to (6.3), we can write the signal $\mathbf{R}$ captured by a mismatched receiver as a function of the estimated channel $\hat{\mathbf{H}}$. As compared to a receiver with PCSI, the detection performance of the mismatched receiver is clearly degraded: the total noise variance is increased, because of the presence of the interference term $\sqrt{E_{S}} \mathbf{E C}$, whereas the useful component is reduced, since it follows from (5.19) that the variance of the estimated channel coefficients is smaller than the unit variance that characterizes a normalized channel. By substituting $\mathbf{R}$ in (6.18) by (6.3), the decision variable can be shown to reduce to

$$
\begin{equation*}
u_{i}=s_{i}+n_{i}, \tag{6.19}
\end{equation*}
$$

which is the sum of the transmitted symbol $s_{i}$ and a disturbance term $n_{i}$. Unlike the case of PCSI, however, the disturbance term $n_{i}$ contains contributions not only from the channel noise $\mathbf{W}$ but also from the channel estimation error E. It is readily verified that $n_{i}$ is given by $n_{i}=e_{i}+w_{i}$, with

$$
\begin{gather*}
e_{i} \triangleq \frac{\operatorname{tr}\left(\mathbf{C}_{i}^{H} \hat{\mathbf{H}}^{H} \mathbf{E C}+\mathbf{C}^{H} \mathbf{E}^{H} \hat{\mathbf{H}} \mathbf{C}_{i}^{\prime}\right)}{\lambda\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}},  \tag{6.20}\\
w_{i} \triangleq \frac{\operatorname{tr}\left(\mathbf{C}_{i}^{H} \hat{\mathbf{H}}^{H} \mathbf{W}+\mathbf{W}^{H} \hat{\mathbf{H}} \mathbf{C}_{i}^{\prime}\right)}{\lambda \sqrt{E_{\mathrm{s}}}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}} \tag{6.21}
\end{gather*}
$$

In case of QAM constellations, the real and imaginary parts of $s_{i}$ can be detected separately from the real and imaginary parts of the decision variable
$u_{i}$, respectively

$$
\begin{gather*}
\hat{s}_{i, \mathrm{R}}=\arg \min _{\tilde{s} \in \Psi^{\prime}}\left|u_{i, \mathrm{R}}-\tilde{s}\right|^{2},  \tag{6.22a}\\
\hat{s}_{i, \mathrm{I}}=\arg \min _{\tilde{s} \in \Psi^{\prime}}\left|u_{i, \mathrm{I}}-\tilde{s}\right|^{2}, \tag{6.22b}
\end{gather*}
$$

where $\Psi^{\prime}$ is defined in (3.14) as the set consisting of the real (and imaginary) parts of the QAM symbols, $\hat{s}_{i, R}=\Re\left[\hat{s}_{i}\right], \hat{s}_{i, \mathrm{I}}=\Im\left[\hat{s}_{i}\right], u_{i, \mathrm{R}}=\Re\left[u_{i}\right]$, and $u_{i, \mathrm{I}}=$ $\Im\left[u_{i}\right]$. Hence, the average BER can be calculated as

$$
\begin{equation*}
P_{\mathrm{b}}=\frac{1}{2 N_{\mathrm{s}}} \sum_{i=1}^{N_{\mathrm{s}}}\left[P_{\mathrm{b}, i, \mathrm{R}}+P_{\mathrm{b}, i, \mathrm{l}}\right], \tag{6.23}
\end{equation*}
$$

where $P_{\mathrm{b}, i, \mathrm{R}}$ and $P_{\mathrm{b}, i \mathrm{I}}$ denote the BERs of the in-phase bits and quadrature bits corresponding to the information symbols $s_{i}$, respectively. Using a similar approach as in section 3.3.3, $P_{\mathrm{b}, i, \mathrm{R}}$ and $P_{\mathrm{b}, i, \mathrm{I}}$ can be obtained from the variances of the real and imaginary parts of $n_{i}$, respectively. To this end, we introduce the $L_{\mathrm{t}} \times L_{\mathrm{t}}$ matrices $\mathbf{C}_{i, R}(\mathbf{s})$ and $\mathbf{C}_{i, I}(\mathbf{s})$ as

$$
\begin{align*}
& \mathbf{C}_{i, \mathrm{R}}(\mathbf{s}) \triangleq \mathbf{C}\left(\mathbf{C}_{i}+\mathbf{C}_{i}^{\prime}\right)^{H},  \tag{6.24a}\\
& \mathbf{C}_{i, 1}(\mathbf{s}) \triangleq \mathbf{C}\left(\mathbf{C}_{i}-\mathbf{C}_{i}^{\prime}\right)^{H}, \tag{6.24b}
\end{align*}
$$

which depend on the information symbol vector $\mathbf{s}$ through the code matrix $\mathbf{C}$. Using (6.24), we show in appendix 6.A. 1 that, when conditioned on $\hat{\mathbf{H}}$ and s, $n_{i}$ is a ZM non-circularly symmetric complex Gaussian RV, the variances of the real and imaginary parts of which are given by

$$
\begin{align*}
& \sigma_{i, \mathrm{R}}^{2}(\hat{\mathbf{H}}, \mathbf{s}) \triangleq \mathbb{E}\left[\left(\Re\left[n_{i}\right]\right)^{2} \mid \hat{\mathbf{H}}, \mathbf{s}\right]=\frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}}\left(1+\frac{\left\|\hat{\mathbf{H}} \mathbf{C}_{i, \mathrm{R}}^{H}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{\lambda\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}\left(\eta K_{\mathrm{p}}+\frac{N_{0}}{E_{\mathrm{s}}}\right)}\right),  \tag{6.25a}\\
& \sigma_{i, \mathrm{I}}^{2}(\hat{\mathbf{H}}, \mathbf{s}) \triangleq \mathbb{E}\left[\left(\Im\left[n_{i}\right]\right)^{2} \mid \hat{\mathbf{H}}, \mathbf{s}\right]=\frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}}\left(1+\frac{\left\|\hat{\mathbf{H}} \mathbf{C}_{i, \mathrm{I}}^{H}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{\lambda\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}\left(\eta K_{\mathrm{p}}+\frac{N_{0}}{E_{\mathrm{s}}}\right)}\right) . \tag{6.25b}
\end{align*}
$$

Since (6.25a) and (6.25b) are obtained for a given $\hat{\mathbf{H}}$ and $\mathbf{s}$, we can easily calculate the conditional BERs of the in-phase and quadrature bits corresponding to the information symbols $s_{i}$, conditioned on $\hat{\mathbf{H}}$ and $\mathbf{s}$. Taking into account that a decision error occurs when $u_{i, \mathrm{q}}$, with $\mathrm{q}=\mathrm{R}$ or $\mathrm{q}=\mathrm{I}$, is located inside the decision area of the projection $b_{\mathrm{q}}$ of a symbol $b$, with $b_{\mathrm{q}} \neq s_{i, \mathrm{q}}$, the conditional BER of the bits allocated to $s_{i, q}$ is given by

$$
\begin{equation*}
P_{\mathrm{b}, i, \mathrm{q}}(\hat{\mathbf{H}}, \mathbf{s})=\frac{1}{\log _{2}(\sqrt{M})} \sum_{b_{\mathrm{q}} \in \Psi^{\prime}} d_{\mathrm{H}}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right) \operatorname{Pr}\left[\hat{\hat{s}}_{i, \mathrm{q}}=b_{\mathrm{q}} \mid \hat{\mathbf{H}}, \mathbf{s}\right], \tag{6.26}
\end{equation*}
$$

where $d_{\mathrm{H}}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)$ denotes the Hamming distance between the bits associated to $s_{i, \mathrm{q}}$ and $b_{\mathrm{q}}$, respectively. Using the distances $D_{1}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)$ and $D_{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)$ between $s_{i, \mathrm{q}}$ and the boundaries of the decision area of $b_{\mathrm{q}}$, as defined in (3.60), the conditional probability in (6.26) is easily shown to reduce to

$$
\begin{align*}
\operatorname{Pr}\left[\hat{s}_{i, \mathrm{q}}=b_{\mathrm{q}} \mid \hat{\mathbf{H}}, \mathbf{s}\right] & =\operatorname{Pr}\left[d_{1}\left(b_{\mathrm{q}}\right) \leq u_{i, \mathrm{q}} \leq d_{2}\left(b_{\mathrm{q}}\right) \mid \hat{\mathbf{H}}, \mathbf{s}\right] \\
& =Q\left(\sqrt{\frac{D_{1}^{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)}{\sigma_{i, \mathrm{q}}^{2}(\hat{\mathbf{H}}, \mathbf{s})}}\right)-Q\left(\sqrt{\frac{D_{2}^{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)}{\sigma_{i, \mathrm{q}}^{2}(\hat{\mathbf{H}}, \mathbf{s})}}\right) \tag{6.27}
\end{align*}
$$

where the variances $\sigma_{i, \mathrm{q}}^{2}(\hat{\mathbf{H}}, \mathbf{s})$ are given by (6.25) and $Q(x)$ is the Gaussian $Q$-function (3.58).

In order to obtain $P_{\mathrm{b}, i, \mathrm{q}}$, the conditional BER (6.26) must be averaged over the statistics of $\hat{\mathbf{H}}$ and $\mathbf{s}$; the latter operation reduces to a finite summation over all $M^{N_{s}}$ possible realisations of $\mathbf{s}$ :

$$
\begin{equation*}
P_{\mathrm{b}, i, \mathrm{q}}=\frac{1}{M^{N_{\mathrm{s}}}} \sum_{\mathbf{s} \in \Psi^{N_{\mathrm{s}}}} \mathbb{E}_{\hat{\mathbf{H}}}\left[P_{\mathrm{b}, i, \mathrm{q}}(\hat{\mathbf{H}}, \mathbf{s})\right] . \tag{6.28}
\end{equation*}
$$

Because of the summation over $\mathbf{s}$, the computational complexity of the BER expression (6.28) is proportional to $M^{N_{s}}$, which may become quite large in case of large symbol constellations and OSTBC matrices comprising many information symbols. If the associated computational complexity is too high, however, the summation over s can always be accurately approximated by Monte-Carlo integration, as will be shown in section 7.3.1.2. The expectation over $\hat{\mathbf{H}}$ of the conditional BER in (6.28) requires averaging the $Q$-functions in (6.27) over the statistics of $\hat{\mathbf{H}}$. Because of the second factor in the right-hand sides of (6.25a) and (6.25b), however, the Q-functions are generally a complicated function of all entries of $\hat{\mathbf{H}}$ and a closed-form BER expression is hard to derive. Although the average BER can be obtained by numerically averaging the $Q$-functions in (6.27) over $\hat{\mathbf{H}}$ through, e.g., Monte-Carlo integration, we present in this section an exact closed-form expression for the special case of square OSTBCs as well as an approximate closed-form BER expression for non-square OSTBCs.

Considering the BER analysis in section 4.2.2 for OSTBCs in case of PCSI, it is understood that a closed-form BER expression for ICSI can be easily obtained if the following two conditions are satisfied:

- the variances (6.25a) and (6.25b) are a function of $\hat{\mathbf{H}}$ through an inverse proportionality to the squared Frobenius norm $\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}$ of the channel estimate only;
- $\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}$ is distributed according to a (scaled) $\chi^{2}$-distribution.

Clearly, the first condition is met when

$$
\begin{align*}
\mathbf{C}_{i, \mathrm{R}}^{H}(\mathbf{s}) \mathbf{C}_{i, \mathrm{R}}(\mathbf{s}) & =\left(\mathbf{C}_{i}+\mathbf{C}_{i}^{\prime}\right) \mathbf{C}^{H} \mathbf{C}\left(\mathbf{C}_{i}+\mathbf{C}_{i}^{\prime}\right)^{H}=\beta_{i, \mathrm{R}}(\mathbf{s}) \mathbf{I}_{L_{\mathrm{t}^{\prime}}}  \tag{6.29a}\\
\mathbf{C}_{i, \mathrm{I}}^{H}(\mathbf{s}) \mathbf{C}_{i, \mathrm{I}}(\mathbf{s}) & =\left(\mathbf{C}_{i}-\mathbf{C}_{i}^{\prime}\right) \mathbf{C}^{H} \mathbf{C}\left(\mathbf{C}_{i}-\mathbf{C}_{i}^{\prime}\right)^{H}=\beta_{i, \mathrm{I}}(\mathbf{s}) \mathbf{I}_{L_{\mathrm{t}}} \tag{6.29b}
\end{align*}
$$

since in this way the variances (6.25a) and (6.25b) reduce to

$$
\begin{align*}
\sigma_{i, \mathrm{R}}^{2}(\hat{\mathbf{H}}, \mathbf{s}) & =\frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}}\left(1+\frac{\beta_{i, \mathrm{R}}(\mathbf{s})}{\lambda\left(\eta K_{\mathrm{p}}+\frac{N_{0}}{E_{\mathrm{s}}}\right)}\right)  \tag{6.30a}\\
\sigma_{i, \mathrm{I}}^{2}(\hat{\mathbf{H}}, \mathbf{s}) & =\frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}}\left(1+\frac{\beta_{i, \mathrm{I}}(\mathbf{s})}{\lambda\left(\eta K_{\mathrm{p}}+\frac{N_{0}}{E_{\mathrm{s}}}\right)}\right) \tag{6.30b}
\end{align*}
$$

The second condition is also satisfied, since it follows from (5.19) that, in case of i.i.d. Rayleigh fading and LMMSE channel estimation, the PDF of $\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}$ is given by (4.68), with $\sigma^{2}=K_{\mathrm{p}} E_{\mathrm{p}} /\left(K_{\mathrm{p}} E_{\mathrm{p}}+N_{0}\right)$. Hence, when (6.29) is satisfied, a closed-form BER expression can be easily found, since (4.11) yields

$$
\begin{align*}
\frac{1}{\left(\sigma^{2}\right)^{L}(L-1)!} & \int_{0}^{\infty} Q(\sqrt{\beta x}) x^{L-1} \exp \left(-\frac{x}{\sigma^{2}}\right) \mathrm{d} x \\
& =\frac{1}{(L-1)!} \int_{0}^{\infty} Q\left(\sqrt{\beta \sigma^{2} x}\right) x^{L-1} \exp (-x) \mathrm{d} x \\
& =\Omega_{L}\left(\beta \sigma^{2}\right) \tag{6.31}
\end{align*}
$$

where $\beta>0$ and $\Omega_{L}(\theta)$ is defined as

$$
\begin{equation*}
\Omega_{L}(x) \triangleq\left[\frac{1}{2}\left(1-\sqrt{\frac{x}{2+x}}\right)\right]^{L} \sum_{k=0}^{L-1}\binom{L-1+k}{k}\left[\frac{1}{2}\left(1+\sqrt{\frac{x}{2+x}}\right)\right]^{k} \tag{6.32}
\end{equation*}
$$

### 6.2.1 Exact BER for Square OSTBCs

Taking (4.36) into account, it is readily verified that (6.29) holds for square OSTBCs

$$
\begin{equation*}
\mathbf{C}_{i, \mathrm{R}}^{H}(\mathbf{s}) \mathbf{C}_{i, \mathrm{R}}(\mathbf{s})=\mathbf{C}_{i, \mathrm{I}}^{H}(\mathbf{s}) \mathbf{C}_{i, \mathrm{I}}(\mathbf{s})=\lambda^{2}\|\mathbf{s}\|^{2} \mathbf{I}_{L_{\mathrm{t}}} \tag{6.33}
\end{equation*}
$$

In this way, the variances of the real and imaginary part of $n_{i}$ are given by

$$
\begin{equation*}
\sigma_{i, \mathrm{R}}^{2}(\hat{\mathbf{H}}, \mathbf{s})=\sigma_{i, \mathrm{I}}^{2}(\hat{\mathbf{H}}, \mathbf{s})=\frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}}\left(1+\frac{\lambda\|\mathbf{s}\|^{2}}{\eta K_{\mathrm{p}}+\frac{N_{0}}{E_{\mathrm{s}}}}\right) \tag{6.34}
\end{equation*}
$$

such that $n_{i}$ is a ZM CSCG RV, the variance of which does not depend on the index $i$. Hence, the BERs of the in-phase and quadrature bits allocated to the information symbols $s_{i}$ are identical and (6.23) reduces to

$$
\begin{equation*}
P_{\mathrm{b}}=P_{\mathrm{b}, i, \mathrm{q}} \tag{6.35}
\end{equation*}
$$

for arbitrary $i$ and q. Again, $P_{\mathrm{b}, i, \mathrm{q}}$ is obtained from (6.26)-(6.28), the only difference being that $\sigma_{i, q}^{2}(\hat{\mathbf{H}}, \mathbf{s})$ in (6.27) is now given by (6.34), such that the argument of the $Q$-functions in (6.27) is proportional to the squared Frobenius norm $\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}$ of the channel estimate. According to (6.31), averaging the $Q$-functions in (6.27) over (4.68), with $\sigma^{2}=K_{\mathrm{p}} E_{\mathrm{p}} /\left(K_{\mathrm{p}} E_{\mathrm{p}}+N_{0}\right)$, yields the following closed-form expression

$$
\begin{equation*}
\mathbb{E}_{\hat{\mathbf{H}}}\left[Q\left(\sqrt{\frac{D_{j}^{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)}{\sigma_{i, \mathrm{q}}^{2}(\hat{\mathbf{H}}, \mathbf{s})}}\right)\right]=\Omega_{L}\left(\frac{2 \lambda D_{j}^{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)}{1+\frac{\lambda\|\mathbf{s}\|^{2}}{\eta K_{\mathrm{p}}}+\frac{1}{\eta K_{\mathrm{p}}} \frac{N_{0}}{E_{\mathrm{s}}}} \frac{E_{\mathrm{s}}}{N_{0}}\right), \tag{6.36}
\end{equation*}
$$

where $j \in\{1,2\}$ and $\Omega_{L}(\cdot)$ is defined in (6.32). Note that the right-hand side of (6.36) is a function of the transmitted symbol vector $s$ through the squared norm $\|\mathbf{s}\|^{2}$ only, such that $\mathbf{s}, \mathbf{s}^{*},-\mathbf{s}$, and $-\mathbf{s}^{*}$ yield the same result. Hence, taking the symmetry of the QAM constellation into account, it follows that the BERs of the bits allocated to $s_{i, \mathrm{q}}$ and $-s_{i, \mathrm{q}}$ are identical, and independent of the sign of the real and imaginary parts of all symbols in the symbol vector s. Therefore, we can restrict the summation over s to constellation points with positive real and imaginary parts only. To this end, we introduce $\Psi_{0}^{\prime}$ as the set consisting of the positive elements of the set $\Psi^{\prime}$, which was defined in (3.14):

$$
\begin{equation*}
\Psi_{0}^{\prime}=\left\{(2 i-1) d_{\mathrm{QAM}}: i=1, \ldots, \sqrt{M} / 2\right\} \tag{6.37}
\end{equation*}
$$

where $d_{\mathrm{QAM}}$ is given by (3.15). Consequently, the set consisting of the QAM symbols with positive real and imaginary parts is given by

$$
\begin{equation*}
\Psi_{0}=\left\{\psi: \Re[\psi], \Im[\psi] \in \Psi_{0}^{\prime}\right\} \tag{6.38}
\end{equation*}
$$

Taking the above considerations into account, it follows from (6.26)-(6.28) and (6.36) that the BER (6.35) for square OSTBCs on i.i.d. Rayleigh fading channels with LMMSE channel estimation can be written in closed-form as

$$
\begin{align*}
& P_{\mathrm{b}}=\left(\frac{4}{M}\right)^{N_{\mathrm{s}}} \frac{1}{\log _{2}(\sqrt{M})} \sum_{\mathrm{s} \in \Psi_{0}^{N_{\mathrm{s}}}} \sum_{\mathrm{b}_{\mathrm{q}} \in \Psi^{\prime}} d_{\mathrm{H}}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right) \\
& \times\left[\Omega_{L}\left(\frac{2 \lambda D_{1}^{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)}{1+\frac{\lambda\|\mathbf{s}\|^{2}}{\eta K_{\mathrm{p}}}+\frac{1}{\eta K_{\mathrm{p}}} \frac{N_{0}}{E_{\mathrm{s}}}} \frac{E_{\mathrm{s}}}{N_{0}}\right)-\Omega_{L}\left(\frac{2 \lambda D_{2}^{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)}{1+\frac{\lambda\|\mathbf{s}\|^{2}}{\eta K_{\mathrm{p}}}+\frac{1}{\eta K_{\mathrm{p}}} \frac{N_{0}}{E_{\mathrm{s}}}} \frac{E_{\mathrm{s}}}{N_{0}}\right)\right], \tag{6.39}
\end{align*}
$$

such that the computational complexity scales with $(M / 4)^{N_{s}}$. Since it follows from (3.65) that for large $x$ (6.32) reduces to

$$
\begin{equation*}
\Omega_{L}(x) \approx\left(\frac{1}{2 x}\right)^{L}\binom{2 L-1}{L} \tag{6.40}
\end{equation*}
$$

it is readily verified that the asymptotic behavior of the BER expression (6.39) is given by

$$
\begin{align*}
P_{\mathrm{b}}^{(\mathrm{as})} \approx\left(\frac{4}{M}\right)^{N_{\mathrm{s}}} \frac{(4 \lambda)^{-L}}{\log _{2}(\sqrt{M})}\binom{2 L-1}{L} & \left(\frac{E_{\mathrm{s}}}{N_{0}}\right)^{-L} \\
& \times \sum_{\mathbf{s} \in \Psi_{0}^{N_{\mathrm{s}}}} \xi_{L}\left(s_{i, \mathrm{q}}\right)\left(1+\frac{\lambda\|\mathbf{s}\|^{2}}{\eta K_{\mathrm{p}}}\right)^{L} \tag{6.41}
\end{align*}
$$

where the function $\xi_{L}\left(s_{\mathrm{q}}\right)$ is defined as

$$
\begin{equation*}
\xi_{L}\left(s_{\mathrm{q}}\right) \triangleq \sum_{b_{\mathrm{q}} \in \Psi^{\prime}} d_{\mathrm{H}}\left(s_{\mathrm{q}}, b_{\mathrm{q}}\right)\left(\frac{1}{D_{1}^{2 L}\left(s_{\mathrm{q}}, b_{\mathrm{q}}\right)}-\frac{1}{D_{2}^{2 L}\left(s_{\mathrm{q}}, b_{\mathrm{q}}\right)}\right) \tag{6.42}
\end{equation*}
$$

Note that, in case of 4-QAM, $\|\mathbf{s}\|^{2}=N_{\mathrm{S}}$, such that the dependence of the terms in (6.39) and (6.41) on the full symbol vector $\mathbf{s}$ reduces to a dependence on $s_{i, q}$ only. In this way, the summation over all possible symbol vectors $s$ reduces to a summation over $s_{i, q}$ only, which significantly simplifies the computational complexity of the BER calculation. Furthermore, it follows from (6.41) that LMMSE channel estimation does not affect the achieved diversity gain, which is also observed in [61]. Hence, the BER curves for perfect and imperfect CSI are parallel for large SNR and the degradation due to pilot-based LMMSE channel estimation can be obtained. Taking (5.8) and the high-SNR approximation (6.41) into account, the ratio of the BER of the mismatched receiver to the BER of the PCSI receiver is easily derived (at high $E_{\mathrm{b}} / N_{0}$ ):

$$
\begin{align*}
\frac{P_{\mathrm{b}, \mathrm{LMMSE}}^{(\mathrm{as})}}{P_{\mathrm{b}, \mathrm{PCSI}}^{(\mathrm{as})}}=\left(1+\frac{\eta K_{\mathrm{p}}}{K}\right)^{L}\left(\frac{4}{M}\right)^{N_{\mathrm{s}}-\frac{1}{2}} & \\
& \times \frac{\sum_{\mathbf{s} \in \Psi_{0}^{N_{\mathrm{s}}}}\left[\xi_{L}\left(s_{i, \mathrm{q}}\right)\left(1+\frac{\lambda\|\mathbf{s}\|^{2}}{\eta K_{\mathrm{p}}}\right)^{L}\right]}{\sum_{s_{\mathrm{q}} \in \Psi_{0}^{\prime}}\left[\xi_{L}\left(s_{\mathrm{q}}\right)\right]} \tag{6.43}
\end{align*}
$$

As stated before, the mismatched receiver must have a larger $E_{\mathrm{b}} / N_{0}$ ratio than the PCSI receiver, in order that both receivers have the same BER. Given that a diversity order of $L=L_{\mathrm{r}} L_{\mathrm{t}}$ is achieved, the amount (in dB) by which the $E_{\mathrm{b}} / N_{0}$ ratio of the mismatched receiver should be increased to obtain the same BER as the PCSI receiver is given by

$$
\begin{equation*}
\Delta_{\mathrm{BER}, \mathrm{~dB}}=\frac{10}{L} \log _{10}\left(\frac{P_{\mathrm{b}, \mathrm{LMMSE}}^{(\mathrm{as})}}{P_{\mathrm{b}, \mathrm{PCSI}}^{(\mathrm{as})}}\right) . \tag{6.44}
\end{equation*}
$$

In case of 4-QAM, (6.44) is easily shown to reduce to

$$
\begin{equation*}
\Delta_{\mathrm{BER}, \mathrm{~dB}}=10 \log _{10}\left[\left(1+\frac{\eta K_{\mathrm{p}}}{K}\right)\left(1+\frac{L_{\mathrm{t}}}{\eta K_{\mathrm{p}}}\right)\right], \tag{6.45}
\end{equation*}
$$

which corresponds to the result from (6.13).

### 6.2.2 Approximate BER for Non-Square OSTBCs

For non-square OSTBCs, (6.29) is not satisfied and an exact closed-form BER expression can not be obtained. However, an approximate closed-form expression can still be derived if we substitute $\left\|\hat{\mathbf{H}} \mathbf{C}_{i, \mathrm{R}}^{H}(\mathbf{s})\right\|_{\mathrm{F}}^{2}$ and $\left\|\hat{\mathbf{H}} \mathbf{C}_{i, \mathrm{R}}^{H}(\mathbf{s})\right\|_{\mathrm{F}}^{2}$ in (6.25) by their expectations over $\hat{\mathbf{H}}$, conditioned on the Frobenius norm $\|\hat{\mathbf{H}}\|_{\mathrm{F}}$ :

$$
\begin{align*}
\mathbb{E}_{\hat{\mathbf{H}} \mid\|\hat{\mathbf{H}}\|_{\mathrm{F}}}\left[\left\|\hat{\mathbf{H}}_{i, \mathrm{R}}^{H}(\mathbf{s})\right\|_{\mathrm{F}}^{2}\right] & =\frac{\left\|\mathbf{C}_{i, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{L_{\mathrm{t}}}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}  \tag{6.46a}\\
\mathbb{E}_{\hat{\mathbf{H}} \mid\|\hat{\mathbf{H}}\|_{\mathrm{F}}}\left[\left\|\hat{\mathbf{H}} \mathrm{C}_{i, \mathrm{I}}^{H}(\mathbf{s})\right\|_{\mathrm{F}}^{2}\right] & =\frac{\left\|\mathbf{C}_{i, \mathrm{I}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{L_{\mathrm{t}}}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2} . \tag{6.46b}
\end{align*}
$$

Note that averaging over $\hat{\mathbf{H}}$, conditioned on the Frobenius norm $\|\hat{\mathbf{H}}\|_{\mathrm{F}}$, implies that the entries of $\hat{\mathbf{H}}$ are considered to be i.i.d. ZM CSCG RVs with variance $\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2} /\left(L_{\mathrm{t}} L_{\mathrm{r}}\right)$. From (6.46), it follows that the variances of the real and imaginary parts of $n_{i}$ reduce to

$$
\begin{align*}
& \sigma_{i, \mathrm{R}}^{2}(\hat{\mathbf{H}}, \mathbf{s}) \approx \frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}}\left(1+\frac{\left\|\mathbf{C}_{i, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{\lambda L_{\mathrm{t}}\left(\eta K_{\mathrm{p}}+\frac{N_{0}}{E_{\mathrm{s}}}\right)}\right),  \tag{6.47a}\\
& \sigma_{i, \mathrm{I}}^{2}(\hat{\mathbf{H}}, \mathbf{s}) \approx \frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}}\left(1+\frac{\left\|\mathbf{C}_{i, \mathrm{I}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{\lambda L_{\mathrm{t}}\left(\eta K_{\mathrm{p}}+\frac{N_{0}}{E_{\mathrm{s}}}\right)}\right) . \tag{6.47b}
\end{align*}
$$

Using the approximate variances in (6.47), the BER for non-square OSTBCs is given by (6.23), where $P_{\mathrm{b}, i, \mathrm{q}}$, with $\mathrm{q}=\mathrm{R}$ or $\mathrm{q}=\mathrm{I}$, is given by

$$
\begin{align*}
& P_{\mathrm{b}, i, \mathrm{q}} \approx \frac{1}{M^{N_{\mathrm{s}}}} \frac{1}{\log _{2}(\sqrt{M})} \sum_{\mathrm{s} \in \Psi^{N_{\mathrm{s}}}} \sum_{b_{\mathrm{q}} \in \Psi^{\prime}} d_{\mathrm{H}}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right) \\
& \times\left[\Omega_{L}\left(\frac{2 \lambda D_{1}^{2}\left(s_{i, \mathrm{q}^{\prime}}, b_{\mathrm{q}}\right)}{1+\frac{\left\|\mathbf{C}_{i, \mathrm{q}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{\lambda L_{\mathrm{q}} \eta K_{\mathrm{p}}}+\frac{1}{\eta K_{\mathrm{p}}} \frac{N_{0}}{E_{\mathrm{s}}}} \frac{E_{\mathrm{s}}}{N_{0}}\right)-\Omega_{L}\left(\frac{2 \lambda D_{2}^{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)}{1+\frac{\left\|\mathbf{C}_{i, \mathrm{q}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{\lambda L_{\mathrm{t}} \eta K_{\mathrm{p}}}+\frac{1}{\eta K_{\mathrm{p}}} \frac{N_{0}}{E_{\mathrm{s}}}} \frac{E_{\mathrm{s}}}{N_{0}}\right)\right], \tag{6.48}
\end{align*}
$$

with $\Omega_{L}(\cdot)$ being given by (6.32). For large SNR, (6.48) reduces to

$$
\begin{align*}
P_{\mathrm{b}, i, \mathrm{q}}^{(\mathrm{as})} \approx \frac{1}{M^{N_{\mathrm{s}}}} \frac{(4 \lambda)^{-L}}{\log _{2}(\sqrt{M})}\binom{2 L-1}{L} & \left(\frac{E_{\mathrm{s}}}{N_{0}}\right)^{-L} \\
& \times \sum_{\mathbf{s} \in \Psi^{N_{\mathrm{s}}}}\left[\xi_{L}\left(s_{i, \mathrm{q}}\right)\left(1+\frac{\left\|\mathbf{C}_{i, \mathrm{q}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{\lambda L_{\mathrm{t}} \eta K_{\mathrm{p}}}\right)^{L}\right] \tag{6.49}
\end{align*}
$$

where $\xi_{L}(\cdot)$ is given by (6.42).
In appendix 6.A.2, we derive $\left\|\mathrm{C}_{i, \mathrm{q}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}$ for various non-square OSTBCs. Due to the symmetry observed in these expressions for different $i$ and q , the BERs for the in-phase and quadrature bits are readily verified to be identical and independent of the index $i$. Moreover, similar to the case of square OSTBCs, the summation in (6.48) over s can be reduced to a summation over the symbols with positive real and imaginary parts only. Hence, for the OSTBCs considered in appendix 6.A.2, the complexity of the BER calculation can be reduced by a factor of $2 N_{\mathrm{s}} 4^{N_{\mathrm{s}}}$, as the BER reduces to

$$
\begin{align*}
& P_{\mathrm{b}} \approx\left(\frac{4}{M}\right)^{N_{\mathrm{s}}} \frac{1}{\log _{2}(\sqrt{M})} \sum_{\mathrm{s} \in \Psi_{0}^{N_{\mathrm{s}}}} \sum_{b_{\mathrm{q}} \in \Psi^{\prime}} d_{\mathrm{H}}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right) \\
& \times\left[\Omega_{L}\left(\frac{2 \lambda D_{1}^{2}\left(s_{i, \mathrm{q}^{\prime}}, b_{\mathrm{q}}\right)}{1+\frac{\left\|\mathbf{C}_{i, \mathrm{q}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{\lambda L_{\mathrm{t}} \eta K_{\mathrm{p}}}+\frac{1}{\eta K_{\mathrm{p}}} \frac{N_{0}}{E_{\mathrm{s}}}} \frac{E_{\mathrm{s}}}{N_{0}}\right)-\Omega_{L}\left(\frac{2 \lambda D_{2}^{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)}{1+\frac{\left\|\mathbf{C}_{i, \mathrm{q}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{\lambda L_{\mathrm{t}} \eta K_{\mathrm{p}}}+\frac{1}{\eta K_{\mathrm{p}}} \frac{N_{0}}{E_{\mathrm{s}}}} \frac{E_{\mathrm{s}}}{N_{0}}\right)\right], \tag{6.50}
\end{align*}
$$

which for high SNR reduces to

$$
\begin{align*}
& P_{\mathrm{b}}^{(\mathrm{as})} \approx\left(\frac{4}{M}\right)^{N_{\mathrm{s}}} \frac{(4 \lambda)^{-L}}{\log _{2}(\sqrt{M})}\binom{2 L-1}{L}\left(\frac{E_{\mathrm{s}}}{N_{0}}\right)^{-L} \\
& \times \sum_{\mathbf{s} \in \Psi_{0}^{N_{\mathrm{s}}}}\left[\xi_{L}\left(s_{i, \mathrm{q}}\right)\left(1+\frac{\left\|\mathbf{C}_{i, \mathrm{q}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{\lambda L_{\mathrm{t}} \eta K_{\mathrm{p}}}\right)^{L}\right] \tag{6.51}
\end{align*}
$$

where the function $\xi_{L}(\cdot)$ is defined in (6.42). From (6.51), the ratio of the BER of the mismatched receiver to the BER of the PCSI receiver is easily derived (at high $E_{\mathrm{b}} / N_{0}$ ):

$$
\begin{align*}
\frac{P_{\mathrm{b}, \mathrm{LMMSE}}^{(\mathrm{as})}}{P_{\mathrm{b}, \mathrm{PCSI}}^{(\mathrm{as})}}=\left(1+\frac{\eta K_{\mathrm{p}}}{K}\right)^{L}\left(\frac{4}{M}\right)^{N_{\mathrm{s}}-\frac{1}{2}} & \\
& \times \frac{\sum_{\mathbf{s} \in \Psi_{0}^{N_{\mathrm{s}}}}\left[\xi_{L}\left(s_{i, \mathrm{q}}\right)\left(1+\frac{\left\|\mathbf{c}_{i, \mathrm{q}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{\lambda L_{\mathrm{t}} \eta K_{\mathrm{p}}}\right)^{L}\right]}{\sum_{\mathrm{s}_{\mathrm{q}} \in \Psi_{0}^{\prime}}\left[\xi_{L}\left(s_{\mathrm{q}}\right)\right]} \tag{6.52}
\end{align*}
$$

The amount (in dB ) by which the $E_{\mathrm{b}} / N_{0}$ ratio of the mismatched receiver should be increased to obtain the same BER as the PCSI receiver is obtained by substituting (6.52) in (6.44).

### 6.2.3 Numerical Results

In this section, we illustrate the closed-form BER expressions derived for OSTBCs with LMMSE channel estimation operating over i.i.d. Rayleigh fading


Figure 6.1: BER of Alamouti's code with 4-QAM.
channels, under the assumption that $E_{\mathrm{p}}=E_{\mathrm{s}}$.
Fig. 6.1 shows the BER curves resulting from the exact closed-form expression (6.39), for a system employing Alamouti's code (4.14) along with 4-QAM transmission. Both results for the PCSI and the mismatched receiver are shown. Using data frames consisting of $K=100$ time slots for data transmission, it follows from (6.14) that the optimal number of pilot symbols is given by $K_{p}=14$, which results in a BER degradation of 1.15 dB , irrespective of $L_{r}$. Also shown in the figure are computer simulation results for the mismatched receiver that confirm the analytical result.

Fig. 6.2 displays the approximate BER curves resulting from (6.50) for the $3 \times 4$ non-square OSTBC given by (4.34). Using a 16-QAM symbol constellation, results are shown for the PCSI receiver (exact result) and the mismatched receiver (analytical approximation and simulation result). The simulations indicate that the closed-form BER expression for the mismatched receiver is very accurate. From (6.44) and (6.52), it follows that the BER degradation amounts to $1.08 \mathrm{~dB}, 1.10 \mathrm{~dB}$ or 1.12 dB when $L_{r}$ equals 1,2 or 3, respectively, whereas the rule of thumb yields a degradation of 1.08 dB .


Figure 6.2: BER of OSTBC given by (4.34) with 16-QAM.

### 6.3 BER Analysis for Correlated Rayleigh Fading

In this section, we extend the results from section 6.2 to arbitrarily correlated Rayleigh fading channels. Using a high-SNR approximation of the channel error covariance matrix, we derive accurate closed-form BER approximations for a mismatched ML receiver that obtains the channel state information through pilot-based LMMSE channel estimation. Moreover, we show that the presented expression yields very accurate BER results for both LMMSE and ML channel estimation, over a wide range of SNRs. As opposed to the BER analysis in section 6.2, we make use of the vector signal model (5.4) in order to derive the BER expressions for correlated fading. In this way, an equivalent vector model can be constructed for ML detection of the matrix B and the resulting decision variables corresponding to the transmitted information symbols. From (5.4), it follows that mismatched ML detection of the matrix $\mathbf{B} \triangleq \mathbf{C}^{T} \otimes \mathbf{I}_{L_{\mathrm{r}}}$ is given by

$$
\begin{equation*}
\hat{\mathbf{B}}=\arg \min _{\tilde{\mathbf{B}}}\left\|\mathbf{r}-\sqrt{E_{\mathrm{S}}} \tilde{\mathbf{B}} \hat{\mathbf{h}}\right\|_{\mathrm{F}}^{2} \tag{6.53}
\end{equation*}
$$

which is shown to reduce to symbol-by-symbol detection

$$
\begin{equation*}
\hat{s}_{i}=\arg \min _{\tilde{s} \in \Psi}\left|u_{i}-\tilde{s}\right|, \quad i=1, \ldots, N_{\mathrm{s}} \tag{6.54}
\end{equation*}
$$

where the vector counterpart of the decision variable (6.18) is given by

$$
\begin{equation*}
u_{i}=\frac{\hat{\mathbf{h}}^{H}\left(\mathbf{C}_{i}^{*} \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \mathbf{r}+\mathbf{r}^{H}\left(\mathbf{C}_{i}^{\prime T} \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \hat{\mathbf{h}}}{\lambda \sqrt{E_{\mathrm{s}}}\|\hat{\mathbf{h}}\|^{2}} \tag{6.55}
\end{equation*}
$$

Using the channel decomposition

$$
\begin{equation*}
\mathbf{h}=\hat{\mathbf{h}}+\varepsilon \tag{6.56}
\end{equation*}
$$

where $\hat{\mathbf{h}}$ denotes the LMMSE channel estimate (5.16), the received signal (5.4) can be written as

$$
\begin{equation*}
\mathbf{r}=\sqrt{E_{\mathrm{s}}} \mathbf{B} \hat{\mathbf{h}}+\sqrt{E_{\mathrm{s}}} \mathbf{B} \boldsymbol{\varepsilon}+\mathbf{w} \tag{6.57}
\end{equation*}
$$

where $\sqrt{E_{\mathrm{S}}} \mathbf{B} \hat{\mathbf{h}}$ is the useful component, $\mathbf{w}$ is the Gaussian channel noise, and $\sqrt{E_{\mathrm{S}}} \mathbf{B} \boldsymbol{\varepsilon}$ is additional noise caused by the channel estimation error; note that $\hat{\mathbf{h}}$ and $\varepsilon$ are independent. Since the channel is assumed to consist of correlated ZM CSCG RVs with covariance matrix $\mathcal{R}$, the elements of the error vector $\boldsymbol{\varepsilon}$ are ZM CSCG RVs with covariance matrix (5.20). Hence, the additional noise vector $\sqrt{E_{\mathrm{s}}} \mathbf{B} \boldsymbol{\varepsilon}$ is Gaussian when conditioned on the data symbol vector $\mathbf{s}$. Using (6.57), the decision variable (6.55) reduces to

$$
\begin{equation*}
u_{i}=s_{i}+n_{i}, \quad 1 \leq i \leq N_{\mathrm{s}} \tag{6.58}
\end{equation*}
$$

where the disturbance term $n_{i}$ contains contributions from the channel noise $\mathbf{w}$ and the channel estimation error $\boldsymbol{\varepsilon}$. It is readily verified that $n_{i}=e_{i}+w_{i}$, with

$$
\begin{align*}
& e_{i}= \hat{\mathbf{h}}^{H}\left(\mathbf{C}_{i}^{*} \mathbf{C}^{T} \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \varepsilon+\varepsilon^{H}\left(\mathbf{C}^{*} \mathbf{C}_{i}^{\prime T} \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \hat{\mathbf{h}}  \tag{6.59}\\
& \lambda\|\hat{\mathbf{h}}\|^{2}  \tag{6.60}\\
& w_{i}=\frac{\hat{\mathbf{h}}^{H}\left(\mathbf{C}_{i}^{*} \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \mathbf{w}+\mathbf{w}^{H}\left(\mathbf{C}_{i}^{\prime T} \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \hat{\mathbf{h}}}{\lambda \sqrt{E_{\mathrm{s}}}\|\hat{\mathbf{h}}\|^{2}}
\end{align*}
$$

Note that $n_{i}$ is Gaussian when conditioned on $\mathbf{s}$ and $\hat{\mathbf{h}}$. In case of PCSI, the channel estimation error $\varepsilon=\mathbf{0}$ and $n_{i}$ is a ZM CSCG RV independent of $\mathbf{s}$ with variance $N_{0} /\left(\lambda E_{\mathrm{S}}\|\mathbf{h}\|^{2}\right)$. Using (6.24), we show in appendix 6.A.3 that, when conditioned on $\hat{\mathbf{h}}$ and $\mathbf{s}, n_{i}$ is a ZM non-circularly symmetric complex Gaussian RV, the variances of the real and imaginary parts of which are given by

$$
\begin{align*}
\sigma_{i, \mathrm{R}}^{2}(\hat{\mathbf{h}}, \mathbf{s}) & =\frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{h}}\|^{2}} \\
& \times\left(1+\frac{E_{\mathrm{s}}}{\lambda N_{0}\|\hat{\mathbf{h}}\|^{2}} \hat{\mathbf{h}}^{H}\left(\mathbf{C}_{i, \mathrm{R}}^{T}(\mathbf{s}) \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \mathcal{R}_{\varepsilon}\left(\mathbf{C}_{i, \mathrm{R}}^{*}(\mathbf{s}) \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \hat{\mathbf{h}}\right), \tag{6.61a}
\end{align*}
$$

$$
\begin{align*}
\sigma_{i, \mathrm{I}}^{2}(\hat{\mathbf{h}}, \mathbf{s}) & =\frac{N_{0}}{2 \lambda E_{S}\|\hat{\mathbf{h}}\|^{2}} \\
& \times\left(1+\frac{E_{\mathrm{S}}}{\lambda N_{0}\|\hat{\mathbf{h}}\|^{2}} \hat{\mathbf{h}}^{H}\left(\mathbf{C}_{i, \mathrm{I}}^{T}(\mathbf{s}) \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \mathcal{R}_{\varepsilon}\left(\mathbf{C}_{i, \mathrm{I}}^{*}(\mathbf{s}) \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \hat{\mathbf{h}}\right), \tag{6.61b}
\end{align*}
$$

where the covariance matrix $\boldsymbol{\mathcal { R }}_{\varepsilon}$ of the channel estimation error $\varepsilon$ is given by (5.20).

Similar to the case of i.i.d. Rayleigh fading, the BER can be written as

$$
\begin{equation*}
P_{\mathrm{b}}=\frac{1}{2 N_{\mathrm{s}}} \sum_{i=1}^{N_{\mathrm{s}}}\left[P_{\mathrm{b}, i, \mathrm{R}}+P_{\mathrm{b}, i, \mathrm{I}}\right], \tag{6.62}
\end{equation*}
$$

where $P_{\mathrm{b}, i, \mathrm{R}}$ and $P_{\mathrm{b}, i, \mathrm{I}}$ denote the BERs of the in-phase and quadrature bits corresponding to the information symbols $s_{i}$, respectively. Since (6.61a) and (6.61b) are obtained for a given $\hat{\mathbf{h}}$ and $\mathbf{s}$, we can easily calculate the conditional BERs of the in-phase and quadrature bits corresponding to the information symbols $s_{i}$, conditioned on $\hat{h}$ and $\mathbf{s}$.

$$
\begin{equation*}
P_{\mathrm{b}, i, \mathrm{q}}(\hat{\mathbf{h}}, \mathbf{s})=\frac{1}{\log _{2}(\sqrt{M})} \sum_{b_{\mathrm{q}} \in \Psi^{\prime}} d_{\mathrm{H}}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right) \operatorname{Pr}\left[\hat{s}_{i, \mathrm{q}}=b_{\mathrm{q}} \mid \hat{\mathbf{h}}, \mathbf{s}\right], \tag{6.63}
\end{equation*}
$$

where $d_{\mathrm{H}}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)$ denotes the Hamming distance between the bits associated to $s_{i, \mathrm{q}}$ and $b_{\mathrm{q}}$, respectively, and $\operatorname{Pr}\left[\hat{s}_{i, \mathrm{q}}=b_{\mathrm{q}} \mid \hat{\mathbf{h}}, \mathbf{s}\right]$ is given by

$$
\begin{align*}
\operatorname{Pr}\left[\hat{s}_{i, \mathrm{q}}=b_{\mathrm{q}} \mid \hat{\mathbf{h}}, \mathbf{s}\right] & =\operatorname{Pr}\left[d_{1}\left(b_{\mathrm{q}}\right) \leq u_{i, \mathrm{q}} \leq d_{2}\left(b_{\mathrm{q}}\right) \mid \hat{\mathbf{h}}, \mathbf{s}\right] \\
& =Q\left(\sqrt{\frac{D_{1}^{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)}{\sigma_{i, \mathrm{q}}^{2}(\hat{\mathbf{h}}, \mathbf{s})}}\right)-Q\left(\sqrt{\frac{D_{2}^{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)}{\sigma_{i, \mathrm{q}}^{2}(\hat{\mathbf{h}}, \mathbf{s})}}\right) \tag{6.64}
\end{align*}
$$

with $D_{1}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)$ and $D_{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)$ being given by (3.60). Finally, $P_{\mathrm{b}, i, \mathrm{R}}$ and $P_{\mathrm{b}, i, \mathrm{I}}$ in (6.62) are obtained by averaging the corresponding conditional BERs (6.63) over $\hat{\mathbf{h}}$ and $\mathbf{s}$

$$
\begin{equation*}
P_{\mathrm{b}, i, \mathrm{q}}=\frac{1}{M^{N_{\mathbf{s}}}} \sum_{\mathbf{s} \in \Psi^{N_{\mathbf{s}}}} \mathbb{E}_{\hat{\mathbf{h}}}\left[P_{\mathrm{b}, i, \mathrm{q}}(\hat{\mathbf{h}}, \mathbf{s})\right] . \tag{6.65}
\end{equation*}
$$

Note that (6.61a) and (6.61b) are complicated functions of the channel estimate $\hat{\mathbf{h}}$, such that the resulting exact conditional BER expressions $P_{\mathrm{b}, i, \mathrm{q}}(\hat{\mathbf{h}}, \mathbf{s})$ can be evaluated by numerical integration over $\hat{\mathbf{h}}$ only. However, if we can approximate (6.61a) and (6.61b) so that they are a function of $\hat{\mathbf{h}}$ through an inverse proportionality to $\|\hat{\mathbf{h}}\|^{2}$ only, a closed-form BER expression can be easily found, since it follows from (4.69) that the PDF of $\|\hat{\mathbf{h}}\|^{2}$ is a weighted sum of $\chi^{2}$-distributions, with $\lambda_{i}, i=1,2, \ldots, \kappa$, being the $i$-th distinct eigenvalue of (5.18), with corresponding algebraic multiplicity $c_{i}$. It is easily seen that if we replace $\mathcal{R}_{\varepsilon}$ in (6.61) by its high-SNR approximation (5.22), the variances
(6.61a) and (6.61b) reduce to

$$
\begin{align*}
& \sigma_{i, \mathrm{R}}^{2}(\hat{\mathbf{h}}, \mathbf{s}) \approx \frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{h}}\|^{2}}\left(1+\frac{\hat{\mathbf{h}}^{H}\left(\mathbf{C}_{i, \mathrm{R}}^{T}(\mathbf{s}) \mathbf{C}_{i, \mathrm{R}}^{*}(\mathbf{s}) \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \hat{\mathbf{h}}}{\lambda \eta K_{\mathbf{p}}\|\hat{\mathbf{h}}\|^{2}}\right),  \tag{6.66a}\\
& \sigma_{i, \mathrm{I}}^{2}(\hat{\mathbf{h}}, \mathbf{s}) \approx \frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{h}}\|^{2}}\left(1+\frac{\hat{\mathbf{h}}^{H}\left(\mathbf{C}_{i, \mathrm{I}}^{T}(\mathbf{s}) \mathbf{C}_{i, \mathrm{I}}^{*}(\mathbf{s}) \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \hat{\mathbf{h}}}{\lambda \eta K_{\mathrm{p}}\|\hat{\mathbf{h}}\|^{2}}\right), \tag{6.66b}
\end{align*}
$$

which is the vector equivalent to the variances (6.25a) and (6.25b) for high SNR. Hence, closed-form BER expressions for correlated Rayleigh fading can be obtained in a similar way as in section 6.2. For correlated Rayleigh fading, however, (6.66a) and (6.66b) are not exact because of the high-SNR approximation of $\mathcal{R}_{\varepsilon}$, whereas there is no approximation involved in the derivation of (6.25a) and (6.25b) for i.i.d. Rayleigh fading.

### 6.3.1 Approximate BER for Square OSTBCs

Taking (6.33) into account, it follows that for square OSTBCs the variances (6.66a) and (6.66b) of the real and imaginary part of $n_{i}$ reduce to

$$
\begin{equation*}
\sigma_{i, \mathrm{R}}^{2}(\hat{\mathbf{h}}, \mathbf{s})=\sigma_{i, \mathrm{I}}^{2}(\hat{\mathbf{h}}, \mathbf{s}) \approx \frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{h}}\|_{\mathrm{F}}^{2}}\left(1+\frac{\lambda\|\mathbf{s}\|^{2}}{\eta K_{\mathrm{p}}}\right) \tag{6.67}
\end{equation*}
$$

such that $n_{i}$ is a ZM CSCG RV, the variance of which does not depend on the index $i$. Hence, the BERs of the in-phase and quadrature bits allocated to the information symbols $s_{i}$ are identical and (6.23) reduces to

$$
\begin{equation*}
P_{\mathrm{b}}=P_{\mathrm{b}, i, \mathrm{q}} \tag{6.68}
\end{equation*}
$$

where $P_{\mathrm{b}, i, \mathrm{q}}$ is obtained from (6.63)-(6.65), with $\sigma_{i, \mathrm{q}}^{2}(\hat{\mathbf{h}}, \mathbf{s})$ in (6.64) being given by (6.67). In this way, it follows from (6.31) and (4.69) that averaging the $Q$-functions in (6.64) over $\hat{\mathbf{h}}$ yields the following closed-form expression

$$
\begin{equation*}
\mathbb{E}_{\hat{\mathbf{h}}}\left[Q\left(\sqrt{\frac{D_{j}^{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)}{\sigma_{i, \mathrm{q}}^{2}(\hat{\mathbf{h}}, \mathbf{s})}}\right)\right]=\sum_{m=1}^{\kappa} \sum_{n=1}^{c_{m}} D_{m, n} \Omega_{n}\left(\frac{2 \lambda \lambda_{m} D_{j}^{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)}{1+\frac{\lambda\|\mathbf{s}\|^{2}}{\eta K_{\mathrm{p}}}} \frac{E_{\mathrm{S}}}{N_{0}}\right) \tag{6.69}
\end{equation*}
$$

where $j \in\{1,2\}, \Omega_{n}(\cdot)$ is defined in (6.32), $\lambda_{m}$ is the $m$-th distinct eigenvalue of (5.18) with algebraic multiplicity $c_{m}$, and the parameters $D_{m, n}$ are given by (4.70). Hence, taking (6.39) into account, the BER (6.68) for square OSTBCs on correlated Rayleigh fading channels with LMMSE channel estimation can be
written in closed-form as

$$
\begin{align*}
P_{\mathrm{b}} & \approx\left(\frac{4}{M}\right)^{N_{\mathrm{s}}} \frac{1}{\log _{2}(\sqrt{M})} \sum_{m=1}^{\kappa} \sum_{n=1}^{c_{m}} D_{m, n} \sum_{\mathrm{s} \in \Psi_{0}^{N_{s}}} \sum_{\mathrm{b}_{\mathrm{q}} \in \Psi^{\prime}} d_{\mathrm{H}}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right) \\
& \times\left[\Omega_{n}\left(\frac{2 \lambda \lambda_{m} D_{1}^{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)}{1+\frac{\left.\lambda\| \|^{2}\right|^{2}}{\eta K_{\mathrm{p}}}} \frac{E_{\mathrm{s}}}{N_{0}}\right)-\Omega_{n}\left(\frac{2 \lambda \lambda_{m} D_{2}^{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)}{1+\frac{\lambda \|\left.\boldsymbol{s}\right|^{2}}{\eta K_{\mathrm{p}}}} \frac{E_{\mathrm{s}}}{N_{0}}\right)\right] \tag{6.70}
\end{align*}
$$

where $\Psi_{0}$ is defined in (6.38). Note that $\|\mathbf{s}\|^{2}=N_{\mathrm{s}}$ in case of 4-QAM, such that the summation over all possible symbol vectors $s$ reduces to a summation over $s_{i, q}$ only.

### 6.3.2 Approximate BER for Non-Square OSTBCs

For non-square OSTBCs, (6.29) is not satisfied and, according to the analysis provided in section 6.2.1, a closed-form BER expression can be obtained by approximating (6.66a) and (6.66b) by their expectations over $\hat{\mathbf{h}}$, conditioned on the norm $\|\hat{\mathbf{h}}\|$

$$
\begin{align*}
\sigma_{i, \mathrm{R}}^{2}(\hat{\mathbf{h}}, \mathbf{s}) & \approx \frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{h}}\|_{\mathrm{F}}^{2}}\left(1+\frac{\left\|\mathbf{C}_{i, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{\lambda L_{\mathrm{t}} \eta K_{\mathrm{p}}}\right),  \tag{6.71a}\\
\sigma_{i, \mathrm{I}}^{2}(\hat{\mathbf{h}}, \mathbf{s}) & \approx \frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{h}}\|_{\mathrm{F}}^{2}}\left(1+\frac{\left\|\mathbf{C}_{i, \mathrm{I}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{\lambda L_{\mathrm{t}} \eta K_{\mathrm{p}}}\right) . \tag{6.71b}
\end{align*}
$$

For the non-square OSTBCs assessed in appendix 6.A.2, the BERs for the inphase and quadrature bits are shown to be identical and independent of the index $i$. Moreover, the summation in (6.48) over $s$ can be reduced to a summation over the symbols with positive real and imaginary parts only. Hence, using (6.71), the BER approximation for the non-square OSTBCs in appendix 6.A. 2 can be written as

$$
\begin{align*}
P_{\mathrm{b}} & \approx\left(\frac{4}{M}\right)^{N_{\mathrm{s}}} \frac{1}{\log _{2}(\sqrt{M})} \sum_{m=1}^{\kappa} \sum_{n=1}^{c_{m}} D_{m, n} \sum_{\mathbf{s} \in \Psi_{0}^{N_{\mathrm{s}}}} \sum_{b_{\mathrm{q}} \in \Psi^{\prime}} d_{\mathrm{H}}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right) \\
& \times\left[\Omega_{n}\left(\frac{2 \lambda \lambda_{m} D_{1}^{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)}{1+\frac{\left\|\mathbf{C}_{i, \mathrm{q}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{\lambda L_{\mathrm{t}} \eta K_{\mathrm{p}}}} \frac{E_{\mathrm{s}}}{N_{0}}\right)-\Omega_{n}\left(\frac{2 \lambda \lambda_{m} D_{2}^{2}\left(s_{i, \mathrm{q},}, b_{\mathrm{q}}\right)}{1+\frac{\left\|\mathbf{C}_{i, \mathrm{q}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{\lambda L_{\mathrm{t}} \eta K_{\mathrm{p}}}} \frac{E_{\mathrm{s}}}{N_{0}}\right)\right] . \tag{6.72}
\end{align*}
$$

### 6.3.3 Numerical Results

In this section, we present numerical results from evaluating the presented closed-form BER expressions under the assumption that $E_{\mathrm{p}}=E_{\mathrm{s}}$. Additionally, Monte-Carlo simulations indicate that the BER expressions yield very


Figure 6.3: BER for Alamouti's code under i.i.d. $(\rho=0)$ and correlated ( $\rho \in$ $\{0.99,0.999,1\}$ ) Rayleigh fading, with and without PCSI.
accurate BER results for square and non-square OSTBCs, for low to moderate SNR, and for both LMMSE and ML channel estimation.

In Fig. 6.3, the impact of highly correlated channels on the BER performance is shown for Alamouti's code (4.14) using 4-QAM over a $2 \times 1 \mathrm{MIMO}$ channel. The covariance matrix $\mathcal{R}$ of the channel is given by

$$
\mathcal{R}=\left[\begin{array}{ll}
1 & \rho  \tag{6.73}\\
\rho & 1
\end{array}\right]
$$

where $\rho$ is assumed to be a positive real-valued correlation coefficient. The BER curves are shown for $\rho \in\{0,0.99,0.999,1\}$, and for PCSI and ICSI. In case of ICSI, it is assumed that $K=100$ and $K_{\mathrm{p}}=14$, and that LMMSE channel estimation is used. For BPSK and 4-QAM constellations under PCSI, it is shown in appendix 6 .A. 4 that as long as the smallest eigenvalue $\lambda_{\text {min }}$ of $\mathcal{R}$ is strictly positive, a full diversity gain of $L$ is achieved in the high-SNR region, i.e., where $E_{\mathrm{b}} / N_{0} \gg 1 /\left(\lambda \lambda_{\min }\right)$. Hence, taking into account that $\lambda=1$ for Alamouti's code and that the eigenvalues of (6.73) are given by $\lambda_{1}=1+\rho$ and $\lambda_{2}=1-\rho$, it is expected that a full diversity gain of 2 will be achieved for high SNR, i.e., $E_{\mathrm{b}} / N_{0} \gg 1 /(1-\rho)$, as long as the correlation coefficient $\rho<1$.


Figure 6.4: Piecewise linear approximation of the BER for Alamouti's code under i.i.d. $(\rho=0)$ and correlated $(\rho \in\{0.99,0.999,1\})$ Rayleigh fading, with PCSI.

Indeed, it is observed from Fig. 6.3 that for highly correlated channels, i.e., $\rho \rightarrow 1$, the BER curves are shown to coincide with the BER for the fully correlated MIMO channel, i.e., $\rho=1$, in case of low to moderate SNR, whereas they achieve full spatial diversity in case of high SNR. In case of full correlation, however, the achieved diversity gain decreases from 2 to 1 , since both MIMO channel coefficients are completely identical and spatial diversity is lost. It is also appreciated from the figure that ICSI results in a shift of the BER curves over an amount which is essentially independent of the correlation coefficient and can be estimated from the rule of thumb (6.13). Hence, although it is shown in Fig. 5.2 that, for low to moderate SNR, the MSE on the LMMSE channel estimate decreases if the correlation coefficient $\rho$ increases, this effect has no significant impact on the BER degradation due to ICSI. Furthermore, it follows from (6.106) in appendix 6.A.4 that, on a logarithmic scale, the BER curves for PCSI in Fig. 6.3 can be roughly approximated by piecewise linear functions, as shown in Fig. 6.4. For low SNR, i.e., $E_{\mathrm{b}} / N_{0}<1 /(1+\rho)$, the BER can be considered to be $1 / 2$ and communication fails completely. For moderate SNR, i.e., $1 /(1+\rho)<E_{\mathrm{b}} / N_{0}<1 /(1-\rho)$, the BER is proportional to
$\left(E_{\mathrm{b}} / N_{0}\right)^{-1}$ such that an effective diversity gain of 1 is achieved. Finally, in the high-SNR region, i.e., $E_{\mathrm{b}} / N_{0}>1 /(1-\rho)$, full diversity is achieved. Similar to the BER degradation (6.1) due to ICSI, we define the BER degradation due to spatial correlation as

$$
\begin{equation*}
\Delta_{\mathrm{BER}, \mathrm{corr}}=\frac{\left(E_{\mathrm{b}} / N_{0}\right)_{\mathrm{corr}}}{\left(E_{\mathrm{b}} / N_{0}\right)_{\mathrm{i} . \mathrm{i.d.}}} \tag{6.74}
\end{equation*}
$$

where it is assumed that a target BER is achieved at $E_{\mathrm{b}} / N_{0}=\left(E_{\mathrm{b}} / N_{0}\right)_{\text {i.i.d. }}$ and $E_{\mathrm{b}} / N_{0}=\left(E_{\mathrm{b}} / N_{0}\right)_{\text {corr }}$ in case of i.i.d. and spatially correlated MIMO channels, respectively. For BPSK and 4-QAM constellations under PCSI, we show in appendix 6.A. 4 that in the high-SNR region, where the BER curves are parallel, $\Delta_{\text {BER,corr }}$ is approximately given by

$$
\begin{equation*}
\Delta_{\mathrm{BER}, \mathrm{corr}} \approx\left(\prod_{i=1}^{L} \lambda_{i}\right)^{-\frac{1}{L}} \tag{6.75}
\end{equation*}
$$

where $\lambda_{i}$ are the eigenvalues of the covariance matrix $\mathcal{R}$ of the channel. Note that the BER degradation due to spatial correlation is the inverse of the geometric mean of the eigenvalues of $\mathcal{R}$ and that (6.75) is not valid for values of $E_{\mathrm{b}} / N_{0}$ smaller than $1 /\left(\lambda \lambda_{\text {min }}\right)$, which is particularly important when $\lambda_{\text {min }}$ becomes very small, i.e., in case of highly correlated channels.

In the next example, we illustrate the tremendous computation time savings that can be achieved by using the presented closed-form BER expressions instead of direct Monte-Carlo simulations. We consider Alamouti's code (4.14) employing $M$-QAM transmission, with $M \in\{4,16,64,256\}$, on a $2 \times 1$ MIMO channel with covariance matrix

$$
\mathcal{R}=\left[\begin{array}{cc}
1 & 0.2  \tag{6.76}\\
0.2 & 1
\end{array}\right]
$$

Furthermore, we assume that $K=100$ and $K_{p}=14$, and that the channel is recovered through LMMSE channel estimation. In order to guarantee a certain accuracy for the simulated BER results, we require that the ratio of the variance of the simulated BER to the square of its expectation is less or equal than a prescribed value $\epsilon^{2}$

$$
\begin{equation*}
\frac{\operatorname{var}\left[\hat{P}_{\mathrm{b}}\right]}{\left(\mathbb{E}\left[\hat{P}_{\mathrm{b}}\right]\right)^{2}} \leq \epsilon^{2} \tag{6.77}
\end{equation*}
$$

where $\hat{P}_{\mathrm{b}}$ denotes the simulated BER, which is given by

$$
\begin{equation*}
\hat{P}_{\mathrm{b}}=\frac{1}{N} \sum_{l=1}^{N} X_{l} \tag{6.78}
\end{equation*}
$$

Note that $N$ and $X_{l}$ in (6.78) denote the number of simulated data frames and the ratio of the number of bit errors counted in the $l$-th frame to the total


Figure 6.5: BER for Alamouti's code employing $M$-QAM transmission, with $M \in\{4,16,64,256\}$, on a correlated $2 \times 1$ MIMO channel with ICSI.
number of bits within one frame, respectively. Since all RVs in the successive data frames are independently generated, $X_{l}$ 's are independent and $\operatorname{var}\left[\hat{P}_{\mathrm{b}}\right]=$ $\frac{1}{N} \operatorname{var}\left[X_{l}\right]$, where $\operatorname{var}\left[X_{l}\right]$ can be approximated by

$$
\begin{equation*}
\operatorname{var}\left[X_{l}\right] \approx \frac{1}{N} \sum_{l=1}^{N} X_{l}^{2}-\left(\frac{1}{N} \sum_{l=1}^{N} X_{l}\right)^{2} \tag{6.79}
\end{equation*}
$$

Taking (6.79) into account and replacing $\mathbb{E}\left[\hat{P}_{\mathrm{b}}\right]$ in (6.77) by $\hat{P}_{\mathrm{b}}$, it follows that for a given accuracy $\epsilon^{2}, N$ needs to satisfy

$$
\begin{equation*}
N \geq \frac{\operatorname{var}\left[X_{l}\right]}{\epsilon^{2} \hat{P}_{\mathrm{b}}^{2}} \tag{6.80}
\end{equation*}
$$

Using the minimal number of frames $N$ required to assure the accuracy associated with $\epsilon^{2}=0.0001$, the BER obtained from Monte-Carlo simulations is shown in Fig. 6.5, along with the curves resulting from the closed-form BER expressions presented in section 6.3.1. It is observed that the simulated BER results are in perfect agreement with the analytically obtained BER curves. The computation times corresponding to the different BER results in


Figure 6.6: Comparison of computation times corresponding to closed-form BER expressions and direct Monte-Carlo simulations.

Fig. 6.5 are shown in Fig. 6.6. It follows from the figure that the computation time increases exponentially with the SNR in case of Monte-Carlo simulations, whereas the computation time associated with the closed-form BER expressions is essentially independent of the SNR. Also, because the computational complexity related to (6.65) is proportional to $M^{N_{s}}$, the computation time corresponding to the closed-form BER expressions increases significantly for larger constellation size $M$. For the case of Monte-Carlo simulations, on the other hand, the computation time decreases for increasing $M$, because the higher resulting BER requires less frames to be simulated in order to obtain a sufficient number of bit errors to assure a certain accuracy. Hence, the computation time savings achieved by the presented closed-form BER expressions are most significant in case of high SNR and small symbol constellations. However, when larger diversity gains are achieved, the difference in computation time between simulations and closed-form BER expressions will be even more dramatic.


Figure 6.7: BER for the $4 \times 4$ OSTBC given by (6.81), with $M$-QAM transmission, correlated and i.i.d. Rayleigh fading, and LMMSE channel estimation.

Fig. 6.7 displays the BER for the $4 \times 4$ OSTBC given by [70]

$$
\mathbf{C}_{4 \times 4}=\frac{2}{\sqrt{3}}\left[\begin{array}{cccc}
s_{1} & -s_{2}^{*} & -s_{3}^{*} & 0  \tag{6.81}\\
s_{2} & s_{1}^{*} & 0 & -s_{3}^{*} \\
s_{3} & 0 & s_{1}^{*} & s_{2}^{*} \\
0 & s_{3} & -s_{2} & s_{1}
\end{array}\right]
$$

where the scaling factor $2 / \sqrt{3}$ is applied in order that (6.81) satisfies (4.31). The BER curves are shown for 4-QAM and 64-QAM constellations under correlated and uncorrelated Rayleigh fading, for $K=200$ and $K_{p}=16$. Also shown are the BER results for PCSI. We consider a single-antenna receiver and a covariance matrix $\mathcal{R}$, which in case of correlated fading is given by

$$
\mathcal{R}=\left[\begin{array}{cccc}
1 & 0.7 & 0.5 & 0.3  \tag{6.82}\\
0.7 & 1 & 0.7 & 0.5 \\
0.5 & 0.7 & 1 & 0.7 \\
0.3 & 0.5 & 0.7 & 1
\end{array}\right]
$$

Note that $\mathcal{R}$ has a Toeplitz structure, which corresponds to the practical situation of an equally spaced linear antenna array without mutual coupling [44].


Figure 6.8: BER for the $3 \times 4$ OSTBC given by (4.34), with $M$-QAM transmission, correlated Rayleigh fading, and LMMSE and ML channel estimation.

Monte-Carlo simulations for the mismatched receiver with LMMSE channel estimation confirm the accuracy of the presented BER expression (6.70). From the figure, we observe that antenna correlation and ICSI both give rise to a horizontal shift of the BER curve at high SNR, and that the amount of degradation due to ICSI is more or less independent of the antenna correlation and the constellation size.

Fig. 6.8 shows the approximate analytical BER results for the $3 \times 4$ OSTBC given by (4.34). Assuming a dual-antenna receiver $\left(L_{r}=2\right)$, Fig. 6.8 shows the BER curves for square $M$-QAM transmission, with $M \in\{4,16,64\}$, under correlated Rayleigh fading, for $K=200$ and $K_{p}=24$. Also shown are the BER results for PCSI. For correlated fading, the covariance matrix $\mathcal{R}$ is assumed to be given by $\boldsymbol{\mathcal { R }}=\boldsymbol{\mathcal { R }}_{\mathrm{t}} \otimes \boldsymbol{\mathcal { R }}_{\mathrm{r}}$, where $\boldsymbol{\mathcal { R }}_{\mathrm{t}}$ and $\boldsymbol{\mathcal { R }}_{\mathrm{r}}$ are given by

$$
\begin{gather*}
\boldsymbol{\mathcal { R }}_{\mathrm{t}}=\left[\begin{array}{ccc}
1 & 0.5+j 0.2 & 0.2-j 0.1 \\
0.5-j 0.2 & 1 & 0.4-j 0.3 \\
0.2+j 0.1 & 0.4+j 0.3 & 1
\end{array}\right]  \tag{6.83a}\\
\boldsymbol{\mathcal { R }}_{\mathrm{r}}=\left[\begin{array}{ccc}
1 & 0.4-j 0.6 \\
0.4+j 0.6 & 1
\end{array}\right] . \tag{6.83b}
\end{gather*}
$$

Monte-Carlo simulations conducted for a mismatched receiver performing either LMMSE or ML channel estimation, indicate that the presented BER expression (6.72) yields very accurate BER results for LMMSE and ML channel estimation, in the range from low to high SNR.

### 6.4 Chapter Summary

In this chapter, we investigated the impact of ICSI on the BER of OSTBCs in Rayleigh fading channels. To this end, we considered a mismatched receiver using LMMSE channel estimation that applies the channel estimate $\hat{\mathbf{H}}$ in the same way as an ML receiver would apply the channel $\mathbf{H}$.

We derived a simple rule of thumb that serves as an indicator for the BER degradation caused by imperfect channel estimation and yields the exact result for high SNR, square OSTBCs with PSK symbols, and i.i.d. Rayleigh fading. Moreover, for square OSTBCs and i.i.d. Rayleigh fading channels, we derived exact closed-form expressions for the BER and the BER degradation due to imperfect channel estimation. For non-square OSTBCs, a very accurate closed-form approximation was provided. For arbitrarily correlated Rayleigh fading channels, we presented closed-form BER approximations for square and non-square OSTBCs which yield very accurate BER results in the low-to-moderate SNR region for both LMMSE and ML channel estimation. For square OSTBCs, the BER expression is asymptotically exact.

## 6.A Appendix

## 6.A. 1 Derivation of the Variance of $n_{i}$ for i.i.d. Rayleigh Fading

Since E and W are uncorrelated, the terms (6.20) and (6.21) are uncorrelated when conditioned on $\hat{\mathbf{H}}$, such that the variance of the real/imaginary part of $n_{i}$ in (6.19) equals the sum of the variances of the real/imaginary parts of $e_{i}$ and $w_{i}$. According to (6.20), the real part of $e_{i}$ is given by

$$
\begin{align*}
\Re\left[e_{i}\right] & =\frac{\Re\left[\operatorname{tr}\left(\mathbf{C}_{i}^{H} \hat{\mathbf{H}}^{H} \mathbf{E C}+\mathbf{C}^{H} \mathbf{E}^{H} \hat{\mathbf{H}} \mathbf{C}_{i}^{\prime}\right)\right]}{\lambda\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}} \\
& =\frac{\operatorname{tr}\left(\mathbf{C}_{i}^{H} \hat{\mathbf{H}}^{H} \mathbf{E C}+\mathbf{C}^{H} \mathbf{E}^{H} \hat{\mathbf{H}} \mathbf{C}_{i}+\mathbf{C}^{H} \mathbf{E}^{H} \hat{\mathbf{H}} \mathbf{C}_{i}^{\prime}+\mathbf{C}_{i}^{\prime H} \hat{\mathbf{H}}^{H} \mathbf{E C}\right)}{2 \lambda\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}} \\
& =\frac{1}{2 \lambda \| \hat{\mathbf{H}}_{2}^{2}} \operatorname{tr}\left(\left(\mathbf{C}_{i}+\mathbf{C}_{i}^{\prime}\right)^{H} \hat{\mathbf{H}}^{H} \mathbf{E C}+\mathbf{C}^{H} \mathbf{E}^{H} \hat{\mathbf{H}}\left(\mathbf{C}_{i}+\mathbf{C}_{i}^{\prime}\right)\right) \\
& =\frac{1}{\lambda\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}} \Re\left[\operatorname{tr}\left(\mathbf{C}\left(\mathbf{C}_{i}+\mathbf{C}_{i}^{\prime}\right)^{H} \hat{\mathbf{H}}^{H} \mathbf{E}\right)\right] . \tag{6.84}
\end{align*}
$$

Taking (6.24a) into account, the real part (6.84) of the disturbance term $e_{i}$ reduces to

$$
\begin{equation*}
\Re\left[e_{i}\right]=\frac{1}{\lambda\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}} \Re\left[\operatorname{tr}\left(\mathbf{C}_{i, \mathrm{R}}(\mathbf{s}) \hat{\mathbf{H}}^{H} \mathbf{E}\right)\right] \tag{6.85}
\end{equation*}
$$

Since it follows from (5.21) that in case of i.i.d. Rayleigh fading, the entries of E in (6.20) are i.i.d. ZM CSCG RVs with variance $N_{0} /\left(K_{\mathrm{p}} E_{\mathrm{p}}+N_{0}\right)$, it is readily verified that $\Re\left[e_{i}\right]$ is a ZM Gaussian RV with variance

$$
\begin{equation*}
\mathbb{E}\left[\left(\Re\left[e_{i}\right]\right)^{2} \mid \hat{\mathbf{H}}, \mathbf{s}\right]=\frac{N_{0}}{K_{\mathrm{p}} E_{\mathrm{p}}+N_{0}} \frac{\left\|\hat{\mathbf{H}} \mathbf{C}_{i, \mathrm{R}}^{H}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{2 \lambda^{2}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{4}} \tag{6.86}
\end{equation*}
$$

Similarly, the imaginary part of (6.20) can be shown to be given by

$$
\begin{align*}
\Im\left[e_{i}\right] & =\frac{\Im\left[\operatorname{tr}\left(\mathbf{C}_{i}^{H} \hat{\mathbf{H}}^{H} \mathbf{E C}+\mathbf{C}^{H} \mathbf{E}^{H} \hat{\mathbf{H}} \mathbf{C}_{i}^{\prime}\right)\right]}{\lambda\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}} \\
& =\frac{1}{\lambda\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}} \Im\left[\operatorname{tr}\left(\mathbf{C}_{i, \mathrm{I}}(\mathbf{s}) \hat{\mathbf{H}}^{H} \mathbf{E}\right)\right], \tag{6.87}
\end{align*}
$$

where the matrix $\mathbf{C}_{i, I}(\mathbf{s})$ is given by (6.24b). Hence, $\Im\left[e_{i}\right]$ is a ZM Gaussian RV with variance

$$
\begin{equation*}
\mathbb{E}\left[\left(\Im\left[e_{i}\right]\right)^{2} \mid \hat{\mathbf{H}}, \mathbf{s}\right]=\frac{N_{0}}{K_{\mathrm{p}} E_{\mathrm{p}}+N_{0}} \frac{\left\|\hat{\mathbf{H}} \mathbf{C}_{i, \mathrm{I}}^{H}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{2 \lambda^{2}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{4}} \tag{6.88}
\end{equation*}
$$

In general, according to (6.86) and (6.88), the variances of the real and imaginary parts of $e_{i}$ are not identical such that $e_{i}$ is not circularly symmetric.

Because the entries of $\mathbf{W}$ in (6.21) are i.i.d. ZM CSCG RVs with variance $N_{0}$, it is readily verified that $w_{i}$ is also a ZM CSCG RV, the variance of the real and imaginary parts of which is given by

$$
\begin{equation*}
\mathbb{E}\left[\left(\Re\left[w_{i}\right]\right)^{2} \mid \hat{\mathbf{H}}\right]=\mathbb{E}\left[\left(\Im\left[w_{i}\right]\right)^{2} \mid \hat{\mathbf{H}}\right]=\frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}} \tag{6.89}
\end{equation*}
$$

Since the total disturbance term $n_{i}$ in (6.19) is the sum of two uncorrelated terms, i.e., the ZM non-circularly symmetric complex Gaussian RV $e_{i}$ and the ZM CSCG RV $w_{i}, n_{i}$ is a ZM non-circularly symmetric complex Gaussian RV. From (6.86), (6.88), and (6.89), the variances of the real and imaginary parts of $n_{i}$ are easily derived as

$$
\begin{align*}
& \mathbb{E}\left[\left(\Re\left[n_{i}\right]\right)^{2} \mid \hat{\mathbf{H}}, \mathbf{s}\right]=\frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}}\left(1+\frac{\left\|\hat{\mathbf{H}} \mathbf{C}_{i, \mathrm{R}}^{H}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{\lambda\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}\left(\eta K_{\mathrm{p}}+\frac{N_{0}}{E_{\mathrm{s}}}\right)}\right),  \tag{6.90a}\\
& \mathbb{E}\left[\left(\Re\left[n_{i}\right]\right)^{2} \mid \hat{\mathbf{H}}, \mathbf{s}\right]=\frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}}\left(1+\frac{\left\|\hat{\mathbf{H}} \mathbf{C}_{i, \mathrm{R}}^{H}(\mathbf{s})\right\|_{\mathrm{F}}^{2}}{\lambda\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}\left(\eta K_{\mathrm{p}}+\frac{N_{0}}{E_{\mathrm{s}}}\right)}\right) . \tag{6.90b}
\end{align*}
$$

## 6.A. 2 Derivation of $\left\|\mathrm{C}_{i, \mathrm{q}}(\mathrm{s})\right\|_{\mathrm{F}}^{2}$ for Various Non-Square OSTBCs

In this section, we compute $\left\|\mathrm{C}_{i, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}$ and $\left\|\mathrm{C}_{i, \mathrm{I}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}$ for various well-known non-square OSTBCs.

- For the $3 \times 8$ OSTBC given by (4.33), we have

$$
\begin{align*}
&\left\|\mathbf{C}_{1, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=12 s_{1, \mathrm{R}}^{2}+8\left(s_{2, \mathrm{R}}^{2}+s_{3, \mathrm{R}}^{2}+s_{4, \mathrm{R}}^{2}\right)  \tag{6.91a}\\
&\left\|\mathbf{C}_{1, \mathrm{I}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=12 s_{1, \mathrm{I}}^{2}+8\left(s_{2, \mathrm{I}}^{2}+s_{3, \mathrm{I}}^{2}+s_{4, \mathrm{I}}^{2}\right)  \tag{6.91b}\\
&\left\|\mathbf{C}_{2, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=12 s_{2, \mathrm{R}}^{2}+8\left(s_{1, \mathrm{R}}^{2}+s_{3, \mathrm{R}}^{2}+s_{4, \mathrm{R}}^{2}\right)  \tag{6.91c}\\
&\left\|\mathbf{C}_{2, \mathrm{I}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=12 s_{2, \mathrm{I}}^{2}+8\left(s_{1, \mathrm{I}}^{2}+s_{3, \mathrm{I}}^{2}+s_{4, \mathrm{I}}^{2}\right)  \tag{6.91d}\\
&\left\|\mathbf{C}_{3, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=12 s_{3, \mathrm{R}}^{2}+8\left(s_{1, \mathrm{R}}^{2}+s_{2, \mathrm{R}}^{2}+s_{4, \mathrm{R}}^{2}\right)  \tag{6.91e}\\
&\left\|\mathbf{C}_{3, \mathrm{I}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=12 s_{3, \mathrm{I}}^{2}+8\left(s_{1, \mathrm{I}}^{2}+s_{2, \mathrm{I}}^{2}+s_{4, \mathrm{I}}^{2}\right)  \tag{6.91f}\\
&\left\|\mathbf{C}_{4, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=12 s_{4, \mathrm{R}}^{2}+8\left(s_{1, \mathrm{R}}^{2}+s_{2, \mathrm{R}}^{2}+s_{3, \mathrm{R}}^{2}\right)  \tag{6.91g}\\
&\left\|\mathbf{C}_{4, \mathrm{I}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=12 s_{4, \mathrm{I}}^{2}+8\left(s_{1, \mathrm{I}}^{2}+s_{2, \mathrm{I}}^{2}+s_{3, \mathrm{I}}^{2}\right) \tag{6.91h}
\end{align*}
$$

- For the $3 \times 4$ OSTBC given by (4.34), we have

$$
\begin{align*}
& \left\|\mathbf{C}_{1, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=\frac{16}{3}\left[s_{1, \mathrm{R}}^{2}+s_{2, \mathrm{I}}^{2}+\frac{2}{3}\left(s_{1, \mathrm{I}}^{2}+s_{2, \mathrm{R}}^{2}+s_{3, \mathrm{R}}^{2}+s_{3, \mathrm{I}}^{2}\right)\right]  \tag{6.92a}\\
& \left\|\mathbf{C}_{1, \mathrm{I}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=\frac{16}{3}\left[s_{1, \mathrm{I}}^{2}+s_{2, \mathrm{R}}^{2}+\frac{2}{3}\left(s_{1, \mathrm{R}}^{2}+s_{2, \mathrm{I}}^{2}+s_{3, \mathrm{R}}^{2}+s_{3, \mathrm{I}}^{2}\right)\right]  \tag{6.92b}\\
& \left\|\mathbf{C}_{2, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=\frac{16}{3}\left[s_{2, \mathrm{R}}^{2}+s_{1, \mathrm{I}}^{2}+\frac{2}{3}\left(s_{1, \mathrm{R}}^{2}+s_{2, \mathrm{I}}^{2}+s_{3, \mathrm{R}}^{2}+s_{3, \mathrm{I}}^{2}\right)\right]  \tag{6.92c}\\
& \left\|\mathbf{C}_{2, I}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=\frac{16}{3}\left[s_{2, \mathrm{I}}^{2}+s_{1, \mathrm{R}}^{2}+\frac{2}{3}\left(s_{1, \mathrm{I}}^{2}+s_{2, \mathrm{R}}^{2}+s_{3, \mathrm{R}}^{2}+s_{3, \mathrm{I}}^{2}\right)\right]  \tag{6.92d}\\
& \left\|\mathbf{C}_{3, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=\frac{16}{3}\left[s_{3, \mathrm{R}}^{2}+s_{3, \mathrm{I}}^{2}+\frac{2}{3}\left(s_{1, \mathrm{R}}^{2}+s_{1, \mathrm{I}}^{2}+s_{2, \mathrm{R}}^{2}+s_{2, \mathrm{I}}^{2}\right)\right]  \tag{6.92e}\\
& \left\|\mathbf{C}_{3, \mathrm{I}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=\frac{16}{3}\left[s_{3, \mathrm{I}}^{2}+s_{3, \mathrm{R}}^{2}+\frac{2}{3}\left(s_{1, \mathrm{I}}^{2}+s_{1, \mathrm{R}}^{2}+s_{2, \mathrm{I}}^{2}+s_{2, \mathrm{R}}^{2}\right)\right] \tag{6.92f}
\end{align*}
$$

- For the $3 \times 4$ OSTBC given by [27, eq. (99)]

$$
\mathbf{C}_{3 \times 4}=\frac{2}{\sqrt{3}}\left[\begin{array}{cccc}
s_{1} & -s_{2}^{*} & -s_{3}^{*} & 0  \tag{6.93}\\
s_{2} & s_{1}^{*} & 0 & -s_{3}^{*} \\
s_{3} & 0 & s_{1}^{*} & s_{2}^{*}
\end{array}\right]
$$

we have

$$
\begin{align*}
&\left\|\mathbf{C}_{1, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=\frac{16}{3}\left[\left|s_{1}\right|^{2}+\frac{2}{3}\left(\left|s_{2}\right|^{2}+\left|s_{3}\right|^{2}\right)\right]  \tag{6.94a}\\
&\left\|\mathbf{C}_{1, \mathrm{I}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=\frac{16}{3}\left[\left|s_{1}\right|^{2}+\frac{2}{3}\left(\left|s_{2}\right|^{2}+\left|s_{3}\right|^{2}\right)\right]  \tag{6.94b}\\
&\left\|\mathbf{C}_{2, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=\frac{16}{3}\left[\left|s_{2}\right|^{2}+\frac{2}{3}\left(\left|s_{1}\right|^{2}+\left|s_{3}\right|^{2}\right)\right]  \tag{6.94c}\\
&\left\|\mathbf{C}_{2, I}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=\frac{16}{3}\left[\left|s_{2}\right|^{2}+\frac{2}{3}\left(\left|s_{1}\right|^{2}+\left|s_{3}\right|^{2}\right)\right]  \tag{6.94d}\\
&\left\|\mathbf{C}_{3, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=\frac{16}{3}\left[\left|s_{3}\right|^{2}+\frac{2}{3}\left(\left|s_{1}\right|^{2}+\left|s_{2}\right|^{2}\right)\right]  \tag{6.94e}\\
&\left\|\mathbf{C}_{3, \mathrm{I}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}=\frac{16}{3}\left[\left|s_{3}\right|^{2}+\frac{2}{3}\left(\left|s_{1}\right|^{2}+\left|s_{2}\right|^{2}\right)\right] \tag{6.94f}
\end{align*}
$$

- For a four-antenna transmitter, we present the following $4 \times 8$ OSTBC [8, eq. (38)]

$$
\mathbf{C}_{3 \times 8}=\left[\begin{array}{cccccccc}
s_{1} & -s_{2} & -s_{3} & -s_{4} & s_{1}^{*} & -s_{2}^{*} & -s_{3}^{*} & -s_{4}^{*}  \tag{6.95}\\
s_{2} & s_{1} & s_{4} & -s_{3} & s_{2}^{*} & s_{1}^{*} & s_{4}^{*} & -s_{3}^{*} \\
s_{3} & -s_{4} & s_{1} & s_{2} & s_{3}^{*} & -s_{4}^{*} & s_{1}^{*} & s_{2}^{*} \\
s_{4} & s_{3} & -s_{2} & s_{1} & s_{4}^{*} & s_{3}^{*} & -s_{2}^{*} & s_{1}^{*}
\end{array}\right],
$$

which yields

$$
\begin{align*}
\left\|\mathbf{C}_{1, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2} & =16\left(s_{1, \mathrm{R}}^{2}+s_{2, \mathrm{R}}^{2}+s_{3, \mathrm{R}}^{2}+s_{4, \mathrm{R}}^{2}\right)  \tag{6.96a}\\
\left\|\mathbf{C}_{1, \mathrm{I}}(\mathbf{s})\right\|_{\mathrm{F}}^{2} & =16\left(s_{1, \mathrm{I}}^{2}+s_{2, \mathrm{I}}^{2}+s_{3, \mathrm{I}}^{2}+s_{4, \mathrm{I}}^{2}\right)  \tag{6.96b}\\
\left\|\mathbf{C}_{2, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2} & =16\left(s_{1, \mathrm{R}}^{2}+s_{2, \mathrm{R}}^{2}+s_{3, \mathrm{R}}^{2}+s_{4, \mathrm{R}}^{2}\right)  \tag{6.96c}\\
\left\|\mathbf{C}_{2, \mathrm{I}}(\mathbf{s})\right\|_{\mathrm{F}}^{2} & =16\left(s_{1, \mathrm{I}}^{2}+s_{2, \mathrm{I}}^{2}+s_{3, \mathrm{I}}^{2}+s_{4, \mathrm{I}}^{2}\right)  \tag{6.96d}\\
\left\|\mathbf{C}_{3, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2} & =16\left(s_{1, \mathrm{R}}^{2}+s_{2, \mathrm{R}}^{2}+s_{3, \mathrm{R}}^{2}+s_{4, \mathrm{R}}^{2}\right)  \tag{6.96e}\\
\left\|\mathbf{C}_{3, \mathrm{I}}(\mathbf{s})\right\|_{\mathrm{F}}^{2} & =16\left(s_{1, \mathrm{I}}^{2}+s_{2, \mathrm{I}}^{2}+s_{3, \mathrm{I}}^{2}+s_{4, \mathrm{I}}^{2}\right)  \tag{6.96f}\\
\left\|\mathbf{C}_{4, \mathrm{R}}(\mathbf{s})\right\|_{\mathrm{F}}^{2} & =16\left(s_{1, \mathrm{R}}^{2}+s_{2, \mathrm{R}}^{2}+s_{3, \mathrm{R}}^{2}+s_{4, \mathrm{R}}^{2}\right)  \tag{6.96g}\\
\left\|\mathbf{C}_{4, \mathrm{I}}(\mathbf{s})\right\|_{\mathrm{F}}^{2} & =16\left(s_{1, \mathrm{I}}^{2}+s_{2, \mathrm{I}}^{2}+s_{3, \mathrm{I}}^{2}+s_{4, \mathrm{I}}^{2}\right) \tag{6.96h}
\end{align*}
$$

Taking into account that all QAM symbols are equally likely and, thus, all $s_{i, \mathrm{q}}$ have equal probability, it follows from the symmetry observed in $\left\|\mathbf{C}_{i, \mathrm{q}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}$ for different $i$ and q , that the resulting BERs for the in-phase and quadrature bits are identical and independent of the index $i$, for the non-square OSTBCs considered in this appendix. Moreover, since $\left\|\mathrm{C}_{i, \mathrm{q}}(\mathbf{s})\right\|_{\mathrm{F}}^{2}$ is given by a weighted sum of $s_{i, q}^{2}$, we can restrict the summation over $\mathbf{s}$ to a summation over the symbols with positive real and imaginary parts only when averaging the conditional BERs over s.

## 6.A. 3 Derivation of the Variance of $n_{i}$ for Correlated Rayleigh Fading

Since $\varepsilon$ and $\mathbf{w}$ are uncorrelated, the terms (6.59) and (6.60) are uncorrelated when conditioned on $\hat{\mathbf{h}}$, such that the variance of the real/imaginary part of $n_{i}$ in (6.58) equals the sum of the variances of the real/imaginary parts of $e_{i}$ and $w_{i}$. According to (6.59), the real part of $e_{i}$ is given by

$$
\begin{align*}
\Re\left[e_{i}\right]= & \frac{1}{2 \lambda\|\hat{\mathbf{h}}\|^{2}}\left[\hat{\mathbf{h}}^{H}\left(\mathbf{C}_{i}^{*} \mathbf{C}^{T} \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \varepsilon+\varepsilon^{H}\left(\mathbf{C}^{*} \mathbf{C}_{i}^{T} \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \hat{\mathbf{h}}\right. \\
& \left.\quad+\varepsilon^{H}\left(\mathbf{C}^{*} \mathbf{C}_{i}^{\prime T} \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \hat{\mathbf{h}}+\hat{\mathbf{h}}^{H}\left(\mathbf{C}_{i}^{*} \mathbf{C}^{T} \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \varepsilon\right] \\
= & \frac{1}{2 \lambda\|\hat{\mathbf{h}}\|^{2}}\left[\hat{\mathbf{h}}^{H}\left(\left(\mathbf{C}_{i}+\mathbf{C}_{i}^{\prime}\right)^{*} \mathbf{C}^{T} \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \varepsilon+\varepsilon^{H}\left(\mathbf{C}^{*}\left(\mathbf{C}_{i}+\mathbf{C}_{i}^{\prime}\right)^{T} \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \hat{\mathbf{h}}\right] \\
= & \frac{1}{\lambda\|\hat{\mathbf{h}}\|^{2}} \Re\left[\hat{\mathbf{h}}^{H}\left(\mathbf{C}_{i, \mathrm{R}}^{T}(\mathbf{s}) \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \varepsilon\right] . \tag{6.97}
\end{align*}
$$

where $\mathbf{C}_{i, R}(\mathbf{s})$ is defined in (6.24a). Since it follows from (5.20) that in case of correlated Rayleigh fading, the entries of $\varepsilon$ are ZM CSCG RVs with covariance
matrix $\boldsymbol{\mathcal { R }}_{\varepsilon}$, it is readily verified that $\Re\left[e_{i}\right]$ is a ZM Gaussian RV with variance

$$
\begin{equation*}
\mathbb{E}\left[\left(\Re\left[e_{i}\right]\right)^{2} \mid \hat{\mathbf{h}}, \mathbf{s}\right]=\frac{1}{2 \lambda^{2}\|\hat{\mathbf{h}}\|^{4}}\left[\hat{\mathbf{h}}^{H}\left(\mathbf{C}_{i, \mathrm{R}}^{T}(\mathbf{s}) \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \mathcal{R}_{\varepsilon}\left(\mathbf{C}_{i, \mathrm{R}}^{*}(\mathbf{s}) \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \hat{\mathbf{h}}\right] . \tag{6.98}
\end{equation*}
$$

Similarly, the imaginary part of $e_{i}$ is given by

$$
\begin{align*}
\Im\left[e_{i}\right] & =\frac{1}{2 j \lambda\|\hat{\mathbf{h}}\|^{2}}\left[\hat{\mathbf{h}}^{H}\left(\left(\mathbf{C}_{i}-\mathbf{C}_{i}^{\prime}\right)^{*} \mathbf{C}^{T} \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \varepsilon-\varepsilon^{H}\left(\mathbf{C}^{*}\left(\mathbf{C}_{i}-\mathbf{C}_{i}^{\prime}\right)^{T} \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \hat{\mathbf{h}}\right] \\
& =\frac{1}{\lambda\|\hat{\mathbf{h}}\|^{2}} \Im\left[\hat{\mathbf{h}}^{H}\left(\mathbf{C}_{i, I}^{T}(\mathbf{s}) \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \varepsilon\right] . \tag{6.99}
\end{align*}
$$

where $\mathbf{C}_{i, \mathrm{I}}(\mathbf{s})$ is defined in (6.24b). Hence, $\Im\left[e_{i}\right]$ is a ZM Gaussian RV with variance

$$
\begin{equation*}
\mathbb{E}\left[\left(\Im\left[e_{i}\right]\right)^{2} \mid \hat{\mathbf{h}}, \mathbf{s}\right]=\frac{1}{2 \lambda^{2}\|\hat{\mathbf{h}}\|^{4}}\left[\hat{\mathbf{h}}^{H}\left(\mathbf{C}_{i, \mathrm{I}}^{T}(\mathbf{s}) \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \boldsymbol{\mathcal { R }}_{\varepsilon}\left(\mathbf{C}_{i, \mathrm{I}}^{*}(\mathbf{s}) \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \hat{\mathbf{h}}\right] \tag{6.100}
\end{equation*}
$$

Similar to the case of i.i.d. Rayleigh fading, the variances of the real and imaginary parts of $e_{i}$ are not identical such that $e_{i}$ is not circularly symmetric.

Since the entries of $\mathbf{w}$ in (6.21) are i.i.d. ZM CSCG RVs with variance $N_{0}$, it is readily verified that $w_{i}$ is also a ZM CSCG RV, the variance of the real and imaginary parts of which is given by

$$
\begin{equation*}
\mathbb{E}\left[\left(\Re\left[w_{i}\right]\right)^{2} \mid \hat{\mathbf{h}}\right]=\mathbb{E}\left[\left(\Im\left[w_{i}\right]\right)^{2} \mid \hat{\mathbf{h}}\right]=\frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{h}}\|^{2}} \tag{6.101}
\end{equation*}
$$

From (6.98), (6.100), and (6.101), it follows that, in case of correlated Rayleigh fading, $n_{i}$ is a ZM non-circularly symmetric complex Gaussian RV, the variances of the real and imaginary parts of which are given by

$$
\begin{align*}
& \mathbb{E}\left[\left(\Re\left[n_{i}\right]\right)^{2} \mid \hat{\mathbf{h}}, \mathbf{s}\right]=\frac{N_{0}}{2 \lambda E_{\mathrm{S}}\|\hat{\mathbf{h}}\|^{2}} \\
& \quad \times\left(1+\frac{E_{\mathrm{s}}}{\lambda N_{0}\|\hat{\mathbf{h}}\|^{2}} \hat{\mathbf{h}}^{H}\left(\mathbf{C}_{i, \mathrm{R}}^{T}(\mathbf{s}) \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \mathcal{R}_{\varepsilon}\left(\mathbf{C}_{i, \mathrm{R}}^{*}(\mathbf{s}) \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \hat{\mathbf{h}}\right),  \tag{6.102a}\\
& \mathbb{E}\left[\left(\Im\left[n_{i}\right]\right)^{2} \mid \hat{\mathbf{h}}, \mathbf{s}\right]=\frac{N_{0}}{2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{h}}\|^{2}} \\
& \quad \times\left(1+\frac{E_{\mathrm{S}}}{\lambda N_{0}\|\hat{\mathbf{h}}\|^{2}} \hat{\mathbf{h}}^{H}\left(\mathbf{C}_{i, \mathrm{I}}^{T}(\mathbf{s}) \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \mathcal{R}_{\varepsilon}\left(\mathbf{C}_{i, \mathrm{I}}^{*}(\mathbf{s}) \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \hat{\mathbf{h}}\right) . \tag{6.102b}
\end{align*}
$$

## 6.A.4 Impact of Spatial Correlation on the BER

Assuming BPSK or 4-QAM constellations under PCSI, it is easily obtained from (3.61) and (4.43) that the conditional BER for OSTBCs, conditioned on
the channel vector $\mathbf{h}$, is given by

$$
\begin{equation*}
P_{\mathrm{b}}(\mathbf{h})=Q\left(\sqrt{2 \lambda \frac{E_{\mathrm{b}}}{N_{0}}|\mathbf{h}|^{2}}\right) . \tag{6.103}
\end{equation*}
$$

Using the Chernoff bound (3.66), the conditional BER (6.103) is bounded by

$$
\begin{equation*}
P_{\mathrm{b}}(\mathbf{h}) \leq \frac{1}{2} \exp \left(\lambda \frac{E_{\mathrm{b}}}{N_{0}}|\mathbf{h}|^{2}\right) \tag{6.104}
\end{equation*}
$$

Since the channel vector $\mathbf{h}$ consists of $L$ correlated ZM CSCG RVs with covariance matrix $\mathcal{R}=\mathbb{E}\left[\mathbf{h} h^{H}\right]$, the joint distribution of the channel coefficients is given by

$$
\begin{equation*}
p(\mathbf{h})=\frac{1}{\pi^{L} \operatorname{det}(\boldsymbol{\mathcal { R }})} \exp \left(-\mathbf{h}^{H} \boldsymbol{\mathcal { R }}^{-1} \mathbf{h}\right) \tag{6.105}
\end{equation*}
$$

where $\operatorname{det}(\mathbf{X})$ denotes the determinant of $\mathbf{X}$. Using (6.104) and (6.105), a bound on the average BER for OSTBCs using BPSK or 4-QAM constellations under PCSI is easily obtained as

$$
\begin{align*}
P_{\mathrm{b}} & \leq \frac{1}{2 \pi^{L} \operatorname{det}(\boldsymbol{\mathcal { R }})} \int_{\mathbf{h}} \exp \left[-\mathbf{h}^{H}\left(\lambda \frac{E_{\mathrm{b}}}{N_{0}} \mathbf{I}_{L}+\boldsymbol{\mathcal { R }}^{-1}\right) \mathbf{h}\right] \mathrm{d} \mathbf{h} \\
& =\frac{1}{2 \operatorname{det}\left(\mathbf{I}_{L}+\lambda \frac{E_{\mathrm{b}}}{N_{0}} \boldsymbol{\mathcal { R }}\right)} \\
& =\frac{1}{2 \prod_{i=1}^{L}\left(1+\lambda \frac{E_{\mathrm{b}}}{N_{0}} \lambda_{i}\right)} \tag{6.106}
\end{align*}
$$

where $\lambda_{i}$ are the eigenvalues of $\mathcal{R}$. Note that $\lambda_{i}$ are real-valued and positive, since $\mathcal{R}$ is a positive semi-definite Hermitian matrix. Defining the high-SNR region as the region where

$$
\begin{equation*}
\frac{E_{\mathrm{b}}}{N_{0}} \gg \frac{1}{\lambda \lambda_{\min }} \tag{6.107}
\end{equation*}
$$

with $\lambda_{\text {min }}$ denoting the smallest eigenvalue of the covariance matrix $\boldsymbol{\mathcal { R }}$, it follows from (6.106) that, in the high-SNR region, the BER for OSTBCs using BPSK or 4-QAM constellations under PCSI is bounded as follows

$$
\begin{equation*}
P_{\mathrm{b}} \leq \frac{1}{2 \prod_{i=1}^{L}\left(\lambda \frac{E_{\mathrm{b}}}{N_{0}} \lambda_{i}\right)} \tag{6.108}
\end{equation*}
$$

which indicates that a diversity gain of $L$ is obtained for high SNR as long as $\lambda_{\text {min }}>0$. Taking the slope of the BER curves in the high-SNR region into account, it is easily obtained from (6.108) that the BER degradation (6.74) due to spatial correlation is given by

$$
\begin{equation*}
\Delta_{\mathrm{BER}, \mathrm{corr}} \approx\left(\prod_{i=1}^{L} \lambda_{i}\right)^{-\frac{1}{L}} \tag{6.109}
\end{equation*}
$$

## 7

## BER Analysis of OSTBCs in Arbitrary Fading

In chapter 6, we have derived exact and approximate closed-form BER expressions for a mismatched OSTBC receiver under Rayleigh fading. In many realistic situations, however, Rayleigh fading does not accurately model the true channel, and other more versatile distributions, such the Nakagami-m distribution, are more suitable. For these distributions, the channel is not Gaussian, and closed-form BER expressions are hard to obtain. Instead, the BER can be calculated through Monte-Carlo simulation, as mentioned in section 3.3.2. However, OSTBCs achieve very low BERs thanks to spatial diversity, such that extremely long simulation times are usually necessary. Moreover, the required simulation time increases dramatically with the SNR, making Monte-Carlo simulations, in general, inappropriate for accurate and efficient BER computations of OSTBCs.

The impact of imperfect channel estimation on the performance of OSTBCs under generalized fading conditions has been studied analytically in [71],
where the SER as well as the decoding error probability (DEP) of square OSTBCs have been examined for M-PSK constellations. However, in [72], we have shown that the method in [71] to compute the error performance includes an approximation and cannot easily be generalized to QAM and PAM constellations. In this chapter, we provide an exact analytical BER analysis for square and non-square OSTBCs with M-ary QAM constellations and ML channel estimation, under the assumption of flat-fading channels with an arbitrary joint PDF. We also show how the exact expressions can be efficiently and accurately evaluated using numerical integration techniques. As the high diversity order resulting from the application of OSTBCs gives rise to small BER values, the numerical evaluation of the presented BER expressions is much faster than straightforward Monte-Carlo simulations. Furthermore, we provide a simple approximate BER expression based on treating the symbol interference due to imperfect channel estimation as white Gaussian noise. Although the resulting BER is in general not asymptotically exact, it yields quite accurate BER results in many practical applications.

This chapter is organized as follows. In section 7.1, we explain briefly how approximate BER curves for OSTBCs under generalized fading conditions can be easily obtained from the rule of thumb, which was derived for i.i.d. Rayleigh fading in section 6.1. The exact BER expressions for OSTBCs under arbitrary fading with ML channel estimation are presented in section 7.2. Section 7.3 deals with the efficient and accurate evaluation of the exact BER expressions, and outlines a method for Monte-Carlo integration with importance sampling. In section 7.4, Monte-Carlo simulations of the receiver operations confirm our efficient numerical evaluation methods and the impact of several system and channel parameters on the BER performance is investigated. Moreover, the accuracy and complexity of the presented BER expressions is discussed. Finally, conclusions are drawn in section 7.5.

### 7.1 Approximate BER Analysis

The BER curve for a PCSI receiver under arbitrary fading conditions can be obtained from the literature, or from averaging the conditional BER resulting from (4.43) over the PDF of the squared channel norm $\|\mathbf{H}\|_{\mathrm{F}}^{2}$. The BER of a mismatched OSTBC receiver can be easily approximated by shifting the BER curve for PCSI to the right over an amount given by the rule of thumb (6.13), which was derived for i.i.d. Rayleigh fading:

$$
\begin{equation*}
\Delta_{\mathrm{BER}, \mathrm{~dB}}=10 \log _{10}\left[\left(1+\frac{\eta K_{\mathrm{p}}}{K}\right)\left(1+\frac{L_{\mathrm{t}}}{\eta K_{\mathrm{p}}}\right)\right] . \tag{7.1}
\end{equation*}
$$

However, it is important to note that for non-Rayleigh fading channels, the channel estimation error $\mathbf{E}=\mathbf{H}-\hat{\mathbf{H}}$ and, consequently, the interference term $\sqrt{E_{s}} \mathrm{EC}$ in (6.3) are not Gaussian. Moreover, the useful terms in (6.2) and (6.3)
do not have similar statistical properties, as $\hat{\mathbf{H}}$ is a linear function of both the channel $\mathbf{H}$ and the Gaussian noise matrix $\mathbf{W}_{\mathrm{p}}$. Hence, the more the fading 'differs' from i.i.d. Rayleigh fading, the less accurate the BER resulting from (7.1) will be.

### 7.2 Exact BER Analysis

In this section, we derive exact BER expressions for square QAM constellations with Gray mapping. Since square $M-Q A M$ with Gray mapping reduces to $\sqrt{M}$-PAM for both the in-phase and quadrature information bits, the BER computation for PAM follows the same lines as the BER computation for the in-phase bits in case of QAM.

By substituting $\mathbf{R}$ in (6.18) by (6.3), we have decomposed in chapter 6 the decision variable $u_{i}$ for the mismatched receiver into the sum of two terms, i.e., the transmitted symbol $s_{i}$ and a disturbance term $n_{i}$, which contains contributions from the channel noise $\mathbf{W}$ and the channel estimation error E. Since the estimation error $\mathbf{E}$ is not Gaussian in case of non-Rayleigh fading, the disturbance term $n_{i}$ in (6.19) is not Gaussian either, and the BER analysis presented in chapter 6 can not straightforwardly be extended to arbitrary fading conditions. Instead, we substitute $\mathbf{R}$ in (6.18) by (6.2), such that the decision variable (6.18) reduces to the sum of a signal $u_{i}^{\prime}=u_{i, \mathrm{R}}^{\prime}+j u_{i, \mathrm{I}}^{\prime}$, with $u_{i, \mathrm{R}}^{\prime}$ and $u_{i, \mathrm{I}}^{\prime}$ denoting the real and imaginary parts of $u_{i}^{\prime}$, respectively, and a Gaussian noise term $w_{i}$

$$
\begin{equation*}
u_{i}=u_{i}^{\prime}+w_{i}, \tag{7.2}
\end{equation*}
$$

where $w_{i}$ is defined in (6.21), and $u_{i}^{\prime}$ is a function of the transmitted symbol vector $\mathbf{s}$ through the code matrix $\mathbf{C}$

$$
\begin{equation*}
u_{i}^{\prime} \triangleq \frac{\operatorname{tr}\left(\mathbf{C}_{i}^{H} \hat{\mathbf{H}}^{H} \mathbf{H C}+\mathbf{C}^{H} \mathbf{H}^{H} \hat{\mathbf{H}} \mathbf{C}_{i}^{\prime}\right)}{\lambda\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}} \tag{7.3}
\end{equation*}
$$

In appendix 6.A.1, we have shown that, when conditioned on the estimated channel $\hat{\mathbf{H}}, w_{i}$ is a ZM CSCG RV, the real and imaginary parts of which have variance $N_{0} /\left(2 \lambda E_{\mathrm{s}}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}\right)$. Taking (4.28) into account, it follows that $u_{i}^{\prime}$ contains a useful term proportional to $s_{i}$ and interference terms containing the data symbols $s_{n}$, with $n \neq i$. If PCSI is available, i.e., $\hat{\mathbf{H}}=\mathbf{H}$, (7.3) reduces to $u_{i}^{\prime}=s_{i}$ because of (4.35), indicating that only a useful signal term is present. Due to the rotational symmetry of the M-QAM constellation and the uniform distribution of the symbol vector $s$, it follows that the BERs related to the in-phase and quadrature bits of $s_{i}$ are identical and irrespective of $i$ in case of PCSI. In case of imperfect channel estimation, however, the BER can be written as

$$
\begin{equation*}
P_{\mathrm{b}}=\frac{1}{2 N_{\mathrm{s}}} \sum_{i=1}^{N_{\mathrm{s}}}\left[P_{\mathrm{b}, i, \mathrm{R}}+P_{\mathrm{b}, i, \mathrm{I}}\right], \tag{7.4}
\end{equation*}
$$

where $P_{\mathrm{b}, i, \mathrm{R}}$ and $P_{\mathrm{b}, i, \mathrm{I}}$ denote the BERs of the in-phase bits and quadrature bits corresponding to the information symbols $s_{i}$, respectively. Since $u_{i}^{\prime}$ is a function of $\mathbf{s}, \mathbf{H}$, and $\hat{\mathbf{H}}$, we can easily obtain the conditional BERs of the in-phase and quadrature bits corresponding to the information symbols $s_{i}$, conditioned on $\mathbf{s}, \mathbf{H}$, and $\hat{\mathbf{H}}$

$$
\begin{equation*}
P_{\mathrm{b}, i, \mathrm{q}}(\mathbf{s}, \mathbf{H}, \hat{\mathbf{H}})=\frac{1}{\log _{2}(\sqrt{M})} \sum_{b_{\mathrm{q}} \in \Psi_{\mathrm{q}}} d_{\mathrm{H}}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right) \operatorname{Pr}\left[\hat{s}_{i, \mathrm{q}}=b_{\mathrm{q}} \mid \mathbf{s}, \mathbf{H}, \hat{\mathbf{H}}\right] \tag{7.5}
\end{equation*}
$$

where $\mathrm{q}=\mathrm{R}$ or $\mathrm{q}=\mathrm{I}$, and $d_{\mathrm{H}}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)$ denotes the Hamming distance between the bits associated to $s_{i, \mathrm{q}}$ and $b_{\mathrm{q}}$. It follows from (3.57) that $\operatorname{Pr}\left[\hat{s}_{i, \mathrm{q}}=b_{\mathrm{q}} \mid \mathbf{s}, \mathbf{H}, \hat{\mathbf{H}}\right]$ is given by

$$
\begin{align*}
\operatorname{Pr}\left[\hat{s}_{i, \mathrm{q}}=b_{\mathrm{q}} \mid \mathbf{s}, \mathbf{H}, \hat{\mathbf{H}}\right] & =\operatorname{Pr}\left[d_{1}\left(b_{\mathrm{q}}\right) \leq u_{i, \mathrm{q}} \leq d_{2}\left(b_{\mathrm{q}}\right) \mid \mathbf{s}, \mathbf{H}, \hat{\mathbf{H}}\right] \\
& =Q_{i, \mathrm{q}, 1}-Q_{i, \mathrm{q}, 2} \tag{7.6}
\end{align*}
$$

where $d_{1}\left(b_{\mathrm{q}}\right)$ and $d_{2}\left(b_{\mathrm{q}}\right)$ denote the boundaries of the decision area of $b_{\mathrm{q}}$, with $d_{1}\left(b_{\mathrm{q}}\right)<d_{2}\left(b_{\mathrm{q}}\right)$, and the quantities $Q_{i, \mathrm{q}, k}$, with $k=1$ and 2 , are given by

$$
\begin{equation*}
Q_{i, \mathrm{q}, k} \triangleq Q\left\{\sqrt{2 \lambda \frac{E_{\mathrm{s}}}{N_{0}}\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}}\left[d_{k}\left(b_{\mathrm{q}}\right)-u_{i, \mathrm{q}}^{\prime}\right]\right\} . \tag{7.7}
\end{equation*}
$$

Note that $Q_{i, \mathrm{q}, k}$ is a function of $\mathbf{s}, \mathbf{H}$, and $\hat{\mathbf{H}}$ through $u_{i, \mathrm{q}}^{\prime}$. Finally, $P_{\mathrm{b}, i, \mathrm{R}}$ and $P_{\mathrm{b}, i, \mathrm{I}}$ in (7.4) are obtained by averaging the corresponding conditional BERs (7.5) over s, $\mathbf{H}$, and $\hat{\mathbf{H}}$

$$
\begin{equation*}
P_{\mathrm{b}, i, \mathrm{q}}=\frac{1}{M^{N_{\mathrm{s}}}} \sum_{\mathbf{s} \in \Psi^{N_{\mathrm{s}}}} \mathbb{E}_{\mathbf{H}, \hat{\mathbf{H}}}\left[P_{\mathrm{b}, i, \mathrm{q}}(\mathbf{s}, \mathbf{H}, \hat{\mathbf{H}})\right] . \tag{7.8}
\end{equation*}
$$

From (7.8), it follows that the evaluation of the BER requires averaging over $4 L$ real-valued continuous RVs, i.e., the real and imaginary parts of the elements of $\mathbf{H}$ and $\hat{\mathbf{H}}$, and over $N_{\mathrm{s}}$ discrete RVs, i.e., the $N_{\mathrm{s}}$ symbols contained in $\mathbf{s}$.

In order to reduce the computational complexity related to the numerical evaluation of the BER, we will decrease the number of RVs involved in the expectation (7.8) by using an appropriate coordinate transformation. To this end, we introduce the $2 L \times 1$ real-valued column vectors $\hat{\mathbf{g}}$ and $\mathbf{g}$, which contain all elements of $\hat{\mathbf{H}}$ and $\mathbf{H}$, respectively, as

$$
\begin{align*}
& \hat{\mathbf{g}} \triangleq\left[\hat{\mathbf{h}}_{1, \mathrm{R}}^{T}, \hat{\mathbf{h}}_{1, \mathrm{I}}^{T}, \hat{\mathbf{h}}_{2, \mathrm{R}}^{T}, \hat{\mathbf{h}}_{2, \mathrm{I}}^{T}, \ldots, \hat{\mathbf{h}}_{L_{\mathrm{t}}, \mathrm{R}}^{T}, \hat{\mathbf{h}}_{L_{\mathrm{t}, \mathrm{I}}}^{T}\right]^{T},  \tag{7.9a}\\
& \mathbf{g} \triangleq\left[\mathbf{h}_{1, \mathrm{R}}^{T}, \mathbf{h}_{1, \mathrm{I}}^{T}, \mathbf{h}_{2, \mathrm{R}}^{T}, \mathbf{h}_{2, \mathrm{I}}^{T}, \ldots, \mathbf{h}_{L_{\mathrm{t}}, \mathrm{R}}^{T}, \mathbf{h}_{L_{\mathrm{t}, \mathrm{I}}}^{T}\right]^{T}, \tag{7.9b}
\end{align*}
$$

with $\hat{\mathbf{h}}_{i, \mathrm{R}}+j \hat{\mathbf{h}}_{i, \mathrm{I}}$ and $\mathbf{h}_{i, \mathrm{R}}+j \mathbf{h}_{i, \mathrm{I}}$ denoting the $i$-th column of $\hat{\mathbf{H}}$ and $\mathbf{H}$, respectively. It can be easily seen that $\|\hat{\mathbf{g}}\|=\|\hat{\mathbf{H}}\|_{\mathrm{F}}$ and $\|\mathbf{g}\|=\|\mathbf{H}\|_{\mathrm{F}}$. Using (7.9), the coordinate transformation derived in appendix 7.A. 1 shows that, under the
assumption of ML channel estimation, the real and imaginary parts of (7.3) reduce to

$$
\begin{equation*}
u_{i, \mathrm{q}}^{\prime}=\frac{1}{\lambda\left(x_{1}^{2}+x_{2}^{2}+z^{2}\right)}\left(\lambda x_{1} s_{i, \mathrm{q}}\|\mathbf{g}\|+x_{2}\left\|\mathbf{M}_{i, \mathrm{q}} \mathbf{g}\right\|\right) \tag{7.10}
\end{equation*}
$$

where $\mathrm{q}=\mathrm{R}$ or $\mathrm{q}=\mathrm{I}$, and the $2 L \times 2 L$ matrix $\mathbf{M}_{i, \mathrm{q}}$, given by (7.43), incorporates the interference from the signal components different from $s_{i, \mathrm{q}}$. When conditioned on $\mathbf{g}$, the real-valued $\operatorname{RVs} x_{1}, x_{2}$, and $z$ are independent and distributed as follows:

- $x_{1}$ is a Gaussian RV with mean $\|\mathbf{g}\|$ and variance $\sigma_{\mathbf{N}}^{2}=N_{0} /\left(2 K_{\mathrm{p}} E_{\mathrm{p}}\right)$;
- $x_{2}$ is a ZM Gaussian RV with variance $\sigma_{\mathbf{N}}^{2}$;
- $z / \sigma_{\mathbf{N}}$ is distributed according to the chi-distribution with $2 L-2$ degrees of freedom [73].

Hence, by substituting $u_{i, q}^{\prime}$ and $\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}$ in (7.7) by (7.10) and (7.47), respectively, the conditional BER given by (7.5) can be rewritten as a function that depends on the actual channel $\mathbf{H}$ through the random vector $\mathbf{g}$ and on the estimated channel $\hat{\mathbf{H}}$ through only 3 RVs: $x_{1}, x_{2}$, and $z$; we denote this function by $\mathrm{B}_{i, \mathrm{q}, 1}\left(\mathbf{s}, \mathbf{g}, x_{1}, x_{2}, z\right)$. Note that the dependence on $\mathbf{g}$ is through $\|\mathbf{g}\|$ and $\left\|\mathbf{M}_{i, \mathrm{q}} \mathbf{g}\right\|$, with $\mathbf{M}_{i, \mathrm{q}}$ depending on $\mathbf{s}$. Due to this substitution, the BER expression given by (7.4) reduces to

$$
\begin{align*}
& P_{\mathrm{b}}=\frac{1}{2 N_{\mathrm{s}}} \frac{1}{M^{N_{\mathrm{s}}}} \sum_{i=1}^{N_{\mathrm{s}}} \sum_{\mathbf{s} \in \Psi^{N_{\mathbf{s}}}} \mathbb{E}_{\mathbf{g}, x_{1}, x_{2}, z}\left[\mathrm{~B}_{i, \mathrm{R}, 1}\left(\mathbf{s}, \mathbf{g}, x_{1}, x_{2}, z\right)\right. \\
&\left.+\mathrm{B}_{i, \mathrm{I}, 1}\left(\mathbf{s}, \mathbf{g}, x_{1}, x_{2}, z\right)\right] \tag{7.11}
\end{align*}
$$

which is an expectation over $2 L+3$ real-valued continuous RVs, i.e., the $2 L$ components of $\mathbf{g}, x_{1}, x_{2}$, and $z$, and $N_{\mathrm{s}}$ discrete RVs, i.e., the components of $\mathbf{s}$.

In the case of square OSTBCs, i.e., $L_{t}=K_{c}$, the BER expression can be considerably simplified. For these OSTBCs, it is shown in appendix 7.A. 2 that the magnitude of $\mathbf{M}_{i, q} \mathbf{g}$ used in (7.10) is given by

$$
\begin{equation*}
\left\|\mathbf{M}_{i, \mathbf{q}} \mathbf{g}\right\|=\lambda\|\mathbf{g}\| \sqrt{\|\mathbf{s}\|^{2}-s_{i, \mathbf{q}}^{2}} \tag{7.12}
\end{equation*}
$$

Hence, by substituting $\left\|\mathbf{M}_{i, \mathrm{q}} \mathbf{g}\right\|$ in $\mathrm{B}_{i, \mathrm{q}, 1}\left(\mathbf{s}, \mathbf{g}, x_{1}, x_{2}, z\right)$ by (7.12), the conditional BER $\mathrm{B}_{i, \mathrm{q}, 1}\left(\mathbf{s}, \mathbf{g}, x_{1}, x_{2}, z\right)$ can be rewritten as a function that depends on $\mathbf{g}$ through only the norm $\|\mathbf{g}\|$ of the channel vector; we denote this function by $\mathrm{B}_{i, \mathrm{q}, 2}\left(\mathbf{s},\|\mathbf{g}\|, x_{1}, x_{2}, z\right)$. It follows from (7.12) that, for square OSTBCs, the dependence of (7.10) on $i$ and q is through $s_{i, \mathrm{q}}$ only. Since the statistical properties of $s_{i, \mathrm{q}}$ depend neither on $i$ nor on q , the BERs related to the in-phase
and quadrature bits of $s_{i}$ are identical and irrespective of $i$, such that (7.11) reduces to

$$
\begin{equation*}
P_{\mathrm{b}}=\left(\frac{4}{M}\right)^{N_{\mathbf{s}}} \sum_{\mathbf{s} \in \Psi_{0}^{N_{\mathrm{s}}}} \mathbb{E}_{\|\mathbf{g}\|, x_{1}, x_{2}, z}\left[\mathrm{~B}_{i, \mathrm{q}, 2}\left(\mathbf{s},\|\mathbf{g}\|, x_{1}, x_{2}, z\right)\right] \tag{7.13}
\end{equation*}
$$

where we have restricted the summation over $\mathbf{s}$ to constellation points with positive real and imaginary parts only, since because of (7.12), the symbol vectors $\mathbf{s}, \mathbf{s}^{*},-\mathbf{s}$, and $-\mathbf{s}^{*}$ yield the same BER result. Note that (7.13) involves the expectation over only 4 real-valued continuous RVs, i.e, the channel norm $\|\mathbf{g}\|=\|\mathbf{H}\|, x_{1}, x_{2}$, and $z$, and $N_{s}$ discrete RVs, i.e., the components of $\mathbf{s}$.

### 7.3 Evaluation of the BER

In section 7.2, we provided an exact analysis of the BER for OSTBCs on fading channels with ICSI, regardless of the fading distribution. In particular, it has been shown that for any OSTBC, the BER can be written as an expectation over $2 L+3$ real-valued continuous RVs and $N_{s}$ discrete RVs, as can be seen from (7.11). Moreover, for square OSTBCs, the BER can be further reduced to an expectation over only 4 real-valued continuous $R V s$ and $N_{s}$ discrete RVs, as can be seen from (7.13). In this section, we deal with the efficient and accurate numerical evaluation of the exact BER expressions (7.11) and (7.13). To this end, two numerical integration techniques will be envisaged: the quadrature rule [74, Sec. 4.1] and Monte-Carlo integration [74, Sec. 7.7] with importance sampling [74, Sec. 7.9.1].

### 7.3.1 Efficient Evaluation of (7.11) and (7.13)

By numerically evaluating (7.11) and (7.13), the BER for OSTBCs can be efficiently obtained with a computation time that increases only very slowly with the SNR. In this section, we briefly describe the quadrature rule and Monte-Carlo integration with importance sampling, point out their benefits and limitations, and apply them to evaluate (7.11) and (7.13).

Let us represent (7.11) and (7.13) by the following generic expectation

$$
\begin{equation*}
\mathrm{BER}=\mathbb{E}_{p}[\mathrm{~B}(\mathbf{v})] \tag{7.14}
\end{equation*}
$$

where $\mathrm{B}(\mathbf{v})$ is a function of a random vector $\mathbf{v}=\left[\mathbf{s}^{T}, \mathbf{u}^{T}\right]^{T}$ consisting of the symbol vector $\mathbf{s}$ and a random column vector $\mathbf{u}$, and the subscript $p$ refers to the joint PDF $p(\mathbf{v})=p(\mathbf{s}) p(\mathbf{u})$ of $\mathbf{v}$, with $\mathbf{s}$ being uniformly distributed over $\Psi^{N_{s}}$. As far as the evaluation of (7.11) is considered, we define

$$
\left\{\begin{array}{l}
\mathbf{u} \triangleq\left[\mathbf{g}^{T}, x_{1}, x_{2}, z\right]^{T}  \tag{7.15}\\
\mathrm{~B}(\mathbf{v}) \triangleq \frac{1}{2 N_{\mathrm{s}}} \sum_{i=1}^{N_{\mathrm{s}}}\left[\mathrm{~B}_{i, \mathrm{R}, 1}(\mathbf{v})+\mathrm{B}_{i, \mathrm{~L}, 1}(\mathbf{v})\right]
\end{array}\right.
$$

such that the expectation (7.11) is given by the following sum of a $J$-fold integral, with $J=2 L+3$ denoting the dimension of $\mathbf{u}$,

$$
\begin{equation*}
\mathbb{E}_{p}[\mathrm{~B}(\mathbf{v})]=\frac{1}{M^{N_{s}}} \sum_{\mathbf{s} \in \Psi^{N_{s}}} \int_{\mathbf{u}} \mathrm{B}(\mathbf{s}, \mathbf{u}) p(\mathbf{u}) \mathrm{d} \mathbf{u} \tag{7.16}
\end{equation*}
$$

Note that the computational complexity associated with the summation in (7.16) is proportional to $M^{N_{s}}$, which increases prohibitively for large constellation size $M$ and/or large number of information symbols $N_{\mathrm{s}}$.

For square OSTBCs, we define

$$
\left\{\begin{array}{l}
\mathbf{u} \triangleq\left[\|\mathbf{g}\|, x_{1}, x_{2}, z\right]^{T}  \tag{7.17}\\
\mathrm{~B}(\mathbf{v}) \triangleq \mathrm{B}_{i, \mathrm{q}, 2}(\mathbf{v})
\end{array}\right.
$$

with $q=R$ or $q=I$, such that the expectation (7.13) is defined by the following sum of a $J$-fold integral, with $J=4$ denoting the dimension of $\mathbf{u}$,

$$
\begin{equation*}
\mathbb{E}_{p}[\mathrm{~B}(\mathbf{v})] \triangleq\left(\frac{4}{M}\right)^{N_{\mathrm{s}}} \sum_{\mathbf{s} \in \Psi_{0}^{N_{\mathrm{s}}}} \int_{\mathbf{u}} \mathrm{B}(\mathbf{s}, \mathbf{u}) p(\mathbf{u}) \mathrm{d} \mathbf{u} \tag{7.18}
\end{equation*}
$$

### 7.3.1.1 The Quadrature Rule

In principle, the $J$-fold integral in (7.16) and (7.18) can be evaluated by repeatedly applying the quadrature rule to each of the integrals, in which case the $J$-fold integral is replaced by a $J$-fold sum. Assuming that each element of $\mathbf{v}$ takes $I$ values in this sum, the computational complexity for computing the $J$-fold integral is proportional to $I^{J}$, which is prohibitively large for large $J$. Therefore, numerical integration is only of practical interest for square OSTBCs, where $\mathbf{u} \triangleq\left[\|\mathbf{g}\|, x_{1}, x_{2}, z\right]^{T}$ and $J=4$. In this case, the joint PDF of $\mathbf{u}$ is easily obtained as the product of the PDFs $p\left(x_{1} \mid\|\mathbf{g}\|\right), p\left(x_{2}\right)$, and $p(z)$, which have been specified in section 7.2, and the PDF of $\|\mathbf{g}\|$, which can be available as an analytical expression or in the form of a histogram (e.g., as the result of a measurement), since $\|\mathbf{g}\|$ is to be discretized.

### 7.3.1.2 Monte-Carlo Integration with Importance Sampling

The exponential dependency of the computational complexity related to (7.16) and (7.18) on $N_{\mathrm{s}}$ and $J$ can be avoided by using Monte-Carlo integration combined with importance sampling. In this way, the expectation in the righthand side of (7.14) is approximated as

$$
\begin{equation*}
G_{\mathcal{N}}\left(\left\{\mathbf{v}_{v}\right\}\right)=\frac{1}{\mathcal{N}} \sum_{v=1}^{\mathcal{N}} \mathrm{B}\left(\mathbf{v}_{v}\right) \frac{p\left(\mathbf{v}_{v}\right)}{q\left(\mathbf{v}_{v}\right)} \tag{7.19}
\end{equation*}
$$

where $\left\{\mathbf{v}_{v}, v=1,2, \ldots, \mathcal{N}\right\}$ are independent samples generated according to a biased PDF $q(\mathbf{v})$. Note that we use the term Monte-Carlo integration for both
discrete and continuous variables, i.e., the symbol vector $\mathbf{s}$ and the vector $\mathbf{u}$, respectively. In case $q\left(\mathbf{v}_{v}\right)=p\left(\mathbf{v}_{v}\right)$, (7.19) reduces to conventional MonteCarlo integration without importance sampling. Defining $e_{\mathcal{N}} \triangleq G_{\mathcal{N}}\left(\left\{\mathbf{v}_{v}\right\}\right)-$ $\mathbb{E}_{p}[\mathrm{~B}(\mathbf{v})]$, it can be shown that $\mathbb{E}\left[e_{\mathcal{N}}\right]=0$ and

$$
\begin{equation*}
\mathbb{E}\left[e_{\mathcal{N}}^{2}\right]=\frac{1}{\mathcal{N}}\left\{\mathbb{E}_{q}\left[\left(\frac{\mathrm{~B}(\mathbf{v}) p(\mathbf{v})}{q(\mathbf{v})}\right)^{2}\right]-\mathbb{E}_{p}^{2}[\mathrm{~B}(\mathbf{v})]\right\} \tag{7.20}
\end{equation*}
$$

where $\mathbb{E}_{q}[\cdot]$ refers to the expectation over the biased PDF $q(\mathbf{v})$. It follows from (7.20) that $\mathbb{E}\left[e_{\mathcal{N}}^{2}\right]$ can be made arbitrarily small by taking $\mathcal{N}$ sufficiently large. However, the smaller the second factor in (7.20), the smaller is the value of $\mathcal{N}$ required to achieve a certain value of $\mathbb{E}\left[e_{\mathcal{N}}^{2}\right]$. By making a judicious choice of $q(\mathbf{v})$, we try to minimize this factor, so that reasonably large values of $\mathcal{N}$ yield very good accuracy. It can be easily verified that $q(\mathbf{v})=\mathrm{B}(\mathbf{v}) p(\mathbf{v}) / \mathbb{E}_{p}[\mathrm{~B}(\mathbf{v})]$ is the optimum biased PDF, as it yields $\mathbb{E}\left[e_{\mathcal{N}}^{2}\right]=0$. However, this choice is not practical, since the optimum $q(\mathbf{v})$ depends on the unknown $\mathbb{E}_{p}[\mathrm{~B}(\mathbf{v})]$. Nevertheless, the optimum $q(\mathbf{v})$ inspires us to take

$$
\begin{equation*}
q(\mathbf{v})=\mathrm{B}_{\mathrm{app}}(\mathbf{v}) p(\mathbf{v}) / \mathbb{E}_{p}\left[\mathrm{~B}_{\mathrm{app}}(\mathbf{v})\right] \tag{7.21}
\end{equation*}
$$

where $B_{\text {app }}(\mathbf{v})$ is a suitable approximation of $B(\mathbf{v})$, i.e., it is chosen in such way that the resulting $q(\mathbf{v})$ allows us to easily generate i.i.d. vectors $\left\{\mathbf{v}_{v}\right\}$.

Let us apply Monte-Carlo integration with importance sampling for evaluating the BER in the case of non-square OSTBCs, where $\mathbf{v} \triangleq\left[\mathbf{s}^{T}, \mathbf{g}^{T}, x_{1}, x_{2}, z\right]^{T}$. For many fading distributions, the vector $\mathbf{g}$ can easily be generated as a transformation $\mathbf{g}=\phi(\gamma)$ of a vector $\gamma$ of auxiliary RVs, distributed according to a joint PDF $p(\gamma)$ which is such that the PDF of $\mathbf{g}=\phi(\gamma)$ is the desired distribution of the real and imaginary parts of the channel coefficients contained in $\mathbf{g}$. Therefore, we redefine $\mathbf{v}$ in terms of the vector of auxiliary RVs instead of the channel vector, i.e., $\mathbf{v} \triangleq\left[\mathbf{s}^{T}, \gamma^{T}, x_{1}, x_{2}, z\right]^{T}$, such that

$$
\begin{equation*}
\mathrm{B}(\mathbf{v}) \triangleq \frac{1}{2 N_{\mathrm{s}}} \sum_{i=1}^{N_{\mathrm{s}}}\left\{\mathrm{~B}_{i, \mathrm{R}, 1}\left[\mathbf{s}, \phi(\gamma), x_{1}, x_{2}, z\right]+\mathrm{B}_{i, \mathrm{I}, 1}\left[\mathbf{s}, \phi(\gamma), x_{1}, x_{2}, z\right]\right\} \tag{7.22}
\end{equation*}
$$

and

$$
\begin{equation*}
p(\mathbf{v})=p(\mathbf{s}) p\left(x_{1} \mid\|\mathbf{g}\|=\|\phi(\gamma)\|\right) p\left(x_{2}\right) p(z) p(\gamma) \tag{7.23}
\end{equation*}
$$

In order to identify an approximate $\mathrm{B}_{\text {app }}(\mathbf{v})$ in (7.21), we will consider $\mathrm{B}(\mathbf{v})$ assuming PCSI, in which case $u_{i, \mathrm{R}}^{\prime}$ from (7.10) is the sum of $s_{i, \mathrm{R}}$ and a noise term with variance $N_{0} /\left[2 \lambda E_{\mathrm{s}}\|\phi(\gamma)\|^{2}\right]$. Denoting by $2 d$ the distance between adjacent QAM constellation points, we approximate the conditional BER assuming PCSI by the following simple expression

$$
\begin{equation*}
\mathrm{B}_{\mathrm{PCSI}}(\mathbf{v}) \approx \frac{N_{\min }}{\log _{2}(M)} Q\left(\sqrt{\frac{2 \lambda E_{\mathrm{s}} d^{2}\|\phi(\gamma)\|^{2}}{N_{0}}}\right) \tag{7.24}
\end{equation*}
$$

with $N_{\min }$ being given by

$$
\begin{equation*}
N_{\min }=\frac{1}{M} \sum_{i=1}^{M} N_{\min }\left(\psi_{i}\right) \tag{7.25}
\end{equation*}
$$

where $N_{\min }\left(\psi_{i}\right)$ denotes the number of neighbors at minimal distance $2 d$ from a constellation point $\psi_{i}$. Hence, $N_{\min }$ denotes the average number of constellation points at minimal distance $2 d$, given a constellation $\Psi$. Considering the bound $Q(x) \leq(1 / 2) \exp \left(-x^{2} / 2\right)$, we select

$$
\begin{equation*}
\mathrm{B}_{\mathrm{app}}(\mathbf{v})=\frac{N_{\min }}{2 \log _{2}(M)} \exp \left(-\lambda d^{2} \frac{E_{\mathrm{s}}}{N_{0}}\|\phi(\gamma)\|^{2}\right) . \tag{7.26}
\end{equation*}
$$

Since $\mathrm{B}_{\text {app }}(\mathbf{v})$ depends only on $\gamma$, the corresponding biased PDF (7.21) is given by

$$
\begin{equation*}
q(\mathbf{v})=p(\mathbf{s}) p\left(x_{1}\|\mathbf{g}\|=\|\phi(\gamma)\|\right) p\left(x_{2}\right) p(z) q(\gamma) \tag{7.27}
\end{equation*}
$$

with

$$
\begin{equation*}
q(\gamma)=C p(\gamma) \exp \left(-\lambda d^{2} \frac{E_{\mathrm{S}}}{N_{0}}\|\phi(\gamma)\|^{2}\right) \tag{7.28}
\end{equation*}
$$

where $C$ is a normalization constant. From (7.27), it follows that only the joint PDF $q(\gamma)$ of the vector $\gamma$ of auxiliary RVs is biased. Hence, $\mathbf{s}, x_{1}, x_{2}$, and $z$ are generated according to their respective distributions, whereas the auxiliary RVs are generated according to the biased joint PDF given by (7.28), which depends on the considered transformation $\phi(\gamma)$.

Note that also a combination of the quadrature rule and Monte-Carlo integration can be used. For example, for square OSTBCs, the integral over $\mathbf{u} \triangleq\left[\|\mathbf{g}\|, x_{1}, x_{2}, z\right]^{T}$ can be evaluated using the quadrature rule, whereas the summation over the symbol vector s can be evaluated through Monte-Carlo integration. In order to further reduce the computational complexity related to the evaluation of $(7.16), \mathrm{B}(\mathbf{v})$ can be approximated by retaining only one term in the summation in (7.22), i.e., $\mathrm{B}_{i, \mathrm{q}, 1}\left[\mathrm{~s}, \phi(\gamma), x_{1}, x_{2}, z\right]$, with $\mathrm{q}=\mathrm{R}$ or $\mathrm{q}=\mathrm{I}$, instead of taking the average over all $2 N_{\mathrm{s}}$ terms. As we have shown that this approximation yields the exact result in case of Rayleigh fading for the non-square OSTBCs considered in appendix 6.A.2, it could be expected that also for non-Rayleigh fading very accurate results could be obtained. In section 7.4, comparisons between analytical BER results and computer simulations confirm this assumption.

In principle, importance sampling can also be used for direct Monte-Carlo simulations of the BER, involving bit error counting. However, finding a suitable biased PDF according to which the RVs (i.e., the data symbols to be transmitted, the channel and the additive Gaussian channel noise) need to be generated in order to decrease the variance of the resulting BER will be much more complicated than what we have shown for the numerical evaluation of (7.16). In the latter case, the integrand is given by a summation of $Q$-functions
and can be approximated by the exponential function (7.26), which allows us to easily find a suitable biased PDF. In the case of direct Monte-Carlo simulations, however, the integrand consists of a summation of indicator functions that indicate whether or not the detected symbols equal the transmitted symbols. These indicator functions will be functions of the data symbols, the channel, the channel estimate and the additive channel noise. Therefore, it is easily understood that finding a suitable approximation of the integrand from which an optimal or suboptimal biased PDF can be found will be a very complicated task.

### 7.3.2 Correlated Nakagami-m Fading Channels

In this section, we apply the theory shown in section 7.3.1 for the evaluation of the BER of OSTBCs with ICSI to the particular case of arbitrarily correlated Nakagami- $m$ fading channels.

### 7.3.2.1 The Quadrature Rule

In section 4.3.2.3, we introduced the Nakagami- $m$ distribution (4.61) as a versatile statistical distribution that accurately models a variety of fading environments by selecting a proper value for the fading parameter $m \geq 1 / 2$. Moreover, the Nakagami- $m$ distribution includes the Rayleigh ( $m=1$ ) and the one-sided Gaussian ( $m=1 / 2$ ) distributions as special cases. The PDF of the squared channel norm $\|\mathbf{H}\|_{\mathrm{F}}^{2}$, which is required to average (7.18) over $\|\mathbf{g}\|=\|\mathbf{H}\|_{\mathrm{F}}$ by means of the quadrature rule, is given in section 4.3.4.2 for i.i.d. and correlated Nakagami-m channels. Assuming PCSI, it follows from [75] that the BER of an OSTBC operating over $L$ i.i.d. Nakagami- $m$ fading channels is proportional to $\left(E_{\mathrm{b}} / N_{0}\right)^{-m L}$ for large $E_{\mathrm{b}} / N_{0}$, which indicates that the BER performance improves with increasing $m$ and/or diversity order $L$.

### 7.3.2.2 Monte-Carlo Integration with Importance Sampling

In section 7.3.1, we showed that the expectation over $\mathbf{g}$ in (7.11) can be efficiently evaluated by means of Monte-Carlo integration with importance sampling, provided that the vector $\gamma$ of auxiliary RVs that is used to obtain the channel vector $\mathbf{g}$ is generated according to the sampling distribution given by (7.28). For integer and identical fading parameters $m_{\ell}=m, \forall \ell$, we showed in section 4.3.3 how $L$ correlated Nakagami-m RVs $\alpha_{\ell}$, with $\mathbb{E}\left[\alpha_{\ell}^{2}\right]=\Omega_{\ell}$ and power correlation matrix $\Sigma$, can be obtained from $2 m$ i.i.d. random vectors $\mathbf{y}_{k}=\left[y_{k, 1}, y_{k, 2}, \ldots, y_{k, L}\right]^{T}$, with $k=1,2, \ldots, 2 m$. The random vectors $\mathbf{y}_{k}$ consist of real-valued ZM Gaussian auxiliary RVs and have a covariance matrix $\mathbf{Q}$ given by (4.67).

In this section, we derive the sampling distribution (7.28) for the auxiliary RVs, in order to enable efficient numerical evaluation of the expectation in (7.11) in case of correlated Nakagami- $m$ fading channels with integer and identical $m_{\ell}=m, \forall \ell, \mathbb{E}\left[\alpha_{\ell}^{2}\right]=\Omega_{\ell}$, and power correlation matrix $\Sigma$. To this end, we select $\gamma=\left[\boldsymbol{\theta}^{T}, \mathbf{y}^{T}\right]^{T}$ as the vector of auxiliary RVs, where $\boldsymbol{\theta}$ contains the phases of the $L$ channel coefficients and $\mathbf{y}=\left[\mathbf{y}_{1}^{T}, \mathbf{y}_{2}^{T}, \ldots, \mathbf{y}_{2 m}^{T}\right]^{T}$. Taking into account that $\mathbf{g}=\phi(\gamma)$ and that the $L$ channel coefficient magnitudes $\alpha_{\ell}$ are obtained from $\mathbf{y}$ according to (4.66), we have $\|\mathbf{g}\|^{2}=\|\phi(\gamma)\|^{2}=\|\boldsymbol{\alpha}\|^{2}=\|\mathbf{y}\|^{2}$. Hence, with $q(\boldsymbol{\theta}, \mathbf{y})=p(\boldsymbol{\theta} \mid \mathbf{y}) q(\mathbf{y})$, it follows from (7.28) that the biased joint PDF of $\mathbf{y}$ is given by

$$
\begin{equation*}
q(\mathbf{y})=C p(\mathbf{y}) \exp \left(-\lambda d^{2} \frac{E_{\mathrm{s}}}{N_{0}}\|\mathbf{y}\|^{2}\right) \tag{7.29}
\end{equation*}
$$

Taking into account that $p(\mathbf{y})$ is the joint PDF of $2 m$ i.i.d. ZM Gaussian vectors $\left\{\mathbf{y}_{k}\right\}$, each having a covariance matrix $\mathbf{Q}$ given by (4.67), we show in appendix 7.A. 3 that $q(\mathbf{y})$ is a similar PDF, but now the vectors $\left\{\mathbf{y}_{k}\right\}$ have a covariance matrix $\mathbf{Q}^{\prime}$ given by

$$
\begin{equation*}
\mathbf{Q}^{\prime}=\mathbf{Q}\left(\mathbf{I}_{L}+2 \lambda d^{2} \frac{E_{\mathrm{s}}}{N_{0}} \mathbf{Q}\right)^{-1} \tag{7.30}
\end{equation*}
$$

Moreover, using (7.29) and (7.30), it is readily verified that the ratio $p(\mathbf{v}) / q(\mathbf{v})$ to be used in (7.19) depends only on $\mathbf{y}$ and is given by

$$
\begin{equation*}
\frac{p(\mathbf{v})}{q(\mathbf{v})}=\frac{\exp \left(\lambda d^{2} \frac{E_{\mathrm{s}}}{N_{0}}\|\mathbf{y}\|^{2}\right)}{\left[\operatorname{det}\left(\mathbf{I}_{L}+2 \lambda d^{2} \frac{E_{\mathrm{s}}}{N_{0}} \mathbf{Q}\right)\right]^{m}} \tag{7.31}
\end{equation*}
$$

### 7.4 Numerical Results

In this section, BER results are presented for correlated Nakagami-m fading channels, under the assumption that $E_{\mathrm{p}}=E_{\mathrm{s}}$ and that the Kronecker channel model $(4.55)$ is valid $[34,36]$, although the analysis is also applicable to arbitrary power correlation matrices. The phases of the channel coefficients are assumed to be uniformly distributed. The accuracy of the curves resulting from the numerical evaluation of the exact BER expressions (7.11) and (7.13) using the tools provided in section 7.3, is illustrated by means of straightforward Monte-Carlo simulation results, which are added to some of the figures for relatively high BER. In Figs. 7.1 and 7.2, the BER curves resulting from the high-SNR approximation given in section 7.1 are also shown, in order to illustrate their accuracy under different fading circumstances.

### 7.4.1 Square OSTBCs

Let us consider Alamouti's code ( $L_{\mathrm{t}}=K_{\mathrm{c}}=N_{\mathrm{s}}=2$ ), which is given by (4.14). In order to obtain the BER curves for this OSTBC, we evaluate the expectation
over $\|\mathbf{g}\|, x_{1}, x_{2}$, and $z$ in (7.13) by means of the quadrature rule, with the distribution of $\|\mathbf{g}\|$ being derived from (4.73) in the case of correlated fading and from (4.61) in the case of i.i.d. fading; the expectation over $\mathbf{s}$ is exactly obtained by means of a finite summation.

In Fig. 7.1(a), we show the BER curves for a $2 \times 1$ Alamouti MIMO scheme with ML channel estimation under i.i.d. Nakagami- $m$ fading. Also the approximate BER curves, resulting from treating the symbol interference due to ICSI as white Gaussian noise, are shown in the figure. The data frames consist of $K=40$ coded data symbols and $K_{p}=4$ pilot symbols per transmit antenna, whereas the symbols belong to a 4-QAM constellation. The BER approximation turns out to be relatively accurate, although it is clearly not asymptotically exact when the fading is not Rayleigh distributed, i.e., for $m>1$. The difference between the Gaussian approximation and the exact result is even larger when BPSK transmission is considered, as shown in Fig. 7.1(b).

Fig. 7.2 shows the BER of Alamouti's code under correlated identically distributed Nakagami- $m$ channels with $m=4$ and $\Omega=1$. We assume that there is no antenna correlation at the transmitter side, whereas the correlation between the receive antennas $\left(L_{r}=3\right)$ can be described by means of a constant correlation model [44] determined by the following power correlation matrix

$$
\boldsymbol{\Sigma}_{\mathrm{r}}=\left[\begin{array}{ccc}
1 & 0.8 & 0.8  \tag{7.32}\\
0.8 & 1 & 0.8 \\
0.8 & 0.8 & 1
\end{array}\right]
$$

For $M \in\{4,16,64,256\}$, the BER results are shown for a PCSI receiver and a receiver using ML channel estimation with $K=80$ and $K_{p} \in\{2,4,12\}$. For the receiver with ICSI, both the exact and approximate BER curves are displayed. It follows from the figure that the accuracy of the approximate BER curves depends on different parameters, such as the constellation size and the number of pilot symbols $K_{p}$. The larger $K_{p}$ is, the smaller the asymptotic difference will be between the approximate and the exact BER curves. This is due to the fact that for large $K_{\mathrm{p}}$, the joint PDFs of $\hat{\mathbf{H}}$ and $\mathbf{H}$ are quite similar such that the useful terms in (6.2) and (6.3) will have similar statistical properties and the approximation from section 7.1 is very accurate.

Fig. 7.3 displays the performance curves resulting from the exact BER expression (7.13) for Alamouti's code along with 4-QAM signaling, operating over correlated identically distributed Nakagami- $m$ channels with $\Omega=1$ and $L_{\mathrm{r}}=3$. The results are shown for both a PCSI receiver and a receiver using ML channel estimation with $K=200$ and $K_{\mathrm{p}}=20$, and for $m \in\{1,3\}$. While no antenna correlation occurs at the transmitter side, three different correlation scenarios are considered at the receiver side. More specifically, the case of uncorrelated fading is compared with two practical receive antenna configurations which are described in more detail in [42]. The correlation matrices corresponding to the considered antenna configurations were obtained in [42]


Figure 7.1: Exact and approximate BER versus $E_{\mathrm{b}} / N_{0}$ for a $2 \times 1$ Alamouti scheme with ML channel estimation under i.i.d. Nakagami- $m$ fading.


Figure 7.2: Exact and approximate BER versus $E_{\mathrm{b}} / N_{0}$ for a $2 \times 3$ Alamouti scheme with ML channel estimation under correlated identically distributed Nakagami- $m$ fading, with $K=80, K_{p} \in\{2,4,12\}$, and $M \in\{4,16,64,256\}$.
using the empirical curves of Lee [76, p. 203]. Briefly, the three correlation scenarios at the receiver side can be described as follows:
i) uncorrelated fading (unc), with $\Sigma_{r}=\mathbf{I}_{L_{\mathrm{r}}}$,
ii) a linear antenna array (lin), the configuration of which, along with the angle of arrival, is depicted in [42, Fig. 4(b)], with power correlation matrix $\Sigma_{\mathrm{r}}$ given by [42, Eq. (38)]

$$
\Sigma_{\mathrm{r}}=\left[\begin{array}{ccc}
1 & 0.795 & 0.605  \tag{7.33}\\
0.795 & 1 & 0.795 \\
0.605 & 0.795 & 1
\end{array}\right]
$$

iii) a triangular antenna array (tri), the configuration of which, along with the angle of arrival, is depicted in [42, Fig. 4(a)], with power correlation matrix $\Sigma_{\mathrm{r}}$ given by [42, Eq. (37)]

$$
\Sigma_{\mathrm{r}}=\left[\begin{array}{ccc}
1 & 0.727 & 0.913  \tag{7.34}\\
0.727 & 1 & 0.913 \\
0.913 & 0.913 & 1
\end{array}\right]
$$



Figure 7.3: BER versus $E_{\mathrm{b}} / N_{0}$ for Alamouti's code with 4-QAM signaling and ML channel estimation under correlated identically distributed Nakagami-m fading channels with $m \in\{1,3\}$ and $\Omega=1$, for uncorrelated (unc), linear (lin), and triangular (tri) correlation matrices at the receiver side ( $L_{r}=3$ ).

In Fig. 7.3, it is shown how $m$, ICSI, and the antenna correlation model affect the BER performance of Alamouti's code. As it is expected, the BER performance improves when $m$ increases. As compared to the case of PCSI and zero correlation, both ICSI and antenna correlation degrade the BER through a horizontal shift of the BER curve for large $E_{\mathrm{b}} / N_{0}$, indicating that, for the correlations considered, the relation BER $\propto\left(E_{\mathrm{b}} / N_{0}\right)^{-m L}$ still holds at large $E_{\mathrm{b}} / N_{0}$. Note that antenna correlation has no significant impact on the BER degradation caused by ICSI only, and that for highly correlated channels, e.g., the triangular correlation model, the BER degradation as compared to zero correlation is much larger than the degradation due to ICSI only.

### 7.4.2 Non-Square OSTBCs

Let us consider the $3 \times 4 \operatorname{OSTBC}\left(L_{\mathrm{t}}=N_{\mathrm{s}}=3, K_{\mathrm{c}}=4\right)$ given by (4.34). The BER curves for this OSTBC are obtained by evaluating the expectation over $\mathbf{s}, \mathbf{g}, x_{1}, x_{2}$, and $z$ in (7.11) by means of Monte-Carlo integration with
importance sampling. In (7.19), however, $\mathrm{B}(\mathbf{v})$ is approximated by retaining only the term $\mathrm{B}_{1, \mathrm{R}, 1}\left[\mathbf{s}, \phi(\gamma), x_{1}, x_{2}, z\right]$ in the summation in (7.22). Moreover, the auxiliary Gaussian RVs yielding the channel coefficients are generated according to (7.29).

Fig. 7.4 shows BER performance evaluation curves in case of a Nakagami$m$ MIMO channel with $m=2$ that is recovered through ML channel estimation with $K=200$ and $K_{\mathrm{p}}=16$. The power correlation matrix $\Sigma_{\mathrm{t}}$ at the transmitter side equals the matrix (7.33), whereas the power correlation matrix $\Sigma_{\mathrm{r}}$ of the dual-antenna receiver $\left(L_{r}=2\right)$ is given by

$$
\Sigma_{\mathrm{r}}=\left[\begin{array}{cc}
1 & 0.4  \tag{7.35}\\
0.4 & 1
\end{array}\right]
$$

The BER performance evaluation results are shown for $M$-QAM signaling, with $M \in\{4,16,64,256\}$ and a $3 \times 2$ Nakagami- $m$ MIMO channel satisfying

$$
\left\{\begin{array}{l}
\Omega_{\ell, 1}=\Omega_{\ell, 2}=1  \tag{7.36}\\
\Omega_{\ell, 3}=t
\end{array}\right.
$$

with $\left.\Omega_{\ell, k}=\left.\mathbb{E}[\mid \mathbf{H})_{\ell, k}\right|^{2}\right], \ell \in\{1,2\}$, and $t \in\{1,0.5\}$. Note that halving the average energy transfer between the third transmit antenna and the receiver causes a BER degradation through a horizontal shift of the BER curve. In order to compare the proposed BER expressions versus direct simulations in terms of accuracy and efficiency, we use (6.77), which requires that the ratio of the variance of the calculated or simulated BER to the square of its expectation is less or equal than a certain prescribed value $\epsilon^{2}$. Note that the BER estimate $\hat{P}_{\mathrm{b}}$ in (6.77) is given by (6.78), and that $N$ and $X_{l}$ denote the number $\mathcal{N}$ of generated sample vectors $\left(\mathbf{v}_{v}\right)$ and the summand in (7.19), respectively, in case the proposed BER expressions are evaluated through Monte-Carlo integration, whereas for direct simulations, $N$ and $X_{l}$ denote the number of simulated data frames and the ratio of the number of bit errors counted in the $l$-th frame to the total number of bits within one frame, respectively. Since $X_{l}{ }^{\prime}$ 's are independently generated, the minimum $N$ required to obtain a certain accuracy $\epsilon^{2}$ for the resulting BER satisfies (6.80). Under the assumption of $M=4$, $t=1$, and a given accuracy of $\epsilon^{2}=0.01$, Fig. 7.5 displays the minimum $N$ as a function of the SNR for both direct simulations and the evaluation of the BER expressions presented in section (7.2) through Monte-Carlo integration. To illustrate the impact of importance sampling, we consider both the case with and without importance sampling, applying the sampling distribution derived in section 7.3.2.2 for arbitrarily correlated Nakagami- $m$ fading both to Monte-Carlo integration and direct simulations. Although the latter sampling distribution does not guarantee optimal performance in case of direct simulations, as mentioned in the last paragraph of section 7.3.1.2, it follows from the figure that importance sampling results in a significant reduction of the


Figure 7.4: BER versus $E_{\mathrm{b}} / N_{0}$ for the $3 \times 4$ OSTBC given by (4.34) operating over a $3 \times 2$ correlated Nakagami- $m$ MIMO fading channel with $m=2$, $\Omega_{\ell, 1}=\Omega_{\ell, 2}=1$, and $\Omega_{\ell, 3}=t$. The results are shown for M-QAM, with $M \in\{4,16,64,256\}$, and for $t \in\{1,0.5\}$.
computation time for both direct simulations and numerical evaluation of the exact BER espressions through Monte-Carlo integration. For moderate to high SNR, the number of required sample vectors $\mathbf{v}_{v}$ in case of Monte-Carlo integration with importance sampling is much less than the number of generated frames in case of direct simulations. However, for low SNR, where the BER is high, it might be more efficient to apply direct simulations. It is also important to note that generating a data frame and detecting it after it has been affected by fading and channel noise, is a much more complex task than evaluating the summand in (7.19). In Fig. 7.4, the BER results from Monte-Carlo integration with importance sampling have been obtained with $\mathcal{N}=5000$ sample vectors $\mathbf{v}_{v}$. Note also that retaining only one term in the summation in (7.22) yields very accurate BER results, i.e., the BERs related to the in-phase and quadrature bits of $s_{i}$ are nearly identical and irrespective of $i$. Because the number of pilot symbols is relatively high $\left(K_{p}=16\right)$, the approximate BER curves resulting from the high-SNR approximation in section 7.1 also provide very accurate BER results. Only for 4-QAM, a clear deviation from the exact


Figure 7.5: Minimum $N$ for both direct simulations and the evaluation of the proposed BER expressions through Monte-Carlo integration for a given accuracy of $\epsilon^{2}=0.01$.

BER can be observed for SNRs below 10 dB . In order not to overload Fig. 7.4, however, the approximate BER curves are omitted in the figure.

Fig. 7.6 illustrates the BER performance versus the number of pilot symbols for the $3 \times 4$ OSTBC given by (4.34), operating over correlated identically distributed Nakagami- $m$ fading channels with $m=2$ and $\Omega=1$, under the assumption that $E_{\mathrm{b}} / N_{0}=10 \mathrm{~dB}$. The power correlation matrix $\Sigma_{\mathrm{r}}$ of the dual-antenna receiver $\left(L_{r}=2\right)$ is assumed to be given by (7.35), the channel is recovered through ML channel estimation, and the transmitted information symbols belong to a 16-QAM constellation. The results are shown for $K \in\{100,400,1000\}$ information symbols and for the power correlation matrix $\Sigma_{\mathrm{t}}$ at the transmitter side being given by either (7.33) (lin) or the identity matrix (unc). From Fig. 7.6, we observe that the optimal number of pilot symbols grows with the number of information symbols $K$ and that antenna correlation does not affect this optimal number. We also notice that for large $K$, obtaining the optimal number of pilot symbols is not very critical as the BER grows only slowly when more pilot symbols are added. According to (6.14), the optimal number of pilot symbols in case of i.i.d. Rayleigh fading


Figure 7.6: BER versus $K_{\mathrm{p}}$ for the $3 \times 4$ OSTBC given by (4.34) with 16 -QAM, operating over identically correlated Nakagami- $m$ fading channels, with $m=$ 2 and $\Omega=1$. The results are shown for $E_{\mathrm{b}} / N_{0}=10 \mathrm{~dB}$, for uncorrelated (unc) and linear (lin) correlation matrices at the transmitter side, and for $K \in$ $\{100,400,1000\}$.
channels, square OSTBCs, and PSK signaling is given by $K_{\mathrm{p}, \mathrm{opt}} \in\{17,35,55\}$ for $K \in\{100,400,1000\}$, respectively (where we have rounded $K_{p, o p t}$ to the nearest integer). From Fig. 7.6, it follows that these approximate optimal values are also useful in case of different fading conditions or transmission parameters.

### 7.5 Chapter Summary

In this chapter, we have investigated the effect of ICSI on the BER performance of OSTBCs under flat-fading channels. For non-square OSTBCs, the resulting exact BER expression can be written as an expectation over $N_{\mathrm{s}}$ discrete RVs and $2 L+3$ real-valued continuous RVs, whereas for square OSTBCs, the resulting exact BER expression reduces to an expectation over $N_{\mathrm{s}}$ discrete RVs and 4 real-valued continuous RVs, regardless of the number of antennas. The exact BER expressions can be efficiently and accurately evaluated by means of
numerical integration techniques, i.e., the quadrature rule and Monte-Carlo integration with importance sampling, or a combination thereof. Additionally, we provided a simple approximation of the BER based on treating the symbol interference due to imperfect channel estimation as white Gaussian noise. Although the resulting expression is in general not asymptotically exact, it yields very accurate BER results when the fading distribution is similar to Rayleigh and when a sufficient number $K_{p}$ of pilot symbols is used. For the case of correlated Nakagami- $m$ fading channels, we elaborated further on the numerical evaluation of the exact BER expressions. Our numerical results illustrate the effect of channel estimation errors and of fading correlation on the BER performance of OSTBCs in the case of Nakagami- $m$ fading.

## 7.A Appendix

## 7.A. 1 Proof of (7.10)

In (7.3), we have defined $u_{i}^{\prime}$ as

$$
\begin{equation*}
u_{i}^{\prime} \triangleq \frac{\operatorname{tr}\left(\mathbf{C}_{i}^{H} \hat{\mathbf{H}}^{H} \mathbf{H C}+\mathbf{C}^{H} \mathbf{H}^{H} \hat{\mathbf{H}} \mathbf{C}_{i}^{\prime}\right)}{\lambda\|\hat{\mathbf{H}}\|_{\mathrm{F}}^{2}} \tag{7.37}
\end{equation*}
$$

With $\mathrm{q}=\mathrm{R}$ and $\mathrm{q}=\mathrm{I}$ referring to the BER computation for the in-phase and quadrature bits, respectively, it can be shown that the real and imaginary parts of (7.37) can be rewritten as

$$
\begin{equation*}
u_{i, \mathrm{q}}^{\prime}=\frac{1}{\lambda\|\hat{\mathbf{g}}\|^{2}} \sum_{n=1}^{N_{\mathrm{s}}}\left\{\left(\hat{\mathbf{g}}^{T} \mathbf{g}_{n, \mathrm{R}}^{(i, \mathbf{q})}\right) s_{n, \mathrm{R}}+\left(\hat{\mathbf{g}}^{T} \mathbf{g}_{n, \mathrm{I}}^{(i, \mathbf{q})}\right) s_{n, \mathrm{I}}\right\} \tag{7.38}
\end{equation*}
$$

where the $2 L \times 1$ column vectors $\mathbf{g}_{n, \mathrm{p}}^{(i, \mathrm{q})}$, with $\mathrm{p}=\mathrm{R}$ or $\mathrm{p}=\mathrm{I}$, are given by

$$
\begin{equation*}
\mathbf{g}_{n, \mathrm{p}}^{(i, \mathrm{q})} \triangleq\left(\mathbf{M}_{n, \mathrm{p}}^{(i, \mathrm{q})} \otimes \mathbf{I}_{L_{\mathrm{r}}}\right) \mathbf{g} \tag{7.39}
\end{equation*}
$$

and the $2 L_{\mathrm{t}} \times 2 L_{\mathrm{t}}$ matrices $\mathbf{M}_{n, \mathrm{p}}^{(i, \mathrm{q})}$ are defined as

$$
\mathbf{M}_{n, \mathrm{p}}^{(i, \mathrm{q})} \triangleq\left[\begin{array}{ccc}
\mathbf{M}_{n, \mathrm{p}}^{(i, \mathrm{q})}(1,1) & \ldots & \mathbf{M}_{n, \mathrm{p}}^{(i, \mathrm{q})}\left(1, L_{\mathrm{t}}\right)  \tag{7.40}\\
\vdots & \ddots & \vdots \\
\mathbf{M}_{n, \mathrm{p}}^{(i, \mathrm{q})}\left(L_{\mathrm{t}}, 1\right) & \ldots & \mathbf{M}_{n, \mathrm{p}}^{(i, \mathrm{q})}\left(L_{\mathrm{t}}, L_{\mathrm{t}}\right)
\end{array}\right]
$$

For all possible combinations of p and q , the $2 \times 2$ matrices $\mathbf{M}_{n, \mathrm{p}}^{(i, \mathrm{q})}(k, \ell)$ in (7.40), with $1 \leq k, \ell \leq L_{\mathrm{t}}$, are defined as follows

$$
\mathbf{M}_{n, \mathrm{R}}^{(i, \mathrm{R})}(k, \ell) \triangleq\left[\begin{array}{cc}
\alpha_{n, \mathrm{R}}^{(i, \mathrm{R})}(k, \ell) & -\beta_{n, \mathrm{R}}^{(i, \mathrm{R})}(k, \ell)  \tag{7.41a}\\
\beta_{n, \mathrm{R}}^{(i, \mathrm{R})}(k, \ell) & \alpha_{n, \mathrm{R}}^{(i, \mathrm{R})}(k, \ell)
\end{array}\right],
$$

with $\alpha_{n, \mathrm{R}}^{(i, \mathrm{R})}(k, \ell)$ and $\beta_{n, \mathrm{R}}^{(i, \mathrm{R})}(k, \ell)$ being the real and imaginary parts of $\left[\left(\mathrm{C}_{n}+\right.\right.$ $\left.\left.\mathbf{C}_{n}^{\prime}\right)\left(\mathbf{C}_{i}+\mathbf{C}_{i}^{\prime}\right)^{H}\right]_{\ell, k}$, respectively;

$$
\mathbf{M}_{n, \mathrm{I}}^{(i, \mathrm{R})}(k, \ell) \triangleq\left[\begin{array}{cc}
-\beta_{n, \mathrm{I}}^{(i, \mathrm{R})}(k, \ell) & -\alpha_{n}^{(i, \mathrm{R})}(k, \ell)  \tag{7.41b}\\
\alpha_{n, \mathrm{I}}^{(i, \mathrm{R})}(k, \ell) & -\beta_{n, \mathrm{I}}^{(i, \mathrm{R})}(k, \ell)
\end{array}\right],
$$

with $\alpha_{n, \mathrm{I}}^{(i, \mathrm{R})}(k, \ell)$ and $\beta_{n, \mathrm{I}}^{(i, \mathrm{R})}(k, \ell)$ being the real and imaginary parts of $\left[\left(\mathbf{C}_{n}-\right.\right.$ $\left.\left.\mathbf{C}_{n}^{\prime}\right)\left(\mathbf{C}_{i}+\mathbf{C}_{i}^{\prime}\right)^{H}\right]_{\ell, k}$, respectively;

$$
\mathbf{M}_{n, \mathrm{R}}^{(i, \mathrm{I})}(k, \ell) \triangleq\left[\begin{array}{cc}
\beta_{n, \mathrm{R}}^{(i, \mathrm{I})}(k, \ell) & \alpha_{n, \mathrm{R}}^{(i, \mathrm{I})}(k, \ell)  \tag{7.41c}\\
-\alpha_{n, \mathrm{R}}^{(i, \mathrm{I}}(k, \ell) & \beta_{n, \mathrm{R}}^{(i, \mathrm{I})}(k, \ell)
\end{array}\right],
$$

with $\alpha_{n, \mathbb{R}}^{(i, \mathrm{I})}(k, \ell)$ and $\beta_{n, \mathrm{R}}^{(i, \mathrm{I})}(k, \ell)$ being the real and imaginary parts of $\left[\left(\mathbf{C}_{n}+\right.\right.$ $\left.\left.\mathbf{C}_{n}^{\prime}\right)\left(\mathbf{C}_{i}-\mathbf{C}_{i}^{\prime}\right)^{H}\right]_{\ell, k}$, respectively;

$$
\mathbf{M}_{n, \mathrm{I}}^{(i, \mathrm{I})}(k, \ell) \triangleq\left[\begin{array}{cc}
\alpha_{n}^{(i, \mathrm{I})}(k, \ell) & -\beta_{n, \mathrm{I}}^{(i, \mathrm{I})}(k, \ell)  \tag{7.41d}\\
\beta_{n, \mathrm{I}}^{(i, \mathrm{I})}(k, \ell) & \alpha_{n, \mathrm{I}}^{(i, \mathrm{I})}(k, \ell)
\end{array}\right]
$$

with $\alpha_{n, \mathrm{I}}^{(i, \mathrm{I})}(k, \ell)$ and $\beta_{n, \mathrm{I}}^{(i, \mathrm{I})}(k, \ell)$ being the real and imaginary parts of $\left[\left(\mathbf{C}_{n}-\right.\right.$ $\left.\left.\mathbf{C}_{n}^{\prime}\right)\left(\mathbf{C}_{i}-\mathbf{C}_{i}^{\prime}\right)^{H}\right]_{\ell, k}$, respectively.

Note that $\alpha_{i, \mathrm{q}}^{(i, \mathrm{q})}(k, \ell)=\lambda \delta_{k-\ell}$ and $\beta_{i, \mathrm{q}}^{(i, \mathrm{q})}(k, \ell)=0$, such that $\mathbf{g}_{i, \mathrm{q}}^{(i, \mathrm{q})}=\lambda \mathbf{g}$. Under the assumption that $\breve{\mathrm{q}}=\mathrm{I}$ or $\breve{\mathrm{q}}=\mathrm{R}$ when $\mathrm{q}=\mathrm{R}$ or $\mathrm{q}=\mathrm{I}$, respectively, it follows from (4.35) that, for any given $i$, the vectors $\mathbf{g}_{n, q}^{(i, q)}$, with $n \neq i$, and $\mathbf{g}_{n, \mathrm{q}}^{(i, \mathrm{q})}$ are orthogonal to $\mathbf{g}_{i, \mathrm{q}}^{(i, \mathrm{q})}$

$$
\begin{gather*}
{\left[\mathbf{g}_{i, \mathrm{q}}^{(i, \mathbf{q})}\right]^{T} \mathbf{g}_{n, \mathrm{q}}^{(i, \mathbf{q})}=\lambda^{2}\|\mathbf{g}\|^{2} \delta_{i-n}}  \tag{7.42a}\\
{\left[\mathbf{g}_{i, \mathrm{q}}^{(i, \mathbf{q})}\right]^{T} \mathbf{g}_{n, \stackrel{\mathrm{q}}{ }}^{(i, \mathbf{q})}=0} \tag{7.42b}
\end{gather*}
$$

In (7.38), the term in $s_{i, \mathrm{q}}$ is the useful term, whereas the terms in $s_{n, \mathrm{q}}$, with $n \neq i$, and the terms in $s_{n, \check{q}}$ represent interference from the symbol components different from $s_{i, q}$. In case of PCSI, i.e., $\hat{\mathbf{g}}=\mathbf{g}$, it is readily verified that $u_{i, \mathrm{q}}^{\prime}=s_{i, \mathrm{q}}$. Let us now introduce the $2 L \times 2 L$ matrix $\mathbf{M}_{i, \mathrm{q}}$ as

$$
\begin{equation*}
\mathbf{M}_{i, \mathrm{q}} \triangleq\left\{\sum_{\substack{n=1 \\ n \neq i}}^{N_{s}} s_{n, \mathrm{q}} \mathbf{M}_{n, \mathrm{q}}^{(i, \mathrm{q})}+\sum_{n=1}^{N_{s}} s_{n, \breve{\mathrm{q}}} \mathbf{M}_{n, \stackrel{\mathrm{q}}{ }}^{(i, \mathrm{q})}\right\} \otimes \mathbf{I}_{L_{\mathrm{r}}} \tag{7.43}
\end{equation*}
$$

which is a function of the transmitted symbol vector $\mathbf{s}$ and the coefficient matrices $\mathbf{C}_{n}$ and $\mathbf{C}_{n}^{\prime}$, with $1 \leq n \leq N_{\mathrm{s}}$. In this way, (7.38) reduces to

$$
\begin{equation*}
u_{i, \mathrm{q}}^{\prime}=\frac{1}{\lambda\|\hat{\mathbf{g}}\|^{2}}\left(\lambda s_{i, \mathrm{q}} \hat{\mathbf{g}}^{T} \mathbf{g}+\hat{\mathbf{g}}^{T} \mathbf{M}_{i, \mathrm{q}} \mathbf{g}\right) \tag{7.44}
\end{equation*}
$$

From (7.39) and (7.42), it follows that $\mathbf{M}_{i, q} \mathbf{g}$ is orthogonal to $\mathbf{g}$. Hence, we can define an orthonormal coordinate system with $2 L$ unit vectors

$$
\begin{equation*}
\left\{\mathbf{e}_{n}^{(i, q)}, n=1,2, \ldots, 2 L\right\} \tag{7.45}
\end{equation*}
$$

where $\mathbf{e}_{1}^{(i, q)}$ and $\mathbf{e}_{2}^{(i, q)}$ are directed along $\mathbf{g}$ and $\mathbf{M}_{i, \mathrm{q}} \mathbf{g}$, respectively. Let us introduce the RVs $x_{n}^{(i, q)}$ as the projections of $\hat{\mathbf{g}}$ on $\mathbf{e}_{n}^{(i, q)}$

$$
\begin{equation*}
x_{n}^{(i, \mathrm{q})} \triangleq \hat{\mathbf{g}}^{T} \mathbf{e}_{n}^{(i, \mathrm{q})} \tag{7.46}
\end{equation*}
$$

Assuming ML channel estimation, it follows from (5.13) that the statistical properties of $x_{n}^{(i, \mathrm{q})}$ are independent of $i$ and q. Therefore, we may drop the superscript $(i, q)$, resulting in

$$
\begin{equation*}
\|\hat{\mathbf{g}}\|^{2}=x_{1}^{2}+x_{2}^{2}+z^{2} \tag{7.47}
\end{equation*}
$$

with

$$
\begin{equation*}
z=\sqrt{\sum_{n=3}^{2 L} x_{n}^{2}} \tag{7.48}
\end{equation*}
$$

Taking the specific choice of $\mathbf{e}_{1}^{(i, q)}$ and $\mathbf{e}_{2}^{(i, q)}$ into account, it follows from (7.46)(7.48) that (7.44) reduces to

$$
\begin{equation*}
u_{i, \mathrm{q}}^{\prime}=\frac{1}{\lambda\left(x_{1}^{2}+x_{2}^{2}+z^{2}\right)}\left(\lambda x_{1} s_{i, \mathrm{q}}\|\mathbf{g}\|+x_{2}\left\|\mathbf{M}_{i, \mathrm{q}} \mathbf{g}\right\|\right) . \tag{7.49}
\end{equation*}
$$

## 7.A. 2 Proof of (7.12)

For square OSTBCs, i.e., $L_{\mathrm{t}}=K_{\mathrm{c}}$, it follows from (4.36) that the vectors $\mathbf{g}_{n, \mathrm{q}}^{(i, \mathbf{q})}$ and $\mathbf{g}_{n, \tilde{q}}^{(i, q)}$, given by (7.39), are mutually orthogonal, such that the orthogonality conditions given by (7.42) reduce to

$$
\begin{gather*}
{\left[\mathbf{g}_{n, \mathrm{p}}^{(i, \mathbf{q})}\right]^{T} \mathbf{g}_{n^{\prime}, \mathrm{p}}^{(i, \mathbf{q})}=\lambda^{2}\|\mathbf{g}\|^{2} \delta_{n-n^{\prime}}}  \tag{7.50a}\\
{\left[\mathbf{g}_{n, \mathbf{q}}^{(i, \mathbf{q})}\right]^{T} \mathbf{g}_{n^{\prime}, \stackrel{\mathrm{q}}{(i, q}}^{(i,}=0} \tag{7.50b}
\end{gather*}
$$

Using the above properties, it can be shown that $\left\|\mathbf{M}_{i, \mathrm{q}} \mathbf{g}\right\|$ which appears in (7.10), with $\mathbf{M}_{i, q}$ given by (7.43), reduces to

$$
\begin{equation*}
\left\|\mathbf{M}_{i, \mathbf{q}} \mathbf{g}\right\|=\lambda\|\mathbf{g}\| \sqrt{\|\mathbf{s}\|^{2}-s_{i, \mathbf{q}}^{2}} \tag{7.51}
\end{equation*}
$$

## 7.A. 3 Derivation of Biased PDF $q(\mathbf{y})$

Since the random vectors $\mathbf{y}_{k}$ consist of real-valued ZM Gaussian RVs with covariance matrix $\mathbf{Q}$, their joint PDF is given by

$$
\begin{equation*}
p\left(\mathbf{y}_{k}\right)=\frac{1}{(2 \pi)^{\frac{L}{2}} \sqrt{\operatorname{det}(\mathbf{Q})}} \exp \left(-\frac{1}{2} \mathbf{y}_{k}^{T} \mathbf{Q}^{-1} \mathbf{y}_{k}\right) \tag{7.52}
\end{equation*}
$$

where $\operatorname{det}(\mathbf{Q})$ denotes the determinant of $\mathbf{Q}$. Because the $2 m$ vectors $\mathbf{y}_{k}$ are i.i.d., the joint PDF of the vector $\mathbf{y}=\left[\mathbf{y}_{1}^{T}, \mathbf{y}_{2}^{T}, \ldots, \mathbf{y}_{2 m}^{T}\right]^{T}$ is given by

$$
\begin{align*}
p(\mathbf{y}) & =\prod_{k=1}^{2 m} p\left(\mathbf{y}_{k}\right) \\
& \propto \prod_{k=1}^{2 m} \exp \left(-\frac{1}{2} \mathbf{y}_{k}^{T} \mathbf{Q}^{-1} \mathbf{y}_{k}\right) \tag{7.53}
\end{align*}
$$

Taking into account that $\|\mathbf{y}\|_{\mathrm{F}}^{2}=\sum_{k=1}^{2 m}\left\|\mathbf{y}_{k}\right\|_{\mathrm{F}}^{2}$, it follows from (7.29) that the biased PDF $q(\mathbf{y})$ is given by

$$
\begin{align*}
q(\mathbf{y}) & \propto \prod_{k=1}^{2 m} \exp \left(-\frac{1}{2} \mathbf{y}_{k}^{T} \mathbf{Q}^{-1} \mathbf{y}_{k}\right) \exp \left(-\lambda d^{2} \frac{E_{\mathrm{s}}}{N_{0}} \mathbf{y}_{k}^{T} \mathbf{y}_{k}\right) \\
& =\prod_{k=1}^{2 m} \exp \left[-\frac{1}{2} \mathbf{y}_{k}^{T}\left(\mathbf{Q}^{-1}+2 \lambda d^{2} \frac{E_{\mathrm{s}}}{N_{0}} \mathbf{I}_{L}\right) \mathbf{y}_{k}\right] \tag{7.54}
\end{align*}
$$

Hence, it follows from the biased PDF (7.54) that the elements of the i.i.d. auxiliary random vectors $\mathbf{y}_{k}$ should be generated as correlated ZM Gaussian RVs with a covariance matrix $\mathbf{Q}^{\prime}$ given by

$$
\begin{equation*}
\mathbf{Q}^{\prime}=\left(\mathbf{Q}^{-1}+2 \lambda d^{2} \frac{E_{\mathrm{S}}}{N_{0}} \mathbf{I}_{L}\right)^{-1} \tag{7.55}
\end{equation*}
$$

Taking the PDFs $p\left(\mathbf{y}_{k}\right)$ and $q\left(\mathbf{y}_{k}\right)$ of the $2 m$ i.i.d. random vectors $\mathbf{y}_{k}$ into account, it follows that the ratio $p(\mathbf{v}) / q(\mathbf{v})$ in (7.19) depends only on $\mathbf{y}$ and is given by

$$
\begin{equation*}
\frac{p(\mathbf{v})}{q(\mathbf{v})}=\frac{\exp \left(\lambda d^{2} \frac{E_{\mathrm{s}}}{N_{0}}\|\mathbf{y}\|^{2}\right)}{\left[\operatorname{det}\left(\mathbf{I}_{L}+2 \lambda d^{2} \frac{E_{\mathrm{s}}}{N_{0}} \mathbf{Q}\right)\right]^{m}} \tag{7.56}
\end{equation*}
$$

## 8

## Food for Future Thought

In this chapter, we illustrate how the expressions and techniques derived in this dissertation can be applied to assess the performance of MIMO OSTBC systems using different or additional transmission or modulation techniques. In section 8.1, we show how the PEP of wireless systems combining orthogonal space-time block coding with channel coding can be obtained using the techniques described in chapter 7. Preliminary performance results of OSTBC systems employing adaptive modulation and coding (AMC) on Rayleigh fading channels are presented in section 8.2. In section 8.3, we briefly touch upon the performance of OSTBC systems using orthogonal frequency division multiplexing (OFDM). A summary of this chapter is provided in section 8.4.

### 8.1 Channel Coding

In uncoded communication, the information bit sequence is modulated onto a carrier wave, transmitted over the channel, and detected by the receiver. In this way, a detection error due to ICSI or channel noise always results in one
or more bit errors. In order to provide additional protection against bit errors, wireless communication systems often use channel coding, which adds controlled redundancy to the information bit sequence. More specifically, each $N_{\mathrm{b}}$-bit information sequence $\mathbf{b}$ is encoded into a unique $N_{\mathrm{c}}$-bit coded bit sequence c, with $N_{\mathrm{c}}>N_{\mathrm{b}}$, before it is mapped to a symbol sequence s. The redundancy in the input bit stream allows the decoder to detect and correct a number of bit errors in the received data stream.

The PEP is defined as the probability that the likelihood of a codeword c which differs from $\mathbf{c}_{0}$ is higher than the likelihood of $\mathbf{c}_{0}$. Let us assume that the symbol vectors $\mathbf{s}_{0}$ and $\mathbf{s}$ corresponding to the codewords $\mathbf{c}_{0}$ and $\mathbf{c}$, respectively, consist of $K$ subvectors $\mathbf{s}_{0, k}$ and $\mathbf{s}_{k}$, with $k=1, \ldots, K$, each consisting of $N_{\mathrm{s}}$ symbols $\mathbf{s}_{0, k, i}$ and $\mathbf{s}_{k, i}$, with $i=1, \ldots, N_{\mathrm{s}}$, respectively. In this way, the subvectors $\mathbf{s}_{0, k}$ and $\mathbf{s}_{k}$ are transformed into the OSTBC matrices $\mathbf{C}_{0, k}$ and $\mathbf{C}_{k}$, respectively. Assuming that all OSTBC matrices are sent within the same data frame $\mathbf{A}_{0} \triangleq\left[\mathbf{C}_{0,1}, \mathbf{C}_{0,2}, \ldots, \mathbf{C}_{0, K}\right]$ and are affected by the same channel matrix $\mathbf{H}$, the received signal is given by

$$
\begin{equation*}
\mathbf{R}=\sqrt{E_{\mathrm{s}}} \mathbf{H} \mathbf{A}_{0}+\mathbf{W}_{0} \tag{8.1}
\end{equation*}
$$

and the PEP in case of imperfect channel estimation is given by

$$
\begin{equation*}
\operatorname{PEP}\left(\mathbf{s}_{0}, \mathbf{s}\right)=\operatorname{Pr}\left[\left\|\mathbf{R}-\sqrt{\mathbf{E}_{s}} \hat{\mathbf{H}} \mathbf{A}\right\|^{2}<\left\|\mathbf{R}-\sqrt{\mathbf{E}_{s}} \hat{\mathbf{H}} \mathbf{A}_{0}\right\|^{2}\right] \tag{8.2}
\end{equation*}
$$

where $\mathbf{A} \triangleq\left[\mathbf{C}_{1}, \mathbf{C}_{2}, \ldots, \mathbf{C}_{K}\right]$. Using a similar coordinate transformation as described in appendix 7.A.1, it can be shown that the conditional PEP reduces to

$$
\begin{array}{r}
\operatorname{PEP}\left(\mathbf{s}_{0}, \mathbf{s} \mid x_{1}, x_{2}, z, \mathbf{g}\right)=\operatorname{Pr}\left[w>\lambda \sqrt{E_{\mathbf{s}}}\left(x_{1}^{2}+x_{2}^{2}+z^{2}\right)\left(\|\mathbf{s}\|^{2}-\left\|\mathbf{s}_{0}\right\|^{2}\right)\right. \\
\left.-2 \sqrt{E_{\mathrm{s}}}\left\{\lambda x_{1} \Re\left[\mathbf{s}_{0}^{H}\left(\mathbf{s}-\mathbf{s}_{0}\right)\right]\|\mathbf{g}\|+x_{2} \sum_{k=1}^{K} \sum_{i=1}^{N_{\mathbf{s}}}\left\|\mathbf{M}_{k, i} \mathbf{g}\right\|\right\}\right] \tag{8.3}
\end{array}
$$

where the vector $\mathbf{g}$ and the $\operatorname{RVs} x_{1}, x_{2}$, and $z$ are introduced in section $7.2, w$ is a zero-mean Gaussian RV with variance

$$
\begin{equation*}
\sigma^{2}=2 \lambda N_{0}\left(x_{1}^{2}+x_{2}^{2}+z^{2}\right)\left\|\mathbf{s}-\mathbf{s}_{0}\right\|^{2} \tag{8.4}
\end{equation*}
$$

and $\mathbf{M}_{k, i}$ is given by

$$
\begin{align*}
& \mathbf{M}_{k, i}=\left(s_{k, i, \mathrm{R}}-s_{0, k, i, \mathrm{R}}\right)\left[\sum_{\substack{n=1 \\
n \neq i}}^{N_{s}} s_{0, k, n, \mathrm{R}} \mathbf{M}_{n, \mathrm{R}}^{(i, \mathrm{R})}+\sum_{n=1}^{N_{s}} s_{0, k, n, \mathrm{I}} \mathbf{M}_{n, \mathrm{I}}^{(i, \mathrm{R})}\right] \otimes \mathbf{I}_{L_{\mathrm{r}}} \\
& +\left(s_{k, i, \mathrm{I}}-s_{0, k, i, \mathrm{I}}\right)\left[\sum_{n=1}^{N_{s}} s_{0, k, n, \mathrm{R}} \mathbf{M}_{n, \mathrm{R}}^{(i, \mathrm{I})}+\sum_{\substack{n=1 \\
n \neq i}}^{N_{s}} s_{0, k, n, \mathrm{I}} \mathbf{M}_{n, \mathrm{I}}^{(i, \mathrm{I})}\right] \otimes \mathbf{I}_{L_{\mathrm{r}}}, \tag{8.5}
\end{align*}
$$

where $s_{0, k, i}=s_{0, k, i, \mathrm{R}}+j s_{0, k, i, \mathrm{I}}$ and $s_{k, i}=s_{k, i, \mathrm{R}}+j s_{k, i, \mathrm{I}}$, and the $2 L_{\mathrm{t}} \times 2 L_{\mathrm{t}}$ matrices $\mathbf{M}_{n, \mathrm{R}}^{(i, \mathrm{R})}, \mathbf{M}_{n, \mathrm{I}}^{(i, \mathrm{I})}, \mathbf{M}_{n, \mathrm{R}}^{(i, \mathrm{I})}$ and $\mathbf{M}_{n, \mathrm{I}}^{(i, \mathrm{I})}$ are obtained from (7.40) and (7.41). From the conditional PEP (8.3), it follows that the PEP is given by averaging the following $Q$-function over $g, x_{1}, x_{2}$ and $z$ :

$$
\begin{align*}
& \operatorname{PEP}\left(\mathbf{s}_{0}, \mathbf{s}\right)=\mathbb{E}\left[Q \left(\sqrt{\frac{E_{\mathrm{s}}}{2 \lambda N_{0}\left(x_{1}^{2}+x_{2}^{2}+z^{2}\right)\left\|\mathbf{s}-\mathbf{s}_{0}\right\|^{2}}}\right.\right. \\
& {\left[\lambda\left(x_{1}^{2}+x_{2}^{2}+z^{2}\right)\left(\|\mathbf{s}\|^{2}-\left\|\mathbf{s}_{0}\right\|^{2}\right)-\right.} 2 \lambda x_{1} \Re\left[\mathbf{s}_{0}^{H}\left(\mathbf{s}-\mathbf{s}_{0}\right)\right]\|\mathbf{g}\| \\
&\left.\left.\left.+2 x_{2} \sum_{k=1}^{K} \sum_{i=1}^{N_{s}}\left\|\mathbf{M}_{k, i} \mathbf{g}\right\|\right]\right)\right] \tag{8.6}
\end{align*}
$$

Unfortunately, unlike the BER computation in chapter 7, the number of RVs in the expectation (8.6) cannot be further decreased for square OSTBCs, since $\left\|\mathbf{M}_{k, i} \mathbf{g}\right\|$ in (8.6) does not reduce to a function of $\|\mathbf{g}\|$ only. The reason for this is that the orthogonality conditions in (7.50) cannot be extended with

$$
\begin{equation*}
\left[\mathbf{g}_{n, \mathrm{q}}^{(i, \mathrm{q})}\right]^{T} \mathbf{g}_{n^{\prime}, \mathrm{p}}^{(i, \breve{\mathrm{q}})}=0 \tag{8.7}
\end{equation*}
$$

where $q$ and p may refer to $R$ or I.

### 8.2 Adaptive Modulation and Coding

In spite of the outstanding average BER performance achieved by OSTBCs, their instantaneous performance is still poor when the channel is in a deep fade. To tackle this problem, we consider the technique of AMC [77], which guarantees a given performance level by adapting one or more transmission parameters, e.g., the constellation size, the coding rate or the transmit power, based on information about the current channel state. In many AMC schemes, the receiver estimates the CSI, updates the transmission parameters accordingly, and sends them to the transmitter via a low-rate feedback channel. Because of the latency introduced by the feedback channel, however, the channel may have changed at the time the updated transmission parameters are applied by the transmitter, a phenomenon which is called outdated feedback. Therefore, in case of rapid fading, channel prediction is sometimes used to avoid this.

In [78], closed-form expressions for the average BER and SE are derived for a rate-adaptive $M$-QAM system employing orthogonal space-time block coding with outdated, finite-rate feedback over i.i.d. flat Rayleigh fading channels. Closed-form expressions for the average SE of two rate-adaptive MIMO schemes, i.e., OSTBCs and spatial multiplexing with zero-forcing receiver
were derived in [79] for i.i.d. Rayleigh fading channels with ICSI. In [80], a similar analysis was given for the case of spatially correlated Rayleigh fading channels with transmit correlation and PCSI. Closed-form expressions of the average BER, SE and outage probability of rate-adaptive OSTBC systems were given in [81] for spatially correlated Rayleigh channels with and without PCSI. However, in [81] the instantaneous BER is approximated by an exponential function, which simplifies the analysis significantly but can be not sufficiently accurate for the performance evaluation in fading channels.

Based on the results from chapter 6, we derive a novel and accurate closedform expression for the average BER of a rate-adaptive MIMO OSTBC system with ICSI in arbitrarily correlated Rayleigh fading channels. For the case of i.i.d. Rayleigh fading channels, an exact closed-form expression is provided. In addition, we provide a simple approximate closed-form expression for the average BER, which yields more accurate results than the approximate expression from [81]. To guarantee a fair comparison between different channel estimation scenarios, the performance is compared given a fixed total energy for both data and pilot symbols, and the achieved spectral efficiency (ASE) is considered rather than the SE. Our analysis enables system designers to choose appropriate system parameters considering the trade-off between SE and outage for a given target BER.

### 8.2.1 Adaptive Modulation Schemes

Let us consider a rate-adaptive MIMO OSTBC system, where the receiver selects a constellation size $M_{j}$ to be used by the transmitter from a finite set of candidates $\mathcal{M}=\left\{M_{0}, M_{1}, \ldots, M_{J}\right\}$ with Gray mapping, and sends this information back to the transmitter over a perfect feedback channel without delay. When fast adaptive modulation (FAM) is applied, the constellation size $M_{j}$ is chosen depending on the value of the estimated instantaneous SNR

$$
\begin{equation*}
\hat{\gamma}=\|\hat{\mathbf{h}}\|^{2} \frac{E_{\mathrm{S}}}{N_{0}} . \tag{8.8}
\end{equation*}
$$

Another approach consists in adapting $M_{j}$ based on tracking of large-scale fading. Analysis of fast and slow adaptive modulation (SAM) with diversity is given in [82]. Here, FAM is applied and the modulation level $M_{j}$ is selected when $\hat{\gamma}$ falls in the interval $\left[\hat{\gamma}_{j}^{\star}, \hat{\gamma}_{j+1}^{\star}\right)$, where $\hat{\gamma}_{J+1}^{\star}=\infty$ and the switching thresholds $\hat{\gamma}_{j}^{\star}, j=0, \ldots, J$, are chosen to provide a given instantaneous target BER $P_{b}^{\star}$

$$
\begin{equation*}
P_{\mathrm{b}, M_{j}}\left(\hat{\gamma}_{j}^{\star}\right)=P_{\mathrm{b}}^{\star} \tag{8.9}
\end{equation*}
$$

where $P_{\mathrm{b}, M_{j}}(\hat{\gamma})$ denotes the BER for $M_{j}$-QAM with imperfect channel estimation as a function of the estimated instantaneous SNR $\hat{\gamma}$. The instantaneous BER $P_{\mathrm{b}, M_{j}}(\hat{\gamma})$ is shown in Fig. 8.1 for $M_{j}$-QAM, with $M_{j} \in\{4,16,64\}$. When


Figure 8.1: Instantaneous BER versus $\hat{\gamma}$ in case of ICSI and $M_{j}$-QAM, with $M_{j} \in\{4,16,64\}$. The thick line shows the instantaneous BER when FAM with target BER $P_{\mathrm{b}}^{\star}=10^{-6}$ is applied.

FAM is applied, the modulation level $M_{j}$ is chosen such that the instantaneous BER is always lower than the target BER. The thick line in Fig. 8.1 shows the instantaneous BER for a target BER of $P_{\mathrm{b}}^{\star}=10^{-6}$. When even the smallest constellation size $M_{0}$ does not meet the target BER, i.e., when $\hat{\gamma}<\hat{\gamma}_{0}^{\star}$, no data are transmitted and the system is in outage. The general behavior of the performance of multidimensional constellation signaling systems is analyzed in [83].

### 8.2.2 Performance Evaluation

In chapters 6 and 7, the performance curves of fixed-rate systems with and without PCSI were plotted as a function of $E_{\mathrm{b}} / N_{0}$, with $E_{\mathrm{b}}$ denoting the energy per information bit. However, since the relation (5.8) between $E_{\mathrm{s}}$ and $E_{\mathrm{b}}$ holds for a specific constellation only, $E_{\mathrm{b}}$ cannot be used as reference energy in case of rate-adaptive systems that switch between different constellation sizes. To this end, we introduce $E_{\mathrm{d}}$ as the average total energy required to transmit one information symbol. Because a data frame reserving $K_{p}$ time slots for
pilot symbols and $K$ time slots for coded data symbols contains $N_{\mathrm{s}} K / K_{\mathrm{c}}$ information symbols, the average total energy devoted to the transmission of one data frame is given by $E_{\text {tot }} \triangleq\left(N_{\mathrm{s}} K / K_{\mathrm{c}}\right) E_{\mathrm{d}}$. As indicated in section 5.1, very accurate channel estimation can be obtained by allocating a large fraction of $E_{\text {tot }}$ to pilot symbols. However, the more energy is devoted to pilot symbols, the less energy remains available for data transmission, calling for a trade-off between resources dedicated to pilot and data symbols for a set of target performance metrics $[60,84]$. The relation between $E_{\mathrm{s}}$ and $E_{\mathrm{d}}$ is given by

$$
\begin{equation*}
E_{\mathrm{s}}=\frac{K}{K+\eta K_{\mathrm{p}}} \rho E_{\mathrm{d}} \tag{8.10}
\end{equation*}
$$

where $\eta \triangleq E_{\mathrm{p}} / E_{\mathrm{s}}$ and $\rho \triangleq N_{\mathrm{s}} /\left(L_{\mathrm{t}} K_{\mathrm{c}}\right)$.
Based on the BER expressions obtained in section 6 for fixed-rate systems under Rayleigh fading channels, we derive closed-form expressions for the resulting average BER, ASE, and bit error outage (BEO) of a rate-adaptive MIMO OSTBC system with LMMSE channel estimation. For square OSTBCs and M-QAM transmission, it follows from sections 6.2.1 and 6.3.1 that the instantaneous conditional BER $P_{\mathrm{b}, j}(\hat{\gamma})$, conditioned on the estimated instantaneous $\operatorname{SNR} \hat{\gamma}$, is given by

$$
\begin{equation*}
P_{\mathrm{b}, M}(\hat{\gamma})=\left(\frac{4}{M}\right)^{N_{\mathrm{s}}} \frac{1}{\log _{2}(\sqrt{M})} \sum_{\mathbf{s} \in \Psi_{0}^{N_{\mathrm{s}}}} \sum_{b_{\mathrm{q}} \in \Psi^{\prime}} d_{\mathrm{H}}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right) \operatorname{Pr}\left[\hat{s}_{i, \mathrm{q}}=b_{\mathrm{q}} \mid \hat{\gamma}, \mathbf{s}\right] \tag{8.11}
\end{equation*}
$$

with $i \in\left\{1, \ldots, N_{\mathrm{s}}\right\}$ and q being given by R or I. In (8.11), $\operatorname{Pr}\left[\hat{s}_{i, \mathrm{q}}=b_{\mathrm{q}} \mid \hat{\gamma}, \mathbf{s}\right]$ is given by

$$
\begin{equation*}
\operatorname{Pr}\left[\hat{s}_{i, \mathrm{q}}=b_{\mathrm{q}} \mid \hat{\gamma}, \mathbf{s}\right]=Q\left(\sqrt{\frac{D_{1}^{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)}{\sigma_{i, \mathrm{q}}^{2}(\hat{\gamma}, \mathbf{s})}}\right)-Q\left(\sqrt{\frac{D_{2}^{2}\left(s_{i, \mathrm{q}}, b_{\mathrm{q}}\right)}{\sigma_{i, \mathrm{q}}^{2}(\hat{\gamma}, \mathbf{s})}}\right) . \tag{8.12}
\end{equation*}
$$

In case of i.i.d. Rayleigh fading, it follows from (6.34) that $\sigma_{i, \mathrm{q}}^{2}(\hat{\gamma}, \mathbf{s})$ is given by

$$
\begin{equation*}
\sigma_{i, \mathrm{q}}^{2}(\hat{\gamma}, \mathbf{s})=\frac{1}{2 \lambda \hat{\gamma}}\left(1+\frac{\lambda\|\mathbf{s}\|^{2}}{\frac{N_{0}}{E_{\mathrm{s}}}+\eta K_{\mathrm{p}}}\right) \tag{8.13}
\end{equation*}
$$

whereas it follows from (6.67) that for arbitrarily correlated Rayleigh fading and $\eta K_{\mathrm{p}} \frac{E_{\mathrm{s}}}{N_{0}} \gg 1, \sigma_{i, \mathrm{q}}^{2}(\hat{\gamma}, \mathbf{s})$ is approximately given by

$$
\begin{equation*}
\sigma_{i, \mathrm{q}}^{2}(\hat{\gamma}, \mathbf{s}) \approx \frac{1}{2 \lambda \hat{\gamma}}\left(1+\frac{\lambda\|\mathbf{s}\|^{2}}{\eta K_{\mathrm{p}}}\right) . \tag{8.14}
\end{equation*}
$$

Note that for non-square OSTBCs, a similar analysis can be provided using the BER expressions resulting from sections 6.2.2 and 6.3.2. For a given target

BER, the switching thresholds can be computed off-line by substituting (8.11) into (8.9) and numerically solving the resulting equation.

Since the constellation size is selected based on the estimated instantaneous SNR $\hat{\gamma}$, the average BER is given by

$$
\begin{equation*}
\overline{\mathrm{BER}}=\frac{\sum_{j=0}^{J} R_{j} \int_{\hat{\gamma}_{j}^{\widehat{\gamma}}}^{\hat{\gamma}_{j+1}^{\star}} P_{\mathrm{b}, M_{j}}(x) p_{\hat{\gamma}}(x) \mathrm{d} x}{\sum_{j=0}^{J} R_{j} \int_{\hat{\gamma}_{j}^{\star}}^{\hat{\gamma}_{j+1}^{\star}} p_{\hat{\gamma}}(x) \mathrm{d} x} . \tag{8.15}
\end{equation*}
$$

where $R_{j} \triangleq \log _{2}\left(M_{j}\right)$ and $p_{\hat{\gamma}}(x)$ is the PDF of $\hat{\gamma}$. The ASE is defined as $\mu$ times the SE, where $\mu \triangleq K /\left(K+K_{\mathrm{p}}\right)$ represents the fraction of the resources that is used for the transmission of the data-dependent portion of the frame [85]. Hence, the average ASE (in bits/s/Hz) is obtained as

$$
\begin{equation*}
\overline{\mathrm{ASE}}=\frac{K}{K+K_{\mathrm{p}}} \frac{N_{\mathrm{s}}}{K_{\mathrm{c}}} \sum_{j=0}^{J} R_{j} \int_{\hat{\gamma}_{j}^{\star}}^{\hat{\gamma}_{j+1}^{\star}} p_{\hat{\gamma}}(x) \mathrm{d} x . \tag{8.16}
\end{equation*}
$$

Since the BEO is given by the probability that the BER corresponding to the smallest constellation size $M_{0}$ exceeds a target value $P_{\mathrm{b}}^{\star}$ [81,86, 87], it results in

$$
\begin{equation*}
\mathrm{BEO}=\int_{0}^{\hat{\gamma}_{0}^{\star}} p_{\hat{\gamma}}(x) \mathrm{d} x . \tag{8.17}
\end{equation*}
$$

It is easily derived from (4.69) that the $\operatorname{PDF} p_{\hat{\gamma}}(x)$ of the estimated instantaneous SNR is given by

$$
\begin{equation*}
p_{\hat{\gamma}}(x)=\sum_{m=1}^{\kappa} \sum_{n=1}^{c_{m}} \frac{D_{m, n}}{(n-1)!\left(\lambda_{m} \frac{E_{\mathrm{s}}}{N_{0}}\right)^{n}} x^{n-1} \exp \left(-\frac{x}{\lambda_{m} \frac{E_{\mathrm{s}}}{N_{0}}}\right), \quad x \geq 0 \tag{8.18}
\end{equation*}
$$

where $\lambda_{m}, m=1,2, \ldots, \kappa$, denotes the $m$-th distinct eigenvalue of the covariance matrix (5.18) of $\hat{\mathbf{h}}$, with corresponding algebraic multiplicity $c_{m}$, and the parameters $D_{m, n}$ are given by (4.70).

In order to obtain the numerator of (8.15), we derive the following closed-
form solution

$$
\begin{align*}
& \frac{1}{c^{n}(L-1)!} \int_{a}^{b} Q(\sqrt{\beta y}) y^{L-1} \exp \left(-\frac{y}{c}\right) \mathrm{d} y \\
& =Q(\sqrt{\beta a}) \exp \left(-\frac{a}{c}\right) \sum_{k=0}^{L-1} \frac{1}{k!}\left(\frac{a}{c}\right)^{k}-Q(\sqrt{\beta b}) \exp \left(-\frac{b}{c}\right) \sum_{k=0}^{L-1} \frac{1}{k!}\left(\frac{b}{c}\right)^{k} \\
& \quad-\sqrt{\frac{\beta c}{2+\beta c}}\left(Q\left(t_{1}\right)-Q\left(t_{2}\right)\right) \sum_{k=0}^{L-1} \frac{1}{2^{k}}\binom{2 k}{k} \frac{1}{(2+\beta c)^{k}} \\
& \quad-\frac{1}{\sqrt{2 \pi}} \sqrt{\frac{\beta c}{2+\beta c}}\left[\sum_{k=1}^{L-1} \frac{1}{2^{k}}\binom{2 k}{k} \frac{1}{(2+\beta c)^{k}}\right. \\
& \left.\quad \times \sum_{i=1}^{k} 2^{i-1} \frac{(i-1)!}{(2 i-1)!}\left(\exp \left(-\frac{t_{1}^{2}}{2}\right) t_{1}^{2 i-1}-\exp \left(-\frac{t_{2}^{2}}{2}\right)^{2 i} t_{2}^{2 i-1}\right)\right] . \tag{8.19}
\end{align*}
$$

where $t_{1}$ and $t_{2}$ are defined as

$$
\begin{align*}
& t_{1} \triangleq \sqrt{\beta a+2 a / c}  \tag{8.20a}\\
& t_{2} \triangleq \sqrt{\beta b+2 b / c} \tag{8.20b}
\end{align*}
$$

In this way, a neat closed-form expression for the integral in the numerator of (8.15) can be obtained from (8.11)-(8.14) and (8.18)-(8.19). Using (8.18), closed-form expressions for the integrals in the denominator of (8.15) and in (8.16)-(8.17) are easily obtained as

$$
\begin{gather*}
\int_{\hat{\gamma}_{j}^{\star}}^{\hat{\gamma}_{j+1}^{\star}} p_{\hat{\gamma}}(x) \mathrm{d} x=\sum_{m=1}^{\kappa} \sum_{n=1}^{c_{m}} \frac{D_{m, n}}{(n-1)!}\left[\gamma\left(n, \frac{\hat{\gamma}_{j+1}^{\star}}{\lambda_{m} \frac{E_{\mathrm{s}}}{N_{0}}}\right)-\gamma\left(n, \frac{\hat{\gamma}_{j}^{\star}}{\lambda_{m} \frac{E_{\mathrm{s}}}{N_{0}}}\right)\right],  \tag{8.21}\\
\int_{0}^{\hat{\gamma}_{0}^{\star}} p_{\hat{\gamma}}(x) \mathrm{d} x=\sum_{m=1}^{\kappa} \sum_{n=1}^{c_{m}} \frac{D_{m, n}}{(n-1)!} \gamma\left(n, \frac{\hat{\gamma}_{0}^{\star}}{\lambda_{m} \frac{E_{s}}{N_{0}}}\right), \tag{8.22}
\end{gather*}
$$

where $\gamma(s, x)$ denotes the lower incomplete gamma function, which is defined as

$$
\begin{equation*}
\gamma(s, x) \triangleq \int_{0}^{x} t^{s-1} \exp (-t) \mathrm{d} t \tag{8.23}
\end{equation*}
$$

The complexity of the resulting closed-form expression for the average BER can be further reduced by replacing $\|\boldsymbol{s}\|^{2}$ in (8.13) or (8.14) by its expectation $\mathbb{E}\left[\|\mathbf{s}\|^{2}\right]=N_{\mathrm{s}}$. In this way, the summation over the data symbol vector $\mathbf{s}$ in (8.11) reduces to a summation over $s_{i, R}$, which reduces the computational complexity of both the instantaneous and average BER considerably. The impact of this approximation on the accuracy of the resulting average BER expression is illustrated in section 8.2.3.


Figure 8.2: Average BER of Alamouti's code under correlated Rayleigh fading, for both a target BER of $P_{\mathrm{b}}^{\star}=10^{-3}$ and $P_{\mathrm{b}}^{\star}=10^{-4}$.

### 8.2.3 Numerical results

We report numerical results for a rate-adaptive MIMO OSTBC system using Alamouti's code (4.14). We assume that $E_{\mathrm{p}}=E_{\mathrm{s}}$, and that the constellation set is given by $\mathcal{M}=\{4,16,64\}$. To maximize the ASE for a given target BER, the number of pilot symbols $K_{\mathrm{p}}$ is optimized with respect to the number of coded data symbols $K$. With larger $K_{p}$, more accurate CSI can be obtained, which reduces the BEO and enables larger constellations to meet the target BER. However, increasing $K_{\mathrm{p}}$ also reduces $E_{\mathrm{s}}$ and the ASE, according to (8.16), because more resources are wasted on pilot symbols. Hence, $K_{p}$ needs to be carefully selected. We choose $K=20$ and $K_{\mathrm{p}}=4$, which can be shown to be a good trade-off between BEO and SE.

In section 8.2.2, we have shown how the average BER can be derived from the instantaneous BER and the PDF of the estimated instantaneous SNR $\hat{\gamma}$. In Fig. 8.2, several analytical average BER curves are presented, corresponding to different approximations of the instantaneous BER $P_{\mathrm{b}, M}(\hat{\gamma})$ : (a) the closed-form expression (8.11) using the approximation in (8.14); (b) the approximation of (8.11) discussed in the last paragraph of section 8.2.2; and (c) the exponential approximation of $P_{\mathrm{b}, M}(\hat{\gamma})$, as used in [81]. The dots in the
figure represent brute-force simulation results, which show a good agreement with the analytical curves. We consider a mismatched receiver with one antenna ( $L_{\mathrm{r}}=1$ ) and assume a covariance matrix $\boldsymbol{\mathcal { R }}$ given by

$$
\mathcal{R}=\left(\begin{array}{cc}
1 & 0.6  \tag{8.24}\\
0.6 & 1
\end{array}\right)
$$

The results are shown for both a target BER of $P_{\mathrm{b}}^{\star}=10^{-3}$ and $P_{\mathrm{b}}^{\star}=10^{-4}$. From the brute-force simulation results, it follows that the presented average BER (a) resulting from (8.11) and (8.14) is very accurate for moderate to high average SNR, whereas the approximations of the instantaneous BER used to obtain (b) and (c) cause a shift of the resulting average BER curves. For a target BER of $P_{\mathrm{b}}^{\star}=10^{-4}$, the average BER curve from (b) turns out to be more accurate than the BER (c) from [81], while both expressions have a similar computational complexity. For low average SNR, the average BER curves from (a), (b), and (c) diverge from the simulations because of the high-SNR approximations used to obtain (8.14) or (8.11). In case of a rate-adaptive system, however, the low SNR region is not of particular interest, as the BEO is very high and, consequently, the resulting average ASE very low. Note that the average BER can be much lower than the given target BER. This is due to the fact that the considered AMC scheme guarantees that the instantaneous BER is always lower than the target BER.

For the remaining numerical results, we will apply a target BER of $P_{\mathrm{b}}^{\star}=$ $10^{-4}$. Fig. 8.3 shows the average BER, ASE, and BEO curves for several values of $L_{\mathrm{r}}$ under the assumption of i.i.d. Rayleigh fading with $\mathcal{R}=\mathbf{I}_{2 L_{\mathrm{r}}}$. Using (8.11) and (8.13), the exact average BER can be obtained. The performance results are shown for both a receiver with PCSI and a mismatched receiver with LMMSE channel estimation. The non-monotonic behavior of the average BER results from the strong peaks in the instantaneous BER, as shown in Fig. 8.1. Also, it is observed from the figure that both ICSI and the number of receive antennas $L_{r}$ have a considerable impact on the ASE and the BEO. ICSI reduces the ASE significantly since channel estimation errors and the energy devoted to pilot symbols give rise to a degradation of the instantaneous BER, such that, compared to the case of PCSI, often a smaller constellation has to be selected in order to satisfy the target BER constraint. Moreover, the transmission of pilot symbols reduces the ASE by a factor $K /\left(K+K_{\mathrm{p}}\right)$ because part of the resources that could be used for data symbols are now occupied by pilot symbols. On the other hand, increasing the number of receive antennas increases the ASE since the provided spatial diversity mitigates fading, such that often a larger constellation can be selected which still satisfies the target BER constraint.

Fig. 8.4 shows the average BER, ASE and BEO for a dual-antenna receiver ( $L_{\mathrm{r}}=2$ ) under correlated Rayleigh fading with $\boldsymbol{\mathcal { R }}=\boldsymbol{\mathcal { R }}_{\mathrm{t}} \otimes \boldsymbol{\mathcal { R }}_{\mathrm{r}}$, where $\boldsymbol{\mathcal { R }}_{\mathrm{t}}$ and


Figure 8.3: Average BER, ASE and BEO of Alamouti's code under i.i.d. Rayleigh fading, for perfect and imperfect CSI, and for $L_{r} \in\{1,2,3\}$.


Figure 8.4: Average BER, ASE and BEO of Alamouti's code under correlated Rayleigh fading with $\rho \in\{0,0.3,0.8\}$, and for perfect and imperfect CSI.
$\mathcal{R}_{\mathrm{r}}$ are given by

$$
\boldsymbol{\mathcal { R }}_{\mathrm{t}}=\boldsymbol{\mathcal { R }}_{\mathrm{r}}=\left(\begin{array}{ll}
1 & \rho  \tag{8.25}\\
\rho & 1
\end{array}\right)
$$

and $\rho$ denotes the correlation factor. The results are shown for both a receiver with PCSI and a mismatched receiver with LMMSE channel estimation, and for $\rho \in\{0,0.3,0.8\}$. Note that $\rho=0$ corresponds to the case of i.i.d. fading. We observe that for low correlation levels, i.e., $\rho<0.3$, the impact of correlation on the ASE and BEO is fairly negligible. For high correlation, however, e.g., $\rho=0.8$, the ASE and BEO are clearly negatively affected by fading correlation.

### 8.3 MIMO-OFDM

In section 3.2.1, we have outlined how complex information symbols produced by the mapper can be converted into a real-valued continuous signal which can be transmitted over the channel. More specifically, it was shown how the successive information symbols are shaped onto a transmit pulse and modulated on a carrier wave with RF frequency $f_{\mathrm{c}}$. As the information bearing signal is modulated on a single carrier wave, this type of communication is called single-carrier communication. Single-carrier technology is particularly appealing on frequency-flat fading channels where the absence of ISI allows to construct low-complexity transmitters and receivers. Note that all BER expressions presented in sections 6 and 7 were derived for single-carrier communication over flat-fading channels. When the channel is frequency-selective, however, the spectrum of the transmitted signal is distorted and computationally complex equalizing techniques are usually required to counter the effect of ISI. Moreover, analytical error analysis becomes vastly complicated and the obtained BER expressions cannot be applied.

By splitting up the information symbol sequence into $N$ low-rate sequences which are modulated on $N$ different subcarriers and transmitted in parallel, the distortion due to frequency-selective fading can be largely avoided. In particular, longer transmit pulses can be used for the $N$ low-rate sequences such that the bandwidth of each of the signals can be made relatively small and the channel can be regarded as frequency-flat over each of the $N$ subbands. The technique where a symbol sequence is modulated on $N$ different subcarriers is called multi-carrier communication.

The most common multi-carrier technique is called orthogonal frequency division multiplexing (OFDM) [88]. As OFDM uses orthogonal subcarriers, inter-carrier interference (ICI) can be avoided and the receiver structure can be kept relatively simple. In combination with MIMO, OFDM is considered an appealing candidate for wireless applications requiring high data rates on frequency-selective channels [89]. Recently, MIMO-OFDM has been adopted in several standards, e.g., IEEE 802.16 (WiMAX) [90] and LTE [12]. In this
section, we briefly explain how the MIMO-OFDM system model reduces to $N$ parallel single-carrier MIMO systems under frequency-flat fading. In this way, the techniques used in chapters 6 and 7 to obtain the presented BER expressions, can also be useful to investigate the performance of MIMO-OFDM systems.

For notational convenience, we assess OFDM for a SISO system. As will be shown later on, the extension of OFDM to MIMO is pretty straightforward. As opposed to a single-carrier system transmitting information symbols at rate $1 / T$, a regular OFDM system uses $N$ subcarriers each transmitting in parallel at a symbol rate $1 /((N+v) T)$, yielding a total symbol rate of

$$
\begin{equation*}
R_{\mathrm{s}, \text { tot }}=\frac{N}{N+v} \frac{1}{T} \tag{8.26}
\end{equation*}
$$

The subcarriers are complex exponentials with frequency $f_{n}=n /(N T)$, with $n=0,1, \ldots, N-1$, such that they are orthogonal over an interval of duration $N T$. Although realistic OFDM systems require transmit pulses with finite bandwidth, we use rectangular transmit pulses $p(t)$ with unit energy for the sake of simplicity

$$
p(t) \triangleq\left\{\begin{array}{ll}
\frac{1}{\sqrt{(N+v) T}} & \text { if }-v T \leq t \leq N T  \tag{8.27}\\
0 & \text { otherwise }
\end{array} .\right.
$$

Hence, by introducing $\mathbf{s}(k)=\left[s_{0}(k), \ldots, s_{N-1}(k)\right]$ as the $k$-th block of $N$ information symbols, the complex-valued OFDM signal is given by

$$
\begin{equation*}
s(t)=\sqrt{E_{\mathrm{s}}} \sum_{k} \sum_{n=0}^{N-1} s_{n}(k) p(t-k(N+v) T) \exp \left(j 2 \pi f_{n} t\right) . \tag{8.28}
\end{equation*}
$$

The contribution of the $n$-th subcarrier to the transmitted signal (8.28) is shown in Fig. 8.5(a) for $k \in\{-1,0,1\}$. Note that the time intervals $[k(N+v) T, k(N+$ $v) T+N T$ ] of length $N T$ are the observation intervals, whereas the intervals $[k(N+v) T-v T, k(N+v) T]$ of length $v T$ are called guard intervals. Assuming that the signal $s(t)$ is transmitted over a frequency-selective channel with impulse response $h_{\mathrm{ch}}(t)=0$ if $t<0$ or $t>T_{\mathrm{ch}}$, and frequency response $H_{\mathrm{ch}}(f)$, it can be shown that within the observation intervals the received signal is given by

$$
\begin{equation*}
r(t)=\sqrt{\frac{E_{\mathrm{S}}}{(N+v) T}} \sum_{k} \sum_{n=0}^{N-1} s_{n}(k) H_{\mathrm{ch}}\left(f_{n}\right) \exp \left(j 2 \pi f_{n} t\right)+w(t), \tag{8.29}
\end{equation*}
$$

provided that $v T \geq T_{\text {ch }}$. In Fig. 8.5(b), the received signal corresponding to the signal transmitted on the $n$-th subcarrier is displayed. It is easily seen from the figure that by making the guard intervals longer than the channel impulse response, inter-block interference (IBI) is avoided in the observation

(a) Transmitted OFDM signal.

(b) Received OFDM signal.

Figure 8.5: Transmitted and received OFDM signals.
intervals. Moreover, due to the specific shape of the pulse, the signals modulated on the different subcarriers are not distorted within the observation intervals, yet only scaled by a complex-valued factor $H_{c h}\left(f_{n}\right)$. Because of the orthogonality of the subcarriers, ICI can be avoided between the contributions corresponding to the different subcarriers in (8.29). To this end, we introduce the variables $r_{n}(k)$, with $n=1, \ldots, N$, as

$$
\begin{equation*}
r_{n}(k)=\frac{1}{\sqrt{N T}} \int_{k(N+v) T}^{k(N+v) T+N T} r(t) \exp \left(-j 2 \pi f_{n} t\right) \mathrm{d} t \tag{8.30}
\end{equation*}
$$

Note that the receiver obtains $r_{n}(k)$ by multiplying the received signal $r(t)$ by $\exp \left(-j 2 \pi f_{n} t\right)$ and integrating the result over the $k$-th observation interval. By substituting (8.29) in (8.30), the variables $r_{n}(k)$ reduce to

$$
\begin{equation*}
r_{n}(k)=\sqrt{E_{\mathrm{s}}} \sqrt{\frac{N}{N+v}} H_{\mathrm{ch}}\left(f_{n}\right) s_{n}(k)+w_{n}(k) \tag{8.31}
\end{equation*}
$$

where $w_{n}(k)$ can be shown to be i.i.d. ZM CSCG RVs with variance $N_{0}$. Hence, the noise contributions corresponding to the different subcarriers are uncorrelated and $r_{n}(k)$ is a function of the corresponding information symbol $s_{n}(k)$ only. Introducing the channel coefficients $h_{n}$ as $h_{n} \triangleq \sqrt{N /(N+v)} H_{\mathrm{ch}}\left(f_{n}\right)$,
with $n=1, \ldots, N$, the system model (8.31) further reduces to

$$
\begin{equation*}
r_{n}(k)=\sqrt{E_{\mathrm{s}}} h_{n} s_{n}(k)+w_{n}(k) \tag{8.32}
\end{equation*}
$$

Hence, considering the system model (3.45), it is readily verified that OFDM reduces to $N$ parallel single-carrier systems operating over frequency-flat fading channels. In realistic OFDM systems, the modulation of the information symbols on the subcarriers is usually implemented by means of the inverse fast Fourier transform (IFFT), whereas the demodulation relies on the fast Fourier transform (FFT) [88]. However, the system model (8.32) is valid irrespective of the specific OFDM implementation.

To enable OFDM on a frequency-selective $L_{r} \times L_{t}$ MIMO channel, a multiantenna system must be equipped with $L_{t}$ OFDM modulators and $L_{r}$ OFDM demodulators [91]. Although MIMO-OFDM can be easily combined with orthogonal space-time block coding by associating to each of the $N$ subcarriers an $L_{t} \times K_{c}$ OSTBC matrix, this strategy introduces a lot of latency as the receiver has to wait for $K_{\mathrm{c}}$ successive OFDM blocks before it can decode the OSTBCs. Therefore, MIMO-OFDM systems usually employ space-frequency coding instead of space-time coding [92,93]. Orthogonal space-frequency block codes (OSFBCs) use identical code matrices as OSTBCs, yet the $K_{c}$ columns of the code matrix are associated to $K_{c}$ adjacent subcarriers instead of $K_{c}$ successive time slots. In this way, $N / K_{c}$ code matrices can be transmitted during one OFDM block. Assuming that the frequency response of the channel remains constant over the $K_{c}$ adjacent subcarriers used to transmit the $n^{\prime}$-th OSFBC matrix $\mathbf{C}_{n^{\prime}}$, with $n^{\prime}=1, \ldots, N / K_{\mathrm{c}}$, the system model corresponding to $\mathbf{C}_{n^{\prime}}$ reduces to

$$
\begin{equation*}
\mathbf{R}_{n^{\prime}}=\sqrt{E_{\mathrm{s}}} \mathbf{H}_{n^{\prime}} \mathbf{C}_{n^{\prime}}+\mathbf{W}_{n^{\prime}} \tag{8.33}
\end{equation*}
$$

where the $L_{r} \times L_{t}$ matrix $\mathbf{H}_{n^{\prime}}$ consists of the channel coefficients representing the channel on the subcarriers used to transmit $\mathbf{C}_{n^{\prime}}$. Note that (8.33) is equivalent to the system model (4.38) for single-carrier MIMO communication on flat-fading channels. This result suggests that the techniques presented in this dissertation could also be useful to investigate the performance of MIMOOFDM systems using orthogonal space-frequency block coding on frequencyselective fading channels with ICSI. However, it is important to note that the resulting expressions will depend on the specific strategy used to estimate the channel coefficients $\mathbf{H}_{n^{\prime}}$ in (8.33) [89,94-97].

### 8.4 Chapter Summary

In this chapter, we showed how we can extend and have already extended the techniques and performance results described in chapters 6 and 7 to a number of interesting MIMO OSTBC applications using different or additional transmission or modulation techniques. Preliminary analytical performance
results were shown for MIMO OSTBC systems using additional channel coding. Furthermore, we investigated the effect of imperfect channel estimation and fading correlation on the performance of rate-adaptive MIMO OSTBC systems. Assuming finite-rate feedback without delay, we presented accurate closed-form expressions for the average BER, ASE, and BEO, which enable the design of such adaptive communication systems. Finally, we indicated that the performance results obtained for single-carrier systems could also be extended to MIMO-OFDM systems employing orthogonal space-frequency block coding.

## 9

## Summary and Conclusions

In order to tackle the detrimental impact of multipath fading on the performance of wireless communication, most wireless systems apply one or more diversity techniques. In this way, the receiver is provided with multiple replicas of the same signal through different, preferably independent, paths, which can be generated in, e.g., frequency, time, or space. In this dissertation, we focused on the exploitation of spatial diversity by systems using multiple antennas at both the transmitter and receiver side. More specifically, these so-called MIMO systems are able to achieve a full diversity order of $L=L_{\mathrm{t}} L_{\mathrm{r}}$, with $L_{\mathrm{t}}$ and $L_{r}$ denoting the number of transmit and receive antennas, respectively, provided that a proper space-time coding scheme is used. All results in this dissertation were obtained for orthogonal space-time block codes, which are considered to be a very appealing transmit diversity technique, since they combine full diversity with a remarkably simple symbol-by-symbol optimal detection algorithm.

In digital communications, the principal figure of merit to evaluate the system performance is the bit error rate, which is defined as the ratio of the number of erroneously received bits to the total number of bits. Since BER
analysis through straightforward Monte-Carlo simulations involves bit error counting, the accuracy of the BER result directly depends on the number of observed bit errors. In case of low average BER, however, extremely long simulation times may be required to generate a sufficient number of bit errors. Therefore, efficient and easy-to-evaluate analytical BER results are to be preferred for communication techniques achieving low BERs at moderate SNRs, such as OSTBCs. Since an elegant analytical performance analysis of OSTBCs is facilitated by their simple symbol-by-symbol detection, we provided in this dissertation several useful BER expressions of OSTBC s under generalized fading conditions, assuming imperfect CSI at the receiver side. To this end, we considered a mismatched receiver using pilot aided channel estimation, which applies the channel estimate in the same way as a receiver with perfect CSI would apply the true channel.

For square OSTBCs and i.i.d. Rayleigh fading channels, we presented exact closed-form expressions for the BER and the BER degradation due to imperfect channel estimation. For non-square OSTBCs, a very accurate closed-form approximation was provided. For arbitrarily correlated Rayleigh fading channels, we derived closed-form BER approximations for square and non-square OSTBCs, yielding very accurate BER results in the low-to-moderate SNR region. For square OSTBCs, the BER approximation is asymptotically exact. In addition, we derived a simple rule of thumb that serves as an indicator for the BER degradation caused by imperfect channel estimation and yields the exact result in case of high SNR, square OSTBCs, PSK symbols, and i.i.d. Rayleigh fading.

Under the assumption of arbitrarily distributed flat-fading channels with ML channel estimation, we have reduced the exact BER expression to an expectation over $N_{\mathrm{s}}$ discrete RVs and $2 L+3$ real-valued continuous RVs, with $N_{\mathrm{s}}$ denoting the number of information symbols in the considered OSTBC matrix. For square OSTBCs, the resulting exact BER expression reduces further to an expectation over $N_{S}$ discrete RVs and 4 real-valued continuous RVs, regardless of the number of antennas. The exact BER expressions can be efficiently and accurately evaluated by means of numerical integration techniques, i.e., the quadrature rule and Monte-Carlo integration with importance sampling, or a combination thereof. The numerical evaluation of the exact BER expressions was specified for the case of correlated Nakagami-m fading channels. Additionally, we provided a computationally simple approximation of the BER based on treating the symbol interference due to imperfect channel estimation as white Gaussian noise. Although the resulting expression is in general not asymptotically exact, it yields very accurate BER results when the fading distribution is similar to Rayleigh and when a sufficient number of pilot symbols is used.

The BER expressions derived in this dissertation allow system designers to assess the effect of channel estimation errors and fading correlation on the
performance of OSTBCs. Moreover, we showed how the presented techniques and performance results can be extended to a number of interesting MIMO OSTBC applications using different or additional transmission or modulation techniques. Preliminary analytical BER results were shown for MIMO OSTBC systems using either additional channel coding or adaptive modulation and coding. Finally, we suggested that the performance results obtained for singlecarrier systems could also be extended to OFDM systems.

## Publications

The majority of the research described in this dissertation was presented in 15 publications: 5 refereed international journal publications and 10 conference publications.

## Journal Publications

- IEEE Communications Letters: [98]
- IEEE Transactions on Communications: $[72,99]$
- IEEE Transactions on Signal Processing: [100]
- IEEE Transactions on Wireless Communications: [101]


## Conference Publications

- Refereed international conference publications: [102-111]


## Bibliography

[1] A. Kaye and D. George, "Transmission of multiplexed PAM signals over multiple channel and diversity systems," IEEE Trans. Commun. Technol., vol. 18, no. 5, pp. 520-526, Oct. 1970.
[2] W. van Etten, "An optimum linear receiver for multiple channel digital transmission systems," IEEE Trans. Commun., vol. 23, no. 8, pp. 828-834, Aug. 1975.
[3] -_, "Maximum likelihood receiver for multiple channel transmission systems," IEEE Trans. Commun., vol. 24, no. 2, pp. 276-283, Feb. 1976.
[4] J. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environment," IEEE J. Select. Areas Commun., vol. 5, no. 5, pp. 871-878, Jun. 1987.
[5] G. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," Bell Labs Tech. J., pp. 41-59, Autumn 1996.
[6] E. Telatar, "Capacity of multi-antenna Gaussian channels," Eur. Trans. Telecomm. ETT, vol. 10, no. 6, pp. 585-596, Nov. 1999.
[7] L. Zheng and D. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," IEEE Trans. Inform. Theory, vol. 49, no. 5, pp. 1073-1096, May 2003.
[8] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," IEEE Trans. Inform. Theory, vol. 45, no. 5, pp. 1456-1467, Jul. 1999.
[9] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," IEEE J. Select. Areas Commun., vol. 16, pp. 1459-1478, Oct. 1998.
[10] E. T. S. Institute, "Digital video broadcasting (DVB); Framing structure, channel coding and modulation for $11 / 12 \mathrm{GHz}$ satellite services," France, ETS 300 421, Dec. 1994.
[11] IEEE 802.11, "Wireless LAN medium access control (MAC) physical layer (PHY) specifications, amendment 1: high-speed physical layer in the 5 GHz band," Jul. 1999.
[12] D. Astely, E. Dahlman, A. Furuskar, Y. Jading, M. Lindstrom, and S. Parkvall, "LTE: the evolution of mobile broadband," IEEE Commun. Mag., vol. 47, no. 4, pp. 44-51, Apr. 2009.
[13] J. Proakis, Digital Communications, 4th ed. McGraw-Hill, 2001.
[14] M. K. Simon and M.-S. Alouini, Digital Communication over Fading Channels, 2nd ed. New York: Wiley, 2005.
[15] K. Cho and D. Yoon, "On the general BER expression of one- and twodimensional amplitude modulations," IEEE Trans. Commun., vol. 50, no. 7, pp. 1074-1080, Jul. 2002.
[16] L. Najafizadeh and C. Tellambura, "BER analysis of arbitrary QAM for MRC diversity with imperfect channel estimation in generalized Ricean fading channels," IEEE Trans. Veh. Technol., vol. 55, no. 4, pp. 1239-1248, Jul. 2006.
[17] F. Lopez-Martinez, E. Martos-Naya, J. Paris, and U. Fernandez-Plazaola, "Generalized BER analysis of QAM and its application to MRC under imperfect CSI and interference in Ricean fading channels," IEEE Trans. Veh. Technol., vol. 59, no. 5, pp. 2598-2604, Jun. 2010.
[18] D. Brennan, "Linear diversity combining techniques," Proc. IRE, vol. 47, no. 6, pp. 1075-1102, Jun. 1959.
[19] R. Price and P. Green, "A communication technique for multipath channels," Proceedings of the IRE, vol. 46, no. 3, pp. 555-570, Mar. 1958.
[20] R. Mallik and G. Karagiannidis, "Equal-gain combining with unequal energy constellations," IEEE Trans. Wireless Commun., vol. 6, no. 3, pp. 1125-1132, Mar. 2007.
[21] Y. G. Kim and S. W. Kim, "Optimum selection combining for M-ary signals in fading channels," in Proc. IEEE Global Telecommunications Conf. (Globecom), vol. 2, Taipei, Taiwan, 17-21 Nov. 2002, pp. 1234-1238.
[22] G. Foschini, G. Golden, R. Valenzuela, and P. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," IEEE J. Select. Areas Commun., vol. 17, no. 11, pp. 1841-1852, Nov. 1999.
[23] A. Wittneben, "Basestation modulation diversity for digital simulcast," in Proc. IEEE Vehicular Technology Conf. (VTC), St. Louis, MO, USA, 19-22 May 1991, pp. 848-853.
[24] _-, "A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation," in Proc. IEEE Int. Conf. Comтип. (ICC), vol. 3, Geneva, Switzerland, 23-26 May 1993, pp. 1630-1634.
[25] N. Seshadri and J. Winters, "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity," Int. J. Wireless Information Networks, vol. 1, no. 1, pp. 49-60, Jan. 1994.
[26] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: Performance analysis and code construction," IEEE Trans. Inform. Theory, vol. 44, no. 2, pp. 744-765, Mar. 1998.
[27] X.-B. Liang, "Orthogonal designs with maximal rates," IEEE Trans. Inform. Theory, vol. 49, no. 10, pp. 2468-2503, Oct. 2003.
[28] J. Radon, "Lineare scharen orthogonaler matrizen," Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg, vol. 1, pp. 1-14, 1922.
[29] X.-B. Liang, "A complex orthogonal space-time block code for 8 transmit antennas," IEEE Commun. Lett., vol. 9, no. 2, pp. 115-117, Feb. 2005.
[30] G. Ganesan and P. Stoica, "Space-time block codes: a maximum SNR approach," IEEE Trans. Inform. Theory, vol. 47, no. 4, pp. 1650-1656, May 2001.
[31] D. Tse and P. Viswanath, Fundamentals of wireless communication. New York, NY, USA: Cambridge University Press, 2005.
[32] P. Samarasinghe, R. Kennedy, T. Abhayapala, and T. Lamahewa, "Spacetime cross correlation between signals at moving receiver," in Proc. Int. Conf. Information, Communications and Signal Processing (ICICS), Macau SAR, China, 7-10 Dec. 2009, pp. 1-5.
[33] H. Ozcelik, M. Herdin, W. Weichselberger, J. Wallace, and E. Bonek, "Deficiencies of 'Kronecker' MIMO radio channel model," Electronics Letters, vol. 39, no. 16, pp. 1209-1210, Aug. 2003.
[34] C. Oestges, "Validity of the Kronecker model for MIMO correlated channels," in Proc. IEEE Vehicular Technology Conf. (VTC-Spring), vol. 6, Melbourne, Australia, 7-10 May 2006, pp. 2818-2822.
[35] M. Ozcelik, N. Czink, and E. Bonek, "What makes a good MIMO channel model?" in Proc. IEEE Vehicular Technology Conf. (VTC-Spring), vol. 1, Stockholm, Sweden, 30 May-1 Jun. 2005, pp. 156-160.
[36] K. I. Pedersen, J. B. Andersen, J. P. Kermoal, and P. Mogensen, "A stochastic multiple-input-multiple-output radio channel model for evaluation of space-time coding algorithms," in Proc. IEEE Vehicular Technology Conference (VTC-Fall), vol. 2, Boston, MA, USA, 24-28 Sep. 2000, pp. 893-897.
[37] M. Nakagami, "The $m$-distribution-A general formula for intensity distribution of rapid fading," in Statistical Methods in Radio Wave Propagation. New York: Pergamon, 1960, pp. 3-36.
[38] M.-S. Alouini, A. Abdi, and M. Kaveh, "Sum of gamma variates and performance of wireless communication systems over Nakagami-fading channels," IEEE Trans. Veh. Technol., vol. 50, no. 6, pp. 1471-1480, Nov. 2001.
[39] G. K. Karagiannidis, N. C. Sagias, and T. A. Tsiftsis, "Closed-form statistics for the sum of squared Nakagami- $m$ variates and its applications," IEEE Trans. Commun., vol. 54, no. 8, pp. 1353-1359, Aug. 2006.
[40] J. Reig, "Performance of maximal ratio combiners over correlated Nakagami- $m$ fading channels with arbitrary fading parameters," IEEE Trans. Wireless Commun., vol. 7, no. 5, pp. 1441-1444, May 2008.
[41] P. Lombardo, G. Fedele, and M. M. Rao, "MRC performance for binary signals in Nakagami fading with general branch correlation," IEEE Trans. Commun., vol. 47, no. 1, pp. 44-52, Jan. 1991.
[42] Q. T. Zhang, "Maximal-ratio combining over Nakagami fading channels with an arbitrary branch covariance matrix," IEEE Trans. Veh. Technol., vol. 48, no. 4, pp. 1141-1150, Jul. 1999.
[43] M. Z. Win, G. Chrisikos, and J. H. Winters, "MRC performance for Mary modulation in arbitrarily correlated Nakagami fading channels," IEEE Commun. Lett., vol. 4, no. 10, pp. 301-303, Oct. 2000.
[44] G. C. Alexandropoulos, N. C. Sagias, F. I. Lazarakis, and K. Berberidis, "New results for the multivariate Nakagami-m fading model with arbitrary correlation matrix and applications," IEEE Trans. Wireless Commип., vol. 8, no. 1, pp. 245-255, Jan. 2009.
[45] N. C. Beaulieu and C. Cheng, "Efficient Nakagami-m fading channel simulation," IEEE Trans. Veh. Technol., vol. 54, no. 2, pp. 413-424, Mar. 2005.
[46] C. Lingzhi and N. C. Beaulieu, "Simple efficient methods for generating independent and bivariate Nakagami- $m$ fading envelope samples," IEEE Trans. Veh. Technol., vol. 56, no. 4, pp. 1573-1579, Jul. 2007.
[47] I.-M. Kim and V. Tarokh, "Variable-rate space-time block codes in M-ary PSK systems," IEEE J. Select. Areas Commun., vol. 21, no. 3, pp. 362-373, Apr. 2003.
[48] S.-H. Kim, I.-S. Kang, and J.-S. No, "Exact bit error probability of orthogonal space-time block codes with QAM," in Proc. IEEE Int. Symp. on Information Theory (ISIT), Yokohama, Japan, 29 Jun.-4 Jul. 2003, p. 63.
[49] C. Siriteanu and S. Blostein, "Performance and complexity analysis of eigencombining, statistical beamforming, and maximal-ratio combining," IEEE Trans. Veh. Technol., vol. 58, no. 7, pp. 3383-3395, Sep. 2009.
[50] I.-M. Kim, "Exact BER analysis of OSTBCs in spatially correlated MIMO channels," IEEE Trans. Commun., vol. 54, no. 8, pp. 1365-1373, Aug. 2006.
[51] A. A. Abu-Dayya and N. C. Beaulieu, "Outage probabilities of cellular mobile radio systems with multiple Nakagami interferers," IEEE Trans. Veh. Technol., vol. 40, no. 4, pp. 757-768, Nov. 1991.
[52] H. Zhang and T. A. Gulliver, "Capacity and error probability analysis for orthogonal space-time block codes over fading channels," IEEE Trans. Wireless Commun., vol. 4, no. 2, pp. 808-819, Mar. 2005.
[53] T. Tran and A. Sesay, "Sum of arbitrarily correlated Gamma variates and performance of wireless communication systems over Nakagami-m fading channels," IET Commun., vol. 1, no. 6, pp. 1133-1137, Dec. 2007.
[54] A. Maaref and S. Aissa, "Performance analysis of orthogonal space-time block codes in spatially correlated MIMO Nakagami fading channels," IEEE Trans. Wireless Commun., vol. 5, no. 4, pp. 807-817, Apr. 2006.
[55] W. Li, H. Zhang, and T. Gulliver, "Capacity and error probability analysis of orthogonal space-time block codes over correlated Nakagami fading channels," IEEE Trans. Wireless Commun., vol. 5, no. 9, pp. 2408-2412, Sep. 2006.
[56] G. Femenias, "BER performance of linear STBC from orthogonal designs over MIMO correlated Nakagami-m fading channels," IEEE Trans. Veh. Technol., vol. 53, no. 2, pp. 307-317, Mar. 2004.
[57] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," IEEE Trans. Veh. Technol., vol. 40, no. 4, pp. 686-693, Nov. 1991.
[58] A. L. Swindlehurst and G. Leus, "Blind and semi-blind equalization for generalized space-time block codes," IEEE Trans. Signal Process., vol. 50, no. 10, pp. 2489-2498, Oct. 2002.
[59] N. Ammar and Z. Ding, "Blind channel identifiability for generic linear space-time block codes," IEEE Trans. Signal Process., vol. 55, no. 1, pp. 202-217, Jan. 2007.
[60] G. Taricco, "Optimum receiver design and performance analysis of arbitrarily correlated Rician fading MIMO channels with imperfect channel state information," IEEE Trans. Inform. Theory, vol. 56, no. 3, pp. 11141134, Mar. 2010.
[61] G. Taricco and E. Biglieri, "Space-time decoding with imperfect channel estimation," IEEE Trans. Wireless Commun., vol. 4, no. 4, pp. 1874-1888, Jul. 2005.
[62] D. Mavares and R. P. Torres, "Channel estimation error effects on the performance of STB codes in flat frequency Rayleigh channels," in Proc. IEEE Vehicular Technology Conf. (VTC-Fall), vol. 1, Orlando, FL, USA, 6-9 Oct. 2003, pp. 647-651.
[63] E. Ko, C. Kang, and D. Hong, "Effect of imperfect channel information on M-QAM SER performance of orthogonal space-time block codes," in Proc. IEEE Vehicular Technology Conf. (VTC-Spring), vol. 1, Jeju Island, Korea, 22-25 Apr. 2003, pp. 722-726.
[64] L. Tao, L. Jianfeng, H. Jianjun, and Y. Guangxin, "Performance analysis for orthogonal space-time block codes in the absence of perfect channel state information," in Proc. IEEE Int. Symp. Personal, Indoor and Mobile Radio Communications (PIMRC), vol. 2, Beijing, China, 7-10 Sep. 2003, pp. 1012-1016.
[65] C. Shan, P. Y. Kam, and A. Nallanathan, "A symbol-by-symbol channel estimation receiver for space-time block coded systems and its performance analysis on the nonselective Rayleigh fading channel," IEEE Trans. Commun., vol. 56, no. 12, pp. 2116-2124, Dec. 2008.
[66] H. Cheon and D. Hong, "Performance analysis of space-time block codes in time-varying Rayleigh fading channels," in Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), vol. 3, Orlando, FL, USA, 13-17 May 2002, pp. 2357-2360.
[67] M. Brehler and M. K. Varanasi, "Asymptotic error probability analysis of quadratic receivers in Rayleigh-fading channels with applications to a unified analysis of coherent and noncoherent space-time receivers," IEEE Trans. Inform. Theory, vol. 47, no. 6, pp. 2383-2399, Sep. 2001.
[68] P. Garg, R. K. Mallik, and H. M. Gupta, "Performance analysis of spacetime coding with imperfect channel estimation," IEEE Trans. Wireless Commun., vol. 4, no. 1, pp. 257-265, Jan. 2005.
[69] R. K. Mallik and P. Garg, "Performance of optimum and suboptimum receivers for space-time coded systems in correlated fading," IEEE Trans. Comтии., vol. 57, no. 5, pp. 1237-1241, May 2009.
[70] W. Su and X.-G. Xia, "Two generalized complex orthogonal space-time block codes of rates $7 / 11$ and $3 / 5$ for 5 and 6 transmit antennas," IEEE Trans. Inform. Theory, vol. 49, no. 1, pp. 313-316, Jan. 2003.
[71] P. Garg, R. K. Mallik, and H. M. Gupta, "Exact error performance of square orthogonal space-time block coding with channel estimation," IEEE Trans. Commun., vol. 54, no. 3, pp. 430-437, Mar. 2006.
[72] L. Jacobs and M. Moeneclaey, "Comments on "exact error performance of square orthogonal space-time block coding with channel estimation"," IEEE Trans. Commun., vol. 57, no. 11, pp. 3249-3251, Nov. 2009.
[73] M. Evans, N. Hastings, and B. Peacock, Statistical Distributions 3rd ed. New York: Wiley, 2000.
[74] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, Numerical Recipes: The Art of Scientific Computing, 3rd ed. New York, NY, USA: Cambridge University Press, 2007.
[75] Z. Chen, Z. Chi, and B. Vucetic, "Error performance of maximal-ratio combining with transmit antenna selection in Nakagami- $m$ fading channels," in Proc. Int. Conf. on Wireless Communications, Networking, and Mobile Computing (WiCOM), Wuhan, China, 22-24 Sep. 2006, pp. 1-3.
[76] W. Lee, Mobile Communications: Design Fundamentals, 2nd ed. New York: Wiley, 1993.
[77] A. Goldsmith and S.-G. Chua, "Variable-rate variable-power MQAM for fading channels," IEEE Trans. Commun., vol. 45, no. 10, pp. 1218-1230, Oct. 1997.
[78] Y. Ko and C. Tepedelenlioglu, "Orthogonal space-time block coded rateadaptive modulation with outdated feedback," IEEE Trans. Wireless Commun., vol. 5, no. 2, pp. 290-295, Feb. 2006.
[79] J. Huang and S. Signell, "Impact of channel estimation error on performance of adaptive MIMO systems," in proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), Las Vegas, NV, USA, 30 Mar.-4 Apr. 2008, pp. 2865-2868.
[80] __, "On spectral efficiency of low-complexity adaptive MIMO systems in Rayleigh fading channel," IEEE Trans. Wireless Commun., vol. 8, no. 9, pp. 4369-4374, Sep. 2009.
[81] __, "On performance of adaptive modulation in MIMO systems using orthogonal space-time block codes," IEEE Trans. Veh. Technol., vol. 58, no. 8, pp. 4238-4247, Oct. 2009.
[82] A. Conti, M. Z. Win, and M. Chiani, "Slow adaptive M-QAM with diversity in fast fading and shadowing," IEEE Trans. Commun., vol. 55, no. 5, pp. 895-905, May 2007.
[83] A. Conti, D. Panchenko, S. Sidenko, and V. Tralli, "Log-concavity property of the error probability with application to local bounds for wireless communications," IEEE Trans. Inform. Theory, vol. 55, no. 6, pp. 27662775, Jun. 2009.
[84] M. Chiani, A. Conti, and C. Fontana, "Improved performance in TDCDMA mobile radio system by optimizing energy partition in channel estimation," IEEE Trans. Commun., vol. 51, no. 3, pp. 352-355, Mar. 2003.
[85] L. Toni and A. Conti, "Does fast adaptive modulation always outperform slow adaptive modulation?" IEEE Trans. Wireless Commun., vol. 10, no. 5, pp. 1504-1513, May 2011.
[86] A. Conti, M. Win, M. Chiani, and J. Winters, "Bit error outage for diversity reception in shadowing environment," IEEE Commun. Lett., vol. 7, no. 1, pp. 15-17, Jan. 2003.
[87] A. Conti, M. Win, and M. Chiani, "On the inverse symbol-error probability for diversity reception," IEEE Trans. Commun., vol. 51, no. 5, pp. 753-756, May 2003.
[88] S. Weinstein, "The history of orthogonal frequency-division multiplexing," IEEE Commun. Mag., vol. 47, no. 11, pp. 26-35, Nov. 2009.
[89] H. Yang, "A road to future broadband wireless access: MIMO-OFDMbased air interface," IEEE Commun. Mag., vol. 43, no. 1, pp. 53-60, Jan. 2005.
[90] IEEE 802.16, "IEEE standard for local and metropolitan area networks part 16: air interface for fixed broadband wireless access systems," May 2009.
[91] Y. Li, J. Winters, and N. Sollenberger, "MIMO-OFDM for wireless communications: signal detection with enhanced channel estimation," IEEE Trans. Commun., vol. 50, no. 9, pp. 1471-1477, Sep. 2002.
[92] K. Lee and D. Williams, "A space-frequency transmitter diversity technique for OFDM systems," in Proc. IEEE Global Telecommunications Conf. (Globecom), vol. 3, San Francisco, CA, USA, 27 Nov.-1 Dec. 2000, pp. 1473-1477.
[93] H. Bolcskei, M. Borgmann, and A. Paulraj, "Impact of the propagation environment on the performance of space-frequency coded MIMOOFDM," IEEE J. Select. Areas Commun., vol. 21, no. 3, pp. 427-439, Apr. 2003.
[94] F. Tufvesson and T. Maseng, "Pilot assisted channel estimation for OFDM in mobile cellular systems," in Proc. IEEE Vehicular Technology Conf. (VTC-Spring), vol. 3, Phoenix, AZ, USA, 4-7 May 1997, pp. 16391643.
[95] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," IEEE Trans. Broadcast., vol. 48, no. 3, pp. 223-229, Sep. 2002.
[96] M.-X. Chang and Y. Su, "Performance analysis of equalized OFDM systems in Rayleigh fading," IEEE Trans. Wireless Commun., vol. 1, no. 4, pp. 721-732, Oct. 2002.
[97] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," IEEE Trans. Signal Process., vol. 51, no. 6, pp. 1615-1624, Jun. 2003.
[98] L. Jacobs and M. Moeneclaey, "BER analysis of square OSTBCs with LMMSE channel estimation in arbitrarily correlated Rayleigh fading channels," IEEE Commun. Lett., vol. 14, no. 7, pp. 626-628, Jul. 2010.
[99] -_, "Effect of MMSE channel estimation on BER performance of orthogonal space-time block codes in Rayleigh fading channels," IEEE Trans. Commun., vol. 57, no. 5, pp. 1242-1245, May 2009.
[100] L. Jacobs, G. Alexandropoulos, M. Moeneclaey, H. Bruneel, and P. Mathiopoulos, "Analysis and efficient evaluation of the BER of OSTBCs with imperfect channel estimation in arbitrarily correlated fading channels," IEEE Trans. Signal Process., vol. 59, no. 6, pp. 2720-2733, Jun. 2011.
[101] P. Van Torre, L. Vallozzi, L. Jacobs, H. Rogier, M. Moeneclaey, and J. Verhaevert, "Characterization of measured indoor off-body MIMO channels with correlated fading, correlated shadowing and constant path loss," IEEE Trans. Wireless Commun., accepted for publication.
[102] L. Jacobs, F. Van Cauter, F. Simoens, and M. Moeneclaey, "BER performance and optimum training strategy for uncoded SIMO and Alamouti space-time block codes with MMSE channel estimation," in Proc. IEEE BENELUX/DSP Valley Signal Processing Symp. (SPS), Antwerp, Belgium, 21-22 Mar. 2007.
[103] L. Jacobs and M. Moeneclaey, "Exact BER analysis for Alamouti's code on arbitrary fading channels with imperfect channel estimation," in Proc. 1st COST2100 Workshop "MIMO and Cooperative Communications", Trondheim, Norway, 3-4 Jun. 2008.
[104] ——, "Accurate BER analysis of orthogonal space-time block codes with MMSE channel estimation," in Proc. IEEE Int. Symp. on Spread Spectrum Techniques and Applications (ISSSTA), Bologna, Italy, 25-28 Aug. 2008, pp. 318-323.
[105] —_, "Exact BER analysis for QAM transmission on arbitrary fading channels with maximal-ratio combining and imperfect channel estimation," in Proc. IEEE Int. Symp. Personal, Indoor and Mobile Radio Communications (PIMRC), Cannes, France, 15-18 Sep. 2008, pp. 1-5.
[106] L. Jacobs, G. Alexandropoulos, M. Moeneclaey, and P. Mathiopoulos, "BER of MRC for M-QAM with imperfect channel estimation over correlated Nakagami-m fading," in NEWCOM++/ACoRN Joint Workshop, Barcelona, Spain, 30 Mar.-1 Apr. 2009.
[107] L. Jacobs and M. Moeneclaey, "Exact BER analysis for square orthogonal space-time block codes on arbitrary fading channels with imperfect channel estimation," in Proc. IEEE Vehicular Technology Conf. (VTCSpring), Barcelona, Spain, 26-29 Apr. 2009, pp. 1-5.
[108] __, "Exact BER analysis for QAM transmission on arbitrary fading channels with maximal-ratio combining and imperfect channel estimation," in Proc. IEEE Int. Symp. Personal, Indoor and Mobile Radio Communications (PIMRC), Tokyo, Japan, 13-16 Sep. 2009.
[109] L. Jacobs, G. Alexandropoulos, M. Moeneclaey, and P. Mathiopoulos, "BER performance of OSTBCs on correlated fading channels with imperfect channel estimation," in Proc. Future Network and Mobile Summit 2010, Florence, Italy, 16-18 Jun. 2010, pp. 1-8.
[110] L. Jacobs and M. Moeneclaey, "Accurate BER analysis of square OSTBCs with imperfect channel estimation in arbitrarily correlated Rayleigh fading," in Proc. IEEE Int. Symp. on Spread Spectrum Techniques and Applications (ISSSTA), Taichung, Taiwan, 17-20 Oct. 2010, pp. 29-33.
[111] L. Jacobs, M. Moeneclaey, and A. Conti, "Rate-adaptive modulation for square OSTBCs in arbitrarily correlated Rayleigh fading with imperfect channel estimation," in Proc. IEEE Int. Workshop on Signal Processing Advances in Wireless Communications (SPAWC), San Francisco, CA, USA, 26-29 Jun. 2011, pp. 406-410.


[^0]:    ${ }^{1}$ We hereby tacitly assume that $\mathbf{x}$ and $\mathbf{y}$ are zero-mean random vectors. If this is not the case, a (linear) estimator of the form $\hat{\mathbf{x}}=\mathbf{M y}+\mathbf{b}$ should be used

[^1]:    ${ }^{1}$ In many textbooks, also a phase shift $\theta_{n}$ is associated to each path. These phase shifts can be taken into account by defining $\alpha_{n}$ in (3.25) as $\alpha_{n} \triangleq \gamma_{n} e^{-j 2 \pi f_{c} \tau_{n}+j \theta_{n}}$, yet this modification of the channel model will not affect its statistical modeling.

