Adaptieve modulatie, codering en vermogensallocatie in cognitieve radionetwerken

Adaptive Modulation, Coding and Power Allocation in Cognitive Radio Networks

Jeroen Van Hecke

Promotoren: prof. dr. ir. M. Moeneclaey, prof. dr. ir. L. Vandendorpe
Proefschrift ingediend tot het behalen van de graad van
Doctor in de ingenieurswetenschappen: elektrotechniek

UNIVERSITEIT
GENT

Vakgroep Telecommunicatie en Informatieverwerking
Voorzitter: prof. dr. ir. H. Bruneel
Faculteit Ingenieurswetenschappen en Architectuur
Academiejaar 2017 - 2018

ISBN 978-94-6355-075-8
NUR 959
Wettelijk depot: D/2017/10.500/110

Members of the Jury

- prof. Marc Moeneclaey
  Ghent University
- prof. Luc Vandendorpe
  Université catholique de Louvain
- prof. Jan Van Campenhout
  Ghent University
- prof. Hendrik Rogier
  Ghent University
- prof. Marc Moonen
  KU Leuven
- prof. François Horlin
  Université libre de Bruxelles
- dr. ir. Lennert Jacobs
  Ghent University

Dankwoord

In de eerste plaats wil ik hier graag Marc Moeneclaey bedanken. Het is bewonderenswaardig om te zien hoe hij een probleem tot zijn kern kan herleiden, en hoe grondig hij het werk van zijn doctoraatsstudenten opvolgt. Hij wist steeds tijd te maken om de papers grondig na te lezen, hoe veraf of hoe nabij de deadlines ook waren. Ten slotte, ben ik hem ook zeer dankbaar voor de kans die hij mij gegeven heeft om bij DIGCOM te mogen doctoreren.

Ik wil ook Patrick Schaillée en Sylvia Moeneclaey bedanken voor hun uitstekende werk op het secretariaat. Verder mag TELIN ook bijzonder trots zijn op de zeer professionele IT-support die Philippe Serbruyns en Davy Moreels de vakgroep bieden. Deze maken ons het leven een stuk eenvoudiger.

Verder wil ik ook mijn collega's bij TELIN uitvoerig bedanken. De vele discussies tijdens de lunch in de Brug, waarbij de ene discussie al iets meer of minder politiek correct was dan de andere, de kubb-trainingen in het park, de diepgaande analyses op dinsdag waar de meest recente aflevering van Game of Thrones onder de loep werd genomen, de vele spelletjesavonden, ... Deze momenten zal ik niet snel vergeten, en zorgen ervoor dat ik met veel plezier terugkijk op de periode dat ik bij TELIN was. Ik hoop oprecht dat ik vele van mijn collega's in de toekomst nog zal zien en dat de meeste van deze activiteiten, met uitzondering van de lunch in de Brug misschien, nog gerust zullen doorlopen nu mijn tijd bij TELIN erop zit.

Mijn ouders ben ik ontzettend dankbaar voor alle kansen en steun die ze me gegeven hebben. Ook mijn familie en vrienden wil ik bedanken, voor de al dan niet oprechte interesse in mijn werk en de leuke ontspannende tijd samen.

En ten slotte wil ik ook heel graag mijn vrouw Nicky bedanken, voor haar eindeloos begrip en steun. Ze is goud waard, geen olympisch maar toch Belgisch ;-), en ik kijk uit om samen met haar op ontdekkingsreis te gaan samen met ons nieuwe avontuur, genaamd Emma...

Gent, December 2017
Jeroen Van Hecke

Table of Contents

List of Abbreviations **vii**

List of Symbols **xi**

Nederlandse Samenvatting **xiii**

English Summary **xvii**

1 Introduction **1-1**
1.1 Background and Motivation 1-1
1.2 Outline 1-4

I Basic Concepts

2 Optimization Theory **2-1**
2.1 A general Optimization Problem 2-1
2.2 Convex Optimization 2-2
2.3 Lagrangian Duality 2-4
2.4 Subgradient Method 2-7
2.5 Stochastic Subgradient Method 2-9
2.6 CVX 2-10
2.7 Optimization Problems with Complex Variables 2-10
2.8 Chapter Summary 2-14

3 Digital Communications Basics **3-1**
3.1 The Wireless Channel 3-1
3.2 Digital Communication System Model 3-4
3.3 MMSE Estimator 3-11
3.4 Performance Metric 3-12
3.5 Resource Allocation 3-15
3.6 Cooperative Network 3-18
3.7 Cognitive Radio 3-22

3.8 Chapter Summary 3-26

4 Packet-Based Transmission Systems and the Goodput Metric 4-1
4.1 A Packet-Based Transmission System 4-2
4.2 Bit-Interleaved Coded Modulation 4-2
4.3 Goodput Metric 4-9
4.4 Effective SNR Mapping 4-11
4.5 Chapter Summary 4-16
4.A Cyclic Redundancy Check 4-17
4.B The Long-Term Average of the Goodput 4-18

II Optimization of Information-Theoretical Metrics

Overview

5 Resource Allocation with Multi-Antenna Relay Selection 5-1
5.1 System Model 5-3
5.2 Resource allocation 5-5
5.3 Minimization of the Link Outage Probability 5-9
5.4 Numerical Results 5-16
5.5 Chapter Summary 5-26
5.A Imperfect CSI 5-28
5.B Proof of Theorem 5.1 5-29
5.C Rank-1 Solution of the Relaxed Optimization Problem 5-30

6 Resource Allocation with Average Interference Constraints 6-1
6.1 System Model 6-2
6.2 Resource Allocation 6-3
6.3 Numerical Results 6-9
6.4 Chapter Summary 6-11
6.A Proof of Theorem 6.1 6-13

7 Resource Allocation with Relay Selection under an AI Constraint 7-1
7.1 System Model 7-2
7.2 Resource Allocation 7-5
7.3 Numerical Results 7-9
7.4 Chapter Summary 7-10

8 Resource Allocation for Multicarrier Networks 8-1
8.1 System Model 8-2
8.2 Resource Allocation 8-8
8.3 Numerical Results 8-10

8.4 Chapter Summary 8-13
8.A Imperfect CSI 8-14
8.B Derivation of Approximation (8.12) 8-15
8.C The Hungarian Algorithm 8-15

III Optimization of the Goodput

9 Accurate Modeling of the κESM with Imperfect Channel State Information 9-1
9.1 System Model 9-2
9.2 Statistical Model of κESM 9-3
9.3 Numerical Results 9-4
9.4 Chapter Summary 9-8
9.A Derivation of the Mean and Variance of Y 9-9

10 Adaptive Coding and Modulation using Imperfect Channel State Information in BIC-OFDM systems 10-1
10.1 Cognitive BIC-OFDM System Model 10-2
10.2 Goodput Performance Metric 10-3
10.3 Goodput Optimization 10-5
10.4 Numerical Results 10-10
10.5 Chapter Summary 10-20
10.A Examples of Different Types of Imperfect CSI at the Transmitter . 10-22
10.B PGP and IC-κESM 10-24

11 Concluding Remarks and Ideas for Future Work 11-1
11.1 Main Conclusions 11-1
11.2 Future Work 11-4
11.3 Publications 11-11

References

Abbreviations

ACM	Adaptive Coding and Modulation
AEGP	Approximate Expected Goodput
AF	Amplify-and-Forward
AI	Average Interference
AWGN	Additive White Gaussian Noise
BICM	Bit-Interleaved Coded Modulation
BICM-ID	Bit-Interleaved Coded Modulation using Iterative Decoding
BPSK	Binary Phase-Shift Keying
CCDF	Complementary Cumulative Distribution Function
CDF	Cumulative Distribution Function
CDI	Channel Distribution Information
CR	Cognitive Radio
CRC	Cyclic Redundancy Check
CSI	Channel State Information
DF	Decode-and-Forward
DN	Destination Node
DT	Direct Transmission
EESM	Exponential Effective SNR Mapping
EGP	Expected Goodput
ESM	Effective SNR Mapping
FBMC	Filter Bank Multicarrier
FFT	Fast Fourier Transform
GB	Greedy Bit
GEV	Generalized Extreme Value
GP	Goodput
ICSI	Imperfect Channel State Information
IP	Interference Probability
ISI	Intersymbol Interference
JS	Jensen-Shannon
κESM	Cumulant Generating Function based Effective SNR Mapping

KKT	Karush-Kuhn-Tucker
KL	Kullback-Leibler
LLR	Log-Likelihood Ratio
LTE	Long Term Evolution
MIESM	Mutual Information based Effective SNR Mapping
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output
ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
MRR	Maximum Ratio Reception
MRT	Maximum Ratio Transmission
MSE	Mean Square Error
ODA	Optimum Dynamic Allocation
OE	Optimized Energy
OFDM	Orthogonal Frequency-Division Multiplexing
OFDMA	Orthogonal Frequency-Division Multiple Access
OPMRT	Orthogonally Projected Maximum Ratio Transmission
PAM	Pulse Amplitude Modulation
PCSI	Perfect Channel State information
pdf	Probability Density Function
PER	Packet Error Rate
PGP	Predicted Goodput
PI	Peak Interference
PSK	Phase-Shift Keying
PU	Primary User
QAM	Quadrature Amplitude Modulation
RB	Robust Beamforming
RBIR	Received Bit Mutual Information Rate
RMS	Root Mean Square
RN	Relay Node
RX	Receiver
SA	Static Allocation
SDA	Suboptimum Dynamic Allocation
SN	Source Node
SNR	Signal-to-Noise Ratio
SU	Secondary User
TM	Transmission Mode
TX	Transmitter
UB	Uniform Bit
UE	Uniform Energy
UFMC	Universal Filtered Multicarrier

WSSUS	Wide-Sense Stationary Uncorrelated Scattering

Symbols

x	a scalar
$\Re(x)$	the real part of x
$\Im(x)$	the imaginary part of x
$\angle x$	the phase angle of x
$\|x\|$	absolute value of x
$[x]^+$	$\max(x, 0)$
$\mathbf{x}$	a column vector
$(\mathbf{x})_i$	denotes the *i*th component of $\mathbf{x}$
$\|\mathbf{x}\|$	Euclidean norm of $\mathbf{x}$
$\mathbf{x} \geq 0$	every element of the vector $\mathbf{x}$ is greater or equal to 0
$\mathbf{X}$	a matrix
$(\mathbf{X})_{i,j}$	denotes the element at the *i*th row and the *j*th column of $\mathbf{X}$
$\mathrm{Tr}\,(\mathbf{X})$	the trace of $\mathbf{X}$
$\mathbf{X}^T$	transpose of $\mathbf{X}$
$\mathbf{X}^*$	conjugate of $\mathbf{X}$
$\mathbf{X}^H$	conjugate transpose of $\mathbf{X}$
$\mathbf{X}^{-1}$	inverse of $\mathbf{X}$
$\|\mathbf{X}\|$	determinant of $\mathbf{X}$
$\|\mathbf{X}\|$	Frobenius norm of $\mathbf{X}$
$\mathbf{X} \succ 0$	a positive-definite matrix
$\mathbf{X} \succeq 0$	a positive semi-definite matrix
$\mathrm{vec}(\mathbf{X})$	this operator produces a vector that contains the columns of the matrix $\mathbf{X}$, stacked below each other
$\mathrm{diag}\,(x_1, \ldots, x_N)$	a diagonal matrix where $x_1, \ldots, x_N$ are the diagonal elements
$\mathbf{I}_n$	identity matrix of size $n \times n$
$\mathbf{e}_i$	ith column of $\mathbf{I}_n$
$\mathbf{z} \sim N_{\mathrm{c}}(\mathbf{0}, \mathbf{\Sigma})$	$\mathbf{z}$ is a zero-mean circularly symmetric complex Gaussian random variable with covariance matrix $\mathbf{\Sigma}$
$\propto$	equal within a certain constant factor
$\otimes$	Kronecker product

$\lceil \cdot \rceil$	ceiling operator
$\lfloor \cdot \rfloor$	flooring operator
$B(z, a, b)$	incomplete beta function
$\delta(x)$	Dirac delta function
δ_n	Kronecker delta function
dom	domain of a function
$\mathbb{E}\left[.\right]$	expectation operator
e^x	exponential function of x
$\inf$	infimum of a set
$J_0(x)$	zeroth-order Bessel function of the first kind
$\mathbf{J}_{i,j}$	single-entry matrix, 1 at $(i,\ j)$ and zero elsewhere
$K_1(x)$	first-order modified Bessel function of the second kind
$\ln$	natural logarithm
$\log_2$	base-2 logarithm
$\log_{10}$	base-10 logarithm
$\max$	maximum of a set
$\min$	minimum of a set
$Q(x)$	probability that a zero-mean unit variance Gaussian random variable exceeds x
$Q(\cdot, \cdot)$	first-order Marcum Q-function
relint	relative interior of a set
$\mathrm{rem}(a(x), b(x))$	the remainder of the modulo-2 polynomial division of $a(x)$ by $b(x)$

Samenvatting

Vandaag is het gebruik van draadloze digitale communicatie niet meer weg te denken uit onze maatschappij. Telefoneren, videogesprekken, sociale media en videostreaming zijn maar enkele van de toepassingen die dankbaar gebruik maken van de talrijke mogelijkheden van draadloze communicatie. Door de flexibiliteit en het gebruiksgemak dat draadloze communicatie biedt, stijgt het aantal toepassingen en toestellen dat hier gebruik van maakt zeer snel. Dit succes heeft echter ook een belangrijke keerzijde: deze toepassingen maken alle gebruik van hetzelfde eindige frequentiespectrum. Om al deze en toekomstige toepassingen een plaats te kunnen blijven geven, is het van het grootste belang dat het draadloze spectrum zo efficiënt mogelijk gebruikt wordt.

In dit doctoraatsonderzoek maken we gebruik van *cognitieve radio* om het spectrum efficiënt te benutten. Cognitieve radio biedt een mogelijke oplossing voor het feit dat bepaalde delen van het spectrum worden onderbenut. Het idee van cognitieve radio bestaat eruit om slimme zenders te gebruiken, die zich bewust zijn van hun omgeving. Deze slimme zenders zijn in staat om het frequentiespectrum te gebruiken van andere netwerken, zonder de werking van de oorspronkelijke, primaire netwerken te schaden. Cognitieve radio maakt het op deze manier dus mogelijk om gebruikers toe te voegen in delen van het frequentiespectrum die voor een conventionele zender al bezet geweest zouden zijn. Hierdoor wordt het spectrum efficiënter gebruikt en kunnen er meer services ondersteund worden in dezelfde bandbreedte. Het cognitieve netwerk wordt ook vaak het secundaire netwerk genoemd.

Het grote gemak dat draadloze communicatie aan de gebruiker biedt, staat echter in schril contrast met de moeilijkheden waarmee een draadloos communicatiesysteem wordt geconfronteerd. Er moeten immers vele hindernissen overwonnen worden, indien men tot een betrouwbaar communicatiekanaal wil komen. Een van de belangrijkste hindernissen is de *fading* van het kanaal: bewegende objecten in de omgeving en het feit dat een signaal meerdere paden kan afleggen tussen zender en ontvanger, leiden ertoe dat de signaalsterkte aan de ontvanger zal variëren in de tijd. Doordat er ook steeds ruis aanwezig is in het ontvangen signaal, kan een sterk geattenueerd signaal leiden tot detectiefouten aan de ontvangerzijde.

Een tweede belangrijke hindernis is interferentie. Indien meerdere zenders

actief zijn in dezelfde frequentieband, kan het gebeuren dat een ontvanger niet enkel het gewenste signaal ontvangt, maar ook ongewenste signalen afkomstig van andere zenders. Interferentie kan de werking van een netwerk ernstig verstoren indien er geen rekening mee gehouden wordt.

Om deze hindernissen te overwinnen, maken we gebruik van verschillende technieken. Eerst en vooral maken we gebruik van zogenaamde coöperatieve netwerken om de *fading* van het kanaal tegen te gaan. Deze netwerken bestaan uit verschillende *nodes*, die samenwerken om de boodschap van de zender naar de ontvanger te krijgen. Dikwijls zijn er meerdere paden tussen de zender en ontvanger in het coöperatieve netwerk. Deze verschillende paden maken het mogelijk om de ontvanger te voorzien van verschillende kopieën van hetzelfde signaal, hetgeen de diversiteit van het netwerk verhoogt. Aangezien de kans dat alle paden naar de ontvanger tegelijk sterk verzwakt worden klein is, heeft diversiteit een positieve invloed op de prestaties van het netwerk.

Ten tweede gaan we ervan uit dat de zender een zekere kanaalkennis tot zijn beschikking heeft. Hierdoor wordt het immers mogelijk voor de zender om diverse transmissieparameters aan te passen in functie van het kanaal om zo de werking van het netwerk te optimaliseren en interferentie met bestaande netwerken te vermijden. We houden er ook rekening mee dat de kanaalkennis aan de zender eventueel verouderd is of estimatiefouten bevat.

In dit doctoraatsonderzoek maken we een onderscheid tussen de theoretische en de meer praktische prestatiemetrieken. De meer theoretische prestatiemetrieken zijn gebaseerd op informatietheoretische grenzen, en stellen ons in staat om de best haalbare prestaties van een netwerk te onderzoeken. Het eerste scenario dat we van naderbij bekijken, bestaat uit een coöperatief cognitief netwerk waar de *relay nodes*, die de zender en ontvanger ondersteunen, meerdere zend- en ontvangstantennes tot hun beschikking hebben. Volgens het principe van cognitieve radio, mag dit secundaire netwerk de werking van het oorspronkelijke, primaire netwerk niet verstoren. Om dit te garanderen zal het secundaire netwerk er steeds op toezien dat de interferentie aan de primaire gebruiker beperkt blijft, terwijl de prestaties van het eigen netwerk geoptimaliseerd worden. Dit is mogelijk doordat het secundaire netwerk informatie heeft over de kanalen naar de *nodes* van zowel het secundaire als het primaire netwerk. Verder kan het secundaire netwerk het zendvermogen aanpassen aan de zender en de *relay nodes* en, door de aanwezigheid van meerdere antennes aan de *relay nodes*, kan het netwerk zelfs *beamforming* toepassen. We onderzoeken verschillende algoritmes die afhankelijk zijn van de hoeveelheid kanaalkennis die aanwezig is bij de *nodes*. Afhankelijk van de kwaliteit van de kanaalkennis wordt het voor de *relay nodes* mogelijk om hun zendrichting weg te sturen van de ontvangers van het primaire netwerk en te richten naar de ontvangers van het secundaire netwerk.

Verder onderzoeken we ook de werking van het secundaire netwerk indien

niet de piekinterferentie beperkt wordt aan de primaire ontvangers, maar wel de gemiddelde interferentie. Aangezien een beperking op de gemiddelde interferentie leidt tot meer flexibiliteit in de toekenning van de transmissieparameters, is er een grote prestatiewinst mogelijk voor het secundaire netwerk. De optimalisatie van deze netwerken kan echter voor eenvoudige netwerken al snel moeilijk worden, waardoor we voor de meer geavanceerde netwerken enkel beperkingen op de piekinterferentie beschouwen.

Naast systemen met één draaggolf, wordt ook de adaptatie van transmissieparameters bij systemen met meerdere draaggolven onderzocht. Het bekendste voorbeeld van modulatie met meerdere draaggolven is "orthogonal frequency division multiplexing (OFDM)". Bij deze modulatie worden de data op verschillende draaggolven geplaatst. Hierdoor wordt het nu ook mogelijk om de energie per draaggolf te optimaliseren. Voor deze systemen beschouwen we ook een meer praktische prestatiemetriek: de *goodput*. Aangezien de *goodput* de verhouding uitdrukt van het verwachte aantal correct ontvangen informatiebits ten opzichte van de werkelijke transmissietijd, geeft deze metriek ons een goed beeld van de werkelijke prestaties van het systeem. Verder stelt deze prestatiemetriek ons in staat om naast de optimalisatie van de energie per draaggolf, ook het aantal bits per draaggolf en het codedebiet te optimaliseren.

Voor de scenario's waar de zender enkel toegang heeft tot onvolmaakte kanaalkennis, stellen we de verwachte *goodput* als prestatiemetriek voor. Deze metriek drukt de verwachtingswaarde van de *goodput* uit, geconditioneerd op de beschikbare kanaalkennis. We tonen immers aan dat dit overeenkomt met de optimale prestatiemetriek die de gemiddelde *goodput* maximaliseert. Voor deze metriek stellen we een accurate benadering voor, die gebruikt wordt om de bijhorende adaptatie-algoritmes af te leiden.

We sluiten dit proefschrift af met een korte samenvatting van de belangrijkste resultaten, en formuleren enkele mogelijkheden om dit werk verder uit te breiden. Zo bespreken we onder andere de mogelijkheid om een offset in de draaggolffrequentie te modelleren, alsook de mogelijke uitbreiding naar een iteratieve ontvanger of het gebruik van turbocodes.

Summary

In today's society, wireless digital communication systems are omnipresent. Mobile telephony, video calls, social media or video streaming are just a few of the many applications that make use of the numerous possibilities of wireless communication. Because of the flexibility and ease of use, the number of wireless applications and devices grows incredibly fast. However, this success comes at a price: all these applications share the same finite frequency spectrum. In order to support all current and future applications, it becomes paramount to use the wireless spectrum as efficiently as possible.

In this dissertation, we have investigated the use of *cognitive radio*, as it offers a possible solution to the underutilization of the frequency spectrum. The idea of cognitive radio consists of the use of intelligent transmitters, that have a certain awareness about their environment. These intelligent transmitters are able to access the frequency spectrum of existing, primary networks, without harming their performance. Cognitive radio is thus able to add additional users to parts of the frequency spectrum that would have been inaccessible to a conventional transmitter. This increases the spectral efficiency, which means that more services can be supported in the same amount of bandwidth. The cognitive network is also called the secondary network.

Although wireless communication systems offer great ease of use, a large number of phenomena have to be overcome in order to allow reliable communication over a wireless channel. A most important phenomenon is fading: moving objects in the environment and the presence of multiple signal paths between the transmitter and the receiver cause time-varying fluctuations in the signal strength at the receiver. Because of the presence of background noise, a strongly attenuated signal will lead to detection errors at the receiver.

A second important phenomenon is interference. When multiple transmitters are active in the same frequency band, it is possible that the destination does not only receive the desired signal, but also the unwanted signals from the other transmitters. When the problem of interference is neglected, it can have a severe impact on the network performance.

To overcome these obstacles, several techniques are applied. First, we introduce a cooperative network to counteract the fading of the wireless channel. A

cooperative network consists of several nodes, which assist the transmitter in delivering its message to the receiver. In most cooperative networks, different paths between the transmitter and the receiver are created. These paths allow the receiver to get multiple copies of the same signal, which increases the diversity of the network. As it is highly unlikely that all the paths are simultaneously in a deep fade, diversity can vastly improve the performance of a network.

Secondly, we assume that the transmitter has a certain amount of channel information at its disposal. This information allows the transmitter to adapt its transmission parameters so it can optimize the performance metric of the network, while avoiding interference to existing networks. We take into account that the available channel information at the transmitter can be outdated or imperfect due to feedback delays or estimation errors, respectively.

In this dissertation we will consider both theoretical and practical performance metrics. The theoretical performance metrics are based on information-theoretical bounds, as these allow us to investigate the best possible performance of a network. The first scenario that we consider, consists of a cooperative cognitive network where the relay nodes, that aid the transmitter and receiver, are equipped with multiple transmit and receive antennas. According to the principle of cognitive radio, the secondary network should not degrade the performance of the existing, primary network. For this reason, the secondary network has to constrain the interference it causes to the primary user receiver, while optimizing the performance of its own network. When the secondary network has some information about the channel gains of the secondary network and about the channel gains towards the primary network, it can adapt the transmission power at the transmitter and relay nodes, and, because of the presence of multiple antennas at the relay nodes, it can also apply beamforming at these nodes. Several algorithms are investigated that depend upon the level of channel information that is available at the nodes. Depending upon the quality of the channel information, the relay nodes can point their transmit beam away from the receivers of the primary network and towards the receivers of the secondary network.

Further, we investigate the scenario where the secondary network limits the average interference at the primary receivers, rather than the peak interference. As these average interference constraints allow for more flexibility in the allocation of the transmission parameters, they can significantly improve the performance of the secondary network. However, it appears that the optimization under these average interference constraints can become quite difficult even for simple networks, which is why we only consider the peak interference constraints for the more advanced networks.

Next to single-carrier systems, the adaptation of transmission parameters in multicarrier systems is also investigated. The most famous example of multicarrier modulation is “orthogonal frequency division multiplexing (OFDM)”. Multi-

carrier modulation transmits its data on different subcarriers, which allows us to optimize the allocated energy per subcarrier. For these systems, we also consider a more practical performance metric called the goodput. As the goodput expresses the ratio of the expected number of correctly received information bits to the actual transmission time, it describes the actual performance of the network under consideration. In addition to the energy per subcarrier, this performance metric allows us to optimize the number of bits per subcarrier and the code rate.

For the scenarios where only imperfect channel information is available at the transmitter, we propose the expected goodput as the performance metric. This metric is defined as the expectation of the goodput, conditioned on the available channel information. We show that this metric corresponds to the optimal performance metric that achieves the highest average goodput. Finally, we propose an accurate approximation for this metric, which allows the transmitter to adapt its transmission parameters.

We conclude this dissertation with a short summary of our main results and present some ideas for future work. For example, we discuss the possibility to incorporate a carrier frequency offset into the analysis, or the possible extension to an iterative receiver or the use of turbo codes.

1 Introduction

In this doctoral thesis, we investigate dynamic resource allocation algorithms for cognitive radio networks. These algorithms allow the transmitter to optimize the network performance by means of a dynamic adaptation of its transmission parameters (power, constellation size, code rate). The proposed algorithms take into account the quality and type of channel information that is available at the transmitter. This thesis consists of three parts.

In the first part of this dissertation, we provide the reader with the necessary background for the subsequent chapters. The second and third part discuss algorithms that optimize information-theoretical and practical performance metrics, respectively.

In section 1.1, we give some background and motivation for this work, while we present the outline of this dissertation in section 1.2.

1.1 Background and Motivation

Today, digital communication is more important than ever. Online streaming services, social media or mobile telephony are just a few of today's applications which rely heavily on the use of digital communications.

In digital communications, the information is represented by sequences of binary values 0 and 1. The node that sends the information is called the *transmitter*, while the node that requests the information is called the *receiver*. The medium

Figure 1.1: A general communication system.

over which this information is transmitted is called *the channel*. The corresponding scheme is shown in Fig. 1.1. For wired communications, the channel is a physical connection such as a twisted-pair cable or optical fiber. However, an increasing amount of devices uses the wireless channel. The latter channel is extremely practical for the user, as it is always present and allows for communication between devices without any physical connection.

However, the convenience of this wireless channel comes at a price: it is extremely challenging to achieve reliable communication over this channel, as there are a large number of phenomena that can adversely affect the communication.

- The *path loss* reduces the received signal power with increasing distance. In addition, the presence of obstacles can further lower the received signal strength. In the presence of noise (e.g., from electronic circuits), small signals can cause the message to be incorrectly received.

- The fact that all wireless transmitters share the same channel gives rise to *interference*. In order to cope with this interference, strict regulations about which service may transmit in which frequency band have been put forward. In Fig. 1.2 we show how the frequency allocation in the United States is regulated. This figure depicts the licensed services that are located in each frequency band, and clearly shows that the wireless spectrum has become very crowded. However, it has been demonstrated that many of these frequency bands are severely underutilized [1]. From this observation, it became clear that the spectral efficiency could be improved by using so-called cognitive radio networks. These networks aim to increase the spectral efficiency by adding additional users to the already crowded spectrum, while protecting the quality of service of the existing users.

- The wireless channel introduces *multipath fading*: the transmitted signal often gets reflected by several objects, which can cause destructive interference of the reflected signals at the receiving antenna. This phenomenon is called a deep fade, and can be detrimental to the performance of a wireless network. Because of moving obstacles in the environment, the fading becomes time-varying. The impact of this time-varying fading can be reduced by creating *diversity*: if the receiver gets multiple copies of the same signal through separate channels, it is highly unlikely that the signal experiences a

Figure 1.2: The frequency allocation in the United States (January 2016, [2]).

deep fade in each channel simultaneously. The diversity of a network can be increased by using cooperative networks. In these networks, additional transmitters relay the original transmitted signal. By selecting the transmitter with the best channel to the receiver, or by allowing all transmitters to be active in non-overlapping time slots, we can increase the reliability of the communication.
Further, the multipath fading can also cause problems if there is a large time difference between the arrival of the first signal and the last signal. This effect is known as the *delay spread* and can cause the signal to interfere with itself.

In order to detect the transmitted message, the receiver needs the channel gain experienced by the transmitted signal. However, as the wireless channel is time-varying, the receiver periodically needs a reliable estimate of the current channel gain. It was shown in [3] that the spectral efficiency can be improved by resource allocation algorithms which dynamically adapt the constellation size, code rate and power at the transmitter as a function of the channel gain. However, the drawback is that the channel state information now also has to be available at the transmitter.

Already many resource allocation algorithms have been proposed in literature. Yet, many of these algorithms assume that perfect channel state information is available at the transmitter. However, in a time-varying environment where background noise is always present, the channel information in a real system will often be a noisy and delayed estimate. In this dissertation, several algorithms will be investigated that optimize the performance of a cognitive radio network when only imperfect channel state information is available.

We will consider both information-theoretical and practical performance metrics. The information-theoretical metrics allow us to investigate the best performance that the network can achieve, while the practical metrics describe the actual performance of the network under consideration. In addition, a practical metric allows us to optimize the code rate and the constellation size, which cannot be achieved by most information-theoretical metrics.

1.2 Outline

This dissertation is organized as follows:

Part I

Chapter 2 provides an overview of some important concepts of optimization theory. The concepts introduced in this chapter will be used to optimize the performance of a wireless network.

Chapter 3 presents the system model that will be used in this dissertation. It discusses the basic principles of digital communications and the wireless channel. This chapter also introduces the minimum mean square error estimator, which will be used at the transmitter for the estimation of the channel. Further, we discuss several performance metrics, which will be optimized by means of a suitable resource allocation algorithm. Finally, we conclude this chapter by presenting several cooperative network protocols and the three main paradigms of cognitive radio: the underlay, overlay and interweave paradigm. As we focus our attention on the underlay paradigm, we explain the different possible formulations of the interference constraints in more detail.

Chapter 4 introduces the necessary background for packet-based transmission systems. Further, we introduce a suitable practical metric, referred to as *goodput*, that describes the performance of such a system, and discuss the concept of effective signal-to-noise ratio in the context of multicarrier transmission.

Part II

Chapter 5 considers a cognitive radio network that uses relay nodes with multiple transmit antennas, under a peak interference constraint. Further, we show how the cognitive radio network can minimize its outage probability for different types of channel state information by dynamically selecting its transmission parameters.

Chapter 6 investigates the performance of a cognitive radio network when the network limits the average interference inflicted on the licensed users instead of the peak interference. An average constraint on the interference is less restrictive than the peak interference constraint that was investigated in the previous chapter, and thus allows for a more flexible allocation of the transmission parameters.

Chapter 7 extends the previous chapter by combining the average interference constraints with a relay selection algorithm. This makes the resource allocation more complicated, as the relay selection alters the distribution of the interference at the licensed users.

Chapter 8 investigates multicarrier networks. A resource allocation algorithm is presented that optimizes the maximal rate of the cognitive radio network by optimizing the subcarrier pairing and transmit energy per subcarrier.

Part III

Chapter 9 provides and investigates an accurate approximation of the goodput, that will be used to describe the performance of a packet-based transmission system when only imperfect channel state information is available.

Chapter 10 derives and analyzes several resource allocation algorithms for a packet-based cognitive radio network. These algorithms use the goodput approximation

proposed in the previous chapter and optimize the transmit energy per subcarrier, the code rate and the bit allocation per subcarrier.
Chapter 11 concludes this dissertation and gives some ideas for future work.

Part I

Basic Concepts

2 Optimization Theory

In this chapter, we give the reader an overview of some important concepts about optimization theory. This chapter should give the reader a basic understanding of many of the concepts that we will rely on in the following chapters.

In section 2.1, we introduce a general notation for a constrained optimization problem and some important definitions. In section 2.2, we restrict our focus to the so called convex optimization problems. Then, a very useful tool, called Lagrangian duality, is introduced in section 2.3. Sections 2.4, 2.5 and 2.6 discuss several numerical methods to solve constrained optimization problems. Finally, we delve into some details about optimization problems with complex-valued variables in section 2.7.

The references [4–10] have been used as the basis for this overview.

2.1 A general Optimization Problem

First, we define a general optimization problem. The following notation

$$\begin{aligned} &\min_{\mathbf{x}} f(\mathbf{x}) \\ &\text{s.t. } h_i(\mathbf{x}) = 0,\ i = 1, \ldots, m \\ &\quad g_j(\mathbf{x}) \leq 0,\ j = 1, \ldots, n, \end{aligned} \tag{2.1}$$

describes an optimization problem where a vector $\mathbf{x} \in \mathbb{R}^{N \times 1}$ has to be found that minimizes the *objective function* $f(\mathbf{x})$. However, this vector $\mathbf{x}$ also has to satisfy

the *equality constraints* $h_i(\mathbf{x}) = 0$, $i = 1, \ldots, m$, and *inequalities constraints* $g_j(\mathbf{x}) \leq 0, j = 1, \ldots, n$. The problem in (2.1) is called a *constrained* optimization problem. When there are no equality and inequality constraints, the optimization problem is called *unconstrained*. The domain $\mathcal{D}$ of (2.1) is given by

$$\mathcal{D} = \text{dom}\, f \bigcap_{i=1}^{m} \text{dom}\, h_i \bigcap_{j=1}^{n} \text{dom}\, g_j. \tag{2.2}$$

A *feasible point* is a point $\mathbf{x} \in \mathcal{D}$ which satisfies the constraints. Further, a point $\mathbf{x}_\text{G} \in \mathcal{D}$ is called a *globally optimal* point when it is feasible and $f(\mathbf{x}_\text{G}) = c$, where c is the *optimal value* of optimization problem (2.1). The *optimal value* c is defined as

$$c \triangleq \inf \{ f(\mathbf{x}) | h_i(\mathbf{x}) = 0,\ i = 1, \ldots, m,\ g_j(\mathbf{x}) \leq 0,\ j = 1, \ldots, n \}. \tag{2.3}$$

A point $\mathbf{x}_\text{L} \in \mathcal{D}$ which satisfies the constraints and for which there exists a value of $R > 0$ such that

$$\begin{aligned} f(\mathbf{x}_\text{L}) \quad = \quad & \inf \big\{ f(\mathbf{y}) | h_i(\mathbf{y}) = 0,\ i = 1, \ldots, m, \\ & g_j(\mathbf{y}) \leq 0,\ j = 1, \ldots, n,\ \|\mathbf{y} - \mathbf{x}_\text{L}\| \leq R) \big\}, \end{aligned} \tag{2.4}$$

is called a *locally optimal* point.

2.2 Convex Optimization

As a general optimization problem (2.1) can have multiple locally optimal points, it can become extremely difficult to solve such a problem. A class of optimization problems, which are easier to solve, are the convex optimization problems. These optimization problems have the nice property that any locally optimal point is also a globally optimal point. Before we introduce the definition of a convex optimization problem, we will first introduce the concept of a convex set and convex function.

A set $\mathcal{C}$ is convex, when for any x and y in $\mathcal{C}$ and any $\theta \in [0, 1]$, we have

$$\theta x + (1 - \theta) y \in \mathcal{C}. \tag{2.5}$$

An example of a convex and non-convex set in $\mathbb{R}^2$ is shown in Fig. 2.1.

A function is said to be convex, if its domain is a convex set and if $\forall \mathbf{x}, \mathbf{y} \in \text{dom} f$

$$f(\theta \mathbf{x} + (1 - \theta)\mathbf{y}) \leq \theta f(\mathbf{x}) + (1 - \theta) f(\mathbf{y}), \tag{2.6}$$

where $\theta \in [0, 1]$. Similarly, a function $f(\mathbf{x})$ is concave when $-f(\mathbf{x})$ is convex. An example of a convex and non-convex function is shown in Fig. 2.2.

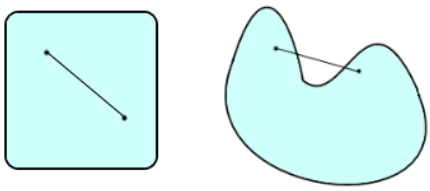


Figure 2.1: An example of a convex (left) and non-convex set (right).

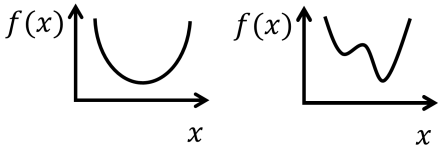


Figure 2.2: An example of a convex (left) and non-convex function (right).

Finally, we also need the following definition:

A function f is affine, if its domain is a convex set and if the function can be written as

$$f(\mathbf{x}) = \mathbf{a}^T\mathbf{x} + c, \tag{2.7}$$

where $\mathbf{a} \in \mathbb{R}^{N\times 1}$ and $c \in \mathbb{R}$.

A general convex optimization problem can be formulated as follows

$$\begin{aligned} &\min_{\mathbf{x}} f(\mathbf{x}) \\ &\text{s.t. } h_i(\mathbf{x}) = 0,\ i = 1,\ldots,m \\ &\quad g_j(\mathbf{x}) \leq 0,\ j = 1,\ldots,n, \end{aligned} \tag{2.8}$$

where $f(x)$ and $g_j(\mathbf{x}), j = 1,\ldots,n$, are convex functions and $h_i(\mathbf{x}), i = 1,\ldots,m$, are affine functions. We note that the maximization of a concave function $f(\mathbf{x})$ with convex constraints can always be rewritten as a convex optimization problem. This follows from the fact that maximizing a concave function $f(\mathbf{x})$ is completely equivalent with minimizing $-f(\mathbf{x})$, which is convex.

A special class of convex optimization problems is called the semidefinite program. These problems are defined as

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{c}^T\mathbf{x} \\ &\text{s.t. } x_1\mathbf{D}_1 + \ldots + x_N\mathbf{D}_N + \mathbf{D}_{N+1} \preceq 0 \\ &\quad \mathbf{A}\mathbf{x} = \mathbf{b}, \end{aligned} \tag{2.9}$$

where $\mathbf{D}_n \in \mathbb{R}^{M\times M}$, $n = 1,\ldots,N, N+1$, is a symmetric matrix, $\mathbf{A} \in \mathbb{R}^{P\times N}$, $\mathbf{b} \in \mathbb{R}^{P\times 1}$ and $\mathbf{c} \in \mathbb{R}^{N\times 1}$. The operator $\mathbf{X} \preceq 0$ denotes that a matrix $\mathbf{X}$ is negative semi-definite.

2.3 Lagrangian Duality

In this section, we will introduce the concept of Lagrangian duality, which is a tool that is often used to solve an optimization problem. In the following, we do not assume the optimization problem to be convex. Taking the constraints in (2.1) into account by augmenting the objective function $f(\mathbf{x})$ with a weighted sum of the equality and inequality constraints, we get

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \triangleq f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i h_i(\mathbf{x}) + \sum_{j=1}^{n} \mu_j g_j(\mathbf{x}). \tag{2.10}$$

This expression is called the *Lagrangian* of (2.1). The quantities $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_m]^T \in \mathbb{R}^{m\times 1}$ and $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_n]^T \in \mathbb{R}^{n\times 1}$ are the *Lagrange multiplier vectors*, and they correspond to the equality and inequality constraints, respectively. The *dual function* $s(\boldsymbol{\lambda}, \boldsymbol{\mu})$ is then defined as

$$s(\boldsymbol{\lambda}, \boldsymbol{\mu}) \triangleq \inf_{\mathbf{x}\in\mathcal{D}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}). \tag{2.11}$$

When the Lagrangian is unbounded from below in $\mathbf{x}$, the dual function is equal to $-\infty$. We can now prove the following:

Theorem 2.1. *The dual function $s(\boldsymbol{\lambda}, \boldsymbol{\mu})$ is always concave.*

Proof. $\forall\theta \in [0, 1]$, we can write the following

$$\begin{aligned}
&s(\theta\boldsymbol{\lambda}_1 + (1-\theta)\boldsymbol{\lambda}_2, \theta\boldsymbol{\mu}_1 + (1-\theta)\boldsymbol{\mu}_2) \\
=&\inf_{\mathbf{x}\in\mathcal{D}} L(\mathbf{x}, \theta\boldsymbol{\lambda}_1 + (1-\theta)\boldsymbol{\lambda}_2, \theta\boldsymbol{\mu}_1 + (1-\theta)\boldsymbol{\mu}_2) \\
=&\inf_{\mathbf{x}\in\mathcal{D}} \Big\{\theta f(\mathbf{x}) + \theta\sum_{i=1}^{m} \lambda_{1,i} h_i(\mathbf{x}) + \theta\sum_{j=1}^{n} \mu_{1,j} g_j(\mathbf{x}) \\
&+ (1-\theta) f(\mathbf{x}) + (1-\theta)\sum_{i=1}^{m} \lambda_{2,i} h_i(\mathbf{x}) + (1-\theta)\sum_{j=1}^{n} \mu_{2,j} g_j(\mathbf{x})\Big\} \\
\geq&\theta \inf_{\mathbf{x}\in\mathcal{D}} L(\mathbf{x}, \boldsymbol{\lambda}_1, \boldsymbol{\mu}_1) + (1-\theta) \inf_{\mathbf{x}\in\mathcal{D}} L(\mathbf{x}, \boldsymbol{\lambda}_2, \boldsymbol{\mu}_2) \\
=&\theta s(\boldsymbol{\lambda}_1, \boldsymbol{\mu}_1) + (1-\theta) s(\boldsymbol{\lambda}_2, \boldsymbol{\mu}_2),
\end{aligned} \tag{2.12}$$

which proves that the dual function $s(\boldsymbol{\lambda}, \boldsymbol{\mu})$ is concave. □

This dual function is important, because of the following theorem.

Theorem 2.2. *The dual function $s(\boldsymbol{\lambda}, \boldsymbol{\mu})$ is always a lower bound on the optimal value c of the optimization problem (2.1). We have*

$$s(\boldsymbol{\lambda}, \boldsymbol{\mu}) \leq c, \tag{2.13}$$

for $\mu_j \geq 0$, $j = 1, \ldots, N$, and any $\boldsymbol{\lambda}$.

Proof. We assume that $\hat{\mathbf{x}}$ is a feasible point of optimization problem (2.1). Considering the constraints, this implies that

$$\sum_{i=1}^{m} \lambda_i h_i(\hat{\mathbf{x}}) + \sum_{j=1}^{n} \mu_j g_j(\hat{\mathbf{x}}) \leq 0, \tag{2.14}$$

for any $\boldsymbol{\lambda}$ and $\mu_j \geq 0$, $j = 1, \ldots, N$. The latter constraint can also be expressed using the following shorthand notation $\boldsymbol{\mu} \geq \mathbf{0}$. The equation shown in (2.14) leads to

$$L(\hat{\mathbf{x}}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \leq f(\hat{\mathbf{x}}). \tag{2.15}$$

If we combine (2.15) with the definition of the dual function (2.11), it follows that $s(\boldsymbol{\lambda}, \boldsymbol{\mu}) \leq f(\hat{\mathbf{x}})$. Because the latter holds for all feasible points, it also holds for the optimal point $\mathbf{x}$ and thus proves (2.13). □

Now we know that the dual function $s(\boldsymbol{\lambda}, \boldsymbol{\mu})$ (for any $\boldsymbol{\mu} \geq \mathbf{0}$ and any $\boldsymbol{\lambda}$) provides a lower bound on the optimal value c of the optimization problem (2.1), we are interested in the value of $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ yielding the best (i.e., the largest) lower bound. This leads us to the following optimization problem, which is called the *Lagrange dual problem*

$$\begin{aligned} &\max_{\boldsymbol{\lambda}, \boldsymbol{\mu}} s(\boldsymbol{\lambda}, \boldsymbol{\mu}) \\ &\text{s.t. } \boldsymbol{\mu} \geq \mathbf{0}. \end{aligned} \tag{2.16}$$

Because the objective function in (2.16) is concave and the constraint is convex, the Lagrange dual problem in (2.16) will always be a convex optimization problem regardless of the properties of the original optimization problem (2.1). The optimal value of the dual problem (2.16) is denoted by d, and the point $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ which attains this value is called *dual optimal*. In a similar fashion, the original optimization problem (2.1) and its solution $\mathbf{x}$ are referred to as the *primal problem* and *primal optimal* point, respectively. The value d is the best lower bound which can be found from the dual function $s(\boldsymbol{\lambda}, \boldsymbol{\mu})$, and thus satisfies the following inequality

$$d \leq c. \tag{2.17}$$

This inequality is also called *weak duality*. An important concept that results from this is the *duality gap*, which is always positive and defined as the difference $c - d$.

When $d = c$, the duality gap is zero, as the best lower bound provided by the dual function $s(\boldsymbol{\lambda}, \boldsymbol{\mu})$ coincides with the optimal value c. In this case we say that *strong duality* holds. Strong duality does not hold in general, but if the original optimization problem is convex (2.8) there is a good chance that strong duality holds. An important result which implies strong duality is Slater's condition. Before we define Slater's condition, we first define the relative interior of a set as [4]

$$\text{relint}\mathcal{D} = \{\mathbf{x} \in \mathcal{D} | B(\mathbf{x}, r) \cap \text{aff}\mathcal{D} \subseteq \mathcal{D} \text{ for some } r > 0\}, \tag{2.18}$$

where $B(\mathbf{x}, r) = \{\mathbf{y} | \|\mathbf{y} - \mathbf{x}\| \leq r\}$ and $\text{aff}\mathcal{D}$ is the affine hull of $\mathcal{D}$, which is given by

$$\text{aff}\mathcal{D} = \{\theta_1 \mathbf{x}_1 + \ldots \theta_N \mathbf{x}_N | \mathbf{x}_1, \ldots, \mathbf{x}_N \in \mathcal{D}, \theta_1 + \ldots + \theta_N = 1\}. \tag{2.19}$$

Theorem 2.3. *Slater's condition. When the optimization problem is convex and* $\exists \mathbf{x} \in \text{relint}\mathcal{D}$ *such that*

$$h_i(\mathbf{x}) = 0, \; i = 1, \ldots, m$$

and

$$g_j(\mathbf{x}) < 0, \; j = 1, \ldots, n,$$

it follows that strong duality holds.

From now on, we will assume that strong duality holds. This means that $f(\mathbf{x}^{\text{opt}}) = s(\boldsymbol{\lambda}^{\text{opt}}, \boldsymbol{\mu}^{\text{opt}})$, where $\mathbf{x}^{\text{opt}}$ and $(\boldsymbol{\lambda}^{\text{opt}}, \boldsymbol{\mu}^{\text{opt}})$ are a primal and dual optimal point, respectively. The following can then be shown

$$\begin{aligned} f(\mathbf{x}^{\text{opt}}) &= s(\boldsymbol{\lambda}^{\text{opt}}, \boldsymbol{\mu}^{\text{opt}}) \\ &\leq f(\mathbf{x}^{\text{opt}}) + \sum_{i=1}^{m} \lambda_i^{\text{opt}} h_i(\mathbf{x}^{\text{opt}}) + \sum_{j=1}^{n} \mu_j^{\text{opt}} g_j(\mathbf{x}^{\text{opt}}) \\ &\leq f(\mathbf{x}^{\text{opt}}). \end{aligned} \tag{2.20}$$

The inequality in the second line follows from the definition of the dual function $s(\boldsymbol{\lambda}, \boldsymbol{\mu})$: the infimum of the Lagrangian will be less or equal to its value at $\mathbf{x} = \mathbf{x}^{\text{opt}}$. Finally, because $\boldsymbol{\mu}^{\text{opt}} \geq 0$, $g_j(\mathbf{x}^{\text{opt}}) \leq 0$, $j = 1, \ldots, n$, and $h_i(\mathbf{x}^{\text{opt}}) = 0$, $i = 1, \ldots, m$, the last inequality follows. As the inequalities in (2.20) are actual equalities, if follows that $\mathbf{x}^{\text{opt}}$ is a minimizer of the Lagrangian $L(\mathbf{x}, \boldsymbol{\lambda}^{\text{opt}}, \boldsymbol{\mu}^{\text{opt}})$ and that

$$\sum_{j=1}^{n} \mu_j^{\text{opt}} g_j(\mathbf{x}^{\text{opt}}) = 0. \tag{2.21}$$

The equality in (2.21) is called *complementary slackness*. From (2.21) we can immediately conclude that

$$\mu_j^{\text{opt}} > 0 \implies g_j(\mathbf{x}^{\text{opt}}) = 0, \tag{2.22}$$

and

$$g_j(\mathbf{x}^{\text{opt}}) < 0 \implies \mu_j^{\text{opt}} = 0. \tag{2.23}$$

We will now derive the conditions which hold for any primal and dual optimal point when strong duality holds, assuming that the objective and constraint functions are differentiable. Because we already know that $\mathbf{x}^{\text{opt}}$ is a minimizer of the Lagrangian $L(\mathbf{x}, \boldsymbol{\lambda}^{\text{opt}}, \boldsymbol{\mu}^{\text{opt}})$, it follows that

$$\nabla_{\mathbf{x}} L(\mathbf{x}^{\text{opt}}, \boldsymbol{\lambda}^{\text{opt}}, \boldsymbol{\mu}^{\text{opt}}) = \mathbf{0}. \tag{2.24}$$

Collecting the above results, we get that $\mathbf{x}^{\text{opt}}$ must satisfy the following equations ($i = 1, \ldots, m; j = 1, \ldots, n$):

$$h_i(\mathbf{x}^{\text{opt}}) = 0, \quad (2.25)$$

$$g_j(\mathbf{x}^{\text{opt}}) \leq 0, \quad (2.26)$$

$$\boldsymbol{\mu}^{\text{opt}} \geq \mathbf{0}, \quad (2.27)$$

$$\mu_j^{\text{opt}} g_j(\mathbf{x}^{\text{opt}}) = 0, \quad (2.28)$$

$$\nabla_{\mathbf{x}} f(\mathbf{x}^{\text{opt}}) + \sum_{i=1}^{m} \lambda_i^{\text{opt}} \nabla_{\mathbf{x}} h_i(\mathbf{x}^{\text{opt}}) + \sum_{j=1}^{n} \mu_j^{\text{opt}} \nabla_{\mathbf{x}} g_j(\mathbf{x}^{\text{opt}}) = \mathbf{0}. \quad (2.29)$$

The necessary conditions (2.25)-(2.29) are called the Karush-Kuhn-Tucker (KKT) conditions.

When the primal problem is assumed to be convex, it is now easy to show that the KKT conditions are also sufficient, i.e., if the points $\mathbf{x}$ and $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ satisfy the KKT conditions, the points are primal and dual optimal and the duality gap is zero. From these results we can conclude the following:

Theorem 2.4. *When a convex optimization problem, with differentiable objective and constraint functions, satisfies Slater's condition, the value of* $\mathbf{x}$ *is primal optimal if and only if there exists a point* $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ *which together with* $\mathbf{x}$ *satisfies the KKT conditions.*

From these results, we can conclude that it is possible to solve the primal optimization problem by solving the dual optimization problem. This can be advantageous when the dual problem is easier to solve. For example, suppose that the dual problem is solved and that we found the optimal dual point $(\boldsymbol{\lambda}^{\text{opt}}, \boldsymbol{\mu}^{\text{opt}})$. When strong duality holds and there exists a solution to the primal optimization problem, the primal optimal point is given by the value of $\mathbf{x}$ which minimizes $L(\mathbf{x}, \boldsymbol{\lambda}^{\text{opt}}, \boldsymbol{\mu}^{\text{opt}})$.

2.4 Subgradient Method

First, we will introduce the definition of a subgradient. A vector $\mathbf{q} \in \mathbb{R}^{N \times 1}$ is a subgradient of a convex function f at a point $\mathbf{x} \in \mathbb{R}^{N \times 1}$ if

$$f(\mathbf{y}) - f(\mathbf{x}) \geq \mathbf{q}^T(\mathbf{y} - \mathbf{x}) \; \forall \mathbf{y} \in \mathbb{R}^{N \times 1}. \quad (2.30)$$

When the function f is differentiable at $\mathbf{x}$, it follows that $\mathbf{q} = \nabla f(\mathbf{x})$.

Let us assume that we want to minimize a function $f(\mathbf{x})$ which is convex. If there are no constraints, we solve this optimization problem by using the following iteration

$$\mathbf{x}(k+1) = \mathbf{x}(k) - \alpha_k \mathbf{q}(k), \quad (2.31)$$

where $\mathbf{x}(k)$ is the estimate of our solution at the kth iteration, $\alpha_k > 0$ is the step size at the kth iteration and $\mathbf{q}(k)$ is a subgradient of f at $\mathbf{x}(k)$. As it very well may happen that $f(\mathbf{x}(k+1)) > f(\mathbf{x}(k))$, the algorithm will return at the kth iteration the value of $\mathbf{x}$ which reached

$$f^{\text{opt}} = \min\{f(\mathbf{x}(1)), \ldots, f(\mathbf{x}(k))\}. \tag{2.32}$$

The subgradient method can also be used to find the primal optimal point by solving the dual problem. Let us assume we want to solve optimization problem (2.8) and the duality gap is zero. If $\mathbf{x}(\boldsymbol{\lambda}, \boldsymbol{\mu})$ represents the value of $\mathbf{x}$ for which the Lagrangian (2.10) reaches its minimum for a given value of $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$, the dual function can be written as

$$s(\boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}(\boldsymbol{\lambda}, \boldsymbol{\mu})) + \sum_{i=1}^{m} \lambda_i h_i(\mathbf{x}(\boldsymbol{\lambda}, \boldsymbol{\mu})) + \sum_{j=1}^{n} \mu_j g_j(\mathbf{x}(\boldsymbol{\lambda}, \boldsymbol{\mu})). \tag{2.33}$$

Note that $\mathbf{x}(\boldsymbol{\lambda}, \boldsymbol{\mu})$ can be found by an unconstrained minimization of $L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$. Now, if the optimal value $(\boldsymbol{\lambda}^{\text{opt}}, \boldsymbol{\mu}^{\text{opt}})$ of $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ has been found, the primal optimal point is given by $\mathbf{x}(\boldsymbol{\lambda}^{\text{opt}}, \boldsymbol{\mu}^{\text{opt}})$. To solve the dual problem, we have to find a subgradient of $-s(\boldsymbol{\lambda}, \boldsymbol{\mu})$. If $s(\boldsymbol{\lambda}, \boldsymbol{\mu})$ is differentiable, the subgradient with respect to λ_i, $i = 1, \ldots, m$, is

$$\begin{aligned} -\frac{\partial s(\boldsymbol{\lambda}, \boldsymbol{\mu})}{\partial \lambda_i} &= -\left(\nabla_{\mathbf{x}} L(\mathbf{x}(\boldsymbol{\lambda}, \boldsymbol{\mu}), \boldsymbol{\lambda}, \boldsymbol{\mu})\right)^T \frac{\partial \mathbf{x}}{\partial \lambda_i} - h_i(\mathbf{x}(\boldsymbol{\lambda}, \boldsymbol{\mu})) \\ &= -h_i(\mathbf{x}(\boldsymbol{\lambda}, \boldsymbol{\mu})), \end{aligned} \tag{2.34}$$

where the last line follows from $\nabla_{\mathbf{x}} L(\mathbf{x}(\boldsymbol{\lambda}, \boldsymbol{\mu}), \boldsymbol{\lambda}, \boldsymbol{\mu}) = \mathbf{0}$ because $\mathbf{x}(\boldsymbol{\lambda}, \boldsymbol{\mu})$ is a minimizer of the Lagrangian. The same derivation can be done for $\boldsymbol{\mu}$. Thus the Lagrange multiplier vectors and $\mathbf{x}$ are updated according to the following equations

$$\begin{aligned} \boldsymbol{\lambda}(k+1) &= \boldsymbol{\lambda}(k) + \alpha_k \mathbf{h}(\mathbf{x}(k)) \\ \boldsymbol{\mu}(k+1) &= \max(\mathbf{0}, \boldsymbol{\mu}(k) + \alpha_k \mathbf{g}(\mathbf{x}(k))) \\ \mathbf{x}(k+1) &= \mathbf{x}(\boldsymbol{\lambda}(k+1), \boldsymbol{\mu}(k+1)), \end{aligned} \tag{2.35}$$

where $\max(\cdot, \cdot)$ denotes a componentwise maximum, $\mathbf{h}(\mathbf{x}) = [h_1(\mathbf{x}), \ldots, h_m(\mathbf{x})]^T$ and $\mathbf{g}(\mathbf{x}) = [g_1(\mathbf{x}), \ldots, g_n(\mathbf{x}))]^T$. The update equations of $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ in (2.35) have an intuitive interpretation. For example, let us consider $\mu_j(k)$ which corresponds to the constraint $g_j(\mathbf{x})$. If the constraint $g_j(\mathbf{x})$ is violated, which means $g_j(\mathbf{x}) > 0$, the value of μ_j will increase according to (2.35). As we can see in (2.10), a higher value of μ_j will increase the relative weight of $g_j(\mathbf{x})$ in the Lagrangian. As $\mathbf{x}(\boldsymbol{\lambda}, \boldsymbol{\mu})$ minimizes the Lagrangian, it follows that an increase of the parameter μ_j will lead to a decrease of the value $g_j(\mathbf{x})$. Eventually, this algorithm will continue until $g_j(\mathbf{x}) = 0$, which leads to $\mu_j(k+1) = \mu_j(k)$ or until

$\mu_j(k+1) = 0$ which means that the constraint $g_j(\mathbf{x})$ is inactive. For completeness, we mention that the equations in (2.35) can also be used if $s(\boldsymbol{\lambda}, \boldsymbol{\mu})$ is not differentiable.

2.5 Stochastic Subgradient Method

It can happen that the optimization problem is defined in terms of expectations of random variables. To illustrate this, we consider the following optimization problem

$$\min_{\mathbf{x}} \mathbb{E}_{\mathbf{z}} \left[f(\mathbf{x}, \mathbf{z}) \right] \quad , \tag{2.36}$$

where $\mathbf{z}$ is a random variable. If $f(\mathbf{x}, \mathbf{z})$ is convex in $\mathbf{x}$ for each value of $\mathbf{z}$ in the domain of $f(\mathbf{x}, \mathbf{z})$, the problem is a convex optimization problem. We now take a closer look at the objective function of (2.36). This function can be written as

$$\mathbb{E}_{\mathbf{z}} \left[f(\mathbf{x}, \mathbf{z}) \right] = \int f(\mathbf{x}, \mathbf{z}) p(\mathbf{z}) \, \mathrm{d}\mathbf{z}, \tag{2.37}$$

where $p(\mathbf{z})$ denotes the probability density of the random variable $\mathbf{z}$. As it can be very difficult to calculate (2.37) exactly, we will approximate the objective function (2.37) by generating L samples $\mathbf{z}_i$ according to the distribution $p(\mathbf{z})$. The approximation of the objective function (2.37) is given by

$$\mathbb{E}_{\mathbf{z}} \left[f(\mathbf{x}, \mathbf{z}) \right] \approx \frac{1}{L} \sum_{i=1}^{L} f(\mathbf{x}, \mathbf{z}_i). \tag{2.38}$$

The update equation related to (2.36) is very similar to (2.31). However, we will use a noisy unbiased subgradient as an approximation of the actual subgradient. A noisy unbiased subgradient $\tilde{\mathbf{q}}$ can be written as $\tilde{\mathbf{q}} = \mathbf{q} + \mathbf{n}$, where $\mathbf{q}$ is a subgradient and $\mathbb{E}[\mathbf{n}] = \mathbf{0}$. In order to show how we can find the noisy unbiased subgradient for the update equations related to (2.36), we will first introduce the function $\mathbf{k}(\mathbf{x}, \mathbf{z})$ as a subgradient of $f(\mathbf{x}, \mathbf{z})$ with respect to $\mathbf{x}$ for $\forall \mathbf{x}, \mathbf{z} \in \mathbb{R}^{N \times 1}$. It is easy to show that $\mathbb{E}_{\mathbf{z}} \left[\mathbf{k}(\mathbf{x}, \mathbf{z}) \right]$ is a subgradient of the objective function $\mathbb{E}_{\mathbf{z}} \left[f(\mathbf{x}, \mathbf{z}) \right]$. This means that the following expression

$$\frac{1}{L} \sum_{i=1}^{L} \mathbf{k}(\mathbf{x}, \mathbf{z}_i), \tag{2.39}$$

is a noisy unbiased subgradient of $\mathbb{E}_{\mathbf{z}} \left[f(\mathbf{x}, \mathbf{z}) \right]$ at $\mathbf{x}$. This result is valid for every value of L. It is even possible to take $L = 1$, as the iterations of the subgradient method will help to 'average out' the statistical fluctuations.

2.6 CVX

A useful tool we would like to mention is CVX. CVX is a modeling system for convex optimization which is available for free on the web. It is implemented in Matlab and allows to formulate and solve convex optimization problems. For example, consider the following optimization problem where $\mathbf{x}$ is constrained by a componentwise inequality:

$$\begin{aligned} &\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\| \\ &\text{s.t.}\, \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}, \end{aligned} \tag{2.40}$$

where $\mathbf{A} \in \mathbb{R}^{N \times N}$ and $\mathbf{l}, \mathbf{u}, \mathbf{x}, \mathbf{b} \in \mathbb{R}^{N \times 1}$. This problem can be specified in CVX as

```
cvx_begin
     variable x(N)
     minimize (norm(A*x-b))
     subject to
         l<=x<=u
cvx_end
```

When Matlab has run these commands, the variable $\mathbf{x}$ will contain the optimal solution to the optimization problem. This tool makes it very easy to describe and solve convex problems in Matlab. For completeness, we also mention that this tool can be used to find the optimal dual variables of the optimization problem. For more information about CVX, we refer to [5].

2.7 Optimization Problems with Complex Variables

Until now, we have assumed that the variables in the optimization problem are real-valued. However, this restriction is not necessary as there are several ways to solve optimization problems with complex variables. This section gives the reader a short overview of these methods. The first and most straightforward method, is to rewrite a complex optimization problem as a real optimization problem by decomposing the complex variables in their real and imaginary parts. For example, if we take the following objective function

$$\mathbf{z}^H \mathbf{A} \mathbf{z} + 2\Re(\mathbf{b}^H \mathbf{z}) + c, \tag{2.41}$$

where $\mathbf{A} \in \mathbb{C}^{N \times N}$ is a Hermitian matrix (i.e., $\mathbf{A}^H = \mathbf{A}$), $\mathbf{b} \in \mathbb{C}^{N \times 1}$ and $c \in \mathbb{R}$. By decomposing the complex vector $\mathbf{z}$ as $\mathbf{z} = \mathbf{x} + j\mathbf{y}$, we can rewrite (2.41) as

$$\mathbf{v}^T \mathbf{B} \mathbf{v} + 2\mathbf{d}^T \mathbf{v} + c, \tag{2.42}$$

where

$$\mathbf{v} = \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}, \tag{2.43}$$

$$\mathbf{B} = \begin{pmatrix} \Re(\mathbf{A}) & -\Im(\mathbf{A}) \\ \Im(\mathbf{A}) & \Re(\mathbf{A}) \end{pmatrix}, \tag{2.44}$$

$$\mathbf{d} = \begin{pmatrix} \Re(\mathbf{b}) \\ \Im(\mathbf{b}) \end{pmatrix}. \tag{2.45}$$

However, it will not always be necessary to rewrite the complex optimization problems as a real optimization problem. As most optimization problems are solved by calculating derivatives, we will show how we can calculate a derivative of a general complex-valued function with complex variables. If

$$f(z) = u(x, y) + jv(x, y), \tag{2.46}$$

represents a general complex-valued function with $z = x + jy$, where $u(x, y)$ and $v(x, y)$ are real-valued functions, the complex derivative is defined as

$$f'(z) = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}. \tag{2.47}$$

This derivative is well-defined only if it is independent of the direction in which Δz approaches zero in the complex plane; this is only the case when the functions $u(x, y)$ and $v(x, y)$ satisfy the Cauchy-Riemann equations. These equations are given by

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}. \tag{2.48}$$

If these equations are met we say that the function $f(z)$ is *holomorphic*, in which case the complex derivative is well-defined.

However, we are often interested in the optimization of a *real-valued* function $f(z)$ with complex variables. As $v(x, y)$ will be 0 for these functions, it follows that any non-constant function $f(z)$ will not satisfy the Cauchy-Riemann condition (2.48). This means that a non-constant real-valued function $f(z)$ is not holomorphic and thus will not be differentiable. So it seems that in order to minimize, for example, the function $f(z) = |z|^2$, we first have to rewrite $f(z)$ as a function of x and y as follows

$$\begin{aligned} f(z) &= f(z(x, y)) = (x + jy)(x - jy) \\ &= x^2 + y^2. \end{aligned} \tag{2.49}$$

In the second step, we then solve the following equations

$$\frac{\partial f(z)}{\partial x} = 2x = 0, \tag{2.50}$$

$$\frac{\partial f(z)}{\partial y} = 2y = 0. \tag{2.51}$$

However, this approach can quickly become tedious. It would be much more convenient, if we could keep writing $f(z)$ as a function of the complex variable z.

For this reason, we will introduce a 'trick'. For this trick, we view a possibly non-holomorphic function as a function of both z and its conjugate z^*. The Cauchy-Riemann condition (2.48) does not have to be satisfied, but it is required that $f(z)$ is a differentiable function of the real part x and imaginary part y of z. This *real-derivative* can be seen as an extension of the complex derivative, and will reduce to the complex derivative when applied to holomorphic functions. The following pair of partial derivatives can now be formally defined as follows

$$\text{real-derivative of } f(z,z^*) = \left.\frac{\partial f(z,z^*)}{\partial z}\right|_{z^*=\text{const.}}, \quad (2.52)$$

$$\text{conjugate real-derivative of } f(z,z^*) = \left.\frac{\partial f(z,z^*)}{\partial z^*}\right|_{z=\text{const.}}. \quad (2.53)$$

These definitions can also be equivalently written as

$$\frac{\partial f}{\partial z} = \frac{1}{2}\left(\frac{\partial f}{\partial x} - j\frac{\partial f}{\partial y}\right), \quad (2.54)$$

$$\frac{\partial f}{\partial z^*} = \frac{1}{2}\left(\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y}\right). \quad (2.55)$$

We verify these formulas by using the following example. According to (2.52) the real-derivative of $f(z,z^*) = |z|^2 = zz^*$ can be calculated as follows

$$\begin{aligned}\frac{\partial f}{\partial z} &= \frac{1}{2}\left(\frac{\partial f}{\partial x} - j\frac{\partial f}{\partial y}\right)\\ &= \frac{1}{2}\left(\frac{\partial (x^2+y^2)}{\partial x} - j\frac{\partial (x^2+y^2)}{\partial y}\right)\\ &= x - jy\\ &= z^*. \end{aligned} \quad (2.56)$$

However, it is much easier to consider z and z^* as two separate variables. In this case, we can get the same result as follows

$$\begin{aligned}\frac{\partial f}{\partial z} &= z^*\frac{\partial z}{\partial z}\\ &= z^*. \end{aligned} \quad (2.57)$$

We can see that (2.50)-(2.51) are equivalent with putting (2.56) equal to 0. We also note that the Cauchy-Riemann condition (2.48) is equivalent with $\frac{\partial f}{\partial z^*} = 0$. Further, it follows that the real-derivative of a holomorphic function $f(z)$ equals the complex derivative $f'(z)$ in (2.47). In the following, we will assume that the function f is a real-valued function.

In some cases, the function f is dependent on a complex matrix $\mathbf{S} \in \mathbb{C}^{M\times N}$. In this case we will use the following definition

$$\frac{\partial f}{\partial \mathbf{S}} = \begin{pmatrix} \frac{\partial f}{\partial s_{1,1}} & \cdots & \frac{\partial f}{\partial s_{1,N}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial s_{M,1}} & \cdots & \frac{\partial f}{\partial s_{M,N}} \end{pmatrix}, \tag{2.58}$$

where $s_{i,j}$ represents the element on the *i*th row and the *j*th column of $\mathbf{S}$. However, in most cases we will deal with a square complex matrix $\mathbf{S} \in \mathbb{C}^{N\times N}$ which is Hermitian, thus $\mathbf{S}^H = \mathbf{S}$. The matrix $\mathbf{S}$ can be written as

$$\mathbf{S} = \begin{pmatrix} \frac{z_{1,1}+z_{1,1}^*}{2} & \cdots & z_{1,N} \\ \vdots & \ddots & \vdots \\ z_{1,N}^* & \cdots & \frac{z_{N,N}+z_{N,N}^*}{2} \end{pmatrix}. \tag{2.59}$$

As some of the elements of $\mathbf{S}$ are related to each other, the derivative of a function $f(\mathbf{S})$ has to be calculated by using the following chain rule

$$\frac{\partial f(\mathbf{S})}{\partial z_{i,j}} = \mathrm{Tr}\left(\left(\frac{\partial f(\mathbf{S})}{\partial \mathbf{S}}\right)^T \frac{\partial \mathbf{S}}{\partial z_{i,j}}\right), \tag{2.60}$$

where $\frac{\partial f(\mathbf{S})}{\partial \mathbf{S}}$ is given by (2.58) and

$$\frac{\partial \mathbf{S}}{\partial z_{i,j}} = \begin{cases} \mathbf{J}_{i,j}, & \text{if } i < j \\ \frac{1}{2}\mathbf{J}_{i,i}, & \text{if } i = j \\ \mathbf{0}, & \text{if } i > j \end{cases}, \tag{2.61}$$

where $\mathbf{J}_{i,j}$ denotes the single-entry matrix: 1 at $(i,\ j)$ and 0 elsewhere. The derivatives with respect to $z_{i,j}^*$ can be calculated as

$$\frac{\partial f(\mathbf{S})}{\partial z_{i,j}^*} = \mathrm{Tr}\left(\left(\frac{\partial f(\mathbf{S})}{\partial \mathbf{S}}\right)^T \frac{\partial \mathbf{S}}{\partial z_{i,j}^*}\right), \tag{2.62}$$

where

$$\frac{\partial \mathbf{S}}{\partial z_{i,j}^*} = \begin{cases} \mathbf{J}_{j,i}, & \text{if } i < j \\ \frac{1}{2}\mathbf{J}_{i,i}, & \text{if } i = j \\ \mathbf{0}, & \text{if } i > j \end{cases}. \tag{2.63}$$

The optimum of the function $f(\mathbf{S})$ will satisfy

$$\frac{\partial f(\mathbf{S})}{\partial z_{i,j}} = 0 = \frac{\partial f(\mathbf{S})}{\partial z_{i,j}^*} \quad i = 1,\ldots,N; j = 1,\ldots,N. \tag{2.64}$$

However, this is equivalent to setting equal to $\mathbf{0}$ the following Hermitian matrix

$$\begin{pmatrix} \frac{\partial f(\mathbf{S})}{\partial z_{1,1}} + \frac{\partial f(\mathbf{S})}{\partial z_{1,1}^*} & \cdots & \frac{\partial f(\mathbf{S})}{\partial z_{1,N}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(\mathbf{S})}{\partial z_{1,N}^*} & \cdots & \frac{\partial f(\mathbf{S})}{\partial z_{N,N}} + \frac{\partial f(\mathbf{S})}{\partial z_{N,N}^*} \end{pmatrix}. \tag{2.65}$$

By using (2.60) and (2.62), it follows that (2.65) equals $\frac{\partial f}{\partial \mathbf{S}}$, which means that the optimum of $f(\mathbf{S})$ can also be found by the following equivalent equation

$$\frac{\partial f}{\partial \mathbf{S}} = \mathbf{0}. \tag{2.66}$$

For completeness we show the relationships between the elements of $\frac{\partial f}{\partial \mathbf{S}}$ and the derivatives of f with respect to the real and imaginary parts of $z_{i,j}$, $i = 1, \ldots, N$, $j = 1 \ldots, N$. The following equalities can be derived

$$\left(\frac{\partial f}{\partial \mathbf{S}}\right)_{i,j} = \begin{cases} \frac{\partial f}{\partial \Re(z_{i,i})}, & \text{if } i = j \\ \frac{1}{2}\left(\frac{\partial f}{\partial \Re(z_{i,j})} - j\frac{\partial f}{\partial \Im(z_{i,j})}\right), & \text{if } i < j \\ \frac{1}{2}\left(\frac{\partial f}{\partial \Re(z_{j,i})} + j\frac{\partial f}{\partial \Im(z_{j,i})}\right), & \text{if } i > j \end{cases}. \tag{2.67}$$

In order to illustrate this result, we will calculate the derivative of

$$\frac{\partial \text{Tr}\,(\mathbf{S})}{\partial \mathbf{S}} = \mathbf{I}. \tag{2.68}$$

It is easily verified that this result satisfies equation (2.67).

A more involved example is the following

$$\begin{aligned} \frac{\partial \|\mathbf{AS}\|}{\partial \mathbf{S}} &= \frac{\partial \sqrt{\text{Tr}\,(\mathbf{A}\mathbf{S}\mathbf{S}^H\mathbf{A}^H)}}{\partial \mathbf{S}} \\ &= \frac{\partial \sqrt{\text{Tr}\,(\mathbf{A}\mathbf{S}\mathbf{S}\mathbf{A}^H)}}{\partial \mathbf{S}} \\ &= \frac{1}{2\|\mathbf{AS}\|}\left(\mathbf{A}^T\mathbf{A}^*\mathbf{S}^* + \left(\mathbf{A}^T\mathbf{A}^*\mathbf{S}^*\right)^H\right). \end{aligned} \tag{2.69}$$

Again, it can be verified that this result indeed satisfies the relations shown in (2.67).

2.8 Chapter Summary

In this chapter, we have laid the foundation of some very important concepts in optimization theory. Some basic numerical methods have also been introduced. In this dissertation, we will often try to optimize the performance of a communication

network as a function of its transmission parameters. However, the network will have several limitations or restrictions, such as the maximal allowable energy that can be transmitted. As these limitations can be translated into several constraints for the transmit parameters, we try to formulate these problems as constrained optimization problems. This means that the following chapters will rely heavily on the theorems and results that we introduced here.

3

Digital Communications Basics

This chapter introduces the wireless channel and the system model of the communication networks that we will be investigating. Section 3.1 discusses the wireless channel and introduces several channel models that will be used in the subsequent chapters. Section 3.2 explains some of the major building blocks that are used in a digital communication system. Section 3.3 introduces the minimum mean square error (MMSE) estimator, which will be of use in subsequent chapters. In section 3.4, we introduce several metrics that can be used to assess the performance of a network. We then show in section 3.5 how the transmitter can adapt its transmission parameters to optimize these performance metrics. Section 3.6 gives a brief overview of several popular relaying protocols for cooperative networks. Finally, section 3.7 presents the three main paradigms of cognitive radio: the underlay, overlay and interweave paradigm.

3.1 The Wireless Channel

In wireless communications, a transmitter (TX) sends a signal to a receiver (RX) through a wireless channel. Any motion of the TX, RX or surrounding obstacles causes random fluctuations in the channel due to the changing reflections and attenuation of the transmitted signal. The properties of the channel thus change randomly with time, which makes the design of a reliable communication system very hard. Because there are many factors which influence the channel properties,

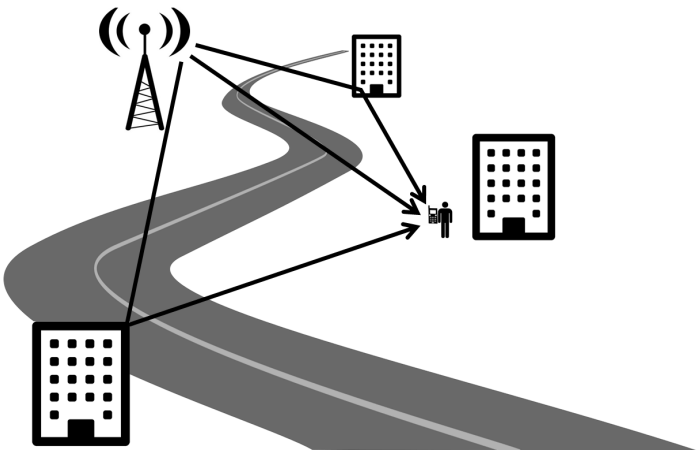


Figure 3.1: Multipath channel between TX (base station) and RX (mobile user).

we use a statistical model to characterize the varying channel. In the following, we provide a description of this statistical model, which is mainly based on [11].

We consider the transmission of a bandpass signal $s_{\rm BP}(t)$, which is represented as $s_{\rm BP}(t) = \sqrt{2}\Re\left[s(t)e^{j2\pi f_{\rm c}t}\right]$, where $s(t)$ and $f_{\rm c}$ denote the complex baseband signal and the carrier frequency, respectively. Denoting by $S_s(f)$ the power spectral density of $s(t)$, we assume that $S_s(f) = 0$ for $|f| > B$, with B the bandwidth of $s(t)$. The corresponding bandpass signal at the output of the wireless channel is denoted $r_{\rm BP}(t)$, with $r_{\rm BP}(t) = \sqrt{2}\Re\left[r(t)e^{j2\pi f_{\rm c}t}\right]$. The complex baseband signals $r(t)$ and $s(t)$ are related by $r(t) = \int h(\tau,t)s(t-\tau){\rm d}\tau$ where $h(\tau,t)$ is the impulse response of the time-varying wireless channel. The impulse response is modeled as

$$h(\tau,t) = \sum_{l=1}^{L} c_l(t)\delta(\tau-\tau_l), \tag{3.1}$$

which describes a multipath channel consisting of L paths: the lth path is characterized by the time-varying complex-valued gain $c_l(t)$ (representing amplitude $|c_l(t)|$ and phase $\angle c_l(t)$) and a delay τ_l. The resulting $r(t)$ is given by

$$r(t) = \sum_{l=1}^{L} c_l(t)s(t-\tau_l), \tag{3.2}$$

which indicates that $r(t)$ is a linear combination (with time-varying complex coefficients $c_l(t)$) of delayed versions of $s(t)$. The multipath channel is illustrated in Fig. 3.1, which shows that due to reflections, the transmitted signal arrives at the RX through different paths.

Assuming wide-sense stationary uncorrelated scattering (WSSUS), the path gains $c_l(t)$ are uncorrelated stationary Gaussian processes.

In the absence of a line-of-sight component, all $c_l(t)$ are considered zero-mean circularly symmetric. The time variations of $c_l(t)$ are characterized by the autocorrelation function $R_l(\tau) = \mathbb{E}\left[c_l(t+\tau)c_l^*(t)\right]$. The power spectrum of $c_l(t)$ is denoted $S_l(f)$, which equals the Fourier transform of $R_l(\tau)$; it can be shown that $S_l(f) = 0$ for $|f| > f_{\rm d}$, where $f_{\rm d}$ is the Doppler frequency of the channel. For a RX moving at speed v, we have $f_{\rm d} = (v/c)f_{\rm c}$ where $c \approx 3 \cdot 10^8$ m/s denotes the speed of light.

3.1.1 Flat/Selective Fading

Assuming that in (3.1) the path delays are ordered such that $\tau_1 < \tau_2 < ... < \tau_L$, one defines the delay spread $\tau_{\max}$ as $\tau_{\max} = \tau_L - \tau_1$, which denotes the delay difference between the last arriving and the first arriving path. The inverse of the delay spread is referred to as the channel coherence bandwidth $B_{\rm coh}$: $B_{\rm coh} = 1/\tau_{\max}$

The effect of the multipath channel on the received signal depends on the ratio $B/B_{\rm coh}$ of the signal bandwidth B to the coherence bandwidth.

- When $B/B_{\rm coh} \ll 1$, we have $s(t-\tau_l) \approx s(t-\tau_1)$ for $l \neq 1$, so that (3.2) reduces to $r(t) = c(t)s(t-\tau_1)$, with $c(t) = \sum_{l=1}^{L} c_l(t)$. The channel is characterized by a single delay τ_1 and a single time-varying gain $c(t)$. The channel is referred to as *frequency-flat*, because all frequency components of $s(t)$ are affected by the channel in the same way: they all experience a gain of magnitude $|c(t)|$. In the absence of a line-of-sight component, $c(t)$ is stationary and zero-mean, with autocorrelation function $R(\tau) = \sum_{l=1}^{L} R_l(\tau)$. One often uses Jakes' model [12], which takes $R(\tau) = R(0)J_0(2\pi f_{\rm d}\tau)$, where $J_0(.)$ is the zero-order Bessel function of the first kind; this correlation function corresponds to the case of a sum of equal-strength paths with angles of arrival uniformly distributed in the interval $(0, 2\pi)$. For given t, the magnitude $z = |c(t)|$ has a Rayleigh distribution, with a probability density function $p_z(x)$ given by

$$p_z(x) = \frac{2x}{R(0)} e^{-\frac{x^2}{R(0)}}, \quad x \geq 0. \tag{3.3}$$

- When $B/B_{\rm coh}$ is in the order of 1 or larger, the differences between path delays can no longer be ignored. The channel is referred to as *frequency-selective*: the channel gain magnitude varies over the signal bandwidth, causing linear distortion to the signal; the gain corresponding to a frequency f has a magnitude $\left|\sum_{l=1}^{L} c_l(t)e^{-j2\pi f\tau_l}\right|$. The gain magnitudes correspond-

ing to the frequencies f_1 and f_2 are essentially the same when $|f_1 - f_2| \ll B_{\mathrm{coh}}$.

3.1.2 Block-Fading Channel

The channel coherence time T_{coh} is a measure of the time interval over which the channel can be considered as time-invariant: when $|t_1 - t_2| \ll T_{\mathrm{coh}}$, we have $h(\tau, t_1) \approx h(\tau, t_2)$. Considering that the path gains $c_l(t)$ have a bandwidth not exceeding the Doppler frequency f_{d}, it is convenient to take $T_{\mathrm{coh}} = 1/f_{\mathrm{d}}$.

Digital information is typically transmitted in blocks, often referred to as packets or frames. In this work we will consider the case where the duration of a frame is much smaller than the channel coherence time, so that the path gains $c_l(t)$ can be considered constant over a frame; this corresponds to a block-fading channel.

3.1.3 Standardized Channel Models

In the simulations, we will make use of different channel models to validate our results, depending on whether the channel is frequency-flat or frequency-selective.

For a frequency-flat channel, we will generate a complex-valued Gaussian channel gain with zero mean (Rayleigh fading). The mean-squared magnitude of the channel gain is set equal to $1/d^{\alpha}$, where d represents the distance between the TX and RX, and α represents the path loss exponent ($\alpha = 2$ in free space).

When simulating a frequency-selective channel, we will use a standardized model for multipath channels [13], which is characterized by the number of paths, the corresponding path delays and the relative power $R_n(0)/R_1(0)$ of the *n*th path with respect to the first path. Table 3.1 shows the parameters of the ITU Pedestrian B and the ITU Vehicular A channel models. The former model describes an environment that is characterized by small cells and low transmit power; it assumes that base stations with small antenna heights are located outdoors, while pedestrian users are located on streets and inside buildings and residences. The latter model describes an environment that is characterized by large cells and high transmit power.

3.2 Digital Communication System Model

In the following we give an overview of the main building blocks that are used in a typical digital communications system. These blocks are:

- channel coding
- symbol mapping
- single-carrier or multicarrier modulation

	ITU Pedestrian B		ITU Vehicular A	
	Path Delay (ns)	Relative power (dB)	Path Delay (ns)	Relative power (dB)
1	0	0.0	0	0.0
2	200	−0.9	310	−1.0
3	800	−4.9	710	−9.0
4	1200	−8.0	1090	−10.0
5	2300	−7.8	1730	−15.0
6	3700	−23.9	2510	−20.0

Table 3.1: ITU Pedestrian B (left) and ITU Vehicular A (right) channel model.

3.2.1 Channel Coding

Channel coding is a means to protect the information bit stream against errors. The channel encoder adds redundant bits to the information bit stream, which enables the RX to correct some bit error patterns. Considering the information bitstream consisting of N_{u} bits, and denoting the length of the corresponding coded bit vector $\mathbf{c} = (c_1, ..., c_{N_{\mathrm{c}}})$ by N_{c}, with $N_{\mathrm{c}} > N_{\mathrm{u}}$, the code rate is defined as the ratio $N_{\mathrm{u}}/N_{\mathrm{c}}$.

3.2.2 Symbol Mapping

The symbol mapping transforms the stream of coded bits $\mathbf{c}$ into a stream of complex-valued data symbols $\mathbf{x} = (x_1, ..., x_N)$. First, the stream of coded bits $\mathbf{c}$ is divided into groups of m bits, and each group is mapped to a data symbol. These symbols are selected from a normalized M-point constellation $\chi = \{\alpha_1, \ldots, \alpha_M\}$, where

$$\frac{1}{M}\sum_{i=1}^{M}|\alpha_i|^2 = 1, \tag{3.4}$$

expresses the normalization condition, and the number of constellation points M is given by

$$M = 2^m. \tag{3.5}$$

The number (N) of data symbols is related to the number (N_{c}) of coded bits by $N = N_{\mathrm{c}}/m$. We will now give an overview of some popular constellations:

- **M-ary Pulse Amplitude Modulation** (M-PAM): The set χ is given by

$$\chi = \{(2k - 1 - M)d(M) : k \in \{1, \ldots, M\}\}, \tag{3.6}$$

where

$$d(M) = \sqrt{\frac{3}{M^2 - 1}}, \tag{3.7}$$

ensures the normalization. When $M = 2$, this constellation is also called binary phase-shift keying (BPSK).

- **M-ary Quadrature Amplitude Modulation** (M-QAM): This constellation consists of the following set of complex symbols

$$\begin{aligned}\chi &= \{(2k-1-\sqrt{M})\frac{d(\sqrt{M})}{\sqrt{2}} \\ &\quad +j(2l-1-\sqrt{M})\frac{d(\sqrt{M})}{\sqrt{2}} : k,l \in \{1,\ldots,\sqrt{M}\}\},\end{aligned} \quad (3.8)$$

  which combines two $\sqrt{M}$-PAM constellations, one for the real part and one for the imaginary part of the M-QAM constellation. Note that m, the number of bits per constellation, has to be even.

- **M-ary Phase-Shift Keying** (M-PSK): In this case, the complex symbols belong to the following set

$$\chi = \{e^{j\frac{2\pi(k-1)}{M}} : k \in \{1,\ldots,M\}\}. \quad (3.9)$$

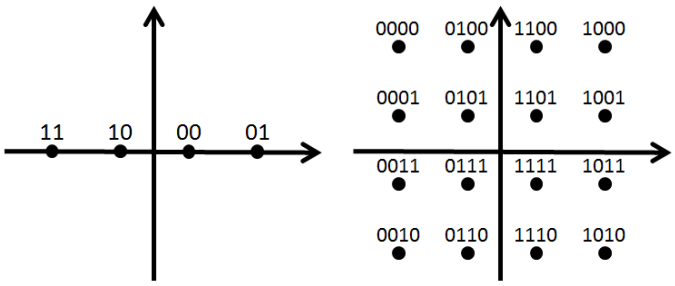
These constellations are illustrated in Fig. 3.2, for the cases of 4-PAM, 16-QAM and 8-PSK. Note that Fig. 3.2 also shows a possible mapping between the different sets of m bits and the constellation points. We have used a *Gray mapping* function. This mapping function has the important property that the constellation points at minimum Euclidean distance differ only in a single bit. This mapping function has the advantage of minimizing the bit error rate for uncoded transmission.

3.2.3 Single-Carrier and Multicarrier Modulation

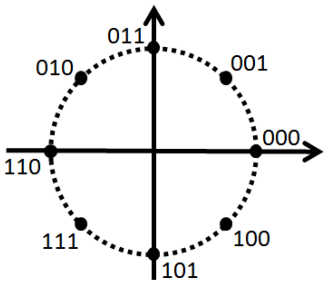
In this work, we consider linear digital modulation. This means that we transmit a sequence $\mathbf{s}$, which is a linear function of the data symbol sequence $\mathbf{x}$. We will consider single-carrier modulation in the case of a frequency-flat channel, and multicarrier modulation in the case of a frequency-selective channel. For both types of modulation, we describe the signal processing operations to be performed at the TX and the RX. We will derive a simple discrete-time model which relates the transmitted data symbols $\mathbf{x}$ to some vector $\mathbf{z}$ which is computed by the RX; in subsequent chapters, this model will be used without further specification of the underlying signal processing operations.

3.2.3.1 Single-Carrier Modulation

In the case of single-carrier modulation, the sample sequence $\mathbf{s}$ is the same as the data symbol sequence $\mathbf{x}$. The transmitted signal corresponding to $\mathbf{x} = [x_1,\ldots,x_N]^T \in$



(a) 4-PAM **(b)** 16-QAM



(c) 8-PSK

Figure 3.2: Examples of several constellations.

χ^N is given by

$$s(t) = \sqrt{E_\mathrm{s}} \sum_{k=1}^{N} x_k p(t - kT), \tag{3.10}$$

where T is the symbol interval (which corresponds to a symbol rate $R_s = 1/T$), and the baseband transmit pulse $p(t)$ satisfies

$$\int_{-\infty}^{+\infty} |p(t)|^2 \, \mathrm{d}t = 1. \tag{3.11}$$

This way, E_s denotes the transmitted energy per data symbol. The pulse $p(t)$ has a bandwidth B, which means that the frequency spectrum $P(f)$ of $p(t)$ is zero for $|f| > B$. The filter $p(t)$ is selected such that $g(t)$, given by

$$g(t) = \int p(t+u) p^*(u) \, \mathrm{d}u, \tag{3.12}$$

is a Nyquist pulse. This means that the following condition holds

$$g(nT) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}. \tag{3.13}$$

Note from (3.12) that $g(t) = g^*(-t)$.

The signal $s(t)$ is sent over a frequency-flat channel. Assuming that the duration of the signal $s(t)$ is much less than the channel coherence time, the received signal can be represented as

$$r(t) = hs(t - \tau_1) + n(t), \tag{3.14}$$

where h is the complex channel gain, which is constant over the duration of $s(t)$, τ_1 is the channel delay, and $n(t)$ is zero-mean complex-valued circularly symmetric additive white Gaussian noise (AWGN) with power spectral density σ^2, which will be denoted as $n(t) \sim N_\mathrm{c}(0, \sigma^2 \delta(u))$.

The received signal $r(t)$ is applied to a filter with impulse response $p^*(-t)$, which is matched to the transmit filter. The matched filter output signal $z(t)$ is sampled at instants $kT + \tau_1$. The resulting samples $z_k = z(kT + \tau_1)$ are given by

$$z_k = \sqrt{E_\mathrm{s}} h x_k + n_k, \tag{3.15}$$

where n_k is a zero-mean complex-valued circularly symmetric Gaussian noise contribution with $\mathbb{E}\left[n_{k+i} n_k^*\right] = \sigma^2 \delta_i$. Dividing both sides of (3.15) by $\sqrt{E_\mathrm{s}} h$, in order to reduce the coefficient of x_k to one, yields the simple observation model

$$y_k = x_k + w_k, \tag{3.16}$$

where w_k is a zero-mean complex-valued circularly symmetric Gaussian noise contribution with $\mathbb{E}\left[w_{k+i} w_k^*\right] = \frac{1}{\gamma} \delta_i$, with $\gamma = E_\mathrm{s} |h|^2 / \sigma^2$ denoting the signal-to-noise ratio (SNR). The RX tries to recover the transmitted information bits from the samples $[y_1, ..., y_N]$.

3.2.3.2 Multicarrier Modulation

In a multicarrier modulation, or more specifically Orthogonal Frequency-Division Multiplexing (OFDM), the symbol vector $\mathbf{x} = [x_1, \ldots, x_N]^T \in \chi^N$ is transformed into the sequence $\mathbf{s}' = [s'_1, \ldots, s'_N]^T$ according to

$$\mathbf{s}' = \frac{1}{\sqrt{N}} \mathbf{F}^H \mathbf{x}, \tag{3.17}$$

where $\mathbf{F} \in \mathbb{C}^{N \times N}$ denotes the Fourier matrix, which is defined as

$$\mathbf{F}_{k,l} = e^{-j2\pi \frac{(k-1)(l-1)}{N}}, \quad k = 1, \ldots, N; l = 1, \ldots, N. \tag{3.18}$$

In order to avoid that the frequency-selective channel causes interference between the data symbols, we construct the vector $\mathbf{s}$ by inserting a cyclic prefix of ν samples in front of the vector $\mathbf{s}'$; this yields

$$\mathbf{s} = \sqrt{\frac{N}{N+\nu}} \begin{pmatrix} s'_{N-\nu+1} \\ \vdots \\ s'_N \\ \mathbf{s}' \end{pmatrix}. \tag{3.19}$$

The contribution of the symbol x_n to s_k consists of the sample of the signal $x_n e^{j2\pi f_n t}$, taken at instant $(k-\nu-1)T$, for $k = 1, ..., N+\nu$, where $f_n = (n-1)/(NT)$, for $n = 1, ..., N$. Hence, the components of $\mathbf{s}$ are the $N+\nu$ samples of the sum of N subcarriers, where the *n*th subcarrier is located at frequency f_n and conveys the data symbol x_n. As the spacing between subcarrier frequencies equals $1/(NT)$, the sampled subcarriers are mutually orthogonal over an interval of N samples.

The $N+\nu$ components of the vector $\mathbf{s}$ are transmitted using the signal $s(t)$, given by

$$s(t) = \sqrt{E_\mathrm{s}} \sum_{k=1}^{N+\nu} s_k p(t - kT), \tag{3.20}$$

where $p(t)$ has the same properties as in (3.10). Considering that we need $N+\nu$ samples $[s_1, ..., s_{N+\nu}]$ to transmit N data symbols $[x_1, ..., x_N]$, the symbol rate is given by $R_s = \frac{N}{N+\nu} \cdot \frac{1}{T}$; because of the scaling factor $\sqrt{\frac{N}{N+\nu}}$ in (3.19), E_s in (3.20) denotes the transmitted energy per data symbol.

The signal $s(t)$ is transmitted over a frequency-selective channel. Assuming that the duration of the transmitted signal $s(t)$ is much smaller than the channel coherence time, the path gains can be considered constant during the transmission of $s(t)$. The resulting received signal is given by

$$r(t) = \sum_{l=1}^{L} c_l s(t - \tau_l) + n(t), \tag{3.21}$$

where $n(t)$ is complex-valued AWGN with power spectral density σ^2.

At the RX, the signal $r(t)$ is applied to a matched filter with impulse response $p^*(-t)$, and the resulting filter output signal $v(t)$ is sampled at rate $1/T$. Let us denote by $h(t)$ the impulse response of the cascade of the transmit filter, the channel and the receive filter, i.e., $h(t) = \sum_{l=1}^{L} c_l g(t - \tau_l)$; defining T_h and t_h as the duration and the starting point of $h(t)$, respectively, we have $T_h = \tau_{\max} + T_g$, and $t_h = \tau_1 - (T_g/2)$, where T_g is the duration of the pulse $g(t)$. Introducing $h_m = h(t_0 + mT)$, where t_0 is a sampling instant between t_h and $t_h + T$, we have $h_m = 0$ for $m < 0$ or $m > L_h - 1$, with $L_h = \text{ceil}(T_h/T)$. The RX stacks the matched filter output samples $v_k = v(t_0 + \nu T + kT)$ with $k = 1, ..., N$, into the vector $\mathbf{v} = [v_1, \ldots, v_N]^T$, and computes the vector $\mathbf{z}$ as

$$\mathbf{z} = \sqrt{\frac{N+\nu}{N}} \frac{1}{\sqrt{N}} \mathbf{F}\mathbf{v}, \tag{3.22}$$

where $\mathbf{F}$ is the Fourier matrix. Note that the samples $v(t_0 + kT)$ with $k = 1, ..., \nu$, which correspond to the cyclic prefix, are ignored by the RX. Provided that $\nu \geq L_h - 1$, it is shown in [14] that the contribution of the vector $\mathbf{s}'$ to $\mathbf{v}$ can be expressed as $\sqrt{\frac{N}{N+\nu}}\mathbf{H}\mathbf{s}'$, where $\mathbf{H}$ is a $N \times N$ cyclic matrix: the first column of $\mathbf{H}$ equals $[h_0, ..., h_{L_h-1}, 0, ..., 0]^T$, and the $(n+1)$th column is a downward cyclic shift of the *n*th column. Because of the cyclic property of $\mathbf{H}$, it can be verified that

$$z_k = \sqrt{E_\text{s}} H_k x_k + n_k, \quad k = 1, \ldots, N, \tag{3.23}$$

where $H_k = H(e^{j2\pi f_k T})$ with $H(e^{j2\pi fT}) = \sum_{m=0}^{L_h-1} h_m e^{-j2\pi fmT}$ denoting the discrete-time Fourier transform of the impulse response samples $[h_0, ..., h_{L_h-1}]$, and n_k is a zero-mean complex-valued circularly symmetric Gaussian noise contribution with $\mathbb{E}[n_{k+i} n_k^*] = \sigma_w^2 \delta_i$, and $\sigma_w^2 = \frac{N+\nu}{N}\sigma^2$. Applying single-tap equalization involves dividing both sides of (3.23) by $\sqrt{E_\text{s}} H_k$; this again yields (3.16), but now with $\mathbb{E}[w_{k+i} w_k^*] = \frac{1}{\gamma_k}\delta_i$, where $\gamma_k = \frac{E_\text{s}|H_k|^2}{\sigma_w^2}$ denotes the SNR on the *k*th subcarrier. The RX tries to recover the transmitted information bits from the samples $[y_1, ..., y_N]$.

We note that a more generic model will be considered in later chapters, which allows every subcarrier to have a different transmit energy and constellation size. The resulting system model is obtained by slightly modifying (3.23), in such a way that both the constellation χ and transmit energy E_s can become dependent on the subcarrier index k.

Multicarrier modulation on a frequency-selective channel and single-carrier modulation on a frequency-flat channel give rise to a similar relation between $\mathbf{y}$ and $\mathbf{x}$, provided that the multicarrier modulation uses a cyclic prefix of sufficient length. Because of the cyclic prefix, each subcarrier "sees" a frequency-flat channel, with a gain depending on the subcarrier index; in spite of the frequency-

selective channel, the orthogonality of the subcarriers is maintained over the interval corresponding to the N samples processed by the RX. Because of the introduction of a cyclic prefix, OFDM experiences some loss of spectral efficiency and power efficiency, caused by the factor $\frac{N}{N+\nu}$; however, typically we have $N \gg \nu$, in which case the loss is very small. For single-carrier modulation, the SNR is the same for all data symbols, whereas for OFDM the SNR depends on the subcarrier index.

3.3 MMSE Estimator

In this section we give the reader a brief introduction to MMSE estimation. Let us consider an observation vector $\mathbf{y}$, of which the statistical properties depend on a parameter a; our aim is to produce an estimator $\hat{a} = \phi(\mathbf{y})$ of a, in such a way that the mean square error (MSE) between $\hat{a}$ and a is minimum. The MSE to be minimized is defined as

$$\mathbb{E}\left[|\hat{a} - a|^2\right] = \mathbb{E}\left[|\phi(\mathbf{y}) - a|^2\right], \tag{3.24}$$

where the expectation is with respect to the joint probability density $p(\mathbf{y}, a)$. It is shown in [15] that the MMSE estimator equals the a posteriori expectation of a, which is given by

$$\hat{a} \quad = \quad \mathbb{E}\left[a|\mathbf{y}\right], \tag{3.25}$$

where the expectation is with respect to the a posteriori density $p(a|\mathbf{y})$. The resulting minimum value of the MSE equals the average (over $\mathbf{y}$) of the a posteriori variance of a, which reduces to

$$\begin{aligned} \mathbb{E}\left[|\hat{a} - a|^2\right]_{\min} &= \mathbb{E}\left[\mathbb{E}\left[|a|^2|\mathbf{y}\right] - |\mathbb{E}\left[a|\mathbf{y}\right]|^2\right] \\ &= \mathbb{E}\left[|a|^2\right] - \mathbb{E}\left[|\mathbb{E}\left[a|\mathbf{y}\right]|^2\right]. \end{aligned} \tag{3.26}$$

In the following, we describe how we can use the MMSE estimator to predict the value of a random process on a future time instant, based on observations related to previous time instants. We introduce the following model for the observations

$$y_k = x_k + w_k, \quad k = 1, \ldots, N+1, \tag{3.27}$$

where $\mathbf{x} = [x_1, \ldots, x_{N+1}]^T$ and $\mathbf{w} = [w_1, \ldots, w_{N+1}]^T$ are jointly Gaussian with zero mean. Now, we want to predict the value $a = x_{N+1}$ based on only the previous observations y_1, y_2, $\ldots$, y_N. From our discussion above, we know that the minimum MSE is achieved by the a posteriori expectation of a, i.e., $\mathbb{E}\left[a|\mathbf{y}\right]$. Defining $\mathbf{y} = [y_1, \ldots, y_N]^T$ and introducing $\mathbf{R_{yy}} = \mathbb{E}\left[\mathbf{y}\mathbf{y}^H\right]$, $\mathbf{r}_{\mathbf{y}a} = \mathbb{E}\left[\mathbf{y}a^*\right]$

and $R_{aa} = \mathbb{E}\left[aa^*\right]$, we can write the a posteriori probability density $p(a|\mathbf{y})$ in the following form:

$$\begin{aligned} p(a|\mathbf{y}) &\propto p(\mathbf{y}, a) \\ &\propto \exp\left(-\begin{bmatrix} \mathbf{y}^H & a^H \end{bmatrix} \begin{bmatrix} \mathbf{R}_{\mathbf{yy}} & \mathbf{r}_{\mathbf{y}a} \\ \mathbf{r}_{\mathbf{y}a}^H & R_{aa} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{y} \\ a \end{bmatrix}\right), \end{aligned} \tag{3.28}$$

where $\propto$ denotes "equal within a factor not depending on a". Setting

$$\begin{bmatrix} \mathbf{R}_{\mathbf{yy}} & \mathbf{r}_{\mathbf{y}a} \\ \mathbf{r}_{\mathbf{y}a}^H & R_{aa} \end{bmatrix}^{-1} = \frac{1}{\sigma^2_{\mathrm{MSE}}} \begin{bmatrix} \mathbf{P} & -\mathbf{q} \\ -\mathbf{q}^H & 1 \end{bmatrix}, \tag{3.29}$$

we obtain from (3.28) that

$$p(a|\mathbf{y}) \propto \exp\left(-\frac{1}{\sigma^2_{\mathrm{MSE}}} \left(a - \mathbf{q}^H\mathbf{y}\right)^H \left(a - \mathbf{q}^H\mathbf{y}\right)\right). \tag{3.30}$$

It follows from (3.30) that for given $\mathbf{y}$, a is Gaussian with mean $\mathbf{q}^H\mathbf{y}$ and variance σ^2_{MSE}. Solving (3.29) for σ^2_{MSE} and $\mathbf{q}$, we find

$$\mathbf{q}^H = \mathbf{r}_{\mathbf{y}a}^H \mathbf{R}_{\mathbf{yy}}^{-1}, \tag{3.31}$$

$$\sigma^2_{\mathrm{MSE}} = R_{aa} - \mathbf{r}_{\mathbf{y}a}^H \mathbf{R}_{\mathbf{yy}}^{-1} \mathbf{r}_{\mathbf{y}a}. \tag{3.32}$$

Hence, the estimator and the corresponding minimum MSE are given by

$$\hat{a} = \mathbf{q}^H\mathbf{y} = \mathbf{r}_{\mathbf{y}a}^H \mathbf{R}_{\mathbf{yy}}^{-1}\mathbf{y}, \tag{3.33}$$

$$\mathbb{E}\left[|\hat{a} - a|^2\right]_{\min} = \sigma^2_{\mathrm{MSE}} = R_{aa} - \mathbf{r}_{\mathbf{y}a}^H \mathbf{R}_{\mathbf{yy}}^{-1} \mathbf{r}_{\mathbf{y}a}. \tag{3.34}$$

The quantities R_{aa}, $\mathbf{R}_{\mathbf{yy}}$ and $\mathbf{r}_{\mathbf{y}a}$ are easily obtained from the autocorrelation and cross-correlation matrices of $\mathbf{x}$ and $\mathbf{w}$. Note that the MMSE estimator (3.33) is a linear function of $\mathbf{y}$; this is because $p(\mathbf{y}, a)$ is Gaussian. Hence, restricting the MMSE estimator to be linear, i.e., $\hat{a}(\mathbf{y}) = \mathbf{q}^H\mathbf{y}$, and selecting $\mathbf{q}$ to minimize $\mathbb{E}\left[|\mathbf{q}^H\mathbf{y} - a|^2\right]$ would give rise to the same solution (3.33) and performance (3.34).

3.4 Performance Metric

This section presents several performance metrics that we will be using in the following chapters. We emphasize that the performance metrics described in 3.4.1 and 3.4.2 are information-theoretical metrics. They describe the best theoretical performance that the network can achieve. However, as these performances metrics are derived under assumptions which are not met in practice, it is safe to say that an actual network will always achieve a performance which is worse than these

theoretical limits. For example, these metrics are derived under the assumption that the transmitted symbols are distributed according to a continuous Gaussian probability density [16], which is substantially different from the discrete constellations that are commonly used in a practical network.

In section 3.4.3 we introduce a more practical metric, which will be the main focus of the third part of this dissertation.

3.4.1 The Capacity Metric

The *capacity* of a channel represents the highest information rate R, expressed in information bits per channel use (with a channel use denoting the transmission of a data symbol), at which reliable communications is possible [16]. Reliable communications means that there exists a sequence of codes (with growing block length) with corresponding rate R, for which the average error probability goes to zero as the block length of the code goes to infinity.

Let us consider an AWGN channel characterized by $z = \sqrt{E_\mathrm{s}}x + n$, where x is a complex-valued data symbol with $\mathbb{E}\left[|x|^2\right] = 1$, n is a zero-mean circularly symmetric Gaussian noise contribution with $\mathbb{E}\left[|n|^2\right] = \sigma^2$ and z is the input to the RX. The capacity of this AWGN is given by

$$C_\mathrm{AWGN} = \log_2\left(1 + \frac{E_\mathrm{s}}{\sigma^2}\right). \tag{3.35}$$

Introducing a fixed channel gain $h \in \mathbb{C}$, which is known by the RX, the corresponding capacity becomes

$$C_h = \log_2\left(1 + \frac{|h|^2 E_\mathrm{s}}{\sigma^2}\right). \tag{3.36}$$

Next, we assume that the channel gain h is time-varying. If each codeword is long enough to experience all possible channel gain values, the following capacity formula is applicable

$$C_\mathrm{erg} = \mathbb{E}_h\left[\log_2\left(1 + \frac{|h|^2 E_\mathrm{s}}{\sigma^2}\right)\right]. \tag{3.37}$$

The quantity C_erg is known as the *ergodic capacity*. However, in the case where the channel gain h is slowly varying and each codeword experiences only a single realization of the channel gain h, it will be more useful to describe the capacity of the channel as a random variable. This scenario is discussed in the next section.

Finally, we discuss the capacity of a multicarrier network when the channel gains are known by the RX. According to (3.23), a multicarrier system with a sufficiently long cyclic prefix can be modeled as N independent channels; the

capacity of a single channel can be found directly from equation (3.36), i.e. the capacity for the channel seen by the *k*th subcarrier is given by

$$\log_2\left(1+\frac{|H_k|^2E_{\rm s}}{\sigma_w^2}\right). \tag{3.38}$$

As the channels are independent, the capacity $C_{\rm OFDM}$ of a multicarrier network (in bits/OFDM symbol) is given by the sum of the capacities of all the subcarriers:

$$C_{\rm OFDM}=\sum_{k=1}^{N}\log_2\left(1+\frac{|H_k|^2E_{\rm s}}{\sigma_w^2}\right). \tag{3.39}$$

3.4.2 The Outage Capacity and Outage Probability

Assume that we want to transmit a message to the destination at an information rate R. In the case of slowly varying h, we define the *outage probability* as [14]

$$P_{\rm out}\triangleq \Pr\{C_h<R\}, \tag{3.40}$$

where C_h is defined in (3.36). The probability $P_{\rm out}$ is the probability that the capacity C_h is smaller than the fixed transmission rate R. Thus, we assume that the capacity C_h is a random variable, whose distribution depends on the statistics of the channel gain h. We note that the outage probability $P_{\rm out}$ is a more suitable performance metric than the capacity for applications with a fixed data rate, such as video transmission, as the outage probability expresses the reliability of the system, instead of the maximum achievable rate.

A related performance metric is the ϵ-*outage capacity* C_ϵ of the channel. This metric is defined as

$$\begin{aligned} C_\epsilon &= \max_R R \\ &\text{s.t. } P_{\rm out}(R)\le\epsilon, \end{aligned} \tag{3.41}$$

where $\epsilon>0$ and the dependency of $P_{\rm out}$ on the transmission rate R is explicitly shown. The outage capacity C_ϵ denotes the maximal rate for which the outage probability $P_{\rm out}(R)$ is smaller or equal to ϵ. As the outage probability $P_{\rm out}(R)$ is an increasing function of R, it follows that the outage probability $P_{\rm out}(R)$ will be equal to ϵ when $R=C_\epsilon$.

3.4.3 The Goodput Metric

Besides the information-theoretical performance metrics mentioned above, we also consider a more practical metric that is applicable to a packet-based transmission system. This metric is called the *goodput* (expressed in bit/s), and is defined in [17], for a given channel realization, as the following ratio

$$\text{GP}\triangleq\frac{N_{\rm d}}{T_{\rm packet}}, \tag{3.42}$$

where N_{d} denotes the expected number of correctly received information bits (associated with correctly decoded packet) per packet, and T_{packet} denotes the transmission time of a single packet. The goodput metric is a function of the transmit energy per subcarrier, the constellation size and the code rate. We note that the transmit energy and the constellation size can be different for every subcarrier. This metric will be discussed in more detail in section 4.3.

3.5 Resource Allocation

A resource allocation algorithm allows the TX to dynamically select its transmission parameters in such a manner that the performance of the network, which can be described by one of the metrics discussed in section 3.4, is optimized. The algorithm adapts the transmission parameters as a function of the channel information that is available at the TX. In this dissertation, most of the derived resource allocation algorithms are aware that the channel information at the TX can contain errors, which are caused by for example noise or feedback delays. This awareness allows the algorithms to make a better choice regarding the transmission parameters of the TX, which ultimately improves the network performance.

In the following, as an example, we derive a resource allocation algorithm for both a single-carrier and a multicarrier network, using the optimization tools from chapter 2. In both cases, the resource allocation algorithm optimizes the capacity of the link between the TX and the RX. We make the following assumptions:

- The TX has perfect information about the channel gain towards the RX.
- The energy that the TX may transmit is restricted by $E^{(\max)}$.
- The TX can adapt its transmission energy per data symbol.

3.5.1 Resource Allocation in a Single-Carrier Network

For a single-carrier network, the capacity of the link between the TX and the RX is given by (3.36). The resource allocation algorithm can be found by solving the following optimization problem

$$\begin{aligned} E_{\mathrm{s}}^{(\mathrm{opt})} = \ & \arg\max_{E} \ \log_2\left(1+\frac{|h|^2 E_{\mathrm{s}}}{\sigma^2}\right) \\ & \text{s.t. } E_{\mathrm{s}} \leq E^{(\max)}. \end{aligned} \tag{3.43}$$

As the objective function in (3.43) is concave and the constraint is linear, (3.43) represents a convex optimization problem. According to section 2.3, the Lagrangian of this optimization problem is given by

$$L(E_{\mathrm{s}},\mu) = -\log_2\left(1+\frac{|h|^2 E_{\mathrm{s}}}{\sigma^2}\right) + \mu(E_{\mathrm{s}} - E^{(\max)}), \tag{3.44}$$

where μ is the Lagrange multiplier corresponding to the constraint on the transmit energy E_s. According to the KKT condition (2.29), we have to set the derivative of the Lagrangian $L(E_\mathrm{s}^{(\mathrm{opt})}, \mu^{(\mathrm{opt})})$ with respect to E_s equal to 0, which means

$$\begin{aligned}
&\nabla_{E_\mathrm{s}} L(E_\mathrm{s}^{(\mathrm{opt})}, \mu^{(\mathrm{opt})}) \\
&= \frac{\partial}{\partial E_\mathrm{s}} \left(-\log_2 \left(1 + \frac{|h|^2 E_\mathrm{s}^{(\mathrm{opt})}}{\sigma^2} \right) + \mu^{(\mathrm{opt})} (E_\mathrm{s}^{(\mathrm{opt})} - E^{(\mathrm{max})}) \right) \\
&= -\frac{1}{\ln(2)} \frac{1}{1 + \frac{|h|^2 E_\mathrm{s}^{(\mathrm{opt})}}{\sigma^2}} \frac{|h|^2}{\sigma^2} + \mu^{(\mathrm{opt})} = 0.
\end{aligned} \tag{3.45}$$

By solving this equation, we obtain

$$E_\mathrm{s}^{(\mathrm{opt})} \quad = \quad \frac{1}{\ln(2)} \frac{1}{\mu^{(\mathrm{opt})}} - \frac{\sigma^2}{|h|^2}. \tag{3.46}$$

As the optimum solution also has to satisfy the following KKT condition (2.28)

$$\mu^{(\mathrm{opt})} (E_\mathrm{s}^{(\mathrm{opt})} - E^{(\mathrm{max})}) \quad = \quad 0, \tag{3.47}$$

we find the following value for $\mu^{(\mathrm{opt})}$

$$\mu^{(\mathrm{opt})} = \frac{1}{\ln(2)} \frac{1}{\frac{\sigma^2}{|h|^2} + E^{(\mathrm{max})}}. \tag{3.48}$$

By combining (3.46) and (3.48), we find the following optimum

$$E_\mathrm{s}^{(\mathrm{opt})} = E^{(\mathrm{max})}. \tag{3.49}$$

This result could also be found immediately by noting that the objective function in (3.43) is a strictly increasing function of E_s.

3.5.2 Resource Allocation in a Multicarrier Network

For a multicarrier network, the capacity of the link between the TX and the RX is given by (3.39). We get the following optimization problem

$$\begin{aligned}
\mathbf{E}^{(\mathrm{opt})} = \ & \arg \max_{\mathbf{E}} \sum_{k=1}^{N} \log_2 \left(1 + \frac{|H_k|^2 E_k}{\sigma_w^2} \right) \\
\text{s.t.} \ & \sum_{k=1}^{N} E_k \leq E^{(\mathrm{max})} \\
& E_k \geq 0, \quad k = 1, \ldots, N,
\end{aligned} \tag{3.50}$$

where $\mathbf{E} = [E_1, \ldots, E_N]^T$. We have added additional constraints to make sure that the transmit energy per symbol E_k, $k = 1, \ldots, N$, is greater than or equal to 0. As the optimization problem in (3.50) is convex, we can again find the optimum by setting $\nabla_{\mathbf{E}} L(\mathbf{E}^{(\text{opt})}, \mu_1^{(\text{opt})}, \ldots, \mu_{N+1}^{(\text{opt})})$ equal to zero, which leads to

$$\begin{aligned}
&\nabla_{\mathbf{E}} L(\mathbf{E}^{(\text{opt})}, \mu_1^{(\text{opt})}, \ldots, \mu_{N+1}^{(\text{opt})}) \\
&= \nabla_{\mathbf{E}} \left(-\sum_{k=1}^{N} \log_2 \left(1 + \frac{|H_k|^2 E_k^{(\text{opt})}}{\sigma_w^2} \right) \right. \\
&\left. - \sum_{k=1}^{N} \mu_k^{(\text{opt})} E_k^{(\text{opt})} + \mu_{N+1}^{(\text{opt})} (\sum_{k=1}^{N} E_k^{(\text{opt})} - E^{(\max)}) \right) \\
&= 0. \qquad (3.51)
\end{aligned}$$

By solving this equation, we obtain the following value for $E_k^{(\text{opt})}$, $k = 1, \ldots, N$,

$$E_k^{(\text{opt})} = \frac{1}{\ln(2)} \frac{1}{\mu_{N+1}^{(\text{opt})} - \mu_k^{(\text{opt})}} - \frac{\sigma_w^2}{|H_k|^2}. \qquad (3.52)$$

According to (2.28), the value of $\mu_k^{(\text{opt})} E_k^{(\text{opt})}$ is equal to 0 for every subcarrier k. This can either be achieved by setting $\mu_k^{(\text{opt})} = 0$ or $E_k^{(\text{opt})} = 0$. We will first consider the case where $E_k^{(\text{opt})} = 0$. In this case, from (3.52) we get that

$$\mu_k^{(\text{opt})} = \mu_{N+1}^{(\text{opt})} - \frac{1}{\ln(2)} \frac{|H_k|^2}{\sigma_w^2}. \qquad (3.53)$$

In the other case, where $\mu_k^{(\text{opt})} = 0$, the optimum value of $E_k^{(\text{opt})}$ is found as follows

$$E_k^{(\text{opt})} = \frac{1}{\ln(2)} \frac{1}{\mu_{N+1}^{(\text{opt})}} - \frac{\sigma_w^2}{|H_k|^2}. \qquad (3.54)$$

These two cases can be combined as follows

$$E_k^{(\text{opt})} = \left[\frac{1}{\ln(2)} \frac{1}{\mu_{N+1}^{(\text{opt})}} - \frac{\sigma_w^2}{|H_k|^2} \right]^+, \qquad (3.55)$$

where the operator $[.]^+$ represents $\max(., 0)$.

As the objective function is a strictly increasing function of E_k, $k = 1, \ldots, N$, the value of $\mu_{N+1}^{(\text{opt})}$ is found by solving the following equality

$$\sum_{k=1}^{N} \left[\frac{1}{\ln(2)} \frac{1}{\mu_{N+1}^{(\text{opt})}} - \frac{\sigma_w^2}{|H_k|^2} \right]^+ = E^{(\max)}. \qquad (3.56)$$

The formula shown in (3.55) is the well-known water-filling solution. We can think of $\frac{\sigma_w^2}{|H_k|^2}$ as the floor of a segment k and $\frac{1}{\ln(2)} \frac{1}{\mu_{N+1}^{(\text{opt})}}$ as the water level. The

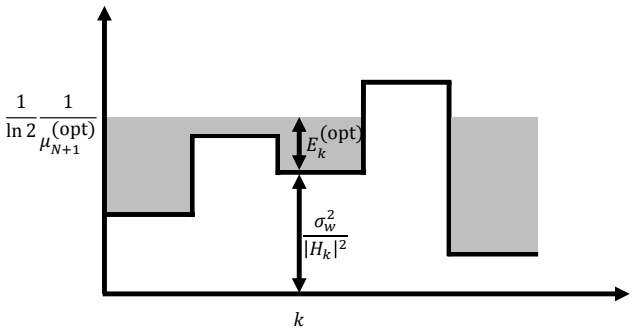


Figure 3.3: Illustration of the water-filling algorithm.

optimal value of $E_k^{(\mathrm{opt})}$ is then given by the height of the water in segment k. This is illustrated in Fig. 3.3.

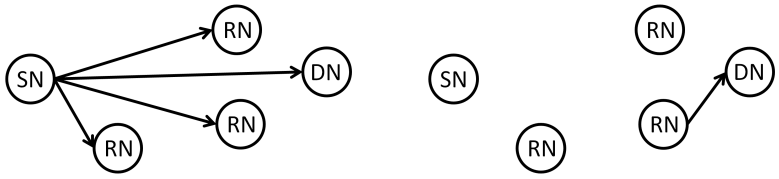
3.6 Cooperative Network

There is an ever growing demand for high data rates and reliable services over the air. Further, new applications for small wireless devices have to combine these high data rates with the requirement of a very low power consumption. Recently, the use of cooperative networks has been proposed as a cost-effective solution to meet these demands in next generations of wireless systems [18–21]. In a cooperative network, there will be one or more relay nodes (RNs) present which aid the TX in conveying its information to the RX. In the following, we call the original TX and RX the source node (SN) and the destination node (DN), respectively. In particular, the use of RNs has been shown to enhance the reliability of the network by improving the spatial diversity of the network. This spatial diversity can be successfully combined with other forms of diversity such as time and frequency diversity, to mitigate signal fading which arises from multipath propagation in a wireless environment [16].

The following are some of the more popular relaying protocols described in literature [22–25]:

- *Amplify-and-Forward*: The RNs amplify the message they receive from the SN, and forward the message to the DN.

- *Decode-and-Forward*: The RNs fully decode, re-encode and retransmit the received message to the DN.



(a) Transmission at the SN. **(b)** Transmission at a RN.

Figure 3.4: A cooperative network.

- *Compute-and-Forward*: The RNs decode linear functions of the messages transmitted by the SN, which the RNs then forward to the DN. Given enough equations, the DN can then recover the original message.

In the following, we will describe in more detail the different relaying protocols we have investigated in our work. Note, that we focus our attention on the case where the SN and the RNs transmit in non-overlapping time slots, in order to avoid mutual interference at the DN. This is illustrated in Fig. 3.4, where we show the transmission of a message by both the SN and by a RN.

Unless stated otherwise, we always assume that the channel gains between the nodes are independent and the number of RNs will be denoted by the variable M.

3.6.1 Direct Transmission Protocol

The most straightforward protocol, is the direct transmission (DT) protocol. In this protocol the RNs are not used for the transmission of the message. The SN simply transmits the message through the direct link between the SN and the DN. The channel between the SN and the DN is denoted by $h_{\mathrm{s,d}}$. We consider *flat Rayleigh fading*, so that the channel gain $h_{\mathrm{s,d}}$ is a zero-mean circularly Gaussian distributed random variable. This scheme is considered as an important alternative to cooperative networks. In contrast to the direct transmission, the transmission of a message x, with $\mathbb{E}\left[|x|^2\right] = 1$, in a cooperative network often requires multiple time slots, which might not always be beneficial for the performance of the network. The signal received at the DN is given by

$$z_{\mathrm{s,d}} = \sqrt{E_0} h_{\mathrm{s,d}} x + n_{\mathrm{s,d}}, \tag{3.57}$$

where E_0 denotes the transmission energy per symbol at the SN and $n_{\mathrm{s,d}} \sim N_{\mathrm{c}}(0, \sigma^2_{\mathrm{s,d}})$. The maximum capacity (bits/channel use) that is possible between the SN and DN for a given channel gain $h_{\mathrm{s,d}}$ is given by

$$C_{\mathrm{DT}} = \log_2\left(1 + E_0 \frac{|h_{\mathrm{s,d}}|^2}{\sigma^2_{\mathrm{s,d}}}\right), \tag{3.58}$$

assuming that the DN knows the channel gain $h_{\mathrm{s,d}}$.

3.6.2 Amplify-and-Forward Protocol

In the amplify-and-forward (AF) protocol, the RNs amplify their received message and forward it to the DN. If each RN participates in the transmission, the transmission of a message will require $M+1$ time slots.

In the first time slot, the SN sequentially broadcasts the components of a symbol vector $\mathbf{x} \in \mathbb{C}^{K\times 1}$, where $\mathbb{E}\left[\mathbf{x}\mathbf{x}^H\right] = \mathbf{I}_K$, to the DN and the RNs of the network. The received signals are given by

$$\mathbf{z}_{\mathrm{s,d}} = \sqrt{E_0}h_{\mathrm{s,d}}\mathbf{x} + \mathbf{n}_{\mathrm{s,d}}, \tag{3.59}$$

$$\mathbf{z}_{\mathrm{s,r}}(m) = \sqrt{E_0}h_{\mathrm{s,r}}(m)\mathbf{x} + \mathbf{n}_{\mathrm{s,r}}(m), \quad m = 1,\ldots,M, \tag{3.60}$$

where $\mathbf{n}_{\mathrm{s,d}} \sim N_{\mathrm{c}}(\mathbf{0}, \sigma^2_{\mathrm{s,d}}\mathbf{I}_K)$ and $\mathbf{n}_{\mathrm{s,r}}(m) \sim N_{\mathrm{c}}(\mathbf{0}, \sigma^2_{\mathrm{s,r}}(m)\mathbf{I}_K)$. The channel gains $h_{\mathrm{s,d}}$ and $h_{\mathrm{s,r}}(m)$ denote the channel gain between the SN and the DN and between the SN and the *m*th RN, respectively.

Then, from the second time slot till the $(M+1)$th time slot, each RN will individually amplify and forward its received message to the DN. Using E_m as the transmission energy per symbol, the *m*th RN amplifies its received signal and normalizes it by

$$\sqrt{E_0|h_{\mathrm{s,r}}(m)|^2 + \sigma^2_{\mathrm{s,r}}(m)}, \tag{3.61}$$

which represents the square-root of the average received energy at the corresponding RN. This means that the RN has to determine the average received energy $\mathbb{E}_{\mathbf{x},\mathbf{n}_{\mathrm{s,r}}(m)}\left[\frac{1}{K}\|\mathbf{z}_{\mathrm{s,r}}(m)\|^2\right]$. The message received by the DN from the *m*th RN during the $(m+1)$th time slot is

$$\mathbf{z}_{\mathrm{r,d}}(m) = \frac{\sqrt{E_m}}{\sqrt{E_0|h_{\mathrm{s,r}}(m)|^2 + \sigma^2_{\mathrm{s,r}}(m)}}h_{\mathrm{r,d}}(m)\mathbf{z}_{\mathrm{s,r}}(m) + \mathbf{n}_{\mathrm{r,d}}(m),\ m = 1,\ldots,M, \tag{3.62}$$

where $\mathbf{n}_{\mathrm{r,d}} \sim N_{\mathrm{c}}(\mathbf{0}, \sigma^2_{\mathrm{r,d}}\mathbf{I}_K)$, and $h_{\mathrm{r,d}}(m)$ denotes the channel gain between the *m*th RN and the DN.

Finally, assuming that the DN applies maximum ratio combining to the signals received in the different time slots, the instantaneous received SNR at the DN can be written as

$$\gamma_{\mathrm{AF}} = \sum_{m=0}^{M}\gamma_m, \tag{3.63}$$

where

$$\gamma_0 = \frac{E_0|h_{\mathrm{s,d}}|^2}{\sigma^2_{\mathrm{s,d}}}, \tag{3.64}$$

and

$$\begin{aligned}\gamma_m &= \frac{E_0 \frac{|h_{\mathrm{s,r}}(m)|^2}{\sigma^2_{\mathrm{s,r}}(m)} E_m \frac{|h_{\mathrm{r,d}}(m)|^2}{\sigma^2_{\mathrm{r,d}}(m)}}{E_0 \frac{|h_{\mathrm{s,r}}(m)|^2}{\sigma^2_{\mathrm{s,r}}(m)} + E_m \frac{|h_{\mathrm{r,d}}(m)|^2}{\sigma^2_{\mathrm{r,d}}(m)} + 1} \\ &= \frac{E_0 E_m |h_{\mathrm{s,r}}(m)|^2 |h_{\mathrm{r,d}}(m)|^2}{E_0 |h_{\mathrm{s,r}}(m)|^2 \sigma^2_{\mathrm{r,d}} + E_m |h_{\mathrm{r,d}}(m)|^2 \sigma^2_{\mathrm{s,r}} + \sigma^2_{\mathrm{s,r}}(m)\sigma^2_{\mathrm{r,d}}(m)}. \quad (3.65)\end{aligned}$$

The maximum rate (bits/channel use) that can be achieved by this protocol is given by

$$R_{\mathrm{AF}} = \frac{1}{M+1} \log_2 \left(1 + \gamma_{\mathrm{AF}}\right), \quad (3.66)$$

where the factor of $1/(M+1)$ comes from using $M+1$ time slots for transmitting the information. As the maximum possible rate that is achievable in a general cooperative network is not yet known [26], we do not use the word capacity to describe (3.66). However, (3.66) provides the maximum rate that is theoretically achievable by the considered protocol.

As the factor $1/(M+1)$ in (3.66) can lead to a poor performance, it is proposed in [27] that only the best RN, which is defined as the RN with the highest SNR γ_m, will assist the SN in the transmission. This selection amplify-and-forward scheme is shown to maintain full diversity order, and achieves the following rate

$$R_{\mathrm{S-AF}} = \frac{1}{2} \log_2 \left(1 + \gamma_{\mathrm{S-AF}}\right), \quad (3.67)$$

where

$$\gamma_{\mathrm{S-AF}} = \gamma_0 + \max_{m \in \{1,\ldots,M\}} \gamma_m. \quad (3.68)$$

3.6.3 Decode-and-Forward Protocol

In the decode-and-forward (DF) protocol, the RNs fully decode their received message and forward their estimation of $\mathbf{x}$ towards the DN. Similar to the selection amplify-and-forward scheme, we assume that only the RN that leads to the highest SNR at the DN is selected to participate in the transmission. Thus, the transmission of a message requires only two time slots.

In the first time slot, the SN broadcasts its message to the DN and RNs of the network. The received signals are again given by (3.59)-(3.60). Then, in the second time slot the selected RN decodes and forwards the received message to the DN. If we assume that the *m*th RN is selected, the signal received at the DN is given by

$$\mathbf{z}_{\mathrm{r,d}}(m) = \sqrt{E_m} h_{\mathrm{r,d}}(m)\mathbf{x} + \mathbf{n}_{\mathrm{r,d}}(m).$$

Finally, assuming that the DN applies maximum ratio combining to the signals received in the different time slots, the maximum rate (bits/channel use) that can

be achieved using the decode-and-forward protocol is [20]

$$R_{\mathrm{DF}} = \frac{1}{2} \max_{m\in\{1,\ldots,M\}} \left\{ \min \left\{ \log_2 \left(1 + \frac{E_0 |h_{\mathrm{s,r}}(m)|^2}{\sigma^2_{\mathrm{s,r}}(m)} \right), \right. \right. \\ \left. \left. \log_2 \left(1 + \frac{E_0 |h_{\mathrm{s,d}}|^2}{\sigma^2_{\mathrm{s,d}}} + \frac{E_m |h_{\mathrm{r,d}}(m)|^2}{\sigma^2_{\mathrm{r,d}}(m)} \right) \right\} \right\}. \quad (3.69)$$

3.7 Cognitive Radio

Because of the increasing demand for wireless services and systems, the electromagnetic radio spectrum has become a very scarce resource. Therefore, it is more important than ever to use it efficiently. Currently, national and international regulations determine which part of the spectrum is dedicated to specific licensed communication services: for instance, the frequency bands for GSM-900 are 890-915 MHz (mobile to base) and 935-960 MHz (base to mobile). Spectrum access in the same band by another communication system is strictly forbidden. However, measurements have shown that the current static spectrum allocation leads to a heavy underutilization of the available bandwidth.

Based on these observations, the idea of cognitive radio (CR) emerged. A CR-TX is a device which uses advanced radio and signal processing to allow unlicensed users to operate in an already crowded spectrum without degrading the performance of the existing licensed users. This idea not only requires additional research to make such a device feasible from a technical point of view, but it also requires novel spectrum allocation policies that allow additional TXs to operate in a licensed frequency band. In the context of cognitive radio, the unlicensed users are also called the secondary users (SUs), while the licensed users are referred to as the primary users (PUs).

Thus, CR has the potential to tackle the demanding problems of both scarcity and underutilization of the frequency spectrum [1,28]. However there are different ways in which the CR can solve this problem. In the following sections, we briefly discuss the three main paradigms in CR:

- The underlay paradigm
- The overlay paradigm
- The interweave paradigm

3.7.1 The Underlay Paradigm

In the underlay paradigm, the SU network may transmit concurrently with the PU network only if the interference caused to the PU network remains under a certain

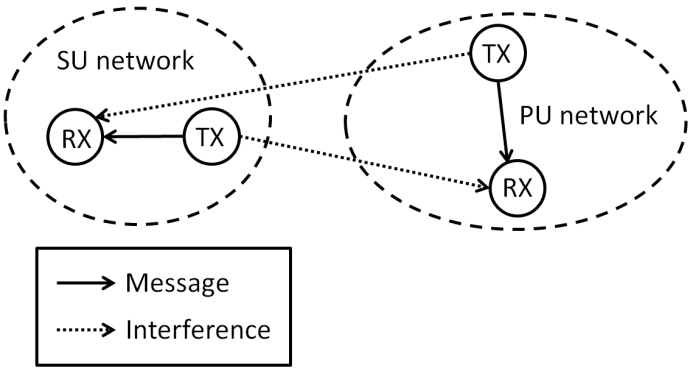


Figure 3.5: Example of an underlay network.

threshold. This threshold expresses a level of interference that is acceptable for the PU network. This means that as long as the interference remains under this threshold, the quality of service of the PU network will still be guaranteed [29–32]. We have illustrated the underlay paradigm in Fig. 3.5. The solid lines represent the intended signal paths, which connects the TX to a RX from the same network. The dashed lines however, represent the interference paths, connecting the TX from one network to a RX from another network, causing interference from the SU network to the PU network, and vice versa. In the following chapters, we will focus our attention to the interference caused at the PU-RX. So we will neglect the interference from the PU-TXs to the SU-RX. If necessary, this interference could be taken into account by increasing the variance of the noise at the SU-RX [33].

In order to keep the interference level below the set threshold, the SU network requires some information about its channel gains to the PU network, in order that the SU network knows how much interference it is generating at the PU-RXs. This information could be obtained from a *band manager* [31, 34], which monitors the spectrum and mediates between the PU and SU network, or, assuming time-division duplexing in the PU network and channel reciprocity, this channel information could be extracted by the SU-TX when the considered PU-RX has switched to a transmission mode.

There are several ways in which the SU network can control the interference caused at the PU-RXs. For example, the SU-TXs can dynamically adapt their transmit energy per symbol as a function of their channel gains towards the PU-RXs. In another example, the SU network can choose a very low transmission power, such that the interference level is almost certainly below threshold for every realistic value of the channel gain. However, the latter approach will only work for short range communications within the SU network, as this approach severely limits the allowable transmission power of the SU network.

The main advantage of the underlay paradigm, is that it requires none or little cooperation from the PU network. The PU-TXs can be completely unaware of the existence of the SU-TXs that simultaneously operate in the same frequency band. Hence, this method also does not require any synchronization between the PU and SU network. However, the disadvantage of this method is that the interference level constraints to be met at the PU-RX limit the achievable data rate performance of the SU network.

We will now discuss some ways to express these interference constraints. We make the distinction between the peak interference constraints and the average interference constraints. Denoting by g the channel gain between the SU-TX and the PU-RX, the peak interference constraint is expressed as

$$E|g|^2 \leq \Gamma, \tag{3.70}$$

where E denotes the transmit energy per symbol and Γ denotes the interference threshold; E is in general a function of all channel gains in the CR network. This approach will be investigated in chapters 5, 8 and 10.

However, another possibility is to control the average interference level at the PU-RXs over all the fading states. This approach results in the following average interference constraint

$$\mathbb{E}\left[E|g|^2\right] \leq \bar{\Gamma}, \tag{3.71}$$

where $\bar{\Gamma}$ denotes the average interference threshold, and the average is over g and all other channel gains E depends on. The average interference constraints will be used in chapters 6 and 7.

Further, notice that in order to satisfy the peak interference constraint (3.70), the SU-TX has to know the instantaneous channel gain towards the PU-RX. However, due to feedback delay and estimation errors it can happen that the channel state information (CSI) at the SU-TX is imperfect. In this case, the SU-TX can no longer guarantee that the peak interference constraint at the PU-RXs is satisfied. Therefore, alternative formulations of the interference constraints are needed that can be satisfied by the SU-TX.

- A first possibility is to satisfy the interference constraints only on average, conditioned on the available imperfect CSI, denoted $\mathbf{ICSI}_{\mathrm{PU}}$. In this case the peak interference constraint is replaced by

$$E\,\mathbb{E}\left[|g|^2\,|\mathbf{ICSI}_{\mathrm{PU}}\right] \leq \Gamma. \tag{3.72}$$

  This approach will be considered in chapters 8 and 10.

- A second possibility is to define the interference constraint by means of uncertainty sets [35, 36]. By defining the following uncertainty set $\mathcal{S}$

$$\mathcal{S} = \left\{g : g = \mathbb{E}\left[g\,|\mathbf{ICSI}_{\mathrm{PU}}\right] + \alpha\epsilon, \|\epsilon\| \leq 1\right\}, \tag{3.73}$$

we can formulate the interference constraint as follows

$$E|g|^2 \leq \Gamma, \quad \forall g \in \mathcal{S} \tag{3.74}$$

where the complex scalar α defines the size of the uncertainty interval, which directly influences the minimum probability that the interference is below the interference threshold Γ. This type of constraint will be investigated in chapter 5.

- A third possibility, used in [37–39], is to neglect the statistical variation of g for given $\mathbf{ICSI}_{\mathrm{PU}}$, which corresponds to simply replacing in (3.70) g by its conditional expectation $\mathbb{E}\left[g\,|\mathbf{ICSI}_{\mathrm{PU}}\right]$. The following interference constraint results:

$$E\left|\mathbb{E}\left[g\,|\mathbf{ICSI}_{\mathrm{PU}}\right]\right|^2 \leq \Gamma. \tag{3.75}$$

We note that the above interference constraints can be linked to the concept of interference probability as defined in [37]. The interference probability (IP) for the PU-RX reads as

$$\mathrm{IP} = \Pr\left(E|g|^2 > \Gamma\right). \tag{3.76}$$

In the case that the dynamically allocated energy E leads to a too large IP, one can substitute Γ in the corresponding interference constraint by $\kappa\Gamma$. The scaling factor κ is chosen such that IP reaches an acceptable value, after finding a new dynamic allocation of the E which satisfies the new constraint.

3.7.2 The Overlay Paradigm

In the overlay paradigm, the SU network actively cooperates with the PU network, which requires the SU network to know the messages and the codebook of the PU network. An example of a cognitive overlay network is shown in Fig. 3.6. Here, a PU-TX sends a message which is received both by the PU-RX and a SU-TX. In the case where the message cannot be decoded by the PU-RX, the SU-TX can act as a RN for the PU network by decoding and forwarding the received message to the PU-RX. This way, the SU network uses a part of its resources to improve the performance of the PU network. However, as the SU network also transmits its own messages, it thereby causes interference at the PU-RX which in turn degrades the performance of the PU network. The overlay paradigm now dictates that the SU network has to guarantee that the resulting performance of the PU network is at least as good as when the SU network would not be present.

3.7.3 The Interweave Paradigm

The interweave paradigm differs from the previous two paradigms, in that the SU network is not allowed to transmit concurrently with the PU network. This mode

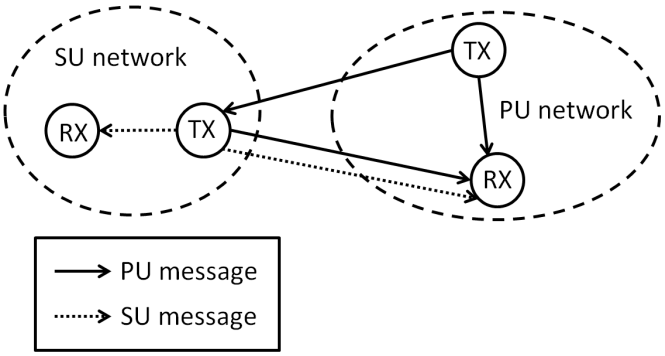


Figure 3.6: Example of an overlay network.

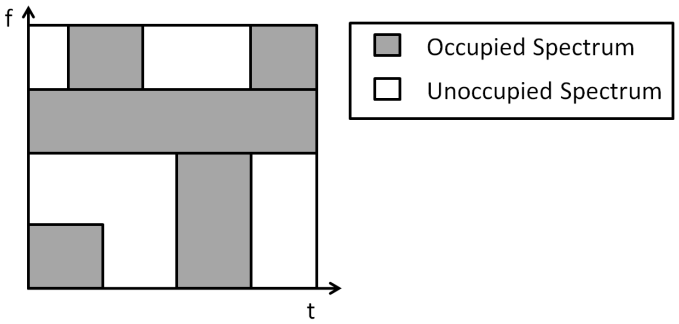


Figure 3.7: Example of an interweave network.

of operation is also known as opportunistic communication. The idea here is that the SU network periodically senses the spectrum, and tries to locate the parts of the spectrum which are currently unoccupied. These unoccupied parts in time and frequency are schematically shown in Fig. 3.7. These parts are often called *spectrum holes*. Once a spectrum hole is detected, the SU network can start its transmission in this unoccupied frequency band. However, as soon as the PU network starts using this frequency band, the SU network has to immediately cease its transmissions in that band.

3.8 Chapter Summary

This chapter has presented the reader with the system model that forms the basis for the next chapters. We introduced the wireless channel models, and discussed the main building blocks that are relevant for the description of the communication

network. The remaining chapters will use both the flat fading and the frequency-selective fading channel models.

A short introduction about the MMSE estimator was presented. This section provides the reader with the necessary background for the channel prediction which we often assume to be performed at the TX.

Further, this chapter gave an overview of the performance metrics that we use in our dissertation to describe the performance of the network and showed how the TX can optimize these metrics by means of a dynamic adaptation of its transmission parameters. We also introduced the two cooperation protocols that will be considered in the subsequent chapters, i.e., amplify-and-forward and decode-and-forward.

Finally, we presented the three main paradigms of cognitive radio: the underlay, overlay and interweave paradigm. The underlay and overlay paradigm allow the SU network to transmit concurrently with the PU network, while the interweave paradigm avoids concurrent transmissions.

The three paradigms require that different amounts of information about the PU network are available at the SU network. For the underlay paradigm, it is necessary to know how much interference the SU network generates at the PU-RXs. The overlay paradigm is clearly the one where the most information is needed. As the SU network cooperates with the PU network, it needs to know the codebook of the PU network and in most cases also the PU messages. For the interweave paradigm, the SU network has to know or sense whether the PU network is currently transmitting in a certain frequency band.

In this dissertation, we will consider the underlay paradigm. The main advantage of this paradigm is that the PU network can be completely oblivious to the presence of the SU network. In contrast to the overlay paradigm, no synchronization or cooperation between the PU and SU networks is required. Further, the underlay paradigm allows the SU network to transmit without interruptions, which contrasts the interweave paradigm.

4

Packet-Based Transmission Systems and the Goodput Metric

In the literature, resource allocation in cognitive underlay networks with imperfect channel state information (CSI) has received much attention. However, there are many papers [37,40–44] that only focus on rather theoretical performance metrics, such as the channel capacity or the SNR at the SU-RX, which have only limited relevance to practical systems using specific codes and constellations. As it was shown in [45] and [46] that resource allocation algorithms can achieve a better performance, when they are aware of the CSI imperfections at the TX, it would be beneficial if these algorithms were also available for a practical coded multicarrier system. These algorithms would extend the results in [45,46], which were obtained for an uncoded Orthogonal Frequency-Division Multiplexing (OFDM) system; in such a system, the probability of a bit error on a subcarrier only depends on the SNR of the considered subcarrier, which considerably simplifies the resource allocation problem. However, the results for uncoded OFDM are of limited use in a practical scenario where channel coding across subcarriers is used.

The third part of this thesis will consider the problem of implementing these resource allocation algorithms in a practical coded multicarrier transmission system, with imperfect CSI at the SU-TX, due to estimation errors or feedback delays. Because the bits are coded and the channel is frequency-selective, the performance of the network depends upon a complicated function of the SNRs on all the sub-

carriers which are used for the transmission. The actual performance metric we consider is the goodput (GP), which will be introduced in more detail. This metric gives us the advantage of allowing the optimization of realistic modulation and coding formats.

This chapter provides the necessary concepts for chapter 10, where we will derive several resource allocation algorithms which optimize the code rate together with uniform or non-uniform bit allocation and uniform or non-uniform energy allocation. The key idea behind the proposed method relies on optimizing the long-term average GP of the SU link, averaged over the realizations of both the actual channel and the available CSI at the SU-TX, under the constraints on the total transmitted energy and on the level of interference affecting the PU-RXs.

In section 4.1, we will introduce the packet-based transmission system. Further, as we make the assumption that these systems use bit-interleaved coded modulation, we explain this coding technique in section 4.2. The GP metric that will be used to describe the performance of these systems will be introduced in section 4.3. Section 4.4 describes a technique called SNR mapping, which makes the analytical description of the performance of these coded multicarrier systems much easier. Finally, the conclusions are drawn in section 4.5.

4.1 A Packet-Based Transmission System

In practice, a transmission system will often send its information bits in small groups or packets. These packets are frequently assumed to contain a fixed number of bits, which simplifies the implementation of the transmission system.

A transmission packet, as illustrated in Fig. 4.1, is divided into N_{p} information bits of payload and N_{CRC} redundant bits for the cyclic redundancy check (CRC), and has a total length of $N_{\mathrm{tot}} = N_{\mathrm{p}} + N_{\mathrm{CRC}}$ bits. The CRC allows the RX to check whether or not the received packet was correctly received. We briefly explain a CRC in appendix 4.A. In the following we will always use the following CRC polynomial [47]

$$g(x) = x^{32}+x^{26}+x^{23}+x^{22}+x^{16}+x^{12}+x^{11}+x^{10}+x^{8}+x^{7}+x^{5}+x^{4}+x^{2}+x^{1}+1. \tag{4.1}$$

To protect the transmission of this packet we will use bit-interleaved coded modulation, which we explain in detail in the next section.

4.2 Bit-Interleaved Coded Modulation

We consider the transmission of data symbols over a frequency-selective channel, using OFDM with N subcarriers (see section 3.2.3.2). We assume that the symbol x_k on the kth subcarrier, which consists of $m_k \in \mathcal{D}_m$ coded bits, belongs to a

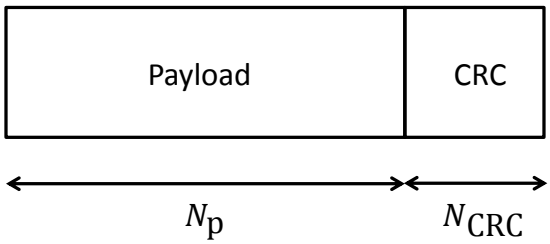


Figure 4.1: A transmission packet.

normalized square 2^{m_k}-QAM constellation; hence, the OFDM symbol represents $\sum_{k\in\mathcal{N}} m_k$ bits, where $\mathcal{N} \triangleq \{1,\ldots,N\}$. The kth component of the properly scaled fast Fourier transform (FFT) output at the RX can be represented as $y_k = x_k + w_k$, with $w_k \sim N_{\mathrm{c}}(0, \frac{1}{\gamma_k})$, where $\gamma_k = \frac{E_k|H_k|^2}{\sigma_w^2}$ denotes the SNR on the kth subcarrier; E_k and H_k are the transmit energy per symbol and the channel gain, both related to the kth subcarrier, $\sigma_w^2 = \frac{N+\nu}{N}\sigma^2$ where σ^2 is the noise spectral density at the RX input and ν is the length of the cyclic prefix (in sampling intervals of duration T).

The transmission of the packets is protected by channel coding. Here we consider bit-interleaved coded modulation (BICM) which uses a convolutional code. BICM has been proposed for the first time in [48], as a more flexible alternative to trellis-coded modulation, allowing the code and the constellation mapping to be selected independently.

The application of BICM in the considered setting involves encoding the payload and CRC bits from the packets by means of a rate r convolutional code. The resulting m_{tot} coded bits are interleaved. The bits at the output of the interleaver are Gray-mapped to N_{s} data symbols, which are transmitted over a frequency-selective fading channel as a frame of $N_{\mathrm{OFDM}} = \lceil N_{\mathrm{s}}/N \rceil$ consecutive OFDM symbols.

Based on its observation, the RX tries to recover the coded bits generated by the convolutional encoder. Using the one-to-one correspondence between the input and output bits from the convolutional encoder, the N_{tot} bits from each packet are derived from the m_{tot} recovered coded bits. We consider the decoding proposed by [48], which yields a RX consisting of the following building blocks: (i) a soft demapper computes the log-likelihood ratio (LLR) for each coded bit from the OFDM symbol; (ii) a de-interleaver applies to the m_{tot} LLRs a permutation which is the inverse of the permutation from the interleaver at the TX; and (iii) a Viterbi decoder operates on the de-interleaved LLRs to recover the coded bits.

The resulting system block diagram is shown in Fig. 4.2, where "channel" represents the concatenation of the OFDM modulator, the frequency-selective channel

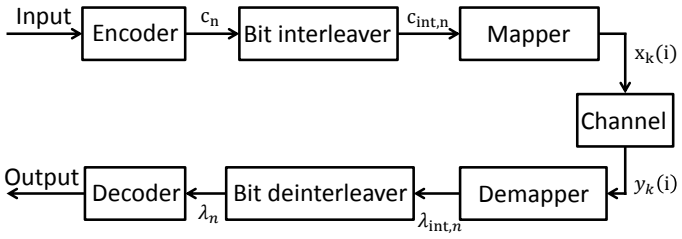


Figure 4.2: Building blocks of a BICM system.

and the OFDM demodulator. These blocks and their corresponding variables are discussed in more detail below.

4.2.1 Convolutional Encoder

In order to protect the information bits against noise and channel fading, we assume that the TX employs a convolutional encoder with code rate $r \in \mathcal{D}_r$, where $\mathcal{D}_r$ denotes the set of possible code rates. An example of a rate-$1/2$ convolutional encoder, with input sequence $\mathbf{b}_1$ and output sequences $\mathbf{c}_1$ and $\mathbf{c}_2$, is shown in Fig. 4.3. In this example the encoder consists of 3 delay cells (each denoted by D); the content of these delay cells represents the encoder state. The constraint length of the code is defined as 1 plus the number of delay cells (i.e., constraint length 4 for the considered example). Starting from the all-zero state, the coded bit sequences $\mathbf{c}_1$ and $\mathbf{c}_2$ are obtained as the modulo-2 convolution between $\mathbf{b}_1$ and the impulse responses $\mathbf{g}_1 = (0, 1, 0, 1)$ and $\mathbf{g}_2 = (1, 1, 1, 0)$, respectively. The impulse responses have a length equal to the constraint length; the 1s and 0s in $\mathbf{g}_1$ and $\mathbf{g}_2$ indicate the presence and absence, respectively, of a connection from the encoder input and the delay cell outputs to the corresponding encoder output. For example, the 4-bit input sequence $\mathbf{b}_1 = (1, 0, 1, 1)$ leads to the 7-bit output sequences $\mathbf{c}_1 = (0, 1, 0, 0, 1, 1, 1)$ and $\mathbf{c}_2 = (1, 1, 0, 0, 0, 1, 0)$. Each of these output sequences consists of a number of bits equal to the number of input bits (4), followed by a number of tail bits equal to the number of delay cells (3); the tail bits result from bringing the encoder state to zero by appending three additional 0s to the input sequence. Assuming statistically independent input bits, it can be verified that coded bits, at positions differing by at least the constraint length, are statistically independent, because they depend on disjoint subsets of information bits. After the encoding, the two coded sequences $\mathbf{c}_1$ and $\mathbf{c}_2$ are multiplexed into the single coded sequence $\mathbf{c} = (0, 1, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 1, 0)$.

When the tail bits are not transmitted, the number of coded bits equals twice the number of input bits, resulting in a code rate of $1/2$. However, including

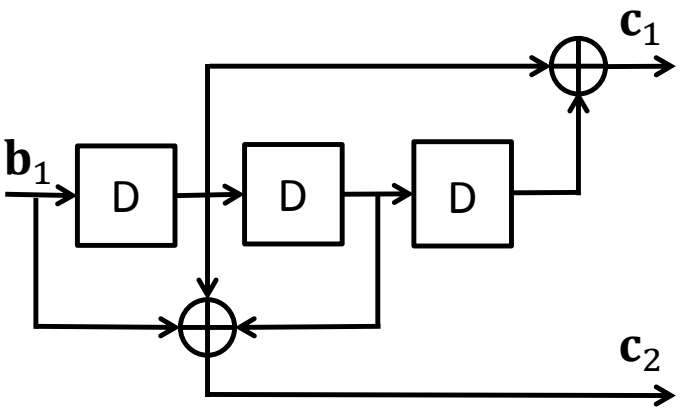


Figure 4.3: A convolutional encoder.

the tail bits in the transmission proves helpful to the decoding at the RX, which then can exploit the fact that the end state of the encoder is the all-zero state; the transmission of the tail bits gives rise to some reduction of the code rate, which becomes negligibly small when the number of input bits is much larger than the constraint length of the code.

In the following, we will define a convolutional code by giving the octal representation of the impulse responses. For example, the encoder in Fig. 4.3 can be defined by the pair $(5, 16)$.

We note that when a higher code rate is required, one option is to choose a different encoder. However, in some cases it can be advantageous to derive a higher-rate code from a given *mother code* by means of a method called puncturing. This method allows the communication system to use the same encoder (and decoder) for different code rates. As an additional benefit, the decoder has a much lower complexity than it would have when designed for a high code rate.

Considering a rate-$1/2$ convolutional code as *mother code*, we can obtain a code of rate $2/3$ by using the following puncturing matrix

$$\mathbf{P} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \tag{4.2}$$

with the first and second row referring to the encoder outputs $\mathbf{c}_1$ and $\mathbf{c}_2$, respectively. The encoder outputs $\mathbf{c}_1$ and $\mathbf{c}_2$ are divided into blocks having a length equal to the number of columns in $\mathbf{P}$. The 1s and 0s in a row of $\mathbf{P}$ indicate which bits from the blocks of the corresponding encoder output must be kept and discarded, respectively. According to the puncturing matrix $\mathbf{P}$ from (4.2), the encoder out-

Code Rate	Puncture Matrix	Hamming Distance
$1/2$	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$	10
$2/3$	$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$	6
$3/4$	$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$	5
$5/6$	$\begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix}$	4

Table 4.1: The optimal puncture matrices and Hamming distances for the code $(171, 133)$ [49].

put sequences are divided into 2-bit blocks; the blocks originating from $\mathbf{c}_1$ remain unaltered, whereas in the blocks originating from $\mathbf{c}_2$ the second bit is removed. Equivalently, from the bit sequence $\mathbf{c}$ obtained by multiplexing $\mathbf{c}_1$ and $\mathbf{c}_2$, every fourth bit is discarded. Applying this puncturing to the previous example results in the coded sequence $\mathbf{c}' = (0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 0)$.

In table 4.1 we show the puncturing matrices which lead to the lowest bit error rate performance under soft Viterbi decoding [49] when puncturing the rate-$1/2$ mother code $(171, 133)$. It can be shown that the mother code has the largest possible minimum Hamming distance for a constraint length of 7.

4.2.2 Bit Interleaver and Mapper

The coded bits resulting from the encoder are fed to a bit interleaver, which applies a random permutation to the m_{tot} coded bits: denoting by c_n and $c_{\text{int},n}$ the *n*th bit at the input and the output, respectively, of the interleaver, these bits are related by $c_n = c_{\text{int},\pi(n)}$, where $(\pi(1), ..., \pi(m_{\text{tot}}))$ represents a permutation of $(1, ..., m_{\text{tot}})$.

We then divide the interleaved bits $\mathbf{c}_{\text{int}}$ into blocks of $\sum_{k\in\mathcal{N}} m_k$ bits. The *i*th ($i = 1, \ldots N_{\text{OFDM}}$) block is then Gray-mapped[1] to a data symbol vector $\mathbf{x}(i) = [x_1(i), \ldots, x_N(i)]^T$ to be transmitted on the subcarriers of the OFDM signal: the first m_1 bits of a block are mapped to the symbol $x_1(i)$, the next m_2 bits are mapped to symbol $x_2(i)$, and so on, until the last m_N bits of a block are mapped to symbol $x_N(i)$.

The task of the interleaver is to permute the coded bits $\mathbf{c}$ in such a way that the bits from $\mathbf{c}_{\text{int}}$ that are mapped to the same symbol originate from positions in the encoder output stream that are far apart, so that these coded bits can be considered statistically independent.

[1] We assume *Gray mapping* as this leads to the best performance in BICM [50].

4.2.3 Soft Demapper and De-Interleaver

The RX first performs soft demapping by computing the LLR for each bit from $\mathbf{c}_{\text{int}}$. Let us denote by $c_{j,k}(i)$ the *j*th bit from the symbol $x_k(i)$; hence, $c_{j,k}(i)$ is the bit from $\mathbf{c}_{\text{int}}$ at position $j+(i-1)\sum_{l=1}^{N} m_l$ (when $k=1$) or $j+\sum_{l=1}^{k-1} m_l + (i-1)\sum_{l=1}^{N} m_l$ (when $k>1$). The LLR associated with $c_{j,k}(i)$ is defined as

$$\text{LLR}_{i,j,k}(y_k(i)) = \ln\left(\frac{p(y_k(i)|c_{j,k}(i)=1,H_k)}{p(y_k(i)|c_{j,k}(i)=0,H_k)}\right), \tag{4.3}$$

where $y_k(i)$ corresponds to the observation of $x_k(i)$ and the channel gain H_k is known by the RX. When $\text{LLR}_{i,j,k}(y_k(i))$ is positive (negative), $c_{j,k}(i)$ is likely to equal 1 (to equal 0). The larger the magnitude $|\text{LLR}_{i,j,k}(y_k(i))|$, the more reliable the information provided by the sign of $\text{LLR}_{i,j,k}(y_k(i))$.

The likelihood $p(y_k(i)|c_{j,k}(i)=b,H_k)$ with $b\in\{0,1\}$ is obtained as

$$p(y_k(i)|c_{j,k}(i)=b,H_k) = \sum_{\alpha\in\chi_k} p(y_k(i)|x_k(i)=\alpha,H_k)\Pr(x_k(i)=\alpha|c_{j,k}(i)=b), \tag{4.4}$$

where χ_k is the set of constellation points related to the constellation of the symbol $x_k(i)$. Assuming that the bits $(c_{1,k}(i),...,c_{m_k,k}(i))$ are statistically independent because of the interleaving, we obtain $\Pr(x_k(i)=\alpha|c_{j,k}(i)=b) = 2^{-(m_k-1)}$ for $\alpha\in\chi_{k,b}^{j}$ and $\Pr(x_k(i)=\alpha|c_{j,k}(i)=b)=0$ for $\alpha\notin\chi_{k,b}^{j}$, where $\chi_{k,b}^{j}$ is the subset of constellation points from χ_k for which $c_{j,k}(i)=b$. This yields

$$\begin{aligned}\text{LLR}_{i,j,k}(y_k(i)) &= \ln\left(\frac{\sum_{\alpha\in\chi_{k,1}^{j}} p(y_k(i)|x_k(i)=\alpha,H_k)}{\sum_{\alpha\in\chi_{k,0}^{j}} p(y_k(i)|x_k(i)=\alpha,H_k)}\right)\\ &= \ln\left(\frac{\sum_{\alpha\in\chi_{k,1}^{j}} \exp\left(-\gamma_k|y_k(i)-\alpha|^2\right)}{\sum_{\alpha\in\chi_{k,0}^{j}} \exp\left(-\gamma_k|y_k(i)-\alpha|^2\right)}\right).\end{aligned} \tag{4.5}$$

The computation of the LLRs can be simplified by keeping in both summations from (4.5) only the largest term, i.e.,

$$\text{LLR}_{i,j,k}(y_k(i)) \approx \gamma_k\left(\min_{\alpha\in\chi_{k,0}^{j}} |y_k(i)-\alpha|^2 - \min_{\alpha\in\chi_{k,1}^{j}} |y_k(i)-\alpha|^2\right). \tag{4.6}$$

This approximation avoids the use of logarithmic and exponential functions. This approximation has minimal impact on the receiver's performance when Gray mapping is being used [51].

The LLRs of the interleaved coded bits $\mathbf{c}_{\text{int}} = (c_{\text{int},1},...,c_{\text{int},m_{\text{tot}}})$ are denoted $(\lambda_{\text{int},1},...,\lambda_{\text{int},m_{\text{tot}}})$. These LLRs are de-interleaved to obtain the LLRs $(\lambda_1,...,\lambda_{m_{\text{tot}}})$ of the encoder output bits $\mathbf{c}=(c_1,...,c_{m_{\text{tot}}})$, i.e., $\lambda_n = \lambda_{\text{int},\pi(n)}$.

4.2.4 Viterbi Decoder

Ideally, we want the RX to detect as many correct packets as possible. When the information bits are equally likely, this can be achieved by maximizing the following maximum likelihood (ML) sequence detection criterion [52]

$$\arg\max_{\mathbf{c}} p(\mathbf{y}(1),\ldots,\mathbf{y}(N_{\text{OFDM}})|\mathbf{c},H_1,\ldots,H_N), \tag{4.7}$$

where $\mathbf{y}(i)=(y_1(i),...,y_N(i))$ denotes the observations related to the *i*th OFDM symbol. However, the typical approach for a BICM system (4.7) is to approximate this criterion as follows [50, 51]

$$\begin{aligned}
&\arg\max_{\mathbf{c}} p(\mathbf{y}(1),\ldots,\mathbf{y}(N_{\text{OFDM}})|\mathbf{c},H_1,\ldots,H_N)\\
&=\arg\max_{\mathbf{c}} \sum_{i=1}^{N_{\text{OFDM}}}\sum_{k=1}^{N} \ln\left(p(y_k(i)|c_{1,k}(i),\ldots,c_{m_k,k}(i),H_k)\right)\\
&\approx\arg\max_{\mathbf{c}} \sum_{i=1}^{N_{\text{OFDM}}}\sum_{k=1}^{N}\sum_{j=1}^{m_k} \ln\left(p(y_k(i)|c_{j,k}(i),H_k)\right).
\end{aligned} \tag{4.8}$$

This step is suboptimal in most cases as an optimal ML decoder requires joint demapping and decoding. However, the advantage of this approach is that it leads to a simple binary decoder, which can be completely unaware of the channel and the type of modulation that was used.

The objective function proposed in (4.8) can be rewritten as a function of the LLRs introduced in (4.5). As the solution of the maximization does not change when we add a constant term or scaling factor to the objective function, we can

write the following

$$
\begin{aligned}
& \arg\max_{\mathbf{c}} \sum_{i=1}^{N_{\text{OFDM}}} \sum_{k=1}^{N} \sum_{j=1}^{m_k} \ln\left(p(y_k(i)|c_{j,k}(i), H_k)\right) \\
= \; & \arg\max_{\mathbf{c}} \sum_{i=1}^{N_{\text{OFDM}}} \sum_{k=1}^{N} \sum_{j=1}^{m_k} 2\ln\left(p(y_k(i)|c_{j,k}(i), H_k)\right) \\
= \; & \arg\max_{\mathbf{c}} \sum_{i=1}^{N_{\text{OFDM}}} \sum_{k=1}^{N} \sum_{j=1}^{m_k} \big(2\ln\left(p(y_k(i)|c_{j,k}(i), H_k)\right) \\
& - \ln\left(p(y_k(i)|c_{j,k}(i)=1, H_k)\right) - \ln\left(p(y_k(i)|c_{j,k}(i)=0, H_k)\right)\big) \\
= \; & \arg\max_{\mathbf{c}} \sum_{i=1}^{N_{\text{OFDM}}} \sum_{k=1}^{N} \sum_{j=1}^{m_k} \left(\ln\left(\frac{p(y_k(i)|c_{j,k}(i), H_k)}{p(y_k(i)|c_{j,k}(i)=1, H_k)}\right)\right. \\
& \left. + \ln\left(\frac{p(y_k(i)|c_{j,k}(i), H_k)}{p(y_k(i)|c_{j,k}(i)=0, H_k)}\right)\right) \\
= \; & \arg\max_{\mathbf{c}} \sum_{i=1}^{N_{\text{OFDM}}} \sum_{k=1}^{N} \sum_{j=1}^{m_k} (2c_{j,k}(i)-1)\text{LLR}_{i,j,k}(y_k(i)) \\
= \; & \arg\max_{\mathbf{c}} \sum_{n=1}^{m_{\text{tot}}} (2c_n - 1)\lambda_n. \qquad (4.9)
\end{aligned}
$$

This means that the decoder has to determine the sequence $\mathbf{c}$ which maximizes the objective function $F(\mathbf{c})$, given by

$$F(\mathbf{c}) = \sum_{n=1}^{m_{\text{tot}}} (2c_n - 1)\lambda_n, \qquad (4.10)$$

over all legitimate encoder output sequences. As the considered code is convolutional, $F(\mathbf{c})$ can be efficiently maximized by means of the Viterbi algorithm [53]. In the case of a punctured convolutional code, the Viterbi algorithm is applied to the mother code, with the LLRs corresponding to the punctured bits simply set to zero.

4.3 Goodput Metric

In this section we will define a performance metric that is capable of describing the performance of the packet-based transmission system introduced in the previous sections. In section 4.1, we defined N_{tot} as the total number of bits in a packet, of which there are N_{p} information bits. The packet is protected against errors by BICM as described in section 4.2.

The bit allocations for every subcarrier are collected in a bit allocation vector $\mathbf{m} \triangleq [m_1, \ldots, m_N]^T$. The selected bit allocation vector $\mathbf{m}$ and code rate r are called a transmission mode (TM) $\phi \triangleq \{\mathbf{m}, r\} \in \mathcal{D}_m^N \times \mathcal{D}_r$. Finally, let us also define the vector $\mathbf{SNR} \triangleq [\gamma_1, \ldots, \gamma_N]$, where γ_k denotes the SNR of the *k*th subcarrier at the FFT output at the RX.

We already briefly introduced the GP in section 3.4.3. As this metric expresses the ratio of the expected number of correctly received information bits (associated with correctly decoded packets) to the actual transmission time, it is clear that this metric has a very clear practical interpretation.

Let us first consider the case where the channel is static, so that $\mathbf{SNR}$ is constant over time. The GP (in bit/s) corresponding to a fixed TM $\phi = \{\mathbf{m}, r\}$ and vector $\mathbf{SNR}$ can be written as

$$\mathrm{GP}(\phi, \mathbf{SNR}) \triangleq \frac{N_\mathrm{p}\left(1 - \mathrm{PER}(\phi, \mathbf{SNR})\right)}{T_\mathrm{packet}}, \tag{4.11}$$

where T_packet denotes the transmission time of a single packet and $\mathrm{PER}(\phi, \mathbf{SNR})$ is the packet error rate (PER) corresponding to the selected $(\phi, \mathbf{SNR})$. Taking into account that the number of coded bits per OFDM symbol equals $\sum_{k\in\mathcal{N}} m_k$, with $\mathcal{N} \triangleq \{1, \ldots, N\}$, and the encoding of a packet gives rise to N_tot/r coded bits, the transmission time T_packet of a single packet is given by

$$T_\mathrm{packet} = \frac{N_\mathrm{tot}}{r\sum_{k\in\mathcal{N}} m_k} T_\mathrm{s}, \tag{4.12}$$

where $T_\mathrm{s} = (N + \nu)T$ denotes the duration of a single OFDM symbol. Normalizing the GP by dividing by N/T_s, the GP in bits per subcarrier per OFDM symbol is expressed as

$$\mathrm{GP}(\phi, \mathbf{SNR}) = \frac{N_\mathrm{p} r}{N_\mathrm{tot} N}\left(\sum_{k\in\mathcal{N}} m_k\right)\left(1 - \mathrm{PER}(\phi, \mathbf{SNR})\right). \tag{4.13}$$

The factor $N_\mathrm{p}/N_\mathrm{tot}$ denotes the packet efficiency; $\frac{1}{N}\sum_{k\in\mathcal{N}} m_k$ is the average number of coded bits per subcarrier.

When the frequency-selective channel is time-varying, with a coherence time much larger than the OFDM symbol duration, the channel can be considered static over a number of OFDM symbols. As we assume that the channel is stationary for the whole packet transmission duration, the OFDM symbol index i can be omitted in the sequel. For each OFDM symbol, the TM $\phi = \{\mathbf{m}, r\}$ is selected according to the actual vector $\mathbf{SNR}$, which depends on the channel realization. As a consequence, the transmission time and error probability of a packet depend on the corresponding channel realization. In appendix 4.B we point out that the resulting GP equals the long-term average of the GP from (4.13), i.e.,

$$\mathrm{GP}_\mathrm{avg} = \mathbb{E}\left[\mathrm{GP}(\phi, \mathbf{SNR})\right], \tag{4.14}$$

where the average is taken over the distribution of the channel coefficients.

So far, we have assumed that the TX selects the TM according to the value of $\mathbf{SNR}$, which depends on the channel realization. However, in practice, the TX has no perfect knowledge about the channel realization: the channel state information available to the TX is imperfect. As a result, the TM is a function of the available channel information, rather than the actual channel realization. It is easily verified that the reasoning from appendix 4.B also applies to the case of imperfect channel state information: the result (4.14) is still valid, but the expectation is now over both the actual channel and the available channel information.

4.4 Effective SNR Mapping

When we aim to select the TM ϕ such that the GP is maximized, we need to know the relationship between ϕ and $\mathrm{GP}(\phi, \mathbf{SNR})$ from (4.13). However, as there is no simple analytical expression for $\mathrm{PER}(\phi, \mathbf{SNR})$, it is not obvious how changing the TM will impact the GP performance.

This problem can be circumvented by introducing an approximation technique which is known as effective SNR mapping (ESM) [54]. This technique maps the vector $\mathbf{SNR}$ to a scalar value γ_{eff}, called effective SNR, by means of (preferably simple) analytical expressions. The meaning of this effective SNR is that a coded system operating at SNR = γ_{eff} on an AWGN channel yields approximately the same PER as the system under consideration, characterized by $(\phi, \mathbf{SNR})$. The PER on the AWGN channel as a function of γ_{eff} is easily obtained in the form of a table, through computer simulations. Dealing with a scalar γ_{eff} instead of the vector $\mathbf{SNR}$ will considerably simplify the description of the resource allocation problem.

Several ways to perform the ESM have been proposed. First, we will briefly describe two general mapping functions, which can be used for any coded multi-carrier system: the exponential ESM (EESM) and the mutual information based ESM (MIESM). Thereafter, we will have a more extended discussion about the cumulant generating function based ESM (κESM), which was specifically designed for a BIC-OFDM system. It is this latter mapping function that will be used in this dissertation. It should be noted that the approximation of the GP, resulting from the application of κESM, will be used only with the purpose of facilitating the resource allocation; when evaluating the GP performance, resulting from the resource allocation obtained through κESM, the actual system will be simulated.

4.4.1 Exponential ESM

The EESM assumes a uniform bit allocation [55], i.e., $m = m_k$ ($\forall k \in \mathcal{N}$); hence, the TM is given by $\phi = (m, r)$. A simple analytic expression is proposed for the

relation between γ_{eff} and $\mathbf{SNR}$, such that

$$\text{PER}(\phi, \mathbf{SNR}) \approx \text{PER}_{\text{AWGN}}(\phi, \gamma_{\text{eff}}), \tag{4.15}$$

where $\text{PER}_{\text{AWGN}}(\phi, \gamma_{\text{eff}})$ is the PER on an AWGN channel, for the same code and the same constellation as the system under consideration. Based on some approximations of the pairwise error probability for coded systems (see [56] for more details), the EESM is given by

$$\gamma_{\text{eff}} = -\beta \ln \left(\frac{1}{N} \sum_{k \in \mathcal{N}} \exp\left(-\frac{\gamma_k}{\beta} \right) \right), \tag{4.16}$$

where γ_k is the *k*th component of the vector $\mathbf{SNR}$, and β is a scaling factor which has to be optimized offline for each TM $\phi = (m, r)$.

The optimized value of β for given ϕ is obtained in the following way. First, independent vectors $\mathbf{SNR}^{(n)}$ are generated according to the channel statistics, and the corresponding $\text{PER}^{(n)}(\phi) = \text{PER}(\phi, \mathbf{SNR}^{(n)})$ are obtained, for $n = 1, ..., I$. Next, $\text{PER}_{\text{AWGN}}(\phi, \gamma_{\text{eff}})$ is determined for a suitable range of γ_{eff}, and $\gamma_{\text{AWGN}}^{(n)}(\phi)$ is calculated by numerically solving $\text{PER}^{(n)}(\phi) = \text{PER}_{\text{AWGN}}(\phi, \gamma_{\text{AWGN}}^{(n)})$, for $n = 1, ..., I$. Finally, the parameter β is selected such that

$$\sum_{n=1}^{I} |10 \log(\gamma_{\text{AWGN}}^{(n)}(\phi)) - 10 \log(\gamma_{\text{eff}}^{(n)}(\beta))|^2, \tag{4.17}$$

is minimized, where $\gamma_{\text{eff}}^{(n)}(\beta)$ results from (4.16) using $\mathbf{SNR} = \mathbf{SNR}^{(n)}$. This optimization must be performed for each TM.

The performance of the BICM can be predicted by looking up the PER of the equivalent system with the same $\phi = (m, r)$ which operates on an AWGN channel with SNR equal to γ_{eff}. These reference curves, denoted $\text{PER}_{\text{AWGN}}(\phi, \gamma_{\text{eff}})$, can be stored in a lookup table for each combination of constellation size and code rate (m, r) from the set $\mathcal{D}_m \times \mathcal{D}_r$.

The EESM is limited to systems with uniform bitloading, whereas we intend to explore also non-uniform bitloading; therefore, EESM will not be further considered in this thesis.

4.4.2 Mutual Information ESM

The MIESM is based on the observation that the relation between the PER after decoding and the *received bit mutual information rate* (RBIR) is nearly independent from the bit allocation vector $\mathbf{m}$ [57]. The RBIR is defined as

$$\text{RBIR} = \frac{\sum_{k \in \mathcal{N}} f_{m_k}(\gamma_k)}{\sum_{k \in \mathcal{N}} m_k}, \tag{4.18}$$

where $f_m(\gamma)$ denotes the mutual information between a 2^m-QAM symbol x and the corresponding observation y, after transmission over an AWGN channel with SNR equal to γ. More specifically, considering equiprobable constellation points, we have

$$f_m(\gamma) = m - \mathbb{E}_{x,y}\left[\log_2\left(\frac{\sum_{x'\in\chi} p(y|x',\gamma)}{p(y|x,\gamma)}\right)\right], \tag{4.19}$$

where χ is the set of the 2^m-QAM constellation points, $p(y|x,\gamma) \sim N_{\mathrm{c}}(x, \frac{1}{\gamma})$ and the expectation is over the joint distribution of x and y. Hence, the system under consideration has nearly the same error performance as a BPSK system using the same encoding and operating on an AWGN channel with SNR equal to γ_{eff}, provided that both systems yield the same RBIR. As the RBIR for the latter system is given by $f_1(\gamma_{\mathrm{eff}})$, the MIESM yields

$$\gamma_{\mathrm{eff}} = f_1^{-1}\left(\frac{\sum_{k\in\mathcal{N}} f_{m_k}(\gamma_k)}{\sum_{k\in\mathcal{N}} m_k}\right). \tag{4.20}$$

The performance of the BICM can be predicted by looking up the PER of the equivalent BPSK system which operates on an AWGN channel with SNR equal to γ_{eff}. These reference curves, denoted $\mathrm{PER}_{\mathrm{AWGN}}(r, \gamma_{\mathrm{eff}})$, can be stored in a lookup table for each code rate r from the set $\mathcal{D}_r$.

The advantage of MIESM over the EESM is that the former is suitable for systems with non-uniform bit allocation. However, as no closed-form expressions are available for the functions $f_m(\gamma)$ the application of MIESM in the context of resource allocation is rather involved, requiring the use of tables of the functions $f_m(\gamma)$.

4.4.3 Cumulant Generating Function based ESM

In [58] a mapping function, called κESM, was introduced specifically for BICM systems that use either convolutional or turbo codes. The mapping function relies on an accurate evaluation of the pairwise error probability, and is shown to be very accurate and simple to use. It is for this reason that we prefer this mapping function over both the EESM, which is not suited for non-uniform bit allocations, and the MIESM, which is more complicated than the κESM.

First, we introduce the cumulant generating function $\kappa(s)$ of the LLR of the codeword bits, which is defined as

$$\kappa(s) \triangleq \ln \mathbb{E}\left[e^{s\mathrm{LLR}_{j,k}(y_k)} | \mathbf{SNR}\right], \tag{4.21}$$

where $\mathrm{LLR}_{j,k}(y_k)$ is the LLR (4.5) related to the *j*th bit from the *k*th symbol; the expectation is over the bit position j in the *k*th symbol, the constellation points of the *k*th symbol, the subcarrier index k and the noise. The value of s is chosen such that $\frac{\mathrm{d}\kappa(s)}{\mathrm{d}s} = 0$; it can be shown that s has to be equal to $1/2$.

The authors in [59] have shown that the pairwise error probability, $\mathrm{PEP} \triangleq \Pr\{p(\mathbf{y}|\mathbf{c}') > p(\mathbf{y}|\mathbf{c})|\mathbf{c}, \mathbf{SNR}\}$, resulting from a BICM system, can for high SNRs be interpreted as the decoding error probability of an equivalent BPSK system operating on an AWGN channel with SNR equal to $-\kappa(1/2)$. In [59], it is also shown that $\kappa(1/2)$ can be approximated as

$$\kappa(1/2) = \ln \mathbb{E}\left[e^{-\gamma_k d^2(x_k, x_k', j)/4} | \mathbf{SNR}\right]. \tag{4.22}$$

Because the system is symmetric, we can assume that a 0 was transmitted. This allows us to calculate the expectation in (4.22) by averaging over the bit position j in the kth symbol, the constellation points x_k from the set $\mathcal{X}_{k,0}^j$, and the subcarrier index k. In (4.22), $d(x_k, x_k', j)$ is the Euclidean distance between the constellation points x_k and x_k'; x_k' is defined as the constellation point from the set $\mathcal{X}_{k,1}^j$ which is closest to x_k. Assuming Gray mapping and QAM constellations, it is shown in [58] that equation (4.22) can be expressed as

$$\kappa(1/2) = \ln\left(\frac{1}{\sum_{k\in\mathcal{N}} m_k} \sum_{k\in\mathcal{N}} \frac{1}{2^{m_k-1}} \sum_{\mu=1}^{\frac{\sqrt{2^{m_k}}}{2}} \psi^k(\mu) e^{-\gamma_k(\mu d_{k,\min})^2/4}\right), \tag{4.23}$$

where $d_{k,\min}$ denotes the minimum Euclidean distance of the constellation used on the kth subcarrier, and $\psi^k(\mu)$ is linked to the number of neighbors at distance $\mu d_{k,\min}$. The function $\psi^k(\mu)$ is given by

$$\psi^k(\mu) = \begin{cases} 4\delta_{\mu-1}, & \text{if } m_k = 2 \text{ (4-QAM)} \\ 24\delta_{\mu-1} + 8\delta_{\mu-2} & \text{if } m_k = 4 \text{ (16-QAM)} \\ 112\delta_{\mu-1} + 48\delta_{\mu-2} + 16\delta_{\mu-3} + 16\delta_{\mu-4} & \text{if } m_k = 6 \text{ (64-QAM)} \end{cases}. \tag{4.24}$$

Based on these results, the κESM mapping function is introduced as

$$\gamma_{\text{eff}} = -\beta \ln\left(\frac{1}{\sum_{k\in\mathcal{N}} m_k} \sum_{k\in\mathcal{N}} \Omega_k\right), \tag{4.25}$$

where

$$\Omega_k = \sum_{\mu=1}^{\frac{\sqrt{2^{m_k}}}{2}} \frac{\psi^k(\mu)}{2^{m_k-1}} e^{-\gamma_k(\mu d_{k,\min})^2/4\beta}, \tag{4.26}$$

and β is a scaling factor that can be optimized across all possible code rates $r \in \mathcal{D}_r$ [60]. Using a similar reasoning as for EESM, the value of the tuning parameter β is selected to minimize

$$\max_r \sum_{n=1}^{I} \left| 10\log_{10}\left(\gamma_{\text{AWGN}}^{(n)}(r)\right) - 10\log_{10}\left(\gamma_{\text{eff}}^{(n)}(\beta)\right) \right|^2, \tag{4.27}$$

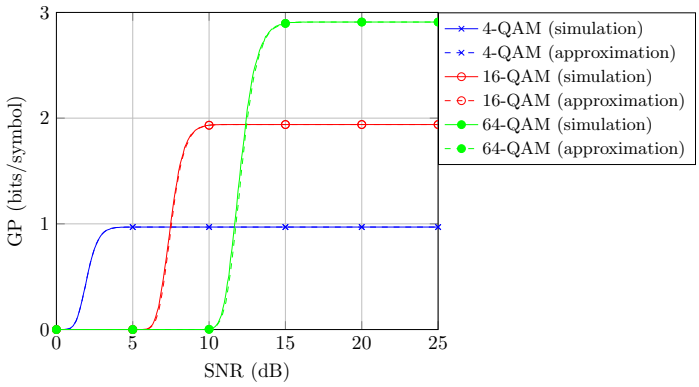


Figure 4.4: Illustration of the simulated and approximated GP in an AWGN channel ($r = 1/2$, $m = 2, 4, 6$).

where $\gamma_{\mathrm{AWGN}}^{(n)}(r)$ has the same meaning as $\gamma_{\mathrm{AWGN}}^{(n)}(\phi)$ from (4.17) with $\phi = (1, r)$ (BPSK constellation), and $\gamma_{\mathrm{eff}}^{(n)}(\beta)$ results from (4.25) instead of (4.16).

The performance of the BICM can be predicted by looking up the PER of the equivalent BPSK system which operates on an AWGN channel with SNR equal to γ_{eff}. These reference curves will be denoted $\mathrm{PER}_{\mathrm{AWGN}}(r, \gamma_{\mathrm{eff}})$ and can be stored in a lookup table for each code rate r from the set $\mathcal{D}_r$.

Fig. 4.4 shows the simulated and approximated GP over an AWGN channel as a function of the SNR $\gamma = E_s/\sigma^2$. The packet is transmitted using a single-carrier modulation, with a code rate r equal to $1/2$. We consider the following constellations: 4-QAM ($m = 2$), 16-QAM ($m = 4$) and 64-QAM ($m = 6$). The total packet length N_{tot} is equal to 1056 bits, where the number of information bits N_{p} and CRC bits N_{CRC} is equal to 1024 and 32, respectively; after convolutional encoding, $m_{\mathrm{tot}} = 2112$ coded bits are obtained. The simulated GP is calculated according to (4.13) in bits/symbol. The approximated GP can be found by substituting $\mathrm{PER}(\phi, \mathbf{SNR})$ in (4.13) by $\mathrm{PER}_{\mathrm{AWGN}}(r, \gamma_{\mathrm{eff}})$. It can be seen that the approximated GP is very close to the simulated GP in all three cases. In fact when $m = 2$, it can be shown that the approximated GP is equal to the simulated GP. This can be understood by noting that: a) a 4-QAM constellation with Gray-mapping is equivalent to a BPSK constellation where the energy per symbol is equal to $\gamma/2$, b) γ_{eff} in (4.25) becomes equal to $\gamma/2$ when $m = 2$. For 16-QAM and 64-QAM we notice a small SNR offset of 0.1 dB or less between the simulated and approximated GP.

4.5 Chapter Summary

In the beginning of this chapter we introduced a packet-based transmission system. Then, we discussed the various building blocks of a system that uses BICM. In the last sections, we introduced the GP metric and discussed several mapping functions which allow us to analyze the dependency of the PER on the vector $\mathbf{SNR}$.

The third part of this dissertation will use these concepts for the optimization of the GP performance of a packet-based SU network.

4.A Cyclic Redundancy Check

The calculation of a CRC of length N_{CRC} is based on a N_{CRC}-degree polynomial $g(x)$, which characterizes the type of CRC. The N_{p} payload bits denoted $(b_{N_p-1}, ..., b_0)$ are represented by the $(N_{\text{p}} - 1)$-degree polynomial $b(x) = \sum_{i=0}^{N_{\text{p}}-1} b_i x^i$, with $b_i \in \{0, 1\}$. The polynomial representation of the bit sequence consisting of the payload bits with N_{CRC} zeros appended is given by $x^{N_{\text{CRC}}} b(x)$. The polynomial representation of the CRC bits is obtained as $\text{CRC}(x) = \text{rem}(x^{N_{\text{CRC}}} b(x), g(x))$; $\text{rem}(a(x), b(x))$ denotes the remainder of the modulo-2 polynomial division of $a(x)$ by $b(x)$, which polynomials have their coefficients taken from the set $\{0, 1\}$. The polynomial representation of the resulting transmitted packet is given by

$$x^{N_{\text{CRC}}} b(x) + \text{CRC}(x), \tag{4.28}$$

which by construction has $g(x)$ as a factor. The RX verifies whether the polynomial representation $P(x)$ of the received packet is divisible by $g(x)$; when $\text{rem}(P(x), g(x)) = 0$ the received packet is considered correct, otherwise the received packet contains errors and is rejected. The computation of the remainder of the polynomial division can be carried out efficiently by means of shift-register operations. For more details, we refer to [61].

Let us illustrate the calculation of the CRC by means of example, where $g(x) = x^2 + 1$ and the payload bits are $(1, 0, 0, 1, 1)$, corresponding to $N_{\text{CRC}} = 2$ and $N_{\text{p}} = 5$. Appending two zeros to the payload bits yields $(1, 0, 0, 1, 1, 0, 0)$, which has the polynomial representation $x^6 + x^3 + x^2$. Considering that

$$x^6 + x^3 + x^2 = (x^4 + x^2 + x)(x^2 + 1) + x \tag{4.29}$$

we obtain $\text{CRC}(x) = x$, which corresponds to the CRC bits $(1, 0)$. Hence, the transmitted packet is given by $(1, 0, 0, 1, 1, 1, 0)$, with the first 5 bits and last 2 bits denoting the payload bits and the CRC bits, respectively; its polynomial representation is $x^6 + x^3 + x^2 + x$, which is divisible by $x^2 + 1$ (the quotient equals $x^4 + x^2 + x$).

Let us assume that the received packet equals $(1, 0, 0, 0, 1, 1, 0)$, i.e., a bit error has occurred on the fourth bit position; this yields $P(x) = x^6 + x^2 + x$. Considering that

$$x^6 + x^2 + x = (x^4 + x^2)(x^2 + 1) + x \tag{4.30}$$

the RX obtains $\text{rem}(P(x), g(x)) = x \neq 0$ and declares that the received packet is erroneous.

4.B The Long-Term Average of the Goodput

Let us consider L consecutive intervals, each equal to the channel coherence time T_{coh}. The channel is assumed to be constant over an interval of duration T_{coh}. The SNR vector and the TM associated with the *l*th interval are denoted $\mathbf{SNR}^{(l)}$ and $\phi^{(l)} = \{\mathbf{m}^{(l)}, r^{(l)}\}$; the latter is some function of $\mathbf{SNR}^{(l)}$; the resulting PER equals $\mathrm{PER}(\phi^{(l)}, \mathbf{SNR}^{(l)})$, where $\mathrm{PER}(\phi, \mathbf{SNR})$ denotes the PER corresponding to a TM ϕ and a SNR vector $\mathbf{SNR}$. During the *l*th interval, the number of packets transmitted equals

$$N_{\mathrm{pack}}^{(l)} = \frac{T_{\mathrm{coh}}}{T_{\mathrm{s}}} \frac{r^{(l)}}{N_{\mathrm{tot}}} \sum_{k \in \mathcal{N}} m_k^{(l)}. \tag{4.31}$$

The corresponding average number of correctly received information bits is given by

$$N_{\mathrm{info}}^{(l)} = N_{\mathrm{p}} N_{\mathrm{pack}}^{(l)} (1 - \mathrm{PER}(\phi^{(l)}, \mathbf{SNR}^{(l)})) \tag{4.32}$$

Considering all L intervals, the GP (in bit/s) is given by $\sum_{l=1}^{L} N_{\mathrm{info}}^{(l)} / (L T_{\mathrm{coh}})$. Multiplying this result by T_{s}/N yields the GP per subcarrier per OFDM symbol, which becomes

$$\mathrm{GP} = \frac{1}{L} \sum_{l=1}^{L} \mathrm{GP}(\phi^{(l)}, \mathbf{SNR}^{(l)}) \tag{4.33}$$

where $\mathrm{GP}(\phi, \mathbf{SNR})$ is given by (4.13). For increasing L, GP from (4.33) converges to the statistical average $\mathrm{GP}_{\mathrm{avg}} = \mathbb{E}\left[\mathrm{GP}(\phi, \mathbf{SNR})\right]$, where the expectation is over the channel statistics.

Part II

Optimization of Information-Theoretical Metrics

Overview

In this part of the dissertation, we will investigate the optimization of various information-theoretical metrics. In the following, we will provide the reader with a summary of the most important assumptions for each chapter. This overview is presented in table A.

From table A, we see that chapters 5, 6 and 7 investigate single-carrier networks that use the amplify-and-forward protocol. These chapters investigate the optimization of the outage probability of a secondary user network and clearly differ from chapter 8, where we investigate a multicarrier network that uses a decode-and-forward protocol. In chapter 8, we investigate the optimization of the transmission rate between the source and destination node of a secondary user network. As the channel state information is assumed to be imperfect, we will maximize the transmission rate of a SU network under a fixed outage probability constraint.

In most chapters, we assume that the number of primary user receivers (PU-RXs) is greater or equal than 1. Even most of the results shown in chapters 6 and 7, where we limit our discussion to the presence of a single primary user receiver, can be extended to a scenario with multiple primary user receivers.

As explained in section 3.6.2, it can be advantageous for the performance of the SU network if only the relay node which leads to the highest SNR at the DN participates in the transmission. Therefore, we always assume that the SU network uses relay selection. However, as the combination of this relay selection scheme with average interference constraints can be difficult to analyze, we first introduce the average interference constraints in chapter 6 without relay selection. We then combine the relay selection with the average interference constraints in chapter 7.

Chapter 5 will investigate the use of relay nodes with multiple transmit and receive antennas. As chapters 6 and 7 rely on an approximation of the outage probability which is not valid for multiple antennas, we limit our discussion in these chapters to relay nodes with a single transmit and receive antenna.

In chapters 6 and 7, we investigate the outage probability of the secondary user network under average interference constraints and compare this performance with the outage probability achieved under peak interference constraints. However, it will become clear that, although the use of average interference constraints leads to a significant performance benefit for the secondary user network, the derivation

of the corresponding resource allocation algorithms become more complex. It is for this reason that we turn our attention solely to the peak interference constraints for the more complicated networks discussed in chapters 5 and 8.

Finally, in this dissertation we investigate three different levels of channel state information (CSI): perfect CSI (PCSI), imperfect CSI (ICSI) and channel distribution information (CDI). We also make a distinction between the level of channel state information that is available at a transmitter about a) its channels to other nodes of the secondary user network and b) its interference channels to the primary user receivers.

In chapter 5, we consider all three cases of channel state information for both types of channels. However, the approximation for the outage probability that is used in chapters 6 and 7 is only valid for the case where the secondary user transmitters have CDI about their channels to the other secondary user nodes.

For the level of channel state information available about the interference channels, we consider both PCSI and CDI in chapter 6. However, our results can easily be extended to the case with ICSI. In chapter 7, we assume that only CDI is available about the interference channels and a possible extension of our results to the case of PCSI or ICSI is not straightforward.

Finally, chapter 8 investigates the case with ICSI for both types of channels. However, the results for PCSI can be derived by considering the case of ICSI and assuming that the variance of the estimation error is zero. In a similar manner, the results for the case with CDI can be found by assuming that the value of the channel estimate is zero and setting the variance of the estimation error equal to the variance of the channel gain.

	Chapter 5	Chapter 6	Chapter 7	Chapter 8
Relaying Protocol	Amplify-and-Forward	Amplify-and-Forward	Amplify-and-Forward	Decode-and-Forward
Type of Modulation	Single-Carrier	Single-Carrier	Single-Carrier	Multicarrier
Performance metric	Outage Probability	Outage Probability	Outage Probability	Transmission Rate
Number of PU-RXs	≥ 1	1	1	≥ 1
Relay Selection	yes	no	yes	yes
Number of antennas	≥ 1	1	1	1
Type of interference constraint	Peak	Peak and Average	Peak and Average	Peak
CSI about SU channels	PCSI, ICSI and CDI	CDI	CDI	ICSI
CSI about interference channels	PCSI, ICSI and CDI	PCSI and CDI	CDI	ICSI

Table A: Overview of the assumptions in chapter 5-8.

5

Distributed Dynamic Resource Allocation for Cooperative Cognitive Radio with Multi-Antenna Relay Selection

In section 3.7, we introduced the underlay paradigm. The biggest advantages of this paradigm is that it requires only a reduced (or null) cooperation between the SU and PU network and that it leads to a more efficient utilization of the spectrum. However, the interference level constraints that have to be met at the PU-RXs limit the achievable data rate performance.

In [33,35,62], a possible solution to alleviate the effect of the interference level constraints on the achievable data rate is proposed: the SU-TXs are equipped with multiple antennas to balance between optimizing the SU performance and avoiding interference to the PU-RXs, thus paving the way to the so-called *cognitive beamforming* concept. In [62], the multiple-input and multiple-output (MIMO) channels from the SU-TX to the SU-RX and PU-RXs are assumed to be perfectly known at the SU-TX. However, this assumption is quite unrealistic since the PU and SU units belong to independent networks, which implies a loose, or even absence of, cooperation between them. In [33], a more robust cognitive beamforming design is proposed which assumes that only the mean and covariance matrix

of the multiple-input and single-output (MISO) channel between the SU-TX and the PU-RX are available due to the loose cooperation between the SU and PU. Robustness is provided by keeping the interference to the PU-RX below a threshold for all realizations of the channel between the SU-TX and the PU-RX within a given uncertainty set. This robust design is extended in [35], where, in addition to the interfering channels, also the channel state information (CSI) about the SU channels is assumed to be imperfect.

Further, it was shown in [63, 64] that the cooperative relaying, discussed in section 3.6, can also be successfully used in CR networks to reduce the outage probability of the SU network. Although the design of a multi-antenna RN in a non-cognitive system was already investigated in [65, 66], these results cannot be used in a cognitive scenario as there is no guarantee that the beamforming will satisfy the interference level constraints at the PU-RXs. The design of the optimal beamforming matrix for a multi-antenna RN in a cognitive network is presented in [67]. Therein, it is assumed that the RNs are equipped with multiple antennas, while the PU-RXs, SU-SN and DN are all equipped with only a single antenna. In [68] and [69], this scenario is extended to the case where all the nodes of the PU and SU network are equipped with multiple antennas. In [69], the beamforming matrices for both the SN and RN are jointly optimized in order to maximize the transmission rate of the SU network. However, the proposed centralized solution requires perfect CSI (PCSI) of all the SU channels, as well as the interfering channels to the PU network. The assumption of PCSI about all the channels in a cooperative network is not realistic, because of estimation errors and feedback delays in a time-varying channel. A more realistic algorithm is proposed for a multi-antenna RN in [70], where the available CSI about the channels to the DN and the PU-RXs is assumed to be imperfect.

In this chapter (which is based on contribution [71]), after introducing the system model and the performance metric in section 5.1, we address the design of a distributed resource allocation algorithm in section 5.2 which minimizes the exact outage probability of a single-carrier SU network based on multi-antenna RNs operating in an amplify-and-forward mode in an underlay scenario. The resource allocation optimizes the power allocation and beamforming scheme for the SU-TXs. This scheme is optimal in the sense that it minimizes the exact outage probability of the SU network with relay selection, under both a transmit power constraint and a constraint on the interference power generated at every PU-RX. Section 5.3 starts by showing that the optimal cognitive beamforming matrix at the RN, considered in [67, 70], can be reduced to a cognitive beamforming vector. Then, several distributed resource allocation algorithms are presented for different levels of CSI at the SU-TXs: perfect CSI and imperfect CSI are considered, along with the case where only channel distribution information (CDI) is available. The algorithms allow the SU network to take into account different levels of CSI

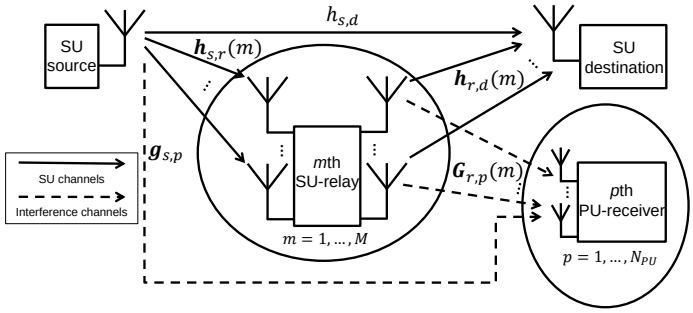


Figure 5.1: The cognitive radio relay network.

available at the SU-TX, (independently) for both the SU-TXs to PU-RXs interference channels and the channels toward the DN. The proposed approach allows to trade off the above levels of CSI, requiring different complexity and message exchange, against the achievable outage probability under the given interference constraints at the PU-RXs. By deriving different algorithms which depend on the available level of CSI, we extend the robust approach used in the single-hop scenarios from [33], [35] to a cooperative multi-antenna multi-relay scenario. In section 5.4, we show the numerical results, which prove that the multi-antenna relays can significantly improve the performance of the SU network, which is otherwise severely limited by the harsh interference constraints. Further, this section also shows how the number of relays, the number of antennas and the level of CSI impact the performance of the SU network. Finally, we point out that the proposed algorithms outperform several algorithms presented in literature. The chapter is concluded in section 5.5.

5.1 System Model

5.1.1 The Cooperative Network

The cognitive radio (CR) scenario we consider is based on a single-carrier SU network, consisting of a SN, a DN and M RNs. The RNs will use the selection amplify-and-forward scheme presented in section 3.6.2. We also assume that N_{PU} PU-RXs are active in the same frequency band. In Fig. 5.1, we present a CR network that is typical for device-to-device communications [70]: the SN and the DN are equipped with a single antenna, while the RNs have N_{a} transmit/receive antennas. It represents, for example, a scenario where two mobile phones can communicate directly with the help of the (multi-antenna) base station of a femtocell. The PU-RXs have K receive antennas.

In this chapter, we assume that all wireless channels follow a flat fading model, which was presented in section 3.1. The random variables $h_{\mathrm{s,d}} \in \mathbb{C}$, $\mathbf{h}_{\mathrm{s,r}}(m) \in \mathbb{C}^{N_{\mathrm{a}}\times 1}$, $\mathbf{h}_{\mathrm{r,d}}(m) \in \mathbb{C}^{N_{\mathrm{a}}\times 1}$, indicated in Fig. 5.1, denote zero-mean circularly symmetric complex Gaussian channel gains between the SN and the DN, between the SN and the N_{a} receive antennas of the *m*th RN, and between the N_{a} transmit antennas of the *m*th RN and the DN, respectively. In the sequel we make use of the covariance matrix of $\mathbf{h}_{\mathrm{r,d}}(m)$, defined as $\mathbf{R}_{\mathbf{h}}(m) \triangleq \mathbb{E}\left[\mathbf{h}_{\mathrm{r,d}}(m)\mathbf{h}_{\mathrm{r,d}}^{H}(m)\right]$. The coefficients of the Rayleigh-fading interference channels from the SN to the *p*th PU-RX and from the *m*th RN to the *p*th PU-RX will be denoted by $\mathbf{g}_{\mathrm{s},p} \in \mathbb{C}^{K\times 1}$ and $\mathbf{G}_{\mathrm{r},p}(m) \in \mathbb{C}^{N_{\mathrm{a}}\times K}$, respectively, with respective covariance matrices $\mathbf{R}_{\mathrm{s},p} \triangleq \mathbb{E}\left[\mathbf{g}_{\mathrm{s},p}\mathbf{g}_{\mathrm{s},p}^{H}\right]$ and $\mathbf{R}_{\mathrm{r},p}(m) \triangleq \mathbb{E}\left[\bar{\mathbf{g}}_{\mathrm{r},p}(m)\bar{\mathbf{g}}_{\mathrm{r},p}^{H}(m)\right]$, where $\bar{\mathbf{g}}_{\mathrm{r},p}(m) = \mathrm{vec}(\mathbf{G}_{\mathrm{r},p}(m))$ $(m = 1,\ldots,M; p = 1,\ldots,N_{\mathrm{PU}})$.

As described in section 3.6.2, the SN transmits its symbols in two time slots. In the first time slot, a symbol x, with $\mathbb{E}\left[|x|^2\right] = 1$, is transmitted to the DN and the M RNs. The signals received by the DN and the *m*th RN, are expressed as

$$z_{\mathrm{s,d}} = \sqrt{E_0}h_{\mathrm{s,d}}x + n_{\mathrm{s,d}}, \tag{5.1}$$

$$\mathbf{z}_{\mathrm{s,r}}(m) = \sqrt{E_0}\mathbf{h}_{\mathrm{s,r}}(m)x + \mathbf{n}_{\mathrm{s,r}}(m), \quad m = 1,\ldots,M, \tag{5.2}$$

where E_0 denotes the transmit energy per symbol used by the SN, and the noise terms $n_{\mathrm{s,d}}$ and $\mathbf{n}_{\mathrm{s,r}}(m)$ are distributed as $N_{\mathrm{c}}(0,\sigma_{\mathrm{s,d}}^2)$ and $N_{\mathrm{c}}(\mathbf{0},\sigma_{\mathrm{s,r}}^2(m)\mathbf{I}_{N_{\mathrm{a}}})$, respectively.

In the second time slot, the DN selects the RN which yields the highest SNR at the DN [27], and the selected *m*th RN, $m = 1,\ldots,M$, multiplies its received signal by a cognitive beamforming matrix $\mathbf{F}(m) \in \mathbb{C}^{N_{\mathrm{a}}\times N_{\mathrm{a}}}$ and forwards it to the DN. The signal received by the DN from the *m*th RN can be written as

$$z_{\mathrm{r,d}}(m) = \mathbf{h}_{\mathrm{r,d}}^{T}(m)\mathbf{F}(m)\mathbf{z}_{\mathrm{s,r}}(m) + n_{\mathrm{r,d}}(m), \tag{5.3}$$

where the noise term $n_{\mathrm{r,d}}(m) \sim N_{\mathrm{c}}(0,\sigma_{\mathrm{r,d}}^2(m))$. Further, the average transmit energy per symbol E_m of the *m*th RN is given by

$$E_m = E_0\|\mathbf{F}(m)\mathbf{h}_{\mathrm{s,r}}(m)\|^2 + \sigma_{\mathrm{s,r}}^2(m)\|\mathbf{F}(m)\|^2. \tag{5.4}$$

5.1.2 The Performance Metric

The metric that will be used to quantify the performance of the SU network is the link outage probability P_{out} between the SN and the DN, which we defined in section 3.4 as

$$P_{\mathrm{out}} \triangleq \Pr\{R_{\mathrm{S-AF}} \leq R\}, \tag{5.5}$$

where R is the desired rate at which the SN wants to transmit the information bits and $R_{\mathrm{S-AF}}$ is the instantaneous maximum rate of the SN-DN channel (including the selected relay channel).

The SNR at the DN associated with the *m*th RN is given by

$$\gamma_m = \frac{E_0 |\mathbf{h}_{\mathrm{r,d}}^T(m)\mathbf{F}(m)\mathbf{h}_{\mathrm{s,r}}(m)|^2}{\sigma_{\mathrm{s,r}}^2(m)\|\mathbf{h}_{\mathrm{r,d}}^T(m)\mathbf{F}(m)\|^2 + \sigma_{\mathrm{r,d}}^2(m)}. \tag{5.6}$$

At the DN, the best RN is selected, i.e., the one yielding the larger SNR γ_m [27]. Hence, after the second time slot (with only one RN transmitting out of M), maximum ratio combining of the signals received from the SN and the selected RN is applied. The overall received SNR at the DN of the amplify-and-forward cooperative network after maximum ratio combining yields

$$\gamma_{\mathrm{S-AF}} = \gamma_0 + \max_{m\in\{1,\ldots,M\}} \gamma_m, \tag{5.7}$$

where $\gamma_0 = E_0 \, |h_{\mathrm{s,d}}|^2/\sigma_{\mathrm{s,d}}^2$. Thus, the corresponding instantaneous maximum rate (in bit per channel use) is given by (3.67)

$$R_{\mathrm{S-AF}} = \frac{1}{2}\log_2(1+\gamma_{\mathrm{S-AF}}). \tag{5.8}$$

The corresponding outage probability is given by

$$P_{\mathrm{out}} = \Pr\{\gamma_{\mathrm{S-AF}} \leq 2^{2R} - 1\}, \tag{5.9}$$

where $\gamma_{\mathrm{S-AF}}$ depends on E_0, the cognitive beamforming matrices $\{\mathbf{F}(m),\, m = 1, ..., M\}$ and all channel gains $h_{\mathrm{s,d}}$, $\{\mathbf{h}_{\mathrm{s,r}}(m),\, m = 1, ..., M\}$ and $\{\mathbf{h}_{\mathrm{r,d}}(m),\, m = 1, ..., M\}$ from the SU network.

5.2 Resource allocation

The resource allocation consists of dynamically selecting the transmit energy E_0 at the SN and the cognitive beamforming matrix $\mathbf{F}(m)$ at the *m*th RN, $m = 1, \ldots, M$, such that P_{out} (5.9) is minimized, under transmit energy constraints at the SU-TXs, and interference constraints at the PU-RXs. The objective function and the constraints depend on the level of available CSI.

5.2.1 Available Channel Information

Considering a generic vector $\mathbf{f}$ of channel gains, PCSI refers to the case where the realization of $\mathbf{f}$ is known. In the case of ICSI, one has access only to an estimate $\hat{\mathbf{f}}$ of $\mathbf{f}$, and to the covariance matrix $\mathbf{R_e}$ of the corresponding estimation error $\mathbf{e} = \mathbf{f} - \hat{\mathbf{f}}$; the error is caused by noise and/or feedback delay (see appendix 5.A). In the case of CDI, only the channel covariance matrix $\mathbf{R_f}$ is known.

We assume that the RNs and the DN always have PCSI regarding their receiving channels, i.e., the *m*th RN and the DN know the realizations of $\mathbf{h}_{\mathrm{s,r}}(m)$ and

$\left\{h_{\mathrm{s,d}}, \mathbf{h}_{\mathrm{r,d}}^T(m)\mathbf{F}(m)\mathbf{h}_{\mathrm{s,r}}(m), m = 1, \ldots, M\right\}$, respectively. In order to perform the maximum ratio combining, the DN also has to know both the noise variances $\sigma_{\mathrm{s,d}}^2$ and $\sigma_{\mathrm{s,r}}^2(m)\|\mathbf{h}_{\mathrm{r,d}}^T(m)\mathbf{F}(m)\|^2 + \sigma_{\mathrm{r,d}}^2(m)$ which correspond to the signals received from the SN and from the selected RN, respectively.

PCSI of the interference and transmission channels is much harder to obtain in a time-varying environment, because it requires feedback from the other SU nodes or even from the PU network. Therefore, we will consider the optimization of P_{out} for the cases where the SN and the RNs have PCSI, ICSI or CDI regarding their interference channels ($\mathbf{g}_{\mathrm{s},p}$ and $\{\mathbf{G}_{\mathrm{r},p}(m),\, m = 1, \ldots, M\}$, respectively) to the PU-RXs and regarding their transmission channels ($\{h_{\mathrm{s,d}},\ \mathbf{h}_{\mathrm{s,r}}(m),\, m = 1, \ldots, M\}$ and $\{\mathbf{h}_{\mathrm{r,d}}(m),\, m = 1, \ldots, M\}$, respectively); later on, we will point out that the SN needs no knowledge about its transmission channels.

5.2.2 Transmit Energy Constraints

We impose the following constraint on the transmit energy of the SN

$$0 \leq E_0 \leq E_0^{(\max)}, \tag{5.10}$$

while the transmit energy of the *m*th RN, $m = 1, \ldots, M$, is constrained by

$$0 \leq E_m \leq E_m^{(\max)}, \tag{5.11}$$

where E_m is given by (5.4). In the above, $E_m^{(\max)}$ denotes the maximal transmit energy per symbol that the *m*th SU node, $m = 0, \ldots, M$, is able or allowed to transmit.

5.2.3 The Interference Constraints

Denoting by $\mathcal{I}_{\mathrm{s},p}$ and $\mathcal{I}_{\mathrm{r},p}(m)$ the interference at the *p*th PU-RX, $p = 1, \ldots, N_{\mathrm{PU}}$, caused by the SN and the *m*th RN, respectively, we have

$$\mathcal{I}_{\mathrm{s},p} = E_0\|\mathbf{g}_{\mathrm{s},p}\|^2, \tag{5.12}$$

$$\mathcal{I}_{\mathrm{r},p}(m) = E_0\|\mathbf{G}_{\mathrm{r},p}^T(m)\mathbf{F}(m)\mathbf{h}_{\mathrm{s,r}}(m)\|^2 + \sigma_{\mathrm{s,r}}^2(m)\|\mathbf{G}_{\mathrm{r},p}^T(m)\mathbf{F}(m)\|^2. \tag{5.13}$$

According to the underlay paradigm, the interference at the PU-RXs should be below a given interference threshold Γ. The formulation of these interference constraints depends upon the level of CSI that is available at the transmitting nodes about their channel gains to the PU-RXs. In practice it will be very difficult for the SU network to obtain PCSI about the channel gains towards the PU-RXs. For this reason, we will also consider ICSI and CDI.

5.2.3.1 Interference Constraints with PCSI

When PCSI is available at the corresponding transmitting node, the interference constraints to be met by the SN and the *m*th RN, $m = 1, \ldots, M$, are expressed as

$$\mathcal{I}_{\mathrm{s},p} \leq \Gamma, \;\; p = 1, \ldots, N_{\mathrm{PU}}, \tag{5.14}$$

$$\mathcal{I}_{\mathrm{r},p}(m) \leq \Gamma, \;\; p = 1, \ldots, N_{\mathrm{PU}}, \tag{5.15}$$

where $\mathcal{I}_{\mathrm{s},p}$ and $\mathcal{I}_{\mathrm{r},p}(m)$ are given by (5.12) and (5.13), with $\mathbf{g}_{\mathrm{s},p}$ and $\mathbf{G}_{\mathrm{r},p}(m)$ denoting the actual channel realizations.

5.2.3.2 Interference Constraints with ICSI

Here we assume that the CSI at each SU-TX about its channel gain to the *p*th PU-RX is imperfect. Similar to $\hat{\mathbf{f}}$ in (5.63), the variables $\hat{\mathbf{g}}_{\mathrm{s},p}$ and $\hat{\bar{\mathbf{g}}}_{\mathrm{r},p}(m)$ denote the estimates of $\mathbf{g}_{\mathrm{s},p}$ and $\bar{\mathbf{g}}_{\mathrm{r},p}(m)$, respectively.

Based on the instantaneous channel gain given by (5.63), we define for given $\hat{\mathbf{g}}_{\mathrm{s},p}$, $\mathbf{B}_{\mathrm{s},p}$, $\hat{\bar{\mathbf{g}}}_{\mathrm{r},p}(m)$ and $\mathbf{B}_{\mathrm{r},p}(m)$ the following ellipsoid channel uncertainty sets [33]

$$\mathcal{U}_{\mathrm{s},p}(\hat{\mathbf{g}}_{\mathrm{s},p}, \mathbf{B}_{\mathrm{s},p}) \triangleq \left\{ \mathbf{g}_{\mathrm{s},p} : \mathbf{g}_{\mathrm{s},p} = \hat{\mathbf{g}}_{\mathrm{s},p} + \mathbf{B}_{\mathrm{s},p}\boldsymbol{\epsilon}_{\mathrm{s},p}, \|\boldsymbol{\epsilon}_{\mathrm{s},p}\|^2 \leq 1 \right\}, \tag{5.16}$$

$$\mathcal{U}_{\mathrm{r},p}(\hat{\bar{\mathbf{g}}}_{\mathrm{r},p}(m), \mathbf{B}_{\mathrm{r},p}(m)) \triangleq \Big\{ \bar{\mathbf{g}}_{\mathrm{r},p}(m) : \bar{\mathbf{g}}_{\mathrm{r},p}(m) = \hat{\bar{\mathbf{g}}}_{\mathrm{r},p}(m) + \mathbf{B}_{\mathrm{r},p}(m)\boldsymbol{\epsilon}_{\mathrm{r},p}(m), \|\boldsymbol{\epsilon}_{\mathrm{r},p}(m)\|^2 \leq 1 \Big\}, \tag{5.17}$$

where $\boldsymbol{\epsilon}_{\mathrm{s},p} \in \mathbb{C}^{K\times 1}$ and $\boldsymbol{\epsilon}_{\mathrm{r},p}(m) \in \mathbb{C}^{KN_{\mathrm{a}}\times 1}$. The variables $\hat{\mathbf{g}}_{\mathrm{s},p}$ and $\hat{\bar{\mathbf{g}}}_{\mathrm{r},p}(m)$ denote the center of the ellipsoids, while the variables $\mathbf{B}_{\mathrm{s},p}$ and $\mathbf{B}_{\mathrm{r},p}(m)$ determine their shape. We now choose $\mathbf{B}_{\mathrm{s},p}$ and $\mathbf{B}_{\mathrm{r},p}(m)$, $m = 1, \ldots, M$, as

$$\mathbf{B}_{\mathrm{s},p} = \sqrt{\frac{\chi^2_\alpha(2K)}{2}}\mathbf{R}^{\frac{1}{2}}_{\mathbf{e},\mathrm{s},p}, \tag{5.18}$$

$$\mathbf{B}_{\mathrm{r},p}(m) = \sqrt{\frac{\chi^2_\alpha(2KN_{\mathrm{a}})}{2}}\mathbf{R}^{\frac{1}{2}}_{\mathbf{e},\mathrm{r},p}(m), \tag{5.19}$$

where $\mathbf{R}_{\mathbf{e},\mathrm{s},p}$ and $\mathbf{R}_{\mathbf{e},\mathrm{r},p}(m)$ denote the error covariance matrix $\mathbf{R}_{\mathbf{e}}$ from appendix 5.A, corresponding to the substitutions $\mathbf{f} = \mathbf{g}_{s,p}$ and $\mathbf{f} = \bar{\mathbf{g}}_{\mathrm{r},p}(m)$, respectively, and $\chi^2_\alpha(l)$ denotes the α-percentile of the χ^2-distribution with l degrees-of-freedom. For the above values of $\mathbf{B}_{\mathrm{s},p}$ and $\mathbf{B}_{\mathrm{r},p}(m)$, the actual channel gains $\mathbf{g}_{\mathrm{s},p}$ and $\bar{\mathbf{g}}_{\mathrm{r},p}(m)$ belong with probability α to their respective sets $\mathcal{U}_{\mathrm{s},p}(\hat{\mathbf{g}}_{\mathrm{s},p}, \mathbf{B}_{\mathrm{s},p})$ and $\mathcal{U}_{\mathrm{r},p}(\hat{\bar{\mathbf{g}}}_{\mathrm{r},p}(m), \mathbf{B}_{\mathrm{r},p}(m))$. Hence, the task of the SU network is that of ensuring that the interference constraints (5.14)-(5.15) hold for *every* channel gain in $\mathcal{U}_{\mathrm{s},p}(\hat{\mathbf{g}}_{\mathrm{s},p}, \mathbf{B}_{\mathrm{s},p})$ and $\mathcal{U}_{\mathrm{r},p}(\hat{\bar{\mathbf{g}}}_{\mathrm{r},p}(m), \mathbf{B}_{\mathrm{r},p}(m))$. The parameter α denotes the desired level of robustness, namely the minimum probability for which the interference constraints (5.14)-(5.15) hold. Thus, the resulting interference constraints are

expressed as

$$\mathcal{I}_{\mathrm{s},p} \leq \Gamma, \;\; p = 1, \ldots, N_{PU}, \; \forall \mathbf{g}_{\mathrm{s},p} \in \mathcal{U}_{\mathrm{s},p}(\hat{\mathbf{g}}_{\mathrm{s},p}, \mathbf{B}_{\mathrm{s},p}), \tag{5.20}$$

$$\mathcal{I}_{\mathrm{r},p}(m) \leq \Gamma, \;\; p = 1, \ldots, N_{\mathrm{PU}}, \; \forall \bar{\mathbf{g}}_{\mathrm{r},p}(m) \in \mathcal{U}_{\mathrm{r},p}(\hat{\bar{\mathbf{g}}}_{\mathrm{r},p}(m), \mathbf{B}_{\mathrm{r},p}(m)). \tag{5.21}$$

5.2.3.3 Interference Constraints with CDI

This scenario can be seen as a special case of the previous section 5.2.3.2. When CDI is available, we only know the distribution of the channel gains $\mathbf{g}_{\mathrm{s},p}$ and $\bar{\mathbf{g}}_{\mathrm{r},p}(m)$. We now exploit the fact that if we fix the values of the estimated channel gains to the mean of the actual channel gains, i.e., $\hat{\mathbf{g}}_{\mathrm{s},p} = \mathbb{E}\left[\mathbf{g}_{\mathrm{s},p}\right] = \mathbf{0}$ and $\hat{\bar{\mathbf{g}}}_{\mathrm{r},p}(m) = \mathbb{E}\left[\bar{\mathbf{g}}_{\mathrm{r},p}(m)\right] = \mathbf{0}$, we can see the covariance matrices $\mathbf{R}_{\mathrm{s},p}$ and $\mathbf{R}_{\mathrm{r},p}(m)$, which correspond to the channel gains $\mathbf{g}_{\mathrm{s},p}$ and $\bar{\mathbf{g}}_{\mathrm{r},p}(m)$, respectively, as the error covariance matrices of the previous section. Therefore, we can still write the interference constraints as in equations (5.20)-(5.21), but with modified uncertainty sets (5.16)-(5.17). The values of $\mathbf{B}_{\mathrm{s},p}$ and $\mathbf{B}_{\mathrm{r},p}(m)$, $m = 1, \ldots, M$, are selected as

$$\mathbf{B}_{\mathrm{s},p} = \sqrt{\frac{\chi^2_\alpha(2K)}{2}} \mathbf{R}_{\mathrm{s},p}^{\frac{1}{2}}, \tag{5.22}$$

$$\mathbf{B}_{\mathrm{r},p}(m) = \sqrt{\frac{\chi^2_\alpha(2KN_{\mathrm{a}})}{2}} \mathbf{R}_{\mathrm{r},p}^{\frac{1}{2}}(m), \tag{5.23}$$

where $\mathbf{R}_{\mathrm{s},p}$ and $\mathbf{R}_{\mathrm{r},p}(m)$ were introduced in section 5.1.1.

5.2.4 The Objective Function

The aim of the resource allocation is to minimize the link outage probability (5.9) by dynamically allocating the SN transmit energy E_0 and the RN beamforming matrices $\mathbf{F}(m)$, $m = 1, ..., M$. Let us stack all channel gains from the SU network (i.e., $h_{\mathrm{s,d}}$, $\{\mathbf{h}_{\mathrm{s,r}}(m),\, m = 1, ..., M\}$ and $\{\mathbf{h}_{\mathrm{r,d}}(m),\, m = 1, ..., M\}$) into the vector $\mathbf{h}_{\mathrm{SU}}$. Denoting by $\mathbf{CSI}$ the channel information which is available at the SN and the RNs regarding the channel vector $\mathbf{h}_{\mathrm{SU}}$, the minimization of P_{out} is equivalent with choosing the value of E_0 and $\mathbf{F}(m)$, $m = 1, ..., M$, which minimizes

$$\Pr\{\gamma_{\mathrm{S-AF}} \leq 2^{2R} - 1 | \mathbf{CSI}\}, \tag{5.24}$$

for a given value of $\mathbf{CSI}$. Although we have presented the objective function in a centralized manner, it will be demonstrated in section 5.3 that for some types of CSI the resulting resource allocation algorithms are distributed: each node only needs CSI about its own channel gains to calculate its transmission parameters; for the other types of CSI where the optimum resource allocation algorithm is centralized, also a (suboptimum) distributed algorithm will be derived.

5.3 Minimization of the Link Outage Probability

In this section several distributed algorithms will be proposed to minimize the outage probability P_{out} of the SU network over the transmit energy E_0 and the beamforming matrices $\mathbf{F}(m)$, $m = 1, \ldots, M$, under the interference and transmit power constraints.

In section 5.3.1, we start by pointing out that the optimum beamforming matrix $\mathbf{F}(m)$ performs maximum ratio combining of the received signals followed by beamforming that is described by a beamforming *vector* $\mathbf{v}(m)$. This result simplifies the expression of the transmission and interference constraints, and it becomes clear that the constraints on E_0 and $\mathbf{F}(m)$, $m = 1, \ldots, M$, are independent of each other. The independence of the constraints, together with the observation that the outage probability P_{out} (5.9) is a monotonically decreasing function of E_0, allows to separate the optimization of E_0 from the optimization of $\mathbf{F}(m)$, $m = 1, \ldots, M$. It is quite obvious that the optimization problem of the SN reduces to the constrained maximization of the transmit energy E_0, irrespective of the level of CSI available at the SN regarding its transmission channel gains. This constrained maximization of E_0 is considered in section 5.3.2. Next, the beamforming vector $\mathbf{v}(m)$ is optimized in section 5.3.3, making a distinction between the different levels of channel information available at the RNs about the channel gains towards the DN. For each of these cases, we also consider three levels of CSI (PCSI, ICSI, CDI) for the interference channels, yielding a total of 9 combinations.

5.3.1 Structure of the Optimal Beamforming Matrix

The following theorem shows that the optimal beamforming matrix $\mathbf{F}(m)$ has a specific structure, which is valid irrespective of the channel knowledge at the RNs regarding their interference channels and channels to the DN.

Theorem 5.1. *Assuming the* m*th RN has a perfect knowledge of its channel vector* $\mathbf{h}_{\text{s,r}}(m)$ *from the SN, the optimal beamforming matrix* $\mathbf{F}(m)$*, which minimizes the outage probability under the transmit constraint (5.11) and the interference constraints that correspond to the level of available information on the interference channels, has the following structure*

$$\mathbf{F}(m) = \frac{\mathbf{v}^*(m)\mathbf{h}_{\text{s,r}}^H(m)/\|\mathbf{h}_{\text{s,r}}(m)\|}{\sqrt{E_0\|\mathbf{h}_{\text{s,r}}(m)\|^2 + \sigma_{\text{s,r}}^2(m)}}, \tag{5.25}$$

where $\mathbf{v}(m) \in \mathbb{C}^{N_{\text{a}} \times 1}$ *can be any complex vector.*

Proof. See appendix 5.B. □

It should be noted that the normalization present in the denominator of (5.25) implies that the RN has to determine the average received energy per symbol for

every frame. It follows from (5.25) that the signal $\mathbf{F}(m)\mathbf{z}_{\mathrm{s,r}}(m)$ transmitted by the *m*th RN can be expressed as $\mathbf{v}^*(m)z'_{\mathrm{s,r}}(m)$, where $\mathbf{v}(m)$ is a beamforming *vector* (to be further optimized), and

$$z'_{\mathrm{s,r}}(m) = \frac{\mathbf{h}^H_{\mathrm{s,r}}(m)\mathbf{z}_{\mathrm{s,r}}(m)}{\|\mathbf{h}_{\mathrm{s,r}}(m)\|\sqrt{E_0\|\mathbf{h}_{\mathrm{s,r}}(m)\|^2+\sigma^2_{\mathrm{s,r}}(m)}}, \tag{5.26}$$

results from applying maximum ratio combining of the signals received by the *m*th RN, followed by the proper scaling.

By substituting (5.25) in (5.4), (5.6) and (5.13), the energy E_m, the SNR γ_m and the interference $\mathcal{I}_{\mathrm{r},p}(m)$, $m = 1, ..., M$, can be expressed as

$$E_m = \|\mathbf{v}(m)\|^2, \tag{5.27}$$

$$\gamma_m = \frac{\gamma_{\mathrm{s,r}}(m)\gamma_{\mathrm{r,d}}(m)}{\gamma_{\mathrm{s,r}}(m)+\gamma_{\mathrm{r,d}}(m)+1}, \tag{5.28}$$

$$\mathcal{I}_{\mathrm{r},p}(m) = \|\mathbf{v}^H(m)\mathbf{G}_{\mathrm{r},p}(m)\|^2, \tag{5.29}$$

where

$$\gamma_{\mathrm{s,r}}(m) = E_0\|\mathbf{h}_{\mathrm{s,r}}(m)\|^2/\sigma^2_{\mathrm{s,r}}(m), \tag{5.30}$$

and

$$\gamma_{\mathrm{r,d}}(m) = |\mathbf{v}^H(m)\mathbf{h}_{\mathrm{r,d}}(m)|^2/\sigma^2_{\mathrm{r,d}}(m). \tag{5.31}$$

The constraints involving the optimum beamforming matrix $\mathbf{F}(m)$ can be expressed in terms of the beamforming vector $\mathbf{v}(m)$. It will be convenient to introduce the rank-1 matrix $\mathbf{S}(m) \triangleq \mathbf{v}(m)\mathbf{v}^H(m)$. The matrix $\mathbf{S}(m)$ is positive semi-definite, which is denoted as $\mathbf{S}(m) \succeq 0$.

- The transmit constraint (5.11) at the *m*th RN, $m = 1, \ldots, M$, becomes

$$\|\mathbf{v}(m)\|^2 \leq E_m^{(\max)}, \tag{5.32}$$

  or, equivalently,

$$\mathrm{Tr}\,(\mathbf{S}(m)) \leq E_m^{(\max)}. \tag{5.33}$$

- The interference constraint (5.15) to be met by the *m*th RN, $m = 1, \ldots, M$, in the case of PCSI reduces to

$$\|\mathbf{v}^H(m)\mathbf{G}_{\mathrm{r},p}(m)\|^2 \leq \Gamma,\ p = 1, \ldots, N_{\mathrm{PU}}, \tag{5.34}$$

  We transform this constraint into

$$\mathrm{Tr}\Big(\mathbf{S}(m)\mathbf{G}_{\mathrm{r},p}(m)\mathbf{G}^H_{\mathrm{r},p}(m)\Big) \leq \Gamma,\ \ p = 1, \ldots, N_{\mathrm{PU}}. \tag{5.35}$$

- In the case of ICSI, the interference constraint (5.21) related to the *m*th RN, $m = 1, \ldots, M$, is given by

$$\begin{aligned} \|\mathbf{v}^H(m)\mathbf{G}_{\mathrm{r},p}(m)\|^2 \quad &\leq \quad \Gamma, \ \forall \bar{\mathbf{g}}_{\mathrm{r},p}(m) \in \mathcal{U}_{\mathrm{r},p}(\hat{\bar{\mathbf{g}}}_{\mathrm{r},p}(m), \mathbf{B}_{\mathrm{r},p}(m)), \\ & \qquad p = 1, \ldots, N_{\mathrm{PU}} \end{aligned} \tag{5.36}$$

with $\mathbf{B}_{\mathrm{r},p}(m)$ given by (5.19). This can be rewritten as

$$\begin{aligned} \bar{\mathbf{g}}^H_{\mathrm{r},p}(m)\mathbf{S}_K(m)\bar{\mathbf{g}}_{\mathrm{r},p}(m) \quad &\leq \quad \Gamma, \ \forall \bar{\mathbf{g}}_{\mathrm{r},p}(m) \in \mathcal{U}_{\mathrm{r},p}(\hat{\bar{\mathbf{g}}}_{\mathrm{r},p}(m), \mathbf{B}_{\mathrm{r},p}(m)), \\ & \qquad p = 1, \ldots, N_{\mathrm{PU}} \end{aligned} \tag{5.37}$$

where $\mathbf{S}_K(m) = \mathbf{I}_K \otimes \mathbf{S}(m)$. Note that for given (m, p) the constraint on $\mathbf{S}(m)$ must be satisfied for a *continuum* of interference channel gains. However, we can convert this constraint by using the following $\mathcal{S}$-lemma [72].

Lemma 5.1. *Let*

$$f_j(\mathbf{z}) = \mathbf{z}^H\mathbf{A}_j\mathbf{z} + 2\Re(\mathbf{b}_j^H\mathbf{z}) + c_j, \quad j = 1, 2,$$

where $\mathbf{z} \in \mathbb{C}^{N\times 1}$, $\mathbf{A}_j \in \mathbb{C}^{N\times N}$ *is a Hermitian matrix,* $\mathbf{b}_j \in \mathbb{C}^{N\times 1}$ *and* $c_j \in \mathbb{R}$. *Suppose that there exists a* $\mathbf{z}_0$ *such that* $f_1(\mathbf{z}_0) < 0$. *Then the following two statements are equivalent:*

1. $f_2(\mathbf{z}) \leq 0$ *for every* $\mathbf{z} \in \mathbb{C}^{N\times 1}$ *for which* $f_1(\mathbf{z}) \leq 0$
2. *There exists some* $\lambda \geq 0$ *such that*

$$\lambda \begin{pmatrix} \mathbf{A}_1 & \mathbf{b}_1 \\ \mathbf{b}_1^H & c_1 \end{pmatrix} - \begin{pmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ \mathbf{b}_2^H & c_2 \end{pmatrix} \succeq 0.$$

Defining $f_1(\boldsymbol{\epsilon}_{\mathrm{r},p}(m)) = \boldsymbol{\epsilon}^H_{\mathrm{r},p}(m)\boldsymbol{\epsilon}_{\mathrm{r},p}(m) - 1$ and

$$\begin{aligned} f_2(\boldsymbol{\epsilon}_{\mathrm{r},p}(m)) = \ & \boldsymbol{\epsilon}^H_{\mathrm{r},p}(m)\mathbf{B}^H_{\mathrm{r},p}(m)\mathbf{S}_K(m)\mathbf{B}_{\mathrm{r},p}(m)\boldsymbol{\epsilon}_{\mathrm{r},p}(m) \\ & + 2\Re((\mathbf{B}^H_{\mathrm{r},p}(m)\mathbf{S}_K(m)\hat{\bar{\mathbf{g}}}_{\mathrm{r},p}(m))^H\boldsymbol{\epsilon}_{\mathrm{r},p}(m)) \\ & + \hat{\bar{\mathbf{g}}}^H_{\mathrm{r},p}(m)\mathbf{S}_K(m)\hat{\bar{\mathbf{g}}}_{\mathrm{r},p}(m) - \Gamma, \end{aligned} \tag{5.38}$$

we can use the S-lemma to rewrite the interference constraint (5.37) for given (m, p) as follows

$$\begin{pmatrix} \lambda_p(m)\mathbf{I}_{KN_a} - \mathbf{B}^H_{\mathrm{r},p}(m)\mathbf{S}_K(m)\mathbf{B}_{\mathrm{r},p}(m) & -\mathbf{B}^H_{\mathrm{r},p}(m)\mathbf{S}_K(m)\hat{\bar{\mathbf{g}}}_{\mathrm{r},p}(m) \\ -\hat{\bar{\mathbf{g}}}^H_{\mathrm{r},p}(m)\mathbf{S}_K(m)\mathbf{B}_{\mathrm{r},p}(m) & \Gamma - \lambda_p(m) - \hat{\bar{\mathbf{g}}}^H_{\mathrm{r},p}(m)\mathbf{S}_K(m)\hat{\bar{\mathbf{g}}}_{\mathrm{r},p}(m) \end{pmatrix} \succeq 0, \tag{5.39}$$

which has to hold for at least one non-negative value of $\lambda_p(m)$.

- In the case of CDI, the constraint related to the mth RN, $m = 1, \ldots, M$, is equal to (5.36) with $\mathbf{B}_{\mathrm{r},p}(m)$ given by (5.23) and $\hat{\bar{\mathbf{g}}}_{\mathrm{r},p}(m) = \mathbf{0}$, which yields

$$\|\mathbf{v}^H(m)\mathbf{G}_{\mathrm{r},p}(m)\|^2 \leq \Gamma,\ p = 1, \ldots, N_{\mathrm{PU}},\ \forall \bar{\mathbf{g}}_{\mathrm{r},p}(m) \in \mathcal{U}_{\mathrm{r},p}(\mathbf{0}, \mathbf{B}_{\mathrm{r},p}(m)), \tag{5.40}$$

or, equivalently, if $\lambda_p(m) \geq 0$ as

$$\begin{pmatrix} \lambda_p(m)\mathbf{I}_{KN_a} - \mathbf{B}_{\mathrm{r},p}^H(m)\mathbf{S}_K(m)\mathbf{B}_{\mathrm{r},p}(m) & \mathbf{0} \\ \mathbf{0} & \Gamma - \lambda_p(m) \end{pmatrix} \succeq 0, \quad p = 1, \ldots, N_{\mathrm{PU}}. \tag{5.41}$$

The latter can be simplified to

$$\mathbf{B}_{\mathrm{r},p}^H(m)\mathbf{S}_K(m)\mathbf{B}_{\mathrm{r},p}(m) \preceq \Gamma\mathbf{I}_{KN_a},\ p = 1, \ldots, N_{\mathrm{PU}}. \tag{5.42}$$

5.3.2 Optimal Transmit Energy E_0

The new formulation of the constraints, shown in section 5.3.1, clearly demonstrates that the optimization of the transmit energy E_0 can be separated from the optimization of $\mathbf{F}(m)$, $m = 1, \ldots, M$, without any performance loss. At the SN, the following optimization problem has to be solved

$$\begin{aligned} E_0^{(\mathrm{opt})} = &\ \arg\max_{E_0} E_0 \\ &\text{s.t. (5.10), intf constraints,} \end{aligned} \tag{5.43}$$

where "intf constraints" refers to the relevant interference constraints corresponding to the level of available knowledge about the interference channels. As the transmit constraint is given by (5.10), the SN needs information only about its interference channels.

When PCSI about the channel gains to the PU-RXs is available, the interference constraints (5.14) yield the following solution

$$E_0^{(\mathrm{opt})} = \min\left\{ \frac{\Gamma}{\|\mathbf{g}_{\mathrm{s},1}\|^2}, \ldots, \frac{\Gamma}{\|\mathbf{g}_{\mathrm{s},N_{\mathrm{PU}}}\|^2}, E_0^{(\max)} \right\}. \tag{5.44}$$

When ICSI is available, the interference constraints are given by (5.20). Using the lemma 5.1, this interference constraint for a given PU-RX p, $p = 1, \ldots, N_{\mathrm{PU}}$, can be rewritten as

$$\begin{pmatrix} \lambda_p\mathbf{I}_K - E_0\mathbf{B}_{\mathrm{s},p}^H\mathbf{B}_{\mathrm{s},p} & -E_0\mathbf{B}_{\mathrm{s},p}^H\hat{\mathbf{g}}_{\mathrm{s},p} \\ -E_0\hat{\mathbf{g}}_{\mathrm{s},p}^H\mathbf{B}_{\mathrm{s},p} & \Gamma - \lambda_p - E_0\|\hat{\mathbf{g}}_{\mathrm{s},p}\|^2 \end{pmatrix} \succeq 0, \quad \lambda_p \geq 0. \tag{5.45}$$

The optimal $E_0^{(\mathrm{opt})}$ is found by solving

$$\begin{aligned} E_0^{(\mathrm{opt})} = &\ \arg\max_{E_0, \lambda_1, \ldots, \lambda_{N_{\mathrm{PU}}}} E_0 \\ &\text{s.t. (5.10), (5.45),} \end{aligned} \tag{5.46}$$

which is a semidefinite program that can be solved in polynomial time using the software package CVX [5].

Finally, when only CDI is available, the solution is obtained by substituting $\hat{\mathbf{g}}_{\mathrm{s},p} = 0$ in (5.45), which can be simplified to

$$E_0 \mathbf{B}_{\mathrm{s},p}^H \mathbf{B}_{\mathrm{s},p} \preceq \Gamma \mathbf{I}_K, \quad p = 1, \ldots N_{\mathrm{PU}}, \tag{5.47}$$

where $\mathbf{B}_{\mathrm{s},p}$, $p = 1, \ldots, N_{\mathrm{PU}}$, is given by (5.22). The optimal $E_0^{(\mathrm{opt})}$ is found as

$$E_0^{(\mathrm{opt})} = \min \left\{ \frac{\Gamma}{\sigma_{\max}^2(1)}, \ldots, \frac{\Gamma}{\sigma_{\max}^2(N_{\mathrm{PU}})}, E_0^{(\max)} \right\}, \tag{5.48}$$

where $\sigma_{\max}(p)$ is the largest singular value of $\mathbf{B}_{\mathrm{s},p}$, $p = 1, \ldots, N_{\mathrm{PU}}$.

5.3.3 Optimal Relay Beamforming

In this subsection we will discuss the optimization problems for the different RNs. A distinction is made between the levels of information (PCSI, ICSI, CDI) at the RNs about their channel gains to the DN. We will first introduce the generic optimization problem

$$\begin{aligned} \mathbf{V}^{(\mathrm{opt})} = & \arg\min_{\mathbf{V}} \ \Pr\{\gamma_{\mathrm{S-AF}} \leq 2^{2R} - 1 | \mathbf{CSI}\} \\ & \text{s.t. (5.32), intf constraints, } m = 1, \ldots, M, \end{aligned} \tag{5.49}$$

where $\mathbf{V} \triangleq [\mathbf{v}(1), \mathbf{v}(2), \ldots, \mathbf{v}(M)]$. Depending upon the available information on the interference channels, the constraint functions (5.34), (5.36) or (5.40) have to be used, for PCSI, ICSI or CDI, respectively. As the constraints on $\mathbf{v}(m)$, $m = 1, \ldots, M$, are independent, we will show that we can split optimization problem (5.49) into M independent subproblems. In order to demonstrate this, we will assume that the RNs also know the value of E_0 and $h_{\mathrm{s,d}}$; however, it will turn out that this requirement is unnecessary for most cases. First, we introduce the conditional cumulative distribution function (CDF)

$$F_m(x) = \Pr\{\gamma_m \leq x | \mathbf{CSI}\}. \tag{5.50}$$

This allows us to rewrite the objective function from optimization problem (5.49) as

$$\Pr\{\max_{m \in \{1,\ldots,M\}} \gamma_m \leq 2^{2R} - 1 - \gamma_0 | \mathbf{CSI}\} = \prod_{m=1}^{M} F_m(2^{2R} - 1 - \gamma_0), \tag{5.51}$$

where the second step comes from the fact that the channels to and from the different RNs are assumed to be independent. The optimization problem for the *m*th

RN becomes, $m = 1, \ldots, M$,

$$\begin{aligned} \mathbf{v}^{(\mathrm{opt})}(m) = \ & \arg \min_{\mathbf{v}(m)} F_m(2^{2R} - 1 - \gamma_0) \\ & \text{s.t. (5.32), intf constraints.} \end{aligned} \tag{5.52}$$

It will become clear that this optimization problem is not necessarily concave [4], which makes it very hard to solve. Therefore, we will rewrite the optimization problem in terms of the rank-1 positive semi-definite matrix $\mathbf{S}(m) = \mathbf{v}(m)\mathbf{v}^H(m)$, which yields

$$\begin{aligned} \mathbf{S}^{(\mathrm{opt})}(m) = \ & \arg \min_{\mathbf{S}(m) \succeq 0, \lambda_1(m) \geq 0, \ldots, \lambda_{N_{\mathrm{PU}}}(m) \geq 0} F_m(2^{2R} - 1 - \gamma_0) \\ & \text{s.t. (5.33), intf constraints,} \end{aligned} \tag{5.53}$$

where the interference constraint functions are (5.35), (5.39) or (5.42). Note that $\lambda_p(m)$ only appears as an optimization variable if (5.39) is used. Further, we have dropped the rank-1 constraint $\mathrm{Rank}(\mathbf{S}(m)) = 1$, which makes optimization problem (5.53) a semidefinite program. However, by dropping the rank constraint, the matrix $\mathbf{S}^{(\mathrm{opt})}(m)$ from (5.53) in general can have a rank higher than 1. The optimal solution to optimization problem (5.52) is only found in the case that $\mathrm{Rank}(\mathbf{S}^{(\mathrm{opt})}(m)) = 1$.

5.3.3.1 Beamforming with PCSI

In the case where the RNs have perfect knowledge of their respective channel gains to the DN, the minimization of the objective function in (5.53) becomes equivalent with the maximization of γ_m for the given channel realizations. As γ_m is a monotonically increasing function of $\left|\mathbf{v}^H(m)\mathbf{h}_{\mathrm{r,d}}(m)\right|^2$, the minimization of $F_m(2^{2R} - 1 - \gamma_0)$ in (5.53) can be substituted by the maximization of $\left|\mathbf{v}^H(m)\mathbf{h}_{\mathrm{r,d}}(m)\right|^2 = \mathrm{Tr}\left(\mathbf{h}_{\mathrm{r,d}}(m)\mathbf{h}_{\mathrm{r,d}}^H(m)\mathbf{S}(m)\right)$. In appendix 5.C, we show that for the three levels of CSI regarding the interference channels, optimization problem (5.53) has a rank-1 solution. Hence, we can write $\mathbf{S}^{(\mathrm{opt})}(m) = \mathbf{v}^{(\mathrm{opt})}(m)(\mathbf{v}^{(\mathrm{opt})}(m))^H$, and $\mathbf{v}^{(\mathrm{opt})}(m)$ is the optimal solution of optimization problem (5.52).

5.3.3.2 Beamforming with ICSI

In the case where the RNs are assumed to have ICSI on their channel gains to the DN, the actual channel gain to the DN can be written as

$$\mathbf{h}_{\mathrm{r,d}}(m) = \hat{\mathbf{h}}_{\mathrm{r,d}}(m) + \mathbf{e}_{\mathrm{r,d}}(m), \tag{5.54}$$

where $\mathbf{e}_{\mathrm{r,d}}(m) \sim N_{\mathrm{c}}(0, \mathbf{R}_{\mathbf{e}}(m))$ and $\mathbf{R}_{\mathbf{e}}(m)$ is defined as $\mathbf{R}_{\mathbf{e}}$ in appendix 5.A, with $\mathbf{f} = \mathbf{h}_{\mathrm{r,d}}(m)$.

The corresponding objective function to be minimized is shown in (5.52), where among the variables affecting γ_m only the error vectors $\mathbf{e}_{\mathrm{r,d}}(m)$ are considered random, and the remaining variables are assumed to be known; for notational convenience, the conditioning on the known vector $\mathbf{CSI}$ will not be shown.

Using (5.28) we can manipulate the CDF (5.50) into

$$F_m(x) = \Pr\left[|\mathbf{v}^H(m)\mathbf{h}_{\mathrm{r,d}}(m)|^2 \leq \eta_m(x)\right], \tag{5.55}$$

when $0 \leq x < \gamma_{\mathrm{s,r}}(m)$ with $\eta_m(x) = \sigma^2_{\mathrm{r,d}}(m)x(\gamma_{\mathrm{s,r}}(m)+1)/(\gamma_{\mathrm{s,r}}(m)-x)$.

If $x \leq 0$ or $\gamma_{\mathrm{s,r}}(m) \leq x$, the objective function $F_m(x)$ is equal to 0 or 1, respectively, and independent of the value of $\mathbf{v}(m)$. When $0 \leq x < \gamma_{\mathrm{s,r}}(m)$, exploiting the fact that $\left|\mathbf{v}^H(m)\mathbf{h}_{\mathrm{r,d}}(m)\right|^2$ is distributed according to a scaled non-central χ^2 distribution, the objective function $F_m(x)$ can be rewritten as [73]

$$F_m(x) = 1 - Q\left(\sqrt{2a(m)}, \sqrt{2b(m,x)}\right), \tag{5.56}$$

where $Q(\cdot,\cdot)$ represents the first-order Marcum Q-function [74, eq. (4.33)], and $a(m)$ and $b(m,x)$ are both real-valued functions given by

$$a(m) = \frac{\left|\mathbf{v}^H(m)\hat{\mathbf{h}}_{\mathrm{r,d}}(m)\right|^2}{\mathbf{v}^H(m)\mathbf{R}_{\mathbf{e}}(m)\mathbf{v}(m)} \tag{5.57}$$

and

$$b(m,x) = \frac{\eta_m(x)}{\mathbf{v}^H(m)\mathbf{R}_{\mathbf{e}}(m)\mathbf{v}(m)}. \tag{5.58}$$

Using [75, eq. (7),(8)] it is easy to see that $Q\left(\sqrt{2a(m)}, \sqrt{2b(m,x)}\right)$ is increasing in $a(m)$ and decreasing in $b(m,x)$. Therefore, setting $\mathbf{v}^H(m)\mathbf{R}_{\mathbf{e}}(m)\mathbf{v}(m) = c$, c being a real-valued variable, the optimization problem (5.52) can be solved by maximizing $a(m)$ for each value of c. Hence, using $\mathbf{S}(m)$, the objective function in optimization problem (5.53) is changed to

$$\mathbf{S}(m,c) = \arg \max_{\mathbf{S}(m)\succeq 0, \lambda_1(m)\geq 0,\ldots,\lambda_{N_{\mathrm{PU}}}(m)\geq 0} \mathrm{Tr}\left(\mathbf{S}(m)\hat{\mathbf{h}}_{\mathrm{r,d}}(m)\hat{\mathbf{h}}^H_{\mathrm{r,d}}(m)\right), \tag{5.59}$$

with an additional constraint: $\mathrm{Tr}\left[\mathbf{R}_{\mathbf{e}}(m)\mathbf{S}(m)\right] = c$. Ideally, this modified optimization problem then has to be solved for every value of c, where $c \in [0, c_{\max}]$. In practice, we divide this interval in a sufficiently large number of points and solve this modified optimization problem for each point. The value of $c_{\max}$ can be found by solving optimization problem (5.53), but with the following objective function

$$c_{\max} = \max_{\mathbf{S}(m)\succeq 0, \lambda_1(m)\geq 0,\ldots,\lambda_{N_{\mathrm{PU}}}(m)\geq 0} \mathrm{Tr}\left[\mathbf{R}_{\mathbf{e}}(m)\mathbf{S}(m)\right]. \tag{5.60}$$

Finally, the optimal value of the original optimization problem (5.53) is found by substituting $\mathbf{S}(m,c)$ in (5.56) and then minimizing (5.56) over c. With c^* minimizing (5.56), the optimal value $\mathbf{S}^{(\mathrm{opt})}(m)$ equals $\mathbf{S}(m,c^*)$.

Because of the equality constraint $\mathrm{Tr}\,[\mathbf{R_e}(m)\mathbf{S}(m)] = c$, the results from appendix 5.C cannot be applied here. This means that in general $\mathrm{Rank}(\mathbf{S}^{(\mathrm{opt})}(m)) \geq 1$. In this case, a suboptimal rank-1 solution can be obtained from the matrix $\mathbf{S}^{(\mathrm{opt})}(m)$ by using the randomization approach described in [76]. First, we generate L independent random vectors $\mathbf{v}_l(m) \sim N_{\mathrm{c}}(0, \mathbf{S}^{(\mathrm{opt})}(m))$, $l = 1,\ldots,L$, which are then scaled such that each of them satisfies with equality the most stringent constraint. Finally, the vector $\mathbf{v}_l(m)$, $l = 1,\ldots,L$, which gives the largest value for the objective function is selected as an approximation for $\mathbf{v}^{(\mathrm{opt})}(m)$.

Note that for the minimization of (5.56) the *m*th RN has to know, apart from CSI about its own channel gains to the DN and PU-RXs, the instantaneous SNR value γ_0 on the direct SN-DN channel. Thus, this solution is not distributed as γ_0 depends on the realization of the channel gain $h_{\mathrm{s,d}}$ and the transmit energy E_0 selected by the SN. Therefore, we also consider a *distributed* version of this optimization problem, which minimizes (5.56) over c under the worst-case assumption on γ_0, that is $\gamma_0 = 0$. The performance of this distributed yet suboptimum solution will be compared to the optimum but centralized solution in section 5.4.4.

5.3.3.3 Beamforming with CDI

In this case, we assume that the RNs only know the distribution of their channel gains to the DN. We follow the same reasoning as in section 5.3.3.2, but now $\mathbf{h}_{\mathrm{r,d}}(m)$ is assumed to be a random variable with mean $\mathbf{0}$ and covariance matrix $\mathbf{R_h}(m)$. This means that in equation (5.56) we have $a(m) = 0$ and $b(m,x) = \eta_m(x)/\mathbf{v}(m)^H\mathbf{R_h}(m)\mathbf{v}(m)$. In section 5.3.3.2, it was shown that $F_m(x)$ is increasing with $b(m,x)$. Hence, the minimization of $F_m(2^{2R}-1-\gamma_0)$ in (5.53) can be substituted by the maximization of $\mathrm{Tr}\,(\mathbf{R_h}(m)\mathbf{S}(m))$. For the scenario with PCSI about the interference channels, we prove in appendix 5.C that we can always find a rank-1 solution to optimization problem (5.53). For the case of ICSI and CDI, it is possible that $\mathbf{S}^{(\mathrm{opt})}(m)$ has a rank higher than 1. In this case, a rank-1 approximation is found by using the randomization approach described in section 5.3.3.2.

5.4 Numerical Results

We consider the configuration as shown in Fig. 5.2. The SN and DN are located at coordinates $(0,0)$ and $(1.375,0)$, respectively, while the supporting RNs are assumed to be uniformly distributed inside an annulus with outer radius 1.25 and inner radius 0.25. The PU-RXs are uniformly distributed inside an annulus with

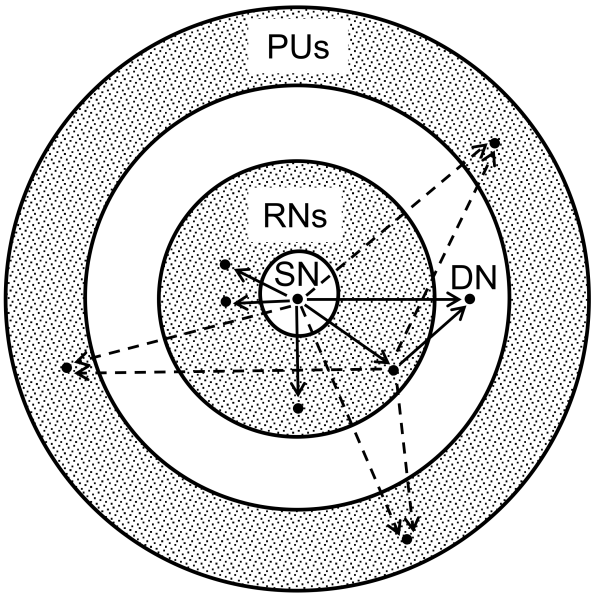


Figure 5.2: The topology of the network.

outer radius 2.5 and inner radius 1.5. The SN broadcasts its message in the first time slot, and in the second time slot only the best RN amplifies and forwards the message to the DN. The solid lines refer to the messages exchanged within the SU network, while the dashed ones refer to the interference from the SN and RNs to the PU-RXs.

The outage probabilities are calculated by means of Monte Carlo simulations. For each channel realization, we randomly select a different location for the RNs and the PU-RXs. We assume that $\mathbb{E}\left[|h_{\mathrm{s,d}}|^2\right] = 1/d_{\mathrm{s,d}}^2$, $\mathbb{E}\left[\mathbf{h}_{\mathrm{s,r}}(m)\mathbf{h}_{\mathrm{s,r}}^H(m)\right] = 1/d_{\mathrm{s,r}}^2(m)\mathbf{I}_{N_{\mathrm{a}}}$, $\mathbf{R}_{\mathbf{h}}(m) = 1/d_{\mathrm{r,d}}^2(m)\mathbf{I}_{N_{\mathrm{a}}}$, $\mathbf{R}_{\mathrm{s},p} = 1/d_{\mathrm{s},p}^2\mathbf{I}_K$ and $\mathbf{R}_{\mathrm{r},p}(m) = 1/d_{\mathrm{r},p}^2(m)\mathbf{I}_{KN_{\mathrm{a}}}$ ($m = 1,\ldots,M$; $p = 1,\ldots,N_{\mathrm{PU}}$), where $d_{\mathrm{s,d}}$, $d_{\mathrm{s,r}}(m)$, $d_{\mathrm{r,d}}(m)$, $d_{\mathrm{s},p}$ and $d_{\mathrm{r},p}(m)$ denote the distances between the corresponding nodes.

We take the same value $E^{(\max)}$ for the maximal transmit energies $E_m^{(\max)}$, $m = 0,\ldots,M$, and set the noise variances on all channels equal to σ^2. This allows to express the outage probabilities of the different scenarios as a function of $E^{(\max)}/\sigma^2$ (dB). In the case of ICSI, the estimation error σ_{e}^2 is chosen equal to $\sigma^2/(2E^{(\max)})$, the number of delayed channel estimates $P = 7$ and the update interval $D = 4096$ symbol intervals. For more information about these variables we refer to appendix 5.A. For all simulations, we take $R = 0.5$ bits/channel use, $\alpha = 0.9$, symbol interval $T = 50$ ns, Doppler frequency $f_{\mathrm{d}} = 144$ Hz and $\Gamma = \sigma^2$.

In the following, the performance of the relay network will be compared to a

SU direct-link network, which optimizes the following expression for the outage probability

$$P_{\text{out}} = \Pr\{\log_2(1 + \frac{E_0 \left|h_{\text{s,d}}\right|^2}{\sigma^2_{\text{s,d}}}) \leq R\}, \tag{5.61}$$

where the optimum value of E_0 is obtained from (5.44), (5.46) or (5.48), depending upon the available CSI on the channel gains to the PU-RXs. The main disadvantage of the relay network compared to the direct-link network is the requirement of an additional time slot.

We also note that in the case where $\text{Rank}(\mathbf{S}(m)) > 1$, a rank-1 solution is found by applying the randomization approach from section 5.3.3.2. The number of generated vectors L is chosen equal to 50. In our numerical results, we have found that the performance loss compared to the multi-rank solution is negligible. Therefore, in the case where $\text{Rank}(\mathbf{S}(m)) > 1$, only the performance curves corresponding to the rank-1 approximation will be shown.

Unless mentioned otherwise, in the case where ICSI about the channels to the DN is available, we will show the performance corresponding to the centralized optimization problem.

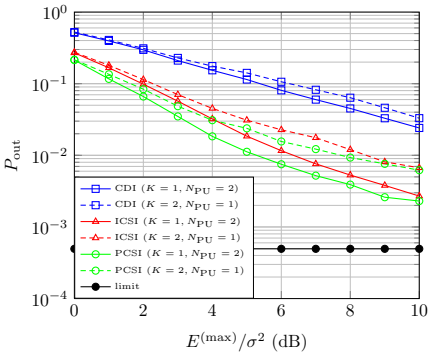
5.4.1 Influence of the Number of Antennas per PU-RX

We now consider the performance of a SU network with 2 RNs ($M = 2$); the RNs are each equipped with 3 transmit and receive antennas ($N_{\text{a}} = 3$). We show the performance in the scenario where there are two PU-RXs present ($N_{\text{PU}} = 2$) with a single antenna ($K = 1$) and the scenario where there is a single PU-RX ($N_{\text{PU}} = 1$) with two antennas ($K = 2$). Fig. 5.3 shows three subfigures, where each subfigure corresponds to a certain level of CSI about the interference channels. Each subfigure then shows the curves for all the levels of CSI about the coefficients between the RNs and the DN: PCSI, ICSI and CDI.

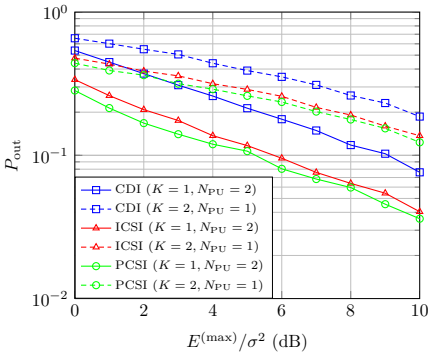
5.4.1.1 PCSI about Interference Channel

Fig. 5.3a shows the exact outage probabilities in the case where the SN and the RNs have PCSI about their interference channels. The availability of the PCSI will allow the RNs to transmit their beams away from the PU-RXs. The performances in the case of PCSI and ICSI about the channel coefficients to the DN are very similar, because both cases allow the RNs to steer their beams towards the DN and at the same time away from the PU antennas. However when only CDI is available about the channel gains to the DN, the RNs are only able to reduce the interference at the PU-RXs. This explains the large performance gap between the performance curves for CDI and PCSI.

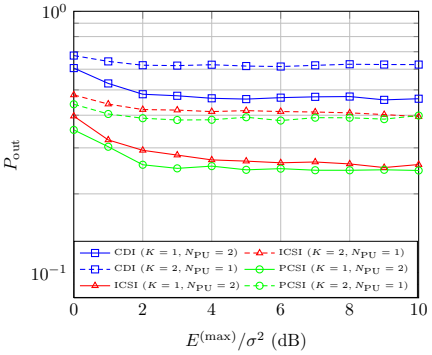
We also notice that the outage probability P_{out} is the highest in the case where $N_{\text{PU}} = 1$ and $K = 2$. This is explained by the fact that we consider an interference



(a) P_{out} versus $E^{(\max)}/\sigma^2$ (PCSI of intf channels).



(b) P_{out} versus $E^{(\max)}/\sigma^2$ (ICSI of intf channels).



(c) P_{out} versus $E^{(\max)}/\sigma^2$ (CDI of intf channels).

Figure 5.3: Comparison between a single PU-RX with $K = 2$ and two PU-RXs with $K = 1$ ($M = 2$, $N_{\text{a}} = 3$).

constraint per PU-RX, which means the interference constraint is harder to satisfy when $K > 1$.

Additionally, it is worth emphasizing that the outage probability P_{out} will eventually converge to a non-zero value for increasing $E^{(\max)}$. To explain this, we take a closer look at (5.7) and (5.28). In the case of PCSI about the channel gains to the DN, when $N_{\mathrm{a}} > KN_{\mathrm{PU}}$ and $E^{(\max)}$ goes to infinity, there are enough degrees of freedom available to select an infinitely large $\mathbf{v}(m)$ which is orthogonal to the columns of $\mathbf{G}_{\mathrm{r},p}(m)$ for $p = 1, ..., N_{\mathrm{PU}}$. This means that the optimal solution of (5.52) will result in a value for $\left|\mathbf{v}^H(m)\mathbf{h}_{\mathrm{r,d}}(m)\right|^2$ which is infinitely large, so that (5.28) can then be written as $\gamma_m = \gamma_{\mathrm{s,r}}(m) = \frac{E_0}{\sigma^2_{\mathrm{s,r}}(m)} \left\|\mathbf{h}_{\mathrm{s,r}}(m)\right\|^2$. Hence, using (5.5) and (5.28), the outage probability for infinite $E^{(\max)}$ can be written as

$$P_{\mathrm{out}} = \Pr\left\{ \frac{E_0^{(\mathrm{opt})} \left|h_{\mathrm{s,d}}\right|^2}{\sigma^2_{\mathrm{s,d}}} + \max_{m\in\{1,\ldots,M\}} \frac{E_0^{(\mathrm{opt})}}{\sigma^2_{\mathrm{s,r}}(m)} \left\|\mathbf{h}_{\mathrm{s,r}}(m)\right\|^2 \leq 2^{2R} - 1 \right\}, \tag{5.62}$$

where $E_0^{(\mathrm{opt})}$ is given by (5.44). This lower limit can be further decreased by adding more RNs to the SU network or by equipping the RNs with more antennas. We have shown also the limiting value (5.62) in Fig. 5.3a.

5.4.1.2 ICSI about Interference Channel

Fig. 5.3b shows the exact outage probabilities in the case where the SN and the RNs have ICSI about their interference channels. Comparing these curves with those from Fig. 5.3a clearly shows that the ICSI about the interference channels causes a significant performance loss. Hence, having only ICSI (instead of PCSI) on the channel gains to the PU antennas deteriorates the performance much more than having ICSI on the channel gains to the DN.

5.4.1.3 CDI about Interference Channel

Finally, Fig. 5.3c shows the exact outage probabilities in the case where the SN and the RNs only have CDI about their interference channels. In this scenario we notice that for increasing $E^{(\max)}/\sigma^2$ the outage probability quickly converges to a non-zero limiting value. This shows that the performance of the SU network is severely limited by the interference constraints: the RNs are unable to steer the transmit beam away from the PU-RXs, which means they do not benefit from the higher allowed transmit energy when $E^{(\max)}/\sigma^2$ is increased. This is even worse for the case where $N_{\mathrm{PU}} = 1$ and $K = 2$.

5.4.2 Influence of the Number of RNs

We compare the performance of a SU network with a single RN ($M = 1$) and with 2 RNs ($M = 2$). Further, we set $N_{\mathrm{a}} = 3$, $N_{\mathrm{PU}} = 2$ and $K = 1$. In Fig. 5.4a, we consider PCSI for the interference channels and the channels to the DN. As a reference we have also shown the performance of the maximum ratio reception (MRR)-orthogonally projected maximum ratio transmission (OPMRT) and MRR-maximum ratio transmission (MRT) schemes proposed in [77]. For the MRT algorithm, the beamforming matrix is given by $\mathbf{F}(m) = a\mathbf{h}_{\mathrm{r,d}}^{*}(m)\mathbf{h}_{\mathrm{s,r}}^{H}(m)$, where $a \in \mathbb{R}$ is chosen such that the constraints (5.11) and (5.13) are satisfied. For the OPMRT algorithm, the beamforming matrix is given by the projection of $a\mathbf{h}_{\mathrm{r,d}}^{*}(m)\mathbf{h}_{\mathrm{s,r}}^{H}(m)$ into the null space of $[\mathbf{G}_{\mathrm{r},1}(m), \ldots, \mathbf{G}_{\mathrm{r},N_{\mathrm{PU}}}(m)]^{T}$, where $a \in \mathbb{R}$ is chosen such that (5.11) is satisfied. It can be clearly seen that a significant performance improvement can be achieved by using the optimal beamforming algorithm from section 5.3.3.1, which we labeled PCSI in Fig. 5.4a. Further we notice that we get a performance gain by increasing the number of RNs M. This performance gain is explained by the fact that the DN is able to select the RN which has the most favorable channel conditions, which are: strong channel gain between SN and RN, a strong channel gain between the RN and the DN and a weak link between the RN and the PU-RXs.

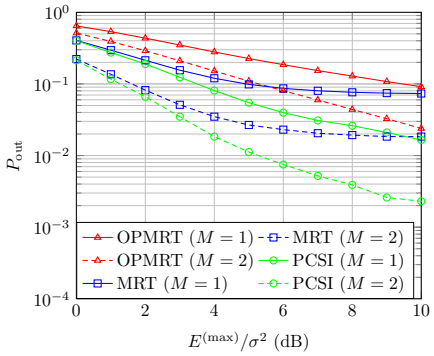
In Fig. 5.4b, we consider ICSI for the interference channels and the channels to the DN. We have also shown the performance of the robust beamforming (RB) proposed in [70]. In order to compare both algorithms in a fair way, the parameters $\mathbf{T}_p$ and $\mathbf{Q}$ described in [70] for the RB algorithm are set to $\mathbf{T}_p = 2\mathbf{R}_{\mathbf{e},\mathrm{r},p}^{-1}(m)/\chi_{\alpha}^{2}(2KN_{\mathrm{a}})$, $p = 1, \ldots, N_{\mathrm{PU}}$, and $\mathbf{Q} = 2\mathbf{R}_{\mathbf{e}}^{-1}(m)/\chi_{\alpha}^{2}(2N_{\mathrm{a}})$. In Fig. 5.4b, we show that by specifically optimizing the outage probability P_{out}, a higher reliability can be achieved compared to the RB algorithm which optimizes the worst-case capacity. Further, we again see a clear gain in performance by increasing the number of RNs from 1 to 2.

5.4.3 Number of Antennas

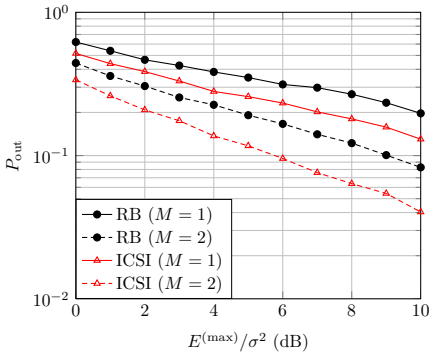
In Fig. 5.5, we take $M = 2$, $N_{\mathrm{PU}} = 2$, $K = 1$ and we compare the various scenarios for a different number of transmit and receive antennas at each RN: $N_{\mathrm{a}} = 2,\ 3$ and 4. As in section 5.4.1, we split the discussion according to the level of CSI about the interference channel coefficients.

5.4.3.1 PCSI about Interference Channel

Fig. 5.5a shows P_{out} in the case where the SU-TXs have PCSI about their interference channel. We also show the limiting values of P_{out} when $E^{(\max)}$ goes to infinity and PCSI is available about the channel gains to the DN. For $N_{\mathrm{a}} = 3, 4$,



(a) Beamforming with PCSI (PCSI of intf channels).



(b) Beamforming with ICSI (ICSI of intf channels).

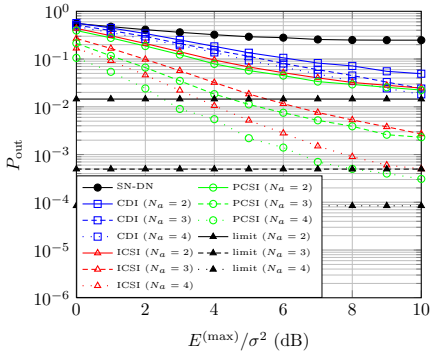
Figure 5.4: Comparison of our proposed algorithms and the MRR-OPMRT and MRR-MRT scheme in [77], and the robust beamforming scheme in [70] ($M = 1$ and 2, $N_{\text{a}} = 3$, $N_{\text{PU}} = 2$, $K = 1$)

the limit is given by (5.62), which is valid for $N_{\text{a}} > KN_{\text{PU}}$ only; for $N_{\text{a}} = 2$, we have solved (5.43) and (5.53) without transmit energy constraint. These lower limits clearly show that the largest performance improvement is achieved by going from 2 to 3 antennas. This is because we need at least $N_{\text{a}} = 3$ to avoid interference at the 2 PU-RXs; for $N_{\text{a}} < 3$, the RNs have to limit their transmit energies in order not to violate the interference constraints.

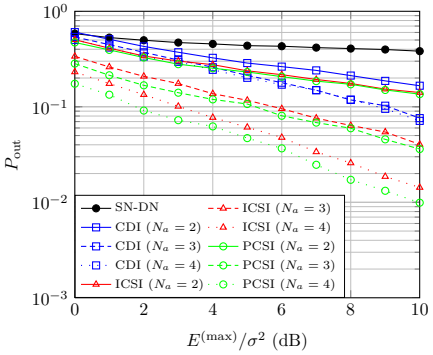
In the case of CDI, we see an improvement by going from $N_{\text{a}} = 2$ to $N_{\text{a}} = 3$, but only a very small improvement by increasing the number of antennas N_{a} from 3 to 4. The latter gain is only caused by the maximum ratio combining at the RN, while the former gain is also caused by the fact that the number of antennas N_{a} exceeds the number of PU-RX antennas KN_{PU}. However, in the case of ICSI and PCSI we keep noticing a large performance improvement by increasing the number of antennas at the RNs. The gain is explained by the maximum ratio combining and beamforming towards the DN. We note that the performance loss between the case of ICSI and PCSI is rather small (a difference in $E^{(\max)}/\sigma^2$ of around $1 - 2$ dB). Finally, Fig. 5.5a also shows the performance of the direct link SN-DN in the case of PCSI about the channel gains to the PU-RXs. The relay network outperforms the direct link network in all three cases.

5.4.3.2 ICSI about Interference Channel

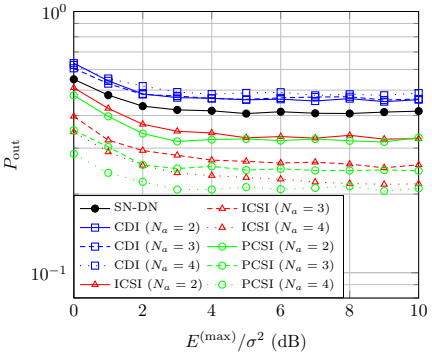
In Fig. 5.5b, the case where the SU-TXs have ICSI about their interference channel is addressed. In the case of CDI, we again notice that there is an improvement by



(a) P_{out} versus $E^{(\max)}/\sigma^2$ (PCSI of intf channels).



(b) P_{out} versus $E^{(\max)}/\sigma^2$ (ICSI of intf channels).



(c) P_{out} versus $E^{(\max)}/\sigma^2$ (CDI of intf channels).

Figure 5.5: Comparison between N_{a} = 2, 3 and 4 ($M = 2$, $N_{\text{PU}} = 2$, $K = 1$).

going from $N_{\mathrm{a}} = 2$ to $N_{\mathrm{a}} = 3$, while there is almost no performance difference between $N_{\mathrm{a}} = 3$ and $N_{\mathrm{a}} = 4$, since (i) transmit beamforming towards the DN is impossible, when only CDI is available about the channel gains to the DN, and (ii) at the RN, the gain offered by the maximum ratio combining is countered by the loss in performance caused by the uncertainty on the interference channels. Further, from (5.19), it is clear that the interference constraints are dependent on the factor $\chi^2_\alpha(2KN_{\mathrm{a}})$. This factor increases with the number of antennas N_{a}, which makes the interference constraints more stringent. In Fig. 5.5b, we also show the performance of the direct link SU network in the case where the SN has ICSI about its interference channel. Again, we can emphasize that the multi-antenna multi-relay network has a significantly better performance than the direct link network.

5.4.3.3 CDI about Interference Channel

In Fig. 5.5c, we consider the case where the SU-TXs have CDI about the interference channel. We observe that when only CDI is available about the channel gains to the DN, increasing the number of antennas at the RNs does not bring any performance gain and can even lead to a degradation caused by the more stringent interference constraints due to the larger value of the factor $\chi^2_\alpha(2KN_{\mathrm{a}})$. In the case of ICSI and PCSI about the channel gains to the DN, we notice only a very small performance improvement. It is clear that interference is the main limiting factor in this scenario, and increasing the number of antennas does not significantly improve the performance. Finally, in Fig. 5.5c we have also shown the performance of the direct link SU network which only has CDI about its interference channels. In this case, the multi-antenna multi-relay network does not provide any performance improvement, compared to the direct link, when only CDI on the channels to the DN is available, since beamforming is impossible and the multiple antennas are not able to avoid interference at the PU-RXs; when using the relays, the gain provided by maximum ratio combining at the DN is offset by the need for two time slots.

5.4.4 Centralized versus Distributed Solutions

In this subsection we assume that the SN and the RNs have ICSI about their channel coefficients to the DN. We make a distinction between PCSI, ICSI and CDI for the knowledge about the interference channels. Choosing $M = 2$, $N_{\mathrm{PU}} = 2$, $K = 1$ and $N_{\mathrm{a}} = 3$, the performance of the centralized and distributed solutions of (5.53) are compared in Fig. 5.6; note that in the centralized solution γ_0 must be known by the RNs, whereas in the distributed solution the RNs assume $\gamma_0 = 0$. In order to investigate the influence of the distance $d_{\mathrm{s,d}}$ on the performance, we show the outage probability for two scenarios: one where the SN and DN are sep-

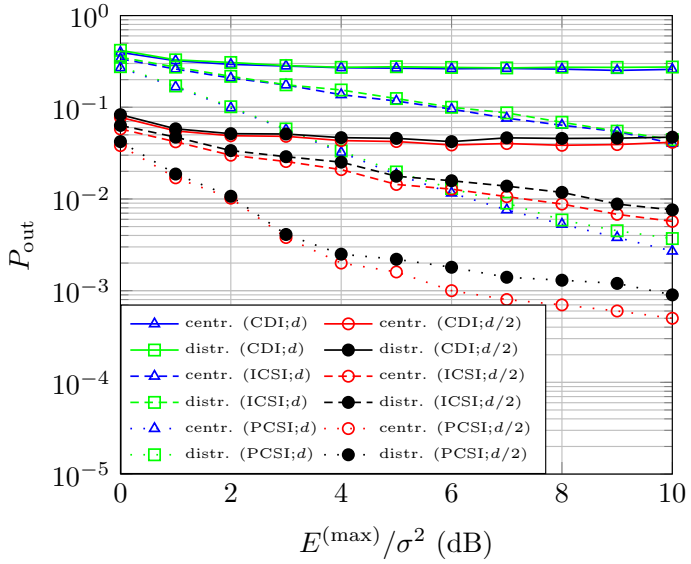


Figure 5.6: Comparison between the centralized and distributed solution ($M = 2$, $N_{\text{PU}} = 2$, $K = 1$, $N_{\text{a}} = 3$, ICSI for channels to DN).

arated by a distance d, and one where they are separated by a distance $d/2$, with $d = 1.375$. Fig. 5.6 shows that the distributed solution yields only a very small performance loss when $d_{\text{s,d}} = d$. In the case where $d_{\text{s,d}} = d/2$, the performance loss is slightly higher. However, as the distributed solution is easier to implement, the slight performance loss of the distributed algorithm is certainly acceptable.

5.4.5 Quality of the ICSI

In Fig. 5.7 the performances of the SU network are displayed for different values of the channel estimation error σ_{e}^2. We consider σ_{e}^2 equal to $\sigma_{\text{e,low}}^2 = \sigma^2/(10E^{(\max)})$, $\sigma_{\text{e,mod}}^2 = \sigma^2/(2E^{(\max)})$ and $\sigma_{\text{e,high}}^2 = 5\sigma^2/(2E^{(\max)})$ and assume that the SN and the RNs have ICSI about their channel coefficients to the DN. In the case of PCSI about the interference channels, the RNs can avoid interference at the PU-RXs; the performance differences for the considered values of σ_{e}^2 can be attributed to the different accuracies of RN beamsteering towards the DN. In the case of ICSI about the interference channels, the RNs cannot avoid interference at the PU-RXs, yielding a worse performance compared to PCSI; the performance differences for the considered values of σ_{e}^2 are larger than for PCSI, because the estimation error variance also applies to the interference channels, hence affecting the amount of interference. Finally, when only CDI about the interfering channels is available,

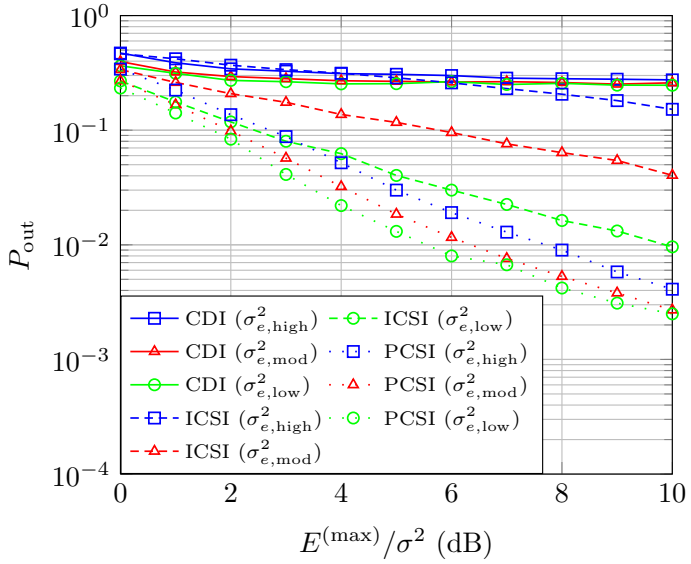


Figure 5.7: Comparison for different estimation error variances ($M = 2$, $N_{\text{PU}} = 2$, $K = 1$, $N_{\text{a}} = 3$, ICSI for channels to DN).

the performance is severely limited by the interference constraints, and the effect of the estimation error variance σ_{e}^2 is very small.

5.5 Chapter Summary

In this chapter, we presented several distributed resource allocation algorithms that succeed in minimizing the exact outage probability of a cooperative SU network with multi-antenna amplify-and-forward relay selection, while protecting the quality of service of the PU network. The derived resource allocation algorithms have the following advantages: (i) they minimize the exact outage probability instead of an approximation as in [27]; (ii) they are distributed, thus meaning that each node of the SU network can independently optimize its transmission parameters. The latter results in more practical algorithms when compared to the more general yet centralized algorithms proposed in [69]; (iii) in the numerical results, they are shown to have a clear performance benefit over other already existing algorithms, introduced in [70, 77].

Further, the numerical results also showed that the introduction of multi-antenna RNs can substantially improve the performance of a SU network which uses the underlay paradigm. The performance of the SU network was investigated for different levels of channel knowledge. The SU network considerably benefits from

having at least imperfect channel knowledge of the channel gains to the PU-RXs, and having RNs with a number of antennas larger than the number of PU antennas. When only CDI about the interference channels is available, we noticed that the interference constraints have a devastating effect on the SU performance.

The level of channel state information available about the channels to the DN is shown to have a smaller impact on the overall performance, as compared to the interference channel state information. But also here substantial gains compared to CDI are achieved when an estimate of the channel gains to the DN is available. The quality of this estimate has less effect than the quality of the estimate of the interference channels.

Finally, we have shown that having multiple RNs available can considerably improve the SU performance, compared to the case with a single relay, because the probability of having poor channel conditions is significantly reduced.

5.A Imperfect CSI

In this section we will describe the case where the CSI available at a SU-TX is imperfect due to feedback delay and estimation errors. The following formulas use a generic channel vector $\mathbf{f} \in \mathcal{C}^{F\times 1}$, where F denotes the number of components of $\mathbf{f}$. Its covariance matrix is denoted by $\mathbf{R_f}$. We also introduce the vector $\mathbf{ICSI}$ which denotes the imperfect CSI available at the SU-TX about the actual channel realization $\mathbf{f}$. We make the assumption that the channel vector $\mathbf{f}$ and the $\mathbf{ICSI}$ are jointly zero-mean circularly symmetric Gaussian. It then follows that $\mathbf{f}$ conditioned on the $\mathbf{ICSI}$ is Gaussian, with expectation $\hat{\mathbf{f}} = \mathbb{E}\left[\mathbf{f}|\mathbf{ICSI}\right]$ and covariance matrix $\mathbf{R_e} = \mathbb{E}\left[\mathbf{f}\mathbf{f}^H|\mathbf{ICSI}\right] - \hat{\mathbf{f}}\hat{\mathbf{f}}^H$. Introducing $\mathbf{e} \sim N_{\mathrm{c}}(\mathbf{0}, \mathbf{R_e})$, the instantaneous channel gain $\mathbf{f}$ can be decomposed as

$$\mathbf{f} = \hat{\mathbf{f}} + \mathbf{e}, \tag{5.63}$$

where $\hat{\mathbf{f}}$ is the MMSE estimate of $\mathbf{f}$ based on $\mathbf{ICSI}$, and the estimation error $\mathbf{e}$ is independent of $\hat{\mathbf{f}}$.

For the numerical results we will make some additional assumptions about $\mathbf{ICSI}$. We consider the case where the channel $\mathbf{f}(t)$ is slowly time-varying. According to Jakes' model [12], described in section 3.1, we take $\mathrm{E}[\mathbf{f}(t+u)\mathbf{f}^H(t)] = J_0(2\pi f_{\mathrm{d}} u)\mathbf{R_f}$ where $J(.)$ denotes the zeroth-order Bessel function of the first kind, and f_{d} denotes the Doppler frequency. At the SU-TX, the information about $\mathbf{f}(t)$ consists of P delayed channel estimates, i.e.,

$$\mathbf{ICSI} = [\tilde{\mathbf{f}}^T(t-\tau_1), \ldots, \tilde{\mathbf{f}}^T(t-\tau_P)]^T. \tag{5.64}$$

.

The estimate $\tilde{\mathbf{f}}(t-\tau_k)$, $k = 1, ..., P$, is given by

$$\tilde{\mathbf{f}}(t-\tau_k) = \mathbf{f}(t-\tau_k) + \tilde{\mathbf{e}}(t-\tau_k), \tag{5.65}$$

where τ_k represents the feedback delay relative to the instantaneous channel vector $\mathbf{f}(t)$. The noise vector $\tilde{\mathbf{e}}(t-\tau_k) \sim N_{\mathrm{c}}(\mathbf{0}, \sigma_{\mathrm{e}}^2\mathbf{I}_F)$, where σ_{e}^2 denotes the variance of the estimation error. In the numerical results we take $\tau_k = kDT$, which indicates that the channel estimates are updated every D symbol intervals T.

Introducing the matrix $\mathbf{J} \in \mathbb{C}^{P\times P}$ as $\mathbf{J}_{k,l} \triangleq J_0(2\pi f_{\mathrm{d}}(\tau_k - \tau_l)), k = 1, \ldots, P; l = 1, \ldots, P$, and using the results from section 3.3, it can now be shown that $\hat{\mathbf{f}} = \mathbf{X}\mathbf{Y}^{-1}\mathbf{ICSI}$, where

$$\mathbf{X} \triangleq [J_0(2\pi f_{\mathrm{d}}\tau_1), J_0(2\pi f_{\mathrm{d}}\tau_2) \ldots, J_0(2\pi f_{\mathrm{d}}\tau_{\mathrm{P}})] \otimes \mathbf{R_f}, \tag{5.66}$$

$$\mathbf{Y} \triangleq \mathbf{J} \otimes \mathbf{R_f} + \mathbf{I}_P \otimes \sigma_{\mathrm{e}}^2\mathbf{I}_F, \tag{5.67}$$

and the covariance matrix is given by $\mathbf{R_e} = \mathbf{R_f} - \mathbf{X}\mathbf{Y}^{-1}\mathbf{X}^H$.

5.B Proof of Theorem 5.1

The following proof is based on [68]. Assuming that the *m*th RN knows $\mathbf{h}_{\mathrm{s,r}}(m)$, the matrix $\mathbf{F}(m)$ can without loss of generality be decomposed as

$$\mathbf{F}(m) = [\mathbf{w}_1(m)\mathbf{W}_2(m)] \left[\frac{\mathbf{h}_{\mathrm{s,r}}(m)}{\|\mathbf{h}_{\mathrm{s,r}}(m)\|} \, \mathbf{H}_{\mathrm{s,r}}^{\perp}(m) \right]^H, \tag{5.68}$$

where $\mathbf{w}_1(m) \in \mathbb{C}^{N_\mathrm{a}\times 1}$, $\mathbf{W}_2(m) \in \mathbb{C}^{N_\mathrm{a}\times(N_\mathrm{a}-1)}$ and $\mathbf{H}_{\mathrm{s,r}}^{\perp}(m) \in \mathbb{C}^{N_\mathrm{a}\times(N_\mathrm{a}-1)}$. The matrix $\mathbf{H}_{\mathrm{s,r}}^{\perp}(m)$ makes $\left[\frac{\mathbf{h}_{\mathrm{s,r}}(m)}{\|\mathbf{h}_{\mathrm{s,r}}(m)\|} \, \mathbf{H}_{\mathrm{s,r}}^{\perp}(m) \right]$ a unitary matrix, which means that $(\mathbf{H}_{\mathrm{s,r}}^{\perp}(m))^H \mathbf{h}_{\mathrm{s,r}}(m) = \mathbf{0}$. Using (5.68), it can be shown that γ_m from (5.6) is given by

$$\gamma_m = \frac{E_0 \|\mathbf{h}_{\mathrm{s,r}}(m)\|^2 |\mathbf{h}_{\mathrm{r,d}}^T(m)\mathbf{w}_1(m)|^2}{\sigma_{\mathrm{s,r}}^2(m) \left(|\mathbf{h}_{\mathrm{r,d}}^T(m)\mathbf{w}_1(m)|^2 + \|\mathbf{h}_{\mathrm{r,d}}^T(m)\mathbf{W}_2(m)\|^2 \right) + \sigma_{\mathrm{r,d}}^2(m)}. \tag{5.69}$$

Combining equation (5.4) and (5.68), we can express the transmit power E_m as

$$E_m = E_0 \|\mathbf{h}_{\mathrm{s,r}}(m)\|^2 \|\mathbf{w}_1(m)\|^2 + \sigma_{\mathrm{s,r}}^2(m)(\|\mathbf{w}_1(m)\|^2 + \|\mathbf{W}_2(m)\|^2). \tag{5.70}$$

In the same way, by combining (5.13) and (5.68), the interference $\mathcal{I}_{\mathrm{r},p}(m)$ to be used in the interference constraints from section 5.2.3 can be written as

$$\begin{aligned} \mathcal{I}_{\mathrm{r},p}(m) \quad = \quad & E_0 \|\mathbf{h}_{\mathrm{s,r}}(m)\|^2 \|\mathbf{G}_{\mathrm{r},p}^T(m)\mathbf{w}_1(m)\|^2 + \sigma_{\mathrm{s,r}}^2(m) \\ & \cdot(\|\mathbf{G}_{\mathrm{r},p}^T(m)\mathbf{w}_1(m)\|^2 + \|\mathbf{G}_{\mathrm{r},p}^T(m)\mathbf{W}_2(m)\|^2), \end{aligned} \tag{5.71}$$

where $\mathbf{G}_{\mathrm{r},p}(m)$ denotes the actual interference channel (PCSI), or a point in the volume $\mathcal{U}_{\mathrm{r},p}(\hat{\bar{\mathbf{g}}}_{\mathrm{r},p}(m), \mathbf{B}_{\mathrm{r},p}(m))$ (ICSI) or $\mathcal{U}_{\mathrm{r},p}(\mathbf{0}, \mathbf{B}_{\mathrm{r},p}(m))$ (CDI). From (5.69) we see that the highest value of γ_m is reached when $\mathbf{W}_2(m) = \mathbf{0}$. From (5.7) and (5.9) it follows that a larger value for γ_m will certainly decrease the outage probability P_{out}, irrespective of which information (PCSI, ICSI or CDI) on $\mathbf{h}_{\mathrm{r,d}}(m)$ is available at the *m*th RN. Further, we note that $\mathbf{W}_2(m) = \mathbf{0}$ will also lead to a lower value of the transmit energy E_m (5.70) and the interference $\mathcal{I}_{\mathrm{r},p}(m)$ (5.71). Therefore we can conclude that the optimal value of $\mathbf{W}_2(m)$ will always be $\mathbf{0}$. Without loss of generality we can define

$$\mathbf{w}_1(m) = \frac{\mathbf{v}(m)^*}{\sqrt{E_0 \|\mathbf{h}_{\mathrm{s,r}}(m)\|^2 + \sigma_{\mathrm{s,r}}^2(m)}}, \tag{5.72}$$

which leads to (5.25).

5.C Rank-1 Solution of the Relaxed Optimization Problem

Here we show that under certain conditions the relaxed optimization problem (5.53) has a solution of rank 1.

5.C.1 A Useful Lemma

In [36], the following lemma is proven

Lemma 5.2. *The following optimization problem is convex in* $\mathbf{S}$ *and always has solutions with Rank(*$\mathbf{S}$*)*≤ 1*:*

$$\begin{aligned} \mathbf{S}^{(\mathrm{opt})} = \; & \arg\max_{\mathbf{S}\succeq\mathbf{0}} Tr\left(\mathbf{A}\mathbf{S}\right) \\ \text{s.t. } & Tr\left(\mathbf{B}_i\mathbf{S}\right) \leq b_i, \quad i = 1, \ldots, I, \end{aligned} \tag{5.73}$$

where $\mathbf{A}$ *is a Hermitian matrix, the matrices* $\mathbf{B}_i$ *are Hermitian with* $\mathbf{B}_i \succeq \mathbf{0}$ *and* $\sum_{i=1}^{I} \mathbf{B}_i \succ \mathbf{0}$*, and the scalars* b_i *satisfy* $b_i \geq 0$*,* $\forall i$*.*

This lemma can be directly applied to optimization problem (5.53) when PCSI is available about the interference channels, and with PCSI or CDI on the channels to the DN.

5.C.2 Rank-1 Solution to Optimization Problem (5.53) with ICSI or CDI on the Interference Channels

In the following, we will prove that optimization problem (5.53) with ICSI or CDI on the interference channels and PCSI on the channels to the DN also has a rank-1 solution. If we let χ denote the optimal value of the objective function $\mathrm{Tr}\left(\mathbf{h}_{\mathrm{r,d}}(m)\mathbf{h}_{\mathrm{r,d}}^{H}(m)\mathbf{S}(m)\right)$, we can rewrite the optimization problem as follows

$$\begin{aligned} \mathbf{S}(m) = \; & \arg \min_{\mathbf{S}(m)\succeq 0, \lambda_1(m)\geq 0, \ldots, \lambda_{N_{\mathrm{PU}}}(m)\geq 0} \mathrm{Tr}\left(\mathbf{S}(m)\right) \\ \text{s.t. } & \mathrm{Tr}\left(\mathbf{h}_{\mathrm{r,d}}(m)\mathbf{h}_{\mathrm{r,d}}^{H}(m)\mathbf{S}(m)\right) \geq \chi - \delta \\ & (5.39) \text{ or } (5.41), \; p = 1, \ldots, N_{\mathrm{PU}}, \end{aligned} \tag{5.74}$$

where δ denotes any small positive value. This optimization problem is convex and it can be easily verified that as δ approaches 0 the solution will approach the same optimal value as optimization problem (5.53). It can now be proven that the solution to optimization problem (5.74) has a rank-1 solution. The following proof is a modified version of the proof given in appendix D in [70]. To make our

proof valid for both the case where ICSI or CDI is available for the interference channels, we will use

$$\mathbf{K}_p \succeq 0, \tag{5.75}$$

to denote either (5.39) or (5.41) in the case of ICSI or CDI, respectively.

As Slater's condition holds, the duality gap of (5.74) will be zero. This means that we can solve this optimization problem by minimizing the following Lagrangian

$$\begin{aligned} L = &\operatorname{Tr}\left(\mathbf{S}(m)\right) + \xi\left(\chi - \delta - \operatorname{Tr}\left(\mathbf{h}_{\mathrm{r,d}}(m)\mathbf{h}_{\mathrm{r,d}}^H(m)\mathbf{S}(m)\right)\right) \\ &- \operatorname{Tr}\left(\mathbf{W}\mathbf{S}(m)\right) - \sum_{p=1}^{N_{\mathrm{PU}}}\left(\operatorname{Tr}\left(\mathbf{V}_p\mathbf{K}_p\right) + \mu_p\lambda_p(m)\right), \end{aligned} \tag{5.76}$$

where $\xi \geq 0$, $\mathbf{W} \succeq 0$, $\mathbf{V}_p \succeq 0$ and $\mu_p \geq 0$, $p = 1, \ldots, N_{PU}$, are the dual variables corresponding to the various constraints. We can rewrite $\mathbf{K}_p$ as

$$\mathbf{K}_p = \mathbf{A}_p - \mathbf{C}_p^H(\mathbf{I}_K \otimes \mathbf{S}(m))\mathbf{C}_p, \tag{5.77}$$

where

$$\mathbf{A}_p = \begin{pmatrix} \lambda_p(m)\mathbf{I}_{KN_a} & \mathbf{0} \\ \mathbf{0} & \Gamma - \lambda_p(m) \end{pmatrix}. \tag{5.78}$$

In the case where ICSI is available for the interference channels, the matrix $\mathbf{C}_p$ is given by

$$\mathbf{C}_p = \begin{pmatrix} \mathbf{B}_{\mathrm{r},p}(m) & \hat{\tilde{\mathbf{g}}}_{\mathrm{r},p}(m) \end{pmatrix}, \tag{5.79}$$

where $\mathbf{B}_{\mathrm{r},p}(m)$ is given by (5.19).

In the case where CDI is available, we get

$$\mathbf{C}_p = \begin{pmatrix} \mathbf{B}_{\mathrm{r},p}(m) & \mathbf{0} \end{pmatrix}, \tag{5.80}$$

where $\mathbf{B}_{\mathrm{r},p}(m)$ is given by (5.23).

The Lagrangian can now be expressed as

$$\begin{aligned} L &= \operatorname{Tr}\left(\mathbf{S}(m)\right) - \operatorname{Tr}\left(\mathbf{W}\mathbf{S}(m)\right) - \operatorname{Tr}\left(\xi\mathbf{h}_{\mathrm{r,d}}(m)\mathbf{h}_{\mathrm{r,d}}^H(m)\mathbf{S}(m)\right) \\ &\quad + \sum_{p=1}^{N_{\mathrm{PU}}}\operatorname{Tr}\left((\mathbf{I}_K \otimes \mathbf{S}(m))\mathbf{C}_p\mathbf{V}_p\mathbf{C}_p^H\right) - \sum_{p=1}^{N_{\mathrm{PU}}}\operatorname{Tr}\left(\mathbf{V}_p\mathbf{A}_p\right) + \phi(\xi, \chi, \delta, \boldsymbol{\mu}, \boldsymbol{\lambda}), \\ &= \operatorname{Tr}\left(\mathbf{S}(m)\right) - \operatorname{Tr}\left(\mathbf{W}\mathbf{S}(m)\right) - \operatorname{Tr}\left(\xi\mathbf{h}_{\mathrm{r,d}}(m)\mathbf{h}_{\mathrm{r,d}}^H(m)\mathbf{S}(m)\right) \\ &\quad + \sum_{k=1}^{K}\sum_{p=1}^{N_{\mathrm{PU}}}\operatorname{Tr}\left(\mathbf{S}(m)\mathbf{U}_p^{(k,k)}\right) - \sum_{p=1}^{N_{\mathrm{PU}}}\operatorname{Tr}\left(\mathbf{V}_p\mathbf{A}_p\right) + \phi(\xi, \chi, \delta, \boldsymbol{\mu}, \boldsymbol{\lambda}), \end{aligned} \tag{5.81}$$

where $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_{N_{\mathrm{PU}}}]^T$, $\boldsymbol{\lambda} = [\lambda_1(m), \ldots, \lambda_{N_{\mathrm{PU}}}(m)]^T$ and $\phi(\xi, \chi, \delta, \boldsymbol{\mu}, \boldsymbol{\lambda})$ denotes a function which is independent of $\mathbf{S}(m)$. In the last step, we introduced

the matrices $\mathbf{U}_p^{(k,k)} \succeq 0$, $k = 1, \ldots, K$, which are submatrices that are located in the block diagonal of $\mathbf{C}_p\mathbf{V}_p\mathbf{C}_p^H$ as follows

$$\mathbf{C}_p\mathbf{V}_p\mathbf{C}_p^H = \begin{pmatrix} \mathbf{U}_p^{(1,1)} & \cdots & \mathbf{U}_p^{(1,K)} \\ \vdots & \ddots & \vdots \\ \mathbf{U}_p^{(K,1)} & \cdots & \mathbf{U}_p^{(K,K)} \end{pmatrix}. \tag{5.82}$$

According to the KKT conditions [4] $\nabla_{\mathbf{S}(m)}L$ has to be equal to zero, which is equivalent to

$$\mathbf{A} - \mathbf{W} - \xi\mathbf{h}_{\mathrm{r,d}}(m)\mathbf{h}_{\mathrm{r,d}}^H(m) = \mathbf{0}, \tag{5.83}$$

where we have introduced the following matrix

$$\mathbf{A} = \mathbf{I}_{N_\mathrm{a}} + \sum_{k=1}^{K}\sum_{p=1}^{N_\mathrm{PU}} \mathbf{U}_p^{(k,k)}. \tag{5.84}$$

It is clear that $\mathbf{A} \succ 0$, which means that $\mathrm{rank}(\mathbf{A}) = N_\mathrm{a}$.

As a second step we use (5.83) and theorem 5.6.1 in [78] to bound the rank of matrix $\mathbf{W}$ as follows

$$\begin{aligned} \mathrm{rank}(\mathbf{A}) &\leq \mathrm{rank}(\mathbf{W}) + \mathrm{rank}(\xi\mathbf{h}_{\mathrm{r,d}}(m)\mathbf{h}_{\mathrm{r,d}}^H(m)) \\ &\leq \mathrm{rank}(\mathbf{W}) + 1, \end{aligned} \tag{5.85}$$

which yields

$$\mathrm{rank}(\mathbf{W}) \geq N_\mathrm{a} - 1. \tag{5.86}$$

In the final step we want to bound the rank of the matrix $\mathbf{S}(m)$. We now use the following KKT condition

$$\mathrm{Tr}\left(\mathbf{S}(m)\mathbf{W}\right) = \mathbf{0}. \tag{5.87}$$

Because both $\mathbf{S}(m)$ and $\mathbf{W}$ are positive semi-definite matrices there exist matrices $\mathbf{C}$ and $\mathbf{D}$ such that

$$\mathbf{W} = \mathbf{C}\mathbf{C}^H, \tag{5.88}$$

$$\mathbf{S}(m) = \mathbf{D}\mathbf{D}^H. \tag{5.89}$$

These can be used to rewrite equation (5.87) as

$$\begin{aligned} \mathrm{Tr}\left(\mathbf{S}(m)\mathbf{W}\right) &= \mathrm{Tr}\left(\mathbf{D}\mathbf{D}^H\mathbf{C}\mathbf{C}^H\right) \\ &= \mathrm{Tr}\left(\mathbf{C}^H\mathbf{D}\mathbf{D}^H\mathbf{C}\right) \\ &= \sum_i\sum_j |(\mathbf{D}^H\mathbf{C})_{i,j}|^2 = 0. \end{aligned} \tag{5.90}$$

This equality can only be satisfied if $\mathbf{D}^H\mathbf{C} = \mathbf{0}$, which leads to

$$\mathbf{S}(m)\mathbf{W} = \mathbf{D}\mathbf{D}^H\mathbf{C}\mathbf{C}^H = \mathbf{0}. \tag{5.91}$$

We now use equation (5.91) and apply Sylvester's rank inequality (theorem 5.6.5 in [78]) to get

$$\begin{aligned} \text{rank}(\mathbf{W}) + \text{rank}(\mathbf{S}(m)) &\leq \text{rank}(\mathbf{S}(m)\mathbf{W}) + N_\text{a} \\ &\leq N_\text{a}, \end{aligned} \tag{5.92}$$

which combined with (5.86) gives

$$\begin{aligned} \text{rank}(\mathbf{S}(m)) &\leq N_\text{a} - \text{rank}(\mathbf{W}) \\ &\leq 1. \end{aligned} \tag{5.93}$$

6

Distributed Dynamic Resource Allocation for Cooperative Cognitive Radio with Average Interference Constraints

In this chapter (which is based on contribution [79]), we investigate a special case of the multi-antenna scenario we discussed in chapter 5. The goal of this chapter is to investigate the performance of the SU network when average interference constraints are considered.

The interference constraint that we discussed in the previous chapter restrict the peak interference at the PU-RXs. These constraints can be seen as a short-term constraint that limits the peak interference level (PI) at each fading state. However, in this chapter we consider the possibility of a long-term constraint that regulates the average interference level (AI) at the PU-RXs over all the fading states [80]. From the cognitive radio (CR) network perspective, the AI constraint is expected to be more favorable than the PI one, since the former enables more flexibility in the dynamic allocation of the transmit energy over the fading states. Under both these types of interference constraints, we propose distributed resource allocation strategies aimed at minimizing the outage probability of a cooperative CR network. However, where the previous chapter used the exact outage probability, we

now use an approximation for the outage probability in order to find an analytical solution for the optimization problem with AI constraints. The outage probability performance of these AI-based algorithms is compared to the performance achieved under the PI constraint.

This chapter is divided as follows. Section 6.1 presents the cooperative system model and defines the link performance metric of interest. Section 6.2 investigates centralized and distributed energy allocation strategies that minimize the outage probability under the AI constraint, when the TXs either have perfect CSI or CDI about the interference channels, and briefly recalls the resource allocation under the PI constraint. In section 6.3, the simulation results obtained in terms of outage probability confirm the effectiveness of the proposed resource allocation algorithms, whereas a few conclusions are drawn in section 6.4.

6.1 System Model

Consider a cognitive relay network where a SU-SN and M SU-RNs transmit to a SU-DN in the same frequency band of a primary network, according to the *underlay* paradigm [1], i.e., by imposing severe constraints on the interference level. We make the same assumptions as in section 5.1, except that

1) All the devices, i.e., the PU-RXs, SN, RNs and DN, are equipped with a single antenna;

2) We do not consider best-relay selection, thus the transmission of a given data frame occurs in $(M+1)$ time slots;

3) We assume that the SN and RNs either have PCSI or CDI about their interference channels;

4) The SN and RNs only have CDI about their SU transmission channels;

5) We limit the number of PU-RXs N_{PU} to 1. For notational brevity, we will therefore drop the subindex p from the corresponding variables.

This leads to several simplifications compared to the previous chapter. The SNR γ_m of (5.28) can be simplified as follows

$$\gamma_m = \frac{E_0 E_m |h_{\mathrm{s,r}}(m)|^2 |h_{\mathrm{r,d}}(m)|^2}{E_0 |h_{\mathrm{s,r}}(m)|^2 \sigma^2_{\mathrm{r,d}}(m) + E_m |h_{\mathrm{r,d}}(m)|^2 \sigma^2_{\mathrm{s,r}}(m) + \sigma^2_{\mathrm{s,r}}(m)\sigma^2_{\mathrm{r,d}}(m)}. \quad (6.1)$$

Further, we will define the following variances $\kappa^2_{\mathrm{s,d}} \triangleq \mathbb{E}\left[|h_{\mathrm{s,d}}|^2\right]$, $\kappa^2_{\mathrm{s,r}}(m) \triangleq \mathbb{E}\left[|h_{\mathrm{s,r}}(m)|^2\right]$ and $\kappa^2_{\mathrm{r,d}}(m) \triangleq \mathbb{E}\left[|h_{\mathrm{r,d}}(m)|^2\right]$.

For a given information transmission rate R, the link outage probability can be formulated as

$$P_{\text{out}} = \Pr\{R_{\text{AF}} \leq R\}, \tag{6.2}$$

where the SN-DN channel rate R_{AF} is given by (3.66).

It has to be remarked, however, that the evaluation of (6.2) for a given set of γ_m, $m = 0, \ldots, M$, reveals a quite demanding task. For this reason we will use an approximation which has been proposed in [27]. The authors of [27] show that the outage probability P_{out} for a cooperative amplify-and-forward scheme, can be asymptotically bounded at high-SNR by

$$P_{\text{out}} \leq \frac{1}{M+1}\left[2^{(M+1)R}-1\right]^{M+1} \alpha_0 \cdot \prod_{m=1}^{M}(\alpha_m+\beta_m), \tag{6.3}$$

$$P_{\text{out}} \geq \frac{1}{(M+1)M^M}\left[2^{(M+1)R}-1\right]^{M+1} \alpha_0 \cdot \prod_{m=1}^{M}(\alpha_m+\beta_m), \tag{6.4}$$

where α_0, α_m and β_m are the parameters of the exponential distribution of $E_0|h_{\text{s,d}}|^2/\sigma^2_{\text{s,d}}$, $E_0|h_{\text{s,r}}(m)|^2/\sigma^2_{\text{s,r}}(m)$ and $E_m|h_{\text{r,d}}(m)|^2/\sigma^2_{\text{r,d}}(m)$, respectively, given by

$$\alpha_0 \triangleq \frac{\sigma^2_{\text{s,d}}}{E_0\,\mathbb{E}\left[|h_{\text{s,d}}|^2\right]} = \frac{\sigma^2_{\text{s,d}}}{E_0\,\kappa^2_{\text{s,d}}}, \tag{6.5}$$

$$\alpha_m \triangleq \frac{\sigma^2_{\text{s,r}}(m)}{E_0\,\mathbb{E}\left[|h_{\text{s,r}}(m)|^2\right]} = \frac{\sigma^2_{\text{s,r}}(m)}{E_0\,\kappa^2_{\text{s,r}}(m)}, \tag{6.6}$$

$$\beta_m \triangleq \frac{\sigma^2_{\text{r,d}}(m)}{E_m\,\mathbb{E}\left[|h_{\text{r,d}}(m)|^2\right]} = \frac{\sigma^2_{\text{r,d}}(m)}{E_m\,\kappa^2_{\text{r,d}}(m)}. \tag{6.7}$$

6.2 Resource Allocation

In order to protect the PU transmission, the interference level at the PU-RX has to be kept under a predefined threshold. The coefficients of the channel gains from the SN to the PU-RX and from the *m*th RN to the PU-RX will be denoted by $g_{\text{s}} \in \mathbb{C}$ and $g_{\text{r}}(m) \in \mathbb{C}$, respectively, with respective variances $\rho^2_{\text{s}} \triangleq \mathbb{E}\left[|g_{\text{s}}|^2\right]$ and $\rho^2_{\text{r}}(m) \triangleq \mathbb{E}\left[|g_{\text{r}}(m)|^2\right]$. The following approaches are pursued:

- The AI constraint regulates the average level of interference as

$$\mathbb{E}\left[E_0\,|g_{\text{s}}|^2\right] \leq \bar{\Gamma}, \tag{6.8}$$

$$\mathbb{E}\left[E_m\,|g_{\text{r}}(m)|^2\right] \leq \bar{\Gamma}, \quad m = 1, \ldots, M, \tag{6.9}$$

  where $\bar{\Gamma}$ denotes a given threshold depending on the characteristics of the PU-RX.

- The PI constraint, which we used in (5.14) and (5.15), forces the peak level of interference below the threshold Γ as

$$E_0 \left|g_{\rm s}\right|^2 \leq \Gamma, \tag{6.10}$$

$$E_m \left|g_{\rm r}(m)\right|^2 \leq \Gamma, \quad m = 1, \ldots, M. \tag{6.11}$$

In addition to the AI or PI constraints, we impose the following transmit constraint

$$0 \leq E_m \leq E_m^{(\max)}, \quad m = 0, \ldots, M, \tag{6.12}$$

where $E_m^{(\max)}$ is the maximum transmit energy admissible for the *m*th SU node.

Hence, choosing the AI constraint and assuming that both SN and RNs have an ideal knowledge of either PCSI or CDI about their channel coefficients to the PU-RX, the optimization problem to search for the optimal $\mathbf{E} \triangleq [E_0, E_1, \ldots, E_M]^T$ can be formulated as

$$\begin{aligned} \mathbf{E}^{(\rm opt)} &= \arg \min_{\mathbf{E}} \left\{\Psi\left(\mathbf{E}\right)\right\} \\ \text{s.t.} \quad & (6.8) - (6.9) \\ & 0 \leq E_m \leq E_m^{(\max)}, \quad m = 0, \ldots, M, \end{aligned} \tag{6.13}$$

where the objective function is defined as

$$\Psi\left(\mathbf{E}\right) \triangleq \mathbb{E}_{\mathbf{g}} \left[\frac{\sigma_{\rm s,d}^2}{E_0\, \kappa_{\rm s,d}^2} \cdot \prod_{m=1}^{M} \left[\frac{\sigma_{\rm s,r}^2(m)}{E_0\, \kappa_{\rm s,r}^2(m)} + \frac{\sigma_{\rm r,d}^2(m)}{E_m\, \kappa_{\rm r,d}^2(m)} \right] \right], \tag{6.14}$$

and the expectation is made over channel coefficients $\mathbf{g} \triangleq [g_{\rm s}, g_{\rm r}(1), \cdots, g_{\rm r}(M)]^T$ between the $M+1$ SU nodes and the PU-RX. As the AI-based optimization problem (6.13) takes care of the long-term constraint of the interference level, it will also consider the long-term objective function (6.14) which is the result of the average over the fading coefficients between the SU-TXs and the PU-RX.

As an alternative, adopting the PI constraint with PCSI, the optimization problem can be put into the form

$$\begin{aligned} \mathbf{E}^{(\rm opt)} &= \arg \min_{\mathbf{E}} \left\{\Phi\left(\mathbf{E}\right)\right\} \\ \text{s.t.} \quad & (6.10) - (6.11) \\ & 0 \leq E_m \leq E_m^{(\max)}, \quad m = 0, \ldots, M, \end{aligned} \tag{6.15}$$

where the objective function $\Psi\left(\mathbf{E}\right)$ has been substituted by

$$\Phi\left(\mathbf{E}\right) \triangleq \frac{\sigma_{\rm s,d}^2}{E_0\, \kappa_{\rm s,d}^2} \cdot \prod_{m=1}^{M} \left[\frac{\sigma_{\rm s,r}^2(m)}{E_0\, \kappa_{\rm s,r}^2(m)} + \frac{\sigma_{\rm r,d}^2(m)}{E_m\, \kappa_{\rm r,d}^2(m)} \right], \tag{6.16}$$

as the objective function $\Phi(\mathbf{E})$ is completely equivalent with $\Psi(\mathbf{E})$ for the PI constraints (6.10)-(6.11).

We will now derive analytic expressions for the transmit energy vector $\mathbf{E}$ in the following cases:

- The SU network has to satisfy AI constraints and PCSI is available about the interference channels;
- The SU network has to satisfy AI constraints and CDI is available about the interference channels;
- The SU network has to satisfy PI constraints and PCSI is available about the interference channels.

6.2.1 AI Constraint under PCSI

6.2.1.1 Centralized Solution

As a reference scenario, we will derive the optimum solution of optimization problem (6.13) for $M = 1$. As the optimization problem in (6.13) is convex and Slater's condition holds, the duality gap of this optimization problem is equal to zero. To solve this optimization problem, we first write down the Lagrangian according to (2.10)

$$\begin{aligned} L(\mathbf{E}, \mu_0, \mu_1) &= \mathbb{E}_{\mathbf{g}}\left[\frac{\sigma^2_{\mathrm{s,d}}}{E_0\,\kappa^2_{\mathrm{s,d}}} \cdot \left[\frac{\sigma^2_{\mathrm{s,r}}(1)}{E_0\,\kappa^2_{\mathrm{s,r}}(1)} + \frac{\sigma^2_{\mathrm{r,d}}(1)}{E_1\,\kappa^2_{\mathrm{r,d}}(1)}\right] + \mu_0(E_0\,|g_{\mathrm{s}}|^2 - \bar{\Gamma}) \right. \\ &\quad \left. + \mu_1(E_1\,|g_{\mathrm{r}}(1)|^2 - \bar{\Gamma})\right], \end{aligned} \tag{6.17}$$

where μ_0 and μ_1 are two Lagrange multipliers which have to satisfy the following equations

$$\mu_0\left(\mathbb{E}_{\mathbf{g}}\left[E_0(\mu_0,\mu_1)\,|g_{\mathrm{s}}|^2\right] - \bar{\Gamma}\right) = 0, \tag{6.18}$$

$$\mu_1\left(\mathbb{E}_{\mathbf{g}}\left[E_1(\mu_0,\mu_1)\,|g_{\mathrm{r}}(1)|^2\right] - \bar{\Gamma}\right) = 0. \tag{6.19}$$

The value of these Lagrange multipliers can be found numerically or by using the stochastic subgradient method as described in section 2.5.

As we have neglected the transmit constraint (6.12), the Lagrangian in (6.17) is only optimal in the case that $0 \leq E_m \leq E_m^{(\max)}, m = 0, 1$. When we have found the solution, we will check if this assumption holds. The solution is found by putting $\nabla_{\mathbf{E}} L(\mathbf{E}, \mu_0, \mu_1)$ equal to $\mathbf{0}$. By solving this equation, we find the fol-

lowing expression for the transmit energy at the SN

$$E_0^{(\mathrm{opt})} = \left(\frac{\sqrt{\mu_1}\frac{\kappa_{\mathrm{s,r}}(1)}{\sigma^2_{\mathrm{s,r}}(1)}|g_{\mathrm{r}}(1)| + \sqrt{8\mu_0\frac{\kappa^2_{\mathrm{r,d}}(1)}{\sigma^2_{\mathrm{r,d}}(1)}|g_{\mathrm{s}}|^2 + \mu_1\frac{\kappa^2_{\mathrm{s,r}}(1)}{\sigma^2_{\mathrm{s,r}}(1)}|g_{\mathrm{r}}(1)|^2}}{2\mu_0\frac{\kappa_{\mathrm{s,d}}}{\sigma_{\mathrm{s,d}}}\frac{\kappa_{\mathrm{s,r}}(1)}{\sigma_{\mathrm{s,r}}(1)}\frac{\kappa_{\mathrm{r,d}}(1)}{\sigma_{\mathrm{r,d}}(1)}|g_{\mathrm{s}}|^2} \right)^{2/3}. \tag{6.20}$$

The transmit energy at the RN can then be easily found as

$$E_1^{(\mathrm{opt})} = \frac{1}{\sqrt{\mu_1 E_0^{(\mathrm{opt})}}\frac{\kappa_{\mathrm{s,d}}}{\sigma_{\mathrm{s,d}}}\frac{\kappa_{\mathrm{r,d}}(1)}{\sigma_{\mathrm{r,d}}(1)}|g_{\mathrm{r}}(1)|}. \tag{6.21}$$

If the value for $E_0^{(\mathrm{opt})}$ and $E_1^{(\mathrm{opt})}$ given by (6.20) and (6.21) satisfies the transmit energy constraint (6.12), we have found the optimal solution. However if either (6.20) or (6.21) violates the transmit energy constraint, we have to check which of the following cases apply:

1. In the case where both $E_0^{(\mathrm{opt})}$ and $E_1^{(\mathrm{opt})}$ violate the maximal transmit energy constraint it can easily be shown that the optimal values are

$$E_0^{(\mathrm{opt})} = E_0^{(\mathrm{max})}, \tag{6.22}$$
$$E_1^{(\mathrm{opt})} = E_1^{(\mathrm{max})}. \tag{6.23}$$

2. In the case where only $E_0^{(\mathrm{opt})} > E_0^{(\mathrm{max})}$, it can be shown that the optimum values are given by

$$E_0^{(\mathrm{opt})} = E_0^{(\mathrm{max})}, \tag{6.24}$$
$$E_1^{(\mathrm{opt})} = \min\left(\frac{1}{\sqrt{\mu_1 E_0^{(\mathrm{max})}}\frac{\kappa_{\mathrm{s,d}}}{\sigma_{\mathrm{s,d}}}\frac{\kappa_{\mathrm{r,d}}(1)}{\sigma_{\mathrm{r,d}}(1)}|g_{\mathrm{r}}(1)|}, E_1^{(\mathrm{max})} \right). \tag{6.25}$$

3. If $E_1^{(\mathrm{opt})} > E_1^{(\mathrm{max})}$ it can be proven that

$$E_0^{(\mathrm{opt})} = \min\left(\frac{1}{3^{2/3}} \frac{3^{1/3}\frac{\kappa^2_{\mathrm{s,r}}(1)}{\sigma^2_{\mathrm{s,r}}(1)} + \frac{\kappa^{2/3}_{\mathrm{s,r}}(1)}{\sigma^{2/3}_{\mathrm{s,r}}(1)}\alpha_1^{2/3}}{\sqrt{\mu_0 E_1^{(\mathrm{max})}}\frac{\kappa_{\mathrm{s,d}}}{\sigma_{\mathrm{s,d}}}\frac{\kappa^{4/3}_{\mathrm{s,r}}(1)}{\sigma^{4/3}_{\mathrm{s,r}}(1)}\frac{\kappa_{\mathrm{r,d}}(1)}{\sigma_{\mathrm{r,d}}(1)}|g_{\mathrm{s}}|\alpha_1^{1/3}}, E_0^{(\mathrm{max})} \right), \tag{6.26}$$
$$E_1^{(\mathrm{opt})} = E_1^{(\mathrm{max})}, \tag{6.27}$$

where

$$\alpha_1 = 9\sqrt{\mu_0}(E_1^{(\mathrm{max})})^{3/2}\frac{\kappa_{\mathrm{s,d}}}{\sigma_{\mathrm{s,d}}}\frac{\kappa^3_{\mathrm{r,d}}(1)}{\sigma^3_{\mathrm{r,d}}(1)}|g_{\mathrm{s}}| + \sqrt{3\alpha_2}, \tag{6.28}$$

$$\alpha_2 = 27\mu_0 (E_1^{(\max)})^3 \frac{\kappa_{\mathrm{s,d}}^2}{\sigma_{\mathrm{s,d}}^2} \frac{\kappa_{\mathrm{r,d}}^6(1)}{\sigma_{\mathrm{r,d}}^6(1)} |g_{\mathrm{s}}|^2 - \frac{\kappa_{\mathrm{s,r}}^4(1)}{\sigma_{\mathrm{s,r}}^4(1)}. \tag{6.29}$$

If $\alpha_2 < 0$, equation (6.26) can be rewritten as

$$E_0^{(\mathrm{opt})} = \min \left(\frac{2}{3} \frac{\sqrt{3} \cos \left(\frac{1}{3} \arctan \left(\frac{1}{9} \frac{\sqrt{-3\alpha_2}}{\sqrt{\mu_0} (E_1^{(\max)})^{3/2} \frac{\kappa_{\mathrm{s,d}}}{\sigma_{\mathrm{s,d}}} \frac{\kappa_{\mathrm{r,d}}^3(1)}{\sigma_{\mathrm{r,d}}^3(1)} |g_{\mathrm{s}}|} \right) \right)}{\sqrt{\mu_0 E_1^{(\max)}} \frac{\kappa_{\mathrm{s,d}}}{\sigma_{\mathrm{s,d}}} \frac{\kappa_{\mathrm{r,d}}(1)}{\sigma_{\mathrm{r,d}}(1)} |g_{\mathrm{s}}|}, E_0^{(\max)} \right). \tag{6.30}$$

6.2.1.2 Distributed Solution

We now search for a distributed solution of optimization problem (6.13). For a given value of E_0, the optimization problem (6.13) turns into M separate optimization problems, one for each RN. The optimization problem (6.13) for the *m*th RN can be put into the form

$$\begin{aligned} E_m^{(\mathrm{opt})} = \arg \min_{E_m} & \left\{ \mathbb{E}_{g_{\mathrm{r}}(m)} \left[\frac{1}{E_m} \right] \right\} \\ \text{s.t.} \quad & \mathbb{E} \left[E_m |g_{\mathrm{r}}(m)|^2 \right] \leq \bar{\Gamma} \\ & 0 \leq E_m \leq E_m^{(\max)}, \end{aligned} \tag{6.31}$$

whose solution can be obtained as stated by the following theorem.

Theorem 6.1. *Under the assumptions of: (i) AI constraint, and (ii) assuming a given energy E_0, the optimal energy allocation for the* m*th RN such that the outage probability is minimized can be written as*

$$E_m^{(\mathrm{opt})} = \min \left\{ \sqrt{\frac{1}{\lambda_m |g_{\mathrm{r}}(m)|^2}}, E_m^{(\max)} \right\}, m = 1, \ldots, M, \tag{6.32}$$

where λ_m is the Lagrange multiplier corresponding to the interference constraint (6.9).

Proof. See appendix 6.A. □

A few remarks are now in order.

- The Lagrange multiplier λ_m has to be evaluated only once, and then, it is applied for all possible channel realizations of $g_{\mathrm{r}}(m)$.

- In order to find a distributed solution, we propose to use the solution of the RNs also for the optimization of the transmit energy E_0 of the SN. Hence, we propose to optimize E_0 according to

$$E_0^{(\mathrm{opt})} = \min\left\{\sqrt{\frac{1}{\lambda_0|g_\mathrm{s}|^2}}, E_0^{(\mathrm{max})}\right\}. \tag{6.33}$$

  This approach is not optimal. However, we will see in section 6.3 that the performance loss compared to the optimal solution is minimal.

6.2.2 AI Constraint under CDI

In the case of the AI-based constraints under CDI, the SU nodes only know the statistics of their channel gain towards the PU-RX. This means that the energy vector $\mathbf{E}$ depends only on ρ_s^2 and $\rho_\mathrm{r}^2(m)$, $m = 1, \ldots, M$. Therefore, the optimization problem (6.13) can be simplified as

$$\begin{aligned} \mathbf{E}^{(\mathrm{opt})} = \arg\ & \min_{\mathbf{E}} \{\Psi(\mathbf{E})\} \\ \text{s.t.}\quad & E_0\rho_\mathrm{s}^2 \leq \bar{\Gamma} \\ & E_m\rho_\mathrm{r}^2(m) \leq \bar{\Gamma}, \quad m = 1, \ldots, M \\ & 0 \leq E_m \leq E_m^{(\mathrm{max})}, \quad m = 0, \ldots, M, \end{aligned} \tag{6.34}$$

It now follows that the optimal value $E_0^{(\mathrm{opt})}$ is obtained as

$$E_0^{(\mathrm{opt})} = \min\left\{\frac{\bar{\Gamma}}{\rho_\mathrm{s}^2}, E_0^{(\mathrm{max})}\right\}, \tag{6.35}$$

and $E_m^{(\mathrm{opt})}$ is given by

$$E_m^{(\mathrm{opt})} = \min\left\{\frac{\bar{\Gamma}}{\rho_\mathrm{r}^2(m)}, E_m^{(\mathrm{max})}\right\}, \quad m = 1, \ldots, M. \tag{6.36}$$

It is worth remarking that: *i*) although the optimization problem (6.34) has a centralized formulation, its solution is distributed; *ii*) unlike the AI-based optimization problem under PCSI, the distributed solution given by (6.35) and (6.36) is globally optimal.

6.2.3 PI Constraint under PCSI

According to the PI constraint, the instantaneous interference level at the PU-RX due to the nodes of the SU network may not exceed a predefined threshold [80]. As the objective function in (6.15) is monotonically decreasing as a function of

E_m, $m = 0, \ldots, M$, the optimal solution $E_m^{(\mathrm{opt})}$ is thus obtained as the maximum energy satisfying the constraints, i.e.,

$$\begin{aligned} E_0^{(\mathrm{opt})} &= \min\left\{\frac{\Gamma}{|g_{\mathrm{s}}|^2}, E_0^{(\max)}\right\}, \\ E_m^{(\mathrm{opt})} &= \min\left\{\frac{\Gamma}{|g_{\mathrm{r}}(m)|^2}, E_m^{(\max)}\right\}, \quad m = 1, \ldots, M. \end{aligned} \tag{6.37}$$

We note that this solution corresponds to the solution we found in the previous chapter. Equation (6.37) corresponds to (5.44) for the SN and to the solution proposed in section 5.3.3.3 for the RN.

6.3 Numerical Results

In this section, the performance of the AI-based resource allocation is compared to the PI-based one in terms of the outage probability that is achieved by the SU network. The channel gain coefficients between the SU nodes and between the SU-TXs to the PU-RX are all assumed to be zero-mean circularly symmetric complex Gaussian random variables with unit variance. The noise variance is set to be equal for every device in the network, namely $\sigma^2 = \sigma^2_{\mathrm{s,d}} = \sigma^2_{\mathrm{s,r}}(m) = \sigma^2_{\mathrm{r,d}}(m)$, $m = 1, \ldots, M$, and we take $\bar{\Gamma} = \Gamma$. We start by considering only a single RN ($M = 1$). The outage probability performance is evaluated by averaging over 10^7 independent channel realizations assuming that the information rate is set to $R = 1$ bit/s/Hz.

Figure 6.1 shows the simulated outage probability given by (6.2) (solid lines) and the outage probability evaluated through the upper bound given by (6.3) (dotted lines), which coincides with the lower bound (6.4) when $M = 1$. The maximum transmit energy admissible for both the SN and RN is chosen as $E^{(\max)} = E_0^{(\max)} = E_1^{(\max)}$. The outage probability is expressed as a function of $E^{(\max)}/\sigma^2$ (dB) and we fix $E^{(\max)}/\Gamma = 10$ dB. It is apparent that the proposed distributed resource allocation algorithms, based on the AI approach with either perfect CSI (triangle marks) or CDI (circle marks), outperform the conventional PI-based energy allocation (square marks) discussed in the previous chapter. As a performance benchmark, we have also shown the outage probability for the centralized solution from section 6.2.1.1. As we can see, there is almost no performance loss by using the distributed solution presented in section 6.2.1.2. Finally, we also show the performance when no PU-RX is active, this illustrates the performance when there exist no constraints on the interference level. The above results show that limiting the peak interference turns out to be more restrictive than setting a threshold on the average value, with the result that the overall CR network makes a more efficient usage of the available resources, i.e., the transmit energy transmitted by the nodes.

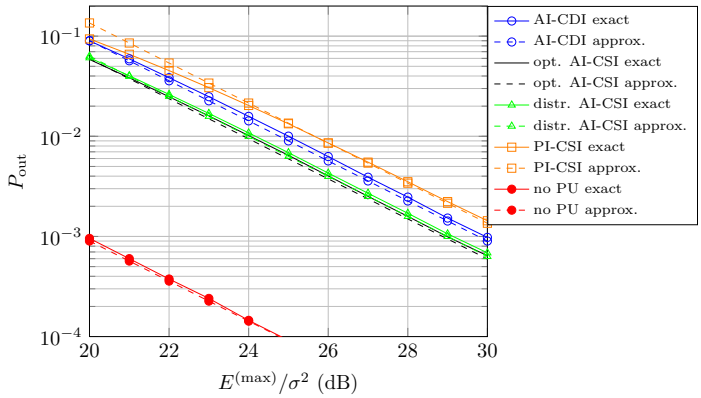


Figure 6.1: Outage probability versus $E^{(\max)}/\sigma^2$ ($E^{(\max)}/\Gamma = 10$ dB).

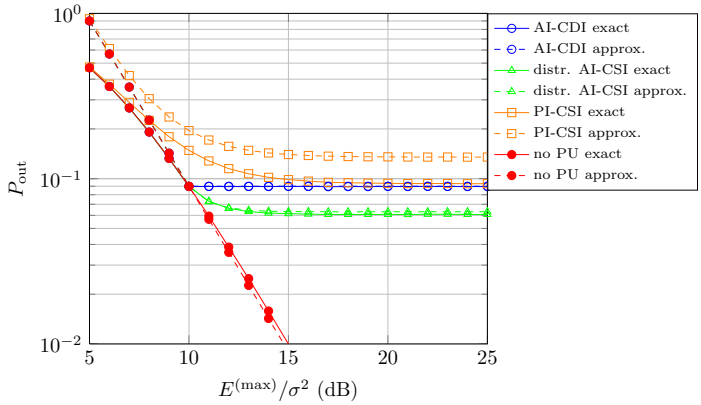


Figure 6.2: Outage probability versus $E^{(\max)}/\sigma^2$ ($\Gamma/\sigma^2 = 10$ dB).

In Fig. 6.2, the outage probability is again expressed as a function of $E^{(\max)}/\sigma^2$ (dB), but the interference threshold is now directly proportional to the noise variance according to $\Gamma/\sigma^2 = 10$ dB. It can be observed that the outage probability for the AI approach with ideal CDI (circle marks) does not change if the CR transmit energy constraint is increased past 10 dB, whereas for smaller values the outage probability collides with the performance of the no-PU case (solid line). The reason that these lines collide for $E^{(\max)}/\sigma^2 < 10$ dB can be seen by substituting $\rho_{\text{s}}^2 = \rho_{\text{r}}^2(m) = 1$ in (6.35) and (6.36), which gives, indeed, $E_m^{(\text{opt})}/\sigma^2 = \min\left\{10,\, E^{(\max)}/\sigma^2\right\} = E^{(\max)}/\sigma^2$, $m = 0, \ldots, M$. Further, the energy allocation based on AI with CSI outperforms the PI-based approach with CSI, although both of them tend to flatten out whenever the AI/PI constraints become more dominant than the maximum transmit energy constraint.

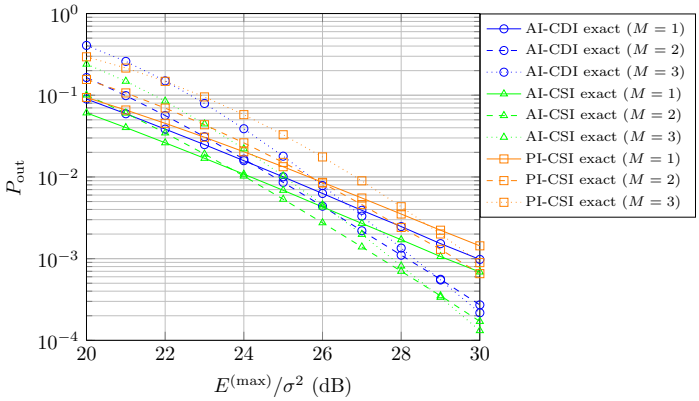


Figure 6.3: Outage probability versus $E^{(\max)}/\sigma^2$ ($M = 1,\ 2$ and 3, $E^{(\max)}/\Gamma = 10$ dB).

In Fig. 6.3 we have shown the exact outage probabilities for a different number of RNs. We consider $M = 1,\ 2$ and 3 and fix $E^{(\max)}/\Gamma = 10$ dB. It is clear that the performance becomes worse for low values of $E^{(\max)}/\sigma^2$ when we increase the number of RNs. This loss in performance is caused by the factor $1/(M+1)$ which decreases the possible rate in (3.66). However, this loss can be avoided by using the selection algorithm used in chapter 5. For high values of $E^{(\max)}/\sigma^2$, we get an improvement in performance as the use of multiple RNs increases the diversity order.

6.4 Chapter Summary

In this chapter, we have focused on the difference in performance between resource allocation strategies with average or peak interference constraints for cooperative CR wireless systems.

First, we introduced bounds for the link outage probability and discussed both the AI and PI constraints which are aimed at protecting the PU-RX. Then, a closed-form distributed energy allocation has been found under the AI approach when either perfect knowledge of the channel (CSI) or its distribution (CDI) are available. Numerical results demonstrate that the AI-based energy allocation algorithm both with perfect CSI and CDI shows an edge in terms of outage probability performance on the traditional PI-based method at the price of a similar computational complexity.

Further, in the case with perfect CSI, the distributed approach was shown to have nearly no performance loss compared to the optimal centralized approach.

Finally, although multiple RNs ($M > 1$) increase the diversity order of the SU

network, there is also a significant performance loss for low values of $E^{(\max)}/\sigma^2$. This performance loss is caused by the fact that each RN needs its own orthogonal time slot. It is possible to avoid this performance loss by using the selection amplify-and-forward scheme introduced in section 3.6.2. However, as we will see in the following chapter, the combination of the AI constraints together with the selection amplify-and-forward scheme is not trivial.

6.A Proof of Theorem 6.1

From (6.31) it can be seen that the objective function and the constraint functions are all convex. As Slater's condition holds, the duality gap will be zero. According to (2.10) the Lagrangian is equal to

$$L(E_m, \lambda_m) = \mathbb{E}_{g_{\mathrm{r}}(m)} \left[\frac{1}{E_m} + \lambda_m (E_m \left| g_{\mathrm{r}}(m) \right|^2 - \bar{\Gamma}) \right]. \tag{6.38}$$

The optimal value of E_m can now be found by calculating the following derivative

$$\nabla_{E_m} L(E_m, \lambda_m) \quad = \quad \mathbb{E}_{g_{\mathrm{r}}(m)} \left[\frac{-1}{E_m^2} + \lambda_m |g_{\mathrm{r}}(m)|^2 \right], \tag{6.39}$$

and setting it to 0. It follows that

$$E_m^{(\mathrm{opt})} = \min \left\{ \sqrt{\frac{1}{\lambda_m |g_{\mathrm{r}}(m)|^2}}, E_m^{(\max)} \right\}, \tag{6.40}$$

where the value of λ_m has to satisfy the following equality

$$\mathbb{E} \left[E_m \left| g_{\mathrm{r}}(m) \right|^2 \right] - \bar{\Gamma} = 0, \tag{6.41}$$

and can be found through the stochastic subgradient method as described in section 2.5 or by solving the following analytical equation

$$\begin{aligned} &\int_0^{\frac{1}{\lambda_m (E_m^{(\max)})^2}} E_m^{(\max)} \frac{x}{\rho_{\mathrm{r}}^2(m)} \exp(-\frac{x}{\rho_{\mathrm{r}}^2(m)}) \, \mathrm{d}x \\ &+ \int_{\frac{1}{\lambda_m (E_m^{(\max)})^2}}^{\infty} \sqrt{\frac{x}{\lambda_m}} \frac{1}{\rho_{\mathrm{r}}^2(m)} \exp(-\frac{x}{\rho_{\mathrm{r}}^2(m)}) \, \mathrm{d}x = \bar{\Gamma}. \end{aligned} \tag{6.42}$$

In the case where the interference constraint (6.9) is inactive, the value of λ_m will be 0.

7

Resource Allocation for Cooperative Cognitive Radio with Best-Relay Selection under an Average Interference Constraint

As we have shown in the previous chapter, the use of multiple RNs increases the diversity order of the SU network [14, 20]. The drawback of this method is the loss in spectral efficiency, since the RNs transmit on orthogonal channels. However this can be partially mitigated by using the selection amplify-and-forward protocol, introduced in section 3.6.2, where only the best RN is chosen to assist in the transmission. This protocol requires only two orthogonal channels: one for the direct link between the SN and the DN, and one for the link between the best RN and the DN. In [27], the authors have shown that this protocol maintains the full diversity order $M + 1$, with M denoting the number of relays. We already investigated the performance of the relay selection algorithm under a PI constraint in chapter 5. However, the performance of the selection amplify-and-forward protocol in combination with the AI constraint still needs to be analyzed. Therefore, we derive in this chapter both an optimum and a low-complexity suboptimum dynamic energy allocation strategy, which minimize the link outage probability between the SN and the DN of the SU network.

In this chapter, we assume that the interference channels from the SU nodes to the PU are affected by fast fading. Therefore, we make the realistic assumption that each SU node has access only to the root mean square (RMS) (rather than the instantaneous value) of the gain of its own interference channel to the PU. We show that both dynamic allocation strategies considerably outperform the optimum static allocation strategy, and that the low-complexity dynamic allocation performs nearly as well as the optimum dynamic allocation. Finally, we confirm that our dynamic algorithms show a significant performance improvement as compared to the optimal dynamic algorithm under a PI constraint.
This chapter (which is based on contribution [81]) is organized as follows. In section 7.1 we describe the system model which consists of the selection amplify-and-forward protocol, the outage probability and the AI constraints. In section 7.2 we derive a static and two dynamic energy allocation strategies. We then discuss the performance of these allocation strategies in section 7.3. Finally, the conclusions are drawn in section 7.4.

7.1 System Model

7.1.1 Selection Amplify-and-Forward

We consider a cooperative network as shown in Fig. 7.1. The SU network is identical to the network considered in chapter 6. However, the transmission of a message or frame consists of only two time slots. During the first slot, the SN transmits, and its signal is received by the DN and the RNs. During the second slot, the best relay forwards to the destination the signal received from the SN during the first slot.

For notational convenience, we introduce the following vector $\boldsymbol{\kappa} \triangleq [\kappa_{\mathrm{s,d}}, \kappa_{\mathrm{s,r}}(1), \ldots, \kappa_{\mathrm{s,r}}(M), \kappa_{\mathrm{r,d}}(1), \ldots, \kappa_{\mathrm{r,d}}(M)]$ which contains the RMS values of the SU channel gains. The mean-square values of the respective channel gain magnitudes are determined by $\kappa^2_{\mathrm{s,d}} = 1/d^{\nu}_{\mathrm{s,d}}$, $\kappa^2_{\mathrm{s,r}}(m) = 1/d^{\nu}_{\mathrm{s,r}}(m)$ and $\kappa^2_{\mathrm{r,d}}(m) = 1/d^{\nu}_{\mathrm{r,d}}(m)$, where ν denotes the path-loss exponent and $d_{\mathrm{s,d}}$, $d_{\mathrm{s,r}}(m)$ and $d_{\mathrm{r,d}}(m)$ are the distances between the corresponding nodes. The number of PU-RXs is limited to 1 ($N_{\mathrm{PU}} = 1$) and the mean-square gains $\rho^2_{\mathrm{s}} = 1/d^{\nu}_{\mathrm{s,PU}}$ and $\rho^2_{\mathrm{r}}(m) = 1/d^{\nu}_{\mathrm{r,PU}}$ correspond to g_{s} and $g_{\mathrm{r}}(m)$, $m = 1, \ldots, M$, respectively.

In this chapter we will assume that the SU-TXs only know their RMS channel gains to the other SU nodes and to the PU-RXs. This is comparable to the scenario with CDI of the previous chapters, with the exception that the positions of the nodes are now no longer fixed. The positions of the nodes are now assumed to be random variables, that remain constant over many frames. As a consequence, the RMS gains of the different channels also have to be considered as random variables. As in the previous chapters, we assume that the fading of all channel

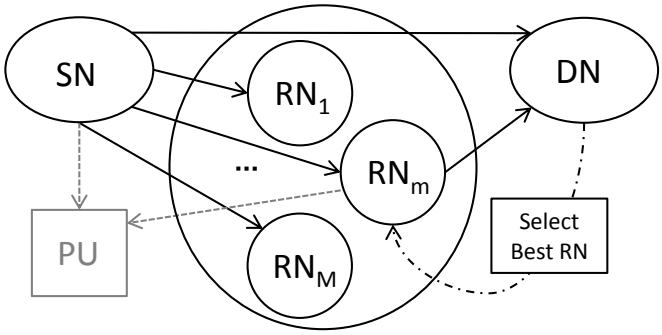


Figure 7.1: The cognitive radio relay network.

gains, for a given position of the nodes, is independent. However, note that the distances between nodes, and, therefore, also the RMS channel gains, are statistically *dependent* random variables.

7.1.2 The Performance Metric

The highest achievable rate between the SN and the DN is given by (3.67), which leads to the following outage probability

$$P_{\text{out}} = \Pr\left\{\frac{1}{2}\log_2\left(1+\gamma_{\text{S}-\text{AF}}\right) < R\right\}, \tag{7.1}$$

where R denotes a given information transmission rate and $\gamma_{\text{S}-\text{AF}}$ is given by

$$\gamma_{\text{S}-\text{AF}} = \gamma_0 + \max_m \gamma_m. \tag{7.2}$$

When the RMS values $\boldsymbol{\kappa}$ and transmit energies are given, the conditional outage probability can be expressed as

$$P_{\text{out}}(\boldsymbol{\kappa},\, E_0,\, \mathbf{E}_r) = \int_0^\infty \prod_{m=1}^{M} \Pr\left[\gamma_m < 2^{2R} - 1 - x\right] p_{\gamma_0}(x)\,\mathrm{d}x, \tag{7.3}$$

where $\mathbf{E}_r \triangleq [E_1, \ldots, E_M]$ and the probability density function (pdf) of γ_0 is given by $p_{\gamma_0}(x) = \alpha_0 e^{-\alpha_0 x}$. The probability $\Pr[\gamma_m < s]$ is given by (see [82])

$$\begin{aligned}\Pr[\gamma_m < s] = 1 - 2e^{-(\alpha_m+\beta_m)s}&\sqrt{\alpha_m\beta_m\,(s+1)\,s} \\ &\cdot\, K_1\left(2\sqrt{\alpha_m\beta_m\,(s+1)\,s}\right),\end{aligned} \tag{7.4}$$

where $K_1(z)$ is a modified Bessel function of the second kind. We note that the variables α_0, α_m and β_m, $m = 1, \ldots, M$, were already defined in (6.5), (6.6) and (6.7), respectively.

However, as the integral in (7.3) cannot be evaluated analytically, we resort to the high-SNR approximation of the conditional outage probability (7.3), derived in [27]:

$$P_{\text{out}}(\boldsymbol{\kappa},\, E_0,\, \mathbf{E}_r) \approx \frac{1}{M+1}\left[2^{2R}-1\right]^{M+1} \cdot\, \alpha_0 \prod_{m=1}^{M}(\alpha_m+\beta_m). \tag{7.5}$$

The approximate unconditional outage probability P_{out} is found by averaging (7.5) over the RMS channel gains of the SU network and the transmit symbol energies (the latter will be selected according to the RMS interference channel gains)

$$P_{\text{out}} \approx \mathbb{E}\left[P_{\text{out}}(\boldsymbol{\kappa},\, E_0,\, \mathbf{E}_{\text{r}})\right]. \tag{7.6}$$

7.1.3 The Interference Constraints

As far as the interference channels are concerned, we assume that each SU node only knows the RMS gain of its own channel towards the PU-RX. As this random variable is constant over many frames, it is reasonable to assume that this parameter can be acquired by the SU node. To keep the interference power at the PU-RX below the threshold $\bar{\Gamma}$, the SU network makes E_0 and E_m a function of the random variables ρ_{s}^2 and $\rho_{\text{r}}^2(m)$, $m=1,\ldots,M$, respectively, so that these transmit energies are also constant over many frames: we set $E_0=\Psi_{\text{s}}(\rho_{\text{s}})$ and $E_m=\Psi_m(\rho_{\text{r}}(m))$, where the functions $\Psi_{\text{s}}(x)$ and $\Psi_m(x)$, $m=1,\ldots,M$, should be determined such that the outage probability is minimized under the following AI constraints:

$$\mathbb{E}\left[E_0\left|g_{\text{s}}\right|^2\right] = \mathbb{E}_{\rho_{\text{s}}}\left[E_0\rho_{\text{s}}^2\right] \le \bar{\Gamma}, \tag{7.7}$$

and

$$\sum_{m=1}^{M}\mathbb{E}\left[E_m\left|g_{\text{r}}(m)\right|^2 I(m)\right] \;=\; \sum_{m=1}^{M}\mathbb{E}_{\rho_{\text{r}}(m)}\left[E_m h(m,\rho_{\text{r}}(m))\right] \;\le\; \bar{\Gamma}, \tag{7.8}$$

where $\mathbb{E}\left[.\right]$ denotes averaging over all channel gains and all RMS channel gains. In (7.8), we have $I(m)=1$ if the *m*th RN is selected by the DN, and $I(m)=0$ otherwise; $h(m,\rho_{\text{r}}(m))$ is a short-hand notation for $\rho_{\text{r}}^2(m)\Pr(m|\rho_{\text{r}}(m))$, where $\Pr(m|\rho_{\text{r}}(m))=\mathbb{E}\left[I(m)|\rho_{\text{r}}(m)\right]$ denotes the probability that the *m*th relay is selected, conditioned on $\rho_{\text{r}}(m)$. The quantity $\mathbb{E}_{\rho_{\text{r}}(m)}\left[E_m h(m,\rho_{\text{r}}(m))\right]$ depends on the functions $\Psi_{\text{s}}(x)$ and $\Psi_m(x)$, $m=1,\ldots,M$, and on the joint pdf of the RMS gains of the M SN-RN channels, the M RN-DN channels, the SN-PU channel and the M RN-PU channels; this joint pdf can be derived from the joint pdf of the SU network node positions.

We make the assumption that the joint pdf of the positions of the SU network nodes is invariant under a permutation of its arguments, and all noise variances are identical ($\sigma^2_{\mathrm{s,d}} = \sigma^2_{\mathrm{s,r}}(m) = \sigma^2_{\mathrm{r,d}}(m) = \sigma^2$ for $m = 1, \ldots, M$); this way, the optimum functions $\Psi_m(x)$ do not depend on the index m. Taking $\Psi_m(x) = \Psi_{\mathrm{r}}(x)$ and $E_m = \Psi_{\mathrm{r}}(\rho_{\mathrm{r}}(m))$ for $m = 1, \ldots, M$, the terms in the summation (7.8) do not depend on the index m of the considered RN; therefore, the constraint (7.8) reduces to

$$\mathbb{E}_{\rho_{\mathrm{r}}(1)}\left[E_1 h(1, \rho_{\mathrm{r}}(1))\right] \leq \frac{\bar{\Gamma}}{M}. \tag{7.9}$$

7.2 Resource Allocation

In this section we derive several energy allocation strategies that minimize the link outage probability between the SN and the DN. Because a closed-form analytical expression for the outage probability is not available, we will instead use the approximated expression for the outage probability (7.6) as the objective function. The optimization will be assisted by a *band manager*, which knows the joint pdf of the positions of the SU network nodes, and derives the optimum functions $\Psi_{\mathrm{s}}(x)$ and $\Psi_{\mathrm{r}}(x)$. Based on these functions and the RMS gains of the interference channels, the energies E_0 and E_m for $m = 1, \ldots, M$ are computed. From these energies, the noise variance σ^2, the M SN-RN channel gains and the M RN-DN channel gains for the current frame, the DN determines which relay provides the largest instantaneous SNR, and instructs the corresponding RN to transmit with the appropriate symbol energy. Note that only the DN needs to acquire instantaneous channel gains for the current frame; the quantities that must be distributed among several nodes (i.e., RMS interference channel gains, symbol energies) are constant over many frames.

7.2.1 Optimal Static Allocation Strategy

First, a static allocation strategy is considered as a reference scenario. In this static strategy the SU nodes use a constant energy allocation, i.e., E_0 and $E_m = E_{\mathrm{r}}$, $m = 1, \ldots, M$, are selected independently of the RMS interference channel gains. Because (7.6) is a decreasing function of both E_0 and E_{r}, these energies are selected such that both interference constraints are fulfilled with equality:

$$\begin{aligned} &E_0 \mathbb{E}_{\rho_{\mathrm{s}}}\left[\rho_{\mathrm{s}}^2\right] = \bar{\Gamma} \\ &E_{\mathrm{r}} \mathbb{E}_{\rho_{\mathrm{r}}(1)}\left[h(1, \rho_{\mathrm{r}}(1))\right] = \tfrac{\bar{\Gamma}}{M}. \end{aligned} \tag{7.10}$$

Note that $\mathbb{E}_{\rho_{\mathrm{r}}(1)}\left[h(1, \rho_{\mathrm{r}}(1))\right]$ is a function of E_0 and E_{r}, that can be computed off-line from the joint pdf of the positions of the SU network nodes.

7.2.2 Optimal Dynamic Allocation Strategy

In this subsection we will derive a dynamic allocation strategy. Because a joint optimization of the SN and the RNs appears very difficult, we will propose an iterative solution method, which alternates between the optimization of the SN and the RNs.

For the SN optimization, we look for the function $\Psi_{\mathrm{s}}(x)$, such that the following expression is minimized:

$$\begin{aligned} &\mathbb{E}\left[\tfrac{1}{E_0\kappa_{\mathrm{s,d}}^2}\prod_{m=1}^{M}\left(\tfrac{1}{E_0\kappa_{\mathrm{s,r}}^2(m)}+\tfrac{1}{E_m\kappa_{\mathrm{r,d}}^2(m)}\right)\right] \\ &\text{s.t. } \mathbb{E}_{\rho_{\mathrm{s}}}\left[E_0\rho_{\mathrm{s}}^2\right]\leq\bar{\Gamma}, \end{aligned} \tag{7.11}$$

where $E_0=\Psi_{\mathrm{s}}(\rho_{\mathrm{s}})$ and $E_m=\Psi_{\mathrm{r}}(\rho_{\mathrm{r}}(m))$ for $m=1,\ldots,M$; $\Psi_{\mathrm{r}}(x)$ is considered to be a known function. This is a convex optimization problem, which can be solved by the method of Lagrange multipliers. Because the objective function is a strictly decreasing function of E_0, the optimum is reached when the interference constraint is fulfilled with equality. The following expression has to be minimized over E_0

$$\begin{aligned} \mathbb{E}_{\rho_{\mathrm{s}}}\left[\mathbb{E}_{\boldsymbol{\kappa},\boldsymbol{\rho}_{\mathrm{r}}|\rho_{\mathrm{s}}}\left[\frac{1}{E_0\kappa_{\mathrm{s,d}}^2}\prod_{m=1}^{M}\left(\frac{1}{E_0\kappa_{\mathrm{s,r}}^2(m)}+\frac{1}{E_m\kappa_{\mathrm{r,d}}^2(m)}\right)\right]\right] \\ +\mu_{\mathrm{s}}\mathbb{E}_{\rho_{\mathrm{s}}}\left[E_0\rho_{\mathrm{s}}^2-\bar{\Gamma}\right], \end{aligned} \tag{7.12}$$

where $\boldsymbol{\rho}_{\mathrm{r}}\triangleq[\rho_{\mathrm{r}}(1),\ldots,\rho_{\mathrm{r}}(M)]$. For given μ_{s}, the optimal $\Psi_{\mathrm{s}}(x)$ is found by minimizing

$$\begin{aligned} \frac{1}{\Psi_{\mathrm{s}}(x)}\mathbb{E}_{\boldsymbol{\kappa},\boldsymbol{\rho}_{\mathrm{r}}|\rho_{\mathrm{s}}=x}\left[\frac{1}{\kappa_{\mathrm{s,d}}^2}\prod_{m=1}^{M}\left(\frac{1}{\Psi_{\mathrm{s}}(x)\kappa_{\mathrm{s,r}}^2(m)}+\frac{1}{E_m\kappa_{\mathrm{r,d}}^2(m)}\right)\right] \\ +\mu_{\mathrm{s}}\left(\Psi_{\mathrm{s}}(x)x^2-\bar{\Gamma}\right), \end{aligned} \tag{7.13}$$

over $\Psi_{\mathrm{s}}(x)$ for every value of x. To solve this minimization one needs the conditional pdf $p(\boldsymbol{\kappa},\boldsymbol{\rho}_{\mathrm{r}}|\rho_{\mathrm{s}}=x)$ for every value of x, where $\boldsymbol{\kappa}$ and $\boldsymbol{\rho}_{\mathrm{r}}$ contain the RMS gains of the channels between the SU network nodes and the RMS gains of the interference channels from the RNs to the PU; this pdf can be computed off-line from the joint pdf of the positions of the SU network nodes. The optimal value of the Lagrange multiplier μ_{s} is found by solving the following equation

$$\mathbb{E}_{\rho_{\mathrm{s}}}\left[\Psi_{\mathrm{s}}(\rho_{\mathrm{s}})\rho_{\mathrm{s}}^2\right]=\bar{\Gamma}. \tag{7.14}$$

This equation can be solved by using a stochastic subgradient method described in section 2.5.

For the RN optimization, we look for the function $\Psi_{\rm r}(x)$, such that the following expression is minimized:

$$\begin{array}{l} \mathbb{E}\left[\frac{1}{E_0\kappa^2_{\rm s,d}}\prod_{m=1}^{M}\left(\frac{1}{E_0\kappa^2_{\rm s,r}(m)}+\frac{1}{E_m\kappa^2_{\rm r,d}(m)}\right)\right] \\ \text{s.t. } \mathbb{E}_{\rho_{\rm r}(1)}\left[E_1 h(1,\rho_{\rm r}(1))\right] \leq \frac{\bar{\Gamma}}{M}, \end{array} \tag{7.15}$$

where $E_1 = \Psi_{\rm r}(\rho_{\rm r}(1))$, $E_0 = \Psi_{\rm s}(\rho_{\rm s})$ and $E_m = \Phi_{\rm r}(\rho_{\rm r}(m))$ for $m = 2, \ldots, M$; $\Psi_{\rm s}(x)$ and $\Phi_{\rm r}(x)$ are considered to be known functions. Because the interference constraint is non-convex this optimization problem can be challenging to solve. However because the cumulative distribution function of $\rho_{\rm r}(m)$ is continuous, it can be proven that the considered optimization problem has zero duality gap. As a consequence the optimization problem can be solved by the method of Lagrange multipliers without loss of optimality. For more details we refer to [83, 84]. We will minimize the following expression as a function of E_1

$$\begin{aligned} \mathbb{E}_{\rho_{\rm r}(1)}\left[\mathbb{E}_{\boldsymbol{\kappa},\rho_{\rm s},\boldsymbol{\rho}_{\rm r}|\rho_{\rm r}(1)}\left[\frac{1}{E_0\kappa^2_{\rm s,d}}\prod_{m=1}^{M}\left(\frac{1}{E_0\kappa^2_{\rm s,r}(m)}+\frac{1}{E_m\kappa^2_{\rm r,d}(m)}\right)\right]\right] \\ + \mu_{\rm r}\mathbb{E}_{\rho_{\rm r}(1)}\left[E_1 h(1,\rho_r(1)) - \frac{\bar{\Gamma}}{M}\right]. \end{aligned} \tag{7.16}$$

In the same manner as for the SN, we find the optimum function $\Psi_{\rm r}(x)$ for given $\mu_{\rm r}$ by minimizing

$$\begin{aligned} \frac{1}{\Psi_{\rm r}(x)}\mathbb{E}_{\boldsymbol{\kappa},\rho_{\rm s},\boldsymbol{\rho}_{\rm r}|\rho_{\rm r}(1)=x}\left[\frac{1}{\kappa^2_{\rm r,d}(1)}\frac{1}{E_0\kappa^2_{\rm s,d}}\prod_{m'=2}^{M}\left(\frac{1}{E_0\kappa^2_{\rm s,r}(m')}+\frac{1}{E_{m'}\kappa^2_{\rm r,d}(m')}\right)\right] \\ + \mu_{\rm r}\left(\Psi_{\rm r}(x)h(1,x) - \frac{\bar{\Gamma}}{M}\right), \end{aligned} \tag{7.17}$$

for every value of x. This requires that for every x one knows the value of $h(1,x)$ as a function of $E_1 = \Psi_{\rm r}(x)$, and the conditional pdf $p(\boldsymbol{\kappa}, \rho_{\rm s}, \boldsymbol{\rho}_{\rm r}|\rho_{\rm r}(1) = x)$; these functions can be computed off-line from the joint pdf of the positions of the SU network nodes. The value of $h(1,x)$ is given by

$$\begin{aligned} h(1,x) &= x^2 \mathrm{Pr}(1|\mathrm{x}) \\ &= x^2\mathbb{E}_{\boldsymbol{\kappa},\rho_{\rm s},\boldsymbol{\rho}_{\rm r}|\rho_{\rm r}(1)=x}\left[\int_0^{\infty}\prod_{m=2}^{M}\Pr\left[\gamma_m < \gamma_1\right]p(\gamma_1)\,\mathrm{d}\gamma_1\right], \end{aligned} \tag{7.18}$$

where $p(\gamma_1)$ can be found from its CDF in (7.4). The optimal value of the Lagrange multiplier $\mu_{\rm r}$ is found by solving the following equation

$$\mathbb{E}_{\rho_{\rm r}(1)}\left[\Psi_{\rm r}(\rho_{\rm r}(1))h(1,\rho_{\rm r}(1))\right] = \frac{\bar{\Gamma}}{M}, \tag{7.19}$$

which can be done by using a stochastic subgradient method.

We will now iteratively alternate between the above algorithms, as described below.

1. Initialize the functions $\Psi_{\mathrm{s}}(x)$ and $\Psi_{\mathrm{r}}(x)$, $\forall x$; initialize μ_{s}, μ_{r}
2. Optimization SN: find $\Psi_{\mathrm{s}}(x)$ and μ_{s} by solving (7.13) and (7.14)
   (a) minimize (7.13) $\forall x$, taking $\Psi_{\mathrm{r}}(x)$ from previous iteration
   (b) update $\mu_{\mathrm{s}} = \max(\mu_{\mathrm{s}} + \epsilon \mathbb{E}_{\rho_{\mathrm{s}}} \left[\Psi_{\mathrm{s}}(\rho_{\mathrm{s}})\rho_{\mathrm{s}}^2 - \bar{\Gamma}\right], 0)$
   (c) if $\mathbb{E}_{\rho_{\mathrm{s}}} \left[\Psi_{\mathrm{s}}(\rho_{\mathrm{s}})\rho_{\mathrm{s}}^2 - \bar{\Gamma}\right] = 0$ go to step 3, otherwise go to 2a
3. Optimization RN: find $\Psi_{\mathrm{r}}(x)$ and μ_{r} by solving (7.17) and (7.19)
   (a) minimize (7.17) $\forall x$, taking $\Psi_{\mathrm{s}}(x)$ from step 2 and $\Phi_{\mathrm{r}}(x)$ equal to $\Psi_{\mathrm{r}}(x)$ from previous iteration
   (b) update the energy allocation of the other RNs:
   take $\Phi_{\mathrm{r}}(x)$ equal to $\Psi_{\mathrm{r}}(x)$ from step 3a
   (c) update μ_{r}
   $= \max(\mu_{\mathrm{r}} + \epsilon \mathbb{E}_{\rho_{\mathrm{r}}(1)} \left[\Psi_{\mathrm{r}}(\rho_{\mathrm{r}}(1))h(1, \rho_{\mathrm{r}}(1)) - \frac{\bar{\Gamma}}{M}\right], 0)$
   (d) if $\mathbb{E}_{\rho_{\mathrm{r}}(1)} \left[\Psi_{\mathrm{r}}(\rho_{\mathrm{r}}(1))h(1, \rho_{\mathrm{r}}(1)) - \frac{\bar{\Gamma}}{M}\right] = 0$ go to step 4, otherwise go to 3a
4. If convergence is reached: exit, otherwise go to step 2

The parameter ϵ denotes the scalar step size used in the subgradient method. From the simulations we have noticed that this algorithm shows a rapid convergence. The change in the functions $\Psi_{\mathrm{s}}(x)$ and $\Psi_{\mathrm{r}}(x)$ is minimal between the different iterations, and convergence is typically reached after only 3 iterations.

7.2.3 Suboptimal Dynamic Allocation Strategy

We will now derive a simple low-complexity dynamic allocation strategy, where the functions $\Psi_{\mathrm{s}}(x)$ and $\Psi_{\mathrm{r}}(x)$ are restricted to take only 2 values, i.e.,

$$\Psi_{\mathrm{s}}(x) = \begin{cases} E_{\mathrm{s}}^{(\max)}, & x < \theta_{\mathrm{s}} \\ E_{\mathrm{s}}^{(\min)}, & x > \theta_{\mathrm{s}} \end{cases}, \tag{7.20}$$

$$\Psi_{\mathrm{r}}(x) = \begin{cases} E_{\mathrm{r}}^{(\max)}, & x < \theta_{\mathrm{r}} \\ E_{\mathrm{r}}^{(\min)}, & x > \theta_{\mathrm{r}} \end{cases}, \tag{7.21}$$

where $E_{\mathrm{s}}^{(\min)}$, $E_{\mathrm{s}}^{(\max)}$, $E_{\mathrm{r}}^{(\min)}$, $E_{\mathrm{r}}^{(\max)}$, θ_{s} and θ_{r} are selected such that outage probability (7.6) is minimized. In contrast with the optimal allocation strategy

from subsection 7.2.2, this suboptimum allocation strategy involves optimizing over only 6 parameters. After the optimization the *band manager* transmits the optimal value of $E_{\mathrm{s}}^{(\min)}$, $E_{\mathrm{s}}^{(\max)}$ and θ_{s} to the SN, and broadcasts the optimal values $E_{\mathrm{r}}^{(\min)}$, $E_{\mathrm{r}}^{(\max)}$ and θ_{r} to every RN. The SU nodes can have a very low complexity: the nodes only have to compare the RMS gain of their channel to the PU with a threshold and apply the corresponding symbol energy, according to (7.20) and (7.21).

To find the optimum we have to minimize the following function over $\mathbf{E}^{(\min)}$, $\mathbf{E}^{(\max)}$ and $\mathbf{\Theta}$

$$\begin{aligned} &\mathbb{E}\left[\tfrac{1}{E_0\kappa_{\mathrm{s,d}}^2}\cdot\prod_{m=1}^{M}\left(\tfrac{1}{E_0\kappa_{\mathrm{s,r}}^2(m)}+\tfrac{1}{E_m\kappa_{\mathrm{r,d}}^2(m)}\right)\right] \\ &\text{s.t. } \mathbb{E}_{\rho_{\mathrm{s}}}\left[E_0\rho_{\mathrm{s}}^2\right]\leq\bar{\Gamma} \\ &\quad\mathbb{E}_{\rho_{\mathrm{r}}(1)}\left[E_1 h(1,\rho_{\mathrm{r}}(1))\right]\leq\tfrac{\bar{\Gamma}}{M}, \end{aligned} \tag{7.22}$$

where $\mathbf{E}^{(\min)}=[E_{\mathrm{s}}^{(\min)},E_{\mathrm{r}}^{(\min)}]$, $\mathbf{E}^{(\max)}=[E_{\mathrm{s}}^{(\max)},E_{\mathrm{r}}^{(\max)}]$ and $\mathbf{\Theta}=[\theta_{\mathrm{s}},\theta_{\mathrm{r}}]$.

We search for solutions where both interference constraints are fulfilled with equality. The optimum is found by using a global optimization algorithm. Note that this optimization can be performed off-line from the joint pdf of the SU network node positions.

1. Global optimization: $\forall\mathbf{E}^{(\min)}$, $\forall\mathbf{E}^{(\max)}$ that belong to a finite set of symbol energy values
   (a) initialize $\theta_{\mathrm{s}}=0$, $\theta_{\mathrm{r}}=0$
   (b) Increase θ_{s} until the interference constraint for the SN is fulfilled with equality
   (c) Increase θ_{r} until the interference constraint for the RNs is fulfilled with equality
   (d) Evaluate P_{out} from (7.6)
2. Return the value of ($\mathbf{E}^{(\min)}$, $\mathbf{E}^{(\max)}$, $\mathbf{\Theta}$) for which P_{out} achieved its minimum value

Note that the algorithm needs to consider only values of $\mathbf{E}^{(\min)}$ and $\mathbf{E}^{(\max)}$ where the interference generated by the SN (RNs) would be strictly larger than $\bar{\Gamma}$ if $\theta_{\mathrm{s}}=\infty$ ($\theta_{\mathrm{r}}=\infty$), and strictly smaller than $\bar{\Gamma}$ if $\theta_{\mathrm{s}}=0$ ($\theta_{\mathrm{r}}=0$).

7.3 Numerical Results

In this section we discuss the numerical results for the different algorithms. For the amplify-and-forward network, we assumed that all SU nodes are uniformly

and independently distributed inside a circle centered at the origin with radius 1. The PU-RX is located at coordinates $(0,\ 1.5)$.

The exact (7.1) and approximated (7.6) outage probabilities are obtained by means of Monte Carlo simulations. The mean-square value of the Rayleigh fading channel gains is derived from the distance between the corresponding nodes. The path loss exponent ν is chosen equal to 3.5 and the information rate is set to $R = 1$ bit/s/Hz. In Fig. 7.2 the outage probabilities are shown as a function of $\bar{\Gamma}/\sigma^2$ for the optimum static allocation (SA), the optimum dynamic allocation (ODA) and the suboptimum dynamic allocation (SDA). We have shown the exact outage probability of the direct link between the SN and DN ($M = 0$), along with the exact and approximated outage probability in the case where the DN selects the best of M RNs ($M = 1,\ M = 2$). For the direct link case, the exact outage probability ($\Pr\{\log_2(1+\gamma_0) < R\}$) was used as the objective function for the derivation of the optimum SA, ODA and SDA.

First, we note that the gain in performance by using multiple RNs is significant. It is also clear that the approximation is tight at high SNR. Further, we notice that our ODA and SDA strategies have a considerably better performance than the SA strategy: the ODA strategy has a minimum gain of 1.2 dB compared to the SA strategy and the SDA strategy is only about 0.3 dB worse than the performance of the ODA strategy.

In Fig. 7.3 we compare our ODA strategy which uses an *average* interference (AI) constraint with the following ODA strategy under a *peak* interference (PI) constraint: $E_0 = \bar{\Gamma}/|g_{\mathrm{s}}|^2$, $E_m = \bar{\Gamma}/|g_{\mathrm{r}}(m)|^2$, $m = 1, \ldots, M$. Fig. 7.3 clearly confirms that the allocation strategy under an AI constraint outperforms the allocation strategy under a PI constraint. So when an AI constraint is reasonable for the PU network, which is the case when the quality of service of the PU network is determined by the average SNR [85], the performance of the SU network can be significantly improved.

7.4 Chapter Summary

In this chapter, we combined the AI constraint with the best-relay selection protocol. Several strategies were investigated that minimize the link outage probability between the SN and the DN of the SU network, and, at the same time, limit the AI power at the PU-RX to a predefined threshold. The allocation strategies assume that the SN and the RNs have access to the mean-square gain of their channel to the PU-RX. Because this information is slowly varying, it is feasible to obtain this information at the SU nodes. In the numerical results, we have shown that both dynamic allocation strategies show a significant performance improvement over the static allocation strategy. Further, it was shown that the performance of the SDA strategy has a very small performance loss compared with the optimal, more

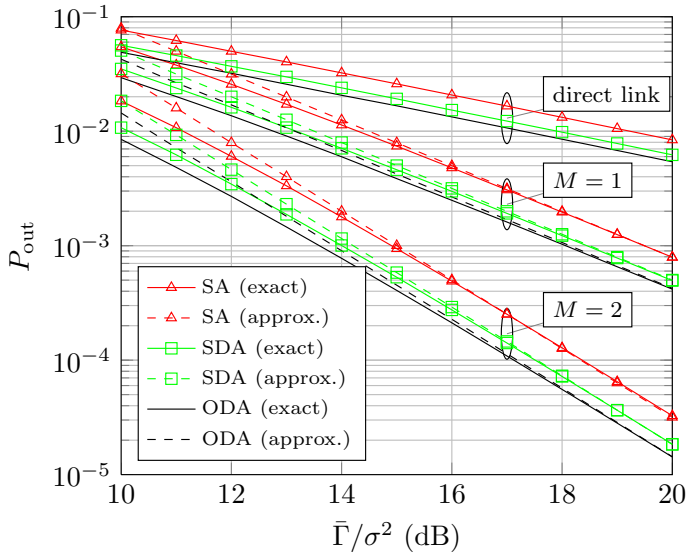


Figure 7.2: Outage probability versus $\bar{\Gamma}/\sigma^2$ (direct link, $M = 1$ and $M = 2$).

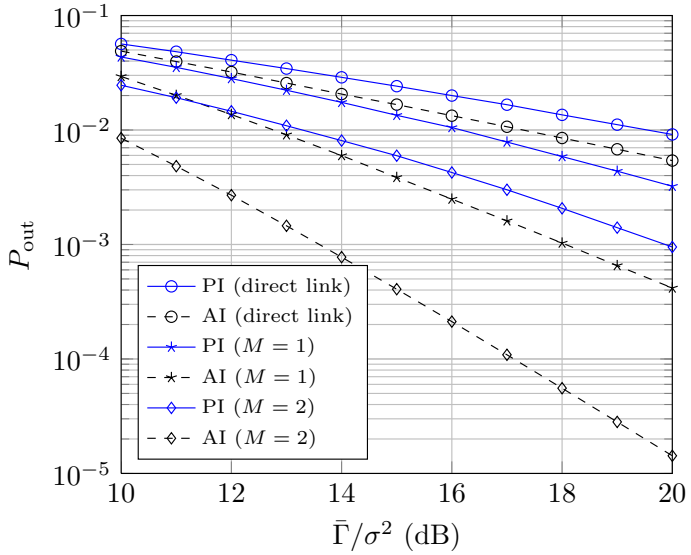


Figure 7.3: Comparison of the performance under an average or peak interference constraint (direct link, $M = 1$ and $M = 2$).

complex dynamic allocation strategy. Our analysis indicates that the dynamic algorithms under an AI constraint show a significant performance improvement as compared to the ODA under a PI constraint.

However, it is clear that combining the AI constraint with the best-relay selection protocol is not an easy task. The optimization for the simple cooperative network, discussed in this chapter, already became quite involved. Further, the knowledge of the joint probability distribution of all the node positions is hard to obtain in a practical network. For these reasons, we will focus our attention again to PI constraints in the next chapters as we begin our discussion about multicarrier networks.

8

Resource Allocation for Multicarrier Cooperative Cognitive Radio Networks

In this chapter (which is based on contribution [86]), we take a closer look at a decode-and-forward cooperative cognitive radio network. Instead of minimizing the outage probability under a fixed rate, as in the previous chapters, we now try to maximize the rate of the SU network under a fixed outage probability constraint. This approach benefits applications that want to maximize their transmission rate under a constraint that guarantees the reliability of the transmission. An example of such an application is video transmission, where the TX can choose the quality of the transmitted stream: when the link is reliable, the video can be transmitted at a high-definition resolution. However, when only a weak link is present, the TX chooses a low-resolution stream so it can guarantee the required reliability of the transmission.

Further, we consider a multicarrier system and introduce an accurate approximation for the outage probability, which is used for the formulation of the resource allocation problem.

The organization of this chapter is as follows: in section 8.1, we introduce the cooperative network and the channel model. Further, we define the constraints on the outage probability, the power and interference at the PU-RXs. Then we introduce the resource allocation problem and its solution in section 8.2. We validate the accuracy of the approximation and the performance of the resource allocation

scheme through numerical simulations in section 8.3. Finally, the conclusions are presented in section 8.4.

8.1 System Model

8.1.1 The Cooperative Network

We consider a cooperative network in a cognitive scenario as shown in Fig. 8.1. The SU cooperative network consists of a SN, a DN and M RNs. We assume that there is no direct link present between the SN and DN. The decode-and-forward protocol is used at the relays. The transmission of a frame occurs in 2 time slots. In the first time slot the SN transmits a frame consisting of several OFDM symbols to the different RNs. Each OFDM symbol has a length $T_{\mathrm{s}} = (N + \nu)T$, where T, N and ν denote the sampling interval, the number of subcarriers and the length of the guard interval (expressed in sampling intervals), respectively. Each RN decodes its received messages, which are then re-encoded and forwarded to the DN. The message received at a RN over a particular subcarrier j is forwarded over a subcarrier k, which might be different from j. This subcarrier pairing is expressed by the binary variable $\pi_{j,k}$, which is equal to 1 when the subcarrier j of the first hop is paired with subcarrier k of the second hop, and 0 otherwise. Each subcarrier is decoded by a single RN only, and each subcarrier j can only be paired with one subcarrier k. This can be summarized as follows

$$\forall j : \sum_{k=1}^{N} \pi_{j,k} = 1, \quad \forall k : \sum_{j=1}^{N} \pi_{j,k} = 1. \tag{8.1}$$

The binary variable $z_{j,k}(m)$ equals 1 when the subcarrier pair (j,k) is allocated to the mth RN, and 0 otherwise. A pair (j, k) can be assigned to only one RN; this can be expressed as

$$\forall j, k \text{ s.t. } \pi_{j,k} = 1 : \sum_{m=1}^{M} z_{j,k}(m) = 1. \tag{8.2}$$

The subcarrier pairing and the relay allocation is also schematically shown in Fig. 8.1, which assumes $N = 5$, $\pi_{2,4} = \pi_{4,1} = \pi_{5,2} = 1$ and $z_{2,4}(1) = z_{4,1}(1) = z_{5,2}(M) = 1$; in the first and second time slot, the subcarriers with $j \in (1, 3)$ and $k \in (3, 5)$ are not used (the corresponding symbol energies are zero).

8.1.2 Channel Model

We assume that the wireless channel between any two nodes of the SU network is a frequency-selective multipath fading channel with a correlation time that is much

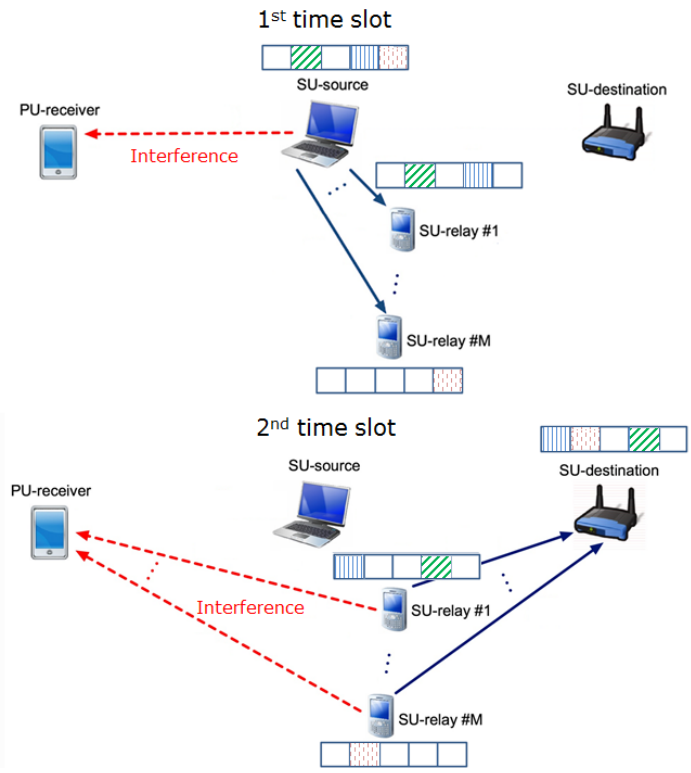


Figure 8.1: Topology of the network.

longer than the OFDM symbol duration $T_{\rm s}$, so that the complex path gains can be considered constant over an interval $T_{\rm s}$. Considering root-raised-cosine transmit and receive filters with roll-off factor β [16], the impulse response $h(u,i)$ of the cascade of the transmit filter, the wireless channel and the receive filter, related to the *i*th OFDM symbol, is expressed as

$$h(u,i) = \sum_{l=1}^{L} c_l(i) g(u-\tau_l), \tag{8.3}$$

where $c_l(i)$ and τ_l denote the gain and delay of the *l*th path, L denotes the number of paths and $g(u)$ is a raised-cosine pulse with roll-off factor β, which is defined as

$$g(u) = \begin{cases} 1, & u = 0 \\ \frac{\beta}{2}\sin\left(\frac{\pi}{2\beta}\right), & u = \pm\frac{T}{2\beta} \\ \frac{\sin\left(\frac{\pi u}{T}\right)}{\frac{\pi u}{T}} \frac{\cos\left(\frac{\pi\beta u}{T}\right)}{1-\left(\frac{2\beta u}{T}\right)^2}, & u \neq \pm\frac{T}{2\beta},\ u \neq 0 \end{cases}. \tag{8.4}$$

We assume that $h(nT,i) = 0$ for $n < 0$ and for $n > \nu$, which is achieved by adding to all path delays a sufficiently large common delay to make $h(u,i)$ causal, and by taking ν sufficiently large. The gains of the different paths are independent zero-mean circularly symmetric Gaussian random variables. According to Jakes' model [12], we take $\mathbb{E}\left[c_l(m+i)c_l^*(i)\right] = J_0(2\pi f_{\rm d} m T_{\rm s})\sigma_l^2$, where $J_0(x)$ represents the zeroth-order Bessel function of the first kind, and $f_{\rm d}$ denotes the Doppler frequency. For later use, we define $\mathbf{R}_{\rm c} = \text{diag}(\sigma_1^2, \ldots, \sigma_L^2)$ and the vector $\mathbf{h}(i) \in \mathbb{C}^{(\nu+1)\times 1}$ containing the $\nu+1$ samples of $h(u,i)$, given by

$$\mathbf{h}(i) \triangleq \mathbf{G}\mathbf{c}(i), \tag{8.5}$$

where $\mathbf{c}(i) \triangleq [c_1(i), \ldots, c_L(i)]^T$ and $\mathbf{G}$ is the $(\nu+1)\times L$ matrix with entries $\mathbf{G}_{k,l} \triangleq g((k-1)T - \tau_l)$, $k = 1, \ldots, \nu+1; l = 1, \ldots, L$. The corresponding frequency response is given by $\mathbf{H}(i) = \mathbf{F}\mathbf{h}(i)$, where the Fourier matrix $\mathbf{F} \in \mathbb{C}^{N\times(\nu+1)}$ is determined by

$$\mathbf{F}_{k,l} \triangleq e^{-j2\pi(k-1)(l-1)/N}, \quad k = 1, \ldots, N; l = 1, \ldots, \nu+1. \tag{8.6}$$

The *k*th component of $\mathbf{H}(i)$ denotes the channel gain seen by the *k*th subcarrier in the *i*th OFDM symbol.

In appendix 8.A, we derive the model for the imperfect CSI. This model is very similar to the model for ICSI presented in appendix 5.A for single-carrier systems. In appendix 8.A, we show that the MMSE estimate of the channel gains for the subcarriers will be denoted by the vector $\mathbf{H}_{\rm p}(i)$. The actual channel gain $\mathbf{H}(i)$ can then be written as

$$\mathbf{H}(i) = \mathbf{H}_{\rm p}(i) + \mathbf{F}\mathbf{e}(i), \tag{8.7}$$

where $\mathbf{e}(i)$ is independent of $\mathbf{H}_{\mathrm{p}}(i)$ and $\mathbf{H}_{\mathrm{p}}(i)$ is a function of the available channel information vector $\mathbf{ICSI}(i)$. For notational convenience, we will from now on omit the time index i.

8.1.3 Definitions of the Constraints

8.1.3.1 Link Quality Constraint

The objective of the resource allocation is to maximize the rate between the SN and the DN. Because the TX has only imperfect CSI available, it can happen that the required rate R cannot be achieved on a certain subcarrier link, because it exceeds the instantaneous capacity. Therefore we impose a quality constraint on each subcarrier link, by limiting the outage probability P_{out}, which is defined as the cumulative distribution of the capacity (in bits per channel use), conditioned on the available channel information vector $\mathbf{ICSI}$

$$P_{\mathrm{out}} \triangleq \Pr\left[\log_2\left(1+|H|^2\frac{E}{\sigma_w^2}\right) \leq R \,\middle|\, \mathbf{ICSI}\right] \tag{8.8}$$

$$= \Pr\left[|H|^2 \leq d \,\middle|\, \mathbf{ICSI}\right]$$

$$= \Pr\left[|H|^2 \leq d \,\middle|\, H_{\mathrm{p}}\right], \tag{8.9}$$

where E denotes the transmit energy per data symbol, H denotes the actual channel gain between a transmitting and receiving node as seen by the considered subcarrier, H_{p} denotes the estimated channel gain, $\sigma_w^2 \triangleq \frac{N+\nu}{N}\sigma^2$, σ^2 denotes the noise variance at the receive filter output and

$$d = \frac{2^R-1}{E}\sigma_w^2. \tag{8.10}$$

The quality constraint imposes a maximum value for the outage probability P_{out} that each subcarrier link is allowed to experience. This outage probability P_{out} denotes the probability that the actual instantaneous channel capacity conditioned on $\mathbf{ICSI}$ is lower than the transmitted rate on that subcarrier.

The actual expression (8.9) for the outage probability is similar to the expression shown in (5.56). However, instead of using the exact expression, we will approximate P_{out} as shown in appendix 8.B. In the appendix, we discuss a general approach that can be used for any distribution of $|H|^2$ conditioned on H_{p} as long as the cumulant generating function is known.

In our scenario, the probability distribution of $|H|^2$ conditioned on the MMSE channel estimate H_{p} is a scaled noncentral χ^2 distribution with 2 degrees of free-

dom. Its cumulant generating function [87] is given by

$$\begin{aligned}\mu(s) &= \ln\left(\mathbb{E}_H\left[e^{s|H|^2}\,\middle|\,H_{\rm p}\right]\right)\\ &= \frac{s|H_{\rm p}|^2}{1-s\sigma_{\rm p}^2} - \ln(1-s\sigma_{\rm p}^2), \quad (\sigma_{\rm p}^2 s < 1),\end{aligned} \tag{8.11}$$

where $\sigma_{\rm p}^2 \triangleq \mathbb{E}_H\left[|H-H_{\rm p}|^2|H_{\rm p}\right]$[1] denotes the conditional mean-square estimation error. According to appendix 8.B, the outage probability (8.9) is approximated as

$$P_{\rm out} = Q(\sqrt{\ddot{\mu}(s^*)}\,|s^*|)e^{\frac{1}{2}\ddot{\mu}(s^*)s^{*2}+\mu(s^*)-s^*\dot{\mu}(s^*)}, \tag{8.12}$$

where $\dot{\mu}(s)$ and $\ddot{\mu}(s)$ are the first and second derivative of $\mu(s)$ with respect to s, s^* is the negative solution of $\dot{\mu}(s^*)=d$, and $Q(x)$ is the probability that a zero-mean unit variance Gaussian random variable exceeds x.

In the following we want to achieve a fixed value for the outage probability $P_{\rm out}$. We first solve (8.12) numerically for s^*, and then compute the corresponding value of d as

$$\begin{aligned}d(H_{\rm p},P_{\rm out}) &= \dot{\mu}(s^*)\\ &= \frac{|H_{\rm p}|^2+\sigma_{\rm p}^2(1-s^*\sigma_{\rm p}^2)}{(1-s^*\sigma_{\rm p}^2)^2}.\end{aligned} \tag{8.13}$$

Taking (8.10) into account, we obtain

$$R = \log_2\left(1+d(H_{\rm p},P_{\rm out})\frac{E}{\sigma_w^2}\right), \tag{8.14}$$

which is the rate (bits/channel use) that yields an outage probability $P_{\rm out}$ when the MMSE estimate equals $H_{\rm p}$. We notice that (8.14) is almost the same expression as the instantaneous capacity, except that $d(H_{\rm p},P_{\rm out})$ replaces the squared magnitude $|H|^2$ of the actual channel gain. Note from (8.13) that $d(H_{\rm p},P_{\rm out}) = |H_{\rm p}|^2 = |H|^2$ in the case of perfect estimation (i.e., $\sigma_{\rm p}^2=0$). The quantity $d(H_{\rm p},P_{\rm out})$ takes into account the imperfect CSI at the TX, and by using (8.14) as the objective function to be maximized, we have automatically satisfied the quality constraint (8.9).

The RNs of the SU network use the decode-and-forward protocol, introduced in section 3.6.3, which means that the maximum instantaneous rate we can transmit on the (j,k) subcarrier pair at the *m*th RN can be written as

$$\begin{aligned}R_{j,k}(m) = \frac{1}{2}\min\Bigg(&\log_2\left(1+d_{{\rm s},j}(m)\frac{E_{{\rm s},j,k}(m)}{\sigma_w^2}\right),\\ &\log_2\left(1+d_{{\rm r},k}(m)\frac{E_{{\rm r},j,k}(m)}{\sigma_w^2}\right)\Bigg),\end{aligned} \tag{8.15}$$

[1]The value of $\sigma_{\rm p}^2$ for a certain subcarrier k is given by the *k*th diagonal element of $\mathbf{F}\mathbf{R}_{\rm e}\mathbf{F}^H$.

where $E_{\mathrm{s},j,k}(m)$ and $E_{\mathrm{r},j,k}(m)$ denote the transmit energy per data symbol for the *j*th subcarrier at the SN paired with the *k*th subcarrier at the *m*th RN, and for the *k*th subcarrier at the *m*th RN paired with the *j*th subcarrier at the SN. The quantities $d_{\mathrm{s},j}(m)$ and $d_{\mathrm{r},k}(m)$ denote the quantity $d(H_{\mathrm{p}}, P_{\mathrm{out}})$ for the channels seen by the *j*th subcarrier from the SN to the *m*th RN, and by the *k*th subcarrier from the *m*th RN to the DN, respectively. Obviously, the rate $R_{j,k}(m)$ is maximized by taking

$$E_{\mathrm{r},j,k}(m) = \eta_{j,k}(m) E_{\mathrm{s},j,k}(m), \tag{8.16}$$

where we have introduced $\eta_{j,k}(m) = \frac{d_{\mathrm{s},j}(m)}{d_{\mathrm{r},k}(m)}$.

8.1.3.2 Transmit Energy Constraints

The total transmit energy at each SU node is assumed to be limited. Denoting by $E_{\mathrm{s}}^{(\max)}$ and $E_{\mathrm{r}}^{(\max)}(m)$ the maximal total transmit energies at the SN and the *m*th RN, $m = 1, \ldots, M$, the transmit energy constraints for the SN and for each RN are expressed as

$$\sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{m=1}^{M} \pi_{j,k} z_{j,k}(m) E_{\mathrm{s},j,k}(m) \leq E_{\mathrm{s}}^{(\max)}, \tag{8.17}$$

$$\sum_{j=1}^{N} \sum_{k=1}^{N} \pi_{j,k} z_{j,k}(m) \eta_{j,k}(m) E_{\mathrm{s},j,k}(m) \leq E_{\mathrm{r}}^{(\max)}(m), \tag{8.18}$$

where in (8.18) we have made use of (8.16).

8.1.3.3 Interference Constraints

The SU network is present in the same bandwidth as N_{PU} PU-RXs. According to the underlay paradigm, the SU network may only gain access to the bandwidth of the PU network, if the SU network keeps the interference at each PU-RX below a corresponding threshold Γ_q, $q = 1, \ldots, N_{\mathrm{PU}}$. As we assume that the exact channel gains towards the PU-RXs are unknown to the SU-TXs, the SU network cannot guarantee that the interference constraints are always satisfied. However, the SU network will make sure that the average interference conditioned on the available channel information is below the threshold Γ_q, $q = 1, \ldots, N_{\mathrm{PU}}$.

Let us denote by $I_j(q)$ the average interference energy experienced by the *q*th PU-RX when a unit energy data symbol is transmitted by the SN on subcarrier j. Assuming that the PU network is active in the frequency band (f_1, f_2), $I_j(q)$ is obtained as the following conditional expectation

$$I_j(q) = \mathbb{E}\left[\int_{f_1}^{f_2} |H_{\mathrm{tr}}(f)|^2 |H_q(f)|^2 S_j(f) \mathrm{d}f \,\middle|\, \mathbf{H}_{\mathrm{p},q} \right], \tag{8.19}$$

where $H_{\rm tr}(f)$ and $H_q(f)$ are the transfer functions of the root-raised-cosine transmit filter and the multipath channel between the SN and the qth PU-RX, respectively, and

$$S_j(f) = \left(\frac{\sin\left(\pi(N+\nu)(fT - \frac{j}{N})\right)}{\sqrt{(N+\nu)}\sin\left(\pi(fT-\frac{j}{N})\right)} \right)^2. \tag{8.20}$$

The expectation in (8.19) is over $H_q(f)$, conditioned on the MMSE estimate $\mathbf{H}_{{\rm p},q}$ of the channel gains for the link between the SN and the qth PU-RX. A similar expression holds for the average interference energy $I_{k,m}(q)$ experienced by the qth PU-RX when a unit energy data symbol is transmitted by the mth RN on subcarrier k. For every PU-RX, $q = 1, \ldots, N_{\rm PU}$, the interference constraints to be satisfied by the SN and the RNs are expressed as follows

$$\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{m=1}^{M} \pi_{j,k} z_{j,k}(m) E_{{\rm s},j,k}(m) I_j(q) \leq \Gamma_q, \tag{8.21}$$

$$\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{m=1}^{M} \pi_{j,k} z_{j,k}(m) \eta_{j,k}(m) E_{{\rm s},j,k}(m) I_{k,m}(q) \leq \Gamma_q. \tag{8.22}$$

Note that the previous chapters did not take the transmit filter into account. However, in these chapters we made the assumptions of a flat-fading channel and single-carrier modulation. In this case, we can show that the interference present in the frequency band (f_1, f_2) scales with

$$\begin{aligned} \int_{f_1}^{f_2} |H_{\rm tr}(f)|^2 |H_q(f)|^2 {\rm d}f &= |H_q(f)|^2 \int_{f_1}^{f_2} |H_{\rm tr}(f)|^2 {\rm d}f \\ &= |H_q(f)|^2. \end{aligned} \tag{8.23}$$

In the last step, we have assumed that $|H_{\rm tr}(f)|$ is only non-zero between f_1 and f_2. The result in (8.23), which is precisely the channel gain that we used in the previous chapters, can thus be seen as a special case of the formulas used in (8.21) and (8.22).

8.2 Resource Allocation

The SU network attempts to maximize the total rate between the SN and the DN, subject to the various constraints. The resource allocation problem will be formulated in a centralized manner, and can be solved by a band manager [31, 34]. In situations where a centralized solution is not desirable, the obtained results can still serve as a useful upper bound for the performance of distributed algorithms.

The optimization problem reads

$$\max_{\mathbf{E},\boldsymbol{\pi},\mathbf{z}} \sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{m=1}^{M} \pi_{j,k} z_{j,k}(m) \log_2\left(1 + d_{\mathrm{s},j}(m)\frac{E_{\mathrm{s},j,k}(m)}{\sigma_w^2}\right) \quad \text{s.t. (8.1), (8.2), (8.17), (8.18), (8.21), (8.22),} \tag{8.24}$$

where $\mathbf{E}$, $\boldsymbol{\pi}$ and $\mathbf{z}$ consist of the values $E_{\mathrm{s},j,k}(m)$, $\pi_{j,k}$ and $z_{j,k}(m)$ $\forall j,k,m$, respectively. The optimization problem (8.24) is a mixed integer programming problem. As this problem is non-convex, the duality gap will in general be non-zero. However, in [88, 89] it is shown that the duality gap will be virtually zero when the number of carriers N is greater than 8. This means that the original problem (8.24) can be solved by the dual problem with negligible performance loss. According to (2.16), we can define the dual problem as follows

$$\min_{\beta\geq 0,\boldsymbol{\kappa}\geq 0,\boldsymbol{\lambda}\geq 0,\boldsymbol{\mu}\geq 0} D(\beta,\boldsymbol{\kappa},\boldsymbol{\lambda},\boldsymbol{\mu}), \tag{8.25}$$

where

$$\begin{aligned} D(\beta,\boldsymbol{\kappa},\boldsymbol{\lambda},\boldsymbol{\mu}) &= \max_{\boldsymbol{\pi}} \Big[\sum_{j=1}^{N}\sum_{k=1}^{N} \pi_{j,k} \max_{\mathbf{z}} \Big[\sum_{m=1}^{M} z_{j,k}(m) \max_{\mathbf{E}} \Big[F(j,k,m) \Big] \Big] \Big] \\ &+ \beta E_{\mathrm{s}}^{(\max)} + \sum_{m=1}^{M} \kappa(m) E_{\mathrm{r}}^{(\max)}(m) + \sum_{q=1}^{N_{\mathrm{PU}}} \lambda_q \Gamma_q + \sum_{q=1}^{N_{\mathrm{PU}}} \mu_q \Gamma_q \\ &\text{s.t. } (8.1) - (8.2), \end{aligned} \tag{8.26}$$

where

$$F(j,k,m) = \log_2\left(1 + d_{\mathrm{s},j}(m)\frac{E_{\mathrm{s},j,k}(m)}{\sigma_w^2}\right) - \alpha(m) E_{\mathrm{s},j,k}(m), \tag{8.27}$$

with

$$\alpha(m) = \beta + \kappa(m)\eta_{j,k}(m) + \sum_{q=1}^{N_{\mathrm{PU}}} \left(\lambda_q I_j(q) + \mu_q \eta_{j,k}(m) I_{k,m}(q)\right). \tag{8.28}$$

In (8.26) we have introduced the dual variables β, $\boldsymbol{\kappa} \triangleq [\kappa_1,\ldots,\kappa_M]^T$, $\boldsymbol{\lambda} \triangleq [\lambda_1,\ldots,\lambda_{N_{\mathrm{PU}}}]^T$ and $\boldsymbol{\mu} \triangleq [\mu_1,\ldots,\mu_{N_{\mathrm{PU}}}]^T$ which are associated with constraints (8.17), (8.18), (8.21) and (8.22), respectively. From equation (8.26), it is clear that this maximization problem can be decomposed into three separate subproblems.

For the first subproblem, the optimization problem is solved for any valid relay assignment $z_{j,k}(m) = 1$ and subcarrier pairing $\pi_{j,k} = 1$. This subproblem is formulated as $F_{\max}(j,k,m) = \max_{E_{\mathrm{s},j,k}(m)} \Big[F(j,k,m) \Big]$; the optimal transmit energy $E_{\mathrm{s},j,k}(m)$ is given by $E_{\mathrm{s},j,k}(m) = \left[\frac{1}{\ln(2)\alpha(m)} - \frac{\sigma_w^2}{d_{\mathrm{s},j}(m)} \right]^+$.

The second subproblem involves the optimization over $z_{j,k}(m)$ for any valid pairing $\pi_{j,k} = 1$; the optimal value of $z_{j,k}(m)$ $(m = 1, \ldots, M)$ is given by

$$z_{j,k}(m) = \begin{cases} 1, & \text{if } m = \arg\max_m F_{\max}(j,k,m) \\ 0, & \text{otherwise} \end{cases}. \tag{8.29}$$

Now that we have found the optimal energy allocation and the relay selection for any given subcarrier pairing $\pi_{j,k} = 1$, we still have to find the optimal subcarrier pairing.

This third problem reduces to a linear assignment problem. Such a problem can be efficiently solved by using a standard Hungarian algorithm, with complexity $O(N^3)$ [90]. We illustrate the Hungarian method by an example in appendix 8.C.

So now we have found the value of $D(\beta, \boldsymbol{\kappa}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ for a specific value of the Lagrange multipliers. The optimal values of the Lagrange multipliers β^*, $\boldsymbol{\kappa}^*$, $\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}^*$ that minimize $D(\beta, \boldsymbol{\kappa}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ are found by means of the subgradient method described in section 2.4.

8.3 Numerical Results

We first investigate the accuracy of our proposed approximation (8.12) of the outage probability P_{out}. This approximation will be compared to the exact (simulated) outage probability P_{out} (8.8) and to a Gaussian approximation proposed in [91]. In [91] the probability density (conditioned on the estimation H_{p}) of the capacity $\log_2(1 + \frac{|H|^2 E}{\sigma_w^2})$ is approximated by a Gaussian distribution with mean $\log_2(1 + \frac{a}{b})$ and variance $\frac{2a}{(\ln(2)(a+b))^2}$, where $a \triangleq |H_{\text{p}}|^2/\sigma_{\text{p}}^2$ and $b \triangleq \sigma_w^2/(E\sigma_{\text{p}}^2)$. In Fig. 8.2a (with $(a, b) = (1,\ 10)$) and 8.2b (with $(a, b) = (100,\ 0.1)$), we show both approximations along with the exact outage probability P_{out} as a function of the rate R (bits/channel use). We notice that the Gaussian approximation has its highest accuracy for R near the mean of the capacity $\log_2(1 + \frac{|H|^2 E}{\sigma_w^2})$ (this mean is 0.250 and 9.967 for Fig. 8.2a and Fig. 8.2b, respectively), however the region of interest corresponds to smaller values of R, yielding small P_{out}. From the figure it is clear that the overall accuracy is much higher for the approximation (8.12) based on the exponentially tilted distribution than for the Gaussian approximation. Further, according to [91] the Gaussian approximation is only useful when a is large, whereas our approximation is accurate for a wide range of a and b values.

The simulation parameters of the resource allocation problem are summarized in table 8.1. For the channel model we have used the ITU vehicular A model, which we described in section 3.1.3. The wireless channel between the nodes is normalized: $\sum_{l=1}^{L} \sigma_l^2 = 1$. The presence of a single PU-RX ($N_{\text{PU}} = 1$) is assumed that occupies the same frequency band as the SU network. The maximal total transmit energies $E_{\text{s}}^{(\max)}$ and $E_{\text{r}}^{(\max)}(m)$, $m = 1, \ldots, M$, are chosen equal

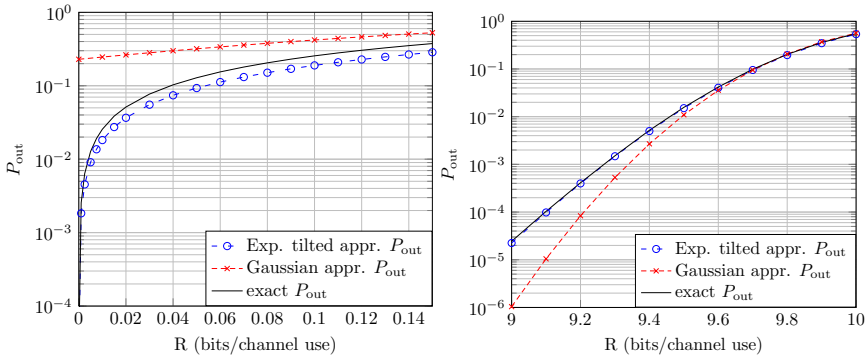


Figure 8.2: Approximations of P_{out}. (a) $a = 1$, $b = 10$. (b) $a = 100$, $b = 0.1$.

to $E^{(\max)}$. Fig. 8.3 shows the rate R (bits/channel use) as function of $E^{(\max)}/\sigma_w^2$ (dB), for Γ/σ_w^2 = 0, 10 and ∞ (dB). The interference threshold Γ has a considerable impact on the performance of the SU network for high values of $E^{(\max)}/\sigma_w^2$. For $\Gamma/\sigma_w^2 = 0$ dB and 10 dB we clearly notice a limiting value of the rate R with increasing $E^{(\max)}/\sigma_w^2$. In this region we say that the performance is interference limited: a further increase in the maximal transmit energy $E^{(\max)}$ will have little influence on the achievable rate R. When the ratio $E^{(\max)}/\sigma_w^2$ has a low value, we notice that the curves for the different ratios of Γ/σ_w^2 are close to each other. In this region the performance is limited by the available transmit energy $E^{(\max)}$, which means that a less restrictive value of the interference threshold Γ will lead to a negligible performance benefit.

We have also illustrated the effect of the memory P of the predictor. When $P = 1$ we notice that the SU suffers a considerable performance loss compared to the case with perfect CSI. However when a predictor is used with memory $P = 4$, we already notice a considerable performance improvement. By using a larger predictor memory we get a better estimate of the actual channel gains and the resulting performance loss because of imperfect CSI is reduced. If we further increase the predictor memory P from 4 to 7, we notice that the obtained performance improvement is negligible. This is explained by the fact that the correlation between the oldest value of the estimated CSI $\hat{\mathbf{h}}$ and the actual impulse response $\mathbf{h}$ becomes very low ($J_0(2\pi f_{\text{d}} 7DT_{\text{s}}) = 0.26$) because of the time-varying channel.

Parameter	Value
Subcarriers (N)	64
Cyclic Prefix (ν)	20
Sampling Interval (T)	$(1.92 \cdot 10^6)^{-1}$s
Doppler Frequency (f_{d})	144 Hz
Estimation Error (σ_{e}^2)	0.01
Outage Probability (P_{out})	0.001
Roll-off factor (β)	0.1
Number of RNs (M)	3
CSI update interval (D)	7

Table 8.1: System parameters.

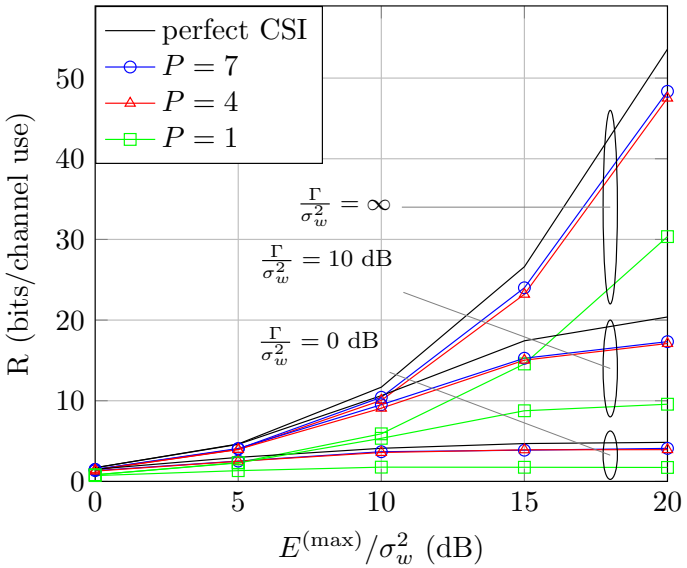


Figure 8.3: Rate R versus $E^{(\max)}/\sigma_w^2$.

8.4 Chapter Summary

In this chapter we proposed an approximation for the outage probability conditioned on the available channel information. In the numerical results, we have shown that this approximation is more accurate than the often used Gaussian approximation.

Based on this approximation we derived an optimal resource allocation algorithm that takes into account imperfect CSI. Simulations have shown that a small performance loss compared to the case with perfect CSI can be achieved, provided that the memory P of the predictor is large enough. Finally, the results provided in this chapter offer an interesting benchmark to which the performance of more practical distributed algorithms can be compared.

In this and previous chapters, we derived several allocation algorithms for cognitive radio networks that rely on the use of theoretic performance metrics. While chapters 5-7 tried to minimize the outage probability of the SU network, this chapter took a different approach, by optimizing the rate of the SU network under a fixed outage probability constraint.

The following chapters take a look at resource allocation algorithms that optimize a more practical performance metric. This performance metric is closely related to the actual performance of a network and is called the goodput.

8.A Imperfect CSI

In order to perform dynamic resource allocation, the transmitting nodes need some form of CSI. Similar as in appendix 5.A, we assume that the transmitting node obtains an estimated impulse response via a feedback channel. The channel estimate $\hat{\mathbf{H}}(i)$ of the frequency response $\mathbf{H}(i)$ is given by

$$\hat{\mathbf{H}}(i) = \mathbf{H}(i) + \mathbf{w}(i), \tag{8.30}$$

where $\mathbf{w}(i) \sim N_{\mathrm{c}}(\mathbf{0}, \mathbf{I}_N \sigma_{\mathrm{e}}^2)$ and σ_{e}^2 represents the estimation error.

However, it is possible to further improve this estimate by using the prior knowledge that $\mathbf{h}(i)$ has only $\nu + 1$ components. By using this information, an estimate $\hat{\mathbf{h}}(i)$ of the impulse response $\mathbf{h}(i)$ can be obtained from the following optimization

$$\min_{\hat{\mathbf{h}}(i)} \|\hat{\mathbf{H}}(i) - \mathbf{F}\hat{\mathbf{h}}(i)\|^2, \tag{8.31}$$

which yields

$$\begin{aligned} \hat{\mathbf{h}}(i) &= (\mathbf{F}^H\mathbf{F})^{-1}\mathbf{F}^H\hat{\mathbf{H}}(i) \\ &= \mathbf{h}(i) + \mathbf{n}(i), \end{aligned} \tag{8.32}$$

where $\mathbf{n}(i) \sim N_{\mathrm{c}}(\mathbf{0}, (\mathbf{F}^H\mathbf{F})^{-1}\sigma_{\mathrm{e}}^2)$. As $\hat{\mathbf{h}}(i)$ contains all the useful information about the channel, it will actually be this estimate that is fed back to the corresponding TX.

However, the TX receives an impulse response estimate only once every D OFDM symbols, which means that the CSI received at the TX will be outdated. The TX uses a MMSE estimator to approximate the actual impulse response. The MMSE impulse response estimate $\hat{\mathbf{h}}_{\mathrm{p}}(i)$ is derived from the P previously received impulse response estimates $\mathbf{ICSI}(i) \triangleq [\hat{\mathbf{h}}^T(i-D), \ldots, \hat{\mathbf{h}}^T(i-DP)]^T$. Similar as in section 3.3, it can be proven that the following relationship holds

$$\begin{aligned} \mathbf{h}(i) &= \mathbb{E}\left[\mathbf{h}(i)|\mathbf{ICSI}(i)\right] + \mathbf{e}(i) \\ &= \mathbf{X}\mathbf{Y}^{-1}\mathbf{ICSI}(i) + \mathbf{e}(i) \\ &= \hat{\mathbf{h}}_{\mathrm{p}}(i) + \mathbf{e}(i), \end{aligned} \tag{8.33}$$

where $\hat{\mathbf{h}}_{\mathrm{p}}(i) \sim N_{\mathrm{c}}(\mathbf{0}, \mathbf{G}\mathbf{R}_{\mathrm{c}}\mathbf{G}^H - \mathbf{R}_{\mathrm{e}})$ and $\mathbf{e}(i) \sim N_{\mathrm{c}}(\mathbf{0}, \mathbf{R}_{\mathrm{e}})$, where

$$\mathbf{R}_{\mathrm{e}} = \mathbf{G}\mathbf{R}_{\mathrm{c}}\mathbf{G}^H - \mathbf{X}\mathbf{Y}^{-1}\mathbf{X}^H. \tag{8.34}$$

Introducing the matrix $\mathbf{J} \in \mathbb{C}^{P\times P}$ with entries $\mathbf{J}_{k,l} \triangleq J_0(2\pi f_{\mathrm{d}} D T_{\mathrm{s}}(k-l))$, $k = 1, \ldots, P; l = 1, \ldots, P$, the matrices $\mathbf{X}$ and $\mathbf{Y}$ are written as follows

$$\mathbf{X} = [J_0(2\pi f_{\mathrm{d}} D T_{\mathrm{s}}), \ldots, J_0(2\pi f_{\mathrm{d}} P D T_{\mathrm{s}})] \otimes \mathbf{G}\mathbf{R}_{\mathrm{c}}\mathbf{G}^H, \tag{8.35}$$

$$\mathbf{Y} = \mathbf{J} \otimes \mathbf{G}\mathbf{R}_{\mathrm{c}}\mathbf{G}^H + \mathbf{I}_P \otimes (\mathbf{F}^H\mathbf{F})^{-1}\sigma_{\mathrm{e}}^2. \tag{8.36}$$

The predicted channel gains for the subcarriers are the components of the vector $\mathbf{H}_{\mathrm{p}}(i) = \mathbf{F}\hat{\mathbf{h}}_{\mathrm{p}}(i)$.

8.B Derivation of Approximation (8.12)

Approximating the probability density $p_{|H|^2}(x|H_{\rm p})$, of $|H|^2$ conditioned on $H_{\rm p}$, by a Gaussian density is expected to yield a poor approximation of (8.9) for small $P_{\rm out}$, because in this case the mean $\mathbb{E}_H\left[|H|^2|H_{\rm p}\right]$ is much larger than d. To put more weight in the neighborhood of d, we will exponentially tilt $p_{|H|^2}(x|H_{\rm p})$ to create a new density $p_Z(x)$, and approximate $p_Z(x)$ by a Gaussian density. The tilted density is given by [15]

$$p_Z(x) = p_{|H|^2}(x|H_{\rm p})e^{sx-\mu(s)}, \quad s \leq 0, \tag{8.37}$$

where $\mu(s)$ is given by (8.11). Its mean and variance equal $\dot{\mu}(s)$ and $\ddot{\mu}(s)$, where $\dot{\mu}(s)$ and $\ddot{\mu}(s)$ are the first and second derivative of $\mu(s)$ with respect to s. We take $s = s^*$ with $\dot{\mu}(s^*) = d$, so that the mean of $p_Z(x)$ equals d. Taking into account that $s^* \leq 0$, we obtain

$$s^* = \frac{2d - \sigma_{\rm p}^2 - \sqrt{\sigma_{\rm p}^4 + 4d|H_{\rm p}|^2}}{2d\sigma_{\rm p}^2}. \tag{8.38}$$

The outage probability $P_{\rm out}$ can be expressed in terms of $p_Z(x)$ as

$$\begin{aligned} P_{\rm out} &= \int_0^d p_{|H|^2}(x|H_{\rm p})\,{\rm d}x \\ &= \int_0^d p_Z(x)e^{\mu(s^*)-s^*x}\,{\rm d}x. \end{aligned} \tag{8.39}$$

Approximating $p_Z(x)$ by a Gaussian density with mean $\dot{\mu}(s^*)$ and variance $\ddot{\mu}(s^*)$, and replacing in (8.39) the integration interval by $(-\infty, d)$ yields the approximation (8.12).

8.C The Hungarian Algorithm

We illustrate the Hungarian algorithm by using an example. Assume that we have the following matrix

$$\begin{pmatrix} 16 & 84 & 87 & 56 \\ 5 & 61 & 34 & 58 \\ 55 & 49 & 20 & 9 \\ 7 & 42 & 42 & 10 \end{pmatrix}. \tag{8.40}$$

The goal is to select from each row a single element, while satisfying the following requirements:

1. Each column has to be selected just once.

2. The sum of all the selected elements has to be minimal.

In the first step we subtract the row minimum from each row. For example, this means that for the first row, we subtract 16 from each element of that row. This gives us the following matrix

$$\begin{pmatrix} 0 & 68 & 71 & 40 \\ 0 & 56 & 29 & 53 \\ 46 & 40 & 11 & 0 \\ 0 & 35 & 35 & 3 \end{pmatrix}. \tag{8.41}$$

It is clear that adding or subtracting a constant to each row does not change our optimum.

In the second step, we will do the same for the columns. We get

$$\begin{pmatrix} 0 & 33 & 60 & 40 \\ 0 & 21 & 18 & 53 \\ 46 & 5 & 0 & 0 \\ 0 & 0 & 24 & 3 \end{pmatrix}. \tag{8.42}$$

We check if the minimum number of lines that cover all zeros is equal to the number of rows of the matrix. We find

$$\begin{pmatrix} 0 & 33 & 60 & 40 \\ 0 & 21 & 18 & 53 \\ \cancel{46} & \cancel{5} & \cancel{0} & \cancel{0} \\ \cancel{0} & \cancel{0} & \cancel{24} & \cancel{3} \end{pmatrix}. \tag{8.43}$$

Which means we need three lines to cover all zeros. As this is lower than the number of rows, we need an additional step that will create additional zeros.

The smallest element that is not crossed out is 18. We subtract this number from all uncovered elements and add it to the elements that are crossed out twice. After these operations, we get the following matrix

$$\begin{pmatrix} 0 & 15 & 42 & 22 \\ 0 & 3 & 0 & 35 \\ 64 & 5 & 0 & 0 \\ 18 & 0 & 24 & 3 \end{pmatrix}. \tag{8.44}$$

We again check the minimum number of lines that is needed to cover all zeros. Now, we find

$$\begin{pmatrix} 0 & 15 & 42 & 22 \\ 0 & 3 & 0 & 35 \\ 64 & 5 & 0 & 0 \\ 18 & 0 & 24 & 3 \end{pmatrix}. \tag{8.45}$$

As there are now 4 lines, the algorithm is finished and we find the following optimum

$$\begin{pmatrix} \boxed{0} & 15 & 42 & 22 \\ 0 & 3 & \boxed{0} & 35 \\ 64 & 5 & 0 & \boxed{0} \\ 18 & \boxed{0} & 24 & 3 \end{pmatrix}, \tag{8.46}$$

which corresponds to

$$\begin{pmatrix} \boxed{16} & 84 & 87 & 56 \\ 5 & 61 & \boxed{34} & 58 \\ 55 & 49 & 20 & \boxed{9} \\ 7 & \boxed{42} & 42 & 10 \end{pmatrix}. \tag{8.47}$$

We note that if the sum of the selected elements needs to be maximal instead of minimal, we can simply apply the same algorithm for the same matrix where all its elements are negated.

Part III

Optimization of the Goodput

9

Accurate Modeling of the κESM with Imperfect Channel State Information

In chapter 4, we discussed the effective SNR mapping (ESM), which is a technique that transforms a vector of subcarrier SNRs into a scalar effective SNR. Our goal now is to apply this technique at the TX, such that the TX can dynamically choose its transmission parameters. However, the calculation of the κESM requires the availability of perfect CSI at the TX, which can be problematic as the available CSI, in a practical system, is often imperfect, due to the presence of noise, and outdated, because of the time-varying channel.

If we condition the κESM on the available imperfect CSI, the κESM becomes a random variable. It appears that its distribution, when only imperfect CSI is available, cannot be expressed in closed form.

There have been several papers [92–95] which apply a moment matching approximation to approximate the pdf, of the effective SNR related to the EESM, with a certain distribution. However, these papers rely on perfect CSI and try to calculate the long-term average performance of a system, while we are interested in providing the TX with an accurate statistical model of the effective SNR so it can make an accurate prediction of the instantaneous link quality. This way, the TX can make a better choice regarding its transmission parameters and thereby avoid packet errors or a loss in spectral efficiency by over- or underestimating the actual channel conditions.

Therefore, this chapter (which is based on contribution [96]) extends the approach in [92–95] to κESM in combination with imperfect CSI. Our proposed approximation will then be used in the following chapter to optimize the goodput when only imperfect CSI is available at the TX.

Section 9.1 briefly describes the system model, while section 9.2 introduces our proposed approximation of the κESM. In section 9.3 we compare our model in terms of approximation accuracy against models based on the Gaussian distribution, the gamma distribution and the more complex Pearson and generalized extreme value (GEV) distributions. This chapter is concluded in section 9.4.

9.1 System Model

A point-to-point OFDM link is considered where a message is transmitted by means of the packet-oriented BIC-OFDM communication system described in chapter 4. At the RX, the *k*th OFDM subcarrier, with $k \in \mathcal{N} \triangleq \{1, \ldots, N\}$ is observed as

$$z_k \triangleq \sqrt{E_k} H_k x_k + n_k, \tag{9.1}$$

where E_k is the transmit energy on the *k*th subcarrier, H_k is the corresponding channel coefficient, x_k is the constellation symbol transmitted on subcarrier k corresponding to $m_k \in \mathcal{D}_m$ coded bits and $\mathbb{E}\left[|x_k|^2\right] = 1$, and $n_k \in N_{\mathrm{c}}\left(0, \sigma_w^2\right)$ is the ambient noise.

The received signal-to-noise ratio (SNR) associated with the *k*th subcarrier can then be defined as

$$\gamma_k \triangleq \frac{E_k |H_k|^2}{\sigma_w^2}. \tag{9.2}$$

In the following chapter, the TX will use the κESM to optimize the energy allocation per subcarrier E_k and the transmission mode ϕ.

The time-varying multipath channel is modeled as in section 8.1.2. The only difference is that we now no longer assume that $g(u)$ is a raised-cosine pulse, and take $g(u) = \delta(u)$. This means that the samples $h(mT, i)$, $m = 0, \ldots, \nu$, are now independent circularly symmetric zero-mean Gaussian complex random variables and $\mathbb{E}\left[h(mT, i+j)h^*(mT, i)\right] = \mathrm{J}_0(2\pi f_{\mathrm{d}} j T_s)\sigma_m^2$. For later reference, we define[1] $\mathbf{R}_h \triangleq \mathrm{diag}(\sigma_0^2, \ldots, \sigma_\nu^2)$ and the channel impulse vector $\mathbf{h}(i) \triangleq [h(0, i), h(T, i), \ldots, h(\nu T, i)]^T$.

[1] Note that if we have a number of paths $L < \nu + 1$, only L diagonal elements of $\mathbf{R}_h$ are strictly greater than 0.

In order to select the optimal TM ϕ, the TX requires an estimation of the current channel state. As in (8.7), it can be shown that the predicted frequency response $\mathbf{H}_{\mathrm{p}}(i)$ is linked to the actual frequency response $\mathbf{H}(i)$ as follows:

$$\mathbf{H}(i) = \mathbf{H}_{\mathrm{p}}(i) + \mathbf{F}\mathbf{e}(i), \tag{9.3}$$

where $\mathbf{e}(i) \sim N_{\mathrm{c}}(\mathbf{0}, \mathbf{R}_{\mathrm{e}})$ and $\mathbf{F}$ denotes the Fourier matrix introduced in (8.6). As this covariance matrix is not necessary full rank, for later reference we define

$$\mathbf{F}\mathbf{e}(i) = \mathbf{B}\mathbf{s}(i), \tag{9.4}$$

where $\mathbf{s}(i) \sim N_{\mathrm{c}}(\mathbf{0}, \mathbf{I}_q)$, q is the rank of the covariance matrix $\mathbf{F}\mathbf{R}_{\mathrm{e}}\mathbf{F}^H$ and $\mathbf{B}\mathbf{B}^H = \mathbf{F}\mathbf{R}_{\mathrm{e}}\mathbf{F}^H$. The matrix $\mathbf{B} \in \mathbb{C}^{N \times q}$ can be easily found from an eigenvalue decomposition of the matrix $\mathbf{F}\mathbf{R}_{\mathrm{e}}\mathbf{F}^H$. For notational convenience, the index i will be omitted in the sequel.

9.2 Statistical Model of κESM

In this section we propose a simple statistical model for the description of the effective SNR γ_{eff} resulting from κESM, which takes into account the imperfect and outdated CSI. If perfect CSI were available, the TX could make a proper selection of its transmission parameters, based on the corresponding value of the effective SNR. However, when the TX only has outdated and imperfect CSI at its disposal, simply replacing the actual CSI by its prediction when computing the effective SNR can lead to an over- or underestimation of the instantaneous quality of the link. A better selection of the transmission mode would result if the TX knew how γ_{eff} is distributed given the predicted CSI $\mathbf{H}_{\mathrm{p}}$. However, as the κESM is highly non-linear in terms of $\{\gamma_k, k = 1, \ldots, N\}$, the exact pdf conditioned on the prediction $\mathbf{H}_{\mathrm{p}}$ is extremely difficult to obtain [97].

In the context of approximating the distribution of the effective SNR γ_{eff} associated with the EESM in the case of perfect CSI, it is proposed in [92] to approximate $e^{-\gamma_{\mathrm{eff}}/\beta}$ as a beta-distributed random variable. We will now apply this method to the effective SNR resulting from κESM. The main difference with [92] however, is that we will not consider perfect CSI, but derive the distribution of γ_{eff}, which we introduced in (4.25), conditioned on a prediction $\mathbf{H}_{\mathrm{p}}$ of the current channel state. If $\alpha_{k,\mu} \triangleq \frac{1}{\sum_{n \in \mathcal{N}} m_n} \frac{\psi^k(\mu)}{2^{m_k - 1}}$ and $(\mathbf{B}\mathbf{s})_k$ denotes the kth component of the vector $\mathbf{B}\mathbf{s}$, the following random variable

$$Y = \sum_{k \in \mathcal{N}} \sum_{\mu=1}^{\frac{\sqrt{2^{m_k}}}{2}} \alpha_{k,\mu} e^{-\frac{E_k \left| H_{\mathrm{p},k} + (\mathbf{B}\mathbf{s})_k \right|^2 \left(\mu d_{k,\min}\right)^2}{4\beta\sigma_w^2}}, \tag{9.5}$$

for given $H_{\mathrm{p},k}$, $k = 1, \ldots, N$, will be approximated by a beta-distributed random variable. A beta-distributed random variable Z takes values from the interval $[0, 1]$; its cumulative distribution function (CDF) is given by $\Pr[Z < z] = B(z, a, b)/B(1, a, b)$ for $0 \leq z \leq 1$, where

$$B(z, a, b) = \int_0^z u^{a-1}(1-u)^{b-1}\mathrm{d}u, \tag{9.6}$$

denotes the incomplete beta function, and a and b are shaping parameters of the beta distribution. The parameters a en b are related to the mean and variance of Z as follows [92]

$$a = \frac{\mathbb{E}\left[Z\right]\left(\mathbb{E}\left[Z\right] - \mathbb{E}\left[Z\right]^2 - \mathrm{Var}[Z]\right)}{\mathrm{Var}[Z]}, \tag{9.7}$$

$$b = \frac{(1 - \mathbb{E}\left[Z\right])(\mathbb{E}\left[Z\right] - \mathbb{E}\left[Z\right]^2 - \mathrm{Var}[Z])}{\mathrm{Var}[Z]}. \tag{9.8}$$

We will approximate the actual distribution of Y conditioned on a prediction $\mathbf{H}_{\mathrm{p}}$ by a beta distribution, such that the mean and variance of the beta distribution are the same as the mean and variance of the actual distribution. Hence, the shaping parameters a and b of the approximating beta distribution are obtained by replacing in (9.7)-(9.8) $\mathbb{E}\left[Z\right]$ and $\mathrm{Var}[Z]$ by $\mathbb{E}\left[Y\right]$ and $\mathrm{Var}[Y]$, respectively, which are derived in closed form in appendix 9.A (see (9.14), (9.25)).

The beta distribution based approximation of the CDF of γ_{eff} then becomes [92]

$$\Pr[\gamma_{\mathrm{eff}} \leq x] \approx 1 - \frac{B(e^{-\frac{x}{\beta}}, a, b)}{B(1, a, b)}, \tag{9.9}$$

for $x \geq 0$.

9.3 Numerical Results

We consider the communication system characterized by the parameters from table 9.1. These parameters are based on a LTE system which means that only the 1320 inner subcarriers of the 2048 subcarriers are available for data transmission [98]. The sampling frequency is thus equal to $\frac{1}{T} = \frac{2048}{1320}20$ MHz and the energy per symbol $E_{\mathrm{s}} = E_k$, $k = 1, \ldots, N$. For the channel model we use the ITU pedestrian B model [13]. Our beta distribution based approximation model will be compared to other approximation models from literature. The latter include approximations of

- the distribution of γ_{eff} by the (more involved) generalized extreme value (GEV) distributions [94] and Pearson distributions [99], and by the (analytically simpler) lognormal distribution [100] and gamma distribution [93];

Parameter	Value
SNR ($E_{\rm s}/\sigma_w^2$)	10 dB
Bandwidth (B)	20 MHz
FFT size	2048
Length of cyclic prefix (ν)	160
Doppler frequency ($f_{\rm d}$)	144 Hz
CSI update interval (D)	7
Estimation error ($\sigma_{\rm e}^2$)	12.5 dB
Scaling coefficient (β)	1

Table 9.1: System parameters.

- the distribution of $Y = e^{-\gamma_{\rm eff}/\beta}$ by a Gaussian distribution [94] and a log-normal distribution [95]; the latter will be referred to as the log-Gaussian approximation, to clearly make the distinction with the case where the distribution of $\gamma_{\rm eff}$ is approximated by a lognormal distribution [100].

While these approximations from literature pertain to the EESM with perfect CSI, we have applied these approximations to κESM with predicted CSI.

In Fig. 9.1 we show the CDF and the complementary CDF (CCDF) resulting from the empirical distribution of $\gamma_{\rm eff}$ and the various approximations, for a predicted frequency response given by $H_{{\rm p},k} = 0.5 \cdot \frac{(1+i)}{\sqrt{2}}$ ($\forall k$). A 4-QAM constellation is used ($m_k = m = 2$, $k = 1, \ldots, N$) and the number of used subcarriers $N = 64$. The subcarriers are non-contiguous and are evenly spaced across the available bandwidth B. The estimation error is given by $\mathbf{Fe} \sim N_{\rm c}(\mathbf{0}, \mathbf{F}\mathbf{R}_{\rm e}\mathbf{F}^H)$. The empirical results are generated from 10^6 samples obtained by simulation. From Fig. 9.1 we note that both the GEV and Pearson approximation are very close to the simulated CDF and CCDF. We further notice that the Gaussian approximation is very poor in tracking the simulated CCDF, while the log-Gaussian approximation is incapable of tracking the simulated CDF. Finally, we see that the performance of the proposed beta distribution based approximation is very close to the accuracy of the GEV and Pearson approximations, and clearly outperforms the other approximations methods with the exception of the gamma approximation. The latter is shown to be slightly more accurate for the considered realization of the predicted channel frequency response.

A quantitative measure of the approximation accuracy is obtained from the Kullback-Leibler (KL) divergence, which is defined as [101]

$$\mathrm{D}(p||q) = \int_{-\infty}^{\infty} p(x) \ln \frac{p(x)}{q(x)} \, \mathrm{d}x, \tag{9.10}$$

where $p(x)$ is the simulated pdf and $q(x)$ denotes the approximating pdf. This metric is always positive, and is a measure of the similarity between two probabil-

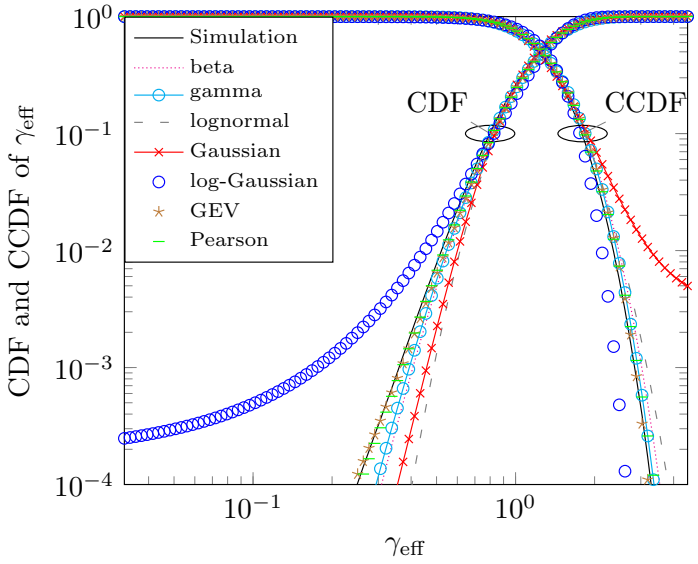


Figure 9.1: CDF and CCDF of γ_{eff} ($N = 64$, $m = 2$).

ity distributions. A smaller value denotes a higher similarity between both pdfs. However, the value of this metric is only defined if, for every value of x, $q(x) = 0$ implies that $p(x) = 0$. Because the GEV and Pearson approximations are equal to 0 in certain intervals, this condition cannot be guaranteed. Therefore we will consider the Jensen-Shannon (JS) divergence [102]. This metric is also a measure of the similarity between two pdfs, but it does not have the restriction of the KL-divergence. The JS-divergence is defined as

$$\text{JSD}(p||q) = \frac{1}{2}\text{D}(p||m) + \frac{1}{2}\text{D}(q||m), \tag{9.11}$$

where $m(x)$ is a pdf which is defined as $m(x) = \frac{1}{2}(p(x) + q(x))$. Table 9.2 shows the JS-divergence of the different approximating distributions, again for $H_{\text{p},k} = 0.5 \cdot \frac{(1+i)}{\sqrt{2}}$ ($\forall k$). The values of the JS-divergence are consistent with the approximation accuracies observed from Fig. 9.1, which confirms the usefulness of the JS-divergence as a similarity metric.

Fig. 9.1 and table 9.2 have shown the performance of the different approximations for a single realization of $\mathbf{H}_{\text{p}}$; however what really matters is which approximation has the best performance averaged over $\mathbf{H}_{\text{p}}$. Therefore, we have generated 1000 independent realizations of $\mathbf{H}_{\text{p}}$ according to the following distribution $N_{\text{c}}(\mathbf{0}, \mathbf{F}(\mathbf{R}_h - \mathbf{R}_{\text{e}})\mathbf{F}^H)$, and averaged the corresponding JS-divergences over these realizations. The results for the various approximating distributions can be found in table 9.3, for the number of used subcarriers $N \in \{8, 64, 128\}$. For

model	JS-divergence
beta	3.19×10^{-4}
gamma	8.49×10^{-5}
lognormal	2.76×10^{-3}
Gaussian	6.09×10^{-3}
log-Gaussian	7.74×10^{-3}
GEV	7.31×10^{-5}
Pearson	4.32×10^{-6}

Table 9.2: JS-divergence ($N = 64$, $m = 2$).

model	$N = 8$	$N = 64$	$N = 128$
beta	8.50×10^{-3}	1.41×10^{-3}	1.35×10^{-3}
gamma	4.37×10^{-3}	1.76×10^{-3}	1.71×10^{-3}
lognormal	3.12×10^{-3}	1.66×10^{-3}	1.79×10^{-3}
Gaussian	2.63×10^{-2}	4.42×10^{-3}	4.37×10^{-3}
log-Gaussian	2.08×10^{-2}	4.42×10^{-3}	3.59×10^{-3}
GEV	1.07×10^{-3}	8.33×10^{-4}	9.83×10^{-4}
Pearson	8.38×10^{-4}	1.15×10^{-4}	9.35×10^{-5}

Table 9.3: Average JS-divergence ($m = 2$).

all three values of N, we observe that the Pearson approximation has the lowest JS-divergence, followed by the GEV approximation. The largest JS-divergence is achieved for the Gaussian and log-Gaussian approximations. When $N = 8$ we note that both the gamma and lognormal approximation are a better choice than the beta approximation. For $N = 64$ or $N = 128$ the JS-divergence of the proposed beta approximation is shown to be lower than for the gamma and lognormal approximation, which indicates the beta approximation becomes more appropriate when the number of subcarriers increases.

In table 9.4 we show the average JS-divergence for a varying number of bits per subcarrier $m_k = m$, $k = 1, \ldots, N$. We compare a 4-QAM, 16-QAM and 64-QAM constellation where E_s/σ_w^2 is equal to 10, 15 and 20 dB, respectively. We take a higher E_s/σ_w^2 for the higher order constellations, in order to reflect that typically the operating SNR increases with the constellation size. The JS-divergences for the different approximations show the same trends as before. Only for 64-QAM we notice some differences, i.e., the beta approximation performs slightly better than the GEV approximation, and the gamma distribution has a slightly better accuracy than the beta distribution; however, the differences in JS-divergence are almost negligible.

model	4-QAM	16-QAM	64-QAM
beta	1.41×10^{-3}	8.46×10^{-4}	6.52×10^{-4}
gamma	1.76×10^{-3}	9.19×10^{-4}	6.31×10^{-4}
lognormal	1.66×10^{-3}	1.31×10^{-3}	1.09×10^{-3}
Gaussian	4.42×10^{-3}	3.28×10^{-3}	2.82×10^{-3}
log-Gaussian	4.42×10^{-3}	2.30×10^{-3}	1.50×10^{-3}
GEV	8.33×10^{-4}	7.36×10^{-4}	7.39×10^{-4}
Pearson	1.15×10^{-4}	4.25×10^{-5}	2.79×10^{-5}

Table 9.4: Average JS-divergence for $m_k = m = 2,\ 4$ and 6, $k = 1, \ldots, N$, with corresponding $E_{\mathrm{s}}/\sigma_w^2$=10, 15 and 20 dB ($N = 64$).

9.4 Chapter Summary

This chapter proposed a new statistical model to approximate the distribution of the effective SNR resulting from κESM when only outdated and imperfect CSI is available at the TX. The proposed beta approximation is shown to have an accuracy which is slightly less than the more involved GEV and Pearson distributions, but is more accurate than other often used approximations like the gamma, Gaussian, log-Gaussian and lognormal distributions. When the number of subcarriers is sufficiently large, the numerical results have shown that this approximation has a high accuracy, independent of the chosen constellation size. This accurate model provides the TX with a simple statistical description of the effective SNR, which will be used in the following chapter to adapt the transmission parameters as a function of the outdated and imperfect channel information.

9.A Derivation of the Mean and Variance of Y

First we calculate the mean of Y as follows

$$\begin{aligned}\mathbb{E}\left[Y\right] &= \mathbb{E}\left[\sum_{k\in\mathcal{N}}\sum_{\mu=1}^{\frac{\sqrt{2^{m_k}}}{2}}\alpha_{k,\mu}e^{-\frac{E_k\left|H_{\mathrm{p},k}+(\mathbf{Bs})_k\right|^2\left(\mu d_{k,\min}\right)^2}{4\beta\sigma_w^2}}\right]\\ &= \sum_{k\in\mathcal{N}}\sum_{\mu=1}^{\frac{\sqrt{2^{m_k}}}{2}}\alpha_{k,\mu}\mathbb{E}\left[e^{-\frac{E_k\chi_k\left(\mu d_{k,\min}\right)^2}{4\beta\sigma_w^2}}\right],\end{aligned} \tag{9.12}$$

where $\chi_k = \left|H_{\mathrm{p},k} + (\mathbf{Bs})_k\right|^2$ is a random variable which is distributed according to a scaled noncentral χ^2 distribution. The moment-generating function of χ_k is

$$\mathbb{E}\left[e^{t\chi_k}\right] = \frac{e^{\frac{t\left|H_{\mathrm{p},k}\right|^2}{1-t(\mathbf{BB}^H)_{k,k}}}}{1-t(\mathbf{BB}^H)_{k,k}}, \tag{9.13}$$

where $(\mathbf{BB}^H)_{k,k}$ denotes the element at the kth row and the kth column of matrix $\mathbf{BB}^H$. If we substitute $t = -\frac{E_k(\mu d_{k,\min})^2}{4\beta\sigma_w^2}$ we obtain

$$\mathbb{E}\left[Y\right] = \sum_{k\in\mathcal{N}}\sum_{\mu=1}^{\frac{\sqrt{2^{m_k}}}{2}}\alpha_{k,\mu}\frac{e^{\frac{-\left|H_{\mathrm{p},k}\right|^2\frac{E_k}{4\beta\sigma_w^2}\left(\mu d_{k,\min}\right)^2}{1+\frac{E_k}{4\beta\sigma_w^2}\left(\mu d_{k,\min}\right)^2(\mathbf{BB}^H)_{k,k}}}}{1+\frac{E_k}{4\beta\sigma_w^2}\left(\mu d_{k,\min}\right)^2(\mathbf{BB}^H)_{k,k}}. \tag{9.14}$$

For the variance we get

$$\mathrm{Var}\left[Y\right] = \mathbb{E}\left[Y^2\right] - \mathbb{E}\left[Y\right]^2, \tag{9.15}$$

so we still have to calculate $\mathbb{E}\left[Y^2\right]$,which gives

$$\begin{aligned}&\mathbb{E}\left[Y^2\right]\\ &= \mathbb{E}\left[\sum_{k,l\in\mathcal{N}}\sum_{\mu=1}^{\frac{\sqrt{2^{m_k}}}{2}}\sum_{\lambda=1}^{\frac{\sqrt{2^{m_l}}}{2}}\alpha_{k,\mu}\alpha_{l,\lambda}e^{-\frac{\gamma_k\left(\mu d_{k,\min}\right)^2}{4\beta}}e^{-\frac{\gamma_l\left(\lambda d_{l,\min}\right)^2}{4\beta}}\right]\\ &= \mathbb{E}\Bigg[\sum_{k,l\in\mathcal{N},k\neq l}\sum_{\mu=1}^{\frac{\sqrt{2^{m_k}}}{2}}\sum_{\lambda=1}^{\frac{\sqrt{2^{m_l}}}{2}}\alpha_{k,\mu}\alpha_{l,\lambda}e^{-\frac{\gamma_k\left(\mu d_{k,\min}\right)^2}{4\beta}}e^{-\frac{\gamma_l\left(\lambda d_{l,\min}\right)^2}{4\beta}}\\ &+ \sum_{k\in\mathcal{N}}\sum_{\mu=1}^{\frac{\sqrt{2^{m_k}}}{2}}\sum_{\lambda=1}^{\frac{\sqrt{2^{m_l}}}{2}}\alpha_{k,\mu}\alpha_{k,\lambda}e^{-\frac{\gamma_k(\mu^2+\lambda^2)d_{k,\min}^2}{4\beta}}\Bigg],\end{aligned} \tag{9.16}$$

where γ_k was defined in (9.2). The second term of this expression can be found by substituting

$$t = -\frac{E_k(\mu^2+\lambda^2)d_{k,\min}^2}{4\beta\sigma_w^2}, \tag{9.17}$$

in (9.13), which gives

$$\begin{aligned} g(\mu,\lambda,k) &= \mathbb{E}\left[e^{-\frac{\gamma_k(\mu^2+\lambda^2)d_{k,\min}^2}{4\beta}}\right] \\ &= \frac{e^{\frac{-\left|H_{\mathrm{p},k}\right|^2\frac{E_k}{4\beta\sigma_w^2}(\mu^2+\lambda^2)d_{k,\min}^2}{1+\frac{E_k}{4\beta\sigma_w^2}(\mu^2+\lambda^2)d_{k,\min}^2(\mathbf{B}\mathbf{B}^H)_{k,k}}}}{1+\frac{E_k}{4\beta\sigma_w^2}(\mu^2+\lambda^2)d_{k,\min}^2(\mathbf{B}\mathbf{B}^H)_{k,k}}. \end{aligned} \tag{9.18}$$

The first term of equation (9.16) can be validated with the aid of the following function

$$f(\mathbf{t}) = \mathbb{E}\left[e^{\sum t_k|H_{\mathrm{p},k}+(\mathbf{B}\mathbf{s})_k|^2}\right], \tag{9.19}$$

where $\mathbf{t} \triangleq [t_1,\ldots,t_N]^T$. If we define

$$\mathbf{A} \triangleq \mathrm{diag}\left(t_1,\ldots,t_N\right), \tag{9.20}$$

the function $f(\mathbf{t})$ can be written as

$$f(t_1,\ldots,t_N) = \int \frac{1}{\pi^q} e^{-\mathbf{s}^H\mathbf{s}} e^{(\mathbf{H}_\mathrm{p}+\mathbf{B}\mathbf{s})^H\mathbf{A}(\mathbf{H}_\mathrm{p}+\mathbf{B}\mathbf{s})}\mathrm{d}\,\mathbf{s}. \tag{9.21}$$

After the introduction of the following vector

$$\mathbf{c} \triangleq (\mathbf{I}_q - \mathbf{B}^H\mathbf{A}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{A}\mathbf{H}_\mathrm{p}, \tag{9.22}$$

this can be evaluated as follows

$$\begin{aligned} f(\mathbf{t}) &= \int \frac{1}{\pi^q} e^{-(\mathbf{s}-\mathbf{c})^H(\mathbf{I}_q-\mathbf{B}^H\mathbf{A}\mathbf{B})(\mathbf{s}-\mathbf{c})} \\ &\quad \cdot e^{\mathbf{H}_\mathrm{p}^H\mathbf{A}\mathbf{H}_\mathrm{p}+\mathbf{c}^H(\mathbf{I}_q-\mathbf{B}^H\mathbf{A}\mathbf{B})\mathbf{c}}\mathrm{d}\,\mathbf{s} \\ &= |\mathbf{I}_q - \mathbf{B}^H\mathbf{A}\mathbf{B}|^{-1} e^{\mathbf{H}_\mathrm{p}^H\mathbf{A}\mathbf{H}_\mathrm{p}+\mathbf{c}^H(\mathbf{I}_q-\mathbf{B}^H\mathbf{A}\mathbf{B})\mathbf{c}} \\ &\quad \cdot \int \frac{1}{\pi^q|\mathbf{I}_q - \mathbf{B}^H\mathbf{A}\mathbf{B}|^{-1}} e^{-(\mathbf{s}-\mathbf{c})^H(\mathbf{I}_q-\mathbf{B}^H\mathbf{A}\mathbf{B})(\mathbf{s}-\mathbf{c})}\mathrm{d}\,\mathbf{s} \\ &= |\mathbf{I}_q - \mathbf{B}^H\mathbf{A}\mathbf{B}|^{-1} e^{\mathbf{H}_\mathrm{p}^H\mathbf{A}\mathbf{H}_\mathrm{p}+\mathbf{H}_\mathrm{p}^H\mathbf{A}\mathbf{B}(\mathbf{I}_q-\mathbf{B}^H\mathbf{A}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{A}\mathbf{H}_\mathrm{p}}, \end{aligned} \tag{9.23}$$

where in the last step we have used the fact that the integral of the complex Gaussian probability density is equal to 1. To calculate the first term of equation (9.16), we have to evaluate $f(\mathbf{t})$ when

$$\mathbf{t} = -\frac{E_k}{4\beta\sigma_w^2}\left(\mu d_{k,\min}\right)^2\mathbf{e}_k - \frac{E_l}{4\beta\sigma_w^2}\left(\lambda d_{l,\min}\right)^2\mathbf{e}_l, \tag{9.24}$$

where $\mathbf{e}_k$ and $\mathbf{e}_l$ denote the kth and lth column of $\mathbf{I}_N$, respectively. Using (9.15), (9.16), (9.18), (9.23) and (9.24), we get

$$\begin{aligned}\mathrm{Var}\left[Y\right] = & \sum_{k,l\in\mathcal{N},k\neq l} \sum_{\mu=1}^{\frac{\sqrt{2^{m_k}}}{2}} \sum_{\lambda=1}^{\frac{\sqrt{2^{m_l}}}{2}} \alpha_{k,\mu}\alpha_{l,\lambda} \\ & \cdot f\left(-\frac{E_k}{4\beta\sigma_w^2}\left(\mu d_{k,\min}\right)^2 \mathbf{e}_k - \frac{E_l}{4\beta\sigma_w^2}\left(\lambda d_{l,\min}\right)^2 \mathbf{e}_l\right) \\ & + \sum_{k\in\mathcal{N}} \sum_{\mu=1}^{\frac{\sqrt{2^{m_k}}}{2}} \sum_{\lambda=1}^{\frac{\sqrt{2^{m_l}}}{2}} \alpha_{k,\mu}\alpha_{k,\lambda} g(\mu,\lambda,k) - \mathbb{E}\left[Y\right]^2 . \end{aligned} \tag{9.25}$$

10

Adaptive Coding and Modulation using Imperfect Channel State Information in BIC-OFDM systems

In this chapter (which is based on contributions [103] and [104]), we investigate resource allocation algorithms under the realistic assumption that the available channel state information (CSI) at the transmitter (TX) is imperfect due to estimation errors and/or feedback delays. First, we introduce an optimal performance metric for the secondary user BIC-OFDM system, called the expected goodput (EGP). By using the accurate approximation derived in chapter 9, we succeed in deriving a tractable and very accurate approximation for the EGP. This approximate EGP (AEGP) is then used for the derivation of several resource allocation algorithms which optimize the code rate, bit and energy allocation under a constraint on the interference caused to the PU network. In the numerical results, we show that the AEGP is far more accurate than previous attempts to model the GP in the presence of imperfect CSI. Further, we verify that, in spite of the imperfect nature of the available CSI, the derived resource allocation algorithms significantly increase the goodput of the SU network, compared to a non-adaptive selection of the transmission parameters.

In section 10.1 we describe the cognitive BIC-OFDM system. In section 10.2, we introduce the EGP metric, and discuss the statistical approximation of the

κESM. The resource allocation algorithms which select the code rate and the energy and bit allocations per subcarrier are derived in section 10.3. The accuracy of the EGP metric and the performance of the resource allocation algorithms are validated in section 10.4. This chapter is concluded in section 10.5.

10.1 Cognitive BIC-OFDM System Model

We consider a SU network, which consists of a point-to-point OFDM link, that occupies the same bandwidth as a PU network containing N_{PU} PU-RXs. Messages are transmitted by means of a packet-oriented BIC-OFDM communication system as described in chapter 4. We assume that the transmit energies are constrained by

$$\sum_{k=1}^{N} E_k \leq E^{(\max)}, \tag{10.1}$$

where $E^{(\max)}$ is the maximal transmit energy per OFDM symbol. We again define the received SNR associated with the *k*th subcarrier at the FFT output as

$$\gamma_k \triangleq \frac{E_k |H_k|^2}{\sigma_w^2}, \tag{10.2}$$

and arrange the received SNRs into a vector $\mathbf{SNR} \triangleq [\gamma_1, \ldots, \gamma_N]$ for further use. Further, as not all N available subcarriers will necessarily be used for the transmission, we make a distinction between the set $\{1, ..., N\}$ of available subcarriers, and the set $\mathcal{N} \subseteq \{1, ..., N\}$ of active subcarriers. When the *k*th subcarrier is not active (i.e., $k \notin \mathcal{N}$), both the transmit energy E_k and the number of bits m_k on that subcarrier are set to 0.

As in the previous chapters, the available CSI at the TX is assumed to be imperfect. To make the description of our proposed approach quite general, we will denote the CSI, which is available at the TX about the actual channel realization $\mathbf{H} \triangleq [H_1, \ldots, H_N]^T$, by the vector $\mathbf{ICSI}$. We make the assumption that $\mathbf{H}$ and $\mathbf{ICSI}$ are jointly zero-mean circularly symmetric Gaussian. It then follows that $\mathbf{H}$ conditioned on $\mathbf{ICSI}$ is Gaussian, with expectation $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{CSI}} = \mathbb{E}_{\mathbf{H}}[\mathbf{H}|\mathbf{ICSI}]$ and covariance matrix $\mathbf{C}_{\mathbf{H}|\mathbf{ICSI}} = \mathbb{E}_{\mathbf{H}}\left[\mathbf{H}\mathbf{H}^H|\mathbf{ICSI}\right] - \boldsymbol{\mu}_{\mathbf{H}|\mathbf{ICSI}}\boldsymbol{\mu}_{\mathbf{H}|\mathbf{ICSI}}^H$; note that $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{ICSI}}$ corresponds to the MMSE estimate (3.25) of $\mathbf{H}$ based on $\mathbf{ICSI}$. Some examples of $\mathbf{ICSI}$ and the associated statistics are given in appendix 10.A.

Denoting by $G_k^{(q)}$ the channel gain from the SU-TX to the *q*th PU-RX, experienced by the *k*th subcarrier, the interference constraints can be expressed as $\sum_{k\in\mathcal{N}} E_k |G_k^{(q)}|^2 \leq \Gamma_q$ for $q \in \mathcal{Q} \triangleq \{1, \ldots, N_{\mathrm{PU}}\}$. Note that this is an approximation compared to the interference constraints used in section 8.1.3.3. However, this approximation is quite accurate for most subcarriers (more specifically, the

inner subcarriers), as the delay spread of the channel gain from the SU-TX to the qth PU-RX is often much smaller than the length of the OFDM symbol T_{s}.

We denote by $\mathbf{ICSI}_{\mathrm{PU}} = \{\mathbf{ICSI}_{\mathrm{PU}}^{(q)}, q \in \mathcal{Q}\}$ the imperfect CSI available at the SU-TX about its channels to the PU-RXs. As only $\mathbf{ICSI}_{\mathrm{PU}}$ and not the exact channel gains $G_k^{(q)}$ are available at the SU-TX, it can happen that the interference constraint at the PU-RXs is violated. Therefore, the interference constraints can be satisfied only on average, i.e. for given $\mathbf{ICSI}_{\mathrm{PU}}^{(q)}$, the constraint is replaced by

$$\mathbb{E}_{\mathbf{G}^{(q)}}\left[\sum_{k\in\mathcal{N}} E_k |G_k^{(q)}|^2 \,|\mathbf{ICSI}_{\mathrm{PU}}^{(q)}\right] \leq \Gamma_q, \quad q \in \mathcal{Q}, \tag{10.3}$$

where $\mathbf{G}^{(q)} \triangleq [G_1^{(q)}, \ldots, G_N^{(q)}]^T$. The expected value in (10.3) can be expressed as

$$\begin{aligned}\mathbb{E}_{\mathbf{G}^{(q)}}\left[\sum_{k\in\mathcal{N}} E_k |G_k^{(q)}|^2 \,|\mathbf{ICSI}_{\mathrm{PU}}^{(q)}\right] = \sum_{k\in\mathcal{N}} E_k (&|(\boldsymbol{\mu}^{(q)}_{\mathbf{G}|\mathbf{ICSI}_{\mathrm{PU}}})_k|^2 \\ &+ (\mathbf{C}^{(q)}_{\mathbf{G}|\mathbf{ICSI}_{\mathrm{PU}}})_{k,k}), \quad q \in \mathcal{Q},\end{aligned} \tag{10.4}$$

where we have assumed that the distribution of $\mathbf{G}^{(q)}$ conditioned on $\mathbf{ICSI}_{\mathrm{PU}}^{(q)}$ is Gaussian with mean $\boldsymbol{\mu}^{(q)}_{\mathbf{G}|\mathbf{ICSI}_{\mathrm{PU}}}$ and covariance matrix $\mathbf{C}^{(q)}_{\mathbf{G}|\mathbf{ICSI}_{\mathrm{PU}}}$.

10.2 Goodput Performance Metric

The secondary user will try to optimize the goodput (GP), which was defined in (4.13) as

$$\mathrm{GP}(\phi, \mathbf{SNR}) = \frac{N_{\mathrm{p}} r}{N_{\mathrm{tot}} N}\left(\sum_{k\in\mathcal{N}} m_k\right) \cdot (1 - \mathrm{PER}(\phi, \mathbf{SNR})). \tag{10.5}$$

As a performance measure of the SU network we consider the long-term average of the goodput $\mathrm{GP}_{\mathrm{avg}}$ (4.14) over many channel realizations.

If perfect CSI were available at the TX (i.e., the TX knows the realizations of its channels to the SU-RX and PU-RXs), the optimal way of selecting the transmission mode ϕ and the energy allocation vector $\mathbf{E} \triangleq [E_1, \ldots, E_N]^T$ as a function of these realizations is to maximize (10.5) under the constraints on the SU transmit energy and the interference at the PU-RXs, for the given realizations $\mathbf{H}$ and $\{\mathbf{G}^{(q)}, q \in \mathcal{Q}\}$. This selection obviously maximizes the long-term average goodput $\mathrm{GP}_{\mathrm{avg}}$ of the system.

However, when only imperfect CSI is available, the transmission parameters $(\phi, \mathbf{E})$ must be selected as functions of $\mathbf{ICSI}$ and $\mathbf{ICSI}_{\mathrm{PU}}$, rather than $\mathbf{H}$ and $\{\mathbf{G}^{(q)}, q \in \mathcal{Q}\}$. Taking into account that for given ϕ and $\mathbf{E}$, GP from (10.5) is

a function of $\mathbf{H}$, and that the joint probability density function of $\mathbf{H}$, $\mathbf{ICSI}$ and $\mathbf{ICSI}_{\mathrm{PU}}$ can be factored as

$$p(\mathbf{H}, \mathbf{ICSI}, \mathbf{ICSI}_{\mathrm{PU}}) = p(\mathbf{H}|\mathbf{ICSI})p(\mathbf{ICSI})p(\mathbf{ICSI}_{\mathrm{PU}}), \tag{10.6}$$

the long-term average goodput can be written as

$$\begin{aligned} \mathrm{GP}_{\mathrm{avg}} = \mathbb{E}_{\mathbf{ICSI},\mathbf{ICSI}_{\mathrm{PU}}} \Bigg[& \frac{N_{\mathrm{p}} r}{N_{\mathrm{tot}} N} \left(\sum_{k \in \mathcal{N}} m_k \right) \\ & \cdot \left(1 - \mathbb{E}_{\mathbf{H}} \left[\mathrm{PER}(\phi, \mathbf{SNR}) | \mathbf{ICSI} \right] \right) \Bigg]. \end{aligned} \tag{10.7}$$

It follows from equation (10.7) that $\mathrm{GP}_{\mathrm{avg}}$ becomes maximum when for given $(\mathbf{ICSI}, \mathbf{ICSI}_{\mathrm{PU}})$ the transmission parameters $(\phi, \mathbf{E})$ maximize the expression between brackets in (10.7), under the constraints (10.1) and (10.3). This is equivalent to maximizing the expected goodput (EGP) metric, given by

$$\begin{aligned} \mathrm{EGP} &= \mathbb{E}_{\mathbf{H}} \left[\mathrm{GP}(\phi, \mathbf{SNR}) | \mathbf{ICSI} \right] \\ &= \frac{N_{\mathrm{p}} r}{N_{\mathrm{tot}} N} \left(\sum_{k \in \mathcal{N}} m_k \right) \cdot \left(1 - \mathbb{E}_{\mathbf{H}} \left[\mathrm{PER}(\phi, \mathbf{SNR}) | \mathbf{ICSI} \right] \right), \end{aligned} \tag{10.8}$$

which is the conditional expectation of GP for given $\mathbf{ICSI}$ and represents the optimal performance metric in terms of $\mathrm{GP}_{\mathrm{avg}}$ when only imperfect CSI is available at the TX.

The EGP from (10.8) will be approximated by replacing $\mathrm{PER}(\phi, \mathbf{SNR})$ by $\mathrm{PER}_{\mathrm{AWGN}}(r, \gamma_{\mathrm{eff}})$, with γ_{eff} given by (4.25). The reference curves $\mathrm{PER}_{\mathrm{AWGN}}(r, \gamma_{\mathrm{eff}})$ can be stored in a lookup table for each code rate r from the set $\mathcal{D}_r$. In order to compute the conditional expectation $\mathbb{E}_{\mathbf{H}} \left[\mathrm{PER}_{\mathrm{AWGN}}(r, \gamma_{\mathrm{eff}}) | \mathbf{ICSI} \right]$, we will use the approximation of $Y \triangleq e^{-\gamma_{\mathrm{eff}}/\beta}$ which we proposed in chapter 9. This means that Y conditioned on $\mathbf{ICSI}$ will be approximated by a random variable Z which follows a beta distribution with shaping parameters a and b, which are given by (9.7) and (9.8), respectively. Note that the distribution of Z depends on the selected bit allocation through the variables $\psi^k(\mu)$, m_k and $d_{k,\min}$. Using this approximating beta distribution, we obtain the approximate EGP (AEGP) given by

$$\mathrm{AEGP} = \frac{N_{\mathrm{p}} r}{N_{\mathrm{tot}} N} \left(\sum_{k \in \mathcal{N}} m_k \right) \cdot \left(1 - \mathbb{E}_Z \left[\mathrm{PER}_{\mathrm{AWGN}}(r, -\beta \ln(Z)) \right] \right). \tag{10.9}$$

The expectation w.r.t. Z in (10.9) can be approximated by means of numerical integration.

10.3 Goodput Optimization

In this section we consider different algorithms the TX can employ to optimize the code rate r, the energy allocation E_k and the bit allocation m_k ($\forall k \in \mathcal{N}$) such that the AEGP from (10.9) is maximized, while satisfying the transmit energy constraint (10.1) and the interference constraints (10.3) at the PU-RXs. These algorithms assume that only imperfect CSI is available at the TX.

10.3.1 Uniform Energy and Bit Allocation

In this first subsection, we make the restriction that the bit and energy allocation are uniform, and that all N available subcarriers are actually used, i.e., $\mathcal{N} = \{1, ..., N\}$. For the bit and energy allocation this means that

$$\forall k \in \mathcal{N} : \; m_k = m, \, E_k = E, \tag{10.10}$$

where $m \in \mathcal{D}_m$. Considering the constraints (10.1) and (10.3), the optimal uniform energy per subcarrier is given by

$$E = \min \left(\min_{q \in \mathcal{Q}} \frac{\Gamma_q}{\mathbb{E}_{\mathbf{G}^{(q)}} \left[\sum_{k \in \mathcal{N}} |G_k^{(q)}|^2 \, | \mathbf{ICSI}_{\mathrm{PU}}^{(q)} \right]}, \frac{E^{(\max)}}{|\mathcal{N}|} \right), \tag{10.11}$$

where the expected value can be found from (10.4) and $|\mathcal{N}|$ denotes the number of active subcarriers. The TX will calculate the AEGP (10.9) for every TM $\phi = \{m, r\}$, and then selects the TM $\phi = \{m, r\}$ which yields the largest AEGP. The pseudo-code of this optimization is outlined in table 10.1.

10.3.2 Optimized Energy and Uniform Bit Allocation

In this subsection, we will adapt the previous algorithm such that the TX optimizes the energy per subcarrier, while the bit allocation remains uniform. As explained further, we will allow some of the subcarriers to be inactive, i.e., $\mathcal{N} \subseteq \{1, 2, ..., N\}$. We first have a closer look at the EGP from (10.8) where $\mathrm{PER}(\phi, \mathbf{SNR})$ is replaced by $\mathrm{PER}_{\mathrm{AWGN}}(r, \gamma_{\mathrm{eff}})$, i.e.,

$$\begin{aligned} \mathrm{EGP} \approx & \frac{N_{\mathrm{p}} r}{N_{\mathrm{tot}} N} \left(\sum_{k \in \mathcal{N}} m_k \right) \\ & \cdot \Big(1 - \mathbb{E}_{\mathbf{H}} \left[\mathrm{PER}_{\mathrm{AWGN}}(r, -\beta \ln(Y(\mathbf{E}))) | \mathbf{ICSI} \right] \Big), \end{aligned} \tag{10.12}$$

where we have explicitly shown the dependence on the energy allocation vector $\mathbf{E}$. Because the PER is a convoluted function of the individual subcarrier energies, an exact optimization of this metric will be very hard to obtain. Therefore

Optimization of E, $\mathbf{m}$ and r

Set $\mathrm{AEGP}_{\mathrm{opt}} = 0$
Set $E = \min\left(\min_{q\in\mathcal{Q}} \frac{\Gamma_q}{\sum_{l\in\mathcal{N}}(|(\boldsymbol{\mu}^{(q)}_{\mathbf{G}|\mathbf{ICSI}})_l|^2+(\mathbf{C}^{(q)}_{\mathbf{G}|\mathbf{ICSI}})_{l,l})}, \frac{E^{(\max)}}{|\mathcal{N}|}\right)$
For $m \in \mathcal{D}_m$
  Set $m_k = m, \quad \forall k \in \mathcal{N}$
  For $r \in \mathcal{D}_r$
    Set AEGP according to (10.9)
    If $\mathrm{AEGP} \geq \mathrm{AEGP}_{\mathrm{opt}}$ **Then**
      Set $\mathrm{AEGP}_{\mathrm{opt}} = \mathrm{AEGP}$
      Set $r_{\mathrm{opt}} = r$
      Set $m_{\mathrm{opt}} = m$
    End If
  End For
End For

Table 10.1: Uniform energy and bit allocation.

we suggest a more computationally efficient method, by optimizing the following simplification of the EGP

$$\begin{aligned}\mathrm{EGP} \approx \frac{N_\mathrm{p} r}{N_\mathrm{tot} N} \left(\sum_{k \in \mathcal{N}} m_k \right) \\ \cdot \left(1 - \mathrm{PER}_\mathrm{AWGN}(r, -\beta \ln(\mathbb{E}_\mathbf{H}\left[Y(\mathbf{E})|\mathbf{ICSI}\right]))\right),\end{aligned} \tag{10.13}$$

where the expectation is now taken inside the logarithm. As $\mathrm{PER}_\mathrm{AWGN}(r, \gamma_\mathrm{eff})$ decreases with increasing γ_eff, the maximization of (10.13) w.r.t. $\mathbf{E}$ is equivalent to the minimization of $\mathbb{E}_\mathbf{H}\left[Y(\mathbf{E})|\mathbf{ICSI}\right]$. The latter function can be obtained analytically as shown in (9.14):

$$\mathbb{E}_\mathbf{H}\left[Y(\mathbf{E})|\mathbf{ICSI}\right] = \frac{1}{\sum_{l \in \mathcal{N}} m_l} \sum_{k \in \mathcal{N}} \sum_{n=1}^{\frac{\sqrt{2^{m_k}}}{2}} g_{k,n}(E_k), \tag{10.14}$$

where

$$g_{k,n}(E_k) = \frac{\psi^k(n)}{2^{m_k - 1}} \frac{\exp\left(\frac{-|(\boldsymbol{\mu}_{\mathbf{H}|\mathbf{ICSI}})_k|^2 \frac{E_k}{4\beta\sigma_w^2} n^2 d_{k,\min}^2}{1 + \frac{E_k}{4\beta\sigma_w^2} n^2 d_{k,\min}^2 (\mathbf{C}_{\mathbf{H}|\mathbf{ICSI}})_{k,k}} \right)}{1 + \frac{E_k}{4\beta\sigma_w^2} n^2 d_{k,\min}^2 (\mathbf{C}_{\mathbf{H}|\mathbf{ICSI}})_{k,k}}. \tag{10.15}$$

So the optimized energy allocation that maximizes the simplified EGP in (10.13) is found by solving the following optimization problem

$$\begin{aligned}&\mathbf{E}^{(\mathrm{opt})} = \arg\min_\mathbf{E} \textstyle\sum_{k \in \mathcal{N}} \sum_{n=1}^{\frac{\sqrt{2^{m_k}}}{2}} g_{k,n}(E_k) \\ &\text{s.t.} \quad \textstyle\sum_{k \in \mathcal{N}} E_k \leq E^{(\max)} \\ &\quad (10.3).\end{aligned} \tag{10.16}$$

According to [4], an optimization problem is convex when both the constraints and the objective function are convex. From (10.16), it is clear that the constraints are convex, as they are linear in the components of $\mathbf{E}$. Further, the convexity of the objective function follows from the fact that the second derivative of $g_{k,n}(E_k)$ with respect to E_k can be shown to be non-negative; hence, each term of the objective function is convex, so that the entire objective function is convex as well. Therefore the optimization problem of (10.16) can be efficiently solved by using optimization tools such as CVX [5].

For the optimization of the EGP, we slightly adapt the algorithm outlined in table 10.1. We start by considering all available subcarriers as active, i.e., $\mathcal{N} = \{1, ..., N\}$. For every possible TM $\phi = \{m, r\}$ the algorithm computes the approximation (10.9) of the EGP, using as energy allocation the solution of optimization problem (10.16). Because the energy allocation now depends on the parameter m, it must now become part of the outer loop of the algorithm. For a

given value of m it might happen that for some k the optimized value of E_k equals 0. In this case, the corresponding subcarriers are removed from the active set $\mathcal{N}$ by putting $m_k = 0$, which also removes the large terms with $E_k = 0$ (i.e., $\gamma_k = 0$) from Y for the considered bit allocation. Finally, the algorithm selects the TM and the corresponding energy allocation yielding the largest value of the AEGP (10.9).

10.3.3 Uniform Energy and Greedy Bit Allocation

In this subsection we consider a uniform energy allocation according to (10.11) and an optimized bit allocation per subcarrier.

We first consider the simplified expression for the EGP (10.13):

$$\begin{aligned}\mathrm{EGP} \approx \frac{N_\mathrm{p} r}{N_\mathrm{tot} N} \left(\sum_{k \in \mathcal{N}} m_k \right) \\ \cdot \left(1 - \mathrm{PER}_\mathrm{AWGN}\left(r, -\beta \ln(\mathbb{E}_\mathbf{H}\left[Y(\mathbf{m}, \mathbf{E}) | \mathbf{ICSI}\right])\right)\right),\end{aligned} \tag{10.17}$$

where now the dependence on the bit and energy allocation vectors $\mathbf{m}$ and $\mathbf{E}$ is explicitly shown. Considering the fact that $Y(\mathbf{m}, \mathbf{E})$ is given by

$$Y(\mathbf{m}, \mathbf{E}) = \frac{1}{\sum_{k \in \mathcal{N}} m_k} \sum_{k \in \mathcal{N}} \Omega_k(m_k, E_k), \tag{10.18}$$

where $\Omega_k(m_k, E_k)$ is introduced in equation (4.26), we notice that the simplified EGP from (10.17) only depends on the bit allocation through the quantity $\sum_{k \in \mathcal{N}} \mathbb{E}_\mathbf{H}\left[\Omega_k(m_k, E_k) | \mathbf{ICSI}\right]$ and the sum $\sum_{k \in \mathcal{N}} m_k \triangleq M(\mathbf{m})$. Because the PER is a decreasing function of the effective SNR γ_eff, the maximal value of the simplified EGP, for a fixed value of $M(\mathbf{m})$, will be achieved for the bit allocation $\mathbf{m}$ and energy allocation $\mathbf{E}$ which minimizes

$$\begin{aligned}&\arg\min_{\mathbf{E},\mathbf{m}} \sum_{k \in \mathcal{N}} \mathbb{E}_\mathbf{H}\left[\Omega_k(m_k, E_k) | \mathbf{ICSI}\right] \\ &= \arg\min_{\mathbf{E},\mathbf{m}} \sum_{k \in \mathcal{N}} \sum_{n=1}^{\frac{\sqrt{2^{m_k}}}{2}} g_{k,n}(m_k, E_k),\end{aligned} \tag{10.19}$$

where $g_{k,n}(m_k, E_k)$ is given by (10.15), and the dependence on m_k is shown explicitly. However this represents a mixed integer programming problem, which is computationally very hard. In order to obtain a computationally efficient solution, we base our algorithm on the iterative suboptimal greedy algorithm described in [105].

In the current iteration we modify the bit allocation from the previous iteration by adding 2 bits (because we restrict our attention to square QAM constellations, representing an even number of bits) to the subcarrier which leads to the smallest

increase of $\sum_{k\in\mathcal{N}} \mathbb{E}_{\mathbf{H}}\left[\Omega_k(m_k, E_k)|\mathbf{ICSI}\right]$. For the resulting bit and energy allocation, we determine the code rate r which leads to the highest AEGP (10.9). The iterative algorithm is initialized with $m_k = 0$ for all available subcarriers (yielding $M(\mathbf{m}) = 0$), and continues until all N available subcarriers have $m_{\max}$ bits (yielding $M(\mathbf{m}) = m_{\max}N$), where $m_{\max}$ is the largest allowed number of bits in the constellation. At that point we select the code rate r, the energy and bit allocation which correspond to the value of $M(\mathbf{m})$ for which the AEGP (10.9) is maximal.

Now we outline how the increase of $\sum_{k\in\mathcal{N}} \mathbb{E}_{\mathbf{H}}\left[\Omega_k(m_k, E_k)|\mathbf{ICSI}\right]$ is evaluated. Let us denote by $\mathbf{m}$ the value of the bit allocation vector and by $\mathcal{N}$ the set of active subcarriers, both referring to the previous iteration. We now introduce the quantity $\delta_k^{(m_k+2)}(\mathbf{m})$ which is defined as the increase of (10.19) when the bit allocation on subcarrier k increases from m_k to $m_k + 2$. If subcarrier k was not active in the previous iteration (i.e., $m_k = 0$), the set of active subcarriers increases from $\mathcal{N}$ (previous iteration) to $\mathcal{N} \cup \{k\}$ (current iteration), yielding the increase

$$\delta_k^{(2)}(\mathbf{m}) = \mathbb{E}_{\mathbf{H}}\Big[\Omega_k(2, E_k(\mathbf{m} + 2\mathbf{e}_k)) + \sum_{l\in\mathcal{N}}(\Omega_l(m_l, E_l(\mathbf{m} + 2\mathbf{e}_k)) - \Omega_l(m_l, E_l(\mathbf{m})))|\mathbf{ICSI}\Big], \quad (10.20)$$

where $\mathbf{E}(\mathbf{m})$ and $\mathbf{E}(\mathbf{m}+2\mathbf{e}_k)$ denote the uniform energy allocations from (10.11) corresponding to the bit allocations $\mathbf{m}$ and $\mathbf{m} + 2\mathbf{e}_k$, respectively, related to the previous and the current iteration; because the corresponding set of active subcarriers has changed, $\mathbf{E}(\mathbf{m})$ and $\mathbf{E}(\mathbf{m} + 2\mathbf{e}_k)$ are different, which makes in (10.20) the summation over l non-zero. If subcarrier k was already active in the previous iteration (i.e., $m_k > 0$), we obtain

$$\delta_k^{m_k+2}(\mathbf{m}) = \mathbb{E}_{\mathbf{H}}\Big[\Omega_k(m_k + 2, E_k(\mathbf{m} + 2\mathbf{e}_k)) - \Omega_k(m_k, E_k(\mathbf{m}))|\mathbf{ICSI}\Big]. \quad (10.21)$$

As in this case the set of active subcarriers equals $\mathcal{N}$ for both the previous and the current iteration, the uniform energy allocation from (10.11) satisfies $\mathbf{E}(\mathbf{m} + 2\mathbf{e}_k) = \mathbf{E}(\mathbf{m})$. In the current iteration, the increments $\delta_k^{m_k+2}(\mathbf{m})$ are computed for all $k \in \{1, \ldots, N\}$; then the subcarrier k which yields the lowest $\delta_k^{m_k+2}(\mathbf{m})$ ($k \in \{1, \ldots, N\}$) is selected, and the bit allocation for this subcarrier and $M(\mathbf{m})$ are both increased by 2, compared to the previous iteration.

10.3.4 Suboptimal Joint Energy and Bit Allocation

The greedy bit allocation algorithm introduced in the previous subsection requires the reevaluation of the values of $\delta_k^{m_k+2}(\mathbf{m})$ ($\forall k \in \{1, \ldots, N\}$) each time the set

$\mathcal{N}$ of active subcarriers is modified. The complexity would increase even further if we combined each step of the greedy bit allocation algorithm with the optimized energy allocation introduced in section 10.3.2, which requires solving a convex optimization algorithm instead of a simple evaluation of equation (10.11).

To circumvent this complexity, we present a faster, less computationally intensive algorithm. We initialize the algorithm with the optimal uniform energy and bit allocation from section 10.3.1. Then, as a first step we calculate for this specific uniform bit allocation the optimized energy allocation vector resulting from optimization problem (10.16), for $\mathcal{N} = \{1, ..., N\}$. In the second step, we optimize the bit allocation and code rate according to the greedy algorithm outlined in 10.3.3. Because during this step the energy allocation vector $\mathbf{E}$ is kept to its value resulting from the previous step, we can drop the dependency of $\delta_k^{m_k+2}$ on $\mathbf{m}$, because $\delta_k^{m_k+2}$ now depends only on m_k for given k and, therefore, has to be evaluated only once for each m_k ($m_k \geq 0$,$\forall k \in \{1, \ldots, N\}$). This considerably reduces the complexity. For more details, we refer to the pseudo-code of this algorithm shown in table 10.2. As a final step, the optimized energy allocation vector $\mathbf{E}$ is recalculated according to section 10.3.2, for the optimized TM resulting from the second step. The resulting values for the code rate r, energy allocation $\mathbf{E}$ and bit allocation $\mathbf{m}$ are then used for the transmission.

10.4 Numerical Results

We consider a communication system characterized by the parameters from table 10.3, which uses orthogonal frequency-division multiple access (OFDMA) to support several users. Here we concentrate on the performance of a user to which 48 data subcarriers are allocated, which is equal to one subchannel in the FUSC permutation mode of WiMax [106]. These subcarriers are considered to be evenly spaced across the 384 inner data subcarriers. The channel impulse responses behave according to the ITU vehicular A model [13], with time variations according to Jakes' model [12]. We consider a single PU-RX (so we can drop the index q) and the channels between the different nodes are characterized by $\mathrm{Tr}(\mathbb{E}\left[\mathbf{h}\mathbf{h}^H\right]) = 1$ and $\mathrm{Tr}(\mathbb{E}\left[\mathbf{g}\mathbf{g}^H\right]) = 10^{-3}$, where $\mathbf{h}$ and $\mathbf{g}$ denote the channel impulse responses corresponding to the channel frequency responses $\mathbf{H}$ and $\mathbf{G}$, respectively; this yields $\mathbb{E}\left[|H_k|^2\right] = 1$ and $\mathbb{E}\left[|G_k|^2\right] = 10^{-3}$ for $k \in \{1, ..., N\}$. In this section, we will consider three types of CSI, i.e., estimated CSI, delayed CSI, and estimated and delayed CSI (see appendix 10.A); we always assume that for both $\mathbf{ICSI}$ and $\mathbf{ICSI}_{\mathrm{PU}}$ the same type of CSI is available at the TX. We note however that this is not a requirement for the proper functioning of our proposed algorithms.

The SNR is defined as

$$\mathrm{SNR} \triangleq \frac{E^{(\max)}}{N\sigma_w^2}. \tag{10.22}$$

Optimization of E, $\mathbf{m}$ and r

Set $\mathrm{AEGP}_{\mathrm{opt}} = 0$
Set r and $\mathbf{m}$ according to section 10.3.1
Set E according to (10.16)
For $k \in \{1, \ldots, N\}$
  For $m_k \in \mathcal{D}_m$
    Set $\delta_k^{m_k}$ according to (10.21)
  End For
  Set $\delta_k^{m_{\max}+2} = \infty$
End For
Set $m_k = 0$ $(\forall k \in \{1, \ldots, N\})$
For $M \in \{2, 4, \ldots, m_{\max} N\}$
  Set $k = \arg\min\{\delta_1^{m_1+2}, \ldots, \delta_N^{m_N+2}\}$
  Set $m_k = m_k + 2$
  Update $\mathcal{N}$
  For $r \in \mathcal{D}_r$
    Set AEGP according to (10.9)
    If $\mathrm{AEGP} \geq \mathrm{AEGP}_{\mathrm{opt}}$ **Then**
      Set $\mathrm{AEGP}_{\mathrm{opt}} = \mathrm{AEGP}$
      Set $r_{\mathrm{opt}} = r$
      Set $\mathbf{m}_{\mathrm{opt}} = \mathbf{m}$
    End If
  End For
End For
Set E according to (10.16)

Table 10.2: Suboptimal joint energy and bit allocation.

Parameter	Value
Data subcarriers (N)	48
Sampling rate ($1/T$)	5.6 MHz
FFT size (N_{car})	512
Length of cyclic prefix (ν)	64
Convolutional code	$(171, 133)$
Code rates ($\mathcal{D}_r$)	$1/2,\ 2/3,\ 3/4,\ 5/6$
Constellation sizes ($\mathcal{D}_m$)	$2,\ 4,\ 6$ bits
Information bits (N_{p})	1024
CRC (N_{CRC})	32

Table 10.3: System parameters.

As a performance indicator for the different resource allocation schemes we will display (10.7), which denotes the average of the actual GP w.r.t. the joint probability density function of $\mathbf{H}$, $\mathbf{ICSI}$ and $\mathbf{ICSI}_{\mathrm{PU}}$. This averaging involves the generation of realizations of $\mathbf{ICSI}$ and $\mathbf{ICSI}_{\mathrm{PU}}$, from which the corresponding $(\mathbf{m}, \mathbf{E}, r)$ are computed. For each such realization of $(\mathbf{m}, \mathbf{E}, r)$, we generate realizations of $\mathbf{H}$ according to the conditional distribution $p(\mathbf{H}|\mathbf{ICSI})$. For each such realization of $\mathbf{H}$ we transmit and decode one packet using the transmission parameters $(\mathbf{m}, \mathbf{E}, r)$, and verify whether a decoding error has occurred; averaging the indicator of a decoding error over the realizations of $\mathbf{H}$ yields $\mathbb{E}_{\mathbf{H}}\left[\mathrm{PER}(\phi, \mathbf{SNR})|\mathbf{ICSI}\right]$ corresponding to the considered realization of $(\mathbf{m}, \mathbf{E}, r)$.

10.4.1 Accuracy of AEGP

In this subsection, we investigate how accurately the AEGP metric (10.9) approximates the EGP from (10.8). As a reference we compare the accuracy with the predicted GP (PGP) and the IC-κESM introduced in [103] and [107], respectively. Both approximations are explained in appendix 10.B. The IC-κESM is an approximation that only applies to delayed CSI. For this reason, we will compare the accuracy of these three metrics for the scenario where the TX only has delayed CSI available (see appendix 10.A.2). The following simulation parameters are used: SNR $= 10$ dB, $\Gamma_q/\sigma_w^2 = 0$ dB and the value of $f_{\mathrm{d}}\tau_{\mathrm{d}}$ is equal to 0.05.

We generate 1000 realizations of $\mathbf{ICSI}$ and $\mathbf{ICSI}_{\mathrm{PU}}$ (see appendix 10.A), and for each realization the corresponding optimum uniform bit and energy allocation and code rate are obtained as described in section 10.3.1. Then for each realization of $\mathbf{ICSI}$, $\mathbf{ICSI}_{\mathrm{PU}}$ and the corresponding (m, E, r), we compute (i) the AEGP from (10.9); (ii) the PGP; (iii) the IC-κESM; (iv) the EGP from (10.8), where the average conditioned on $\mathbf{ICSI}$ is replaced by an arithmetical average over 1000 realizations of $\mathbf{H}$, generated according to the conditional distribution $p(\mathbf{H}|\mathbf{ICSI})$

	AEGP	PGP	IC-κESM
$\mathbb{E}[\epsilon]$	$1.87 \cdot 10^{-2}$	$5.48 \cdot 10^{-1}$	$5.57 \cdot 10^{-1}$
$\sqrt{\mathrm{Var}[\epsilon]}$	$2.07 \cdot 10^{-2}$	$2.45 \cdot 10^{-1}$	$1.95 \cdot 10^{-1}$
$\sqrt{\mathbb{E}[\epsilon^2]}$	$2.79 \cdot 10^{-2}$	$6.00 \cdot 10^{-1}$	$5.90 \cdot 10^{-1}$

Table 10.4: Accuracy of the AEGP, PGP and IC-κESM metric (SNR$=10$ dB, $\Gamma_q/\sigma_w^2 = 0$ dB and $f_\mathrm{d}\tau_\mathrm{d} = 0.05$).

(see appendix 10.A), and for each realization of $\mathbf{H}$ it is verified whether the received packet is correctly decoded; and (v) the differences $\epsilon_{\mathrm{AEGP}} = |\mathrm{AEGP} - \mathrm{EGP}|$, $\epsilon_{\mathrm{PGP}} = |\mathrm{PGP} - \mathrm{EGP}|$ and $\epsilon_{\text{IC-}\kappa\mathrm{ESM}} = |\text{IC-}\kappa\mathrm{ESM} - \mathrm{EGP}|$. Table 10.4 shows the average, the standard deviation and the root mean-squared (RMS) value of ϵ_{AEGP}, ϵ_{PGP} and $\epsilon_{\text{IC-}\kappa\mathrm{ESM}}$, resulting from the simulations; these numbers should be compared to the average of EGP over the CSI, which equals 1.42 bits/subcarrier/OFDM symbol. From table 10.4 we observe that the AEGP is a very accurate estimate of the EGP, outperforming both the PGP and the IC-κESM by about one order of magnitude in terms of RMS value. This result validates the accuracy of both the κESM and our approximation of Y by a beta-distributed random variable. The high accuracy of the AEGP metric makes it a very attractive objective function for the optimization of the SU transmission parameters. Further, we also note that being able to accurately describe the expected performance of a link will also have further benefits for more high level algorithms such as scheduling as the probability, of correctly allocating a user to a channel that satisfies its demands, will be increased.

10.4.2 Uniform Energy and Bit Allocation

The performance of the uniform energy and bit allocation algorithm described in section 10.3.1 is investigated. As a reference we will also show the performance in the case of perfect CSI and also for non-adaptive transmission.

In the case of perfect CSI, the optimal uniform energy allocation is given by

$$E = \min\left(\min_{q\in\mathcal{Q}} \frac{\Gamma_q}{\sum_{k\in\mathcal{N}} |G_k^{(q)}|^2}, \frac{E^{(\max)}}{|\mathcal{N}|}\right). \tag{10.23}$$

Using this uniform energy allocation, the GP metric (10.5) is computed for each possible TM $\{m, r\}$, but with $\mathrm{PER}(\phi, \mathbf{SNR})$ replaced by $\mathrm{PER}_{\mathrm{AWGN}}(r, \gamma_{\mathrm{eff}})$. The TM which corresponds to the largest GP is then considered optimal.

In the case of non-adaptive transmission, the TX has no CSI available. This is equivalent to the case where the pdf of the channel gains conditioned on $\mathbf{ICSI}$ reduces to the unconditional pdf of the channel gains. Hence, the uniform energy

allocation is obtained as

$$E = \min \left(\min_{q \in \mathcal{Q}} \frac{\Gamma_q}{\mathbb{E}_{\mathbf{G}^{(q)}} \left[\sum_{k \in \mathcal{N}} |G_k^{(q)}|^2 \right]}, \frac{E^{(\max)}}{|\mathcal{N}|} \right). \tag{10.24}$$

For the above energy allocation, the TX selects, for the current value of SNR (10.22), the TM $\{m, r\}$ which leads to the highest value of $\mathbb{E}_{\mathbf{H}}[\mathrm{GP}]$, with GP given by (10.5).

Now we will apply the algorithm described in section 10.3.1. As a first example, we assume that the TX only has estimated CSI available (see appendix 10.A.1). The variance of the estimation error related to the PU and SU channels is equal to $\sigma_{\mathrm{e}}^2 = 0,\ 10,\ 20$ and 30 dB. For the interference threshold we consider $\Gamma_q/\sigma_w^2 = 0$ dB. The results are shown in Fig. 10.1. We observe that the performance of the SU network clearly depends on the variance of the estimation error σ_{e}^2. For $\sigma_{\mathrm{e}}^2 = 30$ dB there is almost no gain by exploiting CSI compared to a non-adaptive transmission algorithm, because the CSI is unreliable. However when the value of σ_{e}^2 decreases we consistently see a clear gain in performance by exploiting the CSI. When $\sigma_{\mathrm{e}}^2 = 0$ dB, we notice there is a negligible difference between the algorithm using estimated CSI or perfect CSI. Further, we also note that there is almost no gain compared to non-adaptive transmission for small SNR.

In the following example, the TX only has access to delayed CSI (see appendix 10.A.2). The performance of the SU network is shown in Fig. 10.2 for a value of $f_{\mathrm{d}}\tau_{\mathrm{d}}$ equal to 0.01, 0.05, 0.1 and 0.2. It is clear from Fig. 10.2 that when $f_{\mathrm{d}}\tau_{\mathrm{d}}$ is equal to 0.2 there is almost no gain in performance compared to the non-adaptive transmission algorithm, because the channel variations are too fast. However for lower values of $f_{\mathrm{d}}\tau_{\mathrm{d}}$ the GP of the SU network increases considerably. When $f_{\mathrm{d}}\tau_{\mathrm{d}} = 0.01$ the GP almost equals the performance of the algorithm which uses perfect CSI. In Fig. 10.3 we show the difference in performance between optimizing the AEGP, the PGP (as in [103]) or the IC-κESM (as in [107]). We show the performance for $f_{\mathrm{d}}\tau_{\mathrm{d}}$ equal to 0.05 and 0.2. For $f_{\mathrm{d}}\tau_{\mathrm{d}} = 0.05$ we can see a small performance benefit by optimizing the AEGP compared to the less accurate PGP and IC-κESM. When $f_{\mathrm{d}}\tau_{\mathrm{d}} = 0.2$ we notice that the performance improvement we get by using the AEGP or IC-κESM becomes significantly larger compared to using the PGP. In this case, the performance achieved by using the PGP drops even below the performance we would get by using the non-adaptive approach. This demonstrates that the PGP approximation is unable to accurately describe the expected goodput and is thus not suited as an objective function for the optimization problems, especially in the case of fast channel variations. While optimizing the IC-κESM is shown to achieve a similar performance as the optimization of the AEGP, the IC-κESM is far less general than the proposed AEGP as it can only be used in the scenario with delayed CSI described in appendix 10.A.2.

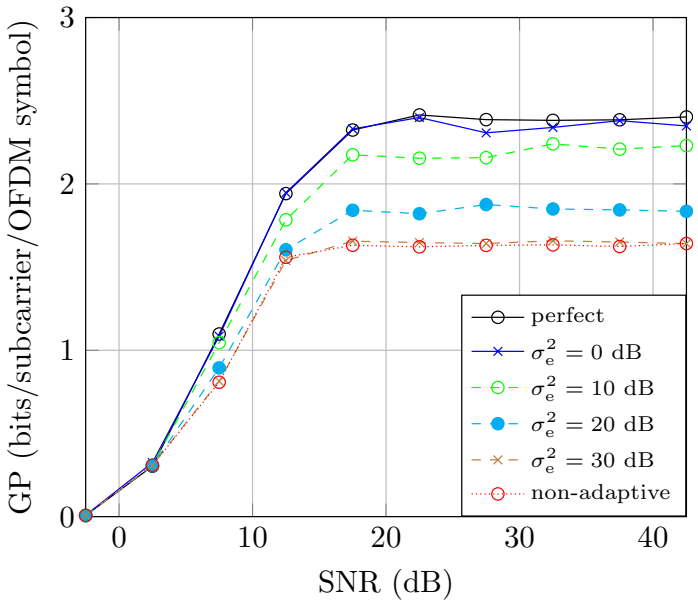


Figure 10.1: GP using estimated CSI ($\sigma_e^2 = 0,\ 10,\ 20$ and 30 dB).

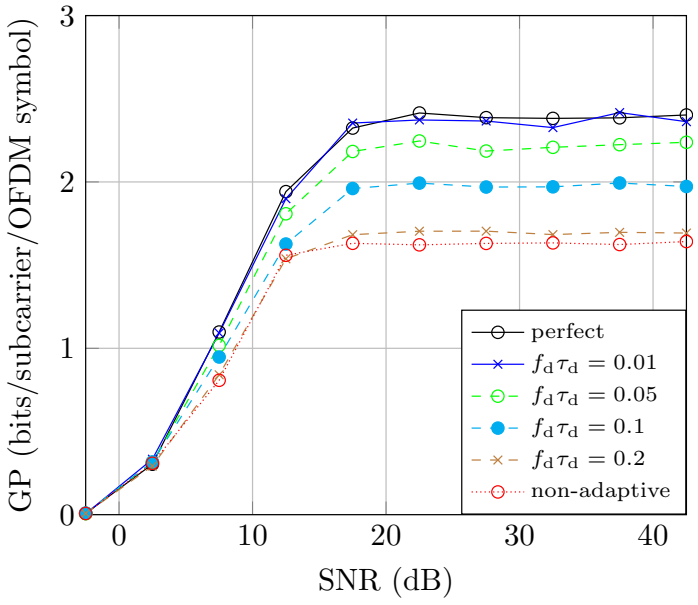


Figure 10.2: GP using delayed CSI ($f_d\tau_d$ =0.01, 0.05, 0.1 and 0.2).

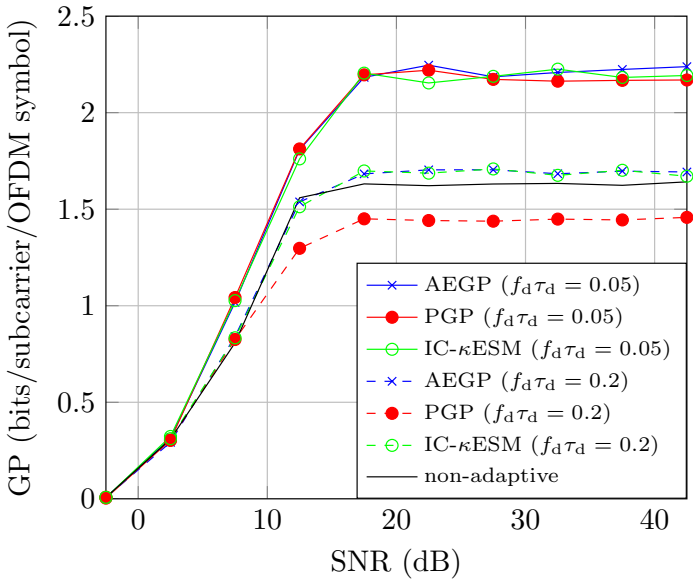


Figure 10.3: Comparison between AEGP, PGP and IC-κESM using delayed CSI ($f_{\mathrm{d}}\tau_{\mathrm{d}}$ =0.05 and 0.2).

In the last example, we combine the delayed CSI with the estimated CSI (see appendix 10.A.3). We choose $f_{\mathrm{d}}\tau_{\mathrm{d}} = 0.2$ and $\sigma_{\mathrm{e}}^2 = 0$ dB. We investigate the performance for a different number (P) of available, delayed channel estimates, with corresponding delays τ_d, $2\tau_d$, ..., $P\tau_d$. The performances are shown in Fig. 10.4 for $P = 1,\ 2,\ 3$ and 4. We observe that the performance of the SU network can be significantly improved when the CSI consists of multiple delayed channel estimates. In this example, the GP increases by about 20 percent when going from $P = 1$ to $P = 4$ for high SNRs. We note that it is not possible to reach the performance of an algorithm with perfect CSI, by increasing the number of estimates. As is clear from Fig. 10.4, there is no noticeable performance gain by going from $P = 3$ to $P = 4$.

In Fig. 10.5 we investigate the impact of the interference threshold. We show the performance of the uniform bit and energy allocation algorithm when $\Gamma_q/\sigma_w^2 = 0,\ 5$ and 10 dB. The resulting goodput is shown for the following simulation variables: $f_{\mathrm{d}}\tau_{\mathrm{d}} = 0.2$, $\sigma_{\mathrm{e}}^2 = 0$ dB and $P = 3$. We observe that the value of the interference threshold has a huge impact on the performance of the SU network. A too conservative value of the interference threshold will severely limit the achievable goodput of the SU network.

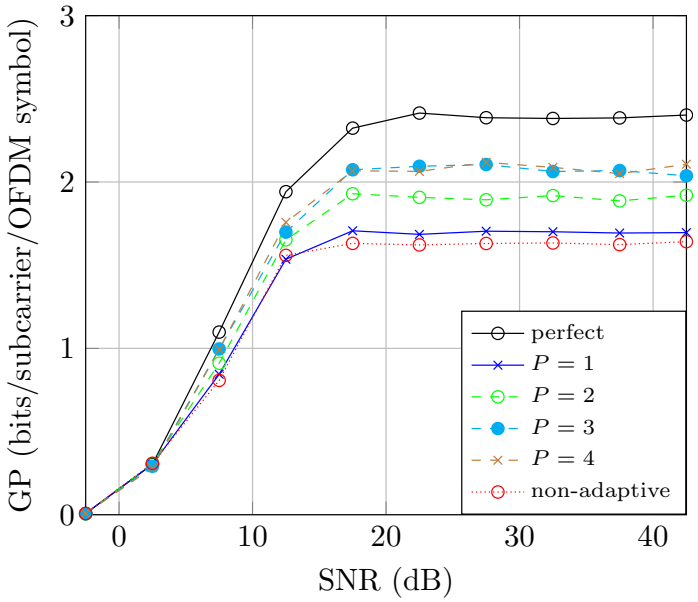


Figure 10.4: GP using estimated and delayed CSI ($\sigma_{\mathrm{e}}^2 = 0$ dB, $f_{\mathrm{d}}\tau_{\mathrm{d}} = 0.2$, $P = 1$, 2, 3 and 4).

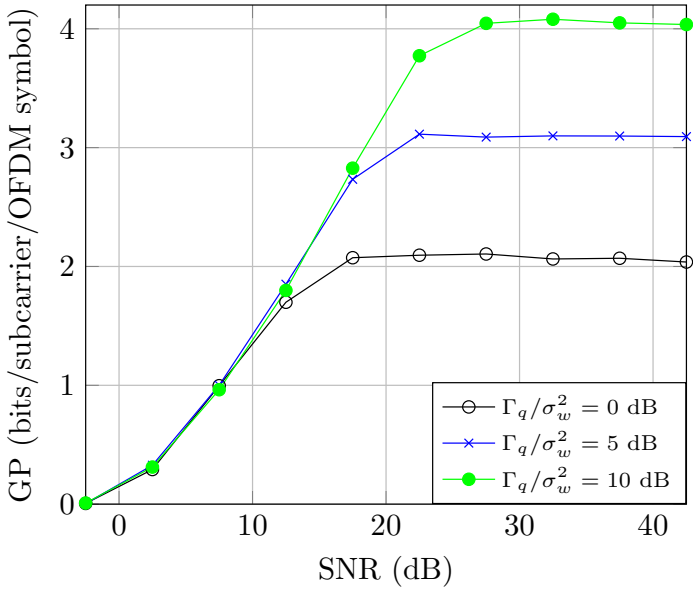


Figure 10.5: GP for different interference thresholds. ($\sigma_{\mathrm{e}}^2 = 0$ dB, $f_{\mathrm{d}}\tau_{\mathrm{d}} = 0.2$, $P = 3$, $\Gamma_q/\sigma_w^2 = 0,\ 5$ and 10 dB).

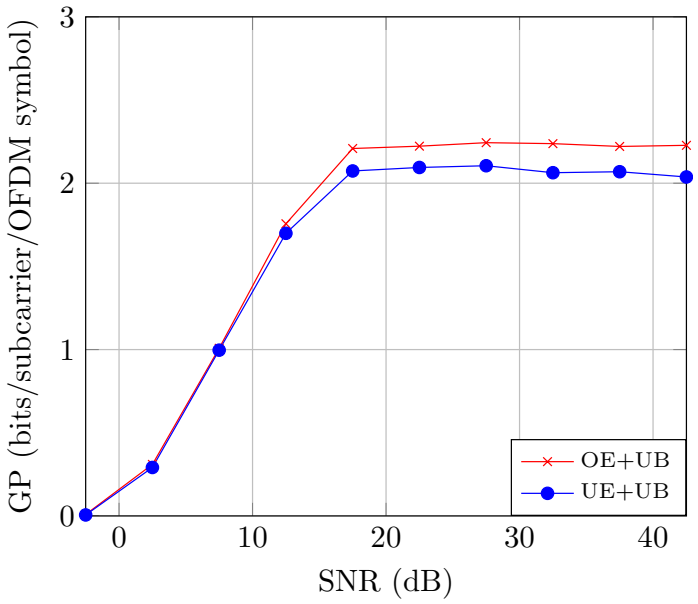


Figure 10.6: GP achieved by optimized and uniform energy allocation ($\sigma_{\mathrm{e}}^2 = 0$ dB, $f_{\mathrm{d}}\tau_{\mathrm{d}} = 0.2$, $P = 3$, $\Gamma_q/\sigma_w^2 = 0$ dB).

10.4.3 Optimized Energy and Uniform Bit Allocation

In this subsection, the optimized energy (OE) allocation from (10.16) and the uniform energy (UE) allocation are compared in terms of goodput. The following simulation parameters are chosen: $\sigma_{\mathrm{e}}^2 = 0$ dB, $f_{\mathrm{d}}\tau_{\mathrm{d}} = 0.2$, $P = 3$ and $\Gamma_q/\sigma_w^2 = 0$ dB. Fig. 10.6 shows the goodput resulting from the uniform energy and bit allocation described in section 10.3.1, along with the goodput corresponding to the OE allocation for the same uniform bit (UB) allocation. We notice that for high SNR the OE allocation improves the goodput by about 8 percent compared to UE allocation.

10.4.4 Greedy Bit Allocation

Now we investigate the performance of the SU network in the case where the SU-TX optimizes the bit allocation per subcarrier. The simulation parameters are chosen as follows: $\sigma_{\mathrm{e}}^2 = 0$ dB, $f_{\mathrm{d}}\tau_{\mathrm{d}} = 0.2$, $P = 3$ and $\Gamma_q/\sigma_w^2 = 0$ dB. We compare the performance of uniform bit and energy allocation (UB+UE), with our algorithm introduced in section 10.3.3 which combines greedy bit allocation with uniform energy allocation (GB+UE). Further, we also consider the performance of the suboptimal algorithm introduced in section 10.3.4 which combines the greedy bit allocation and optimized energy allocation (GB+OE). From Fig. 10.7 we no-

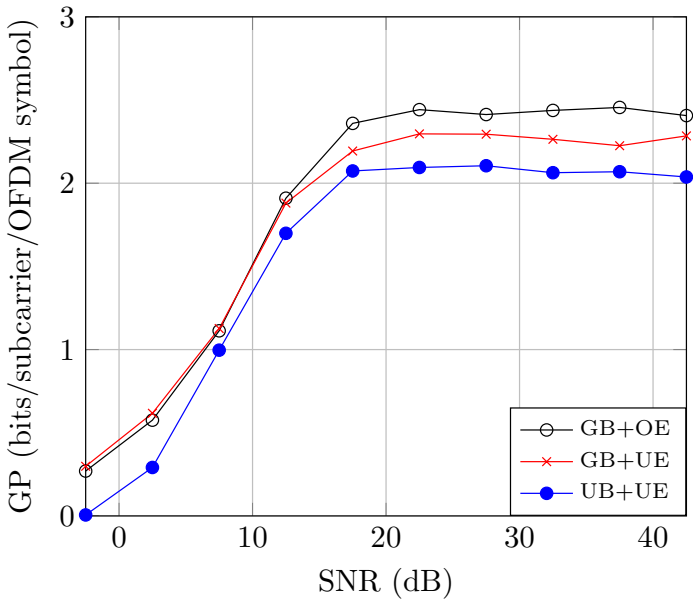


Figure 10.7: Comparison of the goodput achieved by GB and UB allocation ($\sigma_{\mathrm{e}}^2 = 0$ dB, $f_{\mathrm{d}}\tau_{\mathrm{d}} = 0.2$, $P = 3$, $\Gamma_q/\sigma_w^2 = 0$ dB).

tice that there is a considerable increase in GP when we apply GB instead of UB allocation. At low SNR the TX is now capable of deactivating subcarriers with poor instantaneous channel gains, which considerably decreases the PER and improves GP. At higher SNR the TX can now better utilize the full capacity at each subcarrier by allocating a larger number of bits to a subcarrier with favorable channel gains. An even larger gain at higher SNR can be obtained by combining the GB with the OE allocation. In Fig. 10.7, we notice that the gain compared to uniform bit and energy allocation (UB+UE) amounts to 10 percent for greedy bit and uniform energy allocation (GB+UE), and becomes nearly 20 percent for greedy bit and optimized energy allocation (GB+OE). This additional gain is achieved by giving the TX the freedom of reallocating the energy over the subcarriers, which improves the performance in several ways: it can happen for example that subcarriers with less favorable channel gains now receive more energy, or that subcarriers causing strong interference at the PU are switched off to allow for a higher total transmit energy. We do notice however that at lower SNRs the GB+OE algorithm performs slightly worse than the GB+UE algorithm. This is a consequence of our suboptimal approach outlined in section 10.3.4. However the performance loss at low SNR is very small, and an optimal joint bit and energy allocation algorithm would require a much higher complexity.

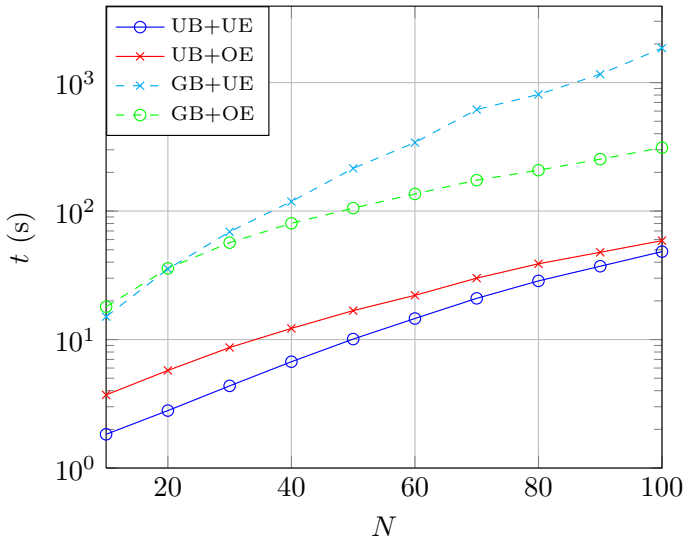


Figure 10.8: Comparison of the simulation time of the different bit and energy allocation algorithms (SNR $= 20$ dB, $\sigma_{\mathrm{e}}^2 = 0$ dB, $f_{\mathrm{d}}\tau_{\mathrm{d}} = 0.2$, $P = 3$, $\Gamma_q/\sigma_w^2 = 0$ dB).

10.4.5 Computational Complexity

To illustrate their complexity, we will compare the average computation times of the different resource allocation algorithms described in section 10.3. The SNR is fixed at 20 dB and the simulation parameters are $\sigma_{\mathrm{e}}^2 = 0$ dB, $f_{\mathrm{d}}\tau_{\mathrm{d}} = 0.2$, $P = 3$ and $\Gamma_q/\sigma_w^2 = 0$ dB. In Fig. 10.8, the computation time of the algorithms is shown as a function of the number of subcarriers N. We notice a slight increase in computation time for the optimized energy allocation (UB+OE) compared to the uniform energy allocation (UB+UE). However, a more significant increase in computation time occurs when implementing the greedy bit allocation. The greedy bit with uniform energy allocation (GB+UE) described in section 10.3.3 clearly becomes unfeasible when the number of subcarriers becomes too high. Compared to GB+UE, the complexity is significantly reduced when using the suboptimal joint energy and bit allocation (GB+OE) described in 10.3.4, whose computation time increases much more slowly with N.

10.5 Chapter Summary

In this chapter, we have considered adaptive coding and modulation in a cognitive BIC-OFDM system, under the realistic assumption that only imperfect CSI is

available. In order to tackle this problem, we introduced an optimum performance metric called the expected goodput (EGP), which is the expectation of the goodput, conditioned on the imperfect CSI.

A major advantage of this metric is that it allows the TX to account for the imperfections of the CSI by selecting its transmission parameters such that the best average goodput is achieved. To make the optimization of the code rate, bit and energy allocation tractable, we proposed a very accurate approximation of this performance metric, referred to as approximate EGP (AEGP)**.** The numerical results clearly show the resource allocation algorithms based on the AEGP have at least the same performance as the non-adaptive algorithms and, in most cases, clearly outperform them. Finally, we also show that, depending upon the quality of the available CSI, the proposed algorithms can come very close to the performance of algorithms with perfect CSI.

10.A Examples of Different Types of Imperfect CSI at the Transmitter

In the following, the impulse response of a generic channel between the SU-TX and any RX of the PU or SU network will be defined as in section 9.1. This means that, for given t, the samples $h(mT,t)$, $m=0,\ldots,\nu$, of the channel impulse vector $\mathbf{h}(t) \triangleq [h(0,t),\ldots,h(\nu T,t)]^T$ are assumed to be independent circularly symmetric zero-mean Gaussian complex random variables; assuming stationarity w.r.t. the variable t, the covariance matrix of $\mathbf{h}(t)$ is given by[1] $\mathbf{R}_h \triangleq \text{diag}(\sigma_0^2,\ldots,\sigma_\nu^2)$. The time variations of the channel are described by Jakes' model [12], which gives $\mathbb{E}\left[h(mT,t+\tau_\text{d})h^*(mT,t)\right] = \text{J}_0(2\pi f_\text{d}\tau_\text{d})\sigma_m^2$, where $\text{J}_0(x)$ represents the zeroth-order Bessel function of the first kind, and f_d denotes the Doppler frequency.

If $\mathbf{F} \in \mathbb{C}^{N_\text{car}\times(\nu+1)}$ denotes the Fourier matrix introduced in (8.6), the time-varying frequency response of the channel can then be written as $\mathbf{H}(t) = \mathbf{F}\mathbf{h}(t)$ which has the covariance matrix $\mathbf{R}_\mathbf{H} = \mathbf{F}\mathbf{R}_h\mathbf{F}^H$. The kth component of $\mathbf{H}(t)$ denotes the channel gain which affects the *k*th subcarrier at time instant t.

In the following subsections, we consider a few possible examples of the type of CSI available at the TX. Each case leads to different expressions for the parameters $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{ICSI}}$ and $\mathbf{C}_{\mathbf{H}|\mathbf{ICSI}}$, which completely describe the random variable $\mathbf{H}(t)$ conditioned on the available $\mathbf{ICSI}$ as follows

$$\mathbf{H}(t) = \boldsymbol{\mu}_{\mathbf{H}|\mathbf{ICSI}}(t) + \mathbf{n}(t), \tag{10.25}$$

where $\mathbf{n}(t) \sim N_\text{c}(\mathbf{0}, \mathbf{C}_{\mathbf{H}|\mathbf{ICSI}})$. The probability density function $p(\mathbf{H}(t)|\mathbf{ICSI})$ is then given by $N_\text{c}(\boldsymbol{\mu}_{\mathbf{H}|\mathbf{ICSI}}(t), \mathbf{C}_{\mathbf{H}|\mathbf{ICSI}})$. If only N of the N_car subcarriers are available at the TX, as is the case in the numerical section, we can define a smaller $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{ICSI}}$ and $\mathbf{C}_{\mathbf{H}|\mathbf{ICSI}}$ which only contain the elements corresponding to the available subcarriers.

10.A.1 Estimated CSI

In this subsection we determine the quantities $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{ICSI}}$ and $\mathbf{C}_{\mathbf{H}|\mathbf{ICSI}}$ in the case of channel estimation errors. The TX only has access to an estimated frequency response $\tilde{\mathbf{H}}(t)$, which means that $\mathbf{ICSI} = \tilde{\mathbf{H}}(t)$. The estimated frequency response $\tilde{\mathbf{H}}(t)$ is decomposed as

$$\tilde{\mathbf{H}}(t) = \mathbf{H}(t) + \tilde{\mathbf{e}}(t), \tag{10.26}$$

where $\tilde{\mathbf{e}}(t)$ and $\mathbf{H}(t)$ are statistically independent, $\tilde{\mathbf{e}}(t) \sim N_\text{c}(\mathbf{0}, \sigma^2\mathbf{I}_{N_\text{car}})$. In section 10.4, we will use the value of the normalized estimation error variance

[1] Note that if we have a number of paths $L < \nu+1$, only L diagonal elements of $\mathbf{R}_h$ are strictly greater than 0.

$\sigma_{\mathrm{e}}^2 \triangleq \sigma^2/\mathrm{Tr}(\mathbb{E}\left[\mathbf{h}\mathbf{h}^H\right])$. Using the results from section 3.3, it can be shown that

$$\boldsymbol{\mu}_{\mathbf{H}|\mathbf{ICSI}} = \mathbf{R}_{\mathbf{H}}(\mathbf{R}_{\mathbf{H}} + \sigma^2\mathbf{I}_{N_{\mathrm{car}}})^{-1}\tilde{\mathbf{H}}(t), \tag{10.27}$$

and

$$\mathbf{C}_{\mathbf{H}|\mathbf{ICSI}} = \mathbf{R}_{\mathbf{H}} - \mathbf{R}_{\mathbf{H}}(\mathbf{R}_{\mathbf{H}} + \sigma^2\mathbf{I}_{N_{\mathrm{car}}})^{-1}\mathbf{R}_{\mathbf{H}}. \tag{10.28}$$

Note that in the case of perfect estimation (i.e., $\sigma^2 = 0$) we obtain perfect CSI, as (10.26), (10.27) and (10.28) reduce to $\tilde{\mathbf{H}}(t) = \mathbf{H}(t)$, $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{ICSI}} = \mathbf{H}(t)$ and $\mathbf{C}_{\mathbf{H}|\mathbf{ICSI}} = 0$.

10.A.2 Delayed CSI

Now we assume that the **ICSI** is outdated, because of a delay in the feedback to the TX. At time instance t, the delayed CSI available at the TX is denoted by $\mathbf{H}(t - \tau_{\mathrm{d}})$, where τ_{d} denotes the delay. In this case, it can be shown that

$$\boldsymbol{\mu}_{\mathbf{H}|\mathbf{ICSI}} = J_0(2\pi f_{\mathrm{d}}\tau_{\mathrm{d}})\mathbf{H}(t - \tau_{\mathrm{d}}), \tag{10.29}$$

and

$$\mathbf{C}_{\mathbf{H}|\mathbf{ICSI}} = (1 - J_0(2\pi f_{\mathrm{d}}\tau_{\mathrm{d}})^2)\mathbf{R}_{\mathbf{H}}. \tag{10.30}$$

When $\tau_d = 0$, we obtain perfect CSI, as (10.29) and (10.30) reduce to $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{ICSI}} = \mathbf{H}(t)$ and $\mathbf{C}_{\mathbf{H}|\mathbf{ICSI}} = 0$.

10.A.3 Estimated and Delayed CSI

In this section we assume that the CSI available at the TX is both delayed and estimated. We also consider the possibility that the TX has access to multiple delayed estimates. With P denoting the number of available estimates, the CSI which is available at the TX is given by

$$\mathbf{ICSI} = [\tilde{\mathbf{H}}(t - \tau_{\mathrm{d}})^T \ldots \tilde{\mathbf{H}}(t - P\tau_{\mathrm{d}})^T]^T, \tag{10.31}$$

where $\tilde{\mathbf{H}}(t - k\tau_{\mathrm{d}})$ ($\forall k \in \{1, \ldots, P\}$) is defined as in (10.26). Defining the matrices

$$\mathbf{X} \triangleq [J_0(2\pi f_{\mathrm{d}}\tau_{\mathrm{d}}), J_0(2\pi 2 f_{\mathrm{d}}\tau_{\mathrm{d}}) \ldots, J_0(2\pi P f_{\mathrm{d}}\tau_{\mathrm{d}})] \otimes \mathbf{R}_{\mathbf{H}}, \tag{10.32}$$

$$\mathbf{Y} \triangleq \mathbf{J} \otimes \mathbf{R}_{\mathbf{H}} + \mathbf{I}_P \otimes \sigma^2\mathbf{I}_{N_{\mathrm{car}}}, \tag{10.33}$$

where $\mathbf{J} \in \mathbb{C}^{P\times P}$ with entries $(\mathbf{J})_{k,l} \triangleq J_0(2\pi f_{\mathrm{d}}\tau_{\mathrm{d}}(k-l))$, $k = 1, \ldots, P; l = 1, \ldots, P$, and $\otimes$ indicates the Kronecker product, it can be shown that

$$\boldsymbol{\mu}_{\mathbf{H}|\mathbf{ICSI}} = \mathbf{X}\mathbf{Y}^{-1}\mathbf{ICSI}, \tag{10.34}$$

and

$$\mathbf{C}_{\mathbf{H}|\mathbf{ICSI}} = \mathbf{R}_{\mathbf{H}} - \mathbf{X}\mathbf{Y}^{-1}\mathbf{X}^{H}. \tag{10.35}$$

10.B PGP and IC-κESM

In the numerical results, we compare our approach to the following two approximations: the predicted goodput metric (PGP) and the IC-κESM. The PGP, introduced in [103], is obtained by neglecting the uncertainty on $\mathbf{H}$ given the vector $\mathbf{ICSI}$, and is calculated by substituting $\mathbf{H}$ by $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{ICSI}}$ in the expression for Y and using this deterministic value of Y to replace the random variable Z in (10.9). Thus, we get

$$\mathrm{PGP} = \frac{N_{\mathrm{p}} r}{N_{\mathrm{tot}} N} \left(\sum_{k \in \mathcal{N}} m_k \right) \cdot \left(1 - \mathrm{PER}_{\mathrm{AWGN}}(r, -\beta \ln(Y_{\mathrm{PGP}}))\right), \tag{10.36}$$

where

$$Y_{\mathrm{PGP}} = \frac{1}{\sum_{l \in \mathcal{N}} m_l} \sum_{k \in \mathcal{N}} \sum_{n=1}^{\frac{\sqrt{2^{m_k}}}{2}} \frac{\psi^k(n)}{2^{m_k - 1}} e^{-\frac{E_k \left| \left(\boldsymbol{\mu}_{\mathbf{H}|\mathbf{ICSI}}\right)_k \right|^2 n^2 d_{k,\min}^2}{4\beta\sigma_w^2}}. \tag{10.37}$$

The IC-κESM metric was introduced in [107] and assumes the following channel model

$$\mathbf{H}(t) = J_0(2\pi f_{\mathrm{d}} \tau_{\mathrm{d}}) \mathbf{H}(t - \tau_{\mathrm{d}}) + \sqrt{1 - J_0(2\pi f_{\mathrm{d}} \tau_{\mathrm{d}})^2}\mathbf{e}, \tag{10.38}$$

where $\mathbf{e} \sim N_{\mathrm{c}}(\mathbf{0}, \mathbf{R}_{\mathbf{H}})$. From appendix 10.A it is clear that this approximation only applies to delayed CSI. We get

$$\mathrm{IC} - \kappa\mathrm{ESM} = \frac{N_{\mathrm{p}} r}{N_{\mathrm{tot}} N} \left(\sum_{k \in \mathcal{N}} m_k \right) \cdot \left(1 - \mathrm{PER}_{\mathrm{AWGN}}(r, -\beta \ln(Y_{\mathrm{IC}-\kappa\mathrm{ESM}}))\right), \tag{10.39}$$

where

$$Y_{\mathrm{IC}-\kappa\mathrm{ESM}} = \frac{1}{\sum_{l \in \mathcal{N}} m_l} \sum_{k \in \mathcal{N}} \sum_{n=1}^{\frac{\sqrt{2^{m_k}}}{2}} \frac{\psi^k(n)}{2^{m_k - 1}} \frac{\exp\left(\frac{-J_0(2\pi f_{\mathrm{d}} \tau_{\mathrm{d}})^2 |(\mathbf{H}(t-\tau_{\mathrm{d}}))_k|^2 \frac{E_k}{4\beta\sigma_w^2} n^2 d_{k,\min}^2}{1 + \frac{E_k}{4\beta\sigma_w^2} n^2 d_{k,\min}^2 (1 - J_0(2\pi f_{\mathrm{d}} \tau_{\mathrm{d}})^2)(\mathbf{R}_{\mathbf{H}})_{k,k}} \right)}{1 + \frac{E_k}{4\beta\sigma_w^2} n^2 d_{k,\min}^2 \left(1 - J_0(2\pi f_{\mathrm{d}} \tau_{\mathrm{d}})^2\right) (\mathbf{R}_{\mathbf{H}})_{k,k}}. \tag{10.40}$$

11

Concluding Remarks and Ideas for Future Work

In this chapter, we formulate the main conclusions from this doctoral thesis (section 11.1) and present some ideas for future work (Section 11.2). Finally, section 11.3 gives an overview of our publications related to this dissertation.

11.1 Main Conclusions

As the demand in wireless applications is continuously increasing, an efficient use of the spectrum is more important than ever. For this reason, we turned our attention towards various resource allocation algorithms in cognitive underlay networks, which are designed to coexist with other networks that are present in the same frequency band.

In this dissertation, we have investigated the performance of several allocation schemes that allow a dynamic adaptation of the transmission parameters, in order to increase the spectral efficiency. However, a dynamic allocation scheme requires the availability of channel state information (CSI) at the transmitter. As providing accurate CSI to the transmitter will lead to additional overhead for the network, we have investigated how the quality of this CSI influences the performance of the network.

First, we focused on the optimization of information-theoretical metrics. We

considered both cooperative networks and multiple antenna systems, as these increase the reliability of the communication over the wireless fading channel.

- For a multi-antenna relay network, we proposed several distributed algorithms that minimize the exact outage probability of the network, while protecting the quality of service of the primary user network. These distributed algorithms optimize the beamforming at the relay nodes and the allocation of the transmit energy at the source node. When the available CSI at the transmitter is not perfect, the proposed resource allocation algorithm allows the transmitter to take these imperfections into account. This method ensures that the secondary user network can maintain the limit on the interference to the primary user network, regardless of the quality of the available CSI, while minimizing its outage probability. The proposed algorithms were shown to outperform several algorithms presented in literature. Further, we found that the use of multi-antenna relays significantly decreases the outage probability of a secondary user network which uses the underlay paradigm. It became clear that the secondary user network considerably benefits from having at least imperfect channel knowledge of the channel gains to the primary user receivers, and having relay nodes with a number of antennas larger than the number of primary user antennas. The level of CSI about the channels to the destination node was found to have a smaller impact on the outage probability.

- We compared the outage probability resulting from resource allocation algorithms that consider an average or a peak interference constraint. It became clear that the outage probability of the secondary user network can be significantly lowered when an average interference constraint is imposed instead of a peak interference constraint. When all relay nodes transmit in turn, the computational complexity of the average interference-based and peak interference-based approaches was found to be comparable. However, even for a simple network, it appeared to be quite difficult to find a resource allocation algorithm that combines the use of average interference constraints with a relay selection algorithm. The latter observation is important, because the use of multiple relay nodes, transmitting in turn, can lead to a significant performance loss for low SNRs.

- We proposed an approximation of the outage probability for a cooperative multicarrier network, which was shown to be more accurate than the often used Gaussian approximation. This approximation is used for the optimization of the transmission rate between the secondary user source, which has imperfect CSI available, and destination node under a fixed outage probability constraint. This approach led to a centralized resource allocation algorithm, which implies that a) each node of the network has to receive a

periodical update from a central unit about its transmission parameters and b) the central unit requires CSI about all the relevant channels in the network. However, even if a centralized solution is not practical, our solution still provides an interesting benchmark to which the spectral efficiency achieved by more practical algorithms can be compared. Further, it was shown that taking a sufficiently large memory P of the channel predictor at the transmitter significantly reduces the loss in spectral efficiency, compared to the case of perfect CSI.

Next, we optimized the goodput of a practical packet-based secondary user transmission system. The goodput is a practical metric that expresses the ratio of the expected number of correctly received information bits to the actual transmission time. We assumed that the packet-based system uses bit-interleaved coded modulation (BICM), as this allows for a high flexibility of the system by separating the encoding and the constellation mapping. Further, we also made use of a technique, called effective SNR mapping, which provides an analytical approximation of the goodput of the system.

- As the effective mapping technique has originally been introduced under the assumption of perfect CSI, we extended this technique to include imperfect CSI, by proposing an approximation based on the beta distribution. We showed that our proposed beta approximation outperforms often used approximations such as the gamma, Gaussian, log-Gaussian and lognormal distributions, while providing an accuracy that is only slightly worse, compared to the more involved generalized extreme value and Pearson distributions.

- We introduced the expected goodput metric, which we defined as the expectation of the goodput, conditioned on the imperfect CSI that is available at the transmitter. By using this metric, it becomes possible to derive resource allocation algorithms that are able to select the code rate, constellation and transmit energy per subcarrier in such a manner that the best average goodput is achieved. In most cases, it was found that our proposed resource allocation algorithms clearly outperform non-adaptive algorithms. Further, by incorporating the knowledge about the imperfect CSI into the optimization problem, we showed that the resulting goodput can come very close to the goodput achieved by algorithms with perfect CSI.

However, there are still some important topics that we did not discuss. In the following section, we will present our ideas about some possibilities for future work.

11.2 Future Work

11.2.1 Imperfect CSI at the Receiver

In practice, the CSI available at the transmitter and at the receiver will both contain errors. Because the CSI errors at the receiver are often smaller than those at the transmitter [3], in this dissertation we made the assumption that the receiver had perfect CSI available, and only the CSI at the transmitter was imperfect. Still, the performance that can be achieved by our algorithms in the scenario where the receiver has imperfect CSI available remains to be investigated. We expect that for relatively small errors in the CSI at the receiver, the performance of the algorithms will still be close to optimal. However, when the errors in the CSI at the receiver become large, it will be necessary to adapt our resource allocation algorithms. We like to mention that even in this case, our proposed algorithms provide a useful upper bound on the performance, which allows a system designer to assess the loss occurred by a receiver with imperfect CSI.

11.2.2 Multiple Secondary Users

In this work, we have restricted our attention to a single secondary user network. However, it is possible to extend the algorithms to a scenario where multiple secondary user networks coexist.

However, when optimizing the performance of multiple users there is no longer a single objective function that we can optimize. For these problems, we come into the domain of multi-objective optimization. As each user has its own objective function, the network has to choose how it will divide its resources between the users. If we want to maximize the worst performance among all users, the following optimization problem can be used

$$\begin{aligned} &\max \min_i \ f_i(\mathbf{E}) \\ &\text{s.t. intf constraints,} \end{aligned} \tag{11.1}$$

where $\mathbf{E} \triangleq [E_1, \ldots, E_K]^T$, $f_i(\mathbf{E})$ denotes the performance metric of the *i*th user, $i = 1, \ldots, K$, as a function of the energy allocation vector $\mathbf{E}$ and "intf constraints" denote the interference constraints that protect the primary user network. It is clear that the function $f_i(\mathbf{E})$ in general does not only depend on the energy allocation E_i of the *i*th user, but also on the energy allocation E_j ($j \neq i$) of the other users. The reason for this is that the different secondary users also cause interference to each other. This means that choosing a higher transmit energy will now not only cause more interference to the primary user network, but also to the other secondary users.

The approach in (11.1) is not always the ideal solution. The proposed optimization problem in (11.1) treats every secondary user equally, however this can lead to an unsatisfactory result. If r_i denotes the minimum required performance of the ith user, $i = 1, \dots K$, it is possible that some users are allocated more resources than strictly required ($r_i < f_i(\mathbf{E})$), while other users might not achieve their required performance criterion ($r_i > f_i(\mathbf{E})$).

Therefore, in some cases the following objective function is more suitable

$$\sum_i \omega_i f_i(\mathbf{E}), \tag{11.2}$$

where ω_i denotes the weight that is given to user i. By tuning these weights, the system designer can prioritize certain users. This allows the network to differentiate between users.

If we want to make sure that each user has a certain minimum performance, the following optimization problem can be used

$$\begin{aligned} &\min_{E_1,\dots,E_K} \sum_i E_i \\ &\text{s.t. } f_i(\mathbf{E}) \geq r_i \\ &\quad \text{intf constraints.} \end{aligned} \tag{11.3}$$

The objective function in (11.3) is the minimization of the total energy consumption of the network, which is for example useful when the network is battery operated. However, the objective function in (11.3) can also be chosen similar to (11.2) such that we can choose a different priority for each user, while guaranteeing the required minimum performance of each user.

11.2.3 Primary User Interference

We have not given much attention to the interference that the secondary user network receives from the primary user network. We briefly mentioned that this interference can be taken into account by increasing the variance of the noise at the secondary user receiver, which requires only a minimal adaptation of the algorithms proposed in this dissertation: a) The algorithms now have to take into account that the combined variance of the noise and interference can be different at each relay and destination node, b) the secondary user transmitters have to know the value of these combined variances.

However, this method is only a good first approximation, if we can assume that the interference coming from the primary user transmitters is approximately normally distributed. This is the case when the interference comes from multiple independent primary user transmitters, as the central limit theorem can be applied in this scenario.

When the interference at the secondary user receivers cannot be considered Gaussian, which can be the case when the number of primary user transmitters is small, we have to take a different approach. One possible solution is to model the SNR at a secondary user receiver as

$$\frac{E_{\mathrm{s}}|h|^2}{E_{\mathrm{p}}|g|^2 + \sigma^2}, \tag{11.4}$$

where σ^2 denotes the noise variance, E_{s} and E_{p} denotes the transmit energy per symbol of the secondary user network and primary user network, respectively. The scalar values h and g denote the channel gains between the secondary user transmitter and secondary user receiver, and between the primary user transmitter and secondary user receiver, respectively. In this approach, the secondary user network requires (an estimate of) the instantaneous channel gain g. However, as it is not always easy to obtain the channel gain g in a practical network, it can be necessary in some cases to take into account that the estimate of the channel gain g is imperfect.

Depending on the amount of interference that is present at the secondary user receivers, there are several approaches that can be pursued. When the interference at the secondary user receivers is low, it can be feasible in some cases to just neglect the interference, in which case a small performance loss will be incurred. However, when the primary user transmitters are relatively close to the secondary user receivers, it can be necessary to take some countermeasures in the secondary user network. For example, when multiple receive antennas are present at the secondary user nodes, as in chapter 5, zero-forcing beamforming instead of the maximum ratio combining can be applied, in order to eliminate the interference caused by the primary user network.

11.2.4 Carrier Frequency Offset

In this dissertation we have made the assumption that there is no carrier frequency offset between a transmitter and a receiver. However, the performance of multicarrier systems is very sensitive to carrier frequency offset. This means that an accurate carrier frequency synchronization between transmitter and receiver is necessary.

However, in some applications low-latency is key. In these cases, a time-consuming synchronization protocol between the transmitter and receiver is not always desirable and the use of orthogonal frequency-division multiplexing (OFDM) may not be the best choice.

For these reasons, other waveforms that are more robust against synchronization errors are being considered for fifth generation (5G) networks. One such waveform is filter bank multicarrier (FBMC) [108]. In this waveform, each subcarrier is individually filtered. This approach makes the multicarrier system much

more robust against intercarrier interference caused by frequency offset. However in order to achieve a steep filtering of the subcarriers, a large filter length is required.

A tradeoff between OFDM and FBMC is universal filtered multicarrier (UFMC), which filters contiguous groups of subcarriers, called subbands, instead of individual subcarriers. This approach leads to an increased robustness against frequency offsets compared to OFDM, without the need for the long filters that FBMC requires. We already made some contributions in the following papers [109, 110] about this topic. In these papers, the resource allocation algorithm from chapter 10 is adapted to a bit interleaved coded UFMC system. These papers also investigate the impact of a carrier frequency offset on the goodput of the network. In [110], the resource allocation takes the carrier frequency offset into account by modeling it as a random variable, with known distribution.

11.2.5 Bit-Interleaved Coded Modulation with Iterative Decoding

In [111], it was shown that the bit error rate of the BICM can be improved by performing iterative decoding (BICM-ID). Reconsidering the equation (4.4) describing the soft demapping, we get

$$p(y_k(i)|c_{j,k}(i)=b,H_k)=\sum_{\alpha\in\mathcal{X}_k} p(y_k(i)|x_k(i)=\alpha,H_k)\Pr(x_k(i)=\alpha|c_{j,k}(i)=b). \tag{11.5}$$

In chapter 4, we set $\Pr(x_k(i)=\alpha|c_{j,k}(i)=b)$ equal to $1/2^{m_k-1}$ for $\alpha\in\chi_{k,b}^j$, where m_k denotes the number of bits on the *k*th subcarrier. This approach assumes that all symbols in $\chi_{k,b}^j$ are equally likely, implying that we do not have any a priori information about the symbols. However, when using iterative decoding, it is possible for the decoder to feed back the estimated values of the probabilities $\Pr(c_{j,k}(i)=b)$ to the demapper. The quantity $\Pr(c_{j,k}(i)=b)$ denotes the probability that the *j*th coded bit of the symbol $x_k(i)$ equals b. This allows the modulator to calculate $\Pr(x_k(i)=\alpha|c_{j,k}(i)=b)$ as follows

$$\Pr(x_k(i)=\alpha|c_{j,k}(i)=b)=\begin{cases}\prod_{l\neq j}\Pr(c_{l,k}(i)=f_l(\alpha)), & \alpha\in\chi_{k,b}^j\\ 0, & \alpha\notin\chi_{k,b}^j\end{cases}, \tag{11.6}$$

where $f_j(\alpha)$ denotes the *j*th bit of the symbol α, and the events $c_{j,k}(i)=f_j(\alpha)$ ($j\in\{1,\ldots,m_k\}$) are assumed to be independent because of the use of a bit interleaver.

For the non-iterative BICM system, we have used Gray mapping. This mapping function has the nice property that the Hamming distances between the bit labels of constellation points at minimum Euclidean distance are minimized. This

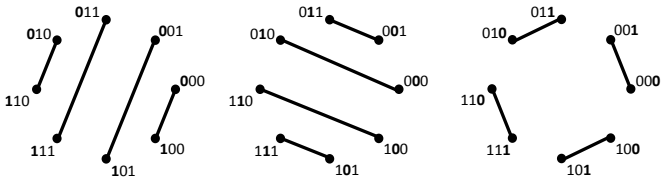


Figure 11.1: Gray mapping with perfect a priori information.

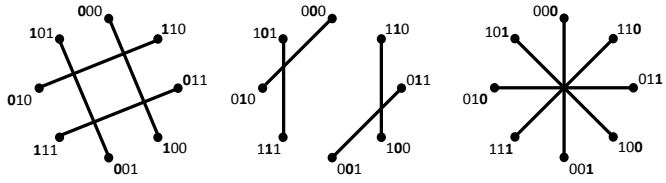


Figure 11.2: Set partitioning with perfect a priori information.

property is perfect for the non-iterative BICM system. However, in the case of the iterative BICM system, the demapper now receives feedback information from the decoder about the coded bits, which has a huge impact on the properties that we require from a mapping function [112]. To illustrate this point, we assume that the demapper has perfect a priori information available about all the bits except the one bit we are trying to detect. In Fig. 11.1 and 11.2 we show the various symbol pairs in a 8-PSK constellation for the Gray mapping and set partitioning, respectively. For example, if we assume that the second and third coded bit are known at the demapper, we can divide the constellation into four separate pairs as shown in the left of Fig. 11.1. The value of the second and third coded bit determines which pair the demapper considers for the detection of the first bit. The center and right part of Fig. 11.1 and 11.2 show the pairs for the detection of the second and third bit, respectively.

From Fig. 11.1 and 11.2, we notice that the Euclidean distance between the two points from a pair is larger for the set partitioning than for the Gray mapping. This means that when a priori information about the coded bits is present at the demapper, the set partitioning can lead to a higher goodput, compared to the Gray mapping. However, when the a priori information at the demapper is wrong, it is possible that error propagation will occur. In this case, set partitioning can lead to a goodput that is lower than the goodput achieved by Gray mapping. However, a good convolutional code can provide a high reliability for the decoded bits, which significantly reduces the probability of error propagation.

However, when a mapping different from the Gray mapping is used, equation (4.23), which was used for the derivation of the cumulant generating function based effective SNR mapping (κESM), does no longer hold. Further, the iterative decoding will also lead to an actual goodput which differs from the goodput predicted by the effective SNR. For these reasons, we have to revise the κESM function for BICM-ID.

A possible way to derive an approximation for BICM-ID, is by considering expression (4.22)

$$\kappa(1/2) = \ln \mathbb{E}\left[e^{-\gamma_k d^2(x_k, x'_k, j)/4} | \mathbf{SNR}\right], \tag{11.7}$$

where $d(x_k, x'_k, j)$ is the Euclidean distance between the symbol x_k and x'_k. We will now assume that the feedback information from the decoder is perfect. This means that we will define the symbol x'_k as the symbol that corresponds to the label of x_k but where the bit in the *j*th position is a 1 instead of a 0. If we then calculate expression (11.7) for the quadrature amplitude modulation (QAM) constellation shown in Fig. 3.2, we get

$$\kappa(1/2) = \ln\left(\frac{1}{\sum_{k\in\mathcal{N}} m_k} \sum_{k\in\mathcal{N}} \frac{1}{2^{m_k-1}} \sum_{\mu=1}^{3} \psi^k(\mu) e^{-\gamma_k (\mu d_{k,\min})^2/4}\right), \tag{11.8}$$

where

$$\psi^k(\mu) = 24\delta_{\mu-1} + 8\delta_{\mu-3}, \tag{11.9}$$

instead of

$$\psi^k(\mu) = 24\delta_{\mu-1} + 8\delta_{\mu-2}, \tag{11.10}$$

in the case of a non-iterative receiver.

However, note that the expression for $\kappa(1/2)$ can become more complicated for other labeling methods. For example, if we consider the random labeling shown in Fig. 11.3. The expression for $\kappa(1/2)$ becomes

$$\kappa(1/2) = \ln\left(\frac{1}{\sum_{k\in\mathcal{N}} m_k} \sum_{k\in\mathcal{N}} \frac{1}{2^{m_k-1}} \sum_{\mu\in\mathcal{S}} \psi^k(\mu) e^{-\gamma_k (\mu d_{k,\min})^2/4}\right), \tag{11.11}$$

where

$$\psi^k(\mu) = 6\delta_{\mu-1}+5\delta_{\mu-\sqrt{2}}+8\delta_{\mu-2}+3\delta_{\mu-\sqrt{5}}+2\delta_{\mu-\sqrt{8}}+4\delta_{\mu-3}+3\delta_{\mu-\sqrt{10}}+\delta_{\mu-\sqrt{13}} \tag{11.12}$$

and $\mathcal{S} = \{1, \sqrt{2}, 2, \sqrt{5}, \sqrt{8}, 3, \sqrt{10}, \sqrt{13}\}$. Notice that the summation of $\kappa(1/2)$ becomes more involved, but the expression itself does not change.

In Fig. 11.4, we show the pdf of the log-likelihood ratio (LLR) defined in (4.3) for a 16-QAM constellation with both Gray and random labeling when perfect

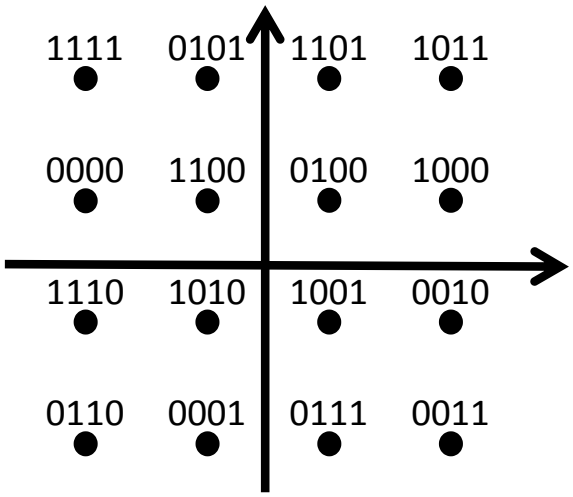


Figure 11.3: 16-QAM with random labeling.

feedback information is assumed and the transmitted bit equals zero. We assume an AWGN channel and a ratio $E_{\mathrm{s}}/\sigma^2 = 12$ dB. In the figure, we also show the Gaussian approximation $N(4\kappa(1/2), -8\kappa(1/2))$, discussed in [59]. We can see that this approximation is tight for the tail of the LLR. As the tail of the LLR is related to the error probability, this tail is exactly the region we are interested in. Further, the fact that the Gaussian distribution is a good approximation of the tail of the distribution of the LLR, is an essential result on which the κESM mapping function is based. Because the Gaussian approximation also appears to be tight in the case with perfect feedback information, it makes us hopeful that the algorithms proposed in chapter 10 can be easily adapted to operate in a system that uses an iterative BICM receiver. However, it remains to be validated how well this approximation will perform when the received feedback from the decoder is not perfect.

11.2.6 BICM with Turbo Codes

In this dissertation, we have investigated the goodput of a BICM system that uses convolutional codes. In [58], where the κESM function was originally proposed, the authors also considered the possible use of a turbo code instead of a convolutional code. Thus, using a turbo code instead of a convolutional code would not drastically change our allocation algorithms.

However, some differences have to be taken into account. First, a small perfor-

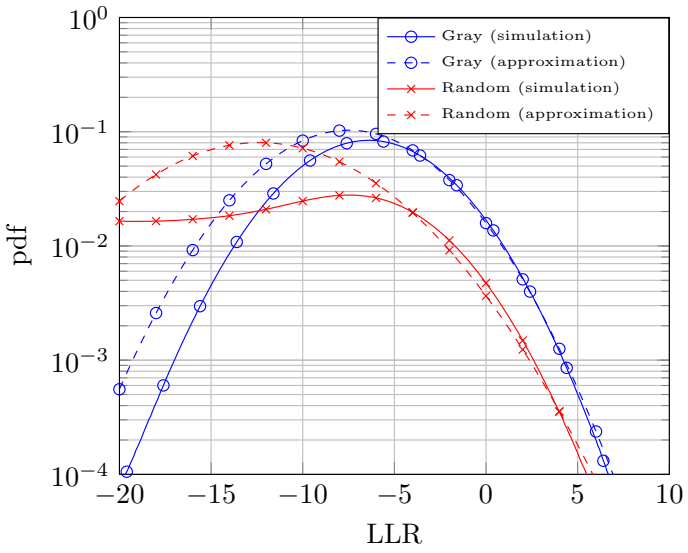


Figure 11.4: Probability density function of the LLR.

mance gain can be achieved when using multiple parallel interleavers, instead of a single interleaver [113]. The use of multiple interleavers allows the transmitter to map the systematic bits to the most protected bit positions in the constellation. As the systematic bits are very important in the iterative decoding of the turbo decoding algorithm, this strategy allows for a small gain in the goodput.

Further, in order to decode the turbo codes, the receiver requires an iterative decoder. Thus, the decoder will be more complicated than a regular convolutional decoder. The advantage of using turbo codes however, is that they achieve a given error performance at a considerably lower SNR, compared to convolutional codes. This could be very beneficial for underlay networks, where the secondary user network has to limit the interference to the primary user network.

Finally, for BICM with turbo codes it is also possible to iterate between the demapper and the decoder as is the case in BICM-ID. However, the possible gain that can be achieved is lower in this case, as the operating point of the turbo codes is already very close to the capacity of the channel [113], which means that there is only very little room for improvement.

11.3 Publications

Our work has been presented in the following refereed international journal and conference publications:

Journal Publications

- IEEE Transactions on Wireless Communications: [71, 114]
- EURASIP Journal on Wireless Communications and Networking: [104]

Conference Publications

- IEEE International Conference on Communications (ICC): [103]
- IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC): [81, 86]
- IEEE Wireless Communications and Networking Conference (WCNC): [79]
- IEEE Symposium on Communications and Vehicular Technologies in the Benelux: [96, 110]
- IEEE Globecom Workshops: [109]
- European Wireless Conference: [107]

References

[1] S. Haykin. *Cognitive radio: brain-empowered wireless communications*. IEEE J. Sel. Areas Commun., 23(2):201 – 220, February 2005.

[2] National Telecommunications & Information Administration. *United States Frequency Allocation Chart*. `https://www.ntia.doc.gov/files/ntia/publications/january_2016_spectrum_wall_chart.pdf`, January 2016.

[3] A. J. Goldsmith and Soon-Ghee Chua. *Variable-rate variable-power MQAM for fading channels*. IEEE Trans. Commun., 45(10):1218–1230, October 1997.

[4] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge, U.K., 2004.

[5] Michael Grant and Stephen Boyd. *CVX: Matlab software for disciplined convex programming, version 2.1*. `http://cvxr.com/cvx`, March 2014.

[6] S. Boyd, L. Xiao, and A. Mutapcic. *Subgradient methods*. Stanford University, October 2003.

[7] S. Boyd and A. Mutapcic. *Stochastic subgradient methods notes for EE364b*. Stanford University, January 2006-07.

[8] Are Hjørungnes. *Complex-valued matrix derivatives : with applications in signal processing and communications*. Cambridge University Press, Cambridge, 2011.

[9] K. B. Petersen and M. S. Pedersen. *The matrix cookbook*, November 2012. Version 20121115.

[10] K. Kreutz-Delgado. *The complex gradient operator and the CR-calculus*. University of California, 2006.

[11] Andrea Goldsmith. *Wireless communications*. Cambridge University Press, New York, NY, USA, 2005.

[12] William C. Jakes and Donald C. Cox, editors. *Microwave mobile communications*. Wiley-IEEE Press, 1994.

[13] ETSI. *Selection procedures for the choice of radio transmission technologies of the UMTS (UMTS 30.03 version 3.1.0)*. Technical report, UMTS.

[14] David Tse and Pramod Viswanath. *Fundamentals of wireless communication*. Cambridge University Press, New York, NY, USA, 2005.

[15] H. L. Van Trees. *Detection, estimation, and modulation theory*. Wiley, 1 edition, September 2001.

[16] J.G. Proakis. *Digital communications*. McGraw-Hill series in electrical and computer engineering. McGraw-Hill Higher Education, 2001.

[17] Daji Qiao, Sunghyun Choi, and K. G. Shin. *Goodput analysis and link adaptation for IEEE 802.11a wireless LANs*. IEEE Trans. Mobile Comput., 1(4):278–292, Oct 2002.

[18] Tho Le-Ngoc (auth.) Leonardo Jiménez Rodríguez, Nghi Tran. *Amplify-and-forward relaying in wireless communications*. SpringerBriefs in Computer Science. Springer International Publishing, 1 edition, 2015.

[19] G. Kramer, M. Gastpar, and P. Gupta. *Cooperative strategies and capacity theorems for relay networks*. IEEE Trans. Inf. Theory, 51(9):3037–3063, September 2005.

[20] J.N. Laneman, D.N.C. Tse, and G.W. Wornell. *Cooperative diversity in wireless networks: efficient protocols and outage behavior*. IEEE Trans. Inf. Theory, 50(12):3062 – 3080, December 2004.

[21] J. N. Laneman and G. W. Wornell. *Distributed space-time coded protocols for exploiting cooperative diversity in wireless networks*. In Proc. IEEE Global Telecommunications Conf. (Globecom), volume 1, pages 77–81 vol.1, November 2002.

[22] Yingwei Yao, Xiaodong Cai, and G. B. Giannakis. *On energy efficiency and optimum resource allocation of relay transmissions in the low-power regime*. IEEE Trans. Wireless Commun., 4(6):2917–2927, November 2005.

[23] Xitirnin Deng and A. M. Haimovich. *Power allocation for cooperative relaying in wireless networks*. IEEE Commun. Lett., 9(11):994–996, November 2005.

[24] M. O. Hasna and M. S. Alouini. *Optimal power allocation for relayed transmissions over Rayleigh-fading channels*. IEEE Trans. Wireless Commun., 3(6):1999–2004, November 2004.

[25] B. Nazer and M. Gastpar. *Compute-and-forward: harnessing interference through structured codes*. IEEE Trans. Inf. Theory, 57(10):6463–6486, October 2011.

[26] T. Cover and A. E. Gamal. *Capacity theorems for the relay channel*. IEEE Trans. Inf. Theory, 25(5):572–584, 1979.

[27] Y. Zhao, R. Adve, and T.J. Lim. *Improving amplify-and-forward relay networks: optimal power allocation versus selection*. IEEE Trans. Wireless Commun., 6(8):3114 –3123, August 2007.

[28] J. Mitola and G. Q. Maguire. *Cognitive radio: making software radios more personal*. IEEE Pers. Commun., 6(4):13–18, August 1999.

[29] A. Goldsmith, S.A. Jafar, I. Maric, and S. Srinivasa. *Breaking spectrum gridlock with cognitive radios: an information theoretic perspective*. Proc. IEEE, 97(5):894 –914, May 2009.

[30] Z. Tian, G. Leus, and V. Lottici. *Joint dynamic resource allocation and waveform adaptation for cognitive networks*. IEEE J. Sel. Areas Commun, 29(2):443–454, February 2011.

[31] Amir Ghasemi and Elvino S. Sousa. *Fundamental limits of spectrum-sharing in fading environments*. IEEE Trans. Wireless Commun., 6(2):649 –658, February 2007.

[32] Zhihui Shu and Wen Chen. *Optimal power allocation in cognitive relay networks under different power constraints*. In Proc. IEEE Wireless Communications, Networking and Information Security (WCNIS), pages 647 –652, June 2010.

[33] Lan Zhang, Ying-Chang Liang, Yan Xin, and H.V. Poor. *Robust cognitive beamforming with partial channel state information*. IEEE Trans. Wireless Commun, 8(8):4143–4153, August 2009.

[34] J.M. Peha. *Approaches to spectrum sharing*. IEEE Commun. Mag., 43(2):10–12, 2005.

[35] Uditha Lakmal Wijewardhana, Marian Codreanu, Matti Latva-aho, and Anthony Ephremides. *A robust beamformer design for underlay cognitive radio networks using worst case optimization*. EURASIP J. Wirel. Commun. Netw., 2014(1):1–16, 2014.

[36] Emil Björnson and Eduard Jorswieck. *Optimal resource allocation in coordinated multi-cell systems*, volume 9. Found. Trends Commun. Inf. Theory, January 2013.

[37] H. Huang, Z. Li, J. Si, and L. Guan. *Underlay cognitive relay networks with imperfect channel state information and multiple primary receivers*. IET Commun., 9(4):460–467, 2015.

[38] X. Zhang, J. Xing, Z. Yan, Y. Gao, and W. Wang. *Outage performance study of cognitive relay networks with imperfect channel knowledge*. IEEE Commun. Lett., 17(1):27–30, January 2013.

[39] K. Ho-Van, P. C. Sofotasios, and S. Freear. *Underlay cooperative cognitive networks with imperfect Nakagami-m fading channel information and strict transmit power constraint: interference statistics and outage probability analysis*. J. Commun. Netw., 16(1):10–17, February 2014.

[40] N. Mokari, S. Parsaeefard, P. Azmi, H. Saeedi, and E. Hossain. *Robust ergodic uplink resource allocation in underlay OFDMA cognitive radio networks*. IEEE Trans. Mobile Comput., 15(2):419–431, February 2016.

[41] D. Tian, J. Zhou, Z. Sheng, and V. C. M. Leung. *Robust energy-efficient MIMO transmission for cognitive vehicular networks*. IEEE Trans. Veh. Technol., 65(6):3845–3859, June 2016.

[42] W. Jaafar, T. Ohtsuki, W. Ajib, and D. Haccoun. *Impact of the CSI on the performance of cognitive relay networks with partial relay selection*. IEEE Trans. Veh. Technol., 65(2):673–684, February 2016.

[43] S. Mallick, R. Devarajan, R. A. Loodaricheh, and V. K. Bhargava. *Robust resource optimization for cooperative cognitive radio networks with imperfect CSI*. IEEE Trans. Wireless Commun., 14(2):907–920, February 2015.

[44] S. Singh, P. D. Teal, P. A. Dmochowski, and A. J. Coulson. *Robust cognitive radio cooperative beamforming*. IEEE Trans. Wireless Commun., 13(11):6370–6381, November 2014.

[45] S. Ye, R. S. Blum, and L. J. Cimini. *Adaptive OFDM systems with imperfect channel state information*. IEEE Trans. Wireless Commun., 5(11):3255–3265, November 2006.

[46] A. Kuhne and A. Klein. *Throughput analysis of multi-user OFDMA-systems using imperfect CQI feedback and diversity techniques*. IEEE J. Sel. Areas Commun., 26(8):1440–1450, October 2008.

[47] Stephen B. Wicker. *Error control systems for digital communication and storage*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1995.

[48] E. Zehavi. *8-PSK trellis codes for a Rayleigh channel*. IEEE Trans. Commun., 40(5):873–884, May 1992.

[49] Y. Yasuda, K. Kashiki, and Y. Hirata. *High-rate punctured convolutional codes for soft decision viterbi decoding*. IEEE Trans. Commun., 32(3):315–319, March 1984.

[50] G. Caire, G. Taricco, and E. Biglieri. *Bit-interleaved coded modulation*. IEEE Trans. Inf. Theory, 44(3):927–946, May 1998.

[51] Alex Alvarado Leszek Szczecinski. *Bit-interleaved coded modulation: fundamentals, analysis and design*. Wiley - IEEE. Wiley, 1 edition, 2015.

[52] Shu Lin and Daniel J. Costello. *Error Control Coding, Second Edition*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2004.

[53] A. Viterbi. *Error bounds for convolutional codes and an asymptotically optimum decoding algorithm*. IEEE Trans. Inf. Theory, 13(2):260–269, April 1967.

[54] S. Nanda and K. M. Rege. *Frame error rates for convolutional codes on fading channels and the concept of effective Eb/N0*. IEEE Trans. Veh. Technol, 47(4):1245–1250, November 1998.

[55] E. Tuomaala and Haiming Wang. *Effective SINR approach of link to system mapping in OFDM/multi-carrier mobile network*. In Proc. Int. Conf. Mobile Technology, Applications and Systems, pages 5 pp.–5, November 2005.

[56] Ericsson. *System-Level evaluation of OFDM - further Considerations*. 3GPP TSG-RAN-1 35, Document R1-031303, Nov 2003.

[57] Lei Wan, Shiauhe Tsai, and M. Almgren. *A fading-insensitive performance metric for a unified link quality model*. In Proc. IEEE Wireless Communications and Networking Conference (WCNC), volume 4, pages 2110–2114, April 2006.

[58] I. Stupia, V. Lottici, F. Giannetti, and L. Vandendorpe. *Link resource adaptation for multiantenna bit-interleaved coded multicarrier systems*. IEEE Trans. Signal Process., 60(7):3644–3656, July 2012.

[59] A. Martinez, A. Guillen i Fabregas, and G. Caire. *Error probability analysis of bit-interleaved coded modulation*. IEEE Trans. Inf. Theory, 52(1):262–271, January 2006.

[60] I. Stupia, F. Giannetti, V. Lottici, R. Andreotti, L. Vandendorpe, and A.N. D'Andrea. *A greedy algorithm for goodput-oriented AMC in turbo-coded OFDM*. In Proc. Future Network and Mobile Summit, pages 1–8, June 2010.

[61] W. W. Peterson and D. T. Brown. *Cyclic codes for error detection*. Proc. IRE, 49(1):228–235, January 1961.

[62] Rui Zhang and Ying-Chang Liang. *Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks*. IEEE J. Sel. Topics Signal Process., 2(1):88 –102, February 2008.

[63] T. Q. Duong, V. N. Q. Bao, and H. j. Zepernick. *Exact outage probability of cognitive AF relaying with underlay spectrum sharing*. Electron. Lett., 47(17):1001–1002, August 2011.

[64] M. Seyfi, S. Muhaidat, and Jie Liang. *Relay selection in underlay cognitive radio networks*. In Proc. IEEE Wireless Communications and Networking Conference (WCNC), pages 283–288, April.

[65] X. Tang and Y. Hua. *Optimal design of non-regenerative MIMO wireless relays*. IEEE Trans. Wireless Commun., 6(4):1398–1407, April 2007.

[66] B.K. Chalise and L. Vandendorpe. *Optimization of MIMO relays for multipoint-to-multipoint communications: nonrobust and robust designs*. IEEE Trans. Signal Process., 58(12):6355–6368, December 2010.

[67] Quanzhong Li, Qi Zhang, Renhai Feng, Liping Luo, and Jiayin Qin. *Optimal relay selection and beamforming in MIMO cognitive multi-relay networks*. IEEE Commun. Lett., 17(6):1188–1191, June 2013.

[68] Q. Li, L. Luo, and J. Qin. *Optimal relay precoder for non-regenerative MIMO cognitive relay systems with underlay spectrum sharing*. Electron. Lett., 48(5):295–297, March 2012.

[69] K. R. Budhathoki, M. Maleki, and H. R. Bahrami. *Iterative source and relay precoder design for non-regenerative MIMO cognitive relay systems*. IEEE Trans. Commun., 63(10):3497–3510, October 2015.

[70] Q. Li, Q. Zhang, and J. Qin. *Robust beamforming for cognitive multi-antenna relay networks with bounded channel uncertainties*. IEEE Trans. Commun., 62(2):478–487, February 2014.

[71] J. Van Hecke, P. Del Fiorentino, V. Lottici, F. Giannetti, L. Vandendorpe, and M. Moeneclaey. *Distributed dynamic resource allocation for cooperative cognitive radio networks with multi-antenna relay selection*. IEEE Trans. Wireless Commun., 16(2):1236–1249, February 2017.

[72] Amir Beck and Yonina C. Eldar. *Strong duality in nonconvex quadratic optimization with two quadratic constraints*. SIAM J. Optim., 17(3):844–860, 2006.

[73] Yonglan Zhu, Yan Xin, and Pooi-Yuen Kam. *Outage probability of rician fading relay channels*. IEEE Trans. Veh. Technol., 57(4):2648–2652, July 2008.

[74] M. K. Simon and M. S. Alouini. *Digital communication over fading channels: a unified approach to performance analysis*. Wiley, New York, second edition, 2004.

[75] A.H. Nuttall. *Some integrals involving the (Q Sub M)-function*. Naval Underwater Systems Center, New London Laboratory, 1974.

[76] Zhi-Quan Luo, Wing-Kin Ma, A.M.-C. So, Yinyu Ye, and Shuzhong Zhang. *Semidefinite relaxation of quadratic optimization problems*. IEEE Signal Process. Mag., 27(3):20–34, May 2010.

[77] K. Jitvanichphaibool, Y. C. Liang, and R. Zhang. *Beamforming and power control for multi-antenna cognitive two-way relaying*. In Proc. IEEE Wireless Communications and Networking Conference (WCNC), pages 1–6, April 2009.

[78] L. Mirsky. *An introduction to linear algebra*. 1961. Reprinted by Dover, New York, 1990.

[79] D. Francesconi, J. Van Hecke, F. Giannetti, V. Lottici, and M. Moeneclaey. *Distributed dynamic resource allocation for cooperative cognitive radio*. In Proc. IEEE Wireless Communications and Networking Conference (WCNC), pages 679 –683, April 2012.

[80] Rui Zhang. *On peak versus average interference power constraints for protecting primary users in cognitive radio networks*. IEEE Trans. Wireless Commun., 8(4):2112 –2120, April 2009.

[81] J. Van Hecke, F. Giannetti, V. Lottici, and M. Moeneclaey. *Outage probability minimization for cooperative cognitive radio with best-relay selection under an average interference power constraint*. In Proc. IEEE Personal Indoor and Mobile Radio Communications (PIMRC), pages 590–595, September 2013.

[82] V. Emamian, P. Anghel, and M. Kaveh. *Multi-user spatial diversity in a shadow-fading environment*. In Proc. IEEE Vehicular Technology Conf. (VTC), volume 1, pages 573 – 576, 2002.

[83] A. Ribeiro and G.B. Giannakis. *Separation principles in wireless networking*. IEEE Trans. Inf. Theory, 56(9):4488–4505, 2010.

[84] A.G. Marques, L.M. Lopez-Ramos, G.B. Giannakis, and J. Ramos. *Resource allocation for interweave and underlay CRs under probability-of-interference constraints*. IEEE J. Sel. Areas Commun., 30(10):1922–1933, November.

[85] Cheng-Xiang Wang, Xuemin Hong, Hsiao-Hwa Chen, and J. Thompson. *On capacity of cognitive radio networks with average interference power constraints*. IEEE Trans. Wireless Commun., 8(4):1620–1625, 2009.

[86] J. Van Hecke, P. Del Fiorentino, F. Giannetti, V. Lottici, L. Vandendorpe, and M. Moeneclaey. *Resource allocation for multicarrier cooperative cognitive radio networks with imperfect channel state information*. In Proc. IEEE Personal Indoor and Mobile Radio Communications (PIMRC), pages 627–632, September 2014.

[87] Hisashi Tanizaki. *Computational methods in statistics and econometrics*. Dekker, Abingdon, 2004.

[88] Wei Yu and R. Lui. *Dual methods for nonconvex spectrum optimization of multicarrier systems*. IEEE Trans. Commun., 54(7):1310 –1322, July 2006.

[89] Kibeom Seong, M. Mohseni, and J.M. Cioffi. *Optimal resource allocation for OFDMA downlink systems*. In Proc. IEEE Int. Symposium on Information Theory (ISIT), pages 1394–1398, July 2006.

[90] Harold W. Kuhn. *The hungarian method for the assignment problem*. Naval Research Logistics Quarterly, 2:83–97, 1955.

[91] Yingwei Yao and G.B. Giannakis. *Rate-maximizing power allocation in OFDM based on partial channel knowledge*. IEEE Trans. Wireless Commun., 4(3):1073–1083, 2005.

[92] J. Francis and N. B. Mehta. *EESM-based link adaptation in point-to-point and multi-cell OFDM systems: modeling and analysis*. IEEE Trans. Wireless Commun., 13(1):407–417, January 2014.

[93] J. Francis and N. B. Mehta. *Characterizing the impact of feedback delays on wideband rate adaptation*. IEEE Trans. Wireless Commun., 14(2):960–971, February 2015.

[94] Hui Song, R. Kwan, and Jie Zhang. *Approximations of EESM effective SNR distribution*. IEEE Trans. Commun., 59(2):603–612, February 2011.

[95] R. Giuliano and F. Mazzenga. *Exponential effective SINR approximations for OFDM/OFDMA-based cellular system planning*. IEEE Trans. Wireless Commun., 8(9):4434–4439, September 2009.

[96] J. Van Hecke, P. Del Fiorentino, R. Andreotti, V. Lottici, F. Giannetti, L. Vandendorpe, and M. Moeneclaey. *Accurate modeling of the predicted kESM-based link performance metric for BIC-OFDM systems*. In Proc. IEEE Symposium on Communications and Vehicular Technology in the Benelux (SCVT), November 2015.

[97] Wan Choi, N. Himayat, S. Talwar, and M. Ho. *The effects of co-channel interference on spatial diversity techniques*. In Proc. IEEE Wireless Communications and Networking Conference (WCNC), pages 1936–1941, March 2007.

[98] Erik Dahlman, Stefan Parkvall, Johan Skold, and Per Beming. *3G evolution, second edition: HSPA and LTE for mobile broadband*. Academic Press, 2 edition, 2008.

[99] R. Kwan and C. Leung. *On the applicability of the pearson method for approximating distributions in wireless communications*. IEEE Trans. Commun., 55(11):2065–2069, November 2007.

[100] S.N. Donthi and N. B. Mehta. *An accurate model for EESM and its application to analysis of CQI feedback schemes and scheduling in LTE*. IEEE Trans. Wireless Commun., 10(10):3436–3448, October 2011.

[101] Thomas M. Cover and Joy A. Thomas. *Elements of information theory 2nd edition (Wiley series in telecommunications and signal processing)*. Wiley-Interscience, July 2006.

[102] Jianhua Lin. *Divergence measures based on the Shannon entropy*. IEEE Trans. Inf. Theory, 37(1):145–151, January 1991.

[103] J. Van Hecke, P. Del Fiorentino, R. Andreotti, V. Lottici, F. Giannetti, L. Vandendorpe, and M. Moeneclaey. *Goodput-maximizing resource allocation in cognitive radio BIC-OFDM systems with DF relay selection*. In Proc. IEEE Int. Conf. on Communications (ICC), pages 1404–1409, June 2015.

[104] J. Van Hecke, P. Del Fiorentino, R. Andreotti, V. Lottici, F. Giannetti, L. Vandendorpe, and M. Moeneclaey. *Adaptive coding and modulation using imperfect CSI in cognitive BIC-OFDM systems*. EURASIP J. Wirel. Commun. Netw., 2016(1):256:1–256:14, 2016.

[105] I. Stupia, F. Giannetti, V. Lottici, and L. Vandendorpe. *A greedy algorithm for goodput-based adaptive modulation and coding in BIC-OFDM systems*. In Proc. European Wireless Conference (EW), pages 608–615, April 2010.

[106] P.L. Nuaymi. *WiMAX: technology for broadband wireless access*. John Wiley & Sons, New York, 2007.

[107] P. Del Fiorentino, R. Andreotti, V. Lottici, F. Giannetti, J. Van Hecke, and M. Moeneclaey. *Link resource adaptation for BIC-OFDM systems with outdated channel state information*. In Proc. European Wireless Conference (EW), pages 1–6, May 2014.

[108] B. Farhang-Boroujeny. *OFDM versus filter bank multicarrier*. IEEE Signal Process. Mag., 28(3):92–112, May 2011.

[109] P. Del Fiorentino, C. Vitiello, V. Lottici, E. Debels, J. Van Hecke, M. Moeneclaey, F. Giannetti, and M. Luise. *Resource allocation in short packets BIC-UFMC transmission for internet of things*. In IEEE Globecom Workshops (GC Wkshps), pages 1–6, December 2016.

[110] E. Debels, P. Del Fiorentino, C. Vitiello, J. Van Hecke, F. Giannetti, M. Luise, V. Lottici, and M. Moeneclaey. *Adaptive modulation and coding for BIC-UFMC and BIC-OFDM systems taking CFO into account*. In Proc. IEEE Symposium on Communications and Vehicular Technology in the Benelux (SCVT), pages 1–5, November 2016.

[111] Xiaodong Li and J. A. Ritcey. *Bit-interleaved coded modulation with iterative decoding*. IEEE Commun. Lett., 1(6):169–171, November 1997.

[112] L. Hanzo, T.H. Liew, B.L. Yeap, R.Y.S. Tee, and S.X. Ng. *Turbo coding, turbo equalisation and space-time coding: EXIT-chart-aided near-capacity designs for wireless channels*. Wiley - IEEE. Wiley, 2011.

[113] Ilan Abramovici and Shlomo Shamai. *On turbo encoded BICM*. Annales Des Télécommunications, 54(3):225–234, 1999.

[114] N. Aerts, I. Avram, J. Van Hecke, H. Bruneel, and M. Moeneclaey. *Iterative SAGE-based channel estimation in a block fading amplify-and-forward relaying network*. IEEE Trans. Wireless Commun., 13(4):1742–1753, April 2014.