UniversidadeVigo

Link Adaptation Techniques for Future Terrestrial and Satellite Communications
2019 DOCTORAL DISSERTATION Anxo Tato Arias

DOCTORAL DISSERTATION

Link Adaptation Techniques for Future Terrestrial and Satellite Communications

Author:
Anxo Tato Arias

Supervisor:
Carlos Mosquera Nartallo

2019

International Mention

UniversidadeVigo

EIDO
Escola Internacional
de Doutoramento

Universidade Vigo

International Doctoral School

Anxo Tato Arias

DOCTORAL DISSERTATION

Link Adaptation Techniques for Future Terrestrial and Satellite Communications

Supervised by:

Carlos Mosquera Nartallo

2019

International Mention

Universidade Vigo

International Doctoral School

Carlos Mosquera Nartallo

DECLARES that the present work, entitled "Link Adaptation Techniques for Future Terrestrial and Satellite Communications," submitted by Anxo Tato Arias to obtain the title of Doctor, was carried out under his supervision in the PhD program on "Information and Communications Technology" (Doc_TIC). Moreover, he authorizes the submission for the International Mention.

Vigo, November 25, 2019.

The supervisor,

Prof. Carlos Mosquera Nartallo

Esta tese foi financiada polas axudas de apoio á etapa predoutoral nas universidades do Sistema universitario galego, nos organismos públicos de investigación de Galicia e noutras entidades do Sistema galego de I+D+i, cuxo financiamento procede do Fondo Social Europeo nun 80 % e no 20 % restante da Secretaría Xeral de Universidades, pertencente á Consellería de Cultura, Educación e Ordenación Universitaria da Xunta de Galicia.

UNIÓN EUROPEA

FONDO SOCIAL EUROPEO

"O FSE inviste no teu futuro"

XUNTA DE GALICIA

CONSELLERÍA DE CULTURA, EDUCACIÓN
E ORDENACIÓN UNIVERSITARIA
Secretaría Xeral de Universidades

Agradecementos

Agora que a etapa do doutoramento está chegando á súa fin é un bo momento para botar a mirada atrás e agradecerlle a todas as persoas que ao longo destes anos me axudaron dunha ou doutra maneira, facendo posible ou facendo máis levadeiro este camiño. Poida que non estean todos os que son, desculpade se se me escapou algunha persoa polas présas por depositar, pero definitivamente son todos os que están.

En primeiro lugar quero darlle as grazas ao meu director de tese, Carlos Mosquera, pola súa xenerosa oferta, pola súa entregada e constante dedicación, por todo o que me ensinou, polo seu bo exemplo e por todo o apoio, axuda e comprensión que me brindou durante todo este tempo que estiven traballando con el.

Grazas tamén á Xunta de Galicia pola gran oportunidade que dou coa axuda predoutoral, que me permitiu dedicarme en exclusiva ao meu proxecto de investigación durante estes anos, sen máis preocupación que intentar facer ben o meu traballo.

I want to thank also the European Space Agency, that by means of the SatNEx project allowed me to do two research stays, one in the CTTC in Catalunya and another in the SnT, in Luxembourg. In both places I met excellent professionals which helped me a lot and fostered new and fruitful research lines. A special mention goes to Stefano Andrenacci, Symeon Chatzinotas and Eva Lagunas, for all the support they gave me while I was working in the SnT, and to Ana Pérez-Neira and Pol Henarejos, for accepting receiving me as a visitor when I was still a young and inexperienced researcher.

Investigar é máis sinxelo cando tes un gran equipo que che dá soporte e facilita o teu traballo, como é o caso de atlanTTic. Grazas especialmente a Carmen, Susana, Fran, María, Gloria... Tampouco me quero esquecer de Laura e Silvia do CTTC, por axudarme tanto coas visitas a Castelldefels e Luxemburgo.

Non quero esquecerme tampouco dos meus compañeiros do GPSC, grazas por deixarvos liar tan facilmente para os proxectos de SDR. Grazas a Tomás, Alberto e a Noelia por todo o apoio e compañía mutua durante o noso período de doutoramento. E grazas tamén ao resto de compañeiros e profesores do grupo de investigación, Roberto, Fernando, Vlad, Miguel, os Javis, David, Vahid, Khawar, Marcos, Iván...

As horas no CUVI pasan máis rápido cando a compañía é boa. Grazas aos compañeiros de Gradiant da Ceremonia e aos colegas de Filoloxía por eses xantares e esas tan boas sobremesas. Thanks to my friends Oltjon, João and all the Researchers' house mates, you helped to make my months in Luxembourg a wonderful time.

Rocío, Xosé, Miguel e Mandianes, sen o voso empuxón igual non tería empezado este camiño, grazas de corazón por animarme a facer a tese. E por último, quero agradecerlle aos meus amigos e amigas e á miña familia por estar comigo ao longo deste tempo, o voso cariño e a vosa compañía son o mellor dos tesouros.

Abstract

The increasing demand of access to data from the users and the enormous number of connected devices requires to enhance the capacity of the wireless networks. They must provide a higher throughput to serve all the requested traffic and they must accommodate the vast number of the Internet of the Things (IoT) devices. In this context, this thesis focus its attention on three different scenarios which have in common that they are a future evolution of current terrestrial and satellite communications systems. These scenarios are Mobile Satellite Systems (MSS), Fixed Satellite Systems (FSS) and next generation 5G networks.

The adoption of Dual Polarization (DP) in MSS along with Multiple Input Multiple Output (MIMO) signal processing techniques allows to double the capacity of previous systems with the same bandwidth and transmit power. On the other hand, the shift to more aggressive frequency reuse patterns in FSS can also provide remarkable gains in the capacity of High Throughput Satellites (HTS) for offering BroadBand Satellite Services (BBS). Linear precoding stands out as a technique to cope with the high level of interference which arises in this scenario. Lastly, energy efficient modulation schemes, like Spatial Modulation (SM) and its many variants, are being proposed for increasing the capacity of future 5G networks since they represent a good trade-off between spectral efficiency, energy efficiency and transmitter complexity.

Adaptive Coding and Modulation (ACM) technology is omnipresent in most of the communication standards since it enables a better exploitation of the system capacity by means of the adaptation of the Modulation and Coding Scheme (MCS). The link adaptation algorithm is responsible for selecting the optimum MCS, as well as other physical layer parameters in some cases, to adapt the transmission bit rate according to the instantaneous channel capacity of the time variant channel. Thus, link adaptation algorithms permit to increase the spectrum efficiency and guarantee a robust communication, adapting the level of redundancy of the coded information

bits and the ruggedness of the modulation scheme.

In this thesis, several link adaptation algorithms are proposed for the three considered scenarios and its effectiveness is supported with simulations. Furthermore, an experimental validation of some algorithms is provided using a real satellite link, implemented with Software Defined Radio (SDR) technology. It is also studied how the carrier detection errors in the Channel State Information (CSI) affect linear precoding in FSS. The errors in the users Signal to Interference and Noise Ratio (SINR) that the gateway calculates to allocate MCS to the users is analyzed statistically and geographically. Moreover, a link adaptation algorithm with an adaptive margin per user is shown to allow a robust communication in the presence of SINR errors whereas the throughput of the system is barely compromised. In addition, a new method for making capacity calculations in SM and Generalized SM systems based on a neural network is proposed, which improves both accuracy and computational complexity with regard to the existing analytical approximations in the literature. With regard to MSS using DP, an adaptation mechanism is proposed in order to select the optimum MIMO mode and MCS which offer the highest throughput. Simulation results in a maritime mobile satellite channel show the dependence of the optimum MIMO mode with the average Signal to Noise Ratio (SNR), and how the spectral efficiency can be maximized whereas a target outage probability can be guaranteed. Lastly, the adaptation in SM systems is also addressed and several methods for deciding the coding rate in SM are given. These include the computation of the capacity prior to the adaptation, and also the use of a deep neural network. The latter offers very good results, with a spectral efficiency very close to the maximum achievable value.

Contents

List of Figures

List of Tables

List of Abbreviations

2G	Second generation
3G	Third generation
4G	Fourth generation
5G	Fifth generation
ACK	Acknowledgment
ACM	Adaptive Coding adn Modulation
ADC	Analog to Digital Converter
AWGN	Additive White Gaussian Noise
BBS	BroadBand Satellite Service
BCH	Bose-Chaudhuri-Hochquenghem
BER	Bit Error Rate
BPCU	Bits Per Channel Use
BS	Base Station
CCDF	Complementary Cumulative Distribution Function
CL	Close lop
CPU	Central Processing Unit
CQI	Channel Quality Indicator
CSI	Channel State Information
CSIR	Channel State Information at the Receiver
CSIT	Channel State Information at the Transmitter
DL	Downlink or Deep Learning
DP	Dual Polarization
DVB	Digital Video Broadcasting
FEC	Forward Error Correction
FECFRAME	FEC Frame
FER	Frame Error Rate
FFR	Full Frequency Reuse
FPGA	Field Programmable Gate Array
FR2	Two color frequency reuse
FR4	Four color frequency reuse
FSS	Fixed Satellite Services
G/T	Gain to Noise Temperature factor
GEO	Geostationary Earth Orbit

GS	Ground Station
GSM	Generalized Spatial Modulation
HTS	High Throughput Satellite
ICT	Information and Communication Technologies
IM	Index Modulations
IoT	Internet of the Things
ITU	International Telecommunication Union
LDPC	Low Density Parity Check
LHCP	Left Hand Circular Polarization
LLR	Log-Likelihood Ratio
LMS	Land Mobile Satellite or Least Mean Squares
LNB	Low Noise Block
MCS	Modulation and Coding Scheme
MEO	Medium Earth Orbit
MFNN	Multilayer Feedforward Neural Network
MI	Mutal Information
MIMO	Multiple Input Multiple Output
ML	Maximum Likelihood or Machine Learning
MMSE	Minimum Mean Square Error
MODCOD	Modulation and Coding Scheme
MPA	Multi-Port Amplifier
MS-GSM	Multi-Symbol Generalized Spatial Modulation
MSE	Mean Square Error
MSS	Mobile Satellite Services
MT	Mobile Terminal
MU-MIMO	Multi-user MIMO
NAK	Non-Acknowledgment
NGO	Non Geostationary Orbit
NN	Neural Network
OFDM	Orthogonal Frequency Division Multiplexing
OL	Open Loop
OLLA	Outer Loop Link Adaptation
OPTBC	Orthogonal Polarization-Time Block Coding
PDF	Probability Density Function
PLA	Physical Layer Abstraction
PLL	Phase-Locked Loop
PLPC	Per Line Power Constraint
PMod	Polarized Modulation
QEF	Quasi Error Free
QPSK	Quadrature Phase Shift Keying
RBIR	Received Bit Mutual Information Rate
RF	Radio Frequency
RHCP	Right Hand Circular Polarization
S-UMTS	Satellite component of UMTS (Universal Mobile Telecommunications System)

SDR	Software Defined Radio
SHT	Sequential Hypothesis Testing
SINR	Signal to Interference and Noise Ratio
SM	Spatial Modulation
SMX	Spatial Multiplexing
SNR	Signal to Noise Ratio
SP	Single Polarization
SPC	Sum Power Constraint
SS-GSM	Single-Symbol Generalized Spatial Modulation
TWTA	Traveling-Wave Tube Amplifier
UAV	Unmanned Aerial Vehicle
USRP	Universal Software Radio Peripheral
UT	User Terminal
V-BLAST	Vertical BLAST (Bell Laboratories Layered Space-Time)
WH	Walsh-Hadamard
XPD	Cross-Polar-Discrimination
ZF	Zero Forcing

Chapter 1

Introduction

According to Cisco's Global Mobile Data Traffic Forecast Update [69], the mobile data traffic grew 71 % in 2017 and a seven-fold increment is expected in the period $2017 - 2022$, with the traffic growing at a Compound Annual Growth Rate (CAGR) of 46 % during that period. Therefore, the capacity of the mobile networks needs to be increased in order to satisfy all the traffic demanded by the users. Furthermore, fixed networks, both terrestrial and satellite, need also to enhance their performance to offer a higher throughput to the users, which demand mainly multimedia content of an increasing quality.

On the other side, the energy consumption of the Information and Communication Technology (ICT) sector represented already around 2 % of the global carbon emissions in 2007, with mobile networks contributing about 0.2 % [31]. Moreover, the footprint of mobile communications is expected to triple that value in 2020 [36]. Apart from the ecological point of view, the energy consumption of Base Stations (BSs) serving mobile users constitutes a significant part of the operational costs of the telecom providers. For all these reasons, it is essential to increase not only the capacity of the networks, but also their energy efficiency.

Regarding the new type of connected devices, the aforementioned Cisco white paper forecasts that the number of Machine-to-Machine (M2M) connections will see a four-fold increment in the $2017 - 2022$ period, reaching 3.9 billion of connections in 2022. Many M2M devices of the Internet of the Things (IoT) only run with batteries which, sometimes, must last during the whole life cycle of the device. This pushes a strong optimization of the power consumption of the devices, including their communications.

Spectrum is a very scarce resource which is shared by many different

players and systems. Therefore, it has to be optimized to provide the best service with the assigned bandwidth. A very common metric to measure the spectrum harnessing is the spectral efficiency, which relates the transmitted throughput in bits/s with the occupied bandwidth in Hz. The objective of the link adaptation algorithms presented in this thesis is precisely the leverage of the time variant channel capacity to increase the spectral efficiency of the communications link.

This thesis addresses then the implementation of algorithms for adaptive communications in three different scenarios, which have in common that they are a future evolution of current terrestrial and satellite communications systems. These scenarios are Mobile Satellite Systems (MSS), Fixed Satellite Systems (FSS) and next generation 5G networks. In the three future scenarios, the capacity of the system is raised by means of different techniques. The adoption of Dual Polarization (DP) in MSS, the shift to Full Frequency Reuse (FFR) in FSS and the adoption of energy efficient multiantenna techniques in 5G networks are considered in this thesis. The latter scenario considers the energy consumption of the 5G communications, not only the capacity enhancement, important for extending the battery life of IoT devices and reducing the carbon footprint of 5G BSs.

In the following paragraphs, it is explained how terrestrial and satellite systems can evolve in order to enhance their performance. The explanation is focused on these three scenarios that are covered later during this thesis. Afterwards, the motivation of the link adaptation techniques in communications is given, highlighting that they allow to increase the spectral efficiency of the systems by means of a better utilization of the channel capacity. Finally, an overview of the thesis is provided.

1.1. Future terrestrial and satellite systems

Communication systems evolve continuously to address people's demands. Thus, transitions from second generation (2G) to fifth generation (5G) in digital cellular technologies happened in the last thirty years, passing through 3G and 4G. On the other hand, the DVB consortium approved the standards DVB-S, DVB-S2 and DVB-S2(X) for satellite communications within the period of nineteen years, from 1995 to 2014. Users demand high speed Internet connections and an ubiquitous connectivity; therefore, the capacity of the networks should be increased to satisfy all the traffic demand and wireless communications systems are required to provide an Internet connection on the move. Moreover, satellite communications are necessary to cover all the places without terrestrial networks, as some rural

and land remote areas, sea and air. Furthermore, communication systems need not only support the human based communications, but also an increasing number of connected devices to embrace the IoT paradigm.

There are several ways to enhance current communication systems in order to satisfy the increasing volume of traffic. This thesis focuses his attention on three enhancements: one for Mobile Satellite Services (MSS) at low frequency bands, another for MSS or Fixed Satellite Services (FSS) which operate at higher frequencies, as Ka band, and a latter consisting on a new modulation scheme proposed for energy efficient 5G networks. These three scenarios are depicted in Fig. 1.1.

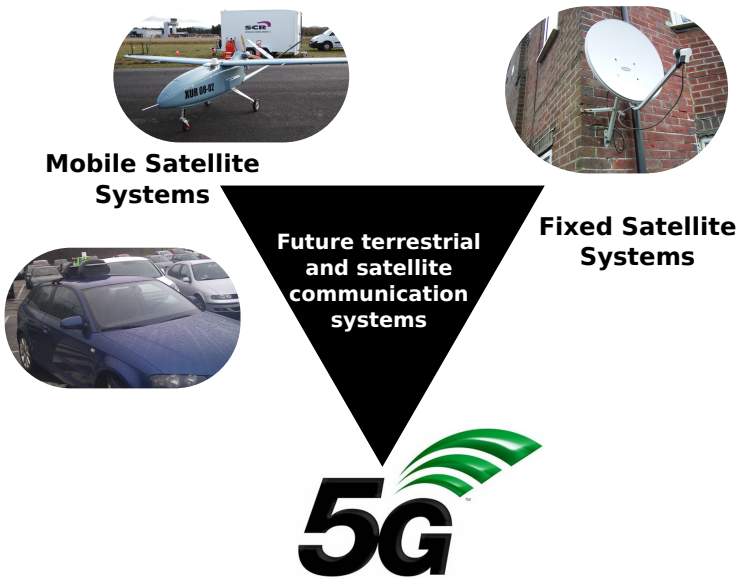


Figure 1.1: The three communication scenarios addressed in this thesis.

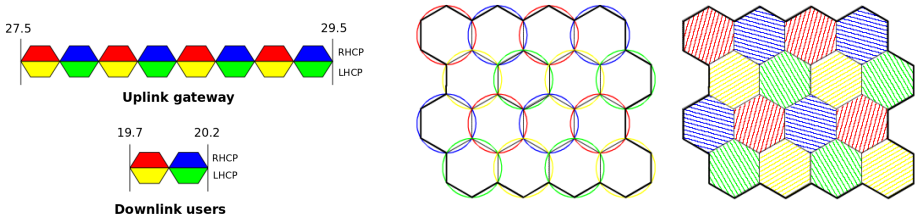
Firstly, MSS have a paramount importance for serving on the move users which are located in places outside the coverage of terrestrial systems. These can be, for example, vessels, commercial aircrafts, Unmanned Aerial Vehicles (UAVs), transportation vehicles or users from media or NGOs. MSS operate in L-band (1-2 GHz) and S-band (2-4 GHz) and they use typically circular polarization, which is preferred at low frequencies over linear polarization to avoid the Faraday rotation effect. MSS, as those offered using the standard S-UMTS family SL [9], usually rely on a single circular polarization, typically Right Hand Circular Polarization (RCHP), due to the fear of small Cross-Polar-Discriminations (XPD) [75]. However, if Dual Polarization (DP) antennas are employed at both satellite and Mobile Terminal (MT), significant capacity gains can be achieved.

The simultaneous use of two orthogonal polarizations, RHCP and Left Hand Circular Polarization (LHCP) in MSS is possible thanks to the development of DP antennas for these bands [74]. A DP satellite system which employs both RHCP and LHCP simultaneously to communicate with the users is analogous to a 2×2 Multiple-Input-Multiple-Output (MIMO) system, simply replacing the spatial by the polarization component. Therefore, the application of signal processing techniques to the DP satellite systems allows to increase their capacity, which can be used to serve more users or increase their throughput, or to extend their operating range, thanks to transmit diversity.

In [48], the use of DP is proposed for mobile satellite systems, with [47] analyzing in more detail a system exploiting a specific type of modulation named Polarized Modulation (PMod). This thesis proposes adaptation techniques for this new DP satellite system, with applications to mobile communications. The adoption of DP can double the capacity of the system at high Signal to Noise Ratios (SNRs) and, on the other hand, can be used to extend the operating SNR range, allowing robust communications at lower SNRs than a Single Polarization (SP) system with the same transmit power. The gains DP satellite systems provide can benefit also IoT devices located in remote places [28].

At higher frequencies, in Ku and Ka-band, High Throughput Satellites (HTS) [62] aim at providing Fixed Satellite Services (FSS), as BroadBand Satellite Service (BBS), by means of many spot beams over the coverage area, delivering a high system capacity and a high user throughput. These systems started to be deployed with a four color frequency reuse (FR4), where the available bandwidth for the user beams is split into two frequency bands and two polarizations, RHCP and LHCP. Fig. 1.2a shows how the available spectrum in the Ka-band is assigned to the gateway and the users in the forward link. On the other hand, Fig. 1.2b depicts a satellite footprint with 16 user beams, where it can be seen how the same color (i.e., the same polarization-frequency), is spatially reused for different non-contiguous beams.

The four color frequency reuse reduces the inter-beam interference reusing the same portion of the spectrum in geographically separated beams; however, in each beam only one fourth of the total available bandwidth is employed. More aggressive frequency reuse schemes are possible, as 2-color frequency reuse (FR2) of full-frequency reuse (FFR). In the latter, all the available spectrum in both polarizations is employed in all the beams. Therefore, the capacity can be ideally multiplied by four. However, due to the radiation pattern of the satellite antennas a high level of inter-beam interference arise with FFR, thus reducing the total capacity. In order to leverage



(a) Forward link frequency plan

(b) Satellite coverage

Figure 1.2: Example of a Ka-band multibeam satellite system with 4 color frequency reuse.

the benefits from FFR and reduce the interference several techniques as precoding or Non Orthogonal Multiple Access (NOMA) are being proposed [106], [73].

Several works propose the application of linear precoding to the forward link of DVB-S2X based systems, as [103], [24], [91], [102] and [100]. Precoding satellite systems require the users to estimate the magnitude and phase of the desired signal and the interfering signals from the neighbor beams and report this Channel State Information (CSI) to the gateway. With this information, the gateway precompensates the transmitted signals in order to reduce the interference and improve the quality of the signal that the users receive. In this case, the satellite scenario is modeled as a Multiuser MIMO (MU-MIMO) channel and the information streams of symbols sent to each beam are pre-multiplied by a matrix calculated with the CSI. Thus, this process is transparent for the satellite terminals, which only need to estimate and report the CSI. The effects of the CSI degradation and how to counteract them will be addressed in this thesis.

Regarding mobile terrestrial communications, in the last years the research community devoted considerable effort to the definition and development of 5G. Under the umbrella of 5G, different communications technologies are foreseen in order to satisfy the diverse requirements of 5G set by the International Telecommunication Union (ITU). These embrace, among others, gigabit connections, very low latency, enormous number of connected devices, reduction in network energy usage by almost 90%, and high battery life especially for low power devices [10]. In this context, a new family of modulation schemes named Spatial Modulation (SM) appear to provide capacity increments with low transmit complexity and high energy efficiency, to reduce consumption at both the base station and terminals [61], [20].

SM is a very simple modulation scheme which forms part of the broad

family of Index Modulations (IM). SM is a multiantenna modulation technique that, contrary to other schemes as Spatial Multiplexing (SMX), it only requires one Radio Frequency (RF) chain in its simplest form. Although with SMX a much higher spectral efficiency can be achieved, since it grows linearly with the number of antennas, SM offers a good trade-off between spectral efficiency, energy efficiency and complexity. Hence, it is being proposed for the future air interfaces of the IoT devices [30]. Fig. 1.3 shows a block diagram of a SM transmitter, where it can be observed that there is only one RF chain which is connected by means of a switch with all the antennas.

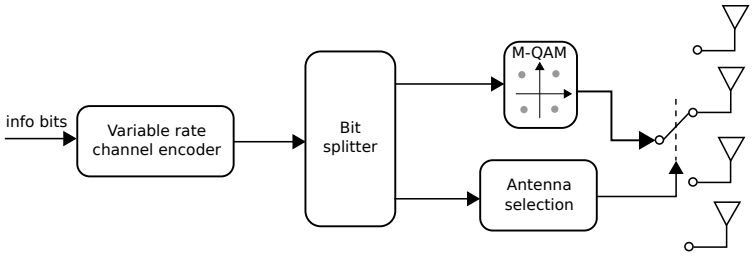


Figure 1.3: Block diagram of an adaptive Spatial Modulation transmitter.

Despite its simplicity, SM presents some interesting challenges. For example, there is not an analytical expression for its capacity, as opposed to SMX, and the adaptation of some physical layer parameters, as the coding rate, was not studied previously. This thesis presents solutions for these two challenges. On the one hand, it proposes a new method for obtaining the capacity of SM based on Machine Learning (ML). And, on the other hand, several coding rate adaptation methods are proposed for SM.

So far, we have outlined how current terrestrial and satellite communications systems are evolving to improve their performance and satisfy the users' demand. In mobile satellite systems which operate at low frequency bands, the jump from single to double polarization permits capacity improvements, very necessary in these congested frequencies. Also in satellite communication but at higher bands, HTS represent an attempt of the satellite industry for offering broadband connections to the consumers users. The adoption of more aggressive frequency reuse patterns in these multibeam satellite systems enables also capacity increments compared with the traditional four color reuse schemes. Lastly, new modulation schemes like SM help future 5G networks to achieve its goals of energy efficiency and capacity enhancements for some use cases.

1.2. Link adaptation

This thesis is focused on the development of algorithms for adaptive communications in the three scenarios outlined in the previous paragraphs: MSS with DP, FSS with FFR and 5G networks using SM. Moreover, it offers an experimental validation of link adaptation algorithms for a baseline MSS with single polarization using a real Medium Earth Orbit (MEO) satellite and implementing the two communication ends with Software Defined Radio (SDR) technology.

Communication systems, wireless specially, need to adapt their transmission parameters in order to offer a good performance since practical channels are time variant. The mobile satellite channel has several states, of good and bad channel conditions, depending on the the good or bad satellite visibility that the Mobile Terminal (MT) has while it is moving [38]. Moreover, if the orbit of the satellite is Low Earth Orbit (LEO) or MEO, the different satellite elevations and the varying distance from the satellite to the MT causes temporal variations of the level of received signal. In FSS operating at Ka band the main source of channel quality variation is the rain attenuation which can reach values as high as 12 dB [52]. Lastly, in 5G cellular systems, not only the level of the received signal changes, but also the paths followed by the signals from the array of transmit antennas to the receive antennas, causing time variations of the MIMO channel matrix when the users are moving.

The Signal to Noise Ratio (SNR) or Signal to Interference and Noise Ratio (SINR) are very common metrics to measure the quality of the received signal. They are defined as the ratio between the power of the received signal and the noise power (together with the power of the interference, in the case of the SINR). From the Shannon capacity formula, which relates the SNR with the capacity C in a Additive White Gaussian Noise (AWGN) channel,

$$C = B \log_2(1 + \text{SNR}) \quad (bits/s), \tag{1.1}$$

it can be seen clearly that a variation in the received SNR affects the capacity of the channel, i.e., the maximum amount of information that the transmitter can send reliably through the communication channel. In order to avoid its dependence with the amount of bandwidth B employed by the system, it is very common to utilize the spectral efficiency η, which is obtained by dividing the capacity by the bandwidth:

$$\eta = \frac{C}{B} = \log_2(1 + \text{SNR}) \quad (bits/s/Hz). \tag{1.2}$$

Moreover, in MIMO channels the capacity depends also on the channel matrix $\mathbf{H}$ and its variations have also an impact on the final channel capacity.

Thus, the capacity of a communications link is also time variant; the amount of information that can be conveyed changes with time. If a transmitter wants to fully exploit the channel, it should adapt the transmission bit rate according to the channel capacity variations. The alternative, the use of a fixed bit rate, is tremendously inefficient since to guarantee the robustness of the communication, the transmission bit rate should take a low value, set by the worst case channel conditions. Therefore, when the capacity of the channel is high, the transmission with that low bit rate implies a waste of the channel capacity, which could be used to provide a link with higher throughput.

A transmitter can employ different mechanisms for adapting the transmission bit rate and follow the time variations of the channel capacity. The standard S-UMTS family SL [9] allows a variation of the transmitted power, the symbol rate (and hence the bandwidth), the modulation order and the coding rate. In other satellite systems, as those using DVB-S2(X) [6], the symbol rate tends to be fixed and the transmitter selects the Modulation and Coding Scheme (MCS) of each frame, i.e. the constellation and the coding rate. In some terrestrial systems, like LTE [8] or Wi-Fi [5], apart from selecting the MCS, the transmitter chooses one of the possible transmission MIMO modes or the number of streams sent.

The term link adaptation refers to the procedure of adjusting some of the physical layers parameters, as the modulation and the coding, in line with the quality of the link, as it is depicted in Fig. 1.4. Adaptive Coding and Modulation (ACM) is another term to refer to the link adaptation, emphasizing that these are the parameters which are adapted. Most of the communication standards consider some sort of ACM, from terrestrial systems as for example Wi-Fi, WiMAX [3] or LTE, to satellite systems, as those using S-UMTS or DVB-S2(X).

With ACM, a transmitter can select the level of protection of the information bits depending on the quality of the link. High order constellations, like 64-QAM, offer high spectral efficiencies which can be used when the receive SNR is high. On the other hand, by using low order constellations, as QPSK, the information bits are more protected and a robust communication is guaranteed in the case of a link with low SNR. Apart from varying the modulation order, the transmitter can change also the coding rate of the channel encoder. Thus, high coding rates, close to 1, allow to send more information bits since the redundancy bits are reduced thanks to the good channel conditions. And, on the other hand, when the channel quality is not so good, the transmitter can reduce the coding rate, which implies to increase the level of redundancy and protect more the information bits at the expense of reducing the spectral efficiency. If a constellation of M symbols

and a coding rate r is employed, the maximum spectral efficiency η that can be achieved is

$$\eta = r \log_2 M \quad (bits/s/Hz). \tag{1.3}$$

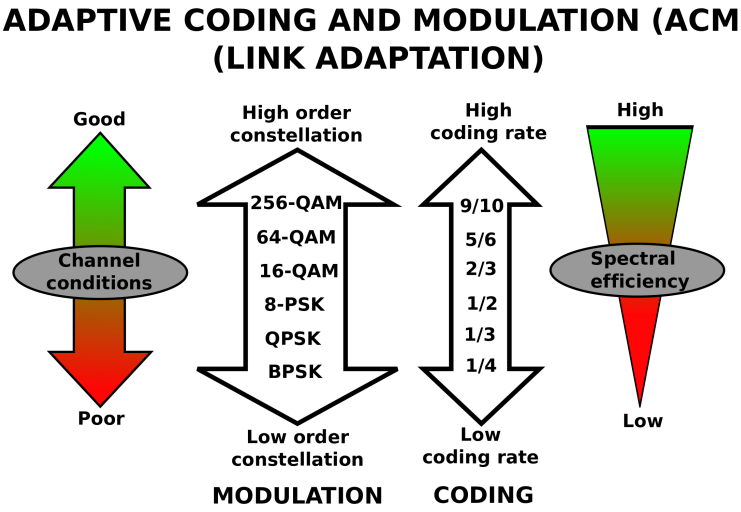


Figure 1.4: Diagram to explain the principles behind the link adaptation.

The implementation of some sort of ACM requires typically some feed-back from the receiver to the transmitter. The fed-back information can be the result of the frame decoding (an Acknowledgment [ACK] if the frame was correctly decoded or a Negative ACK [NAK] if the received could not decode the frame), the SNR of the link, some Channel Quality Indicator (CQI) or the MIMO mode or MCS that the transmitter should use. In some cases, the adaptation can be performed without any feedback if a high level of reciprocity between the forward and return links is assumed in a bidirectional communication. In this thesis, in general we consider that the receiver feeds back some information to help the transmitter in the MCS selection process.

There are different alternatives to perform the link adaptation. A very common approach is named Outer Loop Link Adaptation (OLLA), which consists on the use of some Look-up Tables (LUTs) which map SNR ranges or a CQI to each one of the available MCS, together with a margin which is adapted based on the ACK/NAKs the receiver feeds back. In the classic OLLA algorithm, the margin is increased when errors in the link happen and it is reduced otherwise. This is the approach followed in [16], [76], [78]. Some enhancements to the classic algorithm are provided in [21] and [29] to improve the convergence speed of the algorithm.

In complex systems with many degrees of freedom, it is common to resort to some Machine Learning (ML) algorithms. For example, in [79] a Support Vector Machine (SVM) is employed to select the physical layer parameters of a MIMO-OFDM system. In that case, a supervised learning approach is used. More recently, some works apply the ML paradigm of the reinforcement learning, as [37] and [44]. In this thesis, algorithms based on supervised learning with neural networks and OLLA are applied to the future evolution of terrestrial and satellite communication systems. A radically different approach, which is not explored in this thesis, consists on the formulation of the MCS selection as a multi-armed bandit problem, as it is done in [82].

Now that we have explained the concept of link adaptation and the three scenarios addressed in the thesis, its contents are summarized hereafter. Firstly, this thesis starts providing results from an experimental campaign for validating some link adaptation algorithms introduced in [76]. A complete S-band bidirectional satellite communications link was built and, by means of a real MEO satellite working in transparent mode, a Mobile Terminal (MT) established a connection with a Ground Station (GS). The transceivers were implemented using SDR in two Universal Software Radio Peripherals (USRPs). With them, the performance of different link adaptation algorithms was tested. Apart from using the ACKs and NAKs received from the other end, the algorithms employ different strategies for including the closed loop SNR (the CSI fed back from the other end) and the open loop SNR (the CSI measured directly by the terminal) in the adaptation process.

From the user point of view, the main application of these results is the automatic bit rate adaptation of a mobile satellite link connection which can benefit on the move satcom users. In the field trials, both aeronautical and terrestrial channels were tested, with the MT boarded on top of a car and in an Unmanned Aerial Vehicle (UAV). Although it was not tested, the MT could be also placed on a boat to serve maritime users.

The application of linear precoding for multibeam satellite systems is being proposed for increasing the capacity of High Throughput Satellites (HTS) and compete better with terrestrial alternatives of broadband services. Moreover, only satellite communications can offer Internet connection on the maritime and aeronautical markets. In the last years there has been an intense research work in the field of precoding for HTS [103], [24], [91], [102] and [100]. However, there is a practical issue which was not studied in the literature before: how a strongly degraded Channel State Information (CSI) affects the system. Namely, we characterize statistically and geographically the errors in the users SINR calculated by the gateway with

the imperfect CSI. The main cause of CSI degradation is not the estimation errors, but the nullification of many of the CSI vector components, since the satellite receivers can only estimate a small part of the interfering channels from the surrounding beams.

The errors in the SINR estimated by the gateway are very problematic. If the SINR is underestimated, the gateway assigns MCS with lower spectral efficiency than the optimum MCS for the users channel condition. And, even worst, if the SINR is overestimated, the gateway assigns to the users MCS which cannot be correctly decoded, increasing the Frame Error Ratio (FER) and causing a considerable drop of the system capacity. In this thesis, it is shown how to overcome this problem with a link adaptation algorithm with an adaptive margin per user. Thus, the margin accounts for the SINR errors avoiding to apply global fixed margins, thus preventing the decoding errors but at the expense of a high throughput degradation in the system.

This thesis also addresses the problem of the capacity calculation of some new modulation schemes which are being proposed for next generation communication systems, such as Spatial Modulation (SM) and Generalized Spatial Modulation (GSM). The capacity calculation has a practical interest for developing link adaptation algorithms for future 5G systems using these modulation schemes. If the transmitter knows the capacity of the SM or GSM link, it can adapt its transmit bit rate in order to make a better exploitation of the channel capacity. To obtain the capacity of these systems a novel approach is proposed, which was not used previously for this purpose. A very simple neural network of just one hidden layer is trained to calculate the capacity of SM and GSM with the help of some input features carefully chosen. This method is more accurate than previous analytical approximations from the literature, [49] and [42] and, additionally, it entails a much lower computational complexity.

The SM capacity calculation has applications in both terrestrial and satellite communication systems. On the one hand, low complexity energy efficient mobile devices can benefit from this calculation, allowing the implementation of adaptive SM transmitters. And, on the other hand, these results can be applied to Dual Polarization (DP) mobile satellite systems which employ Polarized Modulation (PMod), a particular case of 2×2 SM where the dimensions represent different orthogonal polarizations, in order to build an adaptive DP system.

The leverage of DP in mobile satellite systems which operate in L and S-bands allows to increase the capacity of these systems without extending the bandwidth, thanks to the application of MIMO signal processing techniques. In this context, a link adaptation algorithm is proposed for selecting both

the optimum MIMO transmission mode and MCS in a DP mobile satellite system. Moreover, for the particular case of a DP system with PMod a novel adaptation architecture is proposed with two parallel channel encoders with independent coding rates. This type of adaptation is possible thanks to the method proposed previously to compute the capacity of SM and PMod. From the point of view of an user, higher throughput communications can be provided to mobile terminals at land, sea or air, and the speed of the connection is optimized in real time.

Lastly, different link adaptation techniques are given for selecting the coding rate in a SM system. The coding rate selection based on the SM capacity calculated by a neural network allows a great improvement in the throughput and the robustness of the communication compared with a fixed coding rate allocation. However, the method which provides the best performance is a Deep Learning (DL) based coding rate selection method, which is explained in the last chapter of the thesis. Thus, mobile devices of future 5G networks can adapt also their transmission bit rate when they employ SM as modulation scheme. Therefore, the DL based adaptation mechanism can benefit specially IoT devices, offering them throughput enhancements but balancing the transmitter complexity and the power consumption.

1.3. Contributions

This thesis has five main scientific contributions, each one developed in a separate chapter. The list of contributions is shown hereafter.

- **C1. Experimental validation of link adaptation algorithms for the return link of mobile satellite systems.**
  An adaptive satellite link is established to communicate a mobile platform with a ground station by means of a MEO satellite. The physical layer of both transmitter and receiver, implemented with SDR technology, is inspired by the specifications of S-UMTS. Several experiments are carried out to test different link adaptation algorithms in the return link. Results show that the inclusion of the open loop SNR is beneficial for the adaptation and that the algorithms succeed in the target of improving the spectral efficiency while maintaining a low frame error rate.

- **C2. Analysis of the impact of strong CSI degradation in Multibeam Satellite Systems with Linear Precoding and countermeasures proposal.**
  To increase the capacity of Multibeam Satellite System, a transition

from four color frequency reuse pattern to full frequency reuse is being proposed. This requires the application of interference mitigation, such as precoding. A problem not addressed in the literature was the impact that carrier detection errors during the CSI estimation have in the satellite precoding system. The estimated CSI is a sparse vector, with many zero components, which cause errors in the SINR that the gateway calculates and employs to allocate resources to the users. A statistical and geographical study of the errors in the SINR calculated by the gateway is presented. Moreover, it is shown how a link adaptation algorithm with an adaptive margin per user can help to reduce the rate of erroneous transmissions that the degraded CSI causes, with a minimum throughput loss.

- **C3. Proposal of a new method for obtaining capacity related metrics of Spatial Modulation and Generalized Spatial Modulation based on a Machine Learning approach.**
  The computation of the capacity of GSM and the capacity of SM constrained to a given constellation is realized with very simple neural networks of just one hidden layer. The selection of a particular set of input features allows neural networks to outperform previous analytical approximations of the capacity for these modulation schemes. Moreover, the capacity calculation with the neural network has a lower computational complexity than previous methods from the literature. This fast and accurate computation of the capacity of SM and GSM enables adaptive systems to compute in real time the achievable rate and adapt the physical layer transmission parameters consequently. To the best of our knowledge, this is the first time that neural networks are employed to make these information theory related calculations for non conventional modulation schemes.

- **C4. A link adaptation algorithm for Mobile Satellite Systems with Dual Polarization.**
  The inclusion of Dual Polarization in Mobile Satellite Systems allows to enhance their capacity by means of MIMO signal processing techniques. An adaptive system is analyzed where, apart from varying the MCS, the MIMO mode is also adapted dynamically, being this the baseline SISO, OPTBC, PMod or V-BLAST. A link adaptation algorithm for selecting both the MIMO mode and the MCS is proposed and its performance is evaluated in a mobile maritime channel by means of physical layer abstraction techniques. This algorithm is based on adaptive margins which are updated using ACK/NAK information received from the other end. In addition, a novel architecture for PMod is presented, with two parallel channel encoders which permit to apply different levels of protection to the bits selecting the polarization and

the modulation symbols.

- **C5. Rate adaptation in Spatial Modulation links assisted with Deep Learning.**
  A solution for adapting the coding rate of a SM transmitter according to the channel conditions using a deep neural network is presented. Simulations results using the whole the receiver detection and decoding chain are provided for a 2×2 SM system with a QPSK constellation. Several adaptation strategies are compared, including a fixed rate solution, a selection based on the theoretical calculated capacity and the proposed deep learning assisted solution. Results show that with the deep learning based coding rate selection, the throughput is very close its maximum achievable value whilst the rate of erroneous frames is maintained at a very low level.

1.4. Publications

The publications that develop each one of the previous contributions of the thesis are shown hereafter:

- **C1. Experimental validation of link adaptation algorithms for the return link of mobile satellite systems.**

  - **A. Tato**, C. Mosquera, I. Gómez. *Link Adaptation in Mobile Satellite Links: Field Trials Results.* In 2016 8th Advanced Satellite Multimedia Systems Conference and the 14th Signal Processing for Space Communications Workshop (ASMS/SPSC), pages 1-8, Sept. 2016, Palma de Mallorca, Spain. [93]

- **C2. Analysis of the impact of strong CSI degradation in Multibeam Satellite Systems with Linear Precoding and countermeasures proposal.**

  - **A. Tato**, S. Andrenacci, S. Chatzinotas, C. Mosquera. *Link Adaptation and Carriers Detection Errors in Multibeam Satellite Systems with Linear Precoding.* In 2018 9th Advanced Satellite Multimedia Systems Conference and the 15th Signal Processing for Space Communications Workshop (ASMS/SPSC), pages 1-8, Sept. 2018, Berlin, Germany. [94]
  - **A. Tato**, S. Andrenacci, E. Lagunas, S. Chatzinotas, C. Mosquera. *Link Adaptation and SINR Errors in Practical Multicast*

Multibeam Satellite Systems with Linear Precoding. International Journal of Satellite Communications and Networking (under revision), 2019. [97]

- **C3. Proposal of a new method for obtaining capacity related metrics of Spatial Modulation and Generalized Spatial Modulation based on a Machine Learning approach.**

  - **A. Tato**, C. Mosquera, P. Henarejos, A. Pérez-Neira. *Neural Network Aided Computation of Mutual Information for Adaptation of Spatial Modulation*. IEEE Transactions on Communications (acceptance pending), 2019. [98]
  - **A. Tato**, C. Mosquera, P. Henarejos, A. Pérez-Neira. *Neural Network Aided Computation of Generalized Spatial Modulation Capacity*. In 2019 27th European Signal Processing Conference (EUSIPCO), Sept. 2019, A Coruña, Spain. [99]

- **C4. A link adaptation algorithm for Mobile Satellite Systems with Dual Polarization.**

  - **A. Tato**, P. Henarejos, C. Mosquera, A. Pérez-Neira. *Link Adaptation Algorithms for Dual Polarization Mobile Satellite Systems*. In P. Pillai, K. Sithamparanathan, G. Giambene, M. Á. Vázquez, and P. D. Mitchell, editors, Wireless and Satellite Systems, pages 52-61, Cham, 2018. Springer International Publishing, Oxford, UK. [95]
  - **A. Tato**, C. Mosquera, P. Henarejos, A. Pérez-Neira. *Practical Implementation of Link Adaptation with Dual Polarized Modulation*. In 2018 11th International Symposium on Communication Systems, Networks Digital Signal Processing (CSNDSP), pages 1-6, July 2018, Budapest, Hungary. **Best paper award**. [96]

- **C5. Rate adaptation in Spatial Modulation links assisted with Deep Learning.**

  - **A. Tato**, C. Mosquera. *Deep Learning Assisted Rate Adaptation in Spatial Modulation Links*. In 16th International Symposium on Wireless Communications Systems (ISWCS), Aug. 2019, Oulu, Finland. [92]

1.5. Strucuture of the Thesis

The thesis is structured in five chapters, each one developing on the previously mentioned contributions. Table 1.1 shows the correspondence between the contributions and the chapters and, in addition, it compares the different contributions. It indicates if the contribution applies to a satellite or a terrestrial system. In the former case, it differentiates Mobile Satellite Services (MSS) with Single Polarization (SP) or Dual Polarization(DP) and Fixed Satellite Services (FSS). In the table it is also indicated the nature of the communication, unicast or multicast, and if the results are obtained by means of simulations or via experimentation. Futhermore, it includes the acronyms of the journals or conferences where the results are published or submitted.

Contribution	**Chapter**	**Journal/conference**	**Scenario**		
C1	2	ASMS 2016	Satellite MSS - SP	Unicast	Experiment
C2	3	ASMS 2018, IJSCN 2019	Satellite FSS - Precoding	Multicast	Simulation
C3	4	TCOM 2019, EUSIPCO 2019	Terrestrial 5G / Satellite MSS - DP	Unicast	Simulation
C4	5	WISATS 2017, CSNDSP 2018	Satellite MSS - DP	Unicast	Simulation
C5	6	ISWCS 2019	Terrestrial 5G / Satellite MSS - DP	Unicast	Simulation

Table 1.1: Contents of the thesis.

Chapter 2

Link Adaptation in Mobile Satellite Links: Field Trials Results

This chapter is adapted, with permission of the coauthors and the editorial, from IEEE: "A. Tato, C. Mosquera, I. Gómez. Link Adaptation in Mobile Satellite Links: Field Trials Results. In 2016 8th Advanced Satellite Multimedia Systems Conference and the 14th Signal Processing for Space Communications Workshop (ASMS/SPSC), pages 1-8, Sept. 2016, Palma de Mallorca, Spain". [93]

2.1. Introduction

Link adaptation is a key element of many terrestrial communication standards such as 3GPPP LTE [1], IEEE 802.16 [2] or IEEE 802.11 [3], and plays also a key role in satellite communication systems such as DVB-S2X [6], DVB-RCS2 [4] and S-UMTS [9]. The SL family of the latter standard, one of the several S-UMTS specifications which are available, is precisely the one used as baseline for the work presented in this chapter. A proprietary version of this standard is employed by the service BGAN (Broadband Global Area Network) of Inmarsat. The field trials described in this chapter were developed to validate some novel link adaptation algorithms for the return link of a mobile satellite channel. Link adaptation, substantiated as Adaptive Coding and Modulation (ACM), involves the selection of the Modulation and Coding Scheme (MCS or MODCOD) of the transmitted frames according to the channel conditions with the aim of maximizing the spectral

efficiency of the communications system while maintaining the error rate, measured as Frame Error Rate (FER), at a target level p_0.

Software Defined Radio (SDR) has been an essential technology for the field trials detailed in this chapter. It was used to implement the physical layer of the Mobile Platform and the Ground Station, with significant effort required to achieve real time operation.

According to the SDR Forum [86], a Software Defined Radio can be defined as a radio in which some or all of the physical layer functions are software defined, so that operating functions within the radio, not just control routines, are processed by software. The radio subsystem consists only of the minimum essential RF (Radio Frequency) parts (antennas, bandpass filters, amplifiers, high frequency oscillators), an ADC (Analog to Digital Converter) and DAC (Digital to Analog Converter), together with programmable devices such as FPGAs (Field Programmable Gate Arrays) and general purpose CPUs.

The SDR platform hosting the radio subsystem was tested by setting a satellite link communicating a Mobile Terminal with a Ground Station (GS). The Mobile Terminal was flied in an UAV (Unmanned Aerial Vehicle) for some of the tests; a car was also used to test the performance for the Land Mobile Satellite (LMS) channel.

The remainder of the chapter is structured in the following way. In Section 2.2 a review of several adaptive schemes for link adaptation is made, with an emphasis on the algorithms implemented in the Mobile Terminal. Then, Section 2.3 deals with the practical setup of the test trials and the details of the prototype development. After that, the results are presented in Section 2.4 and finally Section 2.5 collects the conclusions.

2.2. Adaptive Schemes Review

The use of ACM, by choosing different modulations and FEC (Forward Error Correction) rates, serves to avoid large link margins in time-varying channels, thus increasing the spectral efficiency for a given availability. The operation of the link adaptation schemes requires the exchange of Channel State Information (CSI) between both ends of the communication. In this chapter we focus on the return link, for which the Mobile Terminal (MT) needs information about the reception at the Ground Station (GS) (see Figure 2.1). The algorithms that we will consider exchange the CSI and the decoding outcome in the form of acknowledgements/non-acknowledgements

(ACK/NAK). As CSI, the signal to noise ratio (SNR) is the reference parameter, given its prediction capability on the result of decoding for a given MCS, provided that an accurate estimate of SNR is available. Both ends of the communication can measure the SNR; on the one hand, the closed loop CSI is the SNR measured by the Ground Station and sent back to the MT through a feedback channel. On the other hand, the open loop CSI is the SNR measured directly by the MT itself in the user downlink. If the uplink and downlink frequencies are not too far from each other in a FDD (Frequency Division Duplexing) link, then some correlation between the user uplink and the user downlink channels is expected [112]. Therefore, the open loop CSI is an estimate of the quality of the channel that the transmissions undergo, and which is captured more precisely by the closed loop CSI. One key difference between open loop SNR and closed loop SNR is that while the closed loop CSI arrives to the transmitter after a long delay, specially in satellite communications with GEO satellites, the open loop CSI can be used by the transmitter right after its measurement[1]. Consequently, the closed loop CSI could be completely outdated which means it does not provide useful information about the current channel conditions for rapidly variant channels, i.e., for high and moderate mobile speeds. Instead, open loop CSI, although is less precise, can still be useful in those cases.

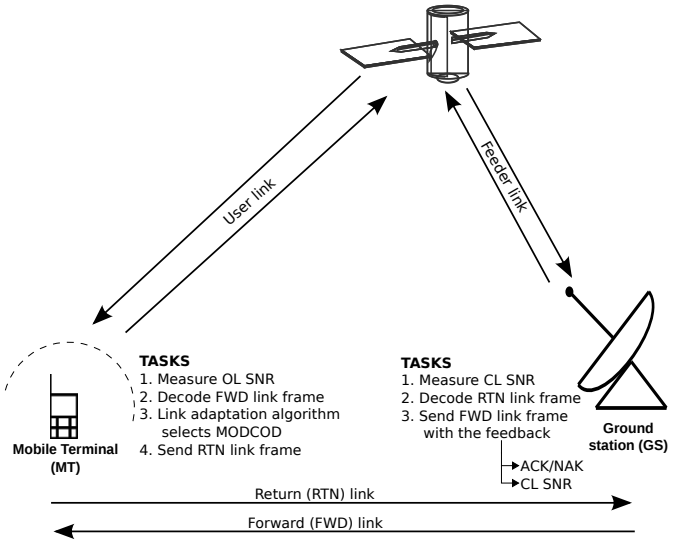


Figure 2.1: Scenario of the link adaptation.

Usually, link adaptation takes place by using look-up tables (LUTs) which assign MCS to ranges of SNR values. These LUTs are built based on simulations for the fine tuning of SNR thresholds; additionally, a back-off

[1]Processing and buffering can delay the application of the open loop information to the transmitted frame.

margin is employed to account for non-ideal effects in real environments. In the algorithms that we developed and that are tested in this work, this back-off margin is computed on the fly by exploiting the exchange of ACK and NAKs between the transmitter and the receiver. Furthermore, both open and closed loop CSI are combined to feed the LUT. The corresponding weight of each SNR contribution is updated with a similar algorithm to that of the margin. Overall, open loop SNR, closed loop SNR and the sequence of ACK/NAKs are used to choose the MCS at each moment, without resorting to any knowledge or assumption about the environment and channel. Note that ACM techniques can be used even in the absence of SNR information, just based on the sequence of ACK/NAKs, as reported in [78], where an algorithm named ARF-M (Automatic Rate Fallback with Memory) is introduced.

Before presenting the adaptive algorithms that we will use, we introduce a formal definition of the communication parameters. The acknowledgement of the correct decoding of the i-th codeword is denoted by ϵ_i, which takes the value 1 when the i-th codeword has not been decoded correctly (NAK) and the value 0 when it was decoded successfully (ACK). Due to the existence of a feedback delay of d frames, at time instant i the transmitter only has access to $\epsilon_0, \ldots, \epsilon_{i-d}$. The other type of feedback received by the transmitter is the closed loop SNR, $\mathrm{SNR}_i^{\mathrm{cl}}$, an estimate of the quality of the channel in time instant $i-d$ but used in time instant i due to the feedback delay. Lastly, the open loop SNR, $\mathrm{SNR}_i^{\mathrm{ol}}$, is the most recent SNR estimation on the forward link, measured in time instant $i-1$ and used in time instant i.

The transmitter selects an MCS m_i for each frame (or codeword) from a set of M available MCSs. An LUT, represented by means of a function Π, maps SNR intervals to MCS. The SNR thresholds of the LUT for each MCS are based on the required SNR needed to achieve a given FER, for example in an AWGN channel. The value of SNR introduced in the LUT for selecting the MCS is obtained by means of an equation that is different for each of the four algorithms considered here. In the sequel, all SNR parameters and margins are provided in dB. For the **closed loop algorithm** the equation is given by

$$m_i = \Pi\left(\mathrm{SNR}_i^{\mathrm{cl}} + c^{\mathrm{cl}}\right) \tag{2.1}$$

and the corresponding **open loop** mapping is

$$m_i = \Pi\left(\mathrm{SNR}_i^{\mathrm{ol}} + c^{\mathrm{ol}}\right), \tag{2.2}$$

where c^{cl} and c^{ol} are the margins used in the closed and open loop algorithms, respectively. In both cases the margins should be obtained adaptively.

Additionally, we consider the **balanced algorithm**, presented in [78],

which selects the MCS as

$$m_i = \Pi\left(\xi^{\mathrm{ol}}\mathrm{SNR}_i^{\mathrm{ol}} + \xi^{\mathrm{cl}}\mathrm{SNR}_i^{\mathrm{cl}} + c\right), \tag{2.3}$$

and the **balanced convex algorithm**, with the weights adding up to one:

$$m_i = \Pi\left(\left(1 - \xi^{\mathrm{cl}}\right)\mathrm{SNR}_i^{\mathrm{ol}} + \xi^{\mathrm{cl}}\mathrm{SNR}_i^{\mathrm{cl}} + c\right). \tag{2.4}$$

In the last two cases the weights ξ^{cl} and ξ^{ol} and the margin c are to be obtained adaptively too.

The adaptation scheme for the weights ξ and the margin c follows from an optimization problem, based on adjusting the frame error rate (FER) to an objective value p_0. The problem is solved using a stochastic gradient descent algorithm which leads to the adaptation rules (2.5), (2.6) and (2.7). Details of their derivation can be found in [78].

The final equation for the adaptation rule of the **balanced algorithm** reads as follows:

$$\begin{bmatrix} c_{i+1} \\ \boldsymbol{\xi}_{i+1} \end{bmatrix} = \begin{bmatrix} c_i \\ \boldsymbol{\xi}_i \end{bmatrix} - \frac{\mu}{\theta^2 + \|\mathbf{SNR}_{i-d}\|^2}\left(\epsilon_{i-d} - \tilde{p}_{0,i}\right)\begin{bmatrix} \theta \\ \mathbf{SNR}_{i-d} \end{bmatrix}, \tag{2.5}$$

where for simplicity we denote $\mathbf{SNR}_i = [\mathrm{SNR}_i^{cl}\ \ \mathrm{SNR}_i^{ol}]^T$ and $\boldsymbol{\xi}_i = [\xi^{cl}\ \xi^{ol}]^T$. In the previous equation μ is the adaptation step, θ is a constant to increasing the converge speed of the margin c and $\tilde{p}_{0,i}$ is the adaptive target FER that will be explained later.

In the case of the **balanced convex algorithm**, for which only one degree of freedom for the SNR weights is available, the expression for the adaptive scheme is:

$$\begin{bmatrix} c_{i+1} \\ \xi_{i+1}^{\mathrm{cl}} \end{bmatrix} = \begin{bmatrix} c_i \\ \xi_i^{\mathrm{cl}} \end{bmatrix} - \frac{\mu}{\theta^2 + \left(\mathrm{SNR}_{i-d}^{\mathrm{cl}} - \mathrm{SNR}_{i-d}^{\mathrm{ol}}\right)^2} \times \left(\epsilon_{i-d} - \tilde{p}_{0,i}\right)\begin{bmatrix} \theta \\ \mathrm{SNR}_{i-d}^{\mathrm{cl}} - \mathrm{SNR}_{i-d}^{\mathrm{ol}} \end{bmatrix}. \tag{2.6}$$

As to the **open and closed loop algorithms**, the adaptation rule of the margin, the only adjustable parameter, is:

$$c_{i+1} = c_i - \frac{\mu}{\theta^2 + (\mathrm{SNR}_{i-d}^{x})^2}(\epsilon_{i-d} - \tilde{p}_{0,i}). \tag{2.7}$$

We need to point out that the selection of the MCS through equations (2.1-2.4) is based on the most recent values of the SNR, $\mathbf{SNR}_i$, whereas the

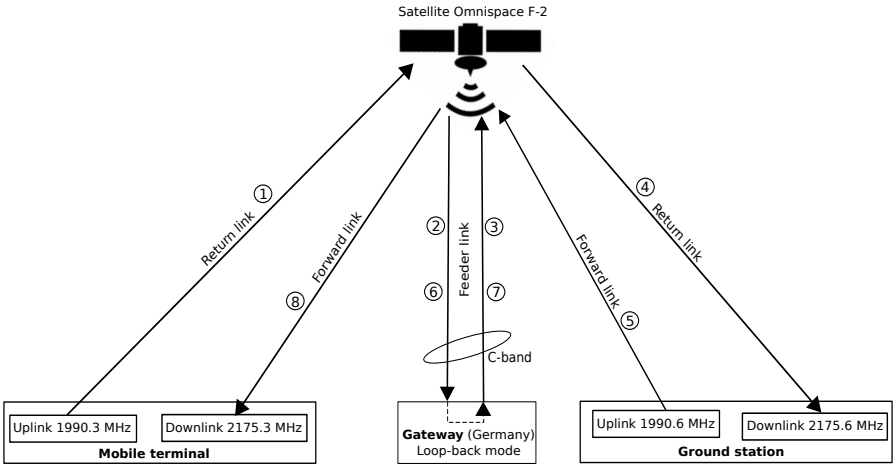


Figure 2.2: Satellite links for the field trials.

update of $\boldsymbol{\xi}$ and c use $\mathbf{SNR}_{i-d}$, the values of SNR employed for selecting the MCS of the frame that the ACK/NAK ϵ_{i-d} refers to. This requires the transmitter to have some memory to store the values of SNR used to set the MCS of each frame until its respective ACK/NAK arrives. Furthermore, an additional loop for the variable $\tilde{p}_{0,i}$ is added, with the aim of achieving a better adjustment of the FER to p_0. The recursion for $\tilde{p}_{0,i}$ is written as

$$\tilde{p}_{0,i+1} = \tilde{p}_{0,i} - \lambda(\epsilon_{i-d} - p_0) \tag{2.8}$$

where $\lambda = p_0/100$. In (2.5), (2.6) and (2.7) $\mu = 1$ and $\theta = 10$, and in (2.7) x reads as $\{ol, cl\}$ for open loop and closed loop, respectively.

The margin c is restricted to lie in the interval $[-8.3, 8.3]$, whereas the weights ξ stay in the range $[0, 1]$. Thus, divergence is avoided even in the case of unattainable target FER due to, for example, poor channel conditions.

Finally, let us remark that the LUT mapping can be referred to as inner loop link adaptation, whereas the adjustment of the LUT mapping itself is known also as outer loop link adaptation in some references [71].

2.3. Practical Setup and Prototype

The field trials deployed a satellite communication link between a Mobile Terminal and a Ground Station. The available MCS schemes for the return link are listed in Table 2.1, taken from [9]. On the other side, the Ground Station always employs the most robust MCS since in our trials we only want to test the algorithms for the return link. In each of the trials, the

selection of the MCS for each frame was determined by one of the four link adaptation algorithms explained in the previous section. The final goal of the field trials was to compare the four algorithms in terms of spectral efficiency (throughput) and FER.

Table 2.1: Coding rate options for the R20T2Q-1B bearer [9], QPSK constellation.

	L8	**L7**	**L6**	**L5**	**L4**	**L3**	**L2**	**L1**	**R**	**H1**
Coding rate	0.33	0.41	0.48	0.55	0.63	0.71	0.78	0.84	0.87	0.91
Rate (spectral efficiency)	0.66	0.82	0.96	1.10	1.26	1.42	1.56	1.68	1.74	1.82
γ_{th} **(dB)**	0.03	1.03	2.03	2.93	3.73	4.83	5.83	6.73	7.43	8.33

2.3.1. Satellite Component

The elements of the satellite link are shown in Figure 2.2. The Mobile Terminal runs the link adaptation algorithms and communicates with the satellite by means of an S-band transceiver. The Ground Station, which also operates in the S-band and remains in a fixed location, sends feedback to the Mobile Terminal about the frames decoding process (ACK/NAK) and the measured SNR. The bent-pipe mode satellite is complemented by an intermediate gateway which operates in loop-back mode and is managed by the satellite operator.

For the field trials some transponder capacity from the F-2 satellite (formerly belonging to ICO) [68] was leased from the satellite operator Omnispace. The satellite operates on 30 MHz in the S band for each direction and, more specifically, it employs the range 1,985 - 2,015 MHz for the uplink and the range 2,170 - 2,200 MHz for the downlink. The communication between the satellite and the gateway is done in C-band. Although the operator owns several gateways, the teleport located in Usingen (Germany) was used for the field trials.

The end to end connection is implemented in a loop-back mode, with the gateway relaying in the uplink the signal received in the corresponding downlink after filtering and amplification. This operation mode implies that the signals travel four times the distance from the satellite to the Earth before arriving to their destination. As a result, a minimum delay of 140 ms and a round-trip-time (RTT) of 280 ms (when the elevation is 90^o) is experienced. Table 2.2 summarizes the main parameters of the involved links whereas Figure 2.2 numbers the sequence of hops from the Mobile Terminal to the Ground Station and the reverse.

Characteristic	Value
Satellite	Omnispace F-2 (former ICO F-2)
Operator	Omnispace LLC
Orbit	MEO (10,500 km) 45^o inclination
Leased bandwidth	200 kHz in each direction
Maximum EIRP (MT and GS)	43 dBm
Minimum Delay (RTT)	140 ms (280 ms)

Table 2.2: Main parameters of the satellite links.

2.3.2. Physical Layer

The reference for the implementation of the physical layer was the S-UMTS standard and, in particular, the SL family [9]. This ETSI standard is conceived for providing mobile communications of third generation using GEO satellites. The frequencies employed in this standard fall in the middle of the L-band, at about 1,500-1,600 MHz. This frequency band is further divided into 200 kHz sub-bands. One important concept in its technical specification is the shared access bearers. They are physical bearers or carriers that support the transfer of data of multiple concurrent connections simultaneously, distinguishing the packets of each connection by means of their address field.

There are 13 bearers for the forward link and 22 for the return link. One bearer differs from the others in the symbol rate (and hence bandwidth), the modulation type and the slot duration. In this way, the transmission parameters can be adapted to the different channel conditions and terminal capabilities. The great flexibility provided by the large number of bearers allows data rates in the range of 3.2 to 858 kbit/s. Channels in the forward direction are allocated on a Time Division Multiplex (TDM) basis, dividing the time in slots of 80 ms. On the return direction, channels are allocated on a Time Division Multiple Access (TDMA) basis, dividing each frequency channel in small time slots being used each one by a different user. The duration of the return link time slots can be 5 ms, 20 ms or 80 ms. The system is single-carrier, and employs a root-raised-cosine (RRC) filter with roll-off factors of 0.25 or 0.13 depending on the bearer type. Regarding channel coding, turbo-codes with variable coding rate are used in S-UMTS.

The adjustment of the communication setting takes place by first assigning a bearer type to each Mobile Terminal, and then a specific coding rate of the turbo-encoder is selected for each frame. The flexibility is somewhat limited, since several users can share a given frame, so the lowest SNR among

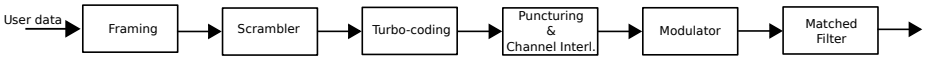


Figure 2.3: Block diagram of the transmitters.

them determines the corresponding MCS. In addition, the SNR is averaged for a long time, which entails a significant inertia for the tracking of the channel. The supported modulations range from QPSK to 64-QAM, with $\pi/4$-QPSK as an additional modulation for return link bearers. The prototype developed for this study includes bearers R20T2Q-1B ($\pi/4$-QPSK) and R20T2X-1B (16-QAM), although only the R20T2Q-1B bearer was used in the field trials for both forward and return links. The roll-off factor of the root raised cosine is 0.25 and the symbol rate is $R_s = 67.5105$ ksymb/s.

Each frame lasts 19.9 ms and contains only one FEC block which conveys 1344 slot symbols. A subgroup of symbols named Unique Word (UW), 40 at the beginning and 24 at the end of the frame, are used for frame detection, SNR estimation and frequency and phase synchronisation. They also signal the coding rate, i.e., the bearer subtype of each frame.

Figure 2.3 shows the block diagram of the MT transmitter. The framing block builds the base-band frames with the data and the header. This is complemented by the scrambler, the channel encoder, puncturing to generate the different coding rates, the channel interleaver, modulator and matched filter. All of these blocks are implemented following the specifications of the standard [9].

Figure 2.4 displays the corresponding block diagram of the MT receiver. The SDR platform has an amplifier whose gain depend on the value set by the AGC (Automatic Gain Control) block. Then, a 12 bit ADC samples the complex signal at a rate of 16 Msamples/s (giving an oversampling factor of 237). Afterwards, a decimator in the FPGA reduces the sample rate by a factor of 79 giving a sample rate at the input of the CPU of 202.532 Ksamples/s (an oversampling factor of 3). Subsequent blocks are the variable bandwidth matched filter and the DC-Offset calibration, which removes the DC component of the signal by subtracting the average of each frame. The AGC sets the power and controls the variable input amplifier. The Timing Recovery block uses the Oerder and Meyer algorithm [67] to set the optimum symbol timing. Then, the samples, now at a rate of 67.5105 Ksamples/s, can follow two different paths. During the acquisition stage, when the receiver is not locked, multiplexers (1) and (2) are in position (b). In this mode, the frame detection is more robust, tolerating higher frequency offsets, due to the fact that the correlations are estimated based on the phase differences, as opposed to the tracking stage, with multiplexers switched to position

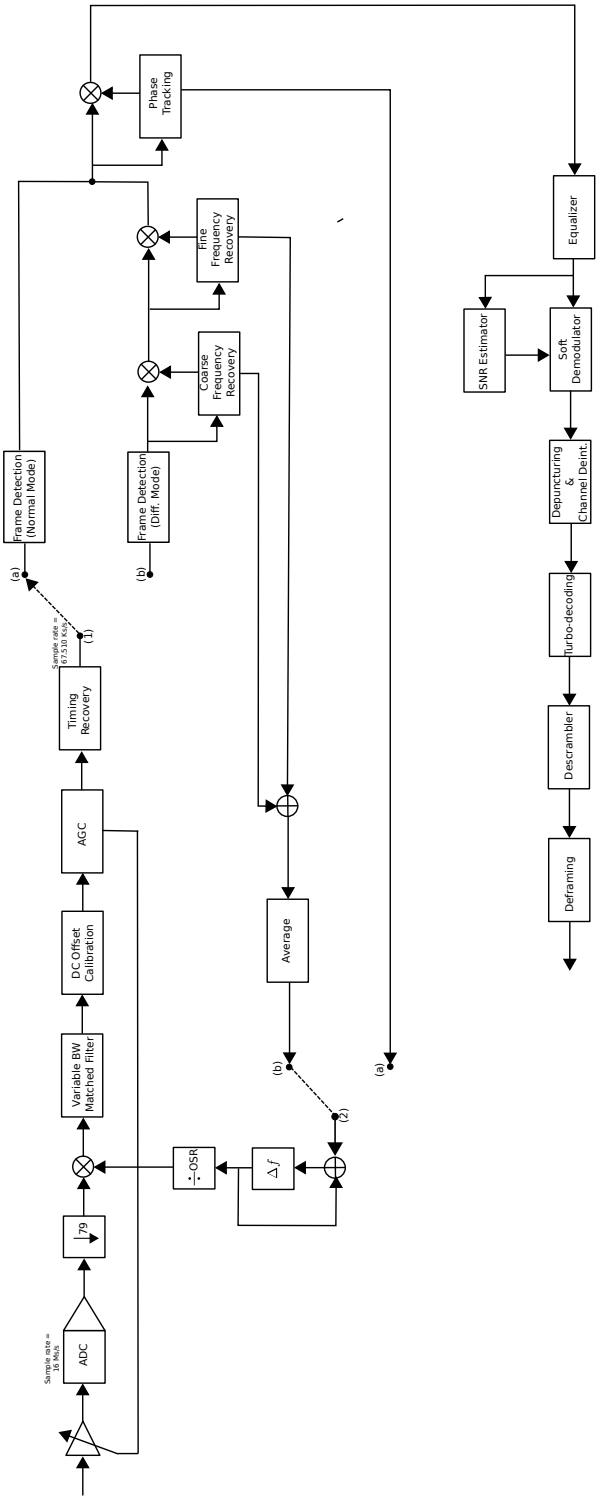


Figure 2.4: Block diagram of the Mobile Terminal receiver. Tracking phase when multiplexers are in position (a) and acquisition phase when they are in position (b).

(a). In the acquisition stage the receiver performs the frequency estimation over each frame by using the algorithm Luise&Reggiannini [58]; an average estimate is computed across several frames. In the phase tracking block, the phase is estimated at the beginning of the frame (using the first UW symbols), at the end (using the end UW symbols), and then a linear interpolation is obtained to compensate the phase error. A short adaptive FIR filter of six coefficients is also included for time equalization, by using the Multi-Modulus Algorithm [107]. The SNR Estimator block calculates the mean SNR of each frame by using a Non-Data-Aided algorithm which makes use of sixth-order statistics [57]. This estimation will be used in the soft-decoding and as a measure of the link quality. Next, the Soft Demodulator performs the soft demodulation in order to obtain the log-probabilities or LLR (Log-Likelihood Ratio). The De-puncturing, Channel De-interleaver, De-scrambler and De-framing carry out the inverse operations of their opposite blocks in the transmitter. Finally, the Turbo-decoding block decodes the signal with the BCJR algorithm. The differences between the Ground Station and the Mobile Terminal receiver lie mainly in the frame detection and the frequency recovery; for instance, in the Mobile Terminal the frame detection is easier because the Ground Station always transmits with the same MCS for this experiment.

2.3.3. Hardware

The main hardware elements of the prototypes for the MT and the GS are the SDR platform and the set of devices which form the external analog front-end. The selected SDR platform was the USRP Ettus E310 [35], which includes a dual core ARM A9 processor and a reconfigurable Xilinx 7 Series FPGA. The bulk of the baseband processing is done in a C/C++ application executed in the processor, while the FPGA implements the Intermediate Frequency up-conversion and baseband down-conversion, the decimation and the interpolation. The RF front-end of the SDR platform, the RFIC (Radio Frequency Integrated Circuit) AD 9361, includes DACs, ADCs, filters, mixers and amplifiers. Apart from the USRP, there is a transmit (TX) filter in the transmission chain, located at the transmission band to reduce the out of band emissions to avoid the contamination of the received signal. It was necessary to insert a driver, since the output power of the SDR platform is not high enough for feeding the 10 W power amplifier. Both the transmission and reception chains share the duplexer and the antenna. The duplexer connects the antenna with both transmission and reception chains. It behaves as two band pass filters, centered at the transmission and reception bands, respectively. In the reception chain there are two LNAs (Low Noise Amplifiers) and a RX filter centered around our reception frequency. This is

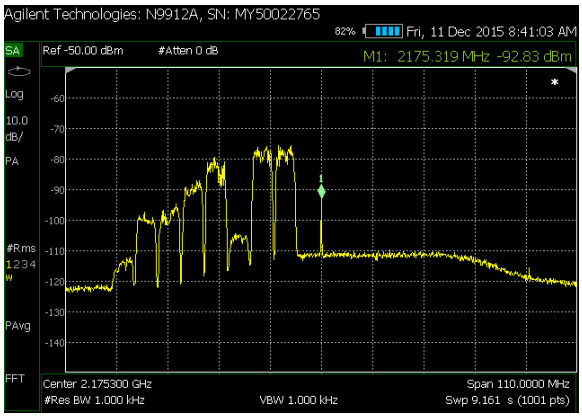


Figure 2.5: Capture of 3G interference and our received signal in the center of the graphic.

necessary for avoiding that the power in the transmission frequency reflected in the antenna can saturate the second LNA. In order to avoid any kind of pointing and tracking in both the Mobile Terminal and the Ground Station, a hemi-directional antenna was used at both ends. It has a beamwidth of 160^ox160^o, a gain of 2 dBiC and a right-hand circular polarization.

During the first tests it was found that the proximity of 3G base stations impaired to a large extent the decoding of the signal. This is due to the presence of strong 3G broadband signals in 2170 MHz with a power 15 dB higher than our received signal, causing the LNA of the SDR platform to saturate. A capture of the spectrum analyzer showing this issue is shown in Figure 2.5. The solution was to insert a fifth order helical passive band-pass filter prior to the input of the USRP able to reject the 3G interference. The frequency response of the filter is shown in Fig. 2.6.

2.3.4. Additional Implementation Details

Correlations

In several functional blocks of the receiver (frame detection and frequency and phase synchronization) it is necessary to compute the correlation between the received and the transmitted UW symbols. These 40 UW symbols take some predefined values according to the standard and therefore are known by the receiver. This burden can be an issue for a real time receiver since each point of the correlation requires the computation of 64 complex products (or 256 real products). By exploiting the properties of the $\pi/4$-QPSK constellation, the calculations could be simplified to a large

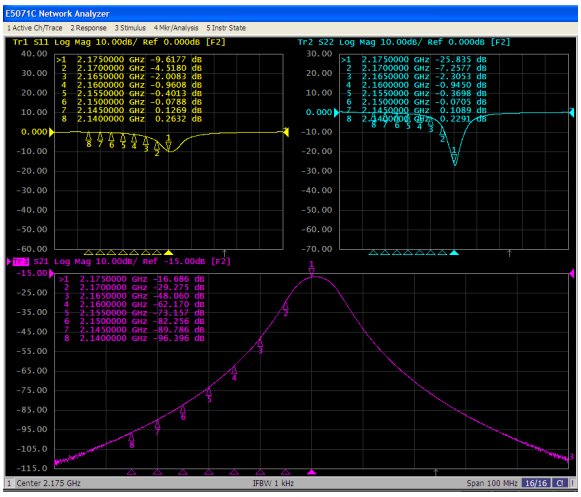


Figure 2.6: Band pass filter.

extent, and only two real products were required for each point.

Variable Bandwidth Matched Filter

Due to the large relative movement between the satellite and the Earth, systems involving MEO satellites suffer from significant Doppler. Although the Doppler of the feeder link is corrected in the gateway of Germany, our system has to cope with the Doppler shift due to the other two Earth-satellite jumps. Each jump causes approximately a shift of ± 10 kHz [56], giving a total shift of ± 20 kHz, a considerable amount if it is compared with the 84 kHz of the bandwidth of our transmitted signal. In reception, the matched filter in Figure 2.4 is actually a variable bandwidth low pass filter. During the acquisition phase, when the receiver is not locked, the cutoff frequency of this filter is extended so that the received signal, with a high Doppler offset, falls within the bandpass of the filter. However, during the tracking phase, when the receiver has estimated the frequency offset and has corrected it, this filter uses the nominal cutoff frequency for reducing the noise.

Link Adaptation Algorithms

The algorithms presented in Section 2.2 assume that a measurement of the open loop and a valid feedback with the closed loop SNR and the decoding outcome are always available. In practice, when the Mobile Terminal is not able to decode a frame, the values of the weights and margin remain

frozen. Also, when the receivers cannot detect a frame the corresponding SNR cannot be extracted, and the value -5 dB is used instead. Additionally, when the closed loop SNR is not available, its last valid value is used in the LUT. Lastly, due to buffering and processing time, the application of the open loop and the closed loop SNRs for MODCOD selection is delayed 6 frames (120 ms) and 30 frames (600 ms), respectively.

Link Layer

In the system, due to the fact that its aim was to test the algorithms, only the essential link layer functions were implemented. Both Mobile Platform and Ground Station send a decoding flag to the other end informing about the decoding status of the frame (ACK/NAK). Future work will include improvements in this layer such as retransmissions.

Tracking of the Mobile Platform

The URSP Ettus E310 includes a GPS receiver and several accelerometers. The information provided by them was included as payload data for a real-time tracking of the Mobile Platform.

2.4. Results

The Mobile Platform was tested as payload of a fixed-wing UAV (see Figure 2.7) and on the top of a car (see Figure 2.8). The location of the ground station was the roof of a building located in the University of Vigo campus, see Fig. 2.9. Three different environments were tested: two terrestrial (highway and semi-rural) and one aeronautical (UAV).

For each test the communication between the Mobile Platform and the Ground Station was active for five minutes. Both ends transmitted frames continuously, with the link adaptation algorithm deciding the MODCOD of each Mobile Terminal transmitted frame, whereas the Ground Station used always the most robust MODCOD for sending its feedback (ACK/NAK and closed loop SNR). The Mobile Terminal was running only one link adaptation algorithm in each trial, always with a target FER of 0.1. The transmitted power, constant during a test, was adjusted automatically at the beginning of each trial to obtain a specific average SNR at the Ground Station (a target closed loop SNR). Five different values of SNR, ranging

Figure 2.7: Picture of the UAV with the Mobile Platform inside and the antenna on top.

Figure 2.8: Picture of the car with the Mobile Platform on top.

Figure 2.9: Picture of the ground station.

from 0 to 12 dB, were tried.

Frame Error Rate (FER) and average spectral efficiency η of the whole transmission were calculated for each trial. The latter is defined as $\frac{1}{N}\sum_{i=1}^{N}(1-\epsilon_i)r_{m_i}$, with r_j the rate of the j-th MCS (as in Table 2.1), and m_i the selected MCS for the transmission of i-th frame. The frames received in the Mobile Terminal with an invalid CRC (Cyclic Redundancy Check) were not taken into account for calculating neither the efficiency nor the FER, since the Mobile Terminal cannot decode correctly the information of the feedback.

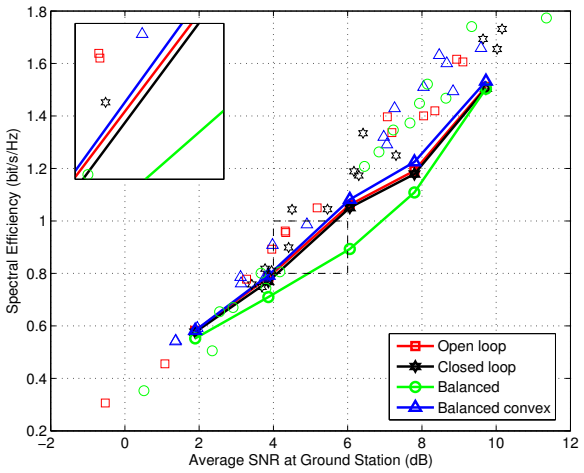
The outcome of the different experiments, measured as spectral efficiency and FER, is plotted in Figure 2.10 as independent markers for both car and UAV. The spectral efficiency of the four adaptive schemes is similar, and all achieve the target FER for SNR values higher than 3 dB. The limitations of our platform were such that only one algorithm could be tested at a time, which made the replication of the same channel conditions a challenging task. Among other factors, the time-varying elevation of the MEO satellite introduced some complications to establish a common operation point, despite the use of the same paths for all the trials and the implementation of an initial power control. Since the time series of open and closed loop SNR values were stored, they were used off-line as input in our link adaptation simulator for a more precise and fair comparison among the four

algorithms. All the algorithms were simulated first with all the available terrestrial datasets, and then the simulations were classified into five groups obtained by applying the *k-means* algorithm [87] to the average closed loop SNRs. After that, the mean efficiency and cumulative FER (the FER during all the experiment) within each group of simulations for each algorithm were calculated. With this, the points connected with lines showed in Figure 2.10 are obtained. The discrepancy between the experimental and the simulated results is deemed to be caused by the abstraction of the simulation setting, working at the SNR level rather than with the waveform and assuming the perfect reception of the ACK/NAK sequence.

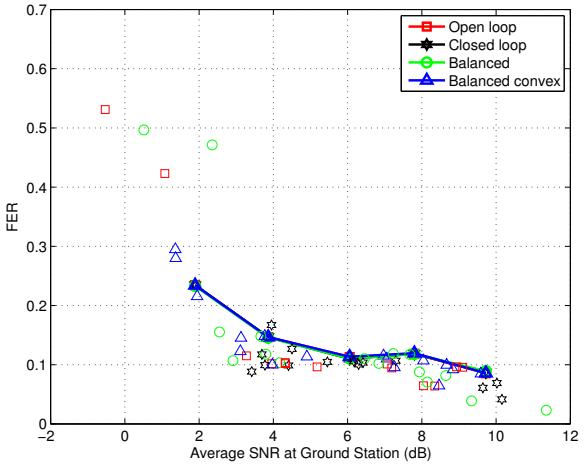
The first significant conclusion is that the balanced algorithm, which adapts independently closed and open loop SNR weights, offers the worst performance in terms of spectral efficiency, whereas the balanced convex version, with the two weights restricted to sum up to one, outperforms slightly the other algorithms. The reason seems to be in the different number of parameters to adapt, three in the balanced scheme and two in the balanced convex one, since a higher number of adaptive parameters slows down the convergence. It can be also noticed that the performance of balanced convex, open loop and closed loop algorithms is very similar, with balanced convex marginally best and closed loop marginally worst. Although the results are not conclusive, it seems that the open loop SNR has a significant correlation with the channel state, at least comparable to that of the closed loop SNR, and possibly slightly higher. One of the potential advantages of the use of open loop SNR is the associated agility to react to quick changes in the received power; this was observed in some of the series of RSSI (Received Signal Strength Indicator) at transitions between two beams. It is also hypothesized that the open loop SNR role could be more beneficial for links with lower multipath levels, for example with directive antennas. In our experiment, the estimated values of the Rice factor (K) ranged between 14 and 16 dB.

On the other side, Figure 2.11 shows the temporal evolution of the closed loop SNR during an UAV flight, together with the selected MODCODs for the corresponding SNR thresholds; note how more efficient MODCODs are chosen for higher SNR values. Although the SNR range is quite low in this specific experiment, SNR values up to 12 dB were tested in other flights by adjusting the power. The right figure shows the longitude of the instantaneous position of the UAV along with the RSSI (Received Signal Strength Indicator) at the Ground Station. In this trial, where the UAV followed an elliptic trajectory, there is a strong correlation between the position of the UAV and the strength of the signal. When the UAV moves to the west, the RSSI falls and when it moves to the east, the RSSI rises. This effect is due to the decrement in the gain of the antenna in the direction of the satellite

when the UAV makes a turn to the East.



(a) Spectral efficiency

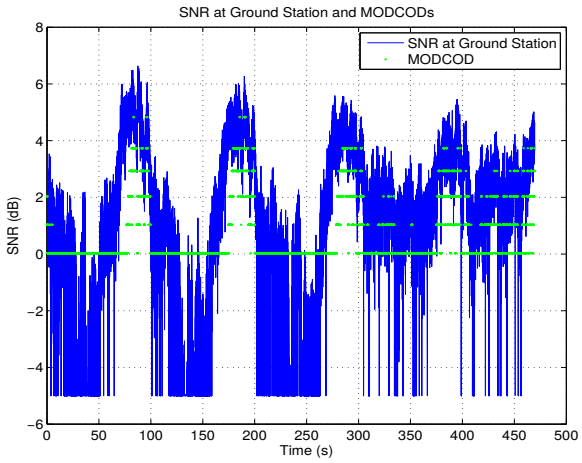


(b) FER

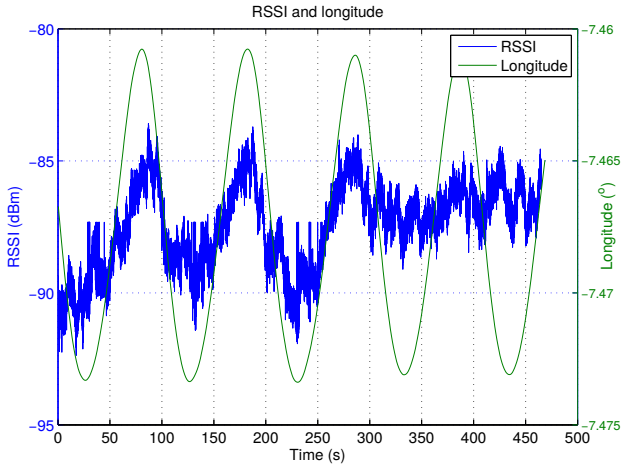
Figure 2.10: Mean spectral efficiency (top) and cumulative FER (bottom) of field trials (independent markers) and simulations (markers connected with lines).

2.5. Conclusions

In this chapter, we have presented several link adaptation algorithms and offer details of the implementation of a satellite communication system



(a) SNR and MODCODs



(b) RSSI and position

Figure 2.11: Evolution of the closed loop SNR and the selected MODCODs (top), and RSSI and longitude of the UAV position (bottom) during one of the trials with the balanced convex algorithm.

that can be used in mobile scenarios such as vehicular and UAV communications. The field trials served to validate the correct operation of the system, and show that the open loop SNR is useful in link adaptation, since the two algorithms with marginally better performance make use of the open loop SNR. The real time implementation could not accommodate the simultaneous test of all the algorithms, and a significant effort was required to try to replicate the operation conditions for all the trials, in part due to the use of a MEO satellite, with a time-varying elevation angle. The different schemes

were able to track the fluctuations of the SNR due to the variable orientation of the UAV with respect to the satellite, adapting the spectral efficiency of the communication accordingly. Lastly, this work shows also the potential of SDR technology to develop an adaptive satellite communication system and test it in real conditions.

Chapter 3

Link Adaptation and SINR Errors in Practical Multicast Multibeam Satellite Systems with Linear Precoding

This chapter is adapted, with permission of the coauthors and the editorial, from IEEE: "A. Tato, S. Andrenacci, S. Chatzinotas, C. Mosquera. Link Adaptation and Carriers Detection Errors in Multibeam Satellite Systems with Linear Precoding. In 2018 9th Advanced Satellite Multimedia Systems Conference and the 15th Signal Processing for Space Communications Workshop (ASMS/SPSC), pages 1-8, Sept. 2018", [94]. And with permission of the coauthors from IJSCN "A. Tato, S. Andrenacci, E. Lagunas, S. Chatzinotas, C. Mosquera. Link Adaptation and SINR errors in Practical Multicast Multibeam Satellite Systems with Linear Precoding[1]*, International Journal of Satellite Communications and Networking, 2019"*, [97].

3.1. Introduction

The use of more aggressive frequency reuse patterns, as full-frequency reuse, allows to increase the capacity of the forward link of multibeam satellite communications systems. This approach is proposed to increase the system capacity of High Throughput Satellites (HTS) operating at Ka-band. Linear precoding techniques reveal as a promising method to mitigate the

[1]Under revision.

higher co-channel interference which arises in this scenario. Indeed, the new superframe format of DVB-S2X [6] enables the use of precoding in DVB broadband satellite systems.

Several works propose the application of linear precoding to the forward link of DVB-S2(X) based systems, as [103], [24], [91], [102] and [100]. Performance analysis and practical issues are addressed; among others, multiplexing of several users per frame, the importance of users scheduling in the final system capacity, the outdated CSIT (Channel State Information at the Transmitter), or CSI estimation errors.

However, there is an important problem related with imperfect CSI in precoding which previous works do not consider, referred to as the nullification effect. To the best of our knowledge this problem is only partially covered in [14] and it is addressed with much more detail and for more realistic assumptions in this thesis. In precoded multibeam satellite systems, the CSI of each user is a vector comprising the link response of the intended and interfering beams. It often happens that receivers are not able to estimate many of these channels since the interfering signals from other beams are received with a very low power level. Therefore, the CSI available at the gateway is a sparse vector with many zero entries. This causes two problems. On the one hand, the calculated precoding matrix is not optimized for the actual channel but for the nullified version of the channel matrix, leading to some performance loss. And, more importantly, the Signal to Interference and Noise Ratio (SINR) calculated by the gateway for each user differs from the actual value. The error in the estimation of the users SINR by the gateway can be high enough to lead, in the case of overestimation, to increase the rate of erroneous frames over the QEF (Quasi Error Free) target of DVB-S2(X) systems, since the estimated SINR is used to perform Modulation and Coding Scheme (MCS or MODCOD) allocation.

This work tries to fill the gap related to the nullification error problem found in the literature, explaining how nullification occurs, characterizing the statistical and geographical distribution of the SINR errors due to the nullification effect, and proposing a method to overcome this problem. Since the nullification effect causes wrong MCS allocation due to the errors in the calculated SINR, a link adaptation algorithm with an adaptive margin is proposed as a countermeasure for those frame errors that the nullification triggers. The improved Outer Loop Link Adaptation (OLLA) algorithm described in [29] is used here, instead of the classic OLLA algorithm employed in [94]. Moreover, we quantify in terms of throughput loss the impact of three different countermeasures for the problem of CSI nullification: a global, a beam-based and a per-user backoff margin. Based on the study carried out in this chapter, our proposal for solving this issue with the mini-

mum impact in the system throughput is to use a link adaptation algorithm with an independent adaptive margin per user.

In this work, realistic assumptions are made in order to simulate a system closer to real practical systems. In particular, we include multicast, a basic user scheduling algorithm and Rician fading. In our simulations we address $M = 2$ users per frame, a number which may seem small compared with old precoding papers, but which is in line with the smaller frames proposed for next DVB-S2X precoding waveforms. Here, two different levels of user scheduling are considered, namely, a geographical inter-beam scheduling, which results from dividing all the beams in sectors, and an intra-beam scheduling, by creating multicast groups of $M = 2$ users in each beam based in the similarity of their channel vectors. Furthermore, in the evaluation of the link adaptation algorithm three scenarios are studied: one without fading, as in [94], and other two with Rician fading and different Rice factors, one for terrestrial and other for aeronautical applications.

This chapter is structured as follows. Section 3.2 presents the equations of the system model, describes the multibeam channel model, the computation of the precoding matrix and precoded SINRs and also the scheduling using in the simulations. Then, Section 3.3 explains the different types of CSI carriers nullification and how this effect is modeled. Afterwards, Section 3.4 presents the system parameters before the main results of the chapter on the analysis of the SINR error due to nullification are collected in Section 3.5. Then, in Section 3.6 the impact in the throughput of different types of margins is commented. Lastly, before the conclusions, Section 3.7 describes the link adaptation algorithm used to overcome the problems caused by the nullification, and provides some simulation results showing how the proposed strategy solves this problem.

3.2. System Model

We consider the forward link of a multibeam satellite communications system with full frequency reuse operating at Ka-band. Linear precoding is used to reduce the co-channel interference which arises in this configuration, significantly higher than that in four-color frequency/polarization configurations [91]. We assume that the satellite operates N beams with a single feed per beam (SFPB) payload. In each frame, data for M users located in the same beam are multiplexed. Therefore, $N \times M$ users are served simultaneously during the duration of a frame, M per beam. With all these considerations the general model for signals received at the User Terminals

(UTs) at a given time instant in matrix notation is

$$\mathbf{y}^{[i]} = \mathbf{H}^{[i]}\mathbf{x} + \mathbf{n}^{[i]} = \mathbf{H}^{[i]}\mathbf{W}\mathbf{s} + \mathbf{n}^{[i]}, \quad i = 1,\ 2, \ldots, M, \tag{3.1}$$

where the vector $\mathbf{y}^{[i]} \in \mathbb{C}^N$ collects the received signals by the i-th set of users in the N beams. $\mathbf{s} \in \mathbb{C}^N$ are the information symbols for the users, which are normalized in terms of energy, $\mathbb{E}\{|s_k|^2\} = 1$, $\mathbf{x} \in \mathbb{C}^N$ are the precoded transmitted signals by each antenna, $\mathbf{W} \in \mathbb{C}^{N\times N}$ is the precoding matrix, $\mathbf{H}^{[i]} \in \mathbb{C}^{N\times N}$ is the channel matrix for $i-$th set of users in all the beams. And $\mathbf{n}^{[i]} \in \mathbb{C}^N$ are the noise terms for the set of each i-th users. $\mathbf{n}^{[i]} \sim \mathcal{CN}(\mathbf{0}, N_0\mathbf{I}_N)$, i.e., the noise measured at each UT receiving antenna is complex zero mean Additive White Gaussian Noise (AWGN).

3.2.1. Channel Model

The general model [23] for the complex channel matrix is

$$\mathbf{H}^{[i]} = \mathbf{\Phi}_p\mathbf{B}^{[i]}\mathbf{\Phi}_{RF} \tag{3.2}$$

with $\mathbf{B}^{[i]}$ an $N \times N$ real matrix which models the link budget of each UT, and $\mathbf{\Phi}_p$ and $\mathbf{\Phi}_{RF}$ two diagonal complex $N \times N$ matrices which contain the phase terms arising in the signal propagation and the satellite RF chains, respectively. The real matrix $\mathbf{B}^{[i]}$ models each UT link budget and contains the satellite antenna radiation pattern, the path loss and the receive antenna gain. Its $m, n-$th entry is given by:

$$b_{mn}^{[i]} = \frac{\lambda}{4\pi d_m^i}\sqrt{G_R G_{mn}^{[i]}}, \tag{3.3}$$

with $d_m^{[i]}$ the distance between the i-th UT in beam m and the satellite (slant range), λ the wavelength, G_R the receiver antenna gain and $G_{mn}^{[i]}$ the multibeam antenna gain between the i-th single antenna UT located in the m-th beam and the n-th on board antenna feed.

The phase of the channel coefficients is divided in two terms. On the one hand, the diagonal matrix $\mathbf{\Phi}_p$ models the phase originated during the propagation. Here we assume that due to the big slant range the signals of all satellite antennas arrive at the UT with the same phase, this is the reason why this matrix multiplies $\mathbf{B}$ by the left. This phase is fixed for each user during all the simulations and it is uniformly distributed between all UTs. Their elements are $[\mathbf{\Phi}_p]_{nn} = e^{j\phi_n}$, where ϕ_n is a uniform random variable in $[0, 2\pi)$.

On the other hand, the diagonal matrix $\mathbf{\Phi}_{RF}$ models the random phase introduced by the local oscillators, which are independent for each payload

chain, and therefore this matrix multiplies $\mathbf{B}$ by the right. The elements of this diagonal matrix are $[\mathbf{\Phi}_{RF}]_{nn} = e^{j\theta_n}$, with $\theta_n \sim \mathcal{N}(0, \sigma_\theta^2)$. A value of 20^o for the variance is used as suggested in [23]. Contrary to ϕ_n, which is fixed during all the simulations, θ_n change with each time realization (transmitted frame).

Lastly, the channel matrix $\mathbf{H}^{[i]}$ for a given time realization is built row by row with the channel vector of the scheduled users, $\mathbf{h}_k^{\perp[i]}$. However, the gateway does not have access to the actual channel matrices $\mathbf{H}^{[i]}$ but to the matrices $\hat{\mathbf{H}}^{[i]}$, which include the nullification effect as well as estimation errors. This will be explained in detail in the next section.

3.2.2. Linear Precoding

Let $\mathsf{snr} = P/N_0$ be the per beam and per polarization transmit SNR (signal-to-noise ratio), with P the average per beam and polarization transmit power and N_0 the noise power at each receiver. The MMSE (Minimum Mean Square Error) precoding matrix is calculated by the gateway by using the average of nullified imperfect CSIT matrices $\hat{\mathbf{H}}^{[i]}$

$$\hat{\mathbf{H}}_{av} = \frac{1}{M}\sum_{i=1}^{M}\hat{\mathbf{H}}^{[i]}. \tag{3.4}$$

Using that matrix, the expression of the MMSE precoder [51] is given by

$$\mathbf{W}_{\text{mmse}} = \hat{\mathbf{H}}_{av}^H\left(\hat{\mathbf{H}}_{av}\hat{\mathbf{H}}_{av}^H + \frac{1}{\mathsf{snr}}\mathbf{I}_N\right)^{-1}. \tag{3.5}$$

In order to set the transmitted power, there are several options to normalize (3.5) such as SPC (Sum Power Constraint), MPC (Maximum Power Constraint) or PLPC (Per Line Power Constraint). SPC, which is used in this chapter, can be applied when the satellite payloads have Multiport Amplifiers (MPA) or Flexible TWTAs, that allow the total power (NP) to be shared among the different beams. In this case, $\mathbf{W}_{\text{mmse}}$ is multiplied by a scaling factor

$$\eta = \sqrt{NP/\text{trace}\{\mathbf{W}_{\text{mmse}}\mathbf{W}_{\text{mmse}}^H\}} \tag{3.6}$$

becoming the final precoding matrix

$$\mathbf{W} = \eta\mathbf{W}_{\text{mmse}}. \tag{3.7}$$

3.2.3. Precoded SINRs

If we refer to the rows of $\mathbf{H}^{[i]}$ and $\hat{\mathbf{H}}^{[i]}$ as $\mathbf{h}_k^{[i]\perp}$ and $\hat{\mathbf{h}}_k^{[i]\perp}$, respectively, and to the columns of $\mathbf{W}$ as $\mathbf{w}_k$ (the precoding vectors), the precoded SINR

for the i-th UT at the k-th beam can be calculated as

$$\mathsf{sinr}_k^{[i]} = \frac{|\mathbf{h}_k^{[i]\perp}\mathbf{w}_k|^2}{\sum_{j\neq k}|\mathbf{h}_k^{[i]\perp}\mathbf{w}_j|^2 + N_0} \tag{3.8}$$

and the SINR which the gateway estimates for the i-th UT at the k-th beam using the Imperfect CSIT is

$$\hat{\mathsf{sinr}}_k^{[i]} = \frac{|\hat{\mathbf{h}}_k^{[i]\perp}\mathbf{w}_k|^2}{\sum_{j\neq k}|\hat{\mathbf{h}}_k^{[i]\perp}\mathbf{w}_j|^2 + N_0}. \tag{3.9}$$

We use the notation $\mathsf{sinr}_k^{[i]}$ and $\hat{\mathsf{sinr}}_k^{[i]}$ to refer to the real and estimated precoded linear SINRs and we reserve the capital letters for their counterpart in logarithmic scale:

$$\mathrm{SINR}_k^{[i]} = 10\log_{10}\mathsf{sinr}_k^{[i]} \quad (dB) \tag{3.10}$$

$$\hat{\mathrm{SINR}}_k^{[i]} = 10\log_{10}\hat{\mathsf{sinr}}_k^{[i]} \quad (dB). \tag{3.11}$$

As we are interested in the SINR errors caused by nullification and their impact on the MCS selection, in the following sections we will characterize the SINR absolute error, defined as

$$e_k^{[i]} = \hat{\mathrm{SINR}}_k^{[i]} - \mathrm{SINR}_k^{[i]} \ (dB). \tag{3.12}$$

This is the difference between the SINR estimated by the gateway for each UT (using $\mathbf{W}$ and $\hat{\mathbf{H}}^{[i]}$) and the real SINR that the UT undergoes (which depends on $\mathbf{W}$ and the actual channel $\mathbf{H}^{[i]}$). If $e_k^{[i]}$ is positive it means that the gateway overestimates the SINR, so that if no appropriate margins are used in the MCS selection, transmission errors will take place as a result, even in the absence of other channel imperfections.

The selected metric to evaluate the nullification problem is the SINR in dB absolute error $e_k^{[i]}$ as defined in (3.12). This has the disadvantage of not being normalized and therefore it depends on the transmitted power; in consequence, for higher P the SINRs and the magnitude of the corresponding SINR absolute errors are higher too. Nevertheless, this metric, contrary to, for example, the relative error, can be translated directly into the margin necessary to subtract from the SINR estimated by the gateway, $\hat{\mathrm{SINR}}_k^{[i]}$, in order to guarantee a robust transmission. To obtain meaningful results with the SINR absolute error, the operation point is selected carefully to reflect a close to reality system.

3.2.4. Scheduling

In order to emulate a realistic scenario, in the simulations a basic scheduling algorithm is included. However, since the purpose of this chapter is to evaluate the SINR errors due to the carrier detection errors in the CSI and to propose a solution to overcome the problems they cause, we have maintained the scheduling part as simple as possible. The interested reader can refer to [55] and [101], for example, where the problem of scheduling in precoded systems is treated in more detail.

In this chapter we consider the application of two levels of scheduling: inter-beam and intra-beam. The main goal of the inter-beam scheduling is to allocate simultaneously transmitted frames to groups of UTs which have channels as orthogonal as possible in order to minimize the interference. This scheduling is done heuristically by following a geographical approach: all the beams are divided similarly in S circular sectors. All the simultaneous N frames aim UTs from homologous sectors in all the N beams, and all the sectors are served sequentially in a round robin fashion in order to schedule transmissions across the beam surface.

Regarding the intra-beam scheduling, its purpose is to group the users of each circular sector in multicast groups of M users which should meet two prescribed channel conditions. On the one hand, if the users have similar channel vectors $\mathbf{h}_k^{[i]}$, the precoding matrix $\mathbf{W}$ built from the average of the channel matrices $\hat{\mathbf{H}}_{av}$ is better fitted to the users channels. And, on the other hand, if the two users have similar SINRs, the user with the highest SINR does not lose so much capacity due to the use of the same MCS that the user with the worst conditions imposes. Inside each circular sector, the available users, i.e., positions from the available radiation pattern, are grouped in groups of M users with the MaxDist algorithm [41] using as features the channel vectors of the UTs and as norm the Euclidean distance.

For the sake of completeness, some details about how the scheduling is implemented to obtain the results of Sections 3.5, 3.6 and 3.7.2 are included here. The multicast groups with all the available positions (UTs) provided by the ESA's radiation pattern are created at the beginning of the simulation. And then, each time a sector is served, a different multicast group is selected following a random sequence which covers all groups inside a circular sector. This guarantees that all the users are served fairly and mimics the randomness nature of the traffic, which causes that UTs are scheduled with different users in different beams each time (although always with UTs from homologous sectors for maintaining the inter-beam scheduling policy).

3.3. Nullification Description

Precoding requires the transmitter to have knowledge of the users channel in the form of CSIT (Channel State Information at the Transmitter), since this is necessary to compute the precoding matrix $\mathbf{W}$ and estimate the precoded SINR of each user. The CSIT available at the gateway is usually imperfect, among other things, due to errors in the estimation of the coefficients of their CSI vectors per the users. However, the main cause of degradation of the CSI is the nullification effect. This is the inability of a receiver to estimate some of the coefficients of the CSI vector due to the low received power of the interfering beams. Thus, a large number of coefficients of the CSI matrix which is used by the gateway are unknown and replaced by zeros. This channel matrix, $\hat{\mathbf{H}}$, is said to have nullification errors.

Regarding the synchronization of the carriers transmitted by the beams, two types of satellite systems can be considered, with different consequences on the CSI estimation and nullification. With network synchronization [13], the waveforms received by the user, the desired carrier and the interfering carriers from adjacent beams, can be considered synchronous, i.e. the time delays between the different waveforms is much lower than the symbol period. In the specific case of DVB-S2X, with orthogonal Walsh-Hadamard (WH) sequences for the P pilots, the detection of the interfering carriers is limited mainly by the noise. The P pilots are sequences of known non-precoded symbols which receivers use to estimate the CSI, since the orthogonal WH sequences enable to differentiate up to 31 different interfering carriers. In this situation, the periodical repetition of the 36-pilots blocks along the bundled PLFRAME allows the detection of interfering carriers whose power is even 15 dB below the noise. If pilots across more than one superframe are averaged, this threshold can be further reduced. When the carriers are not synchronized, the orthogonality of the WH sequences can not be exploited, and the estimation becomes limited fundamentally by the interference [15]. In this asynchronous system, interfering carriers can be estimated even when they are approximately up to 15 dB below the desired carrier. Recent research showed that for levering all the potential of precoding in multibeam satellite systems the carriers should be synchronized to a great extent [13].

Mathematically, for simulating the nullification effect under purely asynchronous or synchronous carriers, we construct the nullified matrix $\hat{\mathbf{H}}$ in two different forms, depending on the system type:

1. **Asynchronous system**: Since our assumption is to work in clear sky conditions, the precoded C/N is relatively high so that the nullification

in the asynchronous sytem is driven mainly by the interference power more than the noise power. Therefore, we assume that the nullification is performed based on an I/C threshold T_h:

$$\hat{h}_{ij} = \begin{cases} 0 & \text{if} \quad 20\log_{10}|h_{ij}/h_{ii}| < T_h \\ h_{ij}, & \text{otherwise.} \end{cases} \tag{3.13}$$

2. **Synchronous system**: the nullification is performed based on an I/N threshold T_h:

$$\hat{h}_{ij} = \begin{cases} 0 & \text{if} \quad 20\log_{10}\sqrt{\mathsf{snr}}|h_{ij}| < T_h \\ h_{ij}, & \text{otherwise.} \end{cases} \tag{3.14}$$

In practice, the satellite networks are not fully asynchronous or synchronous, and the capability of a practical receiver to estimate a given interfering carrier depends on both C/N and C/I. Fig. 3.2 shows the (C/N, C/I) regions for which estimation is possible and for which nullification takes place. This performance of the CSI estimation is achieved by using the architecture shown in Fig. 3.1 and is referred hereafter as real nullification. From that figure it can be inferred that this realistic nullification is similar to a synchronous nullification with an I/N threshold $T_h = -12$ dB.

One of the most challenging and delicate phases for a precoded system to fully exploit its advantages is the proper and accurate recovery of the CSI for the main (reference) and the interfering carriers. The resulting synchronization chain incorporates a series of estimation procedures to recover impairments, ideally for each detectable carrier. Clearly, such an architecture can be optimal for completely asynchronous carriers. On the other hand, the cost for the optimal design is the computational complexity: for example, each timing recovery stage, one for each carrier, requires a digital interpolation which can be demanding for a commercial receiver.

Based on payload specification for a precoded system, which requires the transponder to be quasi aligned in time and frequency to maintain a constant phase amongst channels, a slightly simplified version of the receiver architecture is here used. To reach the aforementioned requirement, a satellite payload based on the shared use of a common Ultra Stable Oscillator, which provides the reference clock for all the down-conversion chains, has been assumed. Since each chain requires a dedicated Phase-Locked Loop (PLL), a residual per-beam phase noise is also expected. The block diagram of the CSI detection and estimation procedure performed by the satellite receivers is shown in Fig. 3.1.

After the coarse frequency recovery, a frame synchronization is performed based on the signature of the reference carrier, while, for each interfering

carrier, a verification stage based on carrier specific signatures is carried out, as can be seen in Fig. 3.1. The architecture assumes each carrier to be synchronous in frequency while not in phase and, as a consequence, it performs a frequency/phase noise tracking on the main carrier (reference waveform) and use the same tracking values to compensate the interfering waveforms. Using this approach, the CSI estimation stage provides a constant (during the channel coherence period) phase value normalized to the main signal phase variation. In fact, while the phase of the superimposed signal is affected not only by the phases of the relative channel vector but also from the receiver impairments (LNB frequency error, phase offset), what the terminal is practical able to estimate is the interfering waveforms phase differences from the tracking phase values of the main carrier, which should be almost constant (regardless of phase noise differences amongst transponders) within a certain coherence window.

In practice, the interfering carriers which are detected suffer also from estimation errors. Those carriers in the shaded area in Fig. 3.2 are estimated with an error which will be a function of both the C/N of the user and the C/I of the particular CSI coefficient. This error, modeled as Gaussian, will be higher for those carriers near the miss-detection boundary. The simulations included later in the chapter will account for these random estimation errors:

$$\hat{h}_{ij} = h_{ij} + e_{ij}, \tag{3.15}$$

with e_{ij} following a Gaussian distribution with parameters detailed in Tab. 3.1. The mean and standard deviation of the phase estimation error are given in absolute terms, whereas they are given in relative terms for the amplitude, to be independent of the channel coefficients magnitude. In both cases, maximum and mode are shown.

Fig. 3.3 contains a map with the number of estimated carriers per position across the satellite coverage using the real nullification model explained in the previous paragraph and for the system parameters of Tab. 3.2. It can be seen that the number of estimated carriers ranges between 1 and 15. In the edge of the satellite coverage typically less than 6 carriers are estimated and far from the edges this number increases, lying between 10 and 15. This last number is the half of the maximum number of carriers which can be differentiated with the WH sequences (32) and much lower than the total number of beams, 245.

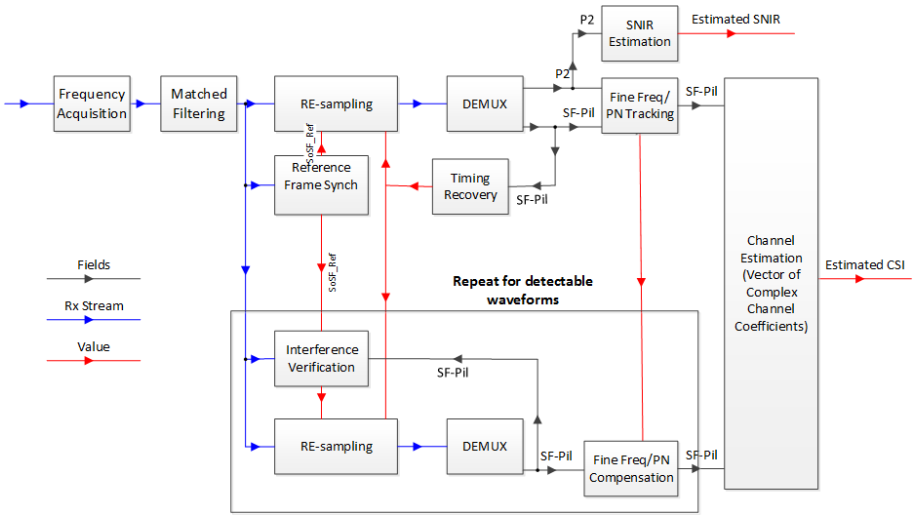


Figure 3.1: Architecture of the receiver for CSI detection and estimation.

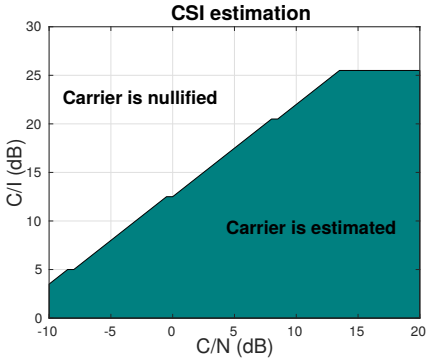


Figure 3.2: Range of values of C/N and C/I which allow carriers estimation in real nullification.

Relative error amplitude $|e_{ij}|$

Mean relative error	Maximum: 0.17	Mode: 0.05
Relative standard deviation	Maximum: 0.02	Mode: 0.005

Error phase $\angle\ e_{ij}(^o)$

Mean error	Maximum: 3^o	Mode: 0^o
Standard deviation	Maximum: 1.2^o	Mode: 0.7^o

Table 3.1: Characterization of the CSI estimation errors.

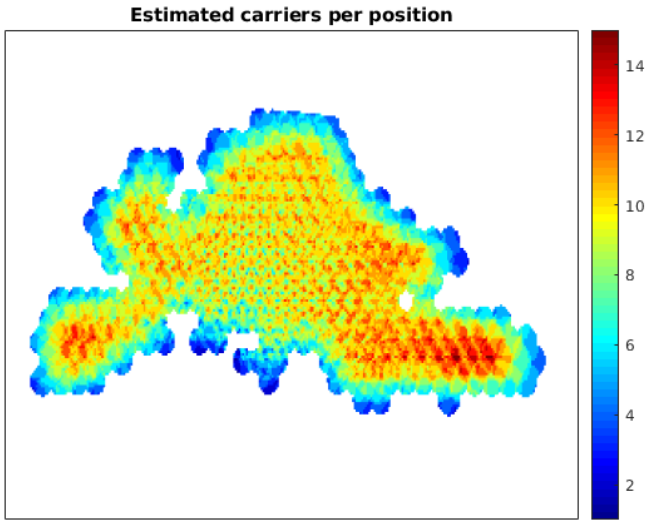


Figure 3.3: Map of the satellite coverage with the number of coefficients of the CSI vector estimated per position when using real nullification.

3.4. System Parameters

The main parameters of the Ka-band multibeam satellite system used in the simulations are collected in Table 3.2. The satellite covers the most of continental Europe as can be seen in Fig. 3.4, which shows the beam boundaries over the map of Europe. A 245 beams multibeam channel, generated based on ESA's radiation pattern [102], is used as basis for the simulations. For each beam a grid with discrete user positions is available, with the number of user positions of the beams ranging from 68 to 108. With the parameters from Table 3.1, the clear sky non-precoded C/N of the users ranges from 7 to 13.75 dB, as can be seen in the histograms of Figs. 3.5-3.7.

Three different system configurations are used in the simulations. One system with all the 245 beams (see Fig. 3.5) and two small size systems which consist of a cluster of only 9 beams, one located in the center of Europe (Fig. 3.6) and other located at the edge of the coverage, in the South West of the Iberian Peninsula (Fig. 3.7).

Parameter	**Value**
Satellite orbit	GEO
Downlink frequency	Ka-band (20 GHz)
Number of beams N	245
Size of multicast groups M	2
# of sectors per beam S (inter-beam sch.)	1, 4 or 6
Receive antenna gain	40 dBi
Noise temperature T	235.3 K
Receiver G/T	16.3 dB/K
Noise bandwidth (roll-off included)	12 MHz
Per beam and polarization transmit power P	0.15 W (21.76 dBm)
Total power	36.75 W (15.65 dBW)

Table 3.2: System parameters.

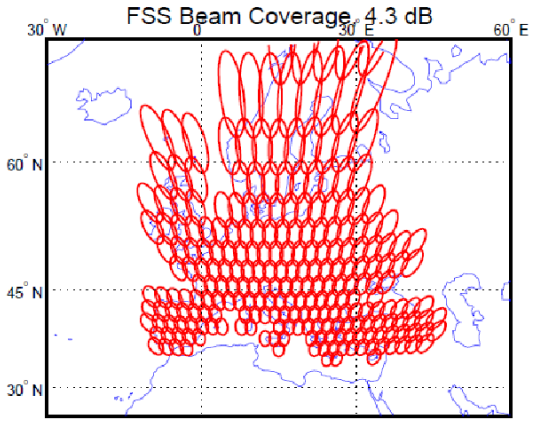
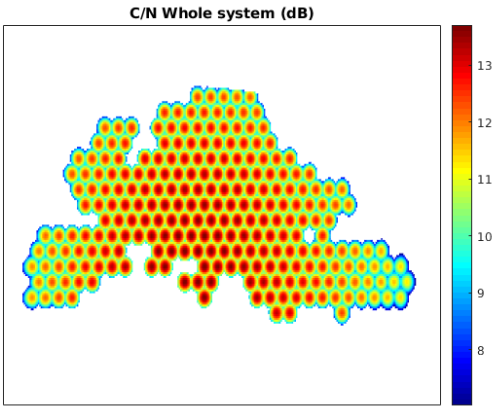


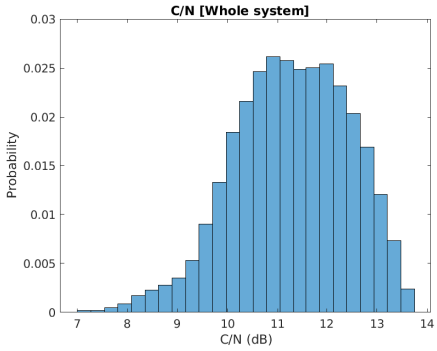
Figure 3.4: Map of the satellite coverage over Europe with the beams boundaries.

3.5. SINR Errors due to the Nullification

In this section simulation results of the SINR absolute error due to nullification and CSI estimation errors will be provided from three different perspectives. Firstly, Subsection 3.5.1 presents several statistics of the aggregated SINR error, testing all user locations on the satellite coverage and showing the global results. Then, Subsection 3.5.2 analyses the SINR absolute error adding the spatial dimension of the data in order to reveal the geographical dependence of the error. Lastly, Subsection 3.5.3 focuses the attention on the statistics of the SINR and the SINR absolute error from the point of view of specific users, i.e., at a fixed location.



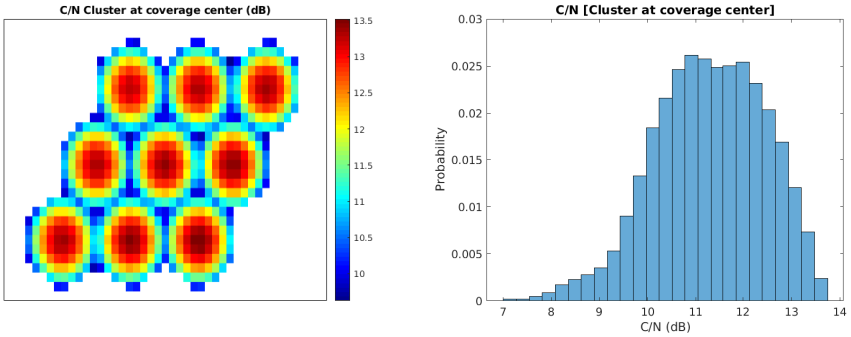
(a) Map of C/N (dB).



(b) Histogram of the C/N distribution.

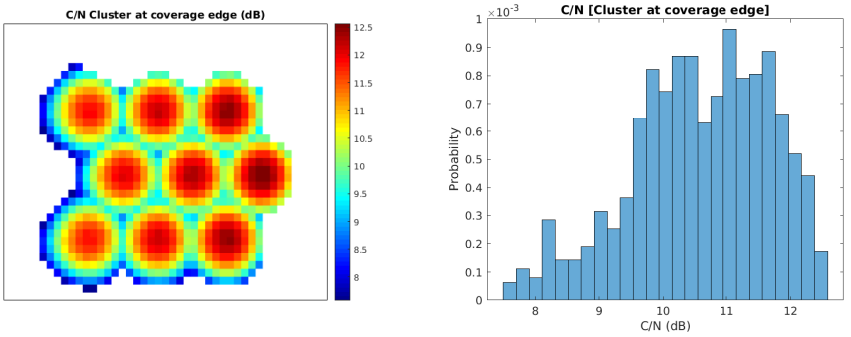
Figure 3.5: Map of the C/N per position in the satellite coverage and histogram with its distribution for the system with all the 245 beams.

Each simulation comprises the transmission of $1,000$ frames. Previously, multicast groups with $M = 2$ users are created in each circular sector of each beam by using the scheduling algorithms detailed in Section 3.2.4 and considering as active all the positions of the beam. Then, once selected the $2N$ users scheduled in each frame, the channel matrices $\mathbf{H}^{[i]}$ are built with the channel vectors of the users, the imperfect CSIT with the nullification and estimation errors $\hat{\mathbf{H}}^{[i]}$ is obtained and the precoding matrix $\mathbf{W}_{mmse}$ is calculated using the average channel matrix $\hat{\mathbf{H}}_{av}$. Afterwards, the actual UT SINRs, $\text{SINR}_k^{[i]}$, and the corresponding estimated value by the gateway, $\hat{\text{SINR}}_k^{[i]}$, are obtained. Lastly, its difference gives the SINR absolute error $e_k^{[i]}$, which is analysed from three different perspectives along the following subsections.



(a) Map of C/N (dB) (b) Histogram of the C/N distribution.

Figure 3.6: Map of the C/N per position and histogram with its distribution for the cluster of 9 beams located at the center of Europe.



(a) Map of C/N (dB) (b) Histogram of the C/N distribution.

Figure 3.7: Map of the C/N per position and histogram with its distribution for the cluster of 9 beams located at the South West of the Iberian Peninsula.

3.5.1. Aggregated Results

In this subsection a collection of results of the SINR absolute error, as defined in equation (3.12), is presented. Firstly, the bar diagrams of Fig. 3.8a show the maximum SINR absolute error by using the simple model for synchronous nullification of equation (3.14). Results are shown for the system with 245 beams, with 6 circular sectors for the inter-beam scheduling and for 2 different values of transmitted power: the nominal operation point of Table 3.2 and also a P 3 dB lower. Three different I/N nullification thresholds are studied in order to link the performance in the interfering carriers detection with the maximum SINR error obtained. Moreover, results for the realistic nullification depicted in Fig. 3.2 are included for comparison. It follows from Fig. 3.8a that, for the maximum SINR error being lower than

1 dB, an I/N close to −20 dB is needed in synchronous systems. This graphic illustrates that, if CSI detection is improved, by averaging the pilot symbols during longer periods, the ensuing SINR error due to the nullification will be lower. Lastly, it can be seen how the maximum SINR absolute error depends also on the operation point. If the power is lower, the SINR will be lower too and the magnitude of the errors will be also reduced.

Now, Fig. 3.8b compares the maximum SINR absolute error when real nullification is used, i.e., applying both C/N and C/I thresholds of Fig. 3.2 to decide if a coefficient of the channel vector is nullified. We show results with the nominal power of Tab. 3.2 for three different systems: one with all the satellite beams ($N = 245$), one consisting of a cluster with only $N = 9$ beams located at the center of Europe, and another with a cluster of also $N = 9$ beams but at the edge of the coverage. We also compare two scenarios, one without inter-beam scheduling (blue bars) and another with a inter-beam scheduling algorithm which creates 6 circular sectors in the beams.

The maximum error for the complete system using inter-beam scheduling is 1.6 dB, which is reduced to 1.4 and 1.0 dB for the systems with a reduced set of beams. The reason why the cluster at the edge has a lower error is because in that location the C/N is smaller compared with the other two scenarios, as it can observed in the histograms of Figs. 3.5-3.7. Regarding the error in the 245-beams system, it is bigger than in the reduced-size clusters because the proportion of nullified coefficients in the channel vector of N is much higher. The accumulation of many small differences among $\mathbf{h}_k^{[i]\perp}\mathbf{w}_j$ and $\hat{\mathbf{h}}_k^{[i]\perp}\mathbf{w}_j$ when calculating the numerator and denominator of the SINR in the Equations (3.8) and (3.9) causes the error to be bigger in the complete system. Lastly, when comparing the blue and red columns we observe that the absence of a proper inter-beam scheduling algorithm increases slightly the error.

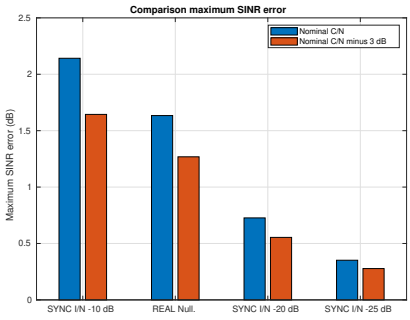
The bottom plot of Fig. 3.8 shows, for the 245-beams system with 6 sectors in the inter-beam scheduling, the experimental Complementary Cumulative Distribution Function (CCDF) of the SINR error with the real nullification. The CCDF provides the probability of the SINR error being larger than a given abscissa value and offers some insight into the value of the margin required for guaranteeing a given target Frame Error Rate (FER) across the satellite footprint. Frame refers to the physical layer unit which contains a codeword with data for M users encoded using the same MCS. For example, from this plot it can be stated that the subtraction of a 1.3 dB margin from the estimated SINR by the gateway should help to guarantee a FER of 10^{-4} for the nominal transmit power case.

These results show that even in the absence of other imperfections, nullification itself can degrade the performance of multibeam satellite systems which use precoding unless proper countermeasures are applied. An error larger than 0 dB means that the gateway is overestimating the actual UT SINR, and it implies that if $\widehat{\text{SINR}}_k^{[i]}$ is used directly to allocate an MCS to the UT at beam k, an error in the transmission will occur whenever the error is bigger than the distance of $\widehat{\text{SINR}}_k^{[i]}$ to the threshold SINR of the MCS used in the transmitted frame. And since the spacing between two DVB-S2X consecutive MCS is typically between 0.3 and 1.0 dB, and as the SINR errors exceed these values with a non-negligible probability, the FER will be higher than the QEF target of 10^{-5}, one erroneous frame out of $100,000$.

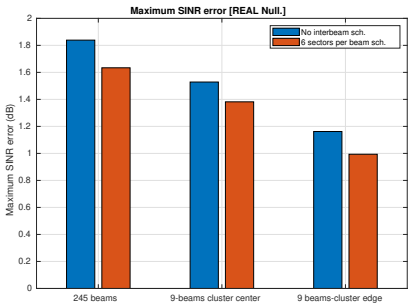
3.5.2. Spatial Analysis

In order to detect a potential geographical dependence of the SINR absolute error $e_k^{[i]}$, we have used the same simulation data employed to plot the results of Fig. 3.8 to create the maps of Fig. 3.9. The map of Fig. 3.9b shows the maximum error per position after the transmission of $1,000$ frames in the complete system, with the nominal power and 6 sectors for the inter-beam scheduling. Although there are some hot points where the error reaches values as high as 1.5 dB, the vast majority of the positions have a lower error slightly higher than 0 dB in the coverage center and even lower than 0 dB in some edge areas. The difference between the lowest and highest maximum error per position is as big as 2.5 dB. Then, the map of Fig. 3.9a sums up the information of the map of Fig. 3.9b showing the maximum error in each beam, aggregating all the beam positions. Again, the same trend is observed: beams in the center of the coverage tend to have higher errors and the maximum SINR absolute error is lower at the edge of the coverage.

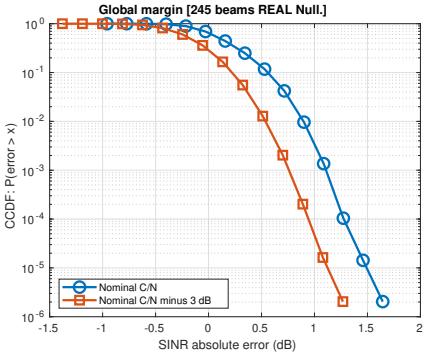
A rough countermeasure for the errors caused by the nullification could be the subtraction of a global margin to the SINRs estimated by the gateway. This margin could be obtained from the CCDF of the SINR absolute error (shown in Fig. 3.8c), by using the target FER. However, the results of the geographical distribution of the error show that a global margin to be applied throughout the whole footprint can be quite inefficient, since the error magnitude differs significantly across the footprint. Furthermore, remarkable differences appear within the same beam from one position to another. As a result, a per user margin seems a better solution.



(a) Effect of operation point (C/N) and detection performance (245-beams system, synchronous nullification).

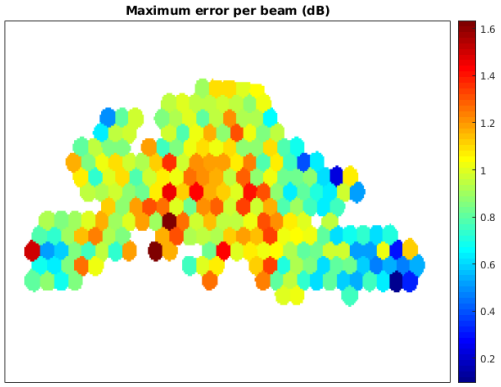


(b) Effect of inter-beam scheduling and system size (real nullification).

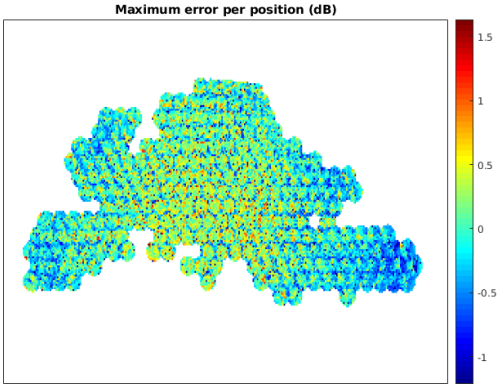


(c) SINR error CCDF with real nullification (245-beam system).

Figure 3.8: Maximum SINR absolute error over 1,000 realizations in (a) and (b), and CCDF of the SINR absolute error with real nullification in (c).



(a) Maximum error per beam.



(b) Maximum error per position.

Figure 3.9: Maps of the SINR maximum error with real nullification for the complete system, nominal power and 6 sectors per beam.

3.5.3. Statistical Analysis at Fixed Locations

This subsection deals with the characterization of the SINR and the associated SINR error in a fixed position. If a pair of users is always scheduled together with the same set of users in all the other $N-1$ beams, the SINR and the SINR absolute error will be constant with time in the absence of other imperfections such as channel fading, phase noise or Gaussian estimation errors in the CSI. However, the user grouping is not kept fixed with time, due to the random nature of the traffic, among other reasons.

Focusing on a specific user, Fig. 3.10 shows the time evolution of the ac-

tual SINR, $\mathsf{SINR}_k^{[i]}$, and the associated error, $e_k^{[i]}$, along with their respective empirical probability density functions. The dashed lines of the upper-left plot represent the threshold SINRs of the DVB-S2X MCS. It can be seen that both SINR and SINR error tend to have a Gaussian distribution in all the positions examined and for three different scenarios: without fading, and adding Rician fading with Rice factors of $K = 25$ and $K = 34$ dB. Furthermore, we observe that both SINR and SINR error show a stationary behavior, with their respective mean and variances remaining quite stable with time.

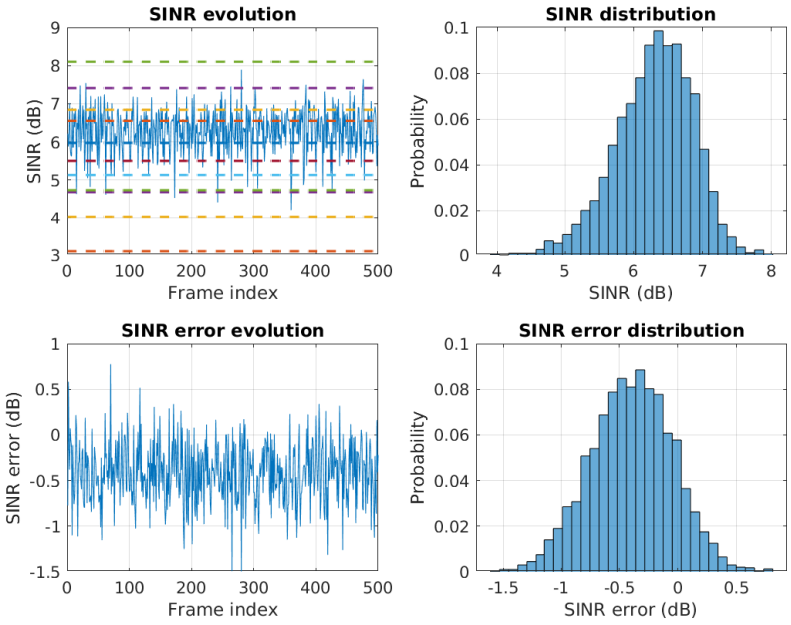


Figure 3.10: Example of SINR and SINR error time evolution and their distribution, at a fixed location.

3.6. System Throughput and Back-off Margin

In the previous section several statistics of the SINR absolute error caused by the nullification have been shown. Although the maximum SINR error in the whole footprint can reach high values (take, for example, 1.6 dB with realistic nullification, nominal transmit power and 6 sectors per beam in the inter-beam scheduling), in general, the error is much lower in the vast majority of the positions, as can be seen in Fig. 3.9b.

When the SINR absolute error of a user is negative, it means that the

gateway is underestimating its actual SINR. Therefore, it will allocate a spectrally underperforming MCS, leading to an under-utilization of the available resources. On the other hand, when the SINR absolute error of a user is positive, the gateway is overestimating its actual SINR. This case is even more critical than the previous one, because the MCS allocated to the frames of the user is very likely beyond its decoding capabilities, leading to an erroneous transmission which increments the FER and implies a waste of resources.

One possible solution to avoid the errors caused by the SINR overestimation consists on applying a safety or back-off margin to the users' SINR estimated by the gateway. The subtraction of a positive margin from the SINR when choosing the MCS of each transmission helps to maintain the FER around a predefined target level. Depending on the granularity of this margin three options can be taken: a unique global margin for all the satellite coverage, a beam-dependent margin, and a user-dependent margin.

In this section we want to find out the throughput degradation due to imperfect and nullified CSIT depending on the type of margin used with respect to a system with perfect CSIT. Fig. 3.11 compares the achieved throughput with perfect CSIT and imperfect CSIT due to the nullification; 6, 4 and 1 circular sectors per beam are considered for the inter-beam scheduling, and the three types of the aforementioned margins are evaluated. For each simulation the average throughput per frame is calculated after $10,000$ realizations, by using the spectral efficiency of DVB-S2X MCS, shown in Fig. 3.12, and grouping $M = 2$ users per frame. Then, the throughput of each simulation is normalized after dividing it by the throughput of the perfect CSIT case with 6 sectors. It can be readily seen that the use of a proper inter-beam scheduling algorithm with 4 or 6 circular sectors per beam increases the throughput. The strong impact on the throughput of the type of margin can also be observed. More specifically, for 6 sectors per beam, if we use a global margin by subtracting the maximum error observed in the system from all the SINRs calculated by the gateway before selecting the MCS, the throughput falls 21% as compared with the perfect CSIT scenario (see purple bar). Instead, if a different margin per beam is applied, the throughput loss is reduced to 16% (see yellow bar). Fig. 3.9a displays differences in the maximum error per beam of about 1.4 dB, giving credit to the benefits when increasing the granularity of the margin.Lastly, the throughput when an independent per user margin is employed is shown with orange bars in Fig. 3.11. In this case, the throughput is obtained after subtracting from each user SINR the maximum SINR error that is observed in its position. As Fig. 3.9b anticipates, the use of an individual margin per UT is clearly the best option to compensate for the errors that the nullification causes in the SINR estimation. By employing a user-dependent margin, only a 6% of

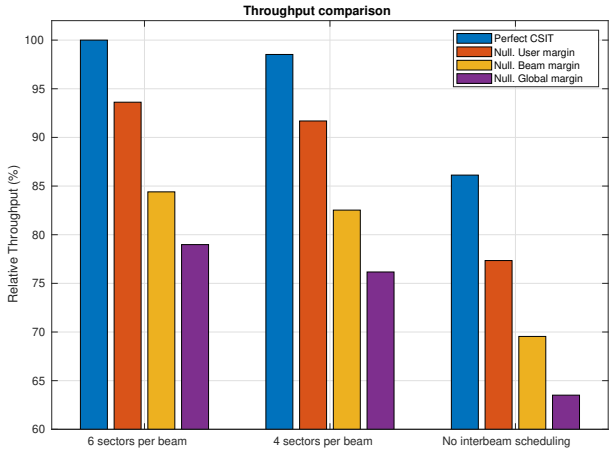


Figure 3.11: Throughput comparison of a 245 beams system with perfect and nullified CSIT depending on the type of margin used.

throughput is lost when compared with the perfect CSIT scenario.

As discussed, nullification is a major impairment affecting the estimation of the SINR at the transmitter. However, additional sources of imperfection affecting also the quality of the available CSIT take place in practice such as the outdated estimation of the SINR, due, for example, to rapid channel variations, at least faster than the tracking ability of the transmit-receive loop. Due to this and other practical reasons, the user-dependent margin should be adaptive, so that the rate of erroneous frames can be properly bounded. We detail in the next section how this margin can be conveniently adapted; a multibeam satellite system with ACM and precoding, severely affected by nullification and Rician fading, will be simulated for illustration purposes.

3.7. Link Adaptation Algorithm

In this section we introduce an enhanced link adaptation algorithm taken from [29], which is based on the use of an adaptive margin as a countermeasure for the errors that nullification and other CSI imperfections can cause. The algorithm performance will be evaluated through simulations. The aim of the algorithm is to select the MCS of each frame for each UT, maintaining the FER around a target value p_0. The link adaptation procedure must be performed at the gateway side rather than at the UTs, since the gateway is the only entity with all the required information to calculate the SINRs,

namely, the channel vector of all scheduled users and the precoding matrix employed in each frame. Due to all the considerations explained previously the gateway will run an independent margin per UT.

3.7.1. Algorithm Description

The adaptation of the margins using the received acknowledgments is commonly known as outer loop link adaptation (OLLA), in contrast to the inner loop, which consists on selecting a MCS using the margins. We assume that for each received frame the UTs feed back the outcome of the frame decoding to the gateway, which is eventually used to update the margin of each user. This variable $\epsilon_{k,t}^{[i]}$ takes the value 1 to indicate that an error has occurred at decoding of the frame number t for i-th user in the k-th beam (NAK), and it takes the value 0 otherwise (ACK). We should note that the ACK/NAK feedback is used only for adapting the margins of the link adaptation algorithms, without any retransmission of the frames at the physical layer due to the long propagation delay. Moreover, the burden of sending these acknowledgments is low compared with all the information a receiver should feed back in a system with precoding.

The MCS selection of frame t intended for the multicast group of M users in the k-th beam, $\mathrm{MCS}_{k,t}$, is performed by means of a Lookup Table (LUT), which receives as inputs the estimation of the UTs SINR, $\widehat{\mathrm{SINR}}_{k,t}^{[i]}$, computed by the gateway with the imperfect CSIT, and the adaptive margins of these users, $m_{k,t}^{[i]}$. The DVB-S2X framing structure imposes the use of a single MCS for the M users whose data is multiplexed in a frame. Therefore, the UT with the worst signal quality will determine the MCS of the multicast group. Mathematically, we can express the MCS selection as

$$\mathrm{MCS}_{k,t} = \prod \left(\min_{i} \ \widehat{\mathrm{SINR}}_{k,t}^{[i]} + m_{k,t}^{[i]} \right), \tag{3.16}$$

where the LUT $\Pi(\cdot)$, which is depicted in Fig. 3.12, is a staircase function which maps SINR intervals to MCS following the performance tables of the DVB-S2X standard [6]. This function is usually referred to as the inner loop for link adaptation [78]. The reverse operation is done by means of the function $\amalg(\cdot)$, which returns the threshold SINR of a given MCS.

For simulation purposes we consider that a decoding error event occurs whenever the actual SINR at the UT is lower than the SINR threshold of the MODCOD used to transmit the corresponding frame. This is expressed

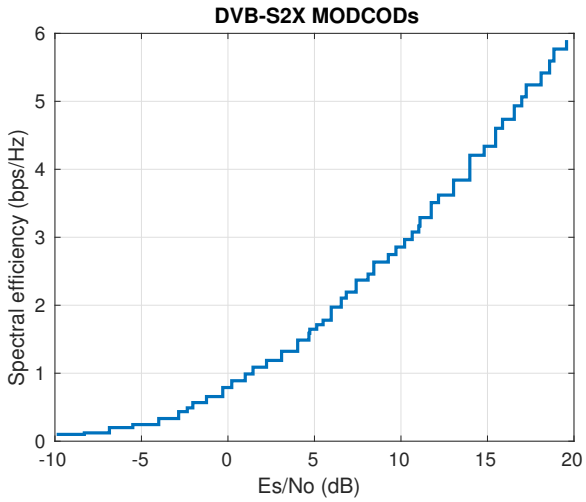


Figure 3.12: Staircase function representing the LUT of the DVB-S2X MCS, with their spectral efficiency and decoding SINR threshold.

by means of the indicator function $[\cdot]$ as

$$\epsilon_{k,t}^{[i]} = [\underbrace{\text{SINR}_{k,t}^{[i]}}_{\text{Actual SINR}} < \underbrace{\text{II}\,(\text{MCS}_{k,t})}_{\text{Threshold SINR of the MCS}}]. \tag{3.17}$$

The most classic OLLA strategy consists on increasing the margin slowly $\Delta_{ack} = \mu \cdot p_0$ each time an ACK is received and decreasing it more drastically, $\Delta_{nak} = \mu \cdot (1 - p0)$, when an error occurs [78]. If the ratio between up and down steps equals $\Delta_{ack}/\Delta_{nak} = p_0/(1-p_0)$, a target FER of p_0 can be guaranteed. Moreover, there is a trade-off when choosing the step size: a high value of μ will increase the convergence speed of the margin at the expense of a significant oscillation at the steady state. In [29] the authors tried to solve this trade-off by proposing a new OLLA which ensures a fast convergence after a transient, occurring for example after a change in the SINR error behavior, while maintaining a good performance during steady state with infrequent margin updates. Hereafter, we provide a brief summary of this enhanced OLLA algorithm, which will be used in the simulations described in the next subsection.

The purpose of this OLLA algorithm is to maintain the FER inside the range $[p_1,\ p_2]$, aiming at having a target FER of p_0. The transmitter (the gateway in our case) should keep an estimation of the instantaneous FER of each user, which is calculated with the acknowledgments of the UTs. If FER $\ll p_1$ or FER $\gg p_2$, then the classic OLLA controller described previously is used to update the margin. The steps have a moderate size in

order to increase the convergence speed and arrive soon to a situation where the FER is close to the target value p_0.

When the FER does not take extreme values (far below p_1 or far above p_2), a sequential hypothesis testing (SHT) is performed and the margin is updated only if its results are conclusive. If hypothesis H0 is true (FER $<$ p_1), the margin is increased by a quantity $\lambda \cdot p_0 \cdot N_{ack}$. On the other hand, if hypothesis H1 is true (FER $> p_2$), the margin is decreased $\lambda \cdot (1 - p_0) \cdot N_{nak}$. In the previous equations N_{nak} represents the number of NAKs required to accept the hypothesis H1 and N_{ack} the number of ACKs required to accept the hypothesis H0. Table 3.3 collects the values of the parameters of OLLA algorithm used in our simulations.

Parameter	**Value**
Target FER p_0	10^{-1}, 10^{-2}, 10^{-3} or 10^{-4}
Desired FER range $[p_1,\ p_2]$	$[0.9 \cdot p_0,\ 1.1 \cdot p_0]$
Classic OLLA range of use	FER $\leq 0.1 \cdot p_0$ or FER $\geq 10 \cdot p_0$
μ	0.05
λ	0.01
N_{nak}	3
N_{ack}	$N_{nak} \cdot (1 - T)/T^2$
Initial margin	0

Table 3.3: OLLA algorithm parameters.

This algorithm entails a relatively low complexity if compared with that of the other tasks that a gateway must perform. All the required mathematical operations are very simple (just a few sums and products), with the complexity scaling linearly with the number of outer loops (the number of users in the system). For each loop, only three variables need to be stored: the value of the margin, the estimation of the FER and the internal state of the SHT controller.

A key point is the margin initialization when a UT gets connected. Ideally, it should be near the stationary margin (if any) to avoid long underperforming transient periods. However, the derivation of the optimum margin is not an easy task since it depends on the Probability Density Function (PDF) of the SINR, the PDF of the SINR error and the MCS SINR thresholds in the UT SINR region. In an ideal system with infinite MCS, which can be emulated in practice using a very large number of MCS very close to each other in terms of SINR thresholds, it can be tested experimentally that the optimum margin depends only on the statistics of the SINR error. Namely, for a target FER of p_0, the optimum margin is $-m^*$ if $\mathbb{P}\left(e_k^{[i]} > m^*\right) = p_0$. However, in a practical DVB-S2X-based system, the threshold SINRs of the

MCS are not equally separated; the distance among contiguous MCS is in the range of the SINR error magnitude, which can span several MCSs, as shown in Fig. 3.10. Thus, predicting the optimum margin for a specific user is not straightforward.

The DVB-S2X standard [Annex D.5] [6] considers that the receiver should feed back the measured precoded SINR (therein referred as CNI). This SINR reported by the UTs, along with the estimation of the users SINR, $\hat{\text{SINR}}_k^{[i]}$, performed by the gateway previously to MCS allocation, allows to calculate the SINR absolute error $e_k^{[i]}$. Therefore, the gateway can obtain some basic statistics of both precoded SINR and SINR error after a moderate number of frames. As possible future work, some supervised Machine Learning (ML) technique [89] could be used to perform a margin re-initialization after some transitory happens, avoiding long convergence times. Since the gateway manages a high number of connections, it can use information such as the margins of connections which comply with the FER constraint and their SINR and SINR error statistics, to train a neural network, for example, and use it to decide the right margin of other connections.

3.7.2. ACM Simulations Results

We have performed ACM simulations of 10 fixed multicast groups of $M = 2$ users for a complete system with $N = 245$ beams. Each multicast group belongs to a different beam, and the 10 beams are located far appart from each other. In the scheduling algorithm $S = 6$ circular sectors have been chosen for the inter-beam scheduling, following Section 3.2.4. The transmission of F frames is simulated, with F equal to 1 000, 5 000, 50 000 and 500 000 for the target FERs of 10^{-1}, 10^{-2}, 10^{-3} and 10^{-4}, respectively. The set of MCS, taken from the DVB-S2X standard, can be seen in Fig. 3.12. For MCS selection purposes, the enhanced outer and inner link adaptation algorithms detailed in the previous subsection for MCS were used. At the end of each test the experimental FER, evaluated during the last half of the simulation, is obtained.

Three different types of simulations were performed. In the first one, the CSIT degradation is only due to the real nullification and CSI estimation errors, as explained in Section 3.3. In the other two, Rician fading is added, with Rice factors of 25 and 34 dB, for modeling the terrestrial and aeronautical channels, following [38] and [6]. We assume that the realizations of the fading coefficients of different users and frames are uncorrelated. As worst case, the fading coefficient f_k of the CSIT vector, which the gateway uses

	Target $p_0 = 10^{-1}$	**Target $p_0 = 10^{-2}$**	**Target $p_0 = 10^{-3}$**	**Target $p_0 = 10^{-4}$**
Without fading	0.1038	0.0099	$9.96 \cdot 10^{-4}$	$1.05 \cdot 10^{-4}$
Rice factor $K = 25$ dB	0.1046	0.0108	$1.00 \cdot 10^{-3}$	$1.02 \cdot 10^{-4}$
Rice factor $K = 34$ dB	0.1038	0.0100	$9.76 \cdot 10^{-4}$	$1.05 \cdot 10^{-4}$

Table 3.4: Average FER of the 10 users with the worse channel conditions of the 10 multicast groups of the ACM simulations.

to compute the estimated SINR,

$$\hat{\mathsf{sinr}}_k = \frac{|f_k \hat{\mathbf{h}}_k^{\perp} \mathbf{w}_k|^2}{\sum_{j \neq k} |f_k \hat{\mathbf{h}}_k^{\perp} \mathbf{w}_j|^2 + N_0}, \tag{3.18}$$

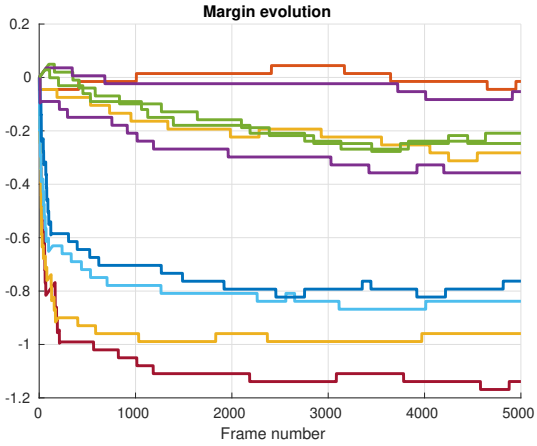
is uncorrelated with the fading coefficient $f_k^{'}$ of the final actual UT SINR,

$$\mathsf{sinr}_k = \frac{|f_k^{'} \mathbf{h}_k^{\perp} \mathbf{w}_k|^2}{\sum_{j \neq k} |f_k^{'} \mathbf{h}_k^{\perp} \mathbf{w}_j|^2 + N_0}, \tag{3.19}$$

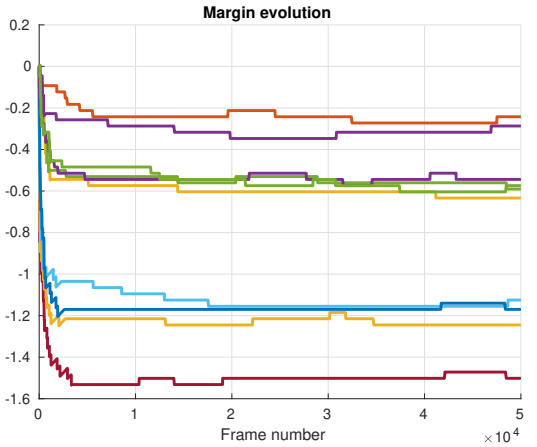
in order to take into account the effect of the delay in the CSI report.

Table 3.4 shows the mean FER of the simulations, obtained after averaging the experimental FER of 10 users of the simulations. It can be seen that in all cases, with and without fading, it is very close to the target value. Moreover, the FER of the users ranges typically between 80% and 120% of the target FER p_0. We should note that, since the multicast groups are fixed during all the simulations and share the same MCS, the user with the lowest SINR achieves the target FER of approximately p_0, whereas the user with the highest SINR has a much lower FER, usually close to 0. When both users of the multicast group have similar SINRs, in both cases the experimental FER is around the target value.

Now, Fig. 3.13 shows the evolution of the margin for those users with the lowest SINR in each multicast group, the ones whose experimental FER matches the target FER. The back-off margins are more conservative in those cases with a lower target FER, as can be noticed in Fig. 3.13. The different behaviour of the link adaptation algorithm depending on the measured FER can also be observed in Fig. 3.13b. Initially, the classic OLLA is active, updating the margin with each received acknowledgment, and using high steps in order to achieve a fast convergence; eventually, SHT comes into play so that the margin update becomes less frequent, and only conclusive hypothesis tests trigger a margin change.



(a) Target FER 0.01



(b) Target FER 0.001

Figure 3.13: Margin evolution for the users with the lowest SINR of each multicast group (simulations with Rician fading with Rice factor $K = 25$ dB).

3.8. Conclusions

The miss detection of interfering carriers was considered in this chapter for a multibeam satellite, analysing the errors on the precoded SINR of the users which the gateway must calculate to perform MCS allocation. This leads to an increment on the rate of erroneous frames if no proper countermeasure is applied. The study of the SINR absolute error revealed that there are important differences on the maximum error among different beams, and even among different positions within the same beam. In or-

der to ensure the robustness of the communications the use of an adaptive margin per user was proposed, which outperformed a global or beam-based margin in terms of throughput loss. The proposed adaptive margin helps to counteract the errors in the SINR of the users which are estimated by the gateway. This margin, independent for each UT, is updated by the gateway based on the ACK/NAK feedback from the users and is employed in the link adaptation for MCS selection. Simulation results showed that by using this link adaptation algorithm, a predefined target FER can be achieved, even in the presence of SINR calculation errors due to the miss detection of some entries of the CSI. It was also checked experimentally that the adaptive margin helped to guarantee the objective FER when Rician fading was added in the simulations on top of the nullification itself.

Chapter 4

Evaluation of the Spatial Modulation Transmission Capacity

This chapter is adapted, with permission of the coauthors, from: "A. Tato, C. Mosquera, P. Henarejos, A. Pérez-Neira. Neural Network Aided Computation of Generalized Spatial Modulation Capacity[1]*, in 2019 27th European Signal Processing Conference (EUSIPCO), Sept. 2019, A Coruña (Spain)", [99]. And from Arxiv: "A. Tato, C. Mosquera, P. Henarejos, A. Pérez-Neira. Neural Network Aided Computation of Mutual Information for Adaptation of Spatial Modulation, Arxiv, 2019", [98].*

4.1. Introduction

The evaluation of the achievable physical layer rate of a given modulation scheme is an important theoretical problem highly relevant in practice. Most modern communication standards, e.g., [8], [5] or [9], incorporate some sort of Adaptive Coding and Modulation (ACM) mechanism, generically known as link adaptation. This consists typically on varying the modulation order and/or the coding rate of the channel encoder to track the varying channel conditions. The ultimate goal is to adjust the transmitted bit rate to the information that the channel can support for a given bit error probability.

Link adaptation makes it necessary for the transmitter to estimate some-

[1]Publication pending.

how the mutual information (MI) between the transmit and received waveforms on a per-frame basis, so that the most efficient Modulation and Coding Scheme (MCS) can be chosen. In most cases, the receiver computes some metric related to the MI and sends it back to the transmitter end. This metric can be in the form of the average or effective Signal to Interference and Noise Ratio (SINR), or some Channel Quality Indicator (CQI) specifically suited to the set of MCS available to the transmitter [83]. In essence, the receiver must estimate the maximum amount of information that can be transmitted reliably through the channel; for all this, the estimation of the MI plays an instrumental role.

A particular family of modulation schemes, known as Index Modulations (IM) [19], has attracted a great deal of interest in the last few years. Among others, we can cite Spatial Modulation (SM) [61]-[111], Generalized Spatial Modulation (GSM) [65]-[113], or Polarized Modulation (PMod) [47]. SM and its more sophisticated variant GSM are proposed for the next generation of wireless networks due to several advantages. In comparison to single-antenna techniques, the spectral efficiency increases, with simpler and more energy efficient hardware than in other multi-antenna techniques [19]. PMod is studied in [47], where the authors propose its use to increase the spectral efficiency of next generation mobile satellite communications; if Multiple-Input-Multiple-Output (MIMO) signal processing techniques are applied to Dual Polarization (DP) satellite systems, the performance of single-antenna (or single polarization) links can be notably enhanced. DP schemes were also highlighted in [66] as a means to improve the satellite coverage in remote areas to serve the increasing number of Internet of Things (IoT) devices.

In this chapter, a novel method based on Machine Learning (ML) is presented, to compute the MI without Channel State Information at the Transmitter (CSIT) of a 2×2 SM system, including also its generalization to an arbitrary number of antennas. This MI is also named the constrained capacity, since it is the capacity constrained to a given constellation. The results are also valid for other types of IM, like PMod. In addition, we also provide a method to calculate the unconstrained capacity of a GSM link, also using ML. The MI and capacity calculations are needed, for example, in adaptive SM or GSM systems, where the transmitter uses these metrics, obtained and fed back by the receiver, to select the proper MCS. Results requiring numerical integration, for example by means of Monte Carlo simulations, can be found in the literature [65] and [49]. One contribution of this chapter is that it explores a radically new approach to solve an essential problem in the practical application of Information Theory: the mutual information of non-conventional modulations is computed by means of ML tools.

The use of ML at the physical layer of communication systems is gaining momentum, see the recent surveys [11] and [89]. In particular, Neural Networks (NN) have been successfully used for channel estimation and equalization [64], signal recognition and modulation classification [34], [63], detection in MIMO Generalized SM [59], and learning of physical layer parameters in Cognitive Radio [37], among others. In [54] and [72] NNs are applied to perform link adaptation in multicarrier systems. In [18], a Multilayer Feedforward Neural Network (MFNN) is used to predict the performance of a WiFi cell. A deep NN is proposed in [90] to decide the optimal power allocation in a wireless resource management problem. In this latter reference the NN is used to obtain the optimal power allocation values, much more efficiently (by speeding up the computational time in several orders of magnitude) than the baseline iterative algorithm which solves the corresponding non-convex optimization problem. Besides NNs, Support Vector Machines (SVM) have also been studied for the selection of physical layer parameters in communication settings with a large number of degrees of freedom [79].

The work of the current chapter applies a one-hidden layer MFNN as a facilitator scheme to compute the MI and the capacity in adaptive SM or GSM links, based on some specifically selected input features which can be easily obtained from the MIMO channel matrix, together with the Signal to Noise Ratio (SNR). Although in the last years several works applying ML to communications have appeared, the application of neural networks to obtain these information theory related metrics of a non-conventional modulation scheme is something completely new. To the best of our knowledge, it is the first time that a NN is proposed to estimate the MI or the capacity of a channel. In the particular scenario of SM or GSM, the evaluation of the capacity, needed for adaptation purposes, is numerically demanding when needed on-the-fly.

The shallow NN proposed to calculate the MI of a generic SM system, valid also for PMod, outperforms recent approximations found in the literature such as [49] and [42], both in terms of estimation accuracy and computational complexity. In order to avoid the numerical evaluation of the involved integrals, these references provide two different approximations of the MI for a specific symbol constellation, with a complexity scaling with the square of the constellation size and the number of antennas. As opposed, the proposed solution has a much lower complexity, which is independent of the size of the constellation.

The main contributions of this chapter are the accurate evaluation of the MI of SM and the capacity of GSM, which have resisted so far those attempts to obtain simple expressions. Moderate size standard neural networks will be seen to be up to the task provided that a careful extraction of channel

parameters and training are performed. The proposed solution enables also to perform a different type of adaptation in SM systems. Due to the lack of accurate MI evaluation methods, works like [111] and [108] only consider the selection of the modulation order in an adaptive system. However, with the accurate MI obtained with the NN, the transmitter could not only adapt the modulation scheme, but also the coding rate of the channel encoder, allowing a fine granularity in the available transmit MCS. Moreover, due to the lower complexity of the proposed method, the receiver can compute the MI more often and then follow faster channel variations, so that the adaptation speed is not necessarily limited by the complexity of the computation of the channel capacity.

The rest of the chapter is structured as follows. Section 4.2 presents the system model, focusing on SM in Section 4.2.1 and on GSM in Section 4.2.2. Afterwards, a review of the equations involved in the calculation of the MI of SM and the capacity of GSM is provided through Sections 4.3.1 and 4.3.2. Section 4.4 explains how to calculate the MI of SM and the capacity of GSM by using a neural network. Then, the simulation results of the SM MI calculation are contained in Section 4.5 whilst the simulation results of the GSM capacity calculation are given in Section 4.6. Finally, Section 4.7 presents the main conclusions and the appendix of Section 4.8 describes briefly the MFNNs employed throughout all the chapter.

The *notation* followed through this chapter is summarized hereafter. Upper (lower) boldface letters denote matrices (vectors). $(\cdot)^H$, $(\cdot)^t$, $\mathbf{I}_N$ and $\mathbf{1}$, denote Hermitian transpose, transpose, $N \times N$ identity matrix and vector of ones, respectively. $\|\cdot\|$ applied to vectors denotes the Euclidean norm. $\mathbb{E}[\cdot]$ is the expected value operator. $\circ$ and $\oslash$ denote the Hadamard (pointwise) matrix product and division. $\Re\{\cdot\}$, $\Im\{\cdot\}$, $(.)^*$ and $|\cdot|$ denote the real part, imaginary part, conjugate and absolute value of a complex number, respectively.

4.2. System Model

This section is divided in two different subsections. The first one describes a 2×2 SM system and the second one a GSM system. Although the first case, SM, is a particular case of GSM, the notation employed in the descriptions of SM is simpler than the other. The reason of doing so is to follow previous work as [49], where the expressions for obtaining the MI of SM are derived.

4.2.1. System Model of 2×2 SM

Traditional digital modulation schemes transmit information modulating only the amplitude, phase and/or frequency of a sinusoidal carrier. However, Index Modulations (IM) benefit from the fact that the transmitter has several building blocks, being these antennas, polarizations or subcarriers, for example, to map additional bits of information to the block selected to transmit the conventional modulated signal [20]. As illustration, consider SM with N_t transmit antennas: in addition to the $\log_2(M)$ bits to index each symbol s in a constellation of M elements, $\log_2(N_t)$ bits can be used to select which particular antenna, from the N_t available antennas, is active at a given instant to transmit the symbol. Therefore, the total spectral efficiency is

$$\eta = \log_2 N_t + \log_2 M. \tag{4.1}$$

Similarly, PMod, by means of the transmit polarization, carries information. In the beginning of this chapter we will focus our attention on the case with $N_t = 2$, so that one input bit is used to select the active transmit dimension, although a latter section will explain how to extend the results for $N_t > 2$. Hereafter, SM is used instead of IM, however results apply to a generic IM, no matter how are interpreted the dimensions (antennas, polarizations, frequencies...).

The system model of a 2×2 SM for a given discrete time instant is

$$\mathbf{y} = \sqrt{\gamma}\mathbf{H}\mathbf{x} + \mathbf{w} \tag{4.2}$$

where $\mathbf{y} \in \mathbb{C}^{2\times 1}$ is the received vector, γ the average Signal to Noise Ratio (SNR), $\mathbf{H} \in \mathbb{C}^{2\times 2}$ the channel matrix, $\mathbf{x} \in \mathbb{C}^{2\times 1}$ the transmitted signal and $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_2)$ the Additive White Gaussian (AWGN) noise vector. Since $\mathbf{x}$ has only one component different from zero (component l) and its value is $s \in \mathbb{C}$, (4.2) can be also expressed as

$$\mathbf{y} = \sqrt{\gamma}\mathbf{h}_l s + \mathbf{w} \tag{4.3}$$

where $\mathbf{h}_l$ denotes the l column of $\mathbf{H}$, $l \in \{1, 2\}$. We assume a unit power constraint, i.e., $\mathbb{E}\left[\mathbf{x}^H\mathbf{x}\right] = \mathbb{E}\left[|s^2|\right] = 1$.

Fig. 4.1 shows a block diagram of an adaptive 2×2 SM system. The transmitter modifies the coding rate r of the channel encoder and the constellation order M of the modulator according to the different mutual information (MI) values calculated and reported by the receiver. In this way, the transmitter adjusts the transmission rate dynamically to the maximum MI that the channel conditions allow at each time instant. Thus, a fine selection of the coding rate is possible due to the accurate calculation of the MI. This

approach differs from previous works, [111] and [108], where the modulation was the only degree of freedom.

The mapping from the achievable rate (MI) reported by the receiver to the MCSs to be used by the transmitter depends on the strength of the channel codes and the receiver implementation. In a practical system a backoff margin should be enforced based on the distance to capacity of the different MCS. An alternative adaptation method is presented in Chapter 6. Instead of using the neural network to calculate the MI, the MCS is selected directly. This requires the training of the network with data from the MCSs performance obtained from extensive Monte Carlo simulations for a vast number of different channel matrices and SNRs.

We should remark again that the model (4.2) and the block diagram of Fig. 4.1 apply to a generic 2×2 IM, with the matrix $\mathbf{H}$ characterizing the channel effects of the specific domain considered, being this polarization, frequency or space. The adaptive SM transmitter requires the knowledge of the MI at each channel realization to perform the link adaptation. In this chapter we do not deal with the MCS selection at the transmitter, since we focus our attention only on the NN aided MI estimation, denoted by the gray block at the receive side. In order to minimize the overhead of the return link, the receiver estimates the SNR γ and the channel matrix $\mathbf{H}$, computes the MI and sends it back to the transmitter. Section 4.3 provides the expressions to compute the MI of a SM system for an arbitrary constellation. A more practical scheme for calculating these MIs, making use of a NN, will be presented later in Section 4.4.

4.2.2. System model of GSM

Generalized Spatial Modulation (GSM) is a family of multi-antenna modulation schemes where information is transmitted not only by modulating the amplitude, phase and/or frequency of a sinusoidal carrier, but also by selecting the group of antennas employed to transmit the modulated symbol(s) [20]. In general, we consider a $N_t \times N_r$ MIMO system, with N_t transmit antennas, N_r receive antennas and R RF (Radio Frequency) chains for conveying one modulated symbol taken from a modulation alphabet $\mathcal{S}$ with cardinality M.

As particular cases, **Spatial Modulation (SM)** activates only one antenna at a time, and requires just one RF chain [20]. As stated previously, 2×2 SM is equivalent to PMod in mathematical terms. On the other hand, **Single Symbol GSM (SS-GSM)** sends the same symbol through all the $(R > 1)$ active antennas during a channel use [113]. The total number of

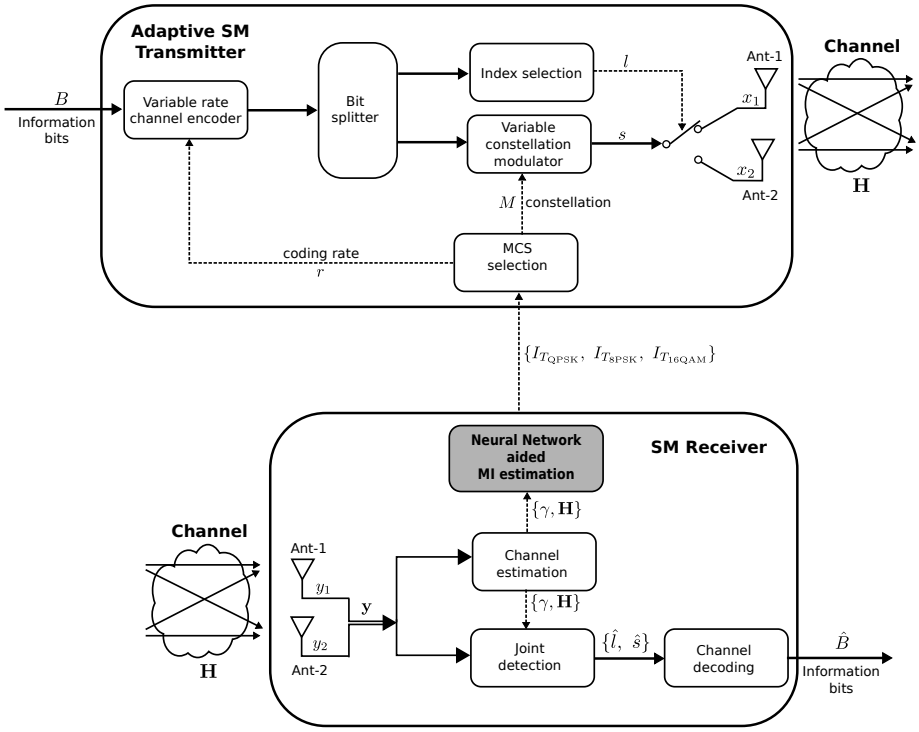


Figure 4.1: Block diagram of an adaptive Spatial Modulation system with Neural Network aided MI calculation at the receiver.

bits conveyed in these two systems is expressed as

$$\eta = \lfloor \log_2 \binom{N_t}{R} \rfloor + \log_2 M. \tag{4.4}$$

When there is only one RF chain, i.e., for $R = 1$, the previous equation reduces to (4.1) and provides the maximum spectral efficiency of SM.

In the most general case, known as **Multi Symbol GSM (MS-GSM)**, the achievable spectral efficiency increases by sending R different symbols through the R active antennas [59]. Note that if $R = N_t$ in MS-GSM, then we have a conventional MIMO system. In this paper we will restrict ourselves to the SM and SS-GSM schemes, without excluding some hints on the evaluation of the capacity for the MS-GSM case.

The base-band samples for a given discrete-time instant of an SM or SS-GSM link can be modeled as

$$\mathbf{y} = \sqrt{\gamma/R}\mathbf{H}\mathbf{x} + \mathbf{w}. \tag{4.5}$$

$\mathbf{y} \in \mathbb{C}^{N_r}$ is the received vector, γ is the average SNR, $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ is the

channel matrix, $\mathbf{x} \in \mathbb{C}^{N_t}$ is the transmitted signal and $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_r})$ is the Additive White Gaussian Noise (AWGN).

Following the notation of [65], the transmit signal $\mathbf{x}$ is constructed with the vector of transmitted symbols $\mathbf{s} \in \mathbb{C}^R$ and an antenna activation pattern matrix $\mathbf{A} \in \mathcal{A}$:

$$\mathbf{x} = \mathbf{A}\mathbf{s} = \mathbf{A}\mathbf{1}s. \tag{4.6}$$

Here $\mathbf{1}$ denotes an $R \times 1$ all ones vector and s the modulation symbol. The set $\mathcal{A}$ contains $N_t \times R$ sparse matrices, with at most one non-zero entry per row. Each column of $\mathbf{A}$ contains a single 1 entry, in the row corresponding to the number of the active antenna. For example, to activate the antennas 1, 2 and 5 in a 6×6 system with 3 RF chains, the corresponding antenna activation pattern matrix would read as

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}^T. \tag{4.7}$$

In general, the number of possible antenna activation choices is not a power of two and the transmitter and receiver have to agree on the use of a set of

$$L = |\mathcal{A}| = 2^{\lfloor \log_2 \binom{N_t}{R} \rfloor} \tag{4.8}$$

antenna activation pattern matrices. The set of all possible transmit signal vectors is then

$$\{\mathbf{x} \colon \mathbf{x} = \mathbf{A}_i \mathbf{1} s_k,\ \mathbf{A}_i \in \mathcal{A},\ s_k \in \mathcal{S}\}, \tag{4.9}$$

where $\mathcal{A}$ is the set of antenna activation pattern matrices and $\mathcal{S}$ is the set of symbols of the constellation. We assume that all $\mathbf{A}_i$ and s_k are equiprobable due to the lack of CSIT (Channel State Information at the Transmitter) and that two independent sequences of information bits are used to select $\mathbf{A}_i$ and s_k.

4.3. Theoretical Expressions of the Capacity and the Mutual Information

In this section, the expression to obtain the MI (the capacity constrained to a given constellation) of a 2×2 SM system will be provided, along with the expression to obtain the capacity of a general GSM system. The equations are taken from [49] and [65], respectively. Moreover, the expressions of two analytical approximations of the MI of a SM system, originally proposed in [49] and [42], are also given, since they will be compared with the method proposed for calculating the MI of SM, which is based on a neural network.

4.3.1. Mutual Information of SM

The expression of the SM capacity conditioned to a given realization of the channel matrix $\mathbf{H}$ is given by

$$C = \max_{f_X(x)} I(\mathbf{x};\ \mathbf{y}\ |\mathbf{H}) = \max_{f_S(s),f_L(l)} I(s,l;\ \mathbf{y}\ |\mathbf{H})\ [bpcu], \tag{4.10}$$

where $I(\mathbf{x};\mathbf{y}|\mathbf{H})$ is the MI between the two random variables $\mathbf{x}$ and $\mathbf{y}$ conditioned to $\mathbf{H}$, and the maximization is performed for all possible distributions of the transmitted signal $\mathbf{x}$ [25]. In (4.10), $f_X(x)$, $f_S(s)$ and $f_L(l)$ denote the probability density functions (PDF) of the random variables of the complex transmit signal $\mathbf{x}$, the complex transmit symbol s, and the hopping index l which selects the antenna/polarization used to transmit the symbols. The units of the capacity in (4.10) are bits per channel use, *bpcu*. The transmitter is expected to operate with only partial CSIT (it only knows the MI, neither $\mathbf{H}$ nor γ), so it will select either index $l = \{1, 2\}$ with the same probability. The capacity is achieved in (4.10) when the transmit symbols belong to a Gaussian codebook [65], i.e., when $s \sim \mathcal{CN}(0, 1)$.

The MI in (4.10) can be expressed as a function of the entropies of the involved random variables:

$$I(\mathbf{x};\ \mathbf{y}\ |\mathbf{H}) = h(\mathbf{y}|\mathbf{H}) - h(\mathbf{y}|\mathbf{x},\mathbf{H}) = h(\mathbf{y}|\mathbf{H}) - h(\mathbf{w}), \tag{4.11}$$

where $h(\cdot)$ is used for the differential entropy, and $h(\mathbf{w})$ is simply written as

$$h(\mathbf{w}) = \log_2 \det(\pi e \mathbf{I}_2). \tag{4.12}$$

As in [65], the received vector $\mathbf{y}$ follows a Gaussian distribution of the form

$$\mathbf{y} \sim \frac{1}{2}\sum_{l=1}^{2} \mathcal{CN}\left(\mathbf{0},\ \mathbf{H}\mathbf{K}_l\mathbf{H}^H + \mathbf{I}_2\right) \triangleq \frac{1}{2}\sum_{l=1}^{2} \mathcal{CN}\left(\mathbf{0},\ \mathbf{\Phi}_l\right), \tag{4.13}$$

where $\mathbf{K}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\mathbf{K}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. With this, the entropy of $\mathbf{y}$ in (4.11) reads as

$$h(\mathbf{y}|\mathbf{H}) = -\frac{1}{2}\sum_{l=1}^{2} \int_{\mathbf{y}} \mathcal{CN}(\mathbf{0},\mathbf{\Phi}_l) \log_2 \left(\frac{1}{2}\sum_{l'=1}^{2} \mathcal{CN}(\mathbf{0},\mathbf{\Phi}_{l'})\right) d\mathbf{y}. \tag{4.14}$$

It is then clear that the computation of the MI $I(\mathbf{x};\ \mathbf{y}\ |\mathbf{H})$ requires the numerical evaluation of the above integral, which can be too demanding for a receiver updating the estimate of the link capacity for adaptation purposes. In this chapter, the MI will be computed by means of Monte Carlo

simulations as a reference for comparing the approximation performed by the neural network.

Practical communication links use symbols from a constellation $\mathcal{S}$ with a finite alphabet. Hereafter, we will refer to the corresponding mutual information, or constrained capacity, simply as total mutual information, I_T, since this includes the information carried by both hopping index l and symbol s:

$$I_T = I(s, l;\ \mathbf{y}\ |\mathbf{H})|_{s\in\mathcal{S}}\,. \tag{4.15}$$

The particularization of (4.11) for a constellation $\mathcal{S}$ with M symbols has been made in [49], and it is replicated in (4.16) for the sake of completeness.

$$I_T = \log_2(2M) - \frac{1}{2M}\sum_{s\in\mathcal{S}}\sum_{l=1}^{2}\mathbb{E}_{\mathbf{W}}\left\{\log_2\left(\sum_{s'\in\mathcal{S}}\sum_{l'=1}^{2} e^{-\gamma\left\|\mathbf{h}_l s-\mathbf{h}_{l'}s'+\frac{\mathbf{w}}{\sqrt{\gamma}}\right\|^2+\|\mathbf{w}\|^2}\right)\right\} \tag{4.16}$$

Monte Carlo integration will be also used to compute (4.16), by generating random values of the noise $\mathbf{w}$.

In an effort to find more convenient expressions to handle in practice, some results have been presented in the literature as approximations to the mutual information (4.16). On the one side, Guo et al [42] used the Jensen's inequality and corrected the ensuing bias to get

$$I_{T_{\text{Jensen}}} = -\log_2\left(\frac{\sum_{\Delta_x\in\mathcal{D}} e^{-\frac{1}{2}\Delta_x^H\mathbf{H}^H\mathbf{H}\Delta_x}}{(2M)^2}\right). \tag{4.17}$$

Here $\mathcal{D}$ is a set with $(2M)^2$ vectors in $\mathbb{C}^{2\times 1}$ of the form

$$\Delta_x = \sqrt{\gamma}\cdot(\mathbf{h}_l s_k - \mathbf{h}_{l'} s_{k'}) \tag{4.18}$$

for $l, l' = 1, 2$ and $k, k' = 1, 2, \ldots, M$, where $\mathbf{h}_l$ are the columns of the channel matrix, and $s_k \in \mathcal{S} \subset \mathbb{C}$ the symbols of the constellation.

A different approach resorting to the Taylor Series Expansion was followed in [49], yielding expression (4.19) as an approximation of I_T. The interested reader is referred to [49] for the definitions of each element of the equation.

$$I_{T_{\text{Taylor}}} = \log_2\left(\frac{2M}{\mathfrak{G}(\mathcal{D}_{sl})}\right) + \\ + \mathfrak{A}\left(\frac{\log_2\left(\mathfrak{G}_{sl}\left(D_{s,l,s',l'}{}^{D_{s,l,s',l'}}\right)\right)}{\mathfrak{A}_{sl}\left(D_{s,l,s',l'}\right)} + \frac{\gamma}{\log(2)\mathcal{D}_{sl}^2}\sum_{m=1}^{2}\left(\mathcal{D}_{m,sl,\Re}^2 + \mathcal{D}_{m,sl,\Im}^2\right)\right) \tag{4.19}$$

One drawback of both (4.17) and (4.19) is that the computational complexity of the MI calculation increases with the square of the constellation order M and the number of antennas N_t.

This chapter develops a more efficient and accurate scheme to compute the MI of a 2×2 SM system, which avoids the quadratic complexity increment with the constellation cardinality. This is especially relevant for practical use, given the need to estimate the channel capacity on the fly for link adaptation purposes. In an adaptive coding and modulation (ACM) system, the receiver must feed back to the transmitter a metric related to the achievable rate, so that the transmitter can select the appropriate MCS[2]. This estimation process needs to be both simple and accurate, since errors will lead to wrong choices of MCSs, with conservative rates or decoding mistakes taking place as a result. The proposed scheme to estimate the achievable rate is based on a simple NN with only one hidden layer, and which provides different outputs, one per constellation in case several are available. The computational burden is much lower than that of any other previously known alternatives, so the MI can be updated more often and faster variations of the channel conditions can be tracked as a benefit.

4.3.2. Capacity of GSM

In this section we review the expression of the capacity of a GSM link with Gaussian signaling, as an upper bound for the use of a finite alphabet or constellation. The GSM channel capacity, for a fixed channel matrix, can be expressed as

$$C_{\text{GSM}} = I(\mathbf{x};\mathbf{y}|\mathbf{H}) = h(\mathbf{y}|\mathbf{H}) - h(\mathbf{y}|\mathbf{x},\mathbf{H}) = h(\mathbf{y}|\mathbf{H}) - h(\mathbf{w}), \tag{4.20}$$

with $I(\cdot;\cdot)$ denoting the MI among two random variables and $h(\cdot)$ the differential entropy. The entropy of the noise is simply $h(\mathbf{w}) = \log_2 \det(\pi e I_{N_r})$

[2]The receiver itself can also make this choice and report back the corresponding index to the transmitter.

and $\mathbf{y}$ follows a Gaussian mixture distribution [65]

$$p(\mathbf{y}) = \sum_{i=1}^{L} p(\mathbf{y}|\mathbf{A}_i)p(\mathbf{A}_i) = \frac{1}{L}\sum_{i=1}^{L} \mathcal{CN}(\boldsymbol{\mu}_i, \boldsymbol{\Phi}_i) \tag{4.21}$$

with parameters:

$$\boldsymbol{\mu}_i = \mathbb{E}\{\mathbf{y}|\mathbf{A}_i\} = \mathbb{E}\{\sqrt{\frac{\gamma}{R}}\mathbf{H}\mathbf{A}_i\mathbf{s} + \mathbf{w}\} = \mathbf{0} \tag{4.22}$$

$$\boldsymbol{\Phi}_i = \mathbb{E}\{\mathbf{y}\mathbf{y}^H|\mathbf{A}_i\} = \frac{\gamma}{R}\mathbf{H}\mathbf{A}_i\mathbb{E}\{\mathbf{s}\mathbf{s}^H\}\mathbf{A}_i^H\mathbf{H}^H + \mathbf{I}_{N_r}, \tag{4.23}$$

where the summation is over the set $\mathcal{A}$, which contains L different antenna activation pattern matrices $\mathbf{A}_i$.

We will assume normalized unit-power symbols, so that the covariance matrix of the symbols vector $\mathbf{s}$ is given for each case by one of the following values:

- **SM**: $\mathbb{E}\{\mathbf{s}\mathbf{s}^H\} = 1$,
- **SS-GSM**: $\mathbb{E}\{\mathbf{s}\mathbf{s}^H\} = \mathbf{1}_{R\times R}$, i.e., a matrix with all ones,
- **MS-GSM**: $\mathbb{E}\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}_R$, i.e., the identity matrix.

The differential entropy of $\mathbf{y}$ requires the evaluation of

$$h(\mathbf{y}|\mathbf{H}) = -\frac{1}{L}\sum_{i=1}^{L}\int_{\mathbf{y}} \mathcal{CN}(\mathbf{0}, \boldsymbol{\Phi}_i)\log_2\left(\frac{1}{L}\sum_{j=1}^{L}\mathcal{CN}(\mathbf{0}, \boldsymbol{\Phi}_j)\right)d\mathbf{y}, \tag{4.24}$$

to obtain the GSM capacity from (4.20), which gives finally

$$C_{\text{GSM}} = -\frac{1}{L}\sum_{i=1}^{L}\int_{\mathbf{y}} \mathcal{CN}(\mathbf{0}, \boldsymbol{\Phi}_i)\log_2\left(\frac{1}{L}\sum_{j=1}^{L}\mathcal{CN}(\mathbf{0}, \boldsymbol{\Phi}_j)\right)d\mathbf{y} - \log_2\det(\pi e I_{N_r}). \tag{4.25}$$

Numerical integration or Monte Carlo simulations are in general required to compute (4.25), which is not practical when trying to estimate the channel capacity on the fly for adaptation purposes.

4.4. Neural Network-based MI and Capacity Estimation

The evaluation of the MI of (4.16) or the capacity of (4.25) can be interpreted as a non-linear mapping from the channel matrix $\mathbf{H}$ and the

SNR γ to the resultant MI or capacity. Multilayer Feedforward Neural Networks (MFNNs), well-known for their fitting capabilities of non-linear functions [39]-[50], will be used to estimate I_T in (4.16) and C_{GSM} in (4.25). In particular, the MFNN to be employed, a one hidden layer network, is shown in Fig. 4.2 and explained with detail in the appendix at the end of this chapter.

The training of the network will be based on an extensive amount of data, by generating a large number of random channels for different values of SNR. The reference true capacity and MI values will be obtained by evaluating (4.16) and (4.25) using Monte Carlo techniques. The neural network input features, $\mathbf{x} = [\mathsf{x}_1, \mathsf{x}_2, \ldots, \mathsf{x}_F]^t$, need to be extracted by means of a function $f(\cdot)$ from the channel matrix $\mathbf{H}$ and the SNR γ, as $\mathbf{x} = f(\gamma, \mathbf{H})$. The input variable selection is highly relevant for the performance of the learning process of the network. Later we will detail how the *feature extraction* for the two cases is applied based on our *domain knowledge*, that is, our knowledge of the particular problem we are addressing.

Alternatively, a deep neural network could be considered, by using directly the channel matrix entries –scaled by $\sqrt{\gamma}$– as inputs rather than a carefully chosen set of input features obtained from our knowledge of the problem. However, we have obtained much better results with single layer networks, requiring much shorter training periods; in addition, smoother training can be expected, in the sense that the parameters of the network converge more smoothly to those values offering a good performance.

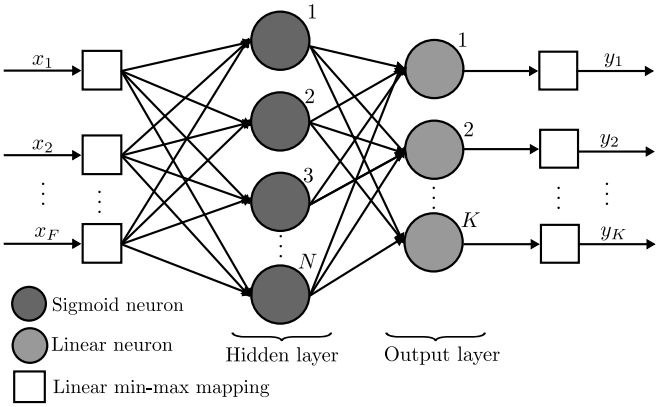


Figure 4.2: Diagram of the neural network.

In the case of SM MI calculation, the neural network has K outputs, each one providing the constrained capacity for a different constellation, as for example QPSK, 8PSK and 16QAM. On the other hand, the network for computing the (unconstrained) capacity of GSM has only $K = 1$ output.

In the following, it will be shown how to obtain the neural network input features for the two cases, MI of SM and capacity of GSM.

4.4.1. Input Variable Selection for MI Calculation of SM

For a Maximum Likelihood receiver, the pairwise error probability (PEP) between (s, l) and $(\hat{s}, \hat{l})$ is given by [108],[114]

$$P_e(s, \hat{s}, l, \hat{l}) = Q\left(\sqrt{\frac{\gamma}{2}}\|\mathbf{h}_l \cdot s - \mathbf{h}_{\hat{l}} \cdot \hat{s}\|\right) \tag{4.26}$$

for the AWGN case. This PEP depends on the distance among supersymbols $\mathbf{h}_l \cdot s$. The set $\mathcal{X}$ of $2M$ different supersymbols for a given channel matrix $\mathbf{H}$ and the transmitted constellation $\mathcal{S} = \{s_k,\ k = 1, 2, \ldots, M\}$ is

$$\mathcal{X} = \{\mathbf{h}_l s_k,\ l = 1, 2,\ k = 1, 2, \ldots, M\}. \tag{4.27}$$

The MI will be also affected by all the involved distances, as inferred from (4.16). For convenience, we put together the squared distances among all the pairs of supersymbols under the matrix $\mathbf{D}$ of size $2M \times 2M$, with the respective entries given by

$$\mathbf{D}[(l-1)M + k, (l'-1)M + k'] = \|\mathbf{h}_l s_k - \mathbf{h}_{l'} s_{k'}\|^2 \tag{4.28}$$

where $l, l' = \{1, 2\}$ and $k, k' = \{1, 2, \ldots, M\}$. Matrix $\mathbf{D}$ can also be expressed as

$$\mathbf{D} = \begin{pmatrix} \|\mathbf{h}_1\|^2\mathbf{D}_S & \mathbf{D}_L \\ \mathbf{D}_L^t & \|\mathbf{h}_2\|^2\mathbf{D}_S \end{pmatrix}. \tag{4.29}$$

The $M \times M$ matrix $\mathbf{D}_S$ on the diagonal is a function of the symbols in the constellation $\mathcal{S}$:

$$\mathbf{D}_S[k, k'] = |s_k - s_{k'}|^2, \tag{4.30}$$

whereas the $M \times M$ matrix $\mathbf{D}_L$ contains all the distances between supersymbols of different antennas/polarizations:

$$\begin{aligned} \mathbf{D}_L[m, n] &= \|\mathbf{h}_1 s_m - \mathbf{h}_2 s_n\|^2 \\ &= \|\mathbf{h}_1 s_m\|^2 + \|\mathbf{h}_2 s_n\|^2 - 2\Re\{s_m^* s_n \mathbf{h}_1^H \mathbf{h}_2\}. \end{aligned} \tag{4.31}$$

Note that the matrix $\mathbf{D}_L$ can be expressed as the sum of four rank-1

matrices:

$$\mathbf{D}_L = \|\mathbf{h}_1\|^2 \begin{pmatrix} |s_1|^2 \\ \vdots \\ |s_M|^2 \end{pmatrix} \mathbf{1}^t \; + \; \|\mathbf{h}_2\|^2 \, \mathbf{1} \begin{pmatrix} |s_1|^2 & \dots & |s_M|^2 \end{pmatrix}$$
$$- \, \mathbf{h}_1^H \mathbf{h}_2 \begin{pmatrix} s_1^* \\ \vdots \\ s_M^* \end{pmatrix} \begin{pmatrix} s_1 & \dots & s_M \end{pmatrix} \; - \; \mathbf{h}_1^t \mathbf{h}_2^* \begin{pmatrix} s_1 \\ \vdots \\ s_M \end{pmatrix} \begin{pmatrix} s_1^* & \dots & s_M^* \end{pmatrix}. \tag{4.32}$$

With this, we have that $\text{rank}\{\mathbf{D}_L\} \leq 4$. Even further, if the constellation of symbols $\{s_n\}$ is known, then only four real values are required to describe the dependence of $\mathbf{D}_L$ and $\mathbf{D}$ with the channel matrix $\mathbf{H}$, namely, $\|\mathbf{h}_1\|^2, \|\mathbf{h}_2\|^2$ and the real and imaginary parts of the scalar product $\mathbf{h}_1^H \mathbf{h}_2$. Alternatively, the scalar product can be expressed as [84]

$$\mathbf{h}_1^H \mathbf{h}_2 = \|\mathbf{h}_1\| \cdot \|\mathbf{h}_2\| \cdot \cos\Theta_H \cdot e^{i\varphi} \tag{4.33}$$

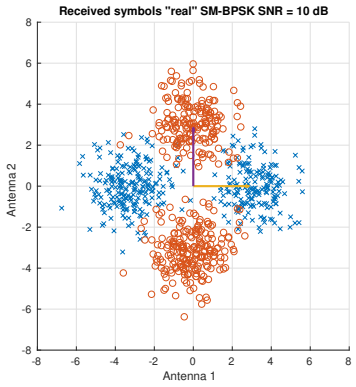
where $\Theta_H \in [0, \pi/2]$ and $\varphi \in [-\pi, \pi]$ denote, respectively, the Hermitian angle and the Kasner's pseudo-angle between two complex vectors. Thus, the four values ($\|\mathbf{h}_1\|$, $\|\mathbf{h}_2\|$, Θ_H and φ) serve to characterize the matrix $\mathbf{D}$.

For illustration purposes, Fig. 4.3 shows the received symbols for a real BPSK case, with real symbols and real noise. Two different SNR values and two different channel matrices are employed to display the clouds of received symbols. Both the SNR and the angle between the column vectors of $\mathbf{H}$ determine the distance among the different color clouds.

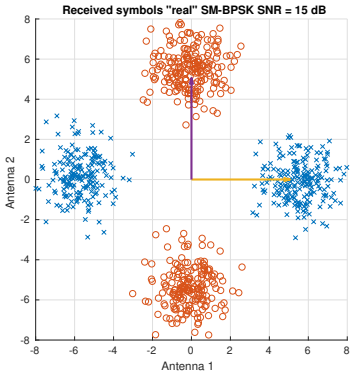
The impact of the two angles Θ_H and φ in the final MI can be grasped with the aid of Fig. 4.4, which shows a 3D representation of the MI as a function of both angles. Monte Carlo simulations were run for a QPSK constellation, with both columns having the same norm, $\|\mathbf{h}_1\| = \|\mathbf{h}_2\| = 1$, and $\gamma = 2$. With this, the structure of the channel matrix $\mathbf{H}$ is the following:

$$\mathbf{H} = \begin{pmatrix} 1 & \cos\Theta_H e^{i\varphi} \\ 0 & \sin\Theta_H \end{pmatrix}. \tag{4.34}$$

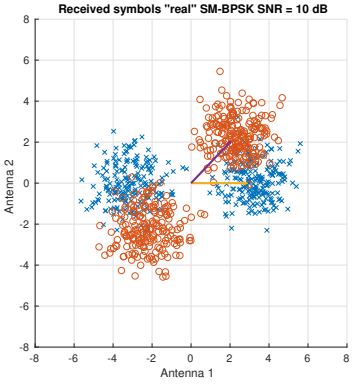
It can be seen that the MI has a strong dependance with the Hermitian angle. If $\Theta_H = \pi/2$ the two columns are orthogonal and the MI is maximum, whereas for $\Theta_H = 0$ the columns are considered parallel, and the MI is reduced. Moreover, the Kasner's pseudoangle φ only affects the MI significantly when Θ_H is close to zero, creating a ripple, due to the radial symmetry of the constellation. However, for $\Theta_H > \pi/3$, the MI is barely



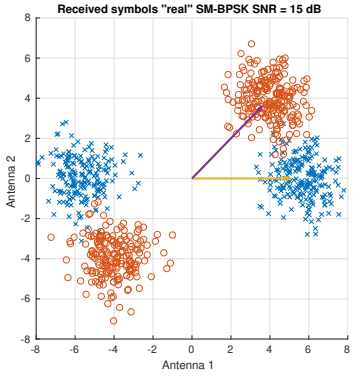
(a) Orthogonal columns, SNR= 10 dB.



(b) Orthogonal columns, SNR= 15 dB.



(c) Non-orthogonal columns, SNR= 10 dB.



(d) Non-orthogonal columns, SNR= 15 dB.

Figure 4.3: Received constellation for 2×2 SM-BPSK system where transmitted symbols, channel matrix and noise are real-valued.

affected by the Kasner's pseudoangle. Instead, if $\Theta_H = 0$, the phase φ determines to which extent the receiver can tell which antenna transmitted the observed symbol.

After extensive training cases, we have observed that performance can be enhanced if, as part of the input parameters, the distances among the supersymbols are also included. Four distances, $\{d_i\}, i = 1, \ldots, 4$, are used; this is the number of different entries of matrix $\mathbf{D}_L$ in (4.31) when a QPSK constellation is employed. It turns out that these four quantities suffice for the neural network to compute a good estimate of the MI for other constellations too, such as 8PSK and 16QAM, even though the number of different entries of the matrix is higher.

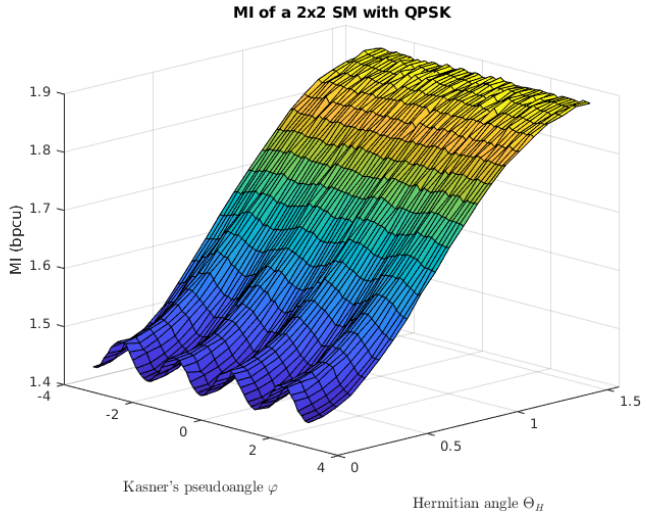


Figure 4.4: MI of a 2×2 SM link with QPSK constellation as a function of the two angles for unit-valued columns norms and 3 dBs of SNR.

In Table 4.1, five different options for the neural network input features are shown. In a later section, data of the performance of a MFNN using those features will be given. Following the above considerations, the input features $\mathbf{x} = f(\gamma, \mathbf{H})$ will correspond to different options to characterize the distance matrix $\mathbf{D}$. Essentially, four inputs is the lowest number of inputs to test following the previous discussion[3]. Some of the values, as the column norms and the distances, are sorted in ascending order given the invariance of the capacity to the labeling of the dimensions and symbols.

Option	Input features F	Description of the features
i	4	$\mathbf{x} = \left[\text{sort}\left([\gamma\|\mathbf{h}_1\|^2,\ \gamma\|\mathbf{h}_2\|^2]\right),\ \Re\left\{\frac{\mathbf{h}_1^H\mathbf{h}_2}{\|\mathbf{h}_1\|\cdot\|\mathbf{h}_2\|}\right\},\ \Im\left\{\frac{\mathbf{h}_1^H\mathbf{h}_2}{\|\mathbf{h}_1\|\cdot\|\mathbf{h}_2\|}\right\}\right]^t$
ii	4	$\mathbf{x} = \left[\text{sort}\left([\gamma\|\mathbf{h}_1\|^2,\ \gamma\|\mathbf{h}_2\|^2]\right),\ \Theta_H,\ \varphi\right]^t$
iii	6	$\mathbf{x} = \left[\text{sort}\left([\gamma\|\mathbf{h}_1\|^2,\ \gamma\|\mathbf{h}_2\|^2]\right),\ \text{sort}([\gamma d_1,\ \gamma d_2,\ \gamma d_3,\ \gamma d_4])\right]^t$
iv	8	$\mathbf{x} = \left[\text{sort}\left([\gamma\|\mathbf{h}_1\|^2,\ \gamma\|\mathbf{h}_2\|^2]\right),\ \text{sort}([\gamma d_1,\ \gamma d_2,\ \gamma d_3,\ \gamma d_4]),\ \Re\left\{\frac{\mathbf{h}_1^H\mathbf{h}_2}{\|\mathbf{h}_1\|\cdot\|\mathbf{h}_2\|}\right\},\ \Im\left\{\frac{\mathbf{h}_1^H\mathbf{h}_2}{\|\mathbf{h}_1\|\cdot\|\mathbf{h}_2\|}\right\}\right]^t$
v	8	$\mathbf{x} = \left[\text{sort}\left([\gamma\|\mathbf{h}_1\|^2,\ \gamma\|\mathbf{h}_2\|^2]\right),\ \text{sort}([\gamma d_1,\ \gamma d_2,\ \gamma d_3,\ \gamma d_4]),\ \Theta_H,\ \varphi\right]^t$

Table 4.1: Different alternatives for selecting the NN input features.

Extension to a higher number of antennas

Simulation results will reveal later that the performance of the neural network improves from the top to the bottom of Table 4.1. However, in order to avoid a too large number of input features in systems with a higher

[3]Note that the number of real values to characterize $\mathbf{H}$ and γ is nine.

number of antennas, we propose to apply option (ii) as a trade-off between performance and complexity. This option makes use of the channel column norms (scaled by the SNR) and the angles.

Whilst the number of norms increases only linearly with the transmit antennas, the number of angles raises rapidly with N_t since there are $2\binom{N_t}{2}$ angles, a tuple (Θ_H, φ) for each possible combination of two transmit antennas. For example, in a system with 16 antennas there are 120 pairs of angles. However, we have found out that it is not necessary to provide the values of all the angles explicitly to the neural network. A few values characterizing the statistical distribution of the angles suffice for the NN to estimate the MI with an MSE similar to those values reported in Table 4.2.

The MI evaluation in a SM system with 4×4 antennas can be easily done with an MFNN trained with the proper dataset, obtained now with 4×4 Rayleigh matrices, and using as input features the four values of $\gamma \|\mathbf{h}_l\|^2$ and the six pairs of angles (θ_H, φ). In the case of an 8×8 IM system we propose to reduce the $\binom{8}{2} = 28$ pairs of angles to just Q values per type of angle (Hermitian and Kasner). These Q values are the quantiles of the angles distribution for Q probabilities taken from 0 to 1 at equal steps. For example, for $Q = 5$ the distribution of the angles is characterized by the minimum, the 25th percentile, the 50th percentile (the median), the 75th percentile and the maximum. Therefore, the MFNN for obtaining the set of MI of an 8×8 IM system has as input features the 8 columns norms $\gamma \|\mathbf{h}_l\|^2$, and the Q quantiles of the Hermitian angle Θ_H and the Kasner's pseudoangle φ, respectively.

4.4.2. Input Variable Selection for Capacity Calculation of GSM

Fig. 4.5 shows the diagram of the neural network for obtaining the capacity of GSM. The single difference with Fig. 4.5 is that in this case there is only one output, compared with the K outputs of the other network, which calculates the MI for K different constellations. As in the case of SM, the true capacity values are obtained by solving the integral of (4.24) with Monte Carlo simulations. The network training is performed off-line, so that the receivers simply have to use the trained net. This entails a much lower complexity than obtaining (4.24), which requires the computation of L integrals of $2N_r$ real variables. The accurate calculation of high-dimensional integrals is a complex problem, since the number of function evaluations required for a given accuracy increases exponentially with the number of dimensions. Then, the use of a slow method such as the Monte Carlo scheme is even faster than traditional numerical integration [85].

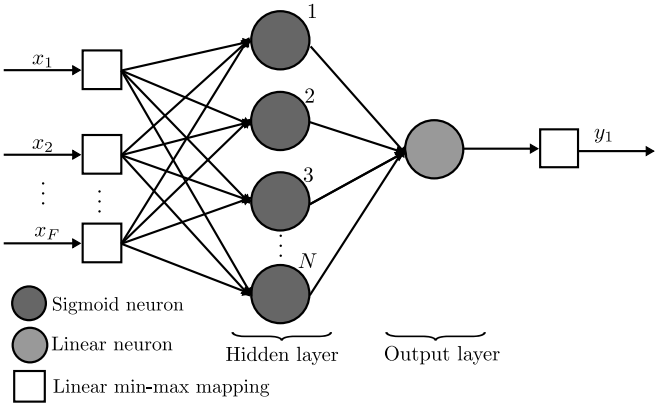


Figure 4.5: Diagram of the neural network for obtaining the capacity of GSM.

Following a similar philosophy, the proposal for selecting the neural network inputs is based on the expressions of the PEP. With Maximum Likelihood decoding, and using s and s' to denote two different modulation symbols, l and l' for two spatial symbols, i.e., two different matrices from $\mathcal{A}$, the expression of the PEP between (s, l) and (s', l') is

$$P_e(l, s, l', s') = Q\left(\sqrt{\frac{\gamma}{2R}}\|\mathbf{H}\mathbf{A}_l\mathbf{1}s - \mathbf{H}\mathbf{A}_{l'}\mathbf{1}s'\|\right). \tag{4.35}$$

This expression, given by [108] for SM, shows the dependence of the PEP on the distance among the received symbols in the absence of noise and with the SNR, which weights them with the noise power.

The square of the distance between two arbitrary noise-free received spatial-modulated symbols is written as

$$d = \sqrt{\gamma/R} \cdot \|\mathbf{H}\mathbf{A}_l\mathbf{1}s - \mathbf{H}\mathbf{A}_{l'}\mathbf{1}s'\|^2 = \|\mathbf{c}_l s - \mathbf{c}_{l'} s'\|^2, \tag{4.36}$$

where we have defined the column vector $\mathbf{c}_l = \sqrt{\gamma/R}\mathbf{H}\mathbf{A}_l\mathbf{1}$. The definition of the antenna activation pattern matrices $\mathbf{A}_l$ is such that the column vectors $\mathbf{c}_l$ consist of the sum of R columns of $\mathbf{H}$ according to the positions of the ones in $\mathbf{A}_l$. As it was done in the previous subsection, we obtain after simple manipulations

$$d = \|\mathbf{c}_l\|^2|s|^2 + \|\mathbf{c}_{l'}\|^2|s'|^2 - 2\Re\{\mathbf{c}_l^H\mathbf{c}_{l'}s^*s'\}. \tag{4.37}$$

This shows that the distances, and hence the capacity in a SS-GSM system, apart from the properties of the constellation, they depend on the norms and scalar product of the sums of columns given by the antenna activation pattern matrices of the set $\mathcal{A}$. Focusing on the influence of the channel

matrix, the scalar product between the two complex column vectors $\mathbf{c}_l$ and $\mathbf{c}_{l'}$ can be expressed as

$$\mathbf{c}_1^H \mathbf{c}_2 = \|\mathbf{c}_1\| \cdot \|\mathbf{c}_2\| \cdot \cos\Theta_H \cdot e^{i\varphi}, \tag{4.38}$$

where the two angles between the complex vectors Θ_H and φ, the so-called Hermitian angle and Kasner's pseudo-angle, respectively, which were introduced before following [84].

Based on all of the above, we know how to proceed to pre-process $\mathbf{H}$ and γ to extract the relevant features to feed the neural network. Thus, Algorithm 1 details all the steps needed to feed the network. Firstly, the vectors $\mathbf{c}_l$ are calculated with the channel matrix $\mathbf{H}$ and the set of antenna activation pattern matrices $\mathcal{A}$. Then, these vectors are employed to obtain the three types of inputs of the neural network: the norms of the vectors $\mathbf{c}_l$ ($\mathcal{N}$), the Hermitian angles ($\mathcal{H}$), and Kasner's pseudo-angles ($\mathcal{K}$) between all the pairs of these vectors $\mathbf{c}_l$.

The number of norms, and specially the angles, increase rapidly with the number of transmit antennas and RF chains. For example, with $N_t = 8$ and $R = 2$ there are $L = 16$ vectors $\mathbf{c}_l$ according to (4.8) and $\binom{L}{R} = 120$ angles. As a consequence, to keep the number of neural network inputs within reasonable bounds, we use a few values characterizing the distribution of both the norms and the angles, rather than feeding the entire set of values. When L or $\binom{L}{R}$ becomes higher than 8 (a value selected empirically), the inputs of the MFNN become the quantiles of norms and angles for some fixed probabilities. We have observed experimentally that the characterization of the Cumulative Distribution Function (CDF) using just $Q = 9$ probabilities (which include, among others, the minimum, the median and the maximum) for obtaining the quantiles, allows to make a good estimation of the GSM capacity while maintaining a reduced number of neural network inputs.

4.5. MI of SM: Simulation results

For performance evaluation, a dataset of 50 000 realizations of the 2×2 channel matrix $\mathbf{H}$ is used, randomly generated following a unit-variance Rayleigh distribution, i.e., $h_{ij} \sim \mathcal{CN}(0, 1)$. Each channel matrix is associated with a different SNR whose value in decibels is drawn from a uniform random variable between -20 and 20 dB. The true MI with QPSK, 8PSK and 16QAM constellations of each realization of $(\gamma, \mathbf{H})$ is calculated with a Monte Carlo simulation using (4.16) with 5 000 realizations of the complex Gaussian noise $\mathbf{w}$. We limit ourselves to these low order constellations, which are more likely to be used with SM. However, other constellations, like

Algorithm 1 Pre-processing of γ and $\mathbf{H}$ to obtain the neural network inputs for GSM capacity calculation

Require: $\gamma,\ \mathbf{H} \in \mathbb{C}^{N_r \times N_t},\ \mathcal{A} = \{\mathbf{A}_l, l = 1, 2, \ldots, L\},\ Q.$
Ensure: $\mathcal{N}$, $\mathcal{H}$, $\mathcal{K}$.
for $l = 1$ to L **do**
  Calculate the column vector $\mathbf{c}_l = \sqrt{\gamma/R}\mathbf{H}\mathbf{A}_l\mathbf{1}$
  Include its norm $\|\mathbf{c}_l\|^2$ in the set $\mathcal{N}$.
end for
for $k = 1$ to $\binom{L}{2}$ **do**
  Calculate Hermitian angle θ_H and Kasner's pseudoangle φ between a pair of column vectors $\mathbf{c}_l$ and $\mathbf{c}_{l'}$.
  Include θ_H in the set $\mathcal{H}$ and φ in the set $\mathcal{K}$.
end for
quants $\leftarrow$ linspace(0, 1, Q) {Equally spaced Q values among 0 and 1}
if $L \leq 8$ **then**
  $\mathcal{N} \leftarrow$ sort($\mathcal{N}$) {Sort in ascending order}
else
  $\mathcal{N} \leftarrow$ quantile($\mathcal{N}$) {Calculate the quantiles of the values of $\mathcal{N}$ for the cumulative probabilities quants}
end if
if $\binom{L}{2} > 8$ **then**
  $\mathcal{H} \leftarrow$ quantile($\mathcal{H}$, quants)
  $\mathcal{K} \leftarrow$ quantile($\mathcal{K}$, quants)
end if

64QAM, could be easily added to the system at the expense of increasing the time required for obtaining the dataset with Monte Carlo simulations -note the two summations over all the constellation symbols in (4.16).

Two different widths, 10 and 20 neurons, are tested for the neural network, which has only one hidden layer. The different parameters of the network will be extracted from supervised learning with the Levenberg-Marquardt (LM) backpropagation algorithm [45] using the Mean Squared Error (MSE) as performance metric. The dataset is divided into two independent parts. 7 500 samples (15%) are reserved for the final test of the performance of the MFNN and the analytical approximations (equations (4.17) and (4.19)). The remaining 35 000 samples (70%) and 7 500 samples (15%) are employed for training and validation of the neural network, respectively. The training runs for 1000 epochs (an epoch consists on a set of iterations of the learning algorithm through which the whole training dataset is used), although it could be halted earlier if the network performance on the validation dataset stopped improving or remained the same for 6 epochs in a row. The default parameters of the *trainlm* function of Matlab® were used for the training.

Firstly, the impact of the selection of the input features in the performance of the MFNN is evaluated. For this, several neural networks are trained using in each case one of the sets of inputs detailed in Table 4.1. Fig. 4.6 shows some histograms with the statistical distribution of the features obtained from the dataset. The distances and the norms include the SNR term and are shown in dB, whilst the unit of the angles is the radian.

Then, the global MSE obtained with the trained NNs when calculating the three MI values is obtained by using the entries of the dataset reserved for testing. Table 4.2 collects the values of MSE for five different selections of the input features and for both numbers of neurons ($N = 10$ and $N = 20$). It shows the best MSE in the testing dataset after 10 trainings with different NN parameters initialization. If the NNs are fed directly with the real and imaginary parts of the channel matrix coefficients, the MSE is very high, in the order of 10^{-1}. Nevertheless, at least two orders of magnitude improvement is achieved when the NNs are fed with the input features detailed in Table 4.1.

Input features option	**Number of features F**	**Global MSE (10 neurons)**	**Global MSE (20 neurons)**
i) Columns norm and projection	4	$1.78 \cdot 10^{-3}$	$6.98 \cdot 10^{-4}$
ii) Columns norm and angles	4	$4.29 \cdot 10^{-4}$	$3.36 \cdot 10^{-4}$
iii) Columns norm and distances	6	$1.79 \cdot 10^{-4}$	$5.21 \cdot 10^{-5}$
iv) Columns norm, distances and projection	8	$1.33 \cdot 10^{-4}$	$4.96 \cdot 10^{-5}$
v) Columns norm, distances and angles	8	$1.00 \cdot 10^{-4}$	$2.97 \cdot 10^{-5}$

Table 4.2: Comparison of the global MSE obtained with the neural network for different input features.

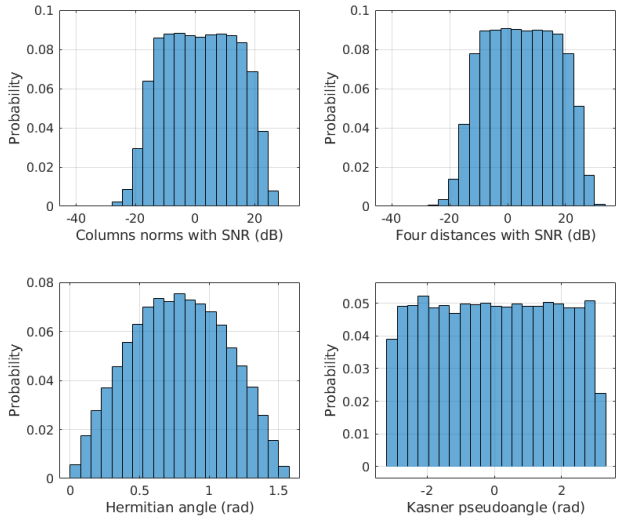


Figure 4.6: Histograms with the distribution of the features (norms, distances and angles) in the dataset of Rayleigh distributed channel matrices and uniformly distributed SNR.

In Table 4.2, it can be observed how the Hermitian angle Θ_H and the Kasner's pseudoangle φ, options (ii) and (v), improve the NN estimation with respect to the use of the real and imaginary parts of the projection, options (i) and (iv). Furthermore, the addition of the four distances to the set of inputs, cases (iii), (iv) and (v), serves to reduce the MSE as compared to cases (i) and (ii). Finally, the MSE reaches a minimum value of about $3 \cdot 10^{-5}$ when the four distances and the two column norms are combined with the two angles for the 20 neurons MFNN.

Secondly, the two NNs with 10 and 20 neurons and the input features selection (v), are compared with the analytical approximations from the literature, (4.17) and (4.19), in Table 4.3. Both Taylor and Jensen based approximations have a similar MSE, around 10^{-2}, which is outperformed by all the NN reported in Table 4.2. Moreover, when we compare the analytical approximations with the best NNs of the table, the improvement in the MSE is about 100 and 600 times with a NN of 10 and 20 neurons, respectively.

As noted previously, the calculation of the MI could be addressed with a deep neural network, a MFNN with several hidden layers, using as inputs the channel matrix coefficients (scaled by the SNR) directly. With this approach, the network is expected to extract the relevant features at the intermediate layers, so that the last layer computes the MI. We have tested this approximation for a number of layers ranging from one to ten, and

	Global MSE	QPSK		8PSK		16QAM	
		3σ	Max. error	3σ	Max. error	3σ	Max. error
Taylor approximation (4.19)	$1.87 \cdot 10^{-2}$	0.330	0.523	0.370	0.492	0.392	0.558
Jensen based approximation (4.17)	$1.21 \cdot 10^{-2}$	0.229	0.300	0.291	0.498	0.300	0.741
MFNN option (v) with 10 neurons	$1.00 \cdot 10^{-4}$	0.020	0.153	0.026	0.140	0.034	0.120
MFNN option (v) with 20 neurons	$2.97 \cdot 10^{-5}$	0.016	0.067	0.015	0.046	0.018	0.105

Table 4.3: Comparison of the performance of the MFNN with the analytical approximations of the literature for calculating the MI of a 2×2 SM.

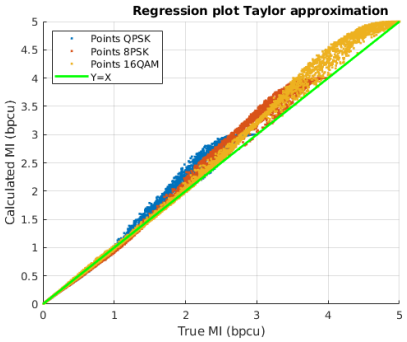
a number of neurons per layer between 20 and 50. It was found that a deep network with at least three layers of 20 neurons can perform better than the Taylor and Jenson approximations, yielding an MSE in the order of 10^{-3}. However, the deep networks do not overcome the peformance of the one-hidden layer MFNN with the input features of row (v) of Table 4.1. In addition, the training of these deep networks is much more time consuming, and the learning algorithm has more difficulties to converge to those parameter values providing a good performance.

As shown in Table 4.3, the analytical approximation suffers a maximum error of 0.741 in the 16-QAM constellation, which is reduced to 0.105 in the case of the MFNN with 20 neurons. The improvement is even more noticeable with the 3σ value: a little bit more than 0.300 for the analytical approximations, and 20 times smaller with the neural network.

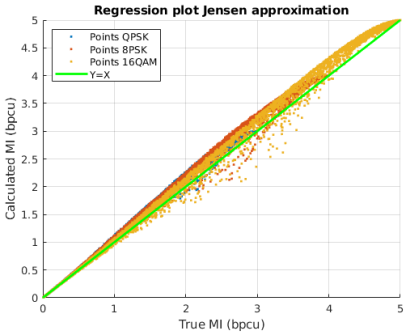
Fig. 4.7 shows a graphical view of the estimated MI values: the scatter plot of the values of the true MI (X axis) are shown together with the values of the MI computed with each method (Y axis) for the three constellations, QPSK, 8PSK and 16QAM. The green line Y=X sets the perfect match of the MI. It can be seen that the analytical approximations provide better results for lower values of MI, while for MIs close to their maximum (3, 4 or 5, depending on the constellation), they have a noticeable positive bias. Remarkably, both MFNNs with 10 and 20 neurons in the hidden layer, match the true value of the MI almost perfectly, clearly outperforming the analytical approximations.

The accuracy achieved by the MFNN has a direct impact on the implementation of adaptive SM links. The quality of the tracking of the MI allows to use smaller back-off margins for the selection of the MCS; large errors make it necessary to use highly conservative margins in the selection of the MCS to guarantee a prescribed error decoding metric, thus reducing the transmission rate.

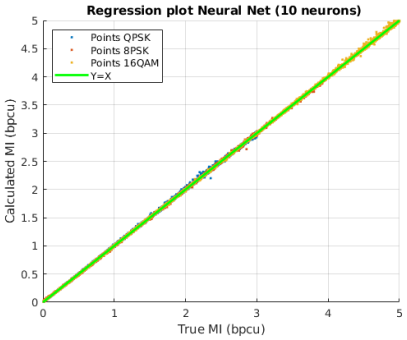
Finally, the ergodic MI of a 2×2 Index Modulation system with QPSK, 8PSK and 16QAM constellations under Rayleigh fading is shown in Fig. 4.8. For each value of SNR, 100 realizations of the channel matrix are generated, similarly to the NN dataset, and the true MI in each case is



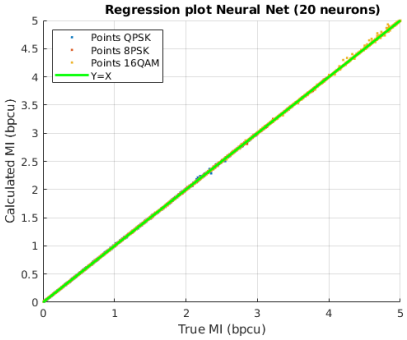
(a) Taylor approximation (4.19).



(b) Jensen based approximation (4.17.)



(c) Neural network option (v), 10 neurons.



(d) Neural network v) 20 neurons

Figure 4.7: Comparison of the scatter plots (true MI vs calculated MI) of the analytical approximation from the literature and the MFNNs with 10 and 20 neurons for the three constellations (QPSK, 8PSK and 16QAM).

calculated with a Monte Carlo simulation with $1,000$ realizations of the noise. Afterwards, the ergodic MI for each SNR point is calculated by averaging the instantaneous values of the MI. The true ergodic MI, shown with circles, is compared with that obtained by averaging the instantaneous MI calculated with each method, the two analytical approximations and the 10 neurons neural network. As it can be seen, the neural network matches perfectly the true ergodic MI, which is overestimated by the other methods for moderate values of the SNR.

In addition to the estimation accuracy, the **computational complexity** is key for practical implementation: the MI computation must be done at the receive side, which has knowledge of the SNR γ and the channel matrix $\mathbf{H}$. This evaluation must be such that an on-line tracking of the channel is made to report the estimated MI back to the receiver.

In this regard, Table 4.4 shows the computational complexity of each

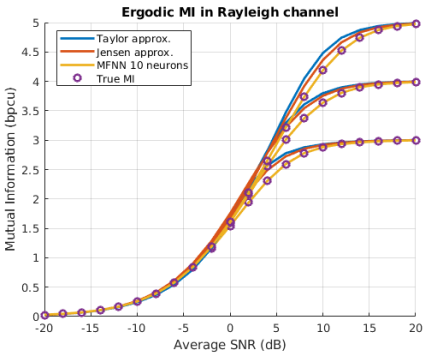


Figure 4.8: Ergodic MI obtained after averaging the instantaneous MI calculated with each method for a Rayleigh channel, from bottom to top QPSK, 8PSK and 16QAM.

	Taylor approximation	**Jensen based approximation**	**MFNN option (v) 20 neurons**
Real products	$7,168$	$32,800$	368
$\exp(\cdot)$	672	$1,344$	20
$\log_2(\cdot)$	112	3	-
Other non-linear operations	$1,344$	-	3
Time for $7,500$ calculations	**41 s**	**76 s**	**0.80 s**

Table 4.4: Comparison of complexity and computational time of the MFNN and the two methods of the literature. Total number of operations for computing the three MI values (QPSK, 8PSK and 16QAM) are given, as well as the computational time required for calculating $7,500$ values of these MIs with Matlab® running in a laptop.

method after counting the number of mathematical operations required to compute the MIs for the three constellations. In the case of the analytical approximations, the table shows a lower bound of the number of operations since it only counts the most demanding instructions, which are repeated $(2M)^2$ times. In the case of the NN, all the required operations are counted, including the preprocessing of γ and $\mathbf{H}$ to calculate the inputs of the NN.

The numbers in Table 4.4 reveal that the MFNN is not only more accurate, but also less computationally demanding. The MFNN requires about 90 times fewer multiplications and non linear operations than the analytical approximations. This is in line with the required time to compute with Matlab® the three MIs for all the testing dataset, in a computer equipped with an i7-4510U 2 GHz processor. From another point of view, the laptop only allows to make an estimation of the MIs every 5.5 ms with the Taylor approximations, which gets reduced to 0.1 ms with the NN. With respect to the off-line training duration, each training of the NN took typically less than 3 minutes.

So far, the focus was put on the 2×2 SM system, although the simulation

	Input features	Global MSE	QPSK		8PSK		16QAM	
			3σ	Max. error	3σ	Max. error	3σ	Max. error
SM 2×2 (20 n.), option (ii)	$2+2\times1=4$	$3.36\cdot10^{-4}$	0.078	0.436	0.041	0.342	0.034	0.322
SM 4×4 (20 n.)	$4+2\times6=16$	$2.40\cdot10^{-4}$	0.047	0.169	0.050	0.200	0.041	0.137
SM 8×8 ($\mathbf{Q=5}$, 20 n.)	$8+2\times5=18$	$5.06\cdot10^{-5}$	0.022	0.050	0.023	0.061	0.018	0.046

Table 4.5: Performance of the MFNNs for obtaining the MI for QPSK, 8PSK and 16QAM of a SM system with 2, 4 and 8 transmit and receive antennas.

results can be easily extended to system with a higher number of antennas. For testing purposes, we have generated two additional datasets, one with $50,000$ 4×4 Rayleigh matrices and another with $25,000$ 8×8 Rayleigh matrices. Again, each channel matrix has associated a random SNR value between -20 and 20 dB, and we have calculated the MI of each pair $(\gamma, \mathbf{H})$ for several constellations (QPSK, 8PSK and 16QAM) using Monte Carlo simulations. Following the same procedure of training and testing, we have obtained two trained MFNNs for calculating the three MI values of 4×4 and 8×8 SM systems, respectively.

Table 4.5 sums up the results obtained with these two neural networks. In the 8×8 system we have used only $Q=5$ quantiles for characterizing the distribution of each angle whilst in the 4×4 system all the 6 angles were used. The MSE obtained are in the order of 10^{-4} to 10^{-5}, similarly to the results obtained for the MI calculation with the MFNN for the 2×2 SM case. If we compare the results of 4 and 8 antennas provided in Table 4.5 with those obtained with the best network for the 2 antennas scenario from Table 4.3, using 20 neurons and input features option (v), the MFNN performs a little worse in the setup with more antennas, although the errors are of the same magnitude. However, a fair comparison with the NN for 2×2 SM which uses the same type of input features, i.e., option (ii) from Table 4.2, reveals that the MSE is slightly better when the number of dimensions grows. In this way, for the 8 antennas setup, the MSE is as low as $5.06\cdot10^{-5}$, the 3σ value is around 0.02 and the maximum error is only about 0.06.

Generalization to Other Channel Distributions

So far, the probability density function of the channel matrices of the training and testing datasets was the same. On the contrary, this section evaluates the robustness of the neural network for calculating the MI of channels with a distribution different from that employed during the training. Namely, MFNN with the same internal parameters, obtained previously with the Rayleigh dataset, are used to calculate the MI of channels with two different distributions. One Dual Polarization (DP) mobile satellite channel and a terrestrial channel with correlation between the antennas at both transmission and reception.

	Global MSE	QPSK		8PSK		16QAM	
		3σ	Max. error	3σ	Max. error	3σ	Max. error
MFNN option (ii) 20 neurons	$3.15 \cdot 10^{-3}$	0.217	0.487	0.072	0.230	0.064	0.196
MFNN option (v) 20 neurons	$7.40 \cdot 10^{-5}$	0.022	0.029	0.024	0.033	0.028	0.045

Table 4.6: Performance of the pre-trained MFNNs for obtaining the MI for QPSK, 8PSK and 16QAM of a Polarized Modulation Dual Polarization Mobile Satellite system.

In Chapter 5, a DP mobile satellite system with Polarized Modulation (PMod) is analyzed. The simulations of that chapter require the continuous calculation of the MI of PMod with a QPSK constellation in that DP channel. Therefore, in this subsection we check if the MFNN trained with the Rayleigh dataset generalizes well to calculate the MI of PMod in that DP satellite scenario.

We have generated a series of $1,000$ channel matrices of a maritime DP satellite channel as explained in Section 5.3, with a random SNR value between -5 and 20 dB assigned to each matrix. Then, the MI of PMod for three constellations, QPSK, 8PSK and 16QAM, is computed using Monte Carlo simulations, similarly to how the ML dataset was generated. Lastly, the MFNNs with the same internal parameters which were obtained previously in the training with the Rayleigh channel matrices were employed to calculate the MI of the satellite channel matrices. Table 4.6 shows the performance of two MFNNs in this new scenario for two input features selections.

The MFNN which uses only the two column norms and the two angles as input parameters (option (ii) of Table 4.1) performs a little bit worse in this new scenario, increasing the global MSE one order of magnitude, as compared with the results of the first row of Table 4.5. However, the best network, which uses also the four distances (option (v) of Table 4.1), has a similar MSE, the variance of error is similar too and, in addition, it reduces the maximum error for each one of the constellations. Figs. 4.9 and 4.10 show, for the MI calculation using the MFNN with the input features option (v), the histogram of the error and a regression plot, respectively, which illustrate the good behavior of the neural network. Therefore, we can conclude that the MFNN with the input features option (v) can be used safely for calculating the MI in the DP satellite channel.

Lastly, the neural networks are going to be tested with channel matrices where correlation between the transmit and receive antennas is added. Some new 2×2 channel matrices $\mathbf{H}$ are generated following the Kronecker model [46]

$$\mathbf{H} = \mathbf{R}_r^{1/2} \mathbf{H}_w \mathbf{R}_t^{H/2}, \tag{4.39}$$

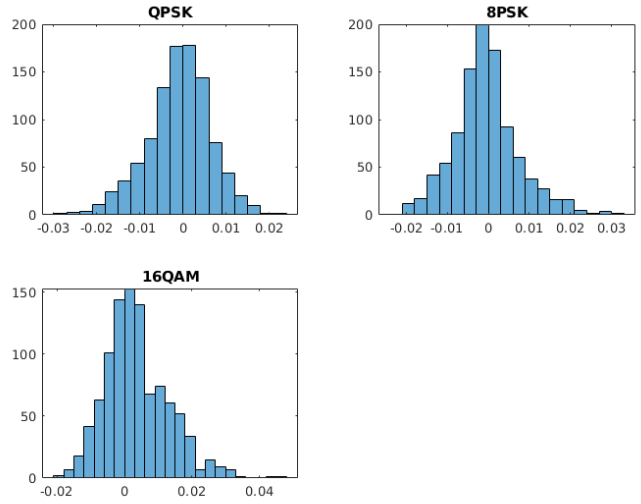


Figure 4.9: Histograms of the error in the MI calculation of PMod in a DP mobile satellite channel with the MFNN trained with Rayleigh matrices.

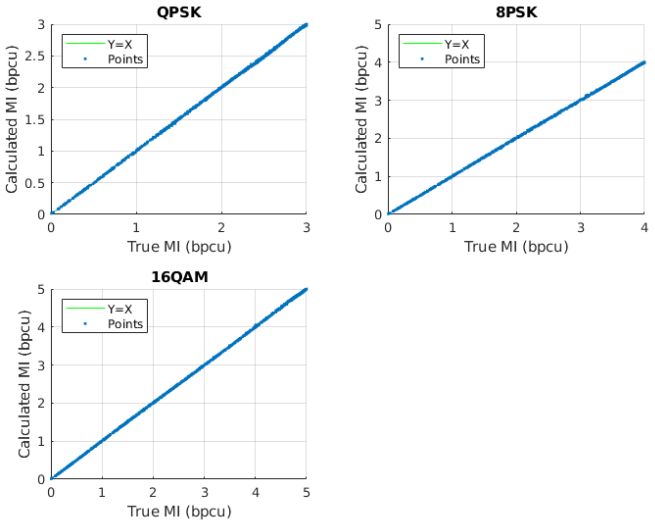


Figure 4.10: Regression plots of the true vs calculated MI of PMod in a DP mobile satellite channel with the MFNN trained with Rayleigh matrices.

where $\mathbf{H}_w$ is a random Rayleigh distributed matrix and $\mathbf{R}_t$ and $\mathbf{R}_r$, the covariance matrices with the correlations between all pairs of transmit and

	MFNN option (ii) 20 neurons	MFNN option (v) 20 neurons
Correlation $\alpha = 0$	$3.36 \cdot 10^{-4}$	$2.97 \cdot 10^{-5}$
Correlation $\alpha = 0.1$	$1.82 \cdot 10^{-3}$	$6.11 \cdot 10^{-5}$
Correlation $\alpha = 0.3$	$8.41 \cdot 10^{-3}$	$2.32 \cdot 10^{-4}$
Correlation $\alpha = 0.6$	$4.05 \cdot 10^{-2}$	$1.11 \cdot 10^{-3}$
Correlation $\alpha = 0.9$	$2.04 \cdot 10^{-1}$	$1.53 \cdot 10^{-1}$

Table 4.7: Performance of the MFNNs trained with Rayleigh matrices for obtaining the MI of 2×2 SM system with antenna correlation.

receive antennas, are equal to

$$\mathbf{R}_t = \begin{pmatrix} 1 & \alpha \\ \alpha^* & 1 \end{pmatrix} \tag{4.40}$$

and

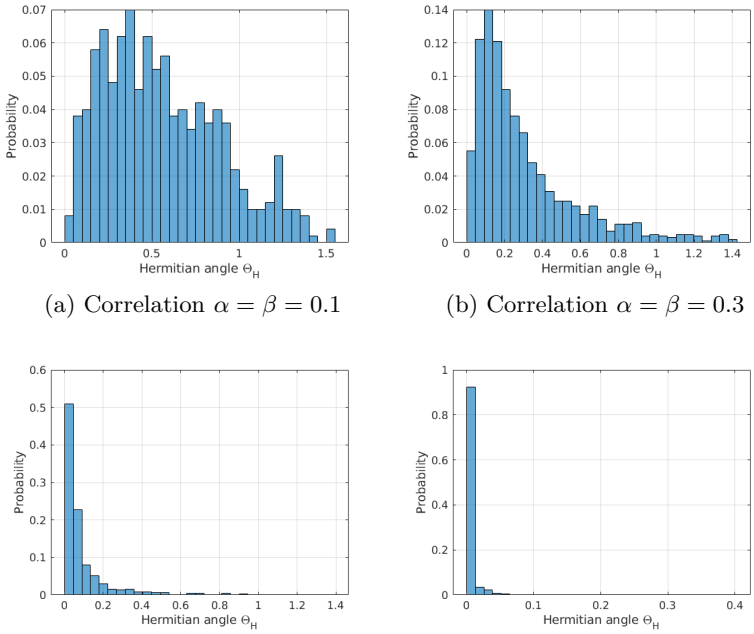
$$\mathbf{R}_r = \begin{pmatrix} 1 & \beta \\ \beta^* & 1 \end{pmatrix}. \tag{4.41}$$

In these new simulations the same value for α and β is employed and four different correlation values are tested, 0.1, 0.3, 0.6 and 0.9. For each correlation value, $1,000$ new matrices $\mathbf{H}$ are generated following (4.39) and a random SNR γ value between -5 and 20 dB is associated to each matrix. Then, the MI for each tuple $(\gamma, \mathbf{H})$ is calculated for the same three constellations by means of Monte Carlo simulations. The same two neural networks tested with the DP satellite channel matrices are employed to calculate the MI for this case with correlation. Table 4.7 shows the global MSE for the four correlation values. To ease the comparison, the first row shows again the MSE results for the uncorrelated case from the original ML dataset.

Table 4.7 reveals that the performance of the neural network gets worse when the correlation between the antennas is higher. This bad generalization performance can be understood with the help of Figs. 4.6 and 4.11. The former shows that, in the original dataset with Rayleigh distributed channel matrices, low and high values of the Hermitian angle Θ_H are underrepresented. Contrary, Fig. 4.11, with the distribution of this angle in the new datasets with antenna correlation, shows that, specially for α greater or equal than 0.6, the probability is concentrated around angles close to zero. Therefore, neural networks, trained in a dataset where small values of this angle are very unlikely, are forced to calculate the MI in a different dataset where this angle tends to be zero, specially for high correlations. As we saw, the neural networks did not generalize very well in this situation since they were not trained with enough examples of matrices whose Hermitian angle is close to zero.

Although the neural networks trained with Rayleigh distributed chan-

nel matrices do not perform as well in the new datasets with high antenna correlation, this case it is not so interesting for SM. If the transmit antenna correlation is high, the capacity decreases due to the fact that the antenna selected to transmit the modulation symbols cannot be distinguished so easily and SM loses interest. However, the training dataset could be augmented with this type of matrices with antenna correlation and, in this way, it is expected that the neural network can learn how to calculate the MI better in the correlated case.



(a) Correlation $\alpha = \beta = 0.1$

(b) Correlation $\alpha = \beta = 0.3$

(c) Correlation $\alpha = \beta = 0.6$

(d) Correlation $\alpha = \beta = 0.9$

Figure 4.11: Histograms with the distribution of the Hermitian angle Θ_H in a dataset of 1 000 random matrices $\mathbf{H}$ with transmission antenna correlation α and reception antenna correlation β.

4.6. Capacity of GSM: Simulation Results

Seven scenarios were simulated to assess the merits of using MFNNs to compute the GSM channel capacity. The number of transmit and receive antennas was identical, ranging from 2 to 8, whereas the number of RF chains R was 1, 2 and 3. The seven scenarios are:

1. SM 2×2 (PMod), SM 4×4 and SM 8×8 (with $R = 1$ RF chain)

2. SS-GSM 6×6 with $R = 2$ and $R = 3$ RF chains
3. SS-GSM 8×8 with $R = 2$ and $R = 3$ RF chains

For each particular scenario a dataset of $50,000$ realizations of $\mathbf{H}$ was generated, with matrices following a unit-variance Rayleigh distribution, i.e., $h_{ij} \sim \mathcal{CN}(0, 1)$. The SNR is drawn from a uniform random variable between -20 and 20 dB. The true capacity of each tuple $(\gamma, \mathbf{H})$ was calculated with (4.25), by using a Monte Carlo simulation with $5,000 \cdot L$ realizations of $\mathbf{y}$, where L denotes the number of spatial symbols.

For each scenario the dataset was divided again into two independent parts. 15% of the samples were reserved for the final test of the performance of the MFNN. The remaining 70% and 15% were employed for training and validation of the neural network, respectively. Each network, one per scenario, was trained 10 times using different random initial values for the weights and biases. Finally, the parameters which provide the lowest MSE of the 10 independent trainings were retained. 10 and 20 neurons in the hidden layer of the MFNN were tested. As a reference of the computing time with Matlab®, each training typically lasted less than 5 minutes, whereas the generation of the entire dataset with Monte Carlo simulations required about 50 hours by using several cores of a processor in parallel.

Table 4.8 shows the results obtained with a one-hidden layer 20-neurons MFNN. In each case, apart from the MSE, we also include the typical error (3 times the standard deviation of the error) and the maximum error, all of them measured on the samples reserved for testing. The number of neurons and inputs of the net, the number spatial symbols L and the number of angles involved in each scenario $\binom{L}{2}$ are also included.

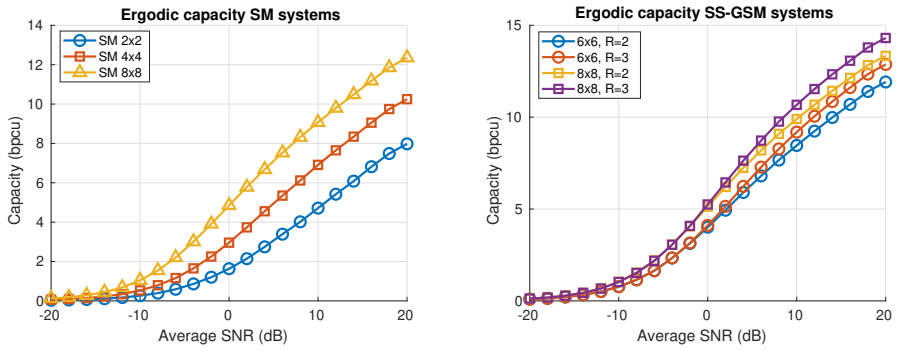
As it can be seen in Table 4.8, the MSE is always in the order of 10^{-4} for all the SM and SS-GSM scenarios. The error, seemingly following a Gaussian distribution, shows a typical and maximum value almost always below 0.07 and 0.10, respectively. This good estimation of the GSM capacity would make it possible for adaptive transmitters to select the Modulation and Coding Scheme (MCS) according to the capacity calculated and fed back by the receiver.

In order to grasp the relative magnitude of the error, Fig. 4.12 shows the ergodic capacity as a function of the average SNR for the different SM and SS-GSM cases, computed with the same datasets, i.e., for Rayleigh distributed channel matrices. Note that the combined transmit symbols $\mathbf{h}_l s$ are not Gaussian, which would be required to achieve the channel capacity of GSM [33]. This is why the capacity that we are computing and displaying in the figure is constrained to the specific selection mechanism of the antennas

described in the chapter. Thus, the capacity curves are not necessarily convex as usual for Gaussian symbols.

The left part of Figs. 4.13 and 4.14 shows the histogram of the absolute error obtained with the MFNN when calculating the capacity of a 4×4 SM system and a SS-GSM system with 8×8 antennas and $R = 3$ RF chains, respectively. The histograms reveal that the error follows approximately a Gaussian distribution. On the other hand, the right side graphics of Figs. 4.13 and 4.14 depicts the quality of the estimation with a regression plot, showing the estimated versus the true value of the capacity. The neural network predictions are almost indistinguishable from the Y=X line due to the low error they have.

The proposed MFNN is a very efficient way of calculating the capacity of SM and GSM systems that, otherwise, would require resorting to long Monte Carlo simulations. For example, obtaining a value of capacity for a SS-GSM 8×8 system with $R = 3$ RF chains requires 5.58 ms (98% for the pre-processing) with the MFNN based capacity estimation. However, the Monte Carlo simulation lasts between 100 and 10 000 times longer, depending on the required level of accuracy. All these computational times are based on computation time in Matlab® run in a computer equipped with an i7-4510U 2 GHz processor.



(a) SM (b) SS-GSM

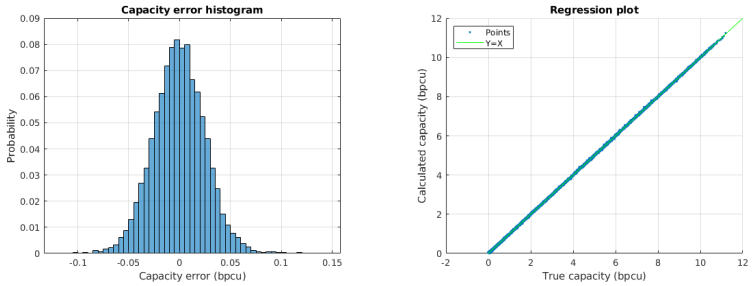
Figure 4.12: Ergodic capacity for Rayleigh channel, $h_{ij} \sim \mathcal{CN}(0, 1)$, of several SM and SS-GSM systems.

As to MS-GSM, with independent symbols transmitted per each RF chain, the norms and angles between the column vectors $\mathbf{c}_l$ are not rich enough to yield a good capacity estimate, since they cannot measure how the different streams of symbols interfere with each other. Although not shown here, we have used the eigenvalues of the submatrices of $\mathbf{H}$ given by the antenna selection. Thus, if the quantiles of the eigenvalues are used as

inputs to the NN, the MSE is around 10^{-3}, one order of magnitude lower than when the same inputs employed for SS-GSM are used instead.

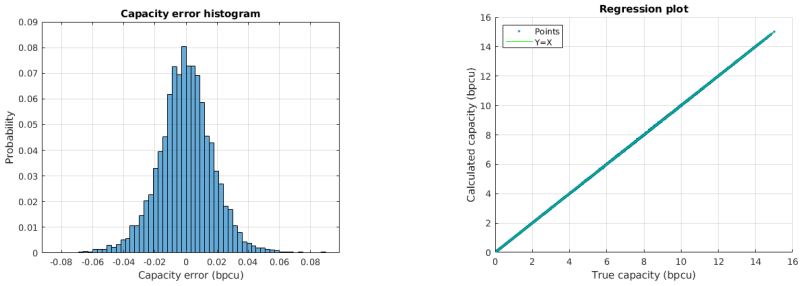
Scenario	MSE	3σ	Max. error	Num. Neurons	NN inputs	L	$\binom{L}{2}$
SM 2×2 (PMod)	$5.27\cdot10^{-4}$	0.069	0.092	20	$4=2+2$	2	1
SM 4×4	$6.53\cdot10^{-4}$	0.077	0.142	20	$16=4+2\times6$	4	6
SM 8×8	$4.85\cdot10^{-4}$	0.066	0.097	20	$26=8+2\times9$	8	28
SS-GSM 6×6, R=2	$4.00\cdot10^{-4}$	0.060	0.082	20	$26=8+2\times9$	8	28
SS-GSM 6×6, R=3	$2.82\cdot10^{-4}$	0.050	0.109	20	$27=9+2\times9$	16	120
SS-GSM 8×8, R=2	$2.75\cdot10^{-4}$	0.050	0.080	20	$27=9+2\times9$	16	120
SS-GSM 8×8, R=3	$3.08\cdot10^{-4}$	0.053	0.090	20	$27=9+2\times9$	32	496

Table 4.8: Performance of a 20-neurons MFNN for calculating the GSM capacity in several scenarios.



(a) Error histogram (b) Regression plot

Figure 4.13: Capacity error histogram and regression plot for a SM system with 4×4 antennas.



(a) Error histogram (b) Regression plot

Figure 4.14: Capacity error histogram and regression plot for a SS-GSM system with 8×8 antennas and $R=3$ RF chains.

4.7. Conclusions

Spatial Modulation (SM) and Generalized Spatial Modulation (GSM) can play an important role in future 5G terrestrial and satellite networks to increase the spectral efficiency while maintaining a reduced number of RF chains and, consequently, a low power consumption. The implementation of next generation adaptive SM and GSM links requires practical mechanisms to estimate the maximum achievable rate that the channel can support. In this context, the calculation of the capacity and the MI of these systems gains traction not only as a problem from the information theory field but also as a problem with practical interest for adaptive systems.

Some analytical approximations were previously known in the literature to calculate the Mutual Information (MI) of SM systems constrained to a given constellation. However, although these expressions avoid running Monte Carlo simulations to make the MI calculation, they still entail a high computational complexity, which increases with the square of both the number of antennas and the constellation size, with limited accuracy.

In this chapter, a new method is proposed for calculating the MI of a SM based on Machine Learning (ML). Namely, a Multilayer Feedforward Neural Network (MFNN) of just one hidden layer is trained using supervised learning; simulation results show that it is able to make a more accurate estimation of the MI with less computational complexity than the previous analytical approximations. Similarly, it is shown that the capacity of GSM can also be calculated with high accuracy using MFNNs. In both cases, some specific features are extracted from the channel matrix and the Signal to Noise Ratio (SNR), which are employed to feed the neural networks and improve their performance.

Adaptive systems can benefit from this accurate and rapid computation of the MI and the capacity. The maximum achievable rate allowed by the channel conditions can be computed in real time, and a transmitter can adapt the constellation order and apply a fine tuning of the coding rate of the channel encoder, providing a better fit to the channel capacity in adaptive SM and GSM systems. In Chapter 5 this approach will be explored in Dual Polarization satellite systems with Polarized Modulation and in Chapter 6 for terrestrial systems with SM.

4.8. Appendix. Multilayer Feedforward Neural Networks

This appendix describes the Multilayer Feedforward Neural Network (MFNN) which is employed to make Mutual Information (MI) and capacity calculations in Spatial Modulation (SM) and Generalized Spatial Modulation (GSM) systems. Fig. 4.15 shows a diagram with the structure of the MFNN. The equations and functions involved in this one hidden layer MFNN are provided hereafter.

In the following, the variables in blue will denote the internal parameters of the neural network, and $\mathbf{a}_i, i = 0, 1, 2$ will be the intermediate internal variables at the i-th stage. We describe a general MFNN with F inputs, N neurons and K outputs.

Each of the F neural network inputs goes through a linear preprocessing block to adjust the neurons input to the range $[-1, +1]$. This initial scaling is expressed as

$$\mathbf{a}_0 = \mathbf{g}_0 \circ (\mathbf{x} - \mathbf{x}_0) - \mathbf{1} \ \in \mathbb{R}^{F\times 1} \tag{4.42}$$

with the gain $\mathbf{g}_0 \in \mathbb{R}^{F\times 1}$ and the offset $\mathbf{x}_0 \in \mathbb{R}^{F\times 1}$. $\circ$ is simply the Hadamard (pointwise) matrix product.

The single hidden layer is made of N neurons, also named processing units, each applying a weighted linear combination of its inputs, a bias and a non-linear function, also known as activation function:

$$\mathbf{a}_1 = g(\mathbf{W}_1 \cdot \mathbf{a}_0 + \mathbf{b}_1) \in \mathbb{R}^{N\times 1}. \tag{4.43}$$

The matrix $\mathbf{W}_1 \in \mathbb{R}^{N\times F}$ and the vector $\mathbf{b}_1 \in \mathbb{R}^{N\times 1}$ collect the weights and the offsets. As activation function we use the hyperbolic tangent:

$$g(x) = \frac{2}{1 + e^{-2x}} - 1.$$

The output layer of K neurons applies a linear processing of the form

$$\mathbf{a}_2 = \mathbf{W}_2 \cdot \mathbf{a}_1 + \mathbf{b}_2 \ \in \mathbb{R}^{K\times 1} \tag{4.44}$$

for matrix $\mathbf{W}_2 \in \mathbb{R}^{K\times N}$ and vector $\mathbf{b}_2 \in \mathbb{R}^{K\times 1}$.

Finally, there is a last stage to accommodate the range of the network outcome:

$$\mathbf{y} = (\mathbf{a}_2 + \mathbf{1}) \oslash \mathbf{g}_3 + \mathbf{y}_0 \ \in \mathbb{R}^{K\times 1} \tag{4.45}$$

with the gain $\mathbf{g}_3 \in \mathbb{R}^{K\times 1}$ and the offset $\mathbf{y}_0 \in \mathbb{R}^{K\times 1}$. $\oslash$ denotes the Hadamard (pointwise) matrix division.

The output vector is expressed as $\mathbf{y} = [y_1, y_2, \ldots, y_K]^t$.

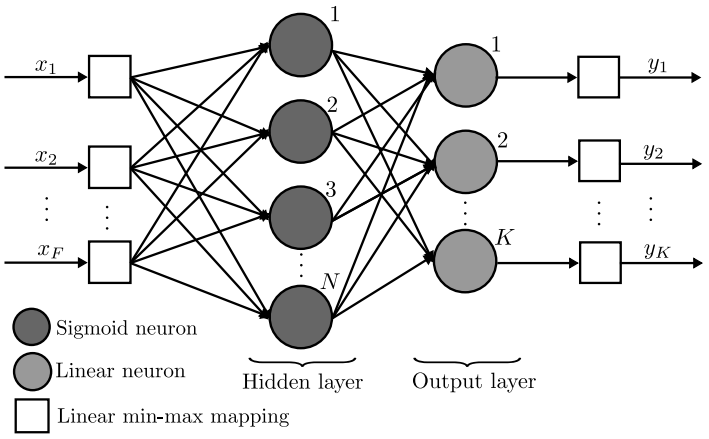


Figure 4.15: Diagram of the neural network.

The different parameters of the network –weights and biases in (4.42)-(4.45)– are to be obtained under supervised learning. The Levenberg-Marquardt (LM) backprogation algorithm [45] is used to minimize the Mean Squared Error (MSE) on the test set:

$$\text{MSE} = \frac{1}{L}\sum_{\ell=1}^{L}\|\mathbf{y}(\ell) - \mathbf{t}(\ell)\|^2, \tag{4.46}$$

with $\mathbf{y}(\ell)$ and $\mathbf{t}(\ell)$ the network output and the target values, respectively, for the training tuples $(\mathbf{x}(\ell), \mathbf{t}(\ell))$, $\ell = 1, \ldots, L$.

Chapter 5

Link Adaptation in Mobile Satellite Systems with Dual Polarization

This chapter is adapted, with permission of the coauthors and the editorial, from Springer: "A. Tato, P. Henarejos, C. Mosquera, A. Pérez-Neira. Link Adaptation Algorithms for Dual Polarization Mobile Satellite Systems. In P. Pillai, K. Sithamparanathan, G. Giambene, M. Á. Vázquez, and P. D. Mitchell, editors, Wireless and Satellite Systems, pages 52-61, Cham, 2018. Springer International Publishing", [95]. And from IEEE: "A. Tato, C. Mosquera, P. Henarejos, A. Pérez-Neira. Practical Implementation of Link Adaptation with Dual Polarized Modulation. In 2018 11th International Symposium on Communication Systems, Networks Digital Signal Processing (CSNDSP), pages 1-6, Budapest, Hungary, July 2018", [96].

5.1. Introduction

In recent years, the spectrum saturation and the increasing demand for higher data rates in a ubiquitous way encourages the engineers to design new techniques in order to increase the capacity of communication systems without resorting to expand the occupied bandwidth. Two of these techniques consist on the leverage of multiple antennas at both transmitter and receiver by means of MIMO (Multiple Input Multiple Output) signal processing techniques and also the so-called Adaptive Coding and Modulation (ACM) or link adaptation. Both are part of many current terrestrial wireless communication standards such as LTE [1] and IEEE 802.16 [2] for cellular

technologies and IEEE 802.11 [3] for wireless local area networks.

The simultaneous use of two polarizations is a means to increase the spectral efficiency that is being explored for mobile satellite communications [48]. At low frequencies, like L- and S-band, only one polarization has been traditionally used in fear of too small Cross-Polar-Discriminations (XPD) [75]. At these bands circular polarizations are preferred over linear polarizations, to avoid the effects of the Faraday rotation [75]. Although Right Hand Circular Polarization (RHCP) was the typical option, it is possible to use simultaneously two orthogonal polarizations, RHCP and Left Hand Circular Polarization (LHCP), to communicate with the users. This is analogous to a 2×2 MIMO system, simply replacing the spatial by the polarization component. Therefore, the application of MIMO signal processing techniques makes it possible to achieve throughput gains while maintaining the same transmit power, even in the presence of interference between the two polarizations.

There are different MIMO modes which can be employed in the scenario of a mobile satellite system with Dual Polarization (DP) but, in practice, only those not requiring Channel State Information at the Transmitter (CSIT) are of interest because of the long propagation delay. Two candidates are Orthogonal Polarization-Time Block Code (OPTBC), based on 2×2 Alamouti Space-Time Coding, and Vertical-Bell Laboratory Layered Space-Time (V-BLAST), a spatial multiplexing technique which transmits two independent streams of symbols, one per each polarization. These two are well-known MIMO modes, which are also part of cellular systems like LTE [1]. An additional mode considered in this chapter is Polarized Modulation (PMod). This is a particular type of an Index Modulation which allows a remarkable gain in spectral efficiency at intermediate SNRs (Signal to Noise Ratios) which cannot be achieved by just employing OPTBC and V-BLAST.

In this chapter a DP mobile satellite system with four different MIMO modes, SISO (Single Input Single Output), OPTBC, PMod and V-BLAST, is analyzed. With SISO we mean a baseline system which only uses one polarization. The transmitter is allowed to switch among the four MIMO modes across frames to maximize the spectral efficiency. Apart from the MIMO mode, a Modulation and Coding Scheme (MCS) must be selected for each frame, yielding two levels of adaptation. Therefore, the link adaptation algorithm, in charge of selecting both MIMO mode and MCS, will be described in this chapter. The proposed algorithm is based on some adaptive margins and the use of Lookup Tables (LUTs).

The particular way of doing the adaptation when PMod is the MIMO

mode selected differs notably from previous works that deal with link adaptation for Index Modulations, like [110] does with Spatial Modulation. Those publications only consider adaptive modulation, i.e., the selection of the modulation order. However, since we have available methods to calculate the constrained capacity of PMod, we can go further and adapt the coding rate and not only the modulation. To make the PMod adaptation, we propose the use of two independent variable rate channel encoders, one encoder for the bit stream which selects the polarization and another for the bit stream which selects the modulation symbol. This means that two LUTs are used in PMod adaptation, one for selecting the coding rate of the polarization bits and other for selecting the MCS of the modulation symbols.

The simulation results provided in the chapter are made using Physical Layer Abstraction (PLA) techniques to avoid the run of all the transmitting and receiving chains. For each simulated transmitted frame a series of the instantaneous channel matrices the frame undergoes is generated, and then the effective SNR and effective rate that the channel supports are calculated. The latter is based on the RBIR (Received Bit Mutual Information Rate) [60]. These metrics allow to decide if the receiver can decode correctly the frame or not, depending on the MIMO mode and MCS used to transmit that frame. In the next chapter, a different approach will be explored, avoiding PLA and running all the reception chain and using a Deep Learning approach to select the MCS in a general Index Modulation system, with application to both satellite systems with PMod or terrestrial 5G systems which use Spatial Modulation.

This chapter is structured as follows. After the Introduction, Section 5.2 provides the general equations of the system model and describes the particular satellite scenario studied in this chapter. Then, in Section 5.3 the channel model and channel generation is described. Section 5.4 explains the PLA assumed along this chapter and then Sections 5.5 and 5.6 introduce the algorithms for selecting the MIMO mode and the MCS, respectively. Lastly, Section 5.7 provides all the simulation results before collecting the main conclusions in Section 5.8.

5.2. System Model

The forward link of a satellite communication system which serves mobile users in the L-band is considered here, with a transmitter (the gateway) communicating with a receiver (the Mobile Terminal). Both sides are equipped with DP antennas which enables the transmission and reception of RHCP and LHCP signals simultaneously. The system, narrow-band and single-

carrier, is inspired on the commercial service BGAN (Broadband Global Area Network), standardized by the ETSI as TS 102 744, [9], but adding as extra feature the use of two polarizations, a feature not included in that standard.

We assume that the satellite system can employ four different MIMO modes. The reason for having more than one is that there is not a single mode which performs better in all channel conditions. Depending on the SNR and the characteristics of the channel matrix, one or another provide the highest spectral efficiency. The MIMO mode used in the frames transmitted by the gateway can be changed from one frame to another. Since the selection of the MIMO mode requires the knowledge of the channel matrix, this will be done by the receiver, which has perfect Channel State Information (CSI). The main reason is that this avoids the feedback of all the channel matrix coefficients to the transmitter. The four MIMO modes are

- **SISO**: The baseline mode where only one polarization, RHCP, is used. Comparison to all the other DP MIMO modes allows to calculate the throughput gain of a DP system.
- **OPTBC**: This mode is based on Alamouti space-time coding, used for achieving transmit diversity and introduced in [12]. The spatial components in Alamouti are replaced here by the two available polarizations in OPTBC. This mode, where the two polarizations are used simultaneously, allows to increase the spectral efficiency for the low SNRs, as it will be shown later.
- **PMod**: This MIMO mode is a particular case of a 2×2 Index Modulation (IM). It is analogous to Spatial Modulation (SM) [47] but with the two dimensions representing different polarizations instead of representing antennas (space). In PMod only one polarization is used at each time instant yielding an up to 50 % gain in spectral efficiency when a QPSK constellation is used.
- **V-BLAST**: This MIMO mode is based on spatial multiplexing techniques [104]. Therefore, two independent streams of symbols are transmitted, one per polarization. The receiver is in charge of reducing the cross-stream interference to perform the detection.

OPTBC, PMod and V-BLAST are MIMO modes which do not require CSI at the transmitter (CSIT), contrary to, for example, precoding, where the transmitter must know the channel matrix. Those three MIMO modes are chosen because their transmitted signal encoding is independent of the channel matrix, a matrix which would change in a round-trip time due to the long propagation delay of GEO satellite channels.

The system model of a general DP system for a given discrete time instant is

$$\mathbf{y} = \sqrt{\gamma}\mathbf{H}\mathbf{x} + \mathbf{w}, \tag{5.1}$$

where $\mathbf{y} \in \mathbb{C}^2$ is the received vector, γ is the average SNR, $\mathbf{H} = (\mathbf{h}_1\ \mathbf{h}_2) \in \mathbb{C}^{2\times 2}$ is the channel matrix, $\mathbf{x} \in \mathbb{C}^2$ is the transmitted signal and $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_2)$ is the Additive White Gaussian Noise (AWGN). For the particular case of PMod, since $\mathbf{x}$ has always a zero component, (5.1) can be expressed as

$$\mathbf{y} = \sqrt{\gamma}\mathbf{h}_l s + \mathbf{w}, \tag{5.2}$$

with $\mathbf{h}_l$ the l-column of the channel matrix, where the index l selects the polarization that transmits the complex symbol s, taken from a constellation $\mathcal{S}$.

Fig. 5.1 shows a diagram of the system where the gateway communicates with a Mobile Terminal (MT) by means of a satellite. In this thesis the attention is focused on the forward link, i.e., on the communications in the direction gateway-MT. Dual polarization (DP) is used in the forward link and an ideal return link is assumed, employed by the receiver to feedback to the gateway all the information needed to make the adaptation.

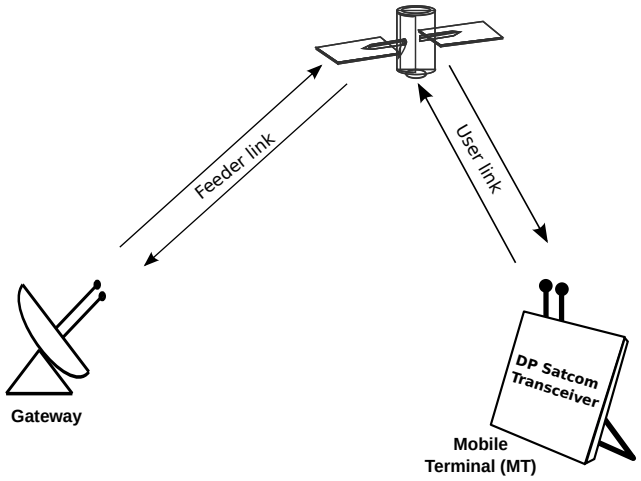
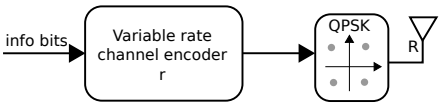


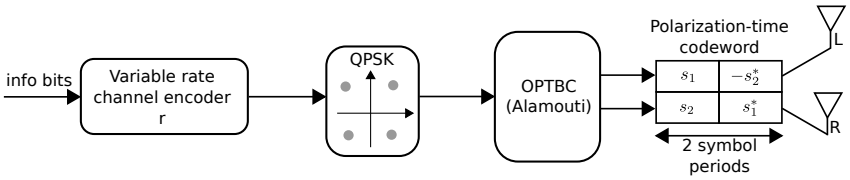
Figure 5.1: Diagram of the mobile satellite communications system.

A frame based communication is assumed in this chapter, where the transmitter groups the symbols into frames or codewords of length N, which are transmitted at a fixed baud rate of f_s symbols per second. Fig. 5.2 shows a block diagram of each transmitter configuration for each one of the four MIMO modes. It can be seen that a channel encoder with variable coding rate is used in all the modes except in PMod, where two channel encoders with independent rates are used, one for the bits selecting the symbols and

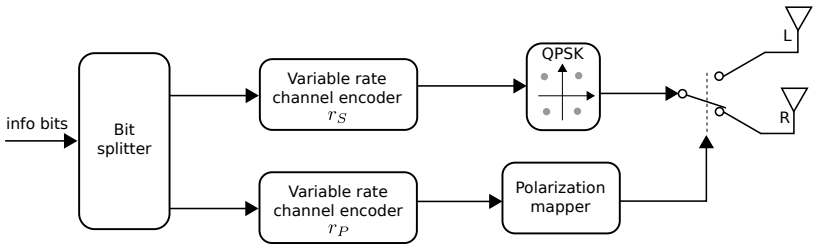
other for the bits selecting the polarization. From the diagram it is clear that SISO, contrary to the other modes, only employs one polarization. Although in this chapter the modulation employed by the transmitter is always the QPSK, one of the options of [9], similar results could be obtained for other constellations or for a combination of several constellations.



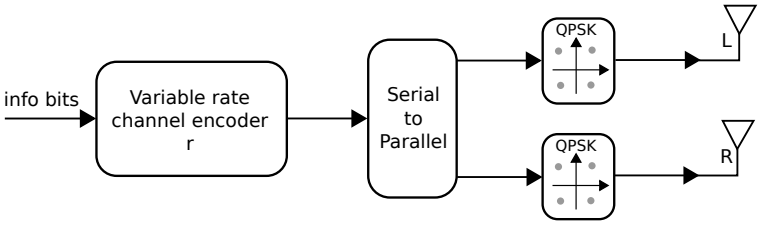
(a) SISO



(b) OPTBC



(c) PMod



(d) V-BLAST

Figure 5.2: Block diagrams of the four configurations of the transmitter for each one of the four MIMO modes.

5.3. Channel Model and Channel Generation

The simulation of the mobile satellite dual polarized channel has been done following the work of [88]. The channel is obtained as the sum of three

different components: the Line-of-Sight (LOS), the specular reflected signal and the diffuse components, produced by the scattering objects near the Mobile Terminal. The specular component, which is important in maritime and aeronautical communications due to the reflection over the calm sea surface, is usually neglected in the terrestrial Land Mobile Satellite (LMS) dual polarized channel. LOS and specular components are modelled as Rician random variables whereas the diffuse component, which causes the fast fading, is Rayleigh distributed. The expression of the channel matrix is, as explained before, the sum of three components:

$$\mathbf{H} = \boldsymbol{\beta} e^{j\phi} \mathbf{K}_L + \boldsymbol{\xi} e^{j\phi} \mathbf{K}_S + \mathbf{D}\mathbf{K}_D \tag{5.3}$$

$\mathbf{K}_L$, $\mathbf{K}_S$ and $\mathbf{K}_D$ are the K-factor matrices of the LOS, specular and diffuse components, respectively and take some fixed values regarding the considered environment (open areas, suburban/rural areas, urban areas or maritime/aeronautical). These matrices depend on the following parameters:

- K_1^L and K_2^L, the Rice factors of the LOS components for each one of the polarizations.
- K_1^S and K_2^S, the Rice factors of the specular components for each one of the polarizations.

The values of the three matrices are

$$\mathbf{K}_L = \begin{pmatrix} \sqrt{\dfrac{K_1^L}{K_1^L + K_1^S + 1}} & 0 \\ 0 & \sqrt{\dfrac{K_2^L}{K_2^L + K_2^S + 1}} \end{pmatrix} \tag{5.4}$$

$$\mathbf{K}_S = \begin{pmatrix} \sqrt{\dfrac{K_1^S}{K_1^L + K_1^S + 1}} & 0 \\ 0 & \sqrt{\dfrac{K_2^S}{K_2^L + K_2^S + 1}} \end{pmatrix} \tag{5.5}$$

$$\mathbf{K}_D = \begin{pmatrix} \sqrt{\dfrac{1}{K_1^L + K_1^S + 1}} & 0 \\ 0 & \sqrt{\dfrac{1}{K_2^L + K_2^S + 1}} \end{pmatrix} \tag{5.6}$$

The two matrices $\boldsymbol{\beta}$ and $\boldsymbol{\xi}$ in (5.3) are also fixed for a given scenario and depend on the following parameters:

- β_1 and β_2 take values between 0 and 0.5 and are inversely related with the XPD of each one of the polarizations for the LOS component.
- ξ_1 and ξ_2 also take values between 0 and 0.5 and are inversely related with the XPD of each one of the polarizations for the specular component.

The expression of these two matrices are

$$\boldsymbol{\beta} = \begin{pmatrix} \sqrt{1-\beta_1} & \sqrt{\beta_2} \\ \sqrt{\beta_1} & \sqrt{1-\beta_2} \end{pmatrix}, \tag{5.7}$$

$$\boldsymbol{\xi} = \begin{pmatrix} \sqrt{1-\xi_1} & \sqrt{\xi_2} \\ \sqrt{\xi_1} & \sqrt{1-\xi_2} \end{pmatrix}. \tag{5.8}$$

Lastly, the components of the diffuse matrix $\mathbf{D}$ in (5.3) are zero mean complex Gaussian random variables. The covariance matrix Σ of the random vector with the components of the matrix $\mathbf{D}$ ($\mathbf{d} = [d_{11}\ d_{12}\ d_{21}\ d_{22}]$) is

$$\Sigma = \begin{pmatrix} 1-\alpha_1 & \rho_{1t}\sqrt{(1-\alpha_1)\alpha_1} & \rho_{1r}\sqrt{(1-\alpha_1)\alpha_2} & 0 \\ \rho_{1t}\sqrt{(1-\alpha_1)\alpha_1} & \alpha_1 & 0 & \rho_{2r}\sqrt{(1-\alpha_2)\alpha_1} \\ \rho_{1r}\sqrt{(1-\alpha_1)\alpha_2} & 0 & \alpha_2 & \rho_{2t}\sqrt{(1-\alpha_2)\alpha_2} \\ 0 & \rho_{2r}\sqrt{(1-\alpha_2)\alpha_1} & \rho_{2t}\sqrt{(1-\alpha_2)\alpha_2} & 1-\alpha_2 \end{pmatrix} \tag{5.9}$$

In the previous covariance matrix α_1 and α_2 have a role similar to β_i and ξ_i, but for the diffuse component, and ρ_{it} and ρ_{ir} are related with the cross-correlation between the copolarized signals [88].

5.3.1. Channel generation

One important aspect of the generated channel series is the time correlation and the Doppler spread. Assuming the Clarke's model, the coherence time can be approximated by

$$\tau_c = \frac{3\lambda}{4v\sqrt{\pi}}, \tag{5.10}$$

being λ the wavelength and v the mobile terminal speed. For obtaining the N channel matrices of a frame, we first generate $Q = \lceil N/(\tau_c f_s) \rceil$ independent realizations of the channel coefficients with the parameters of the selected scenario. Then we make a linear interpolation to obtain the N channel matrices and a low-pass-filter with cut-off frequency equal to the Doppler spread, $D_s = v/\lambda$, is applied. Lastly, the channel matrices are scaled to obtain an average SNR equal to the LOS SNR of the simulation.

In Figure 5.3 is shown a block diagram of the channel time series generator. Firstly, Q independent samples are generated for both the uniform distributed phase term ϕ and for the Gaussian random vector $\boldsymbol{d}$, zero mean and with covariance matrix Σ. Then, the multiplication by the matrices defined previously takes place and the three components are added. Lastly the interpolation and filtering is done.

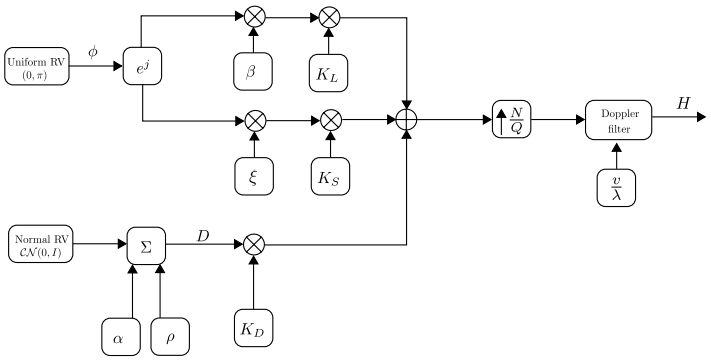


Figure 5.3: Block diagram of the channel time series generator

In a later section simulation results are provided for a maritime scenario. The parameters employed in that channel generation, taken from [88], are $K_1^L = K_2^L = 10$, $K_1^S = K_2^S = 5$, $\beta_1 = \beta_2 = 0.3$, $\alpha_1 = \alpha_2 = 0.4$, $\rho_1^t = \rho_{2t} = 0.5$, $\rho_{1r} = \rho_{2r} = 0.5$, $\xi_1 = \xi_2 = 0.3$. Fig. 5.4 shows the evolution of the four coefficients of the channel matrix during one second (13 frames of 2560 symbols). It can be seen how the diagonal elements are bigger than the other two but the non-diagonal elements still take some non-negligible value.

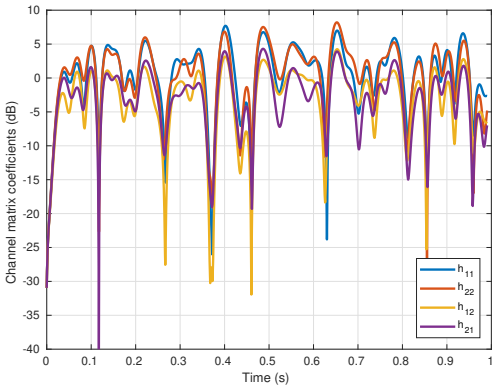


Figure 5.4: Evolution of the channel matrix coefficients amplitude during 1 s (maritime scenario, $f_s = 33,600$ symbols/s, speed 50 km/h)

5.4. Physical Layer Abstraction

In order to avoid long simulations, resorting to Physical Layer Abstraction (PLA) techniques [53] is a common practice. In this way, we avoid to run the entire signal processing chain of both transmitter and receiver PHY layer, which is very time consuming. The simulations contained in this chapter repeat the following procedure for a number of M simulated frames:

1. The transmitter uses the MIMO mode signaled by the receiver and executes the link adaptation algorithm for selecting the MCS for the next transmitted frame, using the information fed back by the receiver (acknowledgments and CSI).

2. The time varying channel is generated for that received frame. The channel state consists on an average SNR γ, fixed for each simulation, and a series of N channel matrices $\{\mathbf{H}_n,\ n = 1, 2, \ldots, N\}$, one matrix per symbol of the received frame.

3. At reception it must be decided if the frame can be decoded correctly or not, given the current channel conditions and the MIMO mode and MCS used in the transmission. It is at this point where the PLA is useful for making this decision without simulating all the decoding process.

4. Lastly, the receiver feeds back to the transmitter the acknowledgment (an ACK if the frame was decoded or a NAK if it was not), the optimum MIMO mode for the next transmission, and the CSI, in the form of the effective SNR of the channel for that particular MIMO mode. In order to emulate the long Round-Trip Time (RTT) with GEO satellites there is a delay in the feedback equivalent to d frames.

To decide the decoding outcome, a metric named Mutual Information effective SNR [78], or simply effective SNR is calculated. This metric condenses in a single value all the channel variations during a frame spanning N symbols. The comparison of the effective SNR with the threshold SNR required for decoding a frame using a particular MCS, SNR_{th}, allows to determine if the result of the decoding is correct or not. Therefore, when the effective SNR of the channel is greater or equal than the threshold SNR of the MCS used to transmit a frame, this is supposed to be correctly decoded. And, if it is lower, we decide that there is an error in the decoding.

Two things remain to be commented. On the one hand, how the threshold SNR of a MCS is obtained. And, on the other hand, how the effective SNR of the channel is calculated, taking into account that this depends on

the particular MIMO mode. The same series of channel matrices provide different values of the effective SNR depending on the particular mode, SISO, OPTBC, PMod or V-BLAST.

Let us start with the threshold SNR. It is assumed that the transmitter can use a set of MCS, each one characterized by a particular choice of the constellation and the coding rate. Table 5.1 shows the set of available MCS which will be used later in the simulations. A QPSK constellation with coding rates r between 0.34 and 0.87 are used, which yield a spectral efficiency of $2r$ bits/s/Hz for SISO mode. The particular selection of the MCS corresponds to the bearer F80T1Q-1B of the Family SL of the Satellite-UMTS system [9]. This is a standard for the 3rd generation mobile satellite communications, which is used by the service BGAN (Broadband Global Area Network) of Inmarsat, for providing Internet access and voice calls to the mobile terminals using a GEO satellite.

Table 5.1: Set of MCS for the symbols with a QPSK constellation.

	L8	**L7**	**L6**	**L5**	**L4**	**L3**	**L2**	**L1**	**R**
Coding rate	0.34	0.40	0.48	0.55	0.63	0.70	0.77	0.83	0.87
Spectral efficiency (bits/s/Hz)	0.68	0.80	0.96	1.10	1.26	1.40	1.54	1.66	1.74
Threshold SNR (SNR_{th}) **(dB)**	-2.15	-1.21	-0.09	0.83	1.84	2.74	3.67	4.54	5.19

Fig. 5.5 shows the curve of the mutual information (MI) $\Theta(\gamma)$ as a function of the SNR γ for a QPSK constellation with some markers placed at the points of the MCS of Table 5.1. Using that curve, the threshold SNR required for each MCS can be easily calculated. To obtain this curve, Monte Carlo simulations must be run using the following expression [53]-[78]

$$\Theta(\gamma) = \log_2 M - \frac{1}{M} \sum_{s \in \mathcal{S}} \mathbb{E} \left[\log_2 \sum_{s' \in \mathcal{S}} e^{-|\gamma(s-s')+w|^2 - |w|^2} \right], \tag{5.11}$$

where M, the constellation cardinality, takes the value 4, $\mathcal{S}$ is the set of constellation symbols, γ is the average SNR and w is a circularly symmetric white Gaussian noise with zero mean and unit variance. However, the curve of MI can be approximated by sums of exponential functions as proposed in [17]. The following expression matches perfectly the MI of the QPSK in AWGN channel, and it will be used recurrently in the system simulator:

$$\Theta(\gamma) \simeq 2 \cdot \left(1 - 0.8551 \cdot e^{-0.5718\gamma} - (1 - 0.8551) \cdot e^{-1.55\gamma} \right). \tag{5.12}$$

Equation (5.12) allows to jump from the SNR to the MI (or spectral efficiency), and its inverse, $\Theta^{-1}(I)$, from the MI domain to the SNR. By using $\Theta^{-1}(I)$, the threshold SNR required for decoding a given MCS of coding rate r can be calculated by doing simply $\mathrm{SNR}_{th} = \Theta^{-1}(2r)$. This

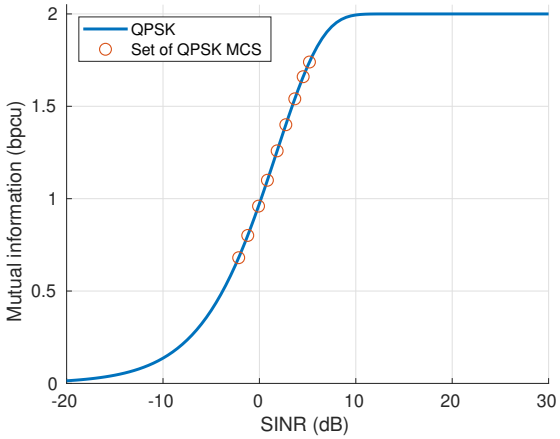


Figure 5.5: Mutual Information in an AWGN channel as function of the SNR for a QPSK constellation, marking with circles the MCS of Table 5.1.

SNR, in natural units, is the same for all the MIMO modes and represents the minimum SNR required for decoding that MCS. It is a theoretical lower bound and, in practice, the codes of the system may not reach that level of performance. However, in this chapter we do not assume any particular channel code and we consider that a frame transmitted with a MCS can be correctly decoded whenever the effective SNR of the channel for the MIMO mode used in the transmission is higher or equal than the threshold SNR of that MCS, which is shown in Table 5.1.

Now, let us focus our attention on how the effective SNR of a received frame is calculated for each one of the four MIMO modes, by using the N realizations of the channel matrices, one per symbol of the frame. For the same channel conditions each MIMO mode has a different effective SNR, depending on how the information is encoded and how the processing is done at the receive side. The effective SNR (or SINR, Signal to Interference and Noise Ratio) is calculated in two steps. The first step consists of the calculation of the SINR γ_n per symbol of the frame. The second step, named SINR compression, takes the N SINR values, γ_n, $n = 1, 2, \ldots, N$, and *compresses* them into a single value, the effective SINR [53]. Although we usually employ the term SNR instead of SINR for simplicity, take into account that specially with V-BLAST mode we are calculating the SINR actually, because apart from the noise there is some residual interference between the two streams due to the use of a Minimum Mean Square Error (MMSE) receiver.

The calculation of the SINR per symbol is done differently for each one of the modes as follows:

1. **SISO**: Only one polarization is used in transmission and reception.

$$\gamma_n = \gamma |[\mathbf{H}_n]_{11}|^2 \tag{5.13}$$

2. **OPTBC**: In this and the following modes the total available power is split between the two polarizations, that is the reason of using $\gamma/2$. This SINR is calculated with the Frobenius norm of the channel matrix.

$$\gamma_n = \frac{\gamma}{2} \|\mathbf{H}_n\|_F^2 \tag{5.14}$$

3. **PMod**: For this mode two different SINRs are calculated, one per polarization k. By $\mathbf{h}_{n,k}$ we represent the k-column of the channel matrix $\mathbf{H}_n$.

$$\gamma_{n,k} = \frac{\gamma}{2} \|\mathbf{h}_{n,k}\|^2, \ k = 1, 2 \tag{5.15}$$

4. **V-BLAST**: In this mode two independent streams of modulated symbols are transmitted, one in each polarization, therefore two SINRs are calculated again, one per polarization k.

$$\gamma_{n,k} = \frac{1}{\left[\left(\mathbf{I}_2 + \frac{\gamma}{2}\mathbf{H}_n^H \mathbf{H}_n\right)^{-1}\right]_{kk}} - 1, \quad k = 1, 2 \tag{5.16}$$

   The previous equation corresponds to the post-processing SINR with a linear MMSE (Minimum Mean Square Error) receiver. If other type of receivers were used, the calculation of the SINR should be changed accordingly. We have selected MMSE due to its simplicity and robustness against noise, when compared with other linear receivers such as the ZF (Zero Forcing).

To summarize the N values of SINR, calculated using the previous equations, into a single metric useful for predicting the outcome of the decoding, a step named *SINR compression* is done. This consists on an average in the domain of the Mutual Information and a final conversion to the SNR domain again, using the function $\Theta(\gamma)$ of equation (5.12) and its inverse $\Theta^{-1}(I)$. Then, the effective SNR for SISO and OPTBC is calculated as follows:

$$\mathrm{SNR}_{\mathrm{eff}} = \Theta^{-1}\left(\frac{1}{N}\sum_{n=1}^{N}\Theta(\gamma_n)\right). \tag{5.17}$$

For the symbols of PMod and for V-BLAST two different effective SNRs are calculated, one for each polarization, and then its minimum is employed as the final effective SNR:

$$\mathrm{SNR}_{\mathrm{eff,k}} = \Theta^{-1}\left(\frac{1}{N}\sum_{n=1}^{N}\Theta(\gamma_{n,k})\right), \ k = 1, \ 2, \tag{5.18}$$

$$\mathrm{SNR}_{\mathrm{eff}} = \min\left(\mathrm{SNR}_{\mathrm{eff},1}, \mathrm{SNR}_{\mathrm{eff},2}\right). \tag{5.19}$$

The reason for doing this can be understood looking at Fig. 5.2, where the same level of protection is applied to the symbols independently of the polarization used to transmit them. Hence, the polarization with the lowest effective SNR will limit the peformance.

The architecture proposed for the PMod MIMO mode consists of two channel encoders in parallel with independent coding rates, as shown in Fig. 5.2. One produces the bits which select the symbols and the other, the bits selecting the polarizations. Hence, in PMod there are two different codewords, one for symbols and another for polarizations, and, consequently, it must be decided independently if each one can be decoded by the receiver or not. The metric related to the codeword of the symbols is the effective SNR of equation (5.19). To decide the decoding outcome of the polarizations bits codeword, a metric we name *effective rate*, and which is some kind of RBIR (Received Bit Mutual Information Rate) [60], is used instead. In this case, we simply avoid the last transformation from MI to SNR of the SINR compresssion step.

The effective rate, which belongs to the interval $[0, 1]$, is obtained with

$$\mathrm{Rate}_{\mathrm{eff}} = \frac{1}{N}\sum_{n=1}^{N} I_P(\sqrt{\gamma}\mathbf{H}_n), \tag{5.20}$$

where $I_P(\sqrt{\gamma}\mathbf{H}_n)$ is the MI of the polarization bits, which is calculated by subtracting from the total MI I_T (which can be calculated using, for example, some of the neural networks proposed in the previous chapter) the MI related with the symbols, I_S, obtained easily with

$$I_S(\sqrt{\gamma}\mathbf{H}_n) = \frac{1}{2}\left(\Theta\left(\gamma\|\mathbf{h}_1\|^2\right) + \Theta\left(\gamma\|\mathbf{h}_2\|^2\right)\right). \tag{5.21}$$

Then, I_P is calculated as

$$I_P(\sqrt{\gamma}\mathbf{H}_n) = I_T(\sqrt{\gamma}\mathbf{H}_n) - I_S(\sqrt{\gamma}\mathbf{H}_n). \tag{5.22}$$

The effective rate works in a similar way to the effective SNR. If in PMod the coding rate selected for the codeword of the polarization bits was r_P, this codeword could be decoded whenever the effective rate of the channel is greater or equal than r_P.

Lastly, Fig. 5.6 intends to show all this graphically. There it is shown the variation of the magnitude of the channel matrix coefficients during a codeword and how these are summarized in five values, the effective SNR of the four modes and the effective rate of the polarization bits of PMod.

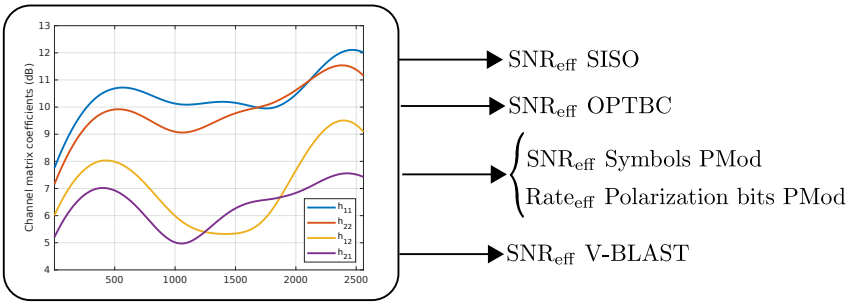


Figure 5.6: Diagram showing how the effective SNR sums up all variations of the channel coefficients during one frame.

5.5. Algorithm for MIMO Mode Selection

With regard to the adaptation of the PHY parameters, in this chapter two degrees of freedom are considered: the MIMO mode and the MCS (or coding rate, since all the MCS use a QPSK constellation). Hence, the link adaptation algorithms must perform these two tasks: the selection of the MIMO mode and the MCS for each frame. These tasks have been split between both sides of the communication, the transmitter (the gateway) and the receiver (the Mobile Terminal, MT). It is advisable that the MIMO mode selection is performed at the receiver, because, otherwise, it would require to feedback much information to the transmitter. On the other hand, we have chosen to let the transmitter select the MCS, although this selection could be done also by the receiver.

Table 5.2 collects the threshold SNR of each MCS along with the spectral efficiency achieved with that MCS depending on the MIMO mode used. Table 5.3 shows the set of coding rates which are available for the polarization bits channel encoder. Nine coding rates between 0.1 and 0.9 are possible. The spectral efficiency for SISO and OPTBC is directly $2r = log_2(4) \cdot r$, with r the coding rate. On the other hand, the spectral efficiency with V-BLAST it is $4r$, since two QPSK symbols are transmitted simultaneously per channel use. With PMod a range is specified in the table since an efficiency between 0.1 and 0.9 can be added to the modulated symbols efficiency depending on the coding rate used for the polarization bits.

A function which relates the effective SNR with the highest coding rate this SNR allows to use is depicted in Fig. 5.7. This stairs function $R(\gamma)$ is obtained from Table 5.2 and it will be used later to specify the algorithm for selecting the MIMO mode.

Table 5.2: Set of MCS with their coding rate, threshold SNR and spectral efficiency for each one of the MIMO modes.

		L8	L7	L6	L5	L4	L3	L2	L1	R
	Coding rate	0.34	0.40	0.48	0.55	0.63	0.70	0.77	0.83	0.87
	Threshold SNR (dB)	-2.15	-1.21	-0.09	0.83	1.84	2.74	3.67	4.54	5.19
Spectral efficiency (bits/s/Hz)	**SISO/OPTBC**	0.68	0.80	0.96	1.10	1.26	1.40	1.54	1.66	1.74
	PMod	0.78-1.58	0.9-1.7	1.06-1.96	1.20-2.10	1.36-2.26	1.50-2.40	1.64-1.44	1.76-2.56	1.84-2.64
	V-BLAST	1.36	1.60	1.92	2.20	2.52	2.80	3.08	3.32	3.48

Table 5.3: Set of coding rates for the polarization bits in PMod.

Coding rate	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9

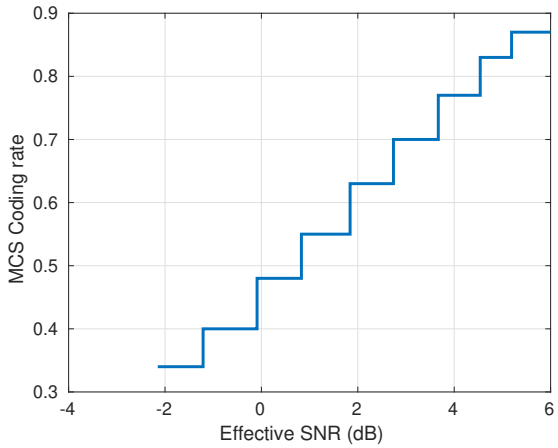


Figure 5.7: Representation of the stair function $R(\gamma)$ which gives the highest coding rate from the available MCS which can be used as a function of the effective SNR.

The steps followed by the MT to decide the MIMO mode are these:

1. We assume that the receiver knows perfectly the channel matrices for all the received symbols of a frame, i.e., we assume perfect CSI at the receiver (CSIR). In a real system a receiver would estimate the channel matrices by using some pilots spread throughout the frame.

2. The receiver must calculate the effective SNR for all the MIMO modes, SISO, OPTBC, PMod and V-BLAST, as well as the effective rate of the PMod polarization bits.

3. Next, the receiver has to calculate the maximum spectral efficiency that could be achieved with each MIMO mode, given the set of effective SNR obtained in the previous step. The function $R(\gamma)$ is used here. The spectral efficiency values for the MIMO modes are

$$s_{\text{SISO}} = 2R(\text{SNR}_{\text{eff}}\ \text{SISO}\ (\text{dB})) \tag{5.23}$$

$$s_{\text{OPTBC}} = 2R(\text{SNR}_{\text{eff}} \text{ OPTBC (dB)}) \tag{5.24}$$

$$s_{\text{PMod}} = 2R(\text{SNR}_{\text{eff}} \text{ PMod (dB)}) + \text{Rate}_{\text{eff}} \tag{5.25}$$

$$s_{\text{V-BLAST}} = 4R(\text{SNR}_{\text{eff}} \text{ V-BLAST (dB)}) \tag{5.26}$$

4. Finally, the MIMO mode which provides the highest spectral efficiency is selected by the MT and it sends this information via a feedback channel to the gateway.

Then, when the gateway receives the feedback from the users about their optimum MIMO mode, it must choose the MCS (only the coding in the proposed scenario since the constellation is fixed) to be applied to the next transmission to each user. The link adaptation algorithm to perform the MCS selection is detailed in the next section. In the feedback, apart from indicating the optimum MIMO mode, the MT sends also the effective SNR of the last frame with that MIMO mode (plus the effective rate in the case PMod is selected) and the acknowledgment of the last frame. This will be explained in detail in the next section.

5.6. Algorithm for MCS Selection

There are several options to perform the MCS selection in a system with ACM. A very popular option consists on the use of Look-up tables which map some metric related with the channel quality, as the SNR, with the set of available MCS. This table could be obtained by means of theoretical calculations or link level simulators in controlled conditions, which lead to the selection of the appropriate thresholds which define the LUT [77]. If the link level simulations are representative of the scenario where the link adaptation protocol is going to operate, then the LUT-based approach will offer a near-optimal performance. However, the inaccuracy of the CSIT due to estimation errors and the long propagation delays of satellite links, which cause the CSIT to be outdated, can degrade the performance of a pure LUT-approach. Moreover, if the working channel statistics differ from those simulated, the threshold SNRs obtained in the simulations could not work in practice. For all these reasons, it is very common that an adaptive backoff margin is applied to the SNR before consulting the LUT to select the MCS [76], [78].

Adaptive margins have not only been proposed in the literature but also several experimental campaigns demonstrated their efficacy. For example,

in [93] we have shown the results of some field trials of link adaptation algorithms in a satellite communications system with a MEO satellite and a MT boarded in both an unmanned aerial vehicle (UAV) and a car. The adaptive margins employed for performing the MCS selection allowed the system to work correctly in the practical setup guaranteeing the reliability of the communications, given by a predefined target Frame Error Rate (FER) p_0.

The solution proposed in this chapter for selecting the MCS is based on the use of a LUT along with an adaptive margin. For SISO, OPTBC and V-BLAST there is a single LUT whereas for PMod there are two LUTs in parallel since, as depicted in Fig. 5.2, two channel encoders with independent coding rates are used for the polarization and symbol bits. Fig. 5.8 has a block diagram for the MCS selection for SISO, OPTBC, V-BLAST and for the symbols coding rate of PMod, and Fig. 5.9 does the same with the coding rate selection for the polarization bits channel encoder of PMod.

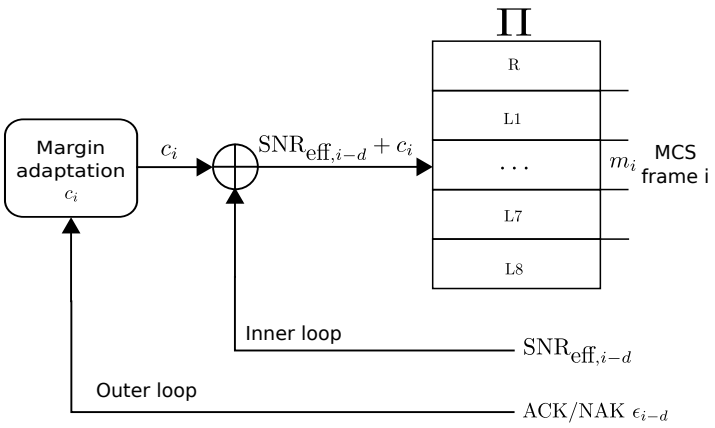


Figure 5.8: MCS selection with a Look-up table and an adaptive margin.

In both Fig. 5.8 and 5.9 there is a LUT, represented by the function $\Pi(\cdot)$ and $\Pi_P(\cdot)$, which maps SNR to MCS or coding rate to an available coding rate, according to Tables 5.2 and 5.3. The inputs of the LUT are obtained by adding a margin, c_i or $c_{P,i}$, to the effective SNR or the effective rate, as can be seen in the diagrams. In this type of adaptation two parts can be differentiated. In the inner loop the CSI fed back by the receiver, in our case the effective SNR for the selected MIMO mode (and also the effective rate if PMod was selected), is used along with the margins for obtaining a valid MCS with the LUT. On the other hand, the outer loop is in charge of adapting the margins with the acknowledgments (ϵ_i and $\epsilon_{P,i}$) received from the other end. Finally, the output of the LUT is m_i, the MCS selected for the i-th transmission frame and r_i, the coding rate selected for PMod polarization bits channel encoder. The integer d subtracted to the variable

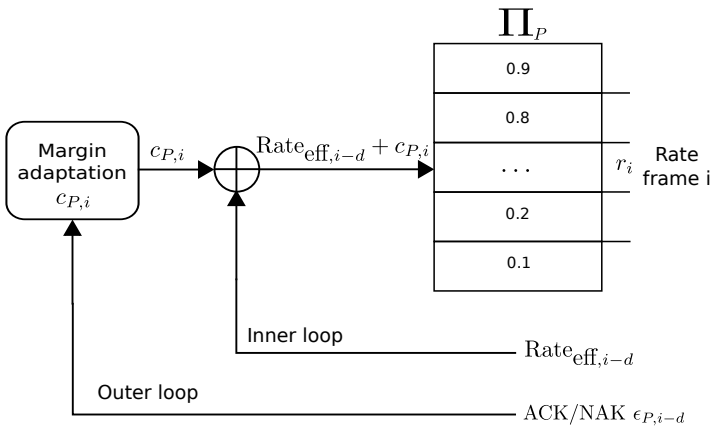


Figure 5.9: Coding rate selection for PMod polarization bits with a Look-up table and an adaptive margin.

i in the variables of the diagram represents a delay expressed as an integer number of frames, to model the RTT of the satellite link.

Mathematically, the selection of the MCS in SISO, OPTBC, V-BLAST and for the symbols MCS of PMod is expressed as

$$m_i = \Pi(\text{SNR}_{\text{eff},i-d} + c_i) \tag{5.27}$$

and the selection of the coding rate for the PMod polarization bits channel encoder as

$$r_i = \Pi_P(\text{Rate}_{\text{eff},i-d} + c_{P,i}). \tag{5.28}$$

Here $\Pi(\cdot)$ and $\Pi_P(\cdot)$ are the two LUTs, $\text{SNR}_{\text{eff},i-d}$ is the effective SNR for the particular MIMO mode under use, $\text{Rate}_{\text{eff},i-d}$ is the effective rate of PMod polarization bits (when PMod MIMO mode is selected), d corresponds to the number of frames that a RTT lasts, c_i and $c_{P,i}$ are the two margins, m_i is the MCS selected and r_i the coding rate for the PMod additional channel encoder.

We define ϵ_i as the error event of the i-th frame, then ϵ_i will be equal to 1 if a decoding error occurs and 0 otherwise. In the case of PMod, since each frame corresponds to two codewords, one per channel encoder, an additional variable $\epsilon_{P,i}$ is defined, as the error event of the output codeword of the polarization bits channel encoder. For PMod, ϵ_i represents the error event of the codeword of the symbols bits channel encoder. These variables ϵ_i and $\epsilon_{P,i}$ are fed back by the receiver.

The link adaptation algorithm detailed in the following paragraphs intends to maximize the spectral efficiency while maintaining the FER around

a low target value p_0. It will select the highest MCS, the most efficient, which allows the quality of service specified by that objective FER p_0. In the case of PMod we consider a frame is received correctly when both codewords are decoded. Therefore, only when ϵ_i and $\epsilon_{P,i}$ are equal to 0 that PMod frame i is counted as successful.

Hereafter, the process to obtain the adaptation rule of the margins will be detailed. Although only the derivation of the margin c_i for SISO, OPTBC and V-BLAST will be provided, the derivation of the two margins c_i and $c_{P,i}$ for PMod is very similar, simply using $p_0/2$ instead of p_0 in the cost function since, for low objective FER p_0

$$P(\text{Erroneous frame}) = 1-P(\epsilon_i = 0 \,\&\, \epsilon_{P,i} = 0) = 1-(1-p_0/2)^2 \simeq p_0 \quad (5.29)$$

In an effort to obtain a flexible solution not requiring specific knowledge about the channel, the margin can be obtained as the solution of an optimization problem for a given FER p_0, as is explained in [77]:

$$\min_c J(c) = \min_c \left|\mathbb{E}[\epsilon] - p_0\right|^2 . \quad (5.30)$$

The choice of a good p_0 is not trivial, since it will have an impact on the final throughput, which will also depend on the existence of retransmission mechanisms. In practical cases $p_0 = 0.1$ is acknowledged to be a good reference [105], although this is not necessarily the case in the setting under study. If we apply a gradient descent operation to solve (5.30), then

$$c_{i+1} = c_i - \mu_i \frac{\partial J}{\partial c}(c_i) \quad (5.31)$$

with μ_k a sequence of positive numbers. The value of the derivative is

$$\frac{\partial J}{\partial c} = \frac{\partial \mathbb{E}[\epsilon]}{\partial c}\left(\mathbb{E}[\epsilon] - p_0\right) \quad (5.32)$$

so the gradient descent is

$$c_{i+1} = c_i - \mu_i \frac{\partial \mathbb{E}\left[\epsilon\right]}{\partial c}\left(\mathbb{E}\left[\epsilon\right] - p_0\right). \quad (5.33)$$

Note that (5.33) depends on the function $\mathbb{E}\left[\epsilon\right]$ and, therefore, cannot be obtained without statistical knowledge of the channel. We propose two modifications to (5.33). First, we substitute $\mathbb{E}\left[\epsilon\right]$ by the error ϵ_{i-d}, available at the transmitter at time instant i, so we transform the gradient descent in a stochastic gradient descent. Second, we interpret the term $\frac{\partial \mathbb{E}[\epsilon]}{\partial c}$ as a positive[1] time varying adaptation weight, so we include its effect in μ_i. The

[1]Note that rigorously speaking $\mathbb{E}\left[\epsilon\right]$ is a non-continuous function.

derivative is positive because an increment of the margin, as it is defined, will cause the selection of a higher MCS, whose error probability is higher than the more robust MCS for the same channel conditions. Therefore, the stochastic adaptation rule is

$$c_{i+1} = c_i - \mu_i \left(\epsilon_{i-d} - p_0\right). \tag{5.34}$$

Let us denote by $p\left(c\right)$ the error probability when the margin is c. Assume that p is a continuously differentiable function, with derivative bounded by $\delta_0 < \frac{\partial}{\partial c} p\left(c\right) < \delta_1 \, \forall c$, with $\delta_0 > 0$. Let us denote by $c_\star$ the value[2] such that $p\left(c_\star\right) = p_0$. Additionally, the stepsize in (5.34) will remain constant and equal to μ. If $d = 0$ in (5.34), then the following convergence results hold.

Theorem 1. *If $\mu < 2/\delta_1$, then $\left|\mathbb{E}\left[c_i\right] - c_\star\right| < \eta^i \left|\mathbb{E}\left[c_0\right] - c_\star\right|$, with $0 < \eta < 1$, which grants an exponential convergence of (5.34). In addition, the expectation of the mean squared error $\mathbb{E}\left[\left(c_i - c_\star\right)^2\right]$ converges as*

$$\lim_{i \to \infty} \mathbb{E}\left[\left(c_i - c_\star\right)^2\right] < \frac{\mu}{2\delta_0 - \mu\delta_1}. \tag{5.35}$$

For the proof, see the Appendix. If $d > 1$ convergence conditions can still be obtained although derivations are more involved.

Lastly, the final adaptation rule for updating the margin c_i for SISO, OPTBC and V-BLAST is

$$\boxed{c_{i+1} = c_i - \mu \left(\epsilon_{i-d} - p_0\right)} \tag{5.36}$$

and for updating the two margins in PMod MIMO mode is

$$\boxed{c_{i+1} = c_i - \mu \left(\epsilon_{i-d} - p_0/2\right)} \tag{5.37}$$

$$\boxed{c_{P,i+1} = c_{P,i} - \mu' \left(\epsilon_{P,i-d} - p_0/2\right)} \tag{5.38}$$

Note that in PMod we allow the adaptation steps μ and μ' to be different, since the magnitude of the effective rate is smaller than that of the effective SNR.

The behavior of equation (5.36) can be summed up in the following two rules:

- When an ACK is received ($\epsilon_i = 0$), the margin is increased slightly by a quantity $\Delta_{\text{ACK}} = \mu p_0$. On the long term, after many ACKs, it will trigger the selection of more efficient MCS.

[2] As p_0 is a monotonic increasing function, this value is unique.

- When a NAK is received due to the fact that there was an error event ($\epsilon_i = 1$), the margin is decreased more dramatically by a quantity $\Delta_{\text{NAK}} = \mu(1 - p_0)$, that for small values of the target FER p_0, as it is usually the case, is approximately equal to μ.

The ratio between the margin increment and decrement following an ACK and a NAK, respectively, depends on the target FER and it is equal to $\Delta_{\text{ACK}}/\Delta_{\text{NAK}} = p_0/(1-p_0)$. This guarantees that when the margin converges to the optimum value which ensures the target FER, it remains oscillating around that value.

5.7. Simulation Results

All the simulation results presented throughout this section are obtained for a maritime scenario with a vessel moving at a constant speed of 50 km/h. The parameters of the channel generator of the DP mobile satellite system are taken from [88] and they were specified in Section 5.3. The carrier frequency in the simulations is $f_c = 1.6$ GHz, in the middle of the L-band. Frames of length 76.19 ms and $N = 2560$ symbols are used in the physical layer, similarly to the bearer F80T1Q-1B of ETSI TS 102-744 [9]. Lastly, the RTT is set to $d = 7$ frames (533 ms) to emulate the feedback delay for geostationary (GEO) satellites. Table 5.4 collects the main parameters, common to all the simulation results provided in this section.

Parameter	**Value**
Satellite orbit	GEO
Polarization	Dual (RHCP and LHCP)
Frequency	1.6 GHz (L-band)
Modulation	QPSK
Frame length	2560 symbols
Frame duration	76.19 ms
RTT	7 frames (533 ms)
Channel	Maritime (50 km/h)

Table 5.4: System parameters.

5.7.1. Mutual Information of PMod in a DP Maritime Channel

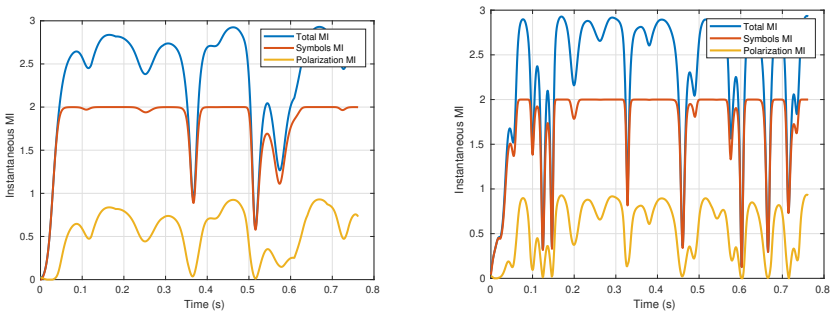
Previously to the simulation of the link adaptation algorithm to select the MIMO mode and MCS, this section contains some graphics related to

the Mutual Information (MI) of PMod in this type of channel. The total MI I_T can be expressed as the sum of two MIs, the polarization MI I_P and the symbols MI conditioned to a known polarization I_S:

$$I_T = I_P + I_S. \tag{5.39}$$

The total MI can be calculated by using some of the neural networks proposed in the previous chapter, the symbols MI I_S can be easily obtained using equation (5.21), and I_P is obtained from the other two.

Firstly, in Fig. 5.10 the evolution of these three MI values is shown during a time spanning 10 frames for two different vessels speeds, 25 and 50 km/h and for a SNR γ of 7 dB. It can be seen how these MI values can present some oscillations during a frame, revealing the need of using some metric, as the effective SNR or the effective rate which compress all this variation in a single value useful to predict the decoding outcome. The total MI is bounded by three bits per channel use (bpcu), the sum of the 2 bpcu carried by the QPSK symbols and the additional bpcu of the polarization selection.



(a) 25 km/h (b) 50 km/h

Figure 5.10: Evolution of the instantaneous Mutual Information of PMod for a DP maritime channel during 10 frames with an average SNR of 7 dB.

Next, the problem of obtaining the ergodic capacity of PMod in this type of channel was addressed. 384,000 realizations of the channel matrix were generated (which correspond to 150 frames) for each value of SNR and the three MIs, I_T, I_S and I_P were calculated and its average obtained. The result is plot in Fig. 5.11 where the three ergodic MI are shown as a function of the SNR. From that graphic it can be concluded that we do not gain one bit of capacity directly with PMod, this only occurs for high SNRs.

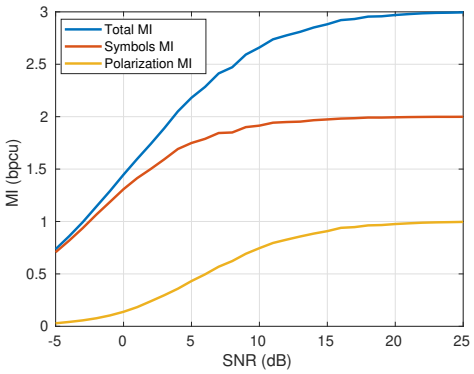


Figure 5.11: Ergodic MI (total, symbols bits and polarization bits) of PMod for a 50 km/h DP maritime channel.

5.7.2. MIMO Mode Selection

The purpose of this section is showing how the MIMO mode selection works in practice, providing results of the MIMO mode which is chosen as a function of the SNR. For a number of SNRs between -5 and 25 dB a simulation was run where the algorithm detailed in Section 5.5 selected the MIMO mode for each one of the $1,000$ frames of the simulation. Afterwards, the frequency of use of the modes per SNR was obtained and its percentages are shown in Fig. 5.12.

In Fig. 5.12 a clear pattern can be observed, where OPTBC was the preferable mode for low SNRs (-5 to -2 dB), PMod for intermediate SNRs (-1 to 8 dB) and V-BLAST for SNRs higher than 9 dB. Since the criterion for choosing the MIMO mode is the spectral efficiency, Fig. 5.12 also indicates which MIMO mode provides the highest spectral efficiency in each SNR range. The inclusion of PMod among the MIMO modes provides an advantage as this figure shows, due to the fact that it can improve the spectral efficiency at intermediate SNRs more than OPTBC or V-BLAST.

5.7.3. Link Adaptation Algorithm Results

This subsection provides results of the average spectral efficiency and the average FER achieved with the two link adaptation algorithms which select the MIMO mode and the MCS. In the simulations the results for each point of SNR are obtained emulating the transmission and reception of $M = 40,000$ frames for a fixed average SNR γ. The average spectral

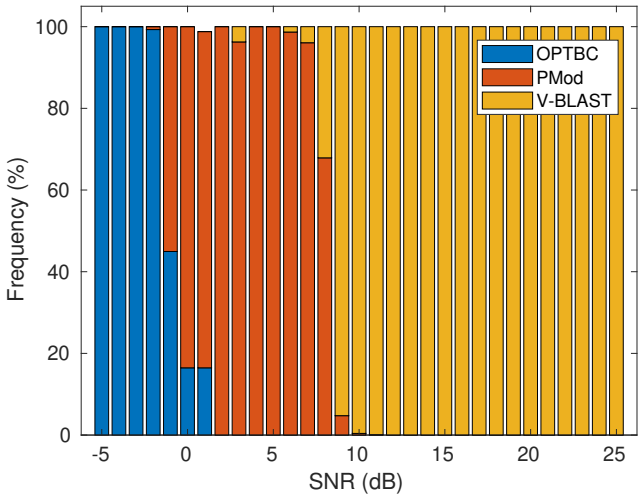


Figure 5.12: Fraction of time where each MIMO mode is the one which provides the highest spectral efficiency as function of the SNR.

efficiency s_{av} is calculated as follows

$$s_{av} = \frac{1}{M}\sum_{i=1}^{M}(1-\epsilon_i)s_{m_i} + (1-\epsilon_{P,i})r_i \quad \text{bits/s/Hz.} \tag{5.40}$$

In the previous expression the second summand it is added only when PMod is selected for frame i, in this case r_i represents the coding rate selected in the polarization bits channel encoder for the i-th frame. On the other hand, s_{m_i} is the spectral efficiency of the MCS m_i for the MIMO mode selected for that frame, which is tabulated in Table 5.2.

The average FER of the simulations is calculated by simple averaging ϵ_i:

$$\text{FER}_{av} = \frac{1}{M}\sum_{i=1}^{M}\epsilon_i \tag{5.41}$$

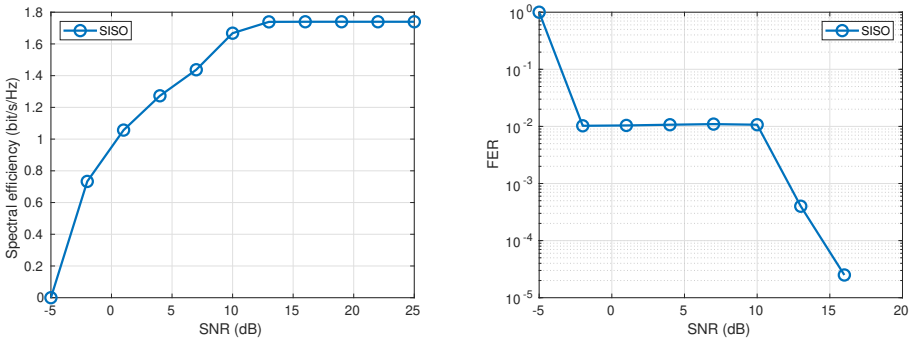
If a frame is transmitted with the PMod MIMO mode, only when both codewords are correctly decoded this frame counts as successful, i.e., when $\epsilon_i = 0$ and $\epsilon_{P,i} = 0$.

In the equations (5.36) and (5.38) for updating the margins the value of the adaptation steps used in the simulations are $\mu = 0.05$ and $\mu' = 0.005$. The step for updating the margin of the coding rate of PMod is smaller since its range is also smaller compared with the range of variation of the effective SNR. The objective FER of all the simulations is $p_0 = 0.01$.

Figures 5.13-5.16 show the average spectral efficiency and the FER as a

function of the SNR for the four MIMO modes, SISO, OPTBC, PMod and V-BLAST, respectively. In these simulations only one of the four modes is allowed in each case, in order to see their individual performance. The MCS is selected with the algorithm introduced in the previous section. It can be observed how the spectral efficiency grows with the SNR until the maximum value is achieved for each MIMO mode (1.74 with SISO and OPTBC, 2.64 with PMod and 3.48 with V-BLAST). The value of SNR at which the spectral efficiency saturates to its maximum value is also different in each mode.

The plots of the FER show that the link adaptation algorithm is able to satisfy the objective FER requirement of 0.01 for a large range of operation points. The constraint is not met when the SNR is very low (because even the most robust MCS does not allow a robust transmission) or when the SNR is very high and the spectral efficiency saturates (because the channel is so good that no errors happen and no higher MCS can be selected).

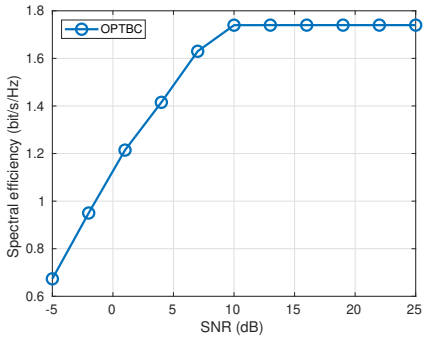


(a) Spectral efficiency (b) FER

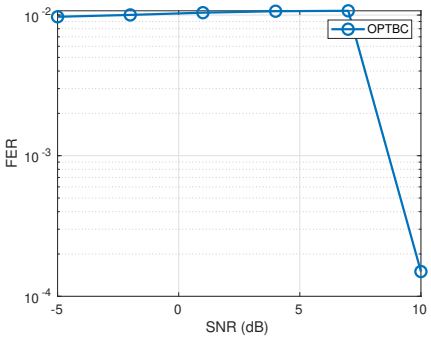
Figure 5.13: Average spectral efficiency and average FER for a DP system which exploits only one polarization (SISO).

Fig. 5.17a plots the average spectral efficiency for three sets of simulations over a range of average SNRs from -5 to 25 dB. In the first simulation only the SISO mode is allowed, in the second the available modes are OPTBC and V-BLAST, whereas in the third simulation PMod is added to the other two MIMO modes. The MIMO mode and the MCS can change from frame to frame. The selection of the MIMO mode is performed as explained in Section 5.5 and the selection of the MCS is done with the algorithm of Section 5.6.

Firstly, the comparison between single polarization (SISO curve) and DP with the three modes in Fig. 5.17a shows that gains up to +100% can be achieved for high SNRs, i.e., the throughput can be doubled. Moreover,

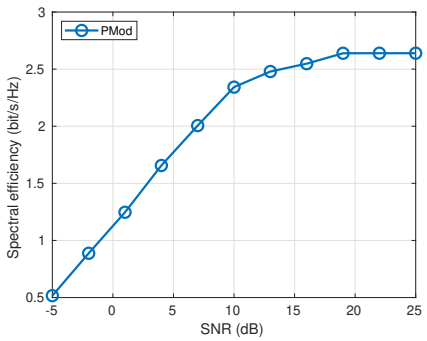


(a) Spectral efficiency

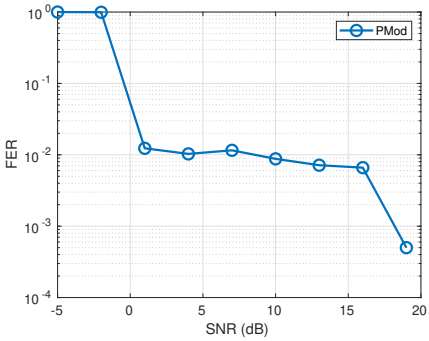


(b) FER

Figure 5.14: Average spectral efficiency and average FER for a DP system which exploits only OPTBC.



(a) Spectral efficiency

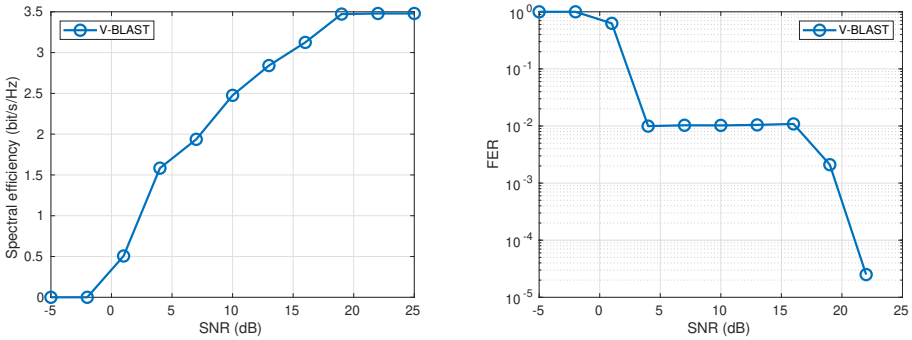


(b) FER

Figure 5.15: Average spectral efficiency and average FER for a DP system which exploits only PMod.

DP allows to extend the operating range for low SNRs. For example, with SISO no communication can be held at a SNR of -5 dB whilst, with DP the OPTBC mode is capable of offering a spectral efficiency of 0.7 bits/s/Hz and meeting the FER constraint. With regard to the low efficiency at 1 dB when the two modes OPTBC and V-BLAST are combined, it is caused by the abrupt switch between the two modes, which leads to some throughput loss. However, when the three MIMO modes are used there is a gain in spectral efficiency for all the SNRs, being the gain typically higher than +30%. This gain could be employed to serve more users or to increase the data rate of the individual users.

Note that the same constellation −QPSK− has been used for both SISO and DP configurations. In order to match the performance of DP, the SISO



(a) Spectral efficiency (b) FER

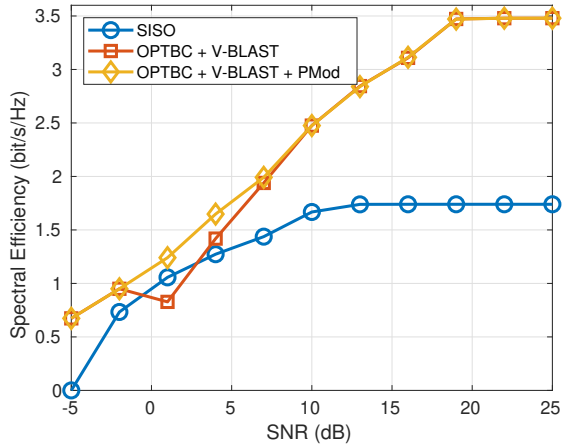
Figure 5.16: Average spectral efficiency and average FER for a DP system which exploits only V-BLAST.

link would require higher order constellations. With 8-PSK, the SISO link would perform as DP-PMod with QPSK at high SNRs, and SISO would require a 16-QAM constellation to match the performance of DP with V-BLAST in the high SNR points.

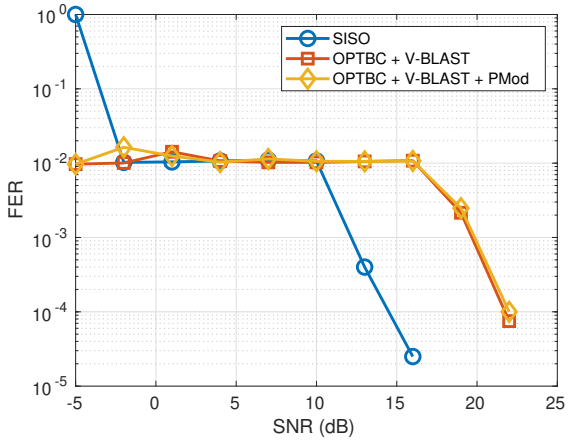
Secondly, SISO is compared with other two DP systems where several MIMO modes are available. In some simulations the modes are only OPTBC and V-BLAST, whereas in other cases these two modes are complemented with PMod. From Fig. 5.17a it is clear that the inclusion of PMod among the MIMO modes of the DP system is worthwhile. For a range of SNRs from -2 to 10 dB, PMod provides a spectral efficiency gain with respect to a configuration where only OPTBC and V-BLAST are used. Thirdly, Fig. 5.17b confirms that the use of the adaptive margin guarantees a FER around the objective value ($p_0 = 0.01$ in this case) for a wide range of SNRs and for all the MIMO modes considered here.

These simulations show the benefits of using two orthogonal polarizations simultaneously to serve a mobile user in low bands of the spectrum. Despite the XPD being lower in these bands than in Ku and Ka bands, it provides a significant throughput gain even with the available power split between the two polarizations, as it is the case.

In this work a novel approach was used for doing the link adaptation with PMod. Instead of using a single channel encoder and then split the bits between the symbols and polarization selection, two independent coding rates are used, one for the stream of bits selecting the polarizations and another for the stream which selects the QPSK symbols. In Fig. 5.18, the most used coding rate for the symbol bits channel encoder and for the



(a) Spectral efficiency



(b) FER

Figure 5.17: Comparison of the average spectral efficiency and average frame error rate (FER) for a SP system (SISO) and a DP system with and without PMod.

polarization bits channel encoder is shown as a function of the average SNR of the channel γ in a PMod simulation. As expected, for lower SNRs more protected MCS are selected, i.e., with lower coding rates, and for higher SNRs more efficient MCS are chosen. It can be seen how the coding rates required for meeting the FER constraint at a particular SNR are different for the two channel encoders, since symbol bits and polarization bits required different levels of protection. All this shows the adequacy of using two independent variable rate channel encoders, one for the stream of bits which selects the constellation symbols and another for the stream of bits which

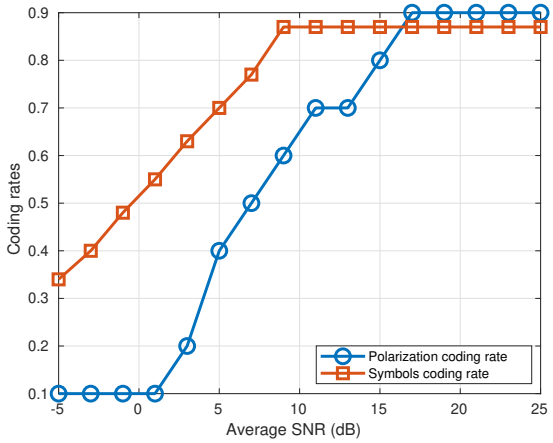
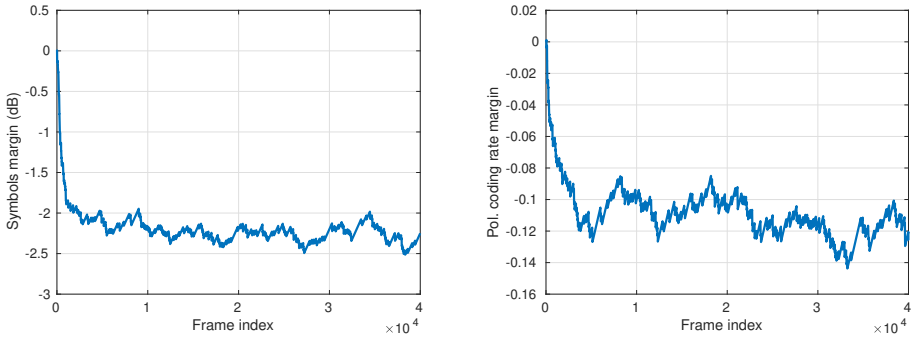


Figure 5.18: Most selected coding rates for polarization and symbol bits channel encoders for a 50 km/h DP maritime channel with adaptive PMod with separate coding.

selects the polarizations hops.

Lastly, in Fig. 5.19 an example of the evolution of the margins is shown. They correspond to the time evolution of the two margins at a SNR of 7 dB when the PMod MIMO mode is selected. Fig. 5.19a plots the evolution of the symbols margin c_i and Fig. 5.19b does the same with the coding rate margin $c_{P,i}$. It can be seen how after some time the margins converge to a value and oscillate around it.



(a) Symbols (b) Polarizations

Figure 5.19: Evolution of the margins used in the coding rates selection for symbols and polarization, for a 50 km/h DP maritime channel with adaptive PMod with separate coding.

5.8. Conclusions

In this chapter a link adaptation algorithm was proposed for a Dual Polarization (DP) mobile satellite system which can alternate among different MIMO modes, namely OPTBC (a polarization-time block code analogous to Alamouti space-time coding), PMod (equivalent to Spatial Modulation but using the domain of the polarization), and V-BLAST (a spatial multiplexing technique but applied here to streams transmitted using different polarizations). Despite the interference between polarizations, gains in the spectral efficiency between +30% and +100% can be achieved. Hence, the capacity can be doubled for high SNRs. In addition, the system can operate at lower SNRs compared with Single Polarization thanks to OPTBC.

The adaptive DP system can change the MIMO mode and the Modulation and Coding Scheme (MCS) from frame to frame. The selection of the mode is based on the maximization of the spectral efficiency given the current channel conditions whereas the MCS selection is based on the use of Lookup-Tables with some adaptive margins, designed for fixing the Frame Error Rate (FER) to a predefined value.

A novel architecture for a Polarized Modulation transmitter was introduced, this is based on two independent channel encoders, one for bits encoding the symbols and another for bits encoding the polarizations. Thanks to the methods for calculating the Mutual Information of PMod introduced in the previous chapter, an optimized coding rate can be used in each channel encoder to match their respective channel capacities.

The simulations of this chapter are obtained using PLA techniques, based on the use of the effective SNR and the effective rate to avoid running all the transmission and reception chains. In the following chapter the adaptation of a general Spatial Modulation system will be addressed, with applications to both terrestrial systems, namely future multiantenna 5G mobile networks, and DP mobile satellite systems. The approach of the next chapter is quite different. On the one hand, instead of PLA all the decoding chain is executed at the receiver, and, on the other hand, the selection of the MCS is done by using a trained neural network. Therefore, the next chapter follows the path of Chapter 4, where Machine Learning was used to compute the Mutual Information and the Capacity of Spatial Modulation and Generalised Spatial Modulation systems. A slightly more complex neural network is employed to go beyond the capacity calculation and predict the optimum MCS constrained to a given receiver performance.

5.9. Appendix. Convergence Analysis

We have performed a convergence analysis of the adaptive outer loop to certify the mean and MSE convergence of the margin c_i in (5.34) for $d = 1$. The proof is based on the results in [26]. We can write the expected value of c_{i+1} conditioned on c_i as

$$\mathbb{E}\left[c_{i+1}|c_i\right] = c_i - \mu\left(p\left(c_i\right) - p_0\right). \tag{5.42}$$

Note that, by definition, $p_0 = p\left(c_\star\right)$. By using the mean value theorem, we can find $\tilde{c}$ between c_i and $c_\star$ such that $p\left(c_i\right) - p_0 = p'\left(\tilde{c}\right)\left(c_i - c_\star\right)$. Therefore,

$$\mathbb{E}\left[c_{i+1}|c_i\right] - c_\star = \left(c_i - c_\star\right)\left(1 - \mu p'\left(\tilde{c}\right)\right). \tag{5.43}$$

If we take the expectation over c_i on both sides of the equation, we have that

$$\left|\mathbb{E}\left[c_{i+1}\right] - c_\star\right| = \left|\mathbb{E}\left[c_i\right] - c_\star\right|\left|1 - \mu p'\left(\tilde{c}\right)\right|. \tag{5.44}$$

Note that the new $\tilde{c}$ is not necessarily the same as before, after applying the first mean value theorem for integration. Convergence is guaranteed if $\left|1 - \mu p'\left(\tilde{c}\right)\right| < 1$ or, equivalently, if $\mu < 2/p'\left(\tilde{c}\right)$. Since the value of the derivative is upper bounded by δ_1, it suffices to choose $\mu < 2/\delta_1$. In such a case,

$$\left|\mathbb{E}\left[c_i\right] - c_\star\right| < \eta^i\left|\mathbb{E}\left[c_0\right] - c_\star\right| \tag{5.45}$$

with $\eta = |1 - \mu\delta_1|$. Thus, exponential convergence of the expectation $\mathbb{E}\left[c_i\right]$ has been proved.

Next we perform a second order analysis. From (5.34) we have

$$\begin{aligned}\mathbb{E}\left[\left(c_{i+1} - c_\star\right)^2\right] = \mathbb{E}\left[\left(c_i - c_\star\right)^2\right] - \\ 2\mu\mathbb{E}\left[\left(c_i - c_\star\right)\left(\hat{p}(c_i) - p_0\right)\right] + \mu^2\mathbb{E}\left[\left(\hat{p}(c_i) - p_0\right)^2\right]\end{aligned} \tag{5.46}$$

where we are referring to ϵ_{i-d} as $\hat{p}(c_i)$. If we write $\hat{p}(c_i) = p(c_i) + w_i$, with w_i a zero-mean random variable, then we have that $\mathbb{E}\left[w_i^2\right] < 1$. Note that we can accommodate more precise estimates of the probability of error in (5.34) by, for example, averaging a sequence of consecutive ACK/NAK; the bound for $\mathbb{E}\left[w_i^2\right]$ could be tightened accordingly. Next we will use also that $\mathbb{E}\left[\left(c_i - c_\star\right)\left(p(c_i) - p_0\right)\right] = p'(\tilde{c})\mathbb{E}\left[\left(c_i - c_\star\right)^2\right]$ after applying the mean value theorem and the first mean value theorem for integration. Analogously, $\mathbb{E}\left[\left(p(c_i) - p_0\right)^2\right] = \left(p'(\bar{c})\right)^2\mathbb{E}\left[\left(c_i - c_\star\right)^2\right]$ for another value $\bar{c}$. With this, and given that $\delta_0 \leq p'(c) \leq \delta_1$, we have that

$$\mathbb{E}\left[\left(c_{i+1} - c_\star\right)^2\right] \quad \leq \quad (1 \; - \; 2\mu\delta_0 \; + \; \mu^2\delta_1^2)\mathbb{E}\left[\left(c_i - c_\star\right)^2\right] \; + \; \mu^2 \tag{5.47}$$

which, in the limit leads to

$$\lim_{i\to\infty} \mathbb{E}\left[(c_i - c_\star)^2\right] < \frac{\mu}{2\delta_0 - \mu\delta_1}. \tag{5.48}$$

Finally, convergence of average BLER $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \epsilon_i$ to the target BLER p_0 can also be proved following a similar reasoning as in [26].

Chapter 6

Deep Learning Assisted Rate Adaptation in Spatial Modulation Links

This chapter is adapted, with permission of the coauthors and the editorial, from: "A. Tato, C. Mosquera. Deep Learning Assisted Rate Adaptation in Spatial Modulation Links. In 16th International Symposium on Wireless Communications Systems (ISWCS), Oulu, Finland, August, 2019", [92].

6.1. Introduction

Spatial Modulation (SM) is being considered for future 5G systems [20], since it can increase the spectral efficiency with respect to single antenna systems, with simpler hardware requirements as compared with other multi-antenna techniques, reducing the power consumption. Its most basic implementation activates only one antenna at a time, with the information encoded into the index of the active antenna and the transmit symbol. This seemingly simple transmission scheme poses some challenges when designing the receiver or computing the achievable information rates.

Moreover, an analogous modulation scheme, named Polarized Modulation (PMod), can be applied to Dual Polarization (DP) mobile satellite links, as it was explained previously in Chapter 5. Mathematically, both SM and PMod are equivalent schemes. The only difference is that in 2×2 SM the dimensions are interpreted in terms of different antennas, whilst in PMod the dimensions represent the signals transmitted or received through different

polarizations of the antennas. As PMod is one of the modulation schemes which can be used in future DP mobile satellite systems, the results of this chapter can find applications in both future terrestrial and satellite systems. Although hereafter we are going to talk only about SM, the reader should bear in mind that the adaptation procedure explained in this chapter is also valid for PMod.

Almost any practical communication system employs link adaptation in order to exploit better the time varying channel capacity. Adaptive Coding and Modulation (ACM) or link adaptation is a technique which consists typically on varying the modulation order and/or the coding rate of the encoder to track the changing channel conditions. The ultimate goal is to adjust the transmitted bit rate as much as possible to the information that the channel can support for a given bit error probability.

Link adaptation makes it necessary for the transmitter-receiver pair to estimate somehow the rate that can be supported by the channel. In most practical cases, the receiver computes some metric which is sent back to the transmitter end. This metric can be in the form of the average or effective Signal to Interference and Noise Ratio (SINR), or some Channel Quality Indicator (CQI) specifically suited to the set of Modulation and Coding Schemes (MCS) available to the transmitter [83].

The authors presented in [49] some analytical approximations to the integral expression of the Mutual Information (MI) in an SM link. Its use in adaptive settings is jeopardized by the error of the approximations, their computational complexity and also by the need to estimate the achievable rate of practical MCSs. In Chapter 4 we have proposed to calculate the MI of a SM system by means of a very simple Multilayer Feedforward Neural Network (MFNN) of just one hidden layer. This method improves the accuracy of the previous analytical approximations given in [49] and, furthermore, the MI calculation requires a much lower computational complexity. This accurate and rapid calculation of the MI with a neural network allows a receiver to estimate the MI of the link in real time and select the MCS of the frames the transmitter sends him. This MCS selection based on the calculation the MI will be explored later in this chapter. As it will be shown, since the codes used by the receiver do not achieve the capacity of the channel, a back-off margin has to be subtracted from the estimated MI previous to the MCS selection in order to have an acceptable outage probability.

This chapter aims to design an adaptive SM system, which can be employed in future terrestrial and satellite communications. In this work we explore the use of Deep Learning (DL) tools to determine at each moment the appropriate MCS to transmit. The use of DL or Machine Learning

(ML) at the physical layer of communication systems is gaining momentum, as recent surveys illustrate [11, 89]. In [27] supervised learning is used to perform the adaptation in MIMO-OFDM (Multiple-Input-Multiple-Output Orthogonal Frequency Division Mutiplexing) systems and in [32] autoencoders together with Support Vector Machines (SVM) are applied to do the MCS selection in MIMO systems. SVM have been also studied for the selection of physical layer parameters in communication settings with a large number of degrees of freedom [79]. In particular, Neural Networks (NN) have been successfully used for channel estimation and equalization [64], signal recognition and modulation classification [34],[63], detection in MIMO Generalized SM [59], and learning of physical layer parameters in Cognitive Radio [37], among others. In [54], and [72] NNs are applied to perform link adaptation in multicarrier systems.

Some publications more related to this work are [80], where the authors use a NN to make codebook selection in SM systems, and [81], where DL is used to estimate the Frame Error Probabilities and to perform coding rate selection in Bit Interleaved Coded Modulation BICM-OFDM systems. The codebook selection on SM consist on deciding which antennas are used and which constellation is used in each antenna. However, in that work, the coding rate is fixed and only the codebook is adapted. Other articles, like [108] and [111], only consider the adaptation of the modulation in SM links, too. This requires the use of fixed and very low coding rate for allowing a robust communication in a wide range of channel conditions. In this chapter, we go further and a finer adaptation is done, adapting also the coding rate in order to enhance the spectral efficiency of the system. Therefore, the coding rate is not a low fixed value designed for the worst channel conditions, and it is adapted dynamically according to the instantaneous channel capacity. The advantage of our approach is that a fine tuning of the coding rate can exploit better the channel capacity over the whole SNR range of a given constellation.

The current work uses a MFNN of three hidden layers to assist the coding rate selection at the receive end. The network employs some specifically selected input features which can be easily obtained from the MIMO channel matrix, together with the Signal to Noise Ratio (SNR). The choice of the input features is based on our work [98] of SM MI calculation with MFNNs. Contrary to the previous chapter, here we use some actual channel codes and we do not make the assumption of having capacity-achieving codes. The neural network is trained off-line with data of many system level simulations of the SM link with different channel matrices. In this way, the neural network learns to select the optimum coding rate for that particular system as a function of the SNR and the channel matrix.

In this chapter, a general procedure for doing coding rate selection in SM systems with the help of DL is explained. Then, this is exemplified for the particular case of a 2×2 SM system with a Quadrature Phase Shift Keying (QPSK) constellation. However, the same procedure can be employed to more complex SM systems by simply performing the pertinent system level simulations and training the neural network with data of the performance of the channel codes in these new systems.

Several neural network architectures are compared and its performance in the coding rate selection task is shown with great detail. In the last part of the chapter, the performance in terms of spectral efficiency and outage probability is calculated for different adaptation strategies, covering fixed coding rate, MI-based coding rate selection and DL-based coding rate selection. The later, the approach proposed in this chapter, clearly outperforms all the other solutions. Both spectral efficiency and outage probability are improved in the DL-based selection when compared with the other approaches.

This chapter is organized in the following way. After the introduction, Section 6.2 provides the general system model and introduces the problem of the coding rate selection. Then, Section 6.3 details the specification of a real 2×2 SM system and Section 6.4 shows the performance of the SM receiver under different channel conditions. Afterwards, Section 6.5 explains all the steps required for doing a DL-based coding rate selection in SM before Section 6.6 finishes the specification of the simulated system. Lastly, Section 6.7 contains the main simulation results of the classification performance obtained with the MFNN for selecting the coding rate. It also provides graphics with the system level performance, in terms of spectral efficiency and outage probability, of different adaptive and non adaptive SM systems. At the end, Section 6.8 collects the main conclusions and signals the future work.

6.2. General System Model

We consider a Spatial Modulation (SM) system, where information is not only conveyed by the selection of a symbol from a constellation $\mathcal{S}$, but also by the antenna chosen for sending that symbol. The system model equation of a SM link with N_t transmit antennas and N_r receive antennas for a given discrete time instant is

$$\mathbf{y} = \sqrt{\gamma}\mathbf{H}\mathbf{x} + \mathbf{w} \tag{6.1}$$

where $\mathbf{y} \in \mathbb{C}^{N_r \times 1}$ is the received vector, γ the average Signal to Noise Ratio (SNR), $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ the channel matrix, $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$ the transmitted signal and $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_r})$ the Additive White Gaussian (AWGN) noise vector.

Due to the specific nature of SM (activation of only one antenna per symbol period), $\mathbf{x}$ has only one component different from zero (component l) and its value is $s \in \mathbb{C}$, a symbol taken from a constellation $\mathcal{S}$ with M symbols. Therefore, (6.1) can be also expressed as

$$\mathbf{y} = \sqrt{\gamma}\mathbf{h}_l s + \mathbf{w} \tag{6.2}$$

where $\mathbf{h}_l$ denotes the l-th column of $\mathbf{H}$, $l \in \{1, 2, \ldots, N_t\}$. We assume a unit power constraint, i.e., $\mathbb{E}\left[\mathbf{x}^H\mathbf{x}\right] = \mathbb{E}\left[|s|^2\right] = 1$.

We assume a packet based transmission where the transmitter encodes a group of information bits into a codeword or FECFRAME. Furthermore, we assume that the channel is block fading, i.e., that the channel matrix $\mathbf{H}$ remains constant during the reception of a frame.

In the system analyzed in this chapter, the coding rate can be chosen from a predefined set of codes for link adaptation purposes. K different coding rates are considered, with respective rates $r_1 < r_2 < \ldots < r_K$. The data is most protected with r_1, the rate closest to zero, although it has also the lowest spectral efficiency. Conversely, if the channel is particularly benevolent, information bits can be encoded with a rate r_K, which provides the highest spectral efficiency at the expense of a lower protection of the information bits. In general, the spectral efficiency attained with a given code of rate r is $\eta = r \log_2(N_t M)$, where M is the number of symbols of the constellation.

The problem of link adaptation in SM systems can be formulated in different ways, such as maximizing the spectral efficiency, minimizing the Bit Error Rate (BER) or maximizing the spectral efficiency but with a BER constraint. In this chapter, the latter option was selected with the constraint of the BER being equal or lower to a predefined (low) value p_0. In this chapter, as the system is assumed to employ a fixed constellation, the link adaptation is reduced to a coding rate selection. Hence, the rate adaptation is expressed as the following optimization problem,

$$\begin{aligned} \underset{r}{\text{maximize}} \quad & r \log_2(N_t M) \\ \text{subject to} \quad & r \in \{r_1, r_2, \ldots, r_K\} \\ & \text{BER}(\gamma; r, \mathbf{H}) \leq p_0. \end{aligned} \tag{6.3}$$

The BER dependence on the selected coding rate r, the channel matrix $\mathbf{H}$ and the SNR γ makes the design of this adaptive SM system very challenging.

In the previous chapter, Physical Layer Abstraction (PLA) techniques were used to predict the decoding outcome of a frame given the series of

time varying channel matrices **H**. The effective SNR and the Received Bit Mutual Information Rate (RBIR), based on the calculation of the average mutual information during a frame, allowed to determine if a given MCS or coding rate could be decoded successfully in a DP satellite system using PMod. There, the underlying assumption was that capacity achieving codes were available. To account for the outdated Channel State Information at the Transmitter (CSIT) due to long the Round Trip Time, adaptive margins were used in the link adaptation algorithms.

Nevertheless, in this chapter a radically different approach is applied. Firstly, we avoid PLA and we simulate the entire physical layer in both transmitter and receiver. Secondly, actual channel codes are used, namely, those taken from the DVB-S2 standard [7] for broadband satellite communications. And thirdly, although the link adaptation algorithm involves also a margin, the core element of the rate selection subsystem is a Multilayer Feedforward Neural Network (MFNN).

In Chapter 4 it is shown how to calculate the MI of SM and the unconstrained capacity of Generalized SM (GSM) systems by means of a MFNN. The proposed Machine Learning (ML) approach outperforms previous analytical approximations for doing this task, not only in terms of accuracy but also reducing significantly the computational cost of calculating these two information theory metrics. The MI calculated by the receiver could be used to select the MCS of the next transmission directly if capacity achieving codes were employed at the receiver side. If not, some back off margins should be applied in order to guarantee a robust transmission. For example, if the MI computed with the MFNN for a given channel condition happens to be 1.5 bits per channel use (bpcu), a MCS with an efficiency a bit lower than that must be used by the transmitter.

In this chapter, a MFNN is trained by means of supervised learning to choose directly the coding rate in a practical SM system. The network is trained offline using data obtained from system level simulations where the BER for the K codes is calculated as a function of the SNR for a large number of different channel matrices. Fig. 6.1 shows the block diagram of a general adaptive $N_t \times N_r$ SM system, with a fixed M-QAM constellation and a variable coding rate channel encoder which performs a joint coding of all the bits, those for choosing the symbols and the antenna indexes. After estimating the channel, the receiver employs a MFNN to decide which coding rate must be employed for the next frame by the transmitter and the transmitter is informed by means of a feedback channel. Next section will describe thoroughly the particular system designed to validate this approach for doing the adaptation.

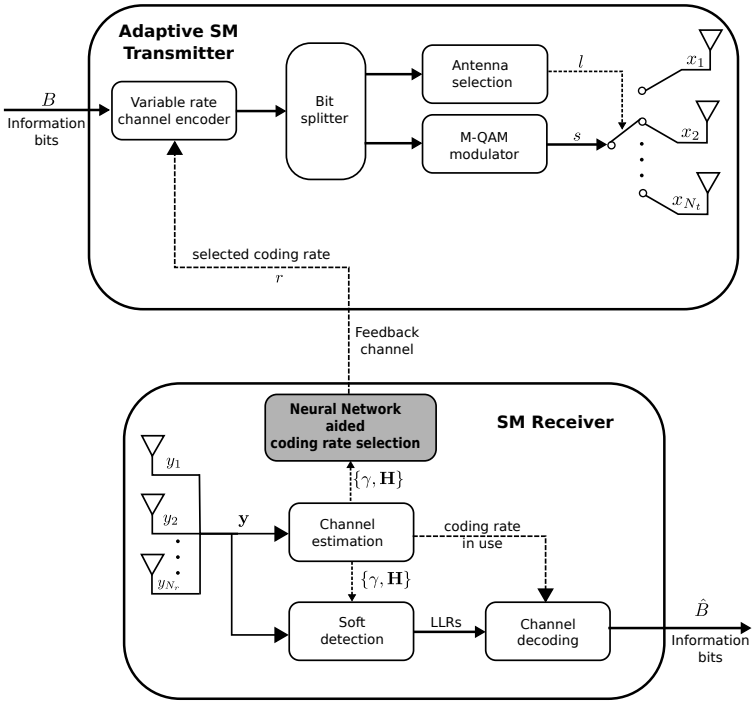


Figure 6.1: Block diagram of an adaptive SM system with variable coding rate.

LDPC code	**BCH Uncoded Block** K_{bch}	**BCH Coded Block** $N_{bch} = K_{ldpc}$	**LDPC Coded Block** N_{ldpc}
1/4	16 008	16 200	64 800
1/3	21 408	21 600	64 800
2/5	25 728	25 920	64 800
1/2	32 208	32 400	64 800
3/5	38 688	38 880	64 800
2/3	43 040	43 200	64 800
3/4	48 408	48 600	64 800
4/5	51 648	51 840	64 800
5/6	53 840	54 000	64 800
8/9	57 472	57 600	64 800
9/10	58 192	58 320	64 800

Table 6.1: Channel encoders parameters of the DVB-S2 FEC.

6.3. Practical SM System Specification

This section details all the specifications of both transmitter and receiver in order to allow reproducibility of the simulation results. The dimensions of the system are simply 2×2, i.e., there are 2 transmit and receive antennas. This example is valid for modeling a 2×2 SM terrestrial system and also a DP satellite system using PMod. The constellation is fixed to a QPSK, then $M = 4$ and the maximum achievable spectral efficiency with this system is $\log_2(N_t M) = \log_2(8) = 3$ bit/s/Hz.

The Forward Error Correction (FEC) of the proposed system is the same as in DVB-S2 [7]. It consists on the concatenation of Bose-Chaudhuri-Hochquenghem (BCH) code with a Low Density Parity Check (LDPC) code providing $K = 11$ different coding rates, ranging from 1/4 to 9/10. In DVB-S2 the codeword size N_{ldpc} is fixed to 64 800 bits, this is the length of the LDPC encoder output, a FECFRAME. Hence, for achieving a variable rate different payload sizes are allowed. Table 6.1 contains the length, in number of bits, of the input of the BCH encoder K_{bch}, its output N_{bch} (which is also the number of bits at the input of the LDPC encoder K_{ldpc}) and the final codeword size N_{ldpc}.

Fig. 6.2 shows a block diagram of the 2 antennas SM transmitter with the two components of the channel coding. Contrary to the diagram of a general SM transmitter of Fig. 6.1, there is not a particular bit which selects the antenna used to transmit the modulation symbol s since an optimized

bits to a SM symbols mapping is used in order to improve the performance, as it will be demonstrated later.

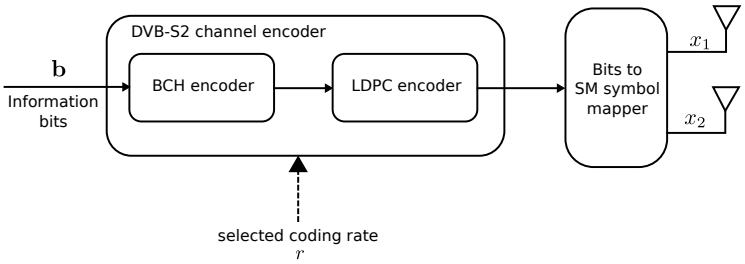


Figure 6.2: Block diagram of the adaptive 2×2 SM transmitter with variable coding rate and DVB-S2 coding.

If we compute all the squared distances between each pair of received symbols after passing through a channel matrix $\mathbf{H} = \mathbf{I}_2$

$$d = \|\mathbf{h}_l s_k - \mathbf{h}_{l'} s_{k'}\|^2 \tag{6.4}$$

the results of Table 6.3 are obtained. There, it can be seen that the squared distance between SM symbols of different antennas is always 2 and that the squared distance between symbols transmitted with the same antenna is either 2 or 4. Therefore, a Gray-coding philosophy can be applied to decide the mapping of bits to SM symbols, assigning bit sequences with a lower Hamming distance to closer SM symbols.

In this way, the bits sequence assignment to SM symbols of Table 6.4 was decided, where complementary bit sequences are assigned to the most distant SM symbols. Hence, the major error event of three wrong bits only occurs in the most unlikely situation, when a given SM symbol is confused with the furthest symbol. On the other hand, a confusion in only the antenna index causes only one erroneous bit out of 3, and, if the right antenna is detected but not the modulation symbol, there can be 1, 2 or 3 erroneous bits. The confusion with the neighbor symbol can cause 1 or 2 erroneous bits whilst, as previously said, the worst case of 3 erroneous bits happens when any of the SM symbols is confused with its furthest SM symbol.

Since this work is focused on just doing the rate adaptation, a basic SM constellation is assumed. However, there are articles which study in detail the design of optimal constellations for SM systems, as for example [109] and [43]. For the sake of simplicity, QPSK is chosen for the modulated symbols and simulations show that, even with this basic constellation, which is not optimized dynamically for the SM channel, good results can be obtained. Interestingly, the procedure for making the adaptation presented in this chapter is general enough to admit also other constellation

Symbols	s_1	s_2	s_3	s_4
Value	1	$1i$	-1	$-1i$

Table 6.2: Symbols of the QPSK constellation $\mathcal{S}$.

		Antenna 1				**Antenna 2**			
		$\boldsymbol{s_1}$	$\boldsymbol{s_2}$	$\boldsymbol{s_3}$	$\boldsymbol{s_4}$	$\boldsymbol{s_1}$	$\boldsymbol{s_2}$	$\boldsymbol{s_3}$	$\boldsymbol{s_4}$
Antenna 1	$\boldsymbol{s_1}$	0	2	4	2	2	2	2	2
	$\boldsymbol{s_2}$		0	2	4	2	2	2	2
	$\boldsymbol{s_3}$			0	2	2	2	2	2
	$\boldsymbol{s_4}$				0	2	2	2	2
Antenna 2	$\boldsymbol{s_1}$					0	2	4	2
	$\boldsymbol{s_2}$						0	2	4
	$\boldsymbol{s_3}$							0	2
	$\boldsymbol{s_4}$								0

Table 6.3: List of the squared distance between each pair of received symbols $\|\mathbf{h}_l s_k - \mathbf{h}_{l'} s_{k'}\|^2$ for the channel matrix $\mathbf{H} = \mathbf{I}_2$.

designs. The incorporation of a different constellation design would require only the modification of the system level simulations to obtain the codes performance.

Regarding the SM receiver design for the simulations of the adaptive system, Fig. 6.3 shows a block diagram of a 2 antennas SM receiver. We assume that perfect CSI is available at the receiver (CSIR), i.e., it knows perfectly the channel conditions: both the SNR γ and the channel matrix $\mathbf{H}$, and we assume that the receiver also knows the coding rate used in the received frame perfectly. With the CSIR the receiver performs the soft detection of the transmitted bits using the Log-Likelihood-Ratios (LLRs) since the channel decoder requires the soft bits as input. The exact LLRs of each bit are calculated following [22].

For each received symbol $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$, three LLRs values have to be calculated for the bits b_2, b_1 and b_0, since three bits per channel use are transmitted in this QPSK 2×2 SM system. We define $\mathcal{X}$ as the SM constellation, whose 8 possible symbol values are tabulated in Table 6.4. We define $\mathcal{X}_m^+$ as the set of SM symbols $\mathbf{x}$ whose $m-$th bit is 1 and $\mathcal{X}_m^-$ as the set of SM symbols $\mathbf{x}$ whose $m-$th bit is 0. These sets, given our bits to SM

Number	**Bits**	**Antenna l**	**Symbol s**	**SM symbol x**
0	000	1	s_1	$\mathbf{x}_0 = [1,\ 0]^t$
1	001	1	s_2	$\mathbf{x}_1 = [1i,\ 0]^t$
2	010	2	s_1	$\mathbf{x}_2 = [0,\ 1]^t$
3	011	2	s_2	$\mathbf{x}_3 = [0,\ 1i]^t$
4	100	2	s_4	$\mathbf{x}_4 = [0,\ -1i]^t$
5	101	2	s_3	$\mathbf{x}_5 = [0,\ -1]^t$
6	110	1	s_4	$\mathbf{x}_6 = [-1i,\ 0]^t$
7	111	1	s_3	$\mathbf{x}_7 = [-1,\ 0]^t$

Table 6.4: List of the 8 SM symbols $\mathbf{x}$ for this 2×2 SM - QPSK system with the bits assignment, which follows a Gray encoding.

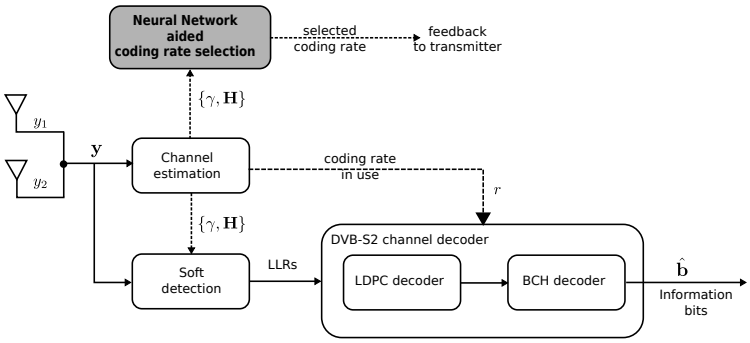


Figure 6.3: Block diagram of the adaptive 2×2 SM receiver with variable coding rate and DVB-S2 coding.

symbols mapping, are:

$$\begin{aligned}
\mathcal{X}_2^+ &= \{\mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7\} \\
\mathcal{X}_2^- &= \{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\} \\
\mathcal{X}_1^+ &= \{\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_6, \mathbf{x}_7\} \\
\mathcal{X}_1^- &= \{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_4, \mathbf{x}_5\} \\
\mathcal{X}_0^+ &= \{\mathbf{x}_0, \mathbf{x}_2, \mathbf{x}_4, \mathbf{x}_6\} \\
\mathcal{X}_0^- &= \{\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_5, \mathbf{x}_7\}.
\end{aligned} \tag{6.5}$$

The LLR of bit b_m is defined as

$$\mathrm{LLR}(b_m) = \log \frac{\mathrm{P}\{b_m = 0|\mathbf{y}\}}{\mathrm{P}\{b_m = 1|\mathbf{y}\}}. \tag{6.6}$$

Applying Bayes theorem and the fact that all SM symbols are equiprobable the previous expression can be developed as

$$\begin{aligned} \mathrm{LLR}(b_m) &= \log \frac{\mathrm{P}\{b_m = 0|\mathbf{y}\}}{\mathrm{P}\{b_m = 1|\mathbf{y}\}} && (6.7)\\ &= \log \frac{f_Y(\mathbf{y}|b_m = 0)\,\mathrm{P}\{b_m = 0\} f_Y(\mathbf{y})}{f_Y(\mathbf{y}|b_m = 1)\,\mathrm{P}\{b_m = 1\} f_Y(\mathbf{y})} && (6.8)\\ &= \log \frac{f_Y(\mathbf{y}|b_m = 0)}{f_Y(\mathbf{y}|b_m = 1)}. && (6.9)\\ & && (6.10) \end{aligned}$$

The Probability Density Function (PDF) of the received signal conditioned to the knowledge of $\mathbf{H}$ and $\mathbf{x}$ is

$$f_Y(\mathbf{y}|\mathbf{H}, \mathbf{x}) = \frac{1}{\pi^2} e^{-\|\mathbf{y} - \sqrt{\gamma}\mathbf{H}\mathbf{x}\|^2}, \tag{6.11}$$

which inserted in equation (6.10) along with the information of the bit mapping gives

$$\mathrm{LLR}(b_m) = \frac{\sum\limits_{\mathbf{x} \in \mathcal{X}_m^-} e^{-\|\mathbf{y} - \sqrt{\gamma}\mathbf{H}\mathbf{x}\|^2}}{\sum\limits_{\mathbf{x} \in \mathcal{X}_m^+} e^{-\|\mathbf{y} - \sqrt{\gamma}\mathbf{H}\mathbf{x}\|^2}}. \tag{6.12}$$

Although the LLRs can be approximated, which is specially interesting for large constellations, in the simulations of this chapter the exact LLRs were calculated. Note that another definition of the LLRs is possible, interchanging numerator and denominator. Our choice is motivated by the implementation of the Matlab® LDPC decoder.

6.4. Receiver Performance

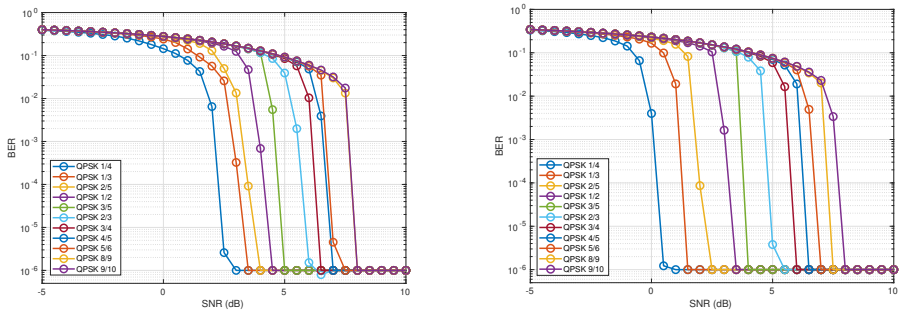
The objective of this section is showing the performance of the receiver described in the previous section for different channel conditions. In the first subsection, the performance for the case of the identity channel matrix will be obtained and, in the next subsection, for the case where the channel matrix follows a Rayleigh distribution.

6.4.1. Identity Matrix Channel

Firstly, we characterize the codes performance in a simple setup, with $\mathbf{H} = \mathbf{I}_2$, for two different bits to SM symbols mapping. In Fig. 6.4b the

BER curves of the 11 coding rates is shown for the assignment of bits to SM symbols previously introduced in Table 6.4. It is compared with an alternative assignment where there exists one particular bit that selects the antenna. However, in that assignment it is not applied a Gray-encoding philosophy for deciding the bits labeling of each SM symbol. Fig. 6.4a contains the BER curves in this alternative case, whose bits assignments are gathered in Table 6.5.

The results of plots from Fig. 6.4 confirm that doing a smart bits to SM symbols assignment helps to improve the system performance, specially for low coding rates. For example, for the code 1/4 there is a gap of about 2.5 dB between both curves. Fig. 6.5 depicts the threshold SNR required for achieving a BER of 10^{-4} with each code for these two different bits to symbols mappings. Moreover, the blue curve represents the normalized Mutual Information (MI), i.e., the capacity constrained to a QPSK constellation divided by the maximum spectral efficiency, three, in order to have a metric similar to the coding rate. There, the performance for both bits mapping can be seen more clearly. And, in addition, the comparison with the MI curve shows that this code does not achieve the capacity of the channel, specially for low coding rates. From this plot it can be concluded that, although the MI were computed with precision, this value cannot be used directly to choose the coding rate, because the codes are not so powerful and they require a SNR higher than the abscissa of the capacity curve to provide a robust commmunication.



(a) Non-Gray (b) Gray

Figure 6.4: BER curves of the DVB-S2 codes in a 2×2 QPSK SM system with $\mathbf{H} = \mathbf{I}_2$ for Non-Gray encoding (left) and Gray encoding (right).

Number	Bits	Antenna l	Symbol s	SM symbol x
0	000	1	s_1	$\mathbf{x}_0 = [1, 0]^t$
1	001	1	s_2	$\mathbf{x}_1 = [1i, 0]^t$
2	010	1	s_3	$\mathbf{x}_2 = [-1, 0]^t$
3	011	1	s_4	$\mathbf{x}_3 = [-1i, 0]^t$
4	100	2	s_1	$\mathbf{x}_4 = [0, 1]^t$
5	101	2	s_2	$\mathbf{x}_5 = [0, 1i]^t$
6	110	2	s_3	$\mathbf{x}_6 = [0, -1]^t$
7	111	2	s_4	$\mathbf{x}_7 = [0, -1i]^t$

Table 6.5: Non-Gray bits-SM symbols mapping.

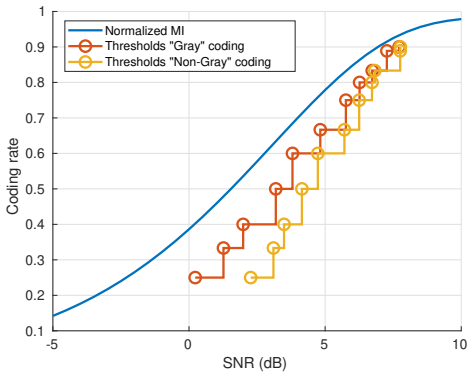
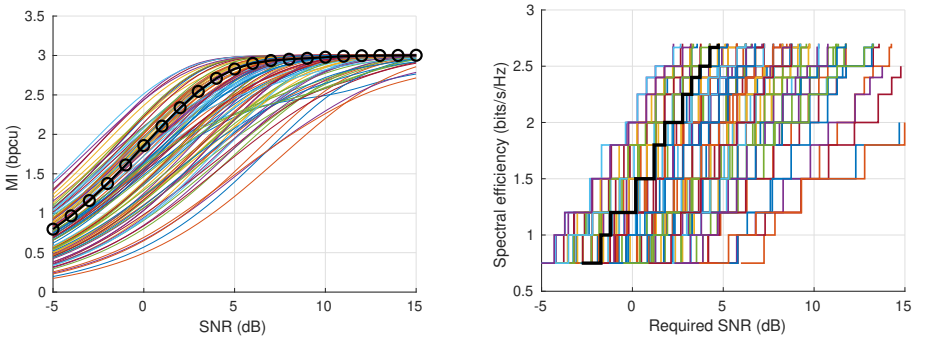


Figure 6.5: Normalized MI and threshold SNRs for a BER of $p_0 = 10^{-4}$ for each coding rate in a MIMO AWGN channel with $\mathbf{H} = \mathbf{I}_2$.

6.4.2. Rayleigh Channel Matrices

Now, in this subsection, the performance of the codes will be evaluated for Rayleigh distributed channel matrices, a typical distribution in terrestrial communications. They are combined with DVB-S2 codes, which come from the satellite communications field, only as an illustration. The coefficients of $\mathbf{H}$ are taken from a complex Gaussian distribution such that $h_{ij} \sim \mathcal{CN}(0, 1)$. Firstly, in the left graphic of Fig. 6.6 the MI is computed for 100 different Rayleigh distributed channel matrices and its dependence with the SNR is shown with the colored lines. The thick black line and the markers represent the MI with $\mathbf{H} = \sqrt{2}\mathbf{I}_2$, for comparison purposes. The scaling factor $\sqrt{2}$ aims to normalize the Fröbenius norm of $\mathbf{H}$ so that it is the same as the average Fröbenius norm of the Rayleigh matrices.

The MI for a given SNR γ changes a lot depending on the particular channel matrix. Fig. 6.6a exemplifies that by showing the MI curve for 100 different Rayleigh channel matrices. Differences in the MI as high as almost 2 bpcu can be observed for some values of SNR. On the other hand, Fig. 6.6b shows the threshold SNR required for attaining a BER equal or lower than 10^{-4} with each coding rate for the same 100 Rayleigh matrices. Again, it is highlighted with the black thick line the stairstep graphic for the reference matrix $\mathbf{H} = \sqrt{2}\mathbf{I}_2$. The difference in the SNR required for decoding a particular coding rate with one matrix or another can reach the astonishingly value of 10 dB, as can be seen in Fig. 6.6b.



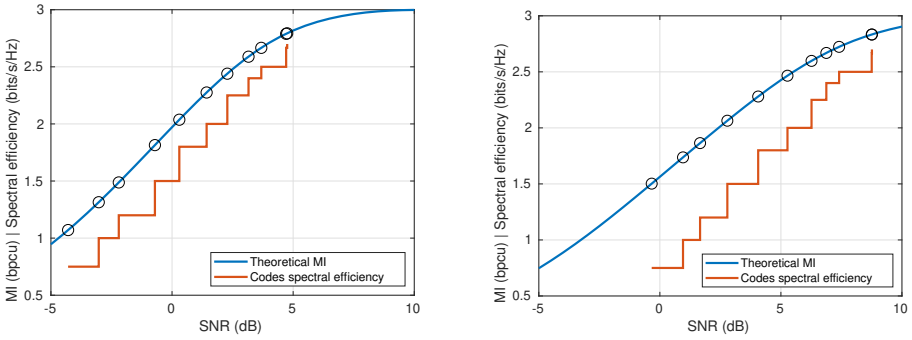
(a) Mutual Information (b) Threshold SNRs for a BER of 10^{-4}

Figure 6.6: Mutual Information and threshold SNR for decoding a frame with a BER of 10^{-4} with Rayleigh distributed channel matrices.

Fig. 6.6b reveals that the selection of the coding rate in a SM system cannot rely only on the SNR γ and that the channel matrix $\mathbf{H}$ has also to be taken into account, since depending on the particular $\mathbf{H}$ the threshold SNR of a given code has a big shift. One may ask the following question. Could

the theoretical MI be used to decide the coding rate? We have developed several methods to obtain the MI and the unconstrained capacity of SM systems in the Chapter 4. There, we suggested that a positive margin Δ could be subtracted from the calculated MI I_T, and that the transmission with a coding rate or MCS with spectral efficiency $I_T - \Delta$ could serve as a way to select the coding rate whilst meeting a BER constraint. In the following, simulation results will demonstrate that a global margin Δ can be obtained. Moreover, they show that there is not an unique margin as a function of the theoretical MI either, i.e., there does not exist a margin as a function of the theoretical MI which works well with all the matrices. The optimum margin has to be matrix and MI dependent. This will lead to a Deep Learning (DL) based solution to perform the coding rate adaptation, which will be explained in the following section.

Firstly, in Fig. 6.7 the theoretical MI is compared with the spectral efficiency that can be achieved with the proposed DVB-S2 codes (for the 10^{-4} BER constraint). The left and right plot are the results for two different matrices. The circles simply denote the theoretical MI at the values of the threshold SNR of each coding rate. It can be observed how the margin Δ required to be applied to the MI, if this is used to select the coding rate, is always higher at low values of the MI, whereas for high values of the MI, the two curves are closer. Moreover, comparing Figs. 6.7a and 6.7b, the margin for the lowest coding rate is very different in these two examples.

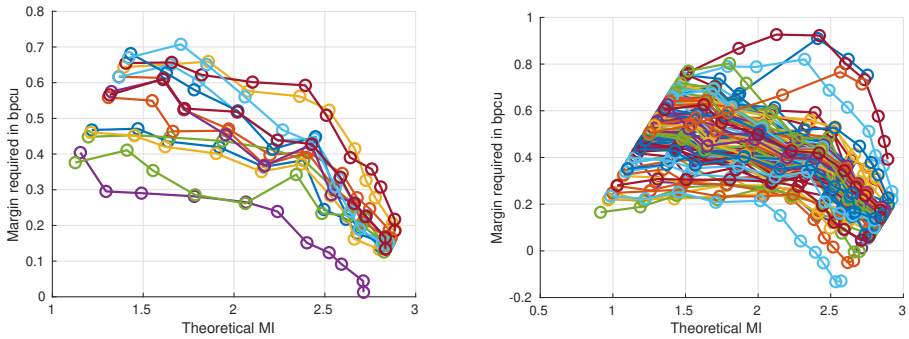


(a) Example matrix 1 (b) Example matrix 2

Figure 6.7: Theoretical MI and threshold SNR required for a BER of 10^{-4} with each coding rate for two different channel matrices drawn from a Rayleigh distribution.

Finally, Fig. 6.8 shows the value of the required margin Δ to subtract from the theoretical MI in order to choose the right coding rate. Fig. 6.8a contains the values of the margin Δ as a function of the theoretical MI for 15 matrices where it can be easily observed how this margin decreases with the

value of the theoretical MI. On the other hand, Fig. 6.8b stacks these plots for 100 different matrices. This allows to observe that with a global margin of $\Delta = 0.8$ bpcu the BER constraint will be met for the vast majority of the channel matrices and theoretical MI values. However, Fig. 6.8 demonstrates that there does not exist a margin as a function of the MI which works well with all the matrices. The optimum margin for a given family of channel codes and a BER constraint depends on the particular channel matrix.



(a) 15 matrices (b) 100 matrices

Figure 6.8: Margin required to subtract to the MI to choose the right coding rate.

6.5. Deep Learning Assisted Coding Rate Selection

In this section, all the steps involved in the design of an adaptive SM system will be detailed; this system will adapt the coding rate of each transmit frame to follow the varying channel conditions. A Deep Learning (DL) approach is applied to solve the problem of the rate adaptation. The proposed solution consists in the offline training of a Multilayer Feedforward Neural Network (MFNN); this will later be used by the SM receiver to select in real time the coding rate that the SM transmitter should apply to the next transmitted frame. This section tries to be general enough so that the principles described here can be applied to any SM system. Then, next section completes the system description initiated in Section 6.3 previous to the simulation results, which are provided in Section 6.7.

Two different phases can be differentiated, the design phase and the operation phase. These comprise the following steps, that will be explained with more detail later:

1. **Design phase**, previous to the system deployment.
   i) *System level simulations.* These aim to evaluate the performance of the channel codes for some fixed transmitter and receiver architecture
   ii) *Extraction of the threshold SNRs.* In this step a target BER is fixed and by using the data of the previous step a collection of threshold SNRs for each channel matrix and coding rate is obtained.
   iii) *Construction of the Machine Learning dataset*
   iv) *Neural network training*
   v) *Neural network performance evaluation*
2. **Operation phase:** *Neural network assisted coding rate selection by the receivers in real time.*

6.5.1. System level simulations

This first step is, by far, the most time consuming, taking even several days of execution time. The performance of the channel codes has to be obtained for the receiver design which will be used by the practical system. Previously in this chapter, it has been shown how the channel matrix has a significant influence on the SNR required for decoding a frame with a given coding rate. Therefore, in the system level simulations a large number of different channel matrices $\mathbf{H}$ should be tried, and these matrices should follow the same distribution as that expected in the practical deployment. The simulations have to be run with all the parameters of the system frozen and these should be the same as in the final system. More specifically, the number of antennas, the family of codes, the set of coding rates, the constellation, the mapping of bits to SM symbols, the architecture of both transmitter and receiver and the detection and decoding algorithms should be fixed.

The output of this step is a collection of performance curves of the channel codes for a large number of channel matrices. These consist of BER curves as a function of the SNR for all the available coding rate options and for each simulated channel matrix. Mathematically, we denote this BER as $\text{BER} = f(\gamma;\ r, \mathbf{H})$. Fig. 6.9 depicts it graphically, showing an example of a collection of BER curves for several channel matrices.

Despite this step being very time consuming, these simulations would have to be done anyway previously to any SM system development. However, they can be easily parallelized and multiple processors or computers

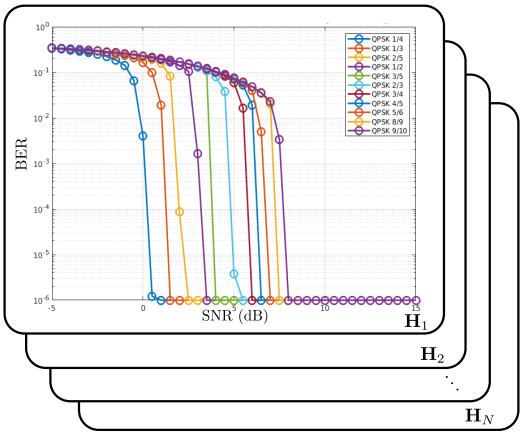


Figure 6.9: The different channel codes performance must be evaluated for a large number of channel matrices.

can work independently to compute the performance for different channel matrices. What is proposed here is to leverage the results of these system level simulations, and employ them to train a neural network which assists later a receiver in the coding rate selection.

6.5.2. Extraction of the threshold SNRs

In the formulation of the optimization problem of equation (6.3), which serves as basis for selecting the coding rate, a target BER constraint is added. Hence, the coding rate which provides the highest spectral efficiency is chosen from the subset of all the coding rates whose BER is equal to or lower than p_0 for the receiver channel conditions, given by γ and $\mathbf{H}$. At this step, a value for p_0 must be assigned; 10^{-4} was used for the simulation results of this chapter. The powerful channel codes selected for the system evaluation, taken from the DVB-S2 standard [7], have a very steep decrement of the BER with the SNR, as it could be seen in Figs. 6.4 and 6.9. As a consequence, the BER is nearly zero for an SNR higher than a given threshold.

In this step, the BER curves obtained in the previous point are to be processed to extract the threshold (minimum) SNR to guarantee a given BER p_0 with a given coding rate for each of the simulated channel matrices. Thus, a collection of stairwise plots, like those of Fig. 6.10, is produced, one per matrix. With these plots the optimum coding rate can be obtained for any value of SNR but only for the set of channel matrices evaluated in the first step. The objective of the learning scheme is to be able to generalize,

in order to obtain the coding rate for any arbitrary channel matrix $\mathbf{H}$, not only those simulated previously.

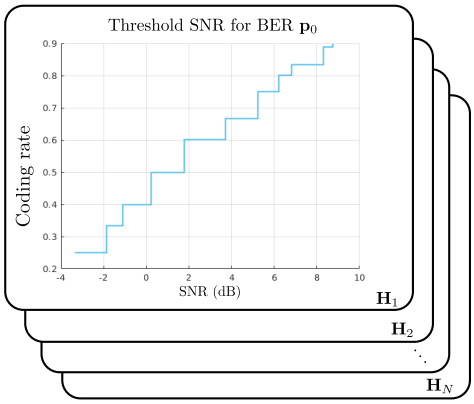


Figure 6.10: For each channel matrix $\mathbf{H}_i$, the minimum required SNR to guarantee a target BER p_0 with each code is extracted.

6.5.3. Construction of the Machine Learning dataset

The next step consists on building the dataset $\mathbb{X}$ used in the training and testing of the neural network. The selection of the neural network input features $\mathbf{x}$ has a paramount importance to obtain a good performance. The vector of features $\mathbf{x}$ is obtained from the SNR γ and the channel matrix $\mathbf{H}$ by means of a transformation $\mathbf{x} = g(\gamma, \mathbf{H})$. In Table 4.1 of Section 4.4, several options for this function $g(\gamma, \mathbf{H})$ are given. There it is shown that the columns norms $\gamma\|\mathbf{h}_l\|^2$ and the angles between each pair of columns of $\mathbf{H}$ (Hermitian angle Θ_H and Kasner's pseudoangle φ) are a good selection for the input features.

The dataset $\mathbb{X} = \{(\mathbf{x}_i, \mathsf{y}_i), i = 1, \ldots, m\}$ is made of all the input-output pairs. The vector $\mathbf{x}_i$ is the neural network input extracted from each tuple $(\gamma_i, \mathbf{H}_i)$, the channel conditions evaluated in the system levels simulations of the first step. The scalar y_i, the neural network desired output or target variable, is a real variable which represents the highest coding rate, from all the available rates, which can be used whilst meeting the BER p_0. Thus, $\mathsf{y}_i = r_k$ for coding rate r_k, and $\mathsf{y}_i = 0$ if no coding rate can guarantee the prescribed BER.

6.5.4. Neural network training

Once the ML dataset $\mathbb{X}$ is built, a neural network architecture has to be selected and the network has to be trained to obtain its internal parameters $\boldsymbol{\theta}$. The goal is to provide good predictions $\hat{\mathsf{y}}$ from the input features $\mathbf{x}$: $\hat{\mathsf{y}} = h(\mathbf{x}; \boldsymbol{\theta})$. The design of the architecture of the network requires the selection of the number of hidden layers (*depth* of the model), of the number of neurons per layer (*width* of the model), and the specification of the output units [40].

The coding rate selection is a classification problem, wherein a discrete class $r \in \{0, r_1, r_2, \ldots, r_K\}$, representing one of the choices for the coding rate, must be assigned to an input vector $\mathbf{x}$. Traditionally, neural networks for doing a multi-class classification have as many outputs as classes and the output layers employ softmax units [40], which give the probability that the input belongs to each class. However, much better results were obtained with a neural network applying regression, which has only one linear output producing an estimation of y. In this case, to obtain a valid coding rate from the output of the neural network $\hat{\mathsf{y}}$, this variable has to be quantized to find the closest valid coding rate.

In this work, tangent hyperbolic is used as the activation function for the hidden layer neurons and the network training is performed with the Levenberg-Marquardt (LM) backpropagation algorithm [45] to minimize the Mean Squared Error (MSE). However, other alternatives for the activation function, the learning algorithm and the cost function could be possible. The training of the neural network is done offline, prior to the system deployment and operation, and it is the second most time consuming task, after the system level simulations. However, it only takes a few minutes, depending on the complexity of the neural network. Once the training is finished and the network parameters $\boldsymbol{\theta}$ have been obtained, the receiver simply uses the neural network with these parameters to make the predictions and selecting the coding rate. The computational complexity of doing so is very low, specially for the small neural networks we propose later in this chapter.

6.5.5. Neural network performance evaluation

The neural network performance evaluation must be conducted on the testing portion of the dataset, which is different from the portion employed for the training. The classification results are fully characterized by the confusion matrix, from which three key metrics can be extracted:

i) the **accuracy**, defined as the ratio between the correct coding rates predictions and the total number of predictions made,

ii) the ratio of **underselections**, defined as the division of the number of cases where a coding rate lower than the target coding rate is selected, i.e., too robust, between the total number of predictions made, and

iii) the ratio of **overestimations**, defined as the division of the number of cases where a coding rate beyond the capabilities of the receiver is selected between the total number of predictions made.

The latter is the most critical since the overestimation causes an outage, with the receiver unable to decode a frame with that coding rate, which results in a loss of throughput. Conversely, an underestimation only causes an underutilization of the link (a reduction of the average spectral efficiency), but the transmission is still successful.

The classification performance depends on the mapping or quantization between the real-valued neural network output $\hat{\mathrm{y}}$ and the set of available coding rates. The simplest way of doing it when the network target output is directly the coding rate is

$$r = Q(\hat{\mathrm{y}}) = \arg\min_{r_k} |\hat{\mathrm{y}} - r_k| \tag{6.13}$$

If the outage probability obtained by selecting the coding rate with equation (6.13) is too high, a positive margin can be substracted from the network output $\hat{\mathrm{y}}$ in order to reduce the probability of coding rate overestimation:

$$r = Q(\hat{\mathrm{y}} - \Delta) = \arg\min_{r_k} |\hat{\mathrm{y}} - \Delta - r_k| \tag{6.14}$$

This, of course, reduces the accuracy and it can reduce the throughput too, but it allows to operate with a lower outage probability, which can be desirable in some scenarios where low latency or ultra reliable communications are mandatory.

6.5.6. Operation phase

Finally, during the system operation, receivers employ the neural network trained in the design phase to perform the coding rate selection. The transmitter receives through a feedback channel the choice made by the receiver, which is used for the selection of the coding rate for the next frame transmissions intended for that receiver. Although the neural network parameters $\boldsymbol{\theta}$ are fixed during the operation phase, the margin Δ, introduced in equation (6.14), could still be adapted with the ACKs/NAKs information,

in a similar way as it was proposed in the previous chapter for DP satellite systems. In this way, the coding rate selection can be made more robust and the adaptive margin can help to reduce the negative consequences of a mismatch between the simulated and working conditions. However, the use of adaptive margins together with neural networks will not be addressed in this thesis.

6.6. System Simulation Setup

In the previous section a procedure to perform coding rate selection with the assistance of DL was proposed. This procedure is general enough to apply to any type of SM system. In this section, some more details will be given for the particular 2×2 SM system which was already presented in Section 6.3. In addition, more information of the neural network and the simulation parameters will be provided. The reference transmitter and receiver are those depicted in Figs. 6.2 and 6.3, respectively.

The simulated SM system has 2 antennas, at both transmission and reception sides, and it employs the QPSK constellation defined in Table 6.2. With regard to the channel coding, the same as in the DVB-S2 standard [7] is employed, which consists on the concatenation of a BCH and a LDPC code. Although other codes could have been used, as for example the turbocodes of LTE, the decision of using DVB-S2 codes was motivated by their ready-to-use implementation in Matlab® and because it was very easy to switch from one coding rate to another. However, we believe that this DL-based adaptation could be applied with any type of channel codes.

In DVB-S2 the 11 different coding rates collected in Table 6.1 are available. Nevertheless, in the system simulated for this chapter two coding rates, 4/5 and 8/9, were deleted since they are very close to other coding rates. This reduction makes the distances between the available coding rates more uniform and eases the classification, whilst nine coding rates still provide enough adaptation steps between the minimum and maximum spectral efficiency. Therefore, there are $K = 9$ available coding rates r_k, ranging from 1/4 to 9/10. Additionally, a class $r_0 = 0$ is considered, with the meaning that no coding rate can be used whilst satisfying the BER constraint, fixed to $p_0 = 10^{-4}$.

As stated previously, the length of the coded frame is fixed to 64 800 bits like in DVB-S2 standard and thus the number of information bits per codeword is variable and depends on the particular coding rate selected. The LDPC maximum number of iterations for decoding was set to 50. The

mapping of bits to SM symbols follows a Gray encoding, by assigning complementary bits sequences to the SM symbols which are more distant. The bits to SM symbol mapping is listed in Table 6.4. We assume perfect Channel State Information (CSI) at the receiver, which includes the SNR γ and the channel matrix $\mathbf{H}$. These are used during the soft detection to provide the exact LLRs, which the channel decoder employs then to decode the information bits.

The system level simulations to obtain the performance of the codes are calculated for $N = 1\,000$ different 2×2 channel matrices $\mathbf{H}$, generated by following a unit-variance Rayleigh distribution, i.e., $h_{ij} \sim \mathcal{CN}(0, 1)$. For each matrix the average BER after the BCH decoding is calculated for 41 equispaced values of SNR between -5 and 15 dB. The average BER is calculated after simulating the transmission of 25 FECFRAMEs. The target BER for coding rate selection is set to $p_0 = 10^{-4}$. Table 6.6 sums up the main parameters of the system.

Parameter	**Value**
Transmit and receive antennas	$N_t = 2$, $N_r = 2$
Constellation	QPSK ($M = 4$)
Channel coding	DVB-S2 codes (BCH + LDPC)
Number of coding rates	$K = 9$
Coding rate options	1/4, 1/3, 2/5, 1/2, 3/5, 2/3, 3/4, 5/6, 9/10
Target BER	$p_0 = 10^{-4}$
Channel matrices	1 000 unit variance Rayleigh distributed
SNR range	-5 to 15 dB (0.5 dB steps)

Table 6.6: Simulated system parameters.

The neural network input features $\mathbf{x}$ are calculated for each tuple $(\gamma, \mathbf{H})$ as

$$\mathbf{x} = g(\gamma, \mathbf{H}) = \left[\text{sort}\left(\gamma\|\mathbf{h}_1\|^2,\ \gamma\|\mathbf{h}_2\|^2\right),\ \ \Theta_H,\ \ \varphi\right]^t \tag{6.15}$$

following the results presented in Chapter 4, where the columns norms are sorted. The parameters Θ_H and φ are two angles obtained from the scalar product between the two complex column vectors of the channel matrix, $\mathbf{h}_1$ and $\mathbf{h}_2$, as

$$\mathbf{h}_1^H \mathbf{h}_2 = \|\mathbf{h}_1\| \cdot \|\mathbf{h}_2\| \cdot \cos\Theta_H \cdot e^{i\varphi}. \tag{6.16}$$

The so-called Hermitian angle Θ_H belongs to the interval $[0, \pi/2]$ whereas φ, named Kasner's pseudo-angle, takes values between $-\pi$ and π [84].

The dataset $\mathbb{X}$ consists of $m = 41\,000$ examples, 41 different values of SNR γ for each one of the 1 000 channel matrices $\mathbf{H}$. In the system level simulations phase, the BER curves as a function of the SNR for the $K = 9$ coding rates is obtained for each one the 1 000 matrices. Then, the threshold SNRs for decoding the frames with each coding rate are obtained for all the

channel matrices. Afterwards, for each example in the dataset, i.e., for each tuple $(\gamma_i, \mathbf{H}_i)$, the target coding rate y_i is obtained, this is the coding rate with the highest spectral efficiency which guarantees the target BER $p_0 = 10^{-4}$.

The dataset $\mathbb{X}$ is divided into three independent parts. 15% of the examples were reserved for the final test of the performance of the neural network. The remaining 70% and 15% were employed for training and validation of the neural network, respectively. Hence, 6 150 examples are used for testing and validation and 28 700 for the training phase.

Several neural networks architectures were tested with a number of hidden layers between one and seven and a number of neurons per layer between 5 and 30. Each neural network has been trained 20 times with different sets of initial parameters. The training runs for 1 000 epochs, although it could be halted earlier if the network performance on the validation dataset stopped improving or remained the same for 6 epochs in a row. The default parameters of the *trainlm* function of Matlab® were used for the training.

6.7. Simulation Results

This section contains the main simulation results of this chapter divided into two parts. The first part presents the performance of several neural networks architecture for selecting the coding rate. The second part shows the performance of the adaptive SM system in terms of spectral efficiency for different cases. These include a non-adaptive baseline SM system with a fixed coding rate and several adaptive SM systems with different coding rate adaptation techniques.

6.7.1. Classification Performance

Selecting the optimum coding rate given some channel conditions can be formulated as a classification problem since the target output takes a finite set of values. Here 10 different classes are considered, no transmission (N/T), meaning that the target BER $p_0 = 10^{-4}$ cannot be satisfied even with the lowest coding rate, and the remaining 9 classes represent each one of the coding rates of the system, which are collected in Table 6.6.

As stated previously, we have proposed to use a neural network for doing curve fitting, i.e., which uses a one-neuron linear output. The discrete class is found simply by finding the closest valid coding rate to the neural network

	10 neurons		20 neurons	
Hidden layers	**MSE**	**Accuracy (%)**	**MSE**	**Accuracy (%)**
1	$1.70 \cdot 10^{-3}$	86.0	$1.19 \cdot 10^{-3}$	90.6
2	$7.56 \cdot 10^{-4}$	91.0	$5.14 \cdot 10^{-4}$	**95.8**
3	$5.12 \cdot 10^{-4}$	94.5	$4.86 \cdot 10^{-4}$	95.7
4	$\mathbf{4.99 \cdot 10^{-4}}$	**95.6**	$4.96 \cdot 10^{-4}$	95.5
5	$5.10 \cdot 10^{-4}$	95.3	$4.90 \cdot 10^{-4}$	95.6
6	$5.13 \cdot 10^{-4}$	95.2	$\mathbf{4.82 \cdot 10^{-4}}$	95.7
7	$\mathbf{4.99 \cdot 10^{-4}}$	95.3	$5.39 \cdot 10^{-4}$	95.4

Table 6.7: Performance of the neural networks for different architectures, ranging from 1 to 7 layers with width of 10 or 20 neurons.

real-valued output $\hat{\mathsf{y}}$, which is expressed mathematically in equation (6.13). We found that this architecture offers a higher accuracy than a traditional neural network for performing classification.

Several architectures have been tested with a number of hidden layers between 1 and 7 and with a number of neurons per layer between 5 and 30. In chapter 4 it was shown that a MFNN with just one hidden layer of 10 or 20 neurons can do a good estimation of the SM MI or the Generalized SM (GSM) capacity. Taking those neural network architectures as starting point, several networks have been trained and their performance calculated in the testing dataset.

Firstly, neural networks with 10 or 20 neurons per hidden layer and a number of hidden layers ranging from 1 to 7 were trained. As this was an exploratory search, only 4 independents trainings were performed for each one of the architectures. Table 6.7 shows the best MSE and classification accuracy obtained with the most performing neural network parameters found in the training. The MSE, calculated as

$$\text{MSE} = \frac{1}{6150} \sum_{i=1}^{6150} \left(\hat{\mathsf{y}}_i - \mathsf{y}_i\right)^2 , \tag{6.17}$$

and the accuracy, all obtained in the testing dataset, are given.

The results of Table 6.7 show that, with the selected input features of equation (6.15) even a neural network with just one hidden layer provides quite acceptable results, being the classification accuracy 86.0% and 90.6% for the cases of 10 and 20 neurons per layer, respectively. For the neural network with a width of 10 neurons there is not significant improvement after a depth of 3 layers and for the network with 20 neurons width this happens after a number of 2 layers. Table 6.7 shows with boldface numbers the best result of each column.

Architecture	**Accuracy**	**Outage**	**Underselection**	**Minimum accuracy per MCS**	**Complexity**
10 + 10 + 10	95.1 %	2.55 %	2.34 %	87.47 %	**250 weights**
20 + 10 + 5	94.9 %	2.67 %	2.43 %	86.1 %	335 weights
20 + 20	**96.2** %	**1.92** %	1.92 %	87.7 %	500 weights
20 + 15 + 10	**96.2** %	2.07 %	**1.72** %	**89.3** %	540 weights
30 + 20 + 10	95.3 %	2.05 %	2.70 %	88.9 %	930 weigths

Table 6.8: Performance of the neural network with 5 different architectures, retaining the data from most performing training out of 20.

Two architectures of Table 6.7, the case of 3 layers of 10 neurons and 2 layers of 20 neurons, were analyzed with more detail along with three other variants of the former. In this case the network of each architecture was trained 20 times with different internal parameters initialization and the best results are shown in Table 6.8. The new three alternatives consist on neural networks of 3 layers with $20 + 10 + 5$, $20 + 15 + 10$ or $30 + 20 + 10$ neurons per layer.

In Table 6.8 the neural networks are ordered according to increasing complexity, measured in terms of the total number of weights of each network. For each architecture another four metrics are given: the total accuracy, the outage probability, the probability of selecting a coding rate below the target one and the minimum accuracy per MCS (or coding rate). As it will be shown later, the accuracy is not evenly distributed among all the target coding rates and, in addition, some of the MCS, namely the lowest, N/T, and the highest, 9/10, are the most likely and their accuracy is also higher. For this reason, the lowest accuracy of a MCS is also tabulated, to give an idea of the minimum performance expected with each network.

There are no significant differences among the architectures of Table 6.8. As can be seen, the total accuracy is always around 95 %, the outage and underselection probabilities have a very small value, typically about 2 %, and the lowest accuracy per coding rate is near 90 %. To select a neural network representative of each architecture from the 20 trained networks the following criteria were followed. On the one hand, it was found a neural network which improves the four metrics, i.e., which maximizes the accuracies and minimizes the outage and underselection probabilities, and that at the same time, if an error in the classification happens, using equation (6.13), the confusion is only with the adjacent coding rates.

Hereafter, two of the architectures of the Table 6.8 will be selected and analyzed with more detail. Namely, the neural network with the smallest complexity, the network of 3 hidden layers of 10 neurons in each layer, and

the best performing neural network, that with three layers of 20, 15 and 10 neurons per layer. The first one, the network with lowest number of weights, will be referred as *architecture 1* (Arch. 1) and the second, the network with a slightly better performance but with a higher number of weights, will be referred as *architecture 2* (Arch. 2).

The first batch of results presented hereafter are obtained with the raw classification, i.e., the selected coding rate or MCS is the valid coding rate closest to the neural network output, following equation (6.13) without using any margin. Figs. 6.11 and 6.12 depict the **confusion matrix** of the raw classification for Arch. 1 and 2, respectively. There, it can be observed how two classes, namely N/T and 9/10, are over-represented compared to the others. In 19.4 % of the simulated conditions the channel is not good enough to allow a robust transmission with the lowest coding rate satisfying the BER constraint. And, in 34.8 % of the cases, the channel conditions allow to use the highest coding rate. Actually, in part of them even other high order constellations, as 8PSK and 16QAM, could be employed due to the high SNR. However, in this thesis we restrict ourselves to the problem of coding rate selection and a more general adaptation, including not only the coding rate but also the constellation, is left as possible future line.

The green diagonal elements of the confusion matrices indicate that the classification is correct whilst all the upper and lower off-diagonal elements represent the cases where a sample of the testing dataset is miss-classified. Inside the matrix the numbers of samples and percentage represent absolute values, referred to the all testing dataset, whereas the bottom row and the right column show the accumulative percentages referred to each target or output class. The green color is employed for the correct classifications percentage and the red color for the errors.

In both confusion matrices of Figs. 6.11 and 6.12 it can be seen how the neural networks performs very well: the right coding rate is selected or, at most, the adjacent codes are chosen. Thus, we have a sparse matrix with only three non-zero diagonals. Other facts can be deduced from the confusion matrix. Since the two most common classes have a very high accuracy, these help to increase the overall performance of the classification. Nevertheless, the minimum accuracy per target MCS is 87.5 %, being for the rest of MCS around 90 %.

Since the two confusion matrices are very sparse, the same information is represented in a more clear way by means of Tables 6.9 and 6.10. There, the accuracy and outage and underselection probabilities are split among all the target coding rates. A comparison of the two tables shows that the Arch. 2 offers a slightly better performance, increasing a little bit the accuracy

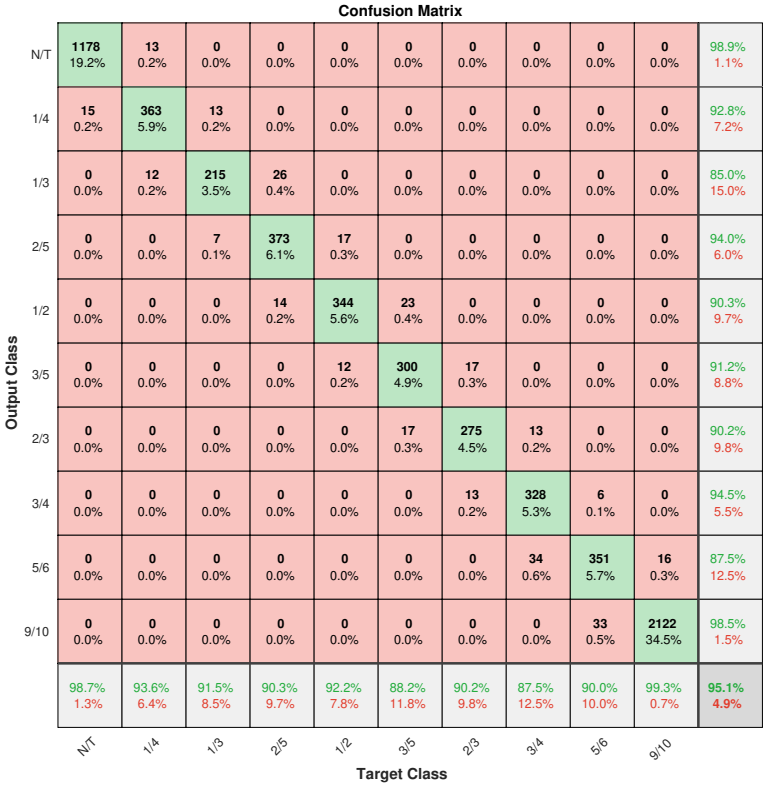


Figure 6.11: Confusion matrix of the raw classification with Arch. 1 (10+10+10 neurons).

and reducing the miss-classifications ratios. This improvement is at the expense of increasing the complexity of the neural network. Although both have three hidden layers, the network with Arch. 1 uses only 250 weights whereas the other has 540 weights, as indicated previously in Table 6.8.

Now, some plots regarding the neural network output are included. The left side of Figs. 6.13 and 6.14 shows a **regression plot** where the target coding rate versus the actual neural network output is represented. Except for the lower classes, N/T and 1/4, the neural network output variable is centered around the target value. However, previous tables show that even for these two target classes the accuracy is very good, more than 95 % and almost 99 % for classes 1/4 and N/T, respectively. On the other hand, the right plots of Figs. 6.13 and 6.14 reflect graphically that the neural network assisted coding rate selection makes the correct choice or it confuses the right MCS with the two adjacent classes.

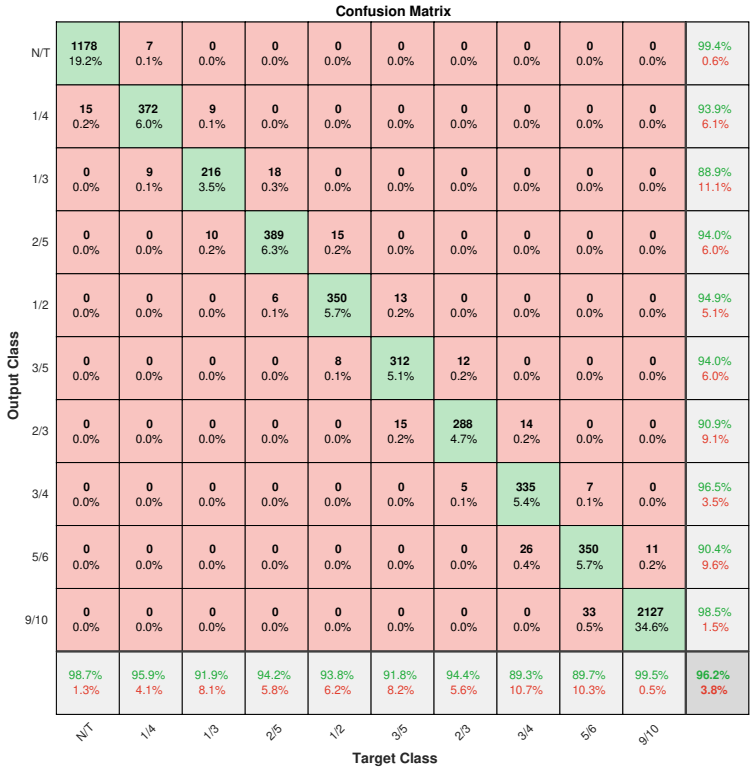
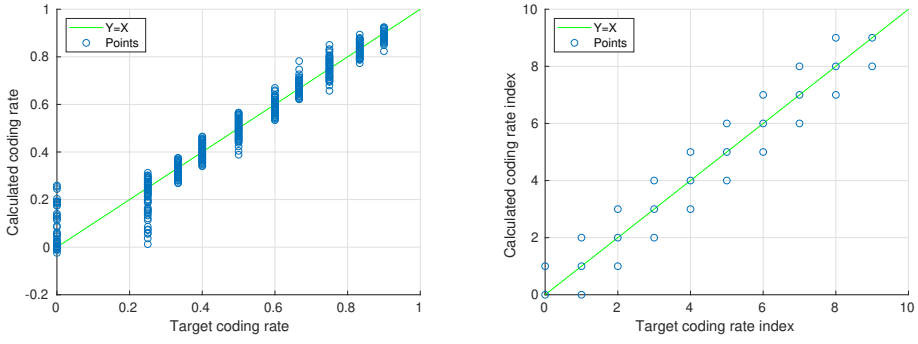


Figure 6.12: Confusion matrix of the raw classification with Arch. 1 (20+15+10 neurons).



(a) Neural network ouput

(b) Selected coding rate index

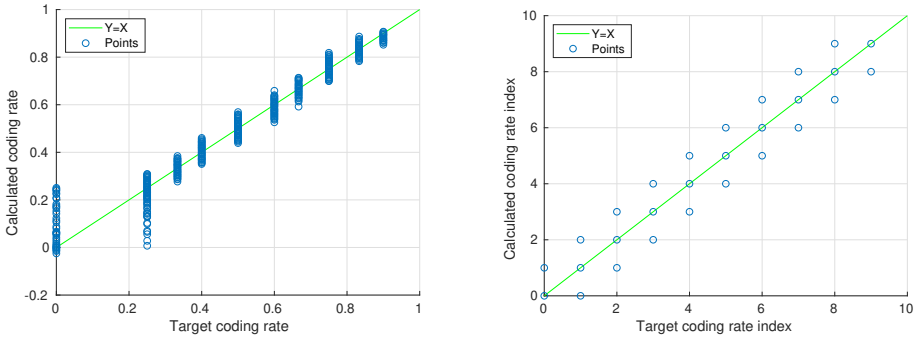
Figure 6.13: Regression plot with target coding rate and calculated coding rate for Arch.1 (10+10+10 neurons).

Target coding rate	**N/T**	**1/4**	**1/3**	**2/5**	**1/2**	**3/5**	**2/3**	**3/4**	**5/6**	**9/10**
Accuracy (%)	98.7	93.6	91.5	90.3	92.2	88.2	90.2	87.5	90.0	99.3
Outage (%)	1.3	3.1	3.0	3.4	3.2	5.0	4.3	9.1	8.5	-
Underselection (%)	-	3.4	5.5	6.3	4.6	6.8	5.6	3.5	1.5	0.7

Table 6.9: Classification performance for Arch. 1 (10+10+10 neurons).

Target coding rate	**N/T**	**1/4**	**1/3**	**2/5**	**1/2**	**3/5**	**2/3**	**3/4**	**5/6**	**9/10**
Accuracy (%)	98.7	95.9	91.9	94.2	93.8	91.8	94.4	89.3	89.7	99.5
Outage (%)	1.3	2.3	4.3	1.5	2.1	4.4	1.6	6.9	8.5	-
Underselection (%)	-	1.8	3.8	4.4	4.0	3.8	3.9	3.7	1.8	0.5

Table 6.10: Classification performance for Arch. 2 (20+15+10 neurons).

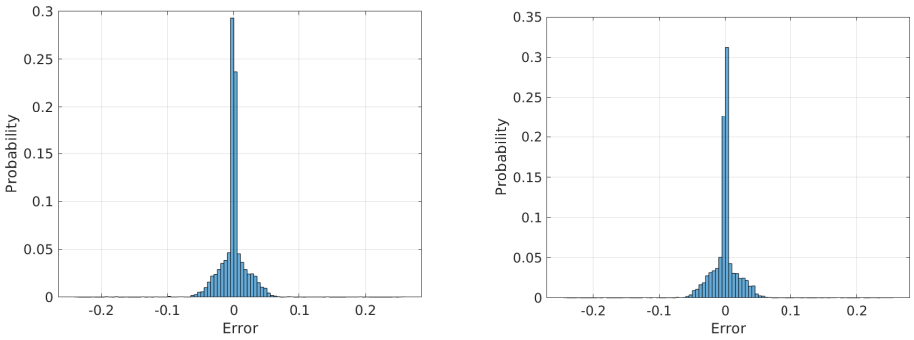


(a) Neural network ouput (b) Selected coding rate index

Figure 6.14: Regression plot with target coding rate and calculated coding rate for Arch. 2 (20+15+10 neurons).

Some last plots regarding the neural network output are shown in Fig. 6.15. There, the **histograms of the neural network output error**, $\hat{y} - y$, are represented for the two architectures. The error is concentrated around zero and, in addition, it can be seen how the vast majority of the probability is concentrated in the ± 0.05 region, some good news taking into account that the minimum separation between coding rates is 0.0667. The histograms spans from -0.25 to 0.25 due to the output of the network for the two lower classes.

Lastly, in Table 6.11 the performance of the two analyzed neural architectures is summarized, giving some additional information, as the MSE, the Mean Absolute Error (MAE), the maximum and minimum error and the value of the three times the standard deviation, which provides an idea of where the 99.7 % of the errors are concentrated. Moreover, we include the mean accuracy per coding rate after excluding the two extreme MCS,



(a) Arch 1. (10+10+10 neurons) (b) Arch 2. (20+15+10 neurons)

Figure 6.15: Histogram of the error for the two architectures, defined as the difference between the neural network output $\hat{y}$ and the target coding rate y.

which are over represented and help to increase the global average accuracy.

Classification for reducing the outage

Before moving to the next subsection, where the performance of adaptive and non adaptive SM systems will be shown in terms of spectral efficiency, some notes about the margin Δ required for avoiding outage situations will be given. The results shown so far were obtained with a raw classification, i.e., without applying any margin to the output of the neural network. The selected coding rate is, then, the valid coding rate closest to the network output. This is expressed mathematically with equation (6.13). However, in a previous paragraph, it was advanced that the calculation of the coding rate can be modified by subtracting a back-off margin Δ, so that the outage probability can be reduced. Therefore, the coding rate is selected in this case as follows,

$$r = Q(\hat{y} - \Delta) = \arg\min_{r_k} |\hat{y} - \Delta - r_k|, \tag{6.18}$$

where $\hat{y}$ is the neural network output, r_k are the $K+1$ valid coding rates, Δ is the positive margin and $Q(\cdot)$ is a quantifying function which finds the closest valid coding rate to its input argument.

In Fig. 6.16 it is shown the margin Δ required for achieving a zero outage probability for each target coding rate. Results for both neural networks architectures are shown there. The x-axis represents the index of coding rate, ranging from 0 (N/T) to 9 (rate 9/10). It can be seen that the target

	Arch. 1 (10+10+10)	**Arch. 2 (20+15+10)**
Total accuracy	95.11 %	96.21 %
Total outage proabability	2.55 %	2.07 %
Total underselection prob.	2.34 %	1.72 %
Minimum accuracy per MCS	87.47 %	89.33 %
Mean accuracy per MCS[1]	90.43 %	92.64 %
MSE	$5.22 \cdot 10^{-4}$	$4.75 \cdot 10^{-4}$
Mean Absolute Error (MAE)	0.0123	0.0118
Maximum error	0.259	0.2507
Minimum error	−0.2370	−0.2427
3σ error	0.0685	0.0653

[1] Without taking into account N/T and 9/10.

Table 6.11: Raw classification performance (without any margin) of the two selected neural network architectures.

coding rate N/T increases the overall required margin, as could be deduced from the previous regression plots of Figs. 6.13 and 6.14. The margin for this case could be disregarded since if the target coding rate is N/T (no transmission) there is not throughput loss even if a frame is transmitted with rate 1/4 (and non detected). Therefore, we can focus our attention on the second maximums of Fig. 6.16. If we look at these values, remarkable differences between the two architectures can be noticed. On the one hand, Arch. 1 requires a margin $\Delta = 0.08$, whereas, on the other hand, Arch. 2 requires a smaller margin of just $\Delta = 0.03$.

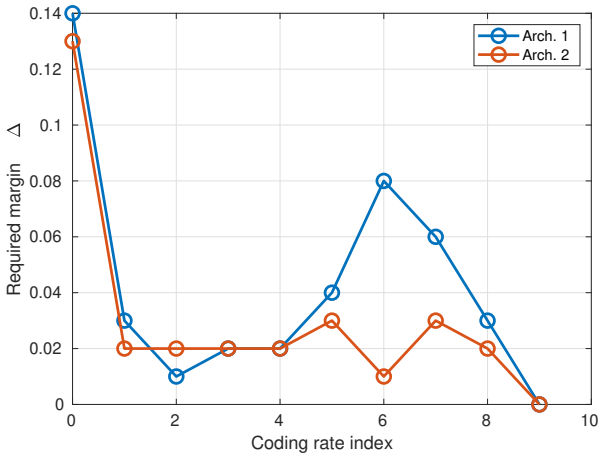


Figure 6.16: Required margin Δ per each target coding rate for having zero outage probability in the testing dataset.

	Arch. 1 (10+10+10)			Arch. 2 (20+15+10)		
Margin	**$\Delta = 0$**	**$\Delta = 0.08$**	**$\Delta = 0.13$**	**$\Delta = 0$**	**$\Delta = 0.03$**	**$\Delta = 0.13$**
Accuracy	95.1 %	28.1 %	20.6 %	96.2 %	80.0 %	21.6 %
Mean accuracy[1]	90.4 %	17.7 %	2.4 %	92.6 %	68.1 %	4.4 %
Outage	2.6 %	0.065 %	0 %	2.0 %	0.21 %	0 %
Underselection	2.3 %	71.8 %	79.4 %	1.7 %	19.8 %	78.4 %

[1] Without taking into account N/T and 9/10.

Table 6.12: Comparison of the performance of the two neural network architectures with and without a margin to reduce the outage probability.

Once the margins required for achieving a zero outage in this dataset have been identified, some final results of the classification performance with these margins are given. Table 6.12 shows the accuracy, the outage probability and rate of selection of coding rates lower than the target for both neural network architectures and for three different values of the margin in each case. The results for the raw classification, i.e., with a zero margin, are shown for comparison purposes. It can be seen that the cost we have to pay for reducing the outage probability is the reduction of the accuracy and the rise of the underselection rate. In the light of the results of Table 6.12, the neural network Arch. 2 offers a good trade-off with the margin $\Delta = 0.03$. In this case, the outage is very low (and even zero if we do not count the target rate N/T) whilst the accuracy is still high, about a 80.0 %. This value contrasts with the very low accuracy of the Arch.1 for a margin $\Delta = 0.08$, only 28.1 %.

For the sake of completeness, we include in Fig. 6.17 the confusion matrix of the classification with Arch. 2 and with a margin of $\Delta = 0.03$. Moreover, Fig. 6.18 contains two regression plots with target versus the selected coding rate for Arch. 2 and for two different margins, 0.03 and 0.13. There it can be understood graphically very well how the addition of a back-off margin for reducing the outage probability reduces the accuracy irretrievably and increases the underselection rate. Despite of this, with $\Delta = 0.03$ the correct coding rate is still selected in the majority of the cases and the miss classifications happen only with the coding rate immediately below the target one.

6.7.2. System Level Performance

In the previous subsection the performance of different neural networks for doing the coding rate selection was analyzed thoroughly. Now, this last subsection includes different graphics of the average spectral efficiency

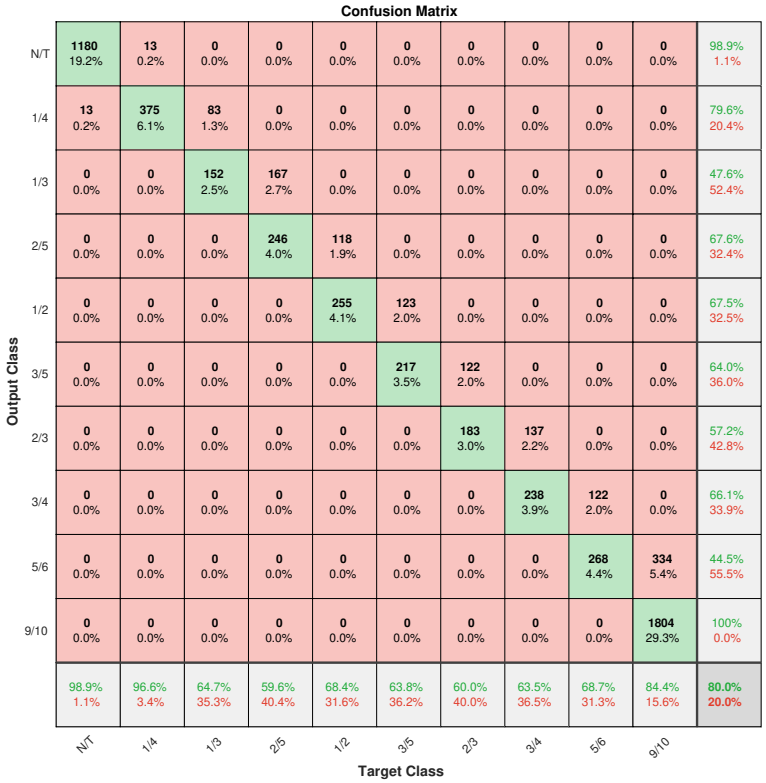
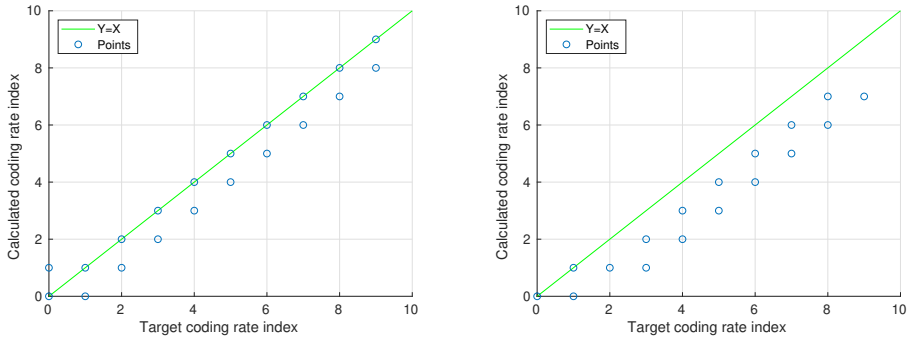


Figure 6.17: Confusion matrix of the classification with Arch. 2 (20+15+10 neurons) using a margin $\Delta = 0.03$ for an almost zero outage probability in the testing dataset (zero disregarding the target rate N/T).

and the average outage probability that a 2×2 SM system can achieve using different adaptation strategies. Several adaptive SM systems and non-adaptive SM systems with fixed coding rate will be compared hereafter.

The three different types of adaptation in SM systems that we are going to consider are the following:

- **Fixed coding rate**
  This is the simplest SM system, without any type of adaptation. The transmitter uses always the same fixed coding rate, which should be low enough to provide an acceptable link availability. If the coding rate is higher, the spectral efficiency of the system is better at the high SNR regime, but at the expense of a higher outage probability.



(a) Margin $\Delta = 0.03$ (b) Margin $\Delta = 0.13$

Figure 6.18: Regression plot with target coding rate and selected coding rate for Arch. 2 (20+15+10 neurons) and two different margins.

Mathematically, the coding rate selection can be expressed as

$$r = r_k, \quad \text{for some fixed } k. \tag{6.19}$$

- **MI-based coding rate selection**
  In this case, we have an adaptive SM system, where the transmitter varies the coding rate of each frame using the information fed back by the receiver. The receiver, assumed to have perfect CSIR, calculates the Mutual Information (MI) that the channel supports I_T, using the method of its preference for doing that calculation, and then applies a fixed back-off margin Δ to that MI, previously to decide the proper coding rate. In Section 6.4, we commented that subtracting a margin of about 0.80 from the MI could serve to select the coding rate. This value, of course, depends on the employed codes, the architecture of the receiver and the channel distribution. Other scenarios would require a different margin value. In the simulations used for this chapter, the MI I_T is calculated with the MFNN of just one hidden layer of 20 neurons which has four inputs, the channel matrix column norms ($\gamma\|\mathbf{h}_l\|^2$) and the two angles (Θ_H and φ). The equation for performing the adaptation is then

$$r = Q\left(\frac{I_T - \Delta}{3}\right) = \arg\min_{r_k} \left|\frac{I_T - \Delta}{3} - r_k\right|, \tag{6.20}$$

  In the previous equation the corrected MI $I_T - \Delta$ is divided by 3 to obtain the estimated coding rate. This value comes from the maximum spectral efficiency of a 2×2 SM system with a QPSK constellation, which is 3 bpcu or bits/s/Hz. Then, the function $Q(\cdot)$ searches the valid coding rate which is closest to the estimated coding rate $\dfrac{I_T - \Delta}{3}$, which comes from the MI calculation.

- **DL-based coding rate selection**
  We can anticipate that this approach, the one proposed in this chapter, outperforms the other two options, the fixed coding rate and the MI-based selection. In this case, the coding rate is chosen with the help of Deep Learning (DL) as it was explained in detail in Section 6.5. The performance of different neural networks for selecting the coding rate was analyzed in the previous subsection. If $\hat{\mathrm{y}}$ is the output of the neural network estimating the optimum coding rate and Δ is an optional margin, applied to reduce the outage probability; the selection of the coding rate can be expressed with equation (6.14), repeated hereafter for the sake of completeness of this part:

$$r = Q\left(\hat{\mathrm{y}} - \Delta\right) = \arg\min_{r_k} \left|\hat{\mathrm{y}} - \Delta - r_k\right|. \tag{6.21}$$

The **average spectral efficiency** η for each SNR value γ is computed using all the ML dataset as

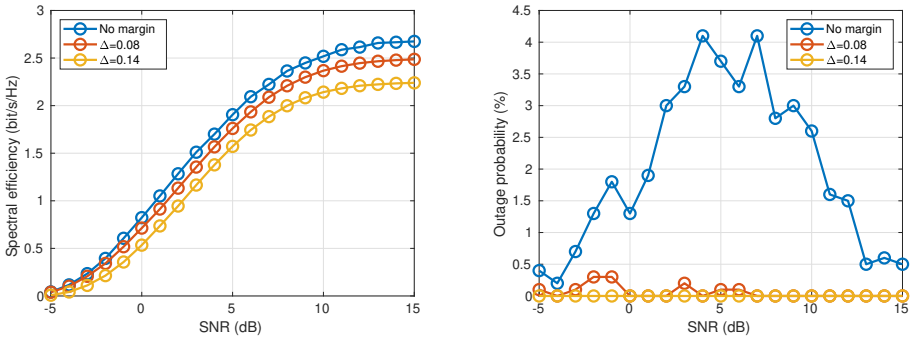
$$\eta = \frac{1}{N}\sum_{i=1}^{N} \log_2(N_t M)\hat{r}_i(1-\epsilon_i) = \frac{1}{N}\sum_{i=1}^{N} 3\hat{r}_i(1-\epsilon_i), \tag{6.22}$$

where $\hat{r}_i$ is the coding rate selected for the channel matrix i and ϵ_i is a binary variable which takes the value 1 in the case of an error event, i.e., when a coding rate higher than the target coding rate is selected, and which takes the value 0 when the correct coding rate or a lower one is selected. In this case, $N = 1\,000$, since one thousand different channel matrices are simulated per SNR value. Hereafter, for all the spectral efficiency and outage probability plots, the complete dataset $\mathbb{X}$ is used and not only the part used in the neural network testing, in order to obtain smoother graphics without increasing the number of channel matrices. With the previous definition of ϵ_i, the **average outage probability** per SNR is formulated as

$$\text{Average outage probability} = \frac{1}{N}\sum_{i=1}^{N} \epsilon_i. \tag{6.23}$$

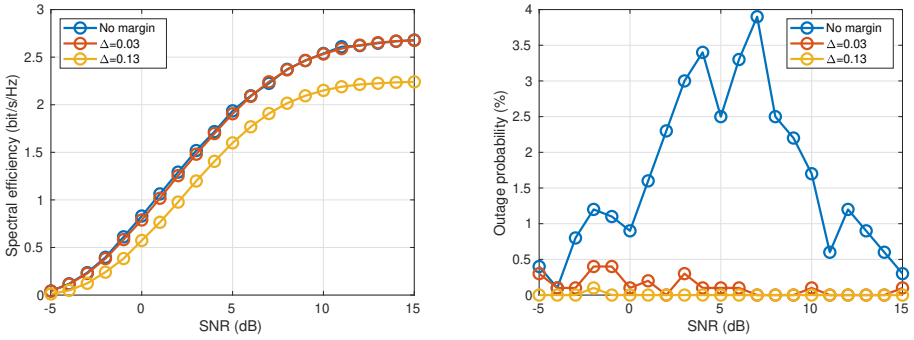
Now, that the three different adaptation strategies have been precised and the two performance metrics defined, a set of 6 figures will show the average spectral efficiency and the average outage probability as a function of the SNR for different cases. Firstly, Figs. 6.19 and 6.20 show the performance of the proposed approach, the **DL-based coding rate selection**, for the two neural networks architectures considered, Arch. 1 and Arch. 2, respectively. In these plots, the results for three different margins Δ are represented. With both architectures it can be observed that an increment of the margin reduces the outage probability and it also reduces the spectral

efficiency. Fig. 6.20 deserves a special comment. Although a margin of $\Delta = 0.03$ is applied to reduce the outage probability, the spectral efficiency is barely compromised whereas the probability of not decoding a frame is almost zero. Therefore, the neural network with the Arch. 2, which has three hidden layers of 20, 15 and 10 neurons, along with a back-off margin of $\Delta = 0.03$ is the best solution we propose to solve the problem of coding rate selection in this 2×2 SM system.



(a) Average spectral efficiency (b) Average outage probability

Figure 6.19: Average spectral efficiency and average outage probability per SNR for an adaptive 2×2 SM system with a QPSK constellation and DL-based coding rate selection with the neural network with Arch. 1 (10+10+10 neurons).

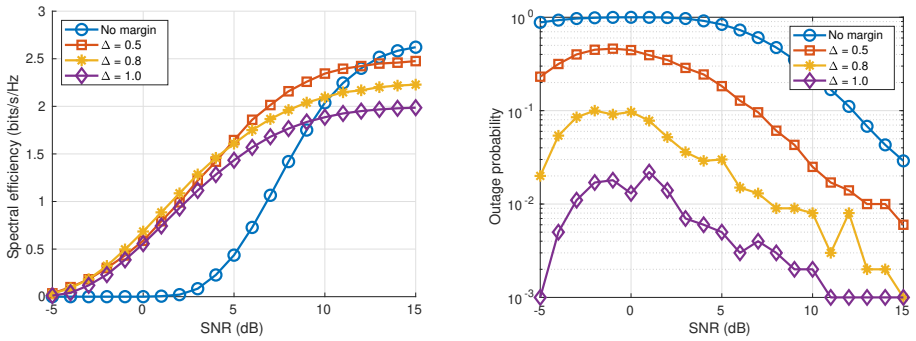


(a) Average spectral efficiency (b) Average outage probability

Figure 6.20: Average spectral efficiency and average outage probability per SNR for an adaptive 2×2 SM system with a QPSK constellation and DL-based coding rate selection with the neural network with Arch. 2 (20+15+10 neurons).

Secondly, Fig. 6.21 shows the results of the **MI-based coding rate selection** for four different values of the margin Δ, from 0 to 1.0. If no margin

is applied the outage probability is extremely high, as can be observed in Fig. 6.21b. A margin of 0.8 is required to reduce the outage probability to 10 % and with a margin of 1.0 the outage lies between 0.1 and 2.0 %, depending on the SNR. As the margin increases, the outage probability gets smaller. However, the behavior of the spectral efficiency with the margin is not as clear. Depending on the average SNR one margin or another provides the highest spectral efficiency. The dependence of the throughput of a system with the outage probability used in the link adaptation was analyzed in [70]. In general, there exists and optimal outage probability where the throughput is maximized. Therefore, not all the margins provide the highest spectral efficiency.

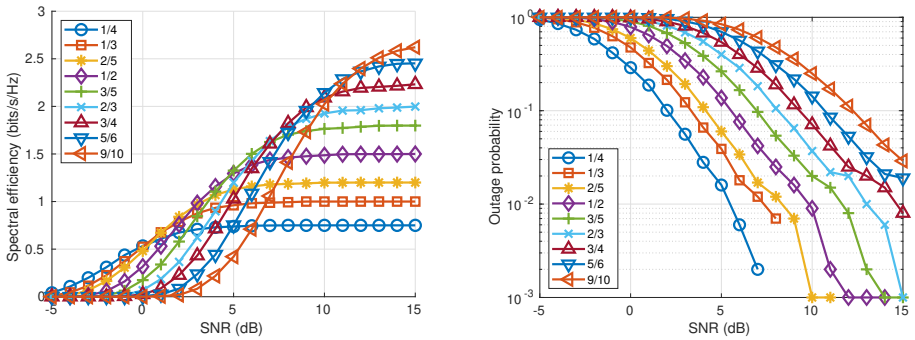


(a) Average spectral efficiency (b) Average outage probability

Figure 6.21: Average spectral efficiency and average outage probability per SNR for an adaptive 2×2 SM system with a QPSK constellation and MI-based coding rate selection.

And thirdly, Fig. 6.22 shows the results obtained with the SM system which uses a **fixed coding rate**. The performance with the 9 codes considered here is shown. A clear trade-off between spectral efficiency and outage can be seen. Whereas a low coding rate offers a smaller outage probability and it is desirable at low SNRs, when the SNR of the system rises, the spectral efficiency saturates to a small value compared with higher coding rates. Therefore, if a system uses a fixed coding rate, it must employ a very low rate to allow the system to work correctly in all the channel conditions. This penalizes the performance at high SNRs, where more efficient coding rates could be selected if an adaptive rate were used.

The last two figures of this section serve to compare the performance of the different approaches: fixed coding rate, MI-based and DL-based coding rate selection. Fig. 6.23 shows all the possible coding rates for the fixed rate setup whereas in Fig. 6.24 only two coding rates, 1/4 and 1/2, are shown for that case. The latter offers a better insight into the problem, since any



(a) Average spectral efficiency

(b) Average outage probability

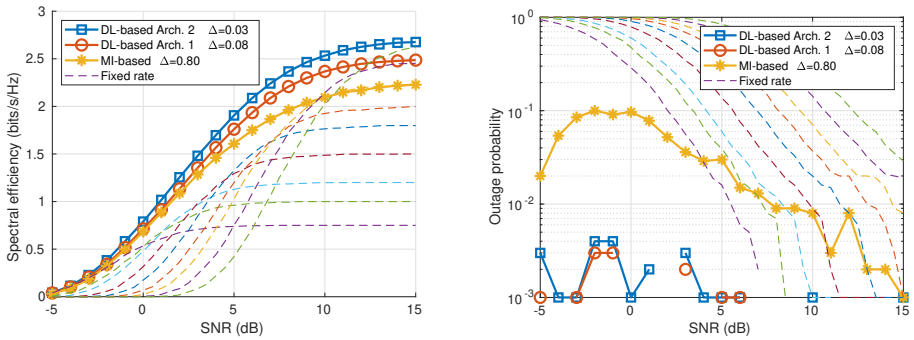
Figure 6.22: Average spectral efficiency and average outage probability per SNR for a non-adaptive 2×2 SM system with a QPSK constellation and a fixed coding rate.

practical SM system with fixed coding would use only very low rate codes.

In Fig. 6.24 the performance of five different SM systems is compared. Two non-adaptive systems with a fixed coding rate of 1/4 and 1/2 and three adaptive SM systems with DL and MI-based coding rate selection. In the case of DL two different neural networks architectures are employed, Arch. 1 with a margin $\Delta = 0.08$ and Arch. 2 with a margin $\Delta = 0.03$. In both cases these margins guarantee and almost zero outage probability (zero if we disregard the N/T target coding rate). On the other hand, in the MI-based system a margin $\Delta = 0.80$ is used, which reduces the outage although this is much higher than with the DL-based approach.

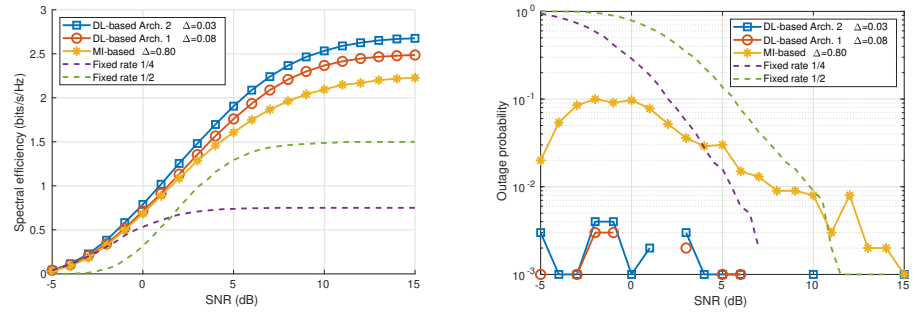
If we look at the spectral efficiency plot of Fig. 6.24a, we can see how the curves for the DL-based selection stay always above the other two approaches: MI-based selection and fixed coding rate. Furthermore, the neural network with Arch. 2 provides the best performance, being slightly better than the neural network with Arch. 1. The second best approach, the MI-based selection, represents also a major improvement with regard to the fixed coding rate. A glance at the outage probability of Fig. 6.24b reveals that the DL-based selection also beats the other two approaches in this metric. The outage probability is zero or it remains very low, in the order of 10^{-3}, between one and two orders of magnitude below the MI-based and the fixed coding rate approaches.

Finally, Fig. 6.25 contains two bar diagrams showing the maximum spectral efficiency at high SNR for the five cases compared in Fig. 6.24. A maximum value of spectral efficiency of 2.68 bits/s/Hz is reached with the DL-based selection using the Arch. 2, very near to the maximum theoretical



(a) Average spectral efficiency (b) Average outage probability

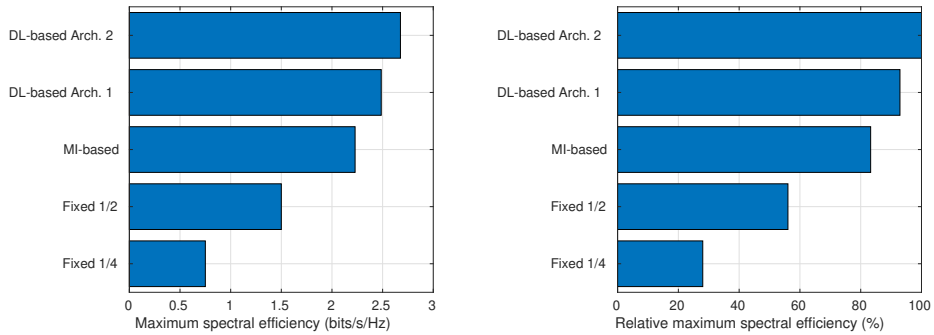
Figure 6.23: Comparison of the average spectral efficiency and average outage probability per SNR for several adaptive and non-adaptive 2×2 SM system with a QPSK constellation.



(a) Average spectral efficiency (b) Average outage probability

Figure 6.24: Comparison of the average spectral efficiency and average outage probability per SNR for several adaptive and non-adaptive 2×2 SM system with a QPSK constellation.

spectral efficiency of 2.70 bits/s/Hz which could be achieved with a 2×2 SM system with QPSK and a 9/10 coding rate. With the Arch. 2, the spectral efficiency is reduced to the 92.9 % of the other architecture, being the maximum value 2.49 bits/s/Hz. The MI-based coding rate selection achieves only a 83.3 % of the spectral efficiency of the DL-based with Arch. 2 and the spectral efficiency is only 2.23 bits/s/Hz. The two non adaptive SM systems with coding rates 1/4 and 1/2 are well below the other adaptive systems, with a spectral efficiency of 0.75 and 1.50 bits/s/Hz, respectively, the 28.0 % and 56.1 % of the maximum value the DL-based approach gets.



(a) Average spectral efficiency (b) Relative spectral efficiency

Figure 6.25: Comparison of the maximum average spectral efficiency of the different adaptation approaches.

6.8. Conclusions and Future Work

The adoption of a link adaptation procedure is necessary for any communication system to make the most of the channel capacity. Spatial Modulation (SM) and its analogous Polarized Modulation (PMod) are being proposed for future 5G systems and for Dual Polarization mobile satellite systems, respectively. Not much previous work was known for the design of adaptive SM systems, in all cases for the adaptation of the modulation order. However, a finer level of adaptation varying the coding rate, which is done in many other communications systems, was not studied. This coding rate adaptation in SM systems was first addressed in this thesis.

This chapter proposes a new adaptive SM system where the coding rate of the channel encoder can be adapted dynamically to follow the channel variations. The coding rate selection is done by the receiver using a Deep Learning approach. In this way, the receiver, assumed to have perfect Channel State Information, uses a deep neural network to select the coding rate of the frames it receives from the other end. The neural network is trained offline, with data obtained from extensive system level simulations for different channel conditions. In this way, by means of tagged examples, the neural network learns to choose the optimum coding rate for each channel condition, represented in terms of the Signal to Noise Ratio (SNR) and the channel matrix.

The procedure for doing the coding rate adaptation is explained in general terms and it can be applied to a SM system with any number of antennas. In this chapter, it was exemplified for a simple 2×2 SM scenario with a QPSK constellation where 9 coding rates were available. A neural

network of just three hidden layers and a number of neurons between 10 and 20 is enough for achieving an accuracy higher than 90 % in the coding rate selection. Moreover, when an erroneous rate is selected, miss-classifications only happen with the two adjacent coding rates.

In the last part of the chapter, the performance of several SM systems is compared. If the SM system employs a fixed coding rate, this has to be low enough to guarantee a reliable communication, leading to an inefficiency use of the spectrum at high SNRs. An improved way of doing the adaptation consists on calculating the Mutual Information (MI), applying a back-off margin to it and using this corrected MI value to select the coding rate. This MI-based coding rate selection increases substantially the spectral efficiency and reduces the outage probability compared with the fixed rate setup. However, the proposed DL-based coding rate selection outperforms the other two methods. It increases the average spectral efficiency even more than the MI-based adaptation mechanism. Moreover, the outage probability with the DL-based selection is close to zero with the help a small back-off margin applied to the neural network output.

This work has a natural extension, firstly, more than one constellation should be included, as is done in most practical communications systems, for example 16QAM and 64QAM. In this way, the adaptation would be in terms of both modulation order and coding rate. The target output of the neural network would have to be changed, and instead of using the network to estimate the optimum coding rate, this should be used to estimate the spectral efficiency. In addition, in the system level simulations more complex scenarios could be emulated, with 4 and 8 antennas. In another enhancement, the adaptive coding and modulation could be merged with some works related with codebook selection in SM, as for example [80], obtaining as result of this a totally adaptive SM system. By doing so the transmitter can play with the coding rate, with the modulation order and with the available antennas in order to have a better adaptation to the channel conditions and improve even more the spectral efficiency of these systems. Lastly, to account the imperfections of the channel model used in the neural network training, the system could be improved by means of using an adaptive margin. Alternatively, Reinforcement Learning could be applied so that the neural network itself is adapted by the receiver with the information of the frames decoding outcome.

Chapter 7

Conclusions and Future Work

The inclusion of Adaptive Coding and Modulation (ACM) technology in modern communication standards provides major improvements in the spectral efficiency, since the communication link is no longer designed for the worst case channel conditions. With the adaptation of the physical layers parameters, the transmitter changes the transmission bit rate according to the time variant channel capacity. The design of the corresponding link adaptation algorithms to perform ACM in future terrestrial communications systems is critical to fully exploit their channel capacity.

Firstly, Chapter 2 described the experimental validation of several link adaptation algorithms employed to make the rate adaptation in the return link of an S-band satellite communications system. The open loop Signal to Noise Ratio (SNR), measured directly by the mobile terminal in the forward link, was complemented with the closed loop SNR, allowing a faster response to changes in the channel conditions, since it has a much lower delay. The algorithm that balances the open and closed loop SNR by means of an adaptive weight, apart from using an adaptive margin, offered the best results in terms of the spectral efficiency. Furthermore, all the tested algorithms demonstrated to be able of following the variations of the terrestrial and aeronautical channel and guarantee a target Framer Error Rate (FER) at the same time that the spectral efficiency is maximized.

Chapter 3 addressed a practical issue which affects multibeam satellite systems which apply linear precoding to reduce the level of inter-beam interference which arises as a consequence of aggressive frequency reuse patterns. The miss-detection of interfering carriers causes a strong degradation on the

Channel State Information (CSI) vector, which has many unknown components which are nullified. Therefore, this lead to errors in the Signal to Interference and Noise Ratio (SINR) of the users that the gateway computes to allocate them a given Modulation and Coding Scheme (MCS). In their turn, these SINR errors cause a severe increment of the number of erroneous frames since MCS beyond the decoding capabilities of the users can be assigned. After characterizing the statistical and geographical distribution of the errors in the user SINR calculated by the gateway, the use of a link adaptation algorithm with an adaptive margin per user was proposed, in order to select the MCS in the presence of SINR errors. This helped to limit the rate of erroneous frames and to cause only a small loss of throughput compared with the systems with perfect CSI.

Chapter 4 dealt with the calculation of the capacity of Spatial Modulation (SM) and Generalized SM (GSM). A novel method, based on a single-hidden layer neural network, was proposed to make these information theory related calculations. With a proper selection of the neural network input features, the proposed method outperformed the analytical approximations existing in the literature for obtaining these capacities, both in terms of accuracy and computational complexity. Therefore, the rapid and precise calculation of the capacity of these schemes allows the implementation of adaptive systems, where the MCS is adapted according to the time evolution of the channel capacity.

Chapter 5 presented link adaptation techniques for Dual Polarization (DP) mobile satellite systems. An algorithm for performing the selection of the Multiple Input Multiple Output (MIMO) mode and the MCS was described. Several MIMO modes were considered, including SISO (single polarization); Orthogonal Polarization-Time Block Code (OPTBC), based on 2×2 Alamouti Space-Time Coding; Vertical-Bell Laboratory Layered Space-Time (V-BLAST), and Polarized Modulation (PMod), a particular case of 2×2 SM. Compared with a single polarization L-band system, simulation results show that despite the interference between polarizations, gains in the spectral efficiency between +30% and +100% can be achieved. Hence, the capacity can be doubled for high SNRs and the system can operate at lower SNRs compared with Single Polarization thanks to OPTBC.

Finally, Chapter 6 proposed two different coding rate adaptation mechanisms for SM links. In this way, the highly inefficient fixed coding rate solution is avoided and important gains in the spectral efficiency can be achieved with the help of the coding rate adaptation. It is shown how the capacity of the SM link calculated with a neural network can be used to select the coding rate, improving the performance with regard to a fixed coding rate alternative. However, the other method proposed for select-

ing the coding rate, based on deep learning, clearly outperforms the others, achieving a spectral efficiency very close to the maximum achievable value with a very low rate of erroneous frames. With this approach, a deep neural network is trained with data obtained from system level simulations and used later to perform the coding rate selection in the deployed system.

As to the future work, the results of Chapters 3-6, which are obtained by means of simulations, could be validated in an experimental setup with the help of SDR technology, in a similar way to the results provided in Chapter 2. In Chapter 4 capacity results are provided for SM and GSM systems by employing the same constellation in all the transmit antennas. However, the use of different constellations per antenna could be recommended in some unbalanced channels. The calculation of the capacity for this case with different constellations could be a possible enhancement of the results of that chapter. A maritime DP mobile satellite system is simulated in Chapter 5, where physical layer abstraction techniques are employed to ease simulations. It could be interesting to replicate other satellite scenarios apart from the maritime channel; in addition, the implementation of the whole transmit and receive chain, using some specific channel codes, could be the next step. Lastly, Chapter 6 simply dealt with the coding rate selection in a 2×2 SM system which employs a QPSK constellation. This study can be further extended to a system with a higher number of antennas, as for example 4 or 8, and where apart from the coding rate, the constellation order is also adapted. In this chapter, which applies Deep Learning to solve the rate adaptation problem, a supervised Learning approach is employed. However, it has the limitation that once the neural network is trained, it stops learning during the operation phase. Alternatively, Reinforcement Learning could be applied so that the neural network itself is adapted by the receiver during the operation, in order to make a better adaptation to the system working conditions.

Appendix A

Resumo en galego

Segundo a actualización de previsión de tráfico global de datos móbiles de Cisco [69], o tráfico de datos móbiles medrou un 71 % en 2017 e agárdase que no período 2017−2022 se incremente sete veces, co o tráfico medrando a unha taxa de crecemento anual composta (CAGR) do 46 % durante ese período. Polo tanto, cómpre aumentar a capacidade das redes móbiles para satisfacer todo o tráfico demandado polos usuarios. Ademais, as redes fixas, tanto terrestres como por satélite, tamén precisan mellorar o seu rendemento para ofrecer un maior rendemento aos usuarios, que demandan principalmente contido multimedia dunha calidade crecente.

Por outro lado, o consumo de enerxía do sector das tecnoloxías da información e as comunicacións (TIC) representaba xa ao redor de 2 % das emisións globais de carbono en 2007, coas redes móbiles contribuíndo preto de 0,2 % [31]. Por outra banda, espérase que a pegada das comunicacións móbiles triplicará ese valor en 2020 [36]. Ademais do punto de vista ecolóxico, o consumo de enerxía das estacións base (BS) que prestan servizo aos usuarios móbiles constitúe unha parte importante dos custos operativos dos provedores de telecomunicacións. Por todos estes motivos, é fundamental aumentar non só a capacidade das redes, senón tamén a súa eficiencia enerxética.

En canto ao novo tipo de dispositivos conectados, o citado libro branco Cisco prevé que o número de conexións Máquina a Máquina (M2M) verá un incremento de catro veces no período de 2017 − 2022, ata alcanzar os 3,9 millóns de dólares de conexións en 2022. Moitos dispositivos M2M da Internet das cousas (IoT) só funcionan con baterías que, ás veces, poden durar durante todo o ciclo de vida do dispositivo. Isto impulsa unha forte optimización do consumo de enerxía dos dispositivos, incluídas as súas

comunicacións.

O espectro é un recurso moi escaso que comparten moitos actores e sistemas diferentes. Polo tanto, ten que ser optimizado para ofrecer o mellor servizo posible co ancho de banda asignado. Unha métrica moi común para medir o aproveitamento do espectro é a eficiencia espectral, que relaciona o throughput transmitido en bits/s co ancho de banda ocupado en Hz. O obxectivo dos algoritmos de adaptación de enlace presentados nesta tese é precisamente o aproveitamento da capacidade das canles variantes de tempo para aumentar a eficiencia espectral do enlace de comunicacións.

Esta tese aborda entón a implementación de algoritmos para comunicacións adaptativas en tres escenarios diferentes, que teñen en común que son unha evolución futura dos sistemas actuais de comunicación terrestre e satélite. Estes escenarios son Sistemas de Satélite Móbil (MSS), Sistemas de Satélite Fixos (FSS) e redes 5G de próxima xeración. Nos tres escenarios futuros, a capacidade do sistema aumenta mediante diferentes técnicas. Nesta tese considérase a adopción de polarización dobre (DP) en MSS, o cambio a Full Frequency Reuse (FFR) en FSS e a adopción de técnicas multiantena con maior eficiencia enerxética en redes 5G. Este último escenario considera o consumo de enerxía das comunicacións 5G, non só o aumento da capacidade, importante para ampliar a duración da batería dos dispositivos IoT e reducir a pegada de carbono das estacións base 5G.

Nos seguintes parágrafos, explícase como poden evolucionar os sistemas terrestres e satélites para mellorar o seu rendemento. A explicación céntrase nestes tres escenarios que se tratarán máis adiante durante esta tese. Despois, dáse a motivación das técnicas de adaptación de enlace nas comunicacións, destacando que permiten aumentar a eficiencia espectral dos sistemas mediante unha mellor utilización da capacidade da canle. Finalmente, ofrécese unha visión xeral da tese.

A.1. Sistemas satélite e terrestres futuros

Os sistemas de comunicación evolucionan continuamente para atender as demandas da xente. Así, as transicións da segunda xeración (2G) á quinta xeración (5G) en tecnoloxías móbiles dixitais sucederon nos últimos trinta anos, pasando por 3G e 4G. Doutra banda, o consorcio DVB aprobou os estándares DVB-S, DVB-S2 e DVB-S2(X) para comunicacións por satélite no prazo de dezanove anos, desde 1995 a 2014. Os usuarios demandan conexións a Internet de alta velocidade e unha conectividade omnipresente; polo tanto, a capacidade das redes debería incrementarse para satisfacer toda a

demanda de tráfico e, ademais, os sistemas de comunicacións sen fíos teñen un papel fundamental para proporcionar unha conexión a Internet a dispositivos móbiles. A maiores, as comunicacións por satélite son necesarias para cubrir todos os lugares sen redes terrestres, como algunhas zonas rurais e terrestres, mar e aire. Por último, os sistemas de comunicación non só soportan as comunicacións das persoas, senón tamén as dun número cada vez maior de dispositivos conectados baixo o paradigma da IoT.

Hai varios xeitos de mellorar os sistemas de comunicación actuais para satisfacer o volume de tráfico cada vez maior. Esta tese centra a súa atención en tres melloras: unha para Servizos de Satélite Móbil (MSS) a baixas frecuencias, outra para MSS ou Fixed Satellite Services (FSS) que funcionan a frecuencias máis altas, como banda Ka, e outra consistente nun novo esquema de modulación proposto para redes 5G cunha maior eficiencia enerxética. Estes tres escenarios están representados na Figura A.1.

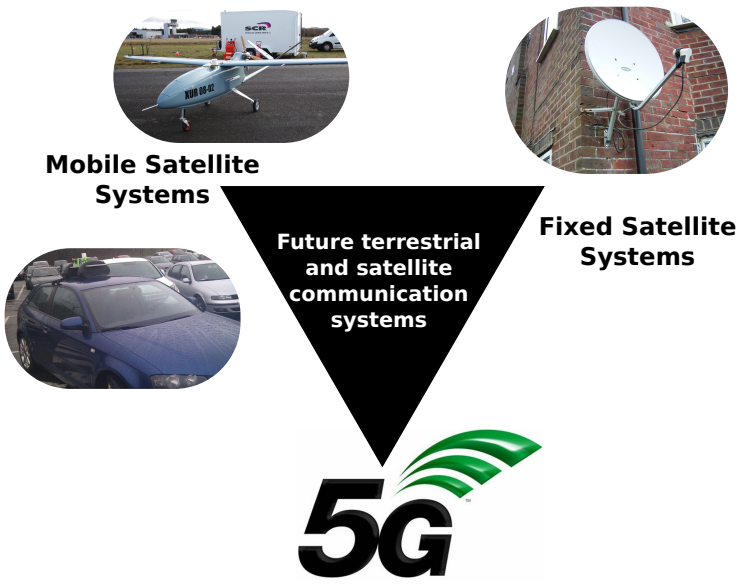


Figura A.1: Os tres escenarios de comunicacións abordados nesta tese.

En primeiro lugar, os MSS teñen unha importancia primordial para servir aos usuarios en movemento que se atopan en lugares fóra da cobertura dos sistemas terrestres. Estes poden ser, por exemplo, buques, avións comerciais, vehículos aéreos non tripulados (UAVs), vehículos de transporte ou usuarios da prensa ou ONGs. Os MSS funcionan en banda L (1-2 GHz) e banda S (2-4 GHz) e usan polarización normalmente circular, que se prefire en frecuencias baixas sobre polarización lineal para evitar o efecto da rotación Faraday. Os MSS, como os que se ofrecen co estándar da familia SL de S-UMTS [9], normalmente empregan unha polarización circular, normal-

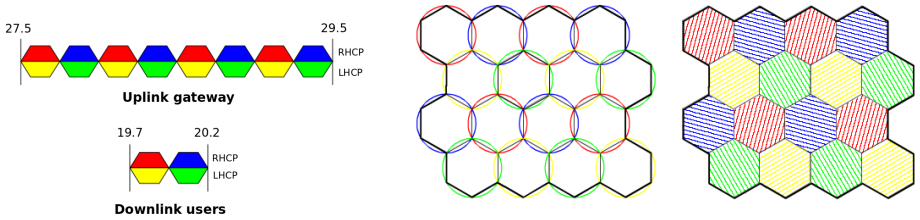
mente a polarización circular á dereitas (RCHP), debido ao temor á pequena Discriminación Cross-Polar (XPD) [75]. Non obstante, se se usan antenas de polarización dobre (DP) tanto no terminal de satélite como no terminal móbil (MT), pódense conseguir ganancias significativas na capacidade.

O uso simultáneo de dúas polarizacións ortogonais, RHCP e Left Hand Circular Polarization (LHCP) en MSS é posible grazas ao desenvolvemento de antenas de dobre polarización para estas bandas [74]. Un sistema de satélite DP que emprega simultaneamente tanto RHCP como LHCP para comunicarse cos usuarios é análogo a un sistema Multiple-Input-Multiple-Output (MIMO) 2×2, simplemente substituíndo a dimensión espacial pola compoñente de polarización. Polo tanto, a aplicación de técnicas de procesamento de sinal aos sistemas de satélite DP permite aumentar a súa capacidade, que se pode empregar para servir a máis usuarios ou aumentar o seu throughput, ou para ampliar o seu rango de funcionamento, grazas á diversidade de transmisión.

En [48], proponse o uso de DP para sistemas de satélite móbiles, con [47] analizando con máis detalle un sistema que explota un tipo de modulación específico chamado Polarized Modulation (PMod). Esta tese propón técnicas de adaptación para este novo sistema de satélite DP, con aplicacións a comunicacións móbiles. A adopción de DP pode duplicar a capacidade do sistema cando a Relación Sinal a Ruído (SNR) é alta e, por outra banda, pódese usar para ampliar o rango de SNR operativo, permitindo comunicacións robustas a SNRs máis baixas que nun sistema con polarización única (SP) e a mesma potencia de transmisión. As vantaxes que proporcionan os sistemas de satélite DP poden beneficiar tamén de dispositivos IoT situados en lugares remotos [28].

A frecuencias máis altas, en banda Ku e Ka, os satélites de alto throughput (HTS) [62] teñen como obxectivo proporcionar servizos de satélite fixo (FSS) de banda ancha, (BroadBand Satellite Service [BBS]) mediante moitos feixes puntuais na área de cobertura, proporcionando un sistema de gran capacidade e cun elevado throughput para os usuarios. Estes sistemas comezaron a despregarse cunha reutilización de frecuencias de catro cores (FR4), onde o ancho de banda dispoñible para os feixes do usuario divídese en dúas bandas de frecuencias e dúas polarizacións, RHCP e LHCP. A Figura A.2a amosa un exemplo de como o espectro dispoñible na banda Ka está asignado ao gateway e aos usuarios no enlace directo. Por outra banda, a Figura A.2b ilustra a pegada na terra dun satélite con 16 feixes de usuario, onde se pode ver como a mesma cor (é dicir, a mesma frecuencia e polarización), é reutilizada espacialmente para feixes non contiguos.

A reutilización de frecuencias de catro cores reduce a interferencia entre



(a) Plan de frecuenicas no enlace directo

(b) Cobertura do satélite

Figura A.2: Exemplo dun sistema satélite multi-feixe en banda Ka con reutilización de frecuencias de 4 cores.

feixes reutilizando a mesma porción do espectro en feixes separados xeograficamente; non obstante, en cada feixe só se emprega un cuarto do ancho de banda dispoñible. Esquemas de reutilización de frecuencias máis agresivas son posibles, como a reutilización de frecuencias de dúas cores (FR2) e a reutilización de frecuencia completa (FFR). Neste último, todo o espectro dispoñible en ambas as dúas polarizacións emprégase en todos os feixes. Polo tanto, a capacidade pode multiplicarse idealmente por catro. Non obstante, debido ao patrón de radiación das antenas do satélite, hai un alto nivel de interferencia entre feixes con FFR, reducindo así a capacidade total. Co fin de aproveitar os beneficios da FFR e reducir a interferencia, varias técnicas como a precodificación ou o acceso múltiple non ortogonal (NOMA) están a propoñerse [106], [73].

Varios traballos propoñen a aplicación de precodificación lineal ao enlace directo dos sistemas baseados en DVB-S2(X), como [103], [24], [91], [102] e [100]. Os sistemas satélite con precodificación requiren que os usuarios estimen a magnitude e fase do sinal desexado e os sinais interferentes dos feixes veciños e informen desta información de estado de canle (CSI) ao gateway. Con esta información, o gateway precompensa os sinais transmitidos para reducir a interferencia e mellorar a calidade do sinal que reciben os usuarios. Neste caso, o escenario satélite está modelado como unha canle MIMO multiusuario (MU-MIMO) e os fluxos de información de símbolos enviados a cada feixe multiplícanse previamente por unha matriz calculada coa CSI. Así, este proceso é transparente para os terminais de satélite, que só precisan estimar e informar da CSI. Nesta tese abordaranse os efectos que ten aa degradación da CSI e o xeito de contrarrestalos.

En canto ás comunicacións terrestres móbiles, nos últimos anos a comunidade investigadora dedicou un esforzo considerable á definición e desenvolvemento do 5G. No marco do 5G, prevese a existencia de diferentes tecnoloxías

de comunicacións para satisfacer os diversos requisitos de 5G establecidos pola Unión Internacional de Telecomunicacións (UIT). Estes abarcan, entre outros, conexións xigabit, moi baixa latencia, un enorme número de dispositivos conectados, redución do consumo de enerxía na rede en case 90 % e longa duración das baterías, especialmente para dispositivos de baixa potencia [10]. Neste contexto, unha nova familia de esquemas de modulación chamada Modulación Espacial, Spatial Modulation (SM), parecen proporcionar aumentos de capacidade con baixa complexidade de transmisión e alta eficiencia enerxética, para reducir o consumo tanto na estación base como nos terminais [61], [20].

SM é un esquema de modulación moi sinxelo que forma parte da ampla familia das Index Modulations (IM). SM é unha técnica de modulación multiantena que, ao contrario doutros esquemas como a multiplexación espacial (SMX), só require unha cadea de radiofrecuencia (RF) na súa forma máis sinxela. Aínda que con SMX pódese conseguir unha eficiencia espectral moito maior, xa que esta crece linealmente co número de antenas, SM ofrece un bo equilibrio entre a eficiencia espectral, a eficiencia enerxética e a complexidade. De aí que se propoña para as futuras interfaces radio dos dispositivos IoT [30]. A Figura A.3 mostra un diagrama de bloques dun transmisor SM, onde se pode observar que hai só unha cadea RF que está conectada por medio dun interruptor con todas as antenas.

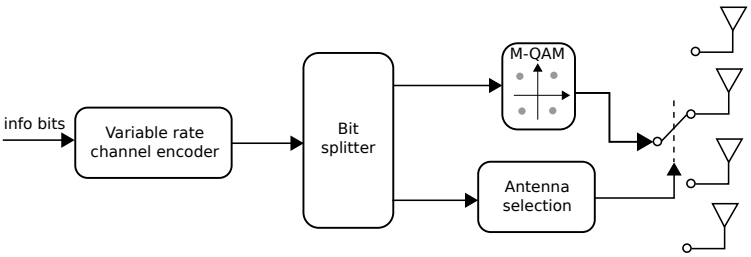


Figura A.3: Diagrama de bloques dun transmisor con Modulación Espacial (SM) adaptativo.

A pesar da súa sinxeleza, SM presenta algúns retos interesantes. Por exemplo, non existe unha expresión analítica para a súa capacidade, ao contrario de SMX, e a adaptación dalgúns parámetros de capa física, como a velocidade de codificación, non se estudou anteriormente. Esta tese presenta solucións para estes dous retos. Por unha banda, propón un novo método para obter a capacidade de SM baseado en Machine Learning (ML). E, por outra banda, propóñense varios métodos de adaptación da taxa de codificación do codificador de canle en SM.

Ata o momento expuxemos como evolucionan os sistemas de comunicacións terrestres e por satélite actuais para mellorar o seu rendemento e

satisfacer a esixente demanda dos usuarios. Nos sistemas de satélite móbiles que operan en bandas de baixa frecuencia, o salto dunha polarización simple á dobre permite mellorar a capacidade, moi necesaria nestas frecuencias tan conxestionadas. Tamén na comunicación por satélite, pero en bandas superiores, os HTS representan un intento da industria de satélites de ofrecer conexións de banda ancha aos usuarios. A adopción de patróns de reutilización de frecuencias máis agresivos nestes sistemas de satélites multi-feixe permite tamén incrementos de capacidade en comparación cos esquemas tradicionais de reutilización de catro cores. Por último, novos esquemas de modulación como SM axudan ás futuras redes 5G a alcanzar os seus obxectivos de eficiencia enerxética e mellora da capacidade para algúns casos de uso.

A.2. Adaptación de enlace

Esta tese céntrase no desenvolvemento de algoritmos de comunicacións adaptativas nos tres escenarios descritos nos parágrafos anteriores: MSS con DP, FSS con FFR e redes 5G que fan uso de SM. Ademais, ofrece unha validación experimental de algoritmos de adaptación de enlace para un MSS cunha única polarización mediante un satélite real de órbita media (MEO) e implementando os dous extremos da comunicación coa tecnoloxía Software Defined Radio (SDR).

Os sistemas de comunicación, especialmente os sen fíos, necesitan adaptar os seus parámetros de transmisión para ofrecer un bo rendemento, xa que calquera canle práctica é variante co tempo. A canle satélite móbil ten varios estados, con boas e malas condicións da canle, dependendo da boa ou mala visibilidade de satélite que ten o terminal móbil (MT) mentres este se move [38]. Ademais, se a órbita do satélite é Low Earth Orbit (LEO) ou MEO, as diferentes elevacións de satélite e a diferente distancia do satélite ao MT provocan variacións temporais do nivel do sinal recibido. En FSS que operan a frecuencias máis altas, como a banda Ka, a principal fonte de variación da calidade da canle é a atenuación da choiva que pode alcanzar valores de ata 12 dB [52]. Por último, nos sistemas celulares 5G, non só cambia o nivel do sinal recibido, senón tamén os camiños seguidos polos sinais desde o array de antenas de transmisoras ata as antenas receptoras, provocando variacións temporais da matriz de canle MIMO mentres que os usuarios están en movemento.

A relación sinal a ruído (SNR) ou a relación sinal a ruído máis interferencia (SINR) son métricas moi comúns para medir a calidade do sinal recibido. Defínense como a relación entre a potencia do sinal recibido e a

potencia de ruído (xunto coa potencia da interferencia, no caso da SINR). Da fórmula de capacidade de Shannon, que relaciona a SNR coa capacidade C nunha canle aditiva con ruído branco Gaussiano (AWGN),

$$C = B \log_2(1 + \text{SNR}) \quad (bits/s), \tag{A.1}$$

pódese ver claramente que unha variación na SNR recibida afecta á capacidade da canle, é dicir, á cantidade máxima de información que o emisor pode enviar de forma fiable a través da canle de comunicación. Para evitar a súa dependencia coa cantidade de ancho de banda B empregado polo sistema, é moi común empregar a eficiencia espectral η, que se obtén dividindo a capacidade polo ancho de banda:

$$\eta = \frac{C}{B} = \log_2(1 + \text{SNR}) \quad (bits/s/Hz). \tag{A.2}$$

Ademais, nas canles MIMO a capacidade depende tamén da matriz de canle $\mathbf{H}$ e as súas variacións tamén repercuten na capacidade final do enlace. Desta maneira, a capacidade dun enlace de comunicacións MIMO tamén é variante de tempo; polo tanto, a cantidade de información que se pode transmitir en sistemas multiantenna tamén cambia co tempo. Se un transmisor quere explotar completamente a canle, debería adaptar a velocidade de transmisión de bits segundo as variacións de capacidade da canle. A alternativa, o uso dunha taxa de bits fixa, é tremendamente ineficiente xa que para garantir a robustez da comunicación, a velocidade binaria de transmisión debe tomar un valor normalmente moi baixo, fixado polas peores condicións da canle. Polo tanto, cando a capacidade da canle é alta, a transmisión con esa taxa binaria baixa implica un malgasto da capacidade da canle, que podería usarse para proporcionar un enlace con maiores prestacións.

Un transmisor pode empregar diferentes mecanismos para adaptar a velocidade de bits de transmisión e seguir as variacións temporais da capacidade da canle. A familia SL do estándar S-UMTS [9] permite unha variación da potencia transmitida, a velocidade de símbolos (e polo tanto o ancho de banda), a orde de modulación e a velocidade de codificación. Noutros sistemas de satélite, como os que usan DVB-S2(X) [6], a velocidade de símbolo tende a ser fixa e o emisor selecciona o Esquema de Modulación e Codificación (MCS, Modulation and Coding Scheme) de cada trama, é dicir, a constelación e a taxa de codificación. Nalgúns sistemas terrestres, como LTE [8] ou Wi-Fi [5], ademais de seleccionar o MCS, o emisor elixe entre un dos modos MIMO de transmisión posibles ou o número de fluxos enviados.

O termo adaptación de enlace refírese ao procedemento de axuste dalgúns dos parámetros da capas física dun transmisor, como a modulación e a codificación, de acordo coa calidade do enlace, como se describe na Figura

A.4. Adaptive Coding and Modulation (ACM) é outro termo para referirse á adaptación da ligazón, destacando que estes son os parámetros que se adaptan. A maioría does estándares de comunicacións consideran algún tipo de ACM, desde sistemas terrestres como por exemplo Wi-Fi, WiMAX [3] ou LTE, ata sistemas vía satélite, como os que utilizan S-UMTS ou DVB-S2(X).

Con ACM, un transmisor pode seleccionar o nivel de protección dos bits de información dependendo da calidade da conexión. As constelacións de orde alta, como 64-QAM, ofrecen unha eficiencia espectral elevada que se pode usar cando a SNR do receptor é alta. Doutra banda, empregando constelacións de baixa orde, como QPSK, os bits de información están máis protexidos e pode garantirse unha comunicación robusta no caso dun enlace con baixa SNR. Ademais de variar a orde de modulación, o transmisor tamén pode cambiar a taxa de codificación do codificador de canle. Así, as taxas de codificación altas, próximas a 1, permiten enviar máis bits de información xa que os bits de redundancia son reducidos grazas ás boas condicións da canle. E, por outra banda, cando a calidade da canle non é tan boa, o transmisor pode reducir a taxa de codificación, o que implica aumentar o nivel de redundancia e protexer máis os bits de información a costa de reducir a eficiencia espectral. Se se emprega unha constelación de M símbolos e unha taxa de codificación r, a eficiencia espectral máxima η que se pode alcanzar é

$$\eta = r \log_2 M \quad (bits/s/Hz). \tag{A.3}$$

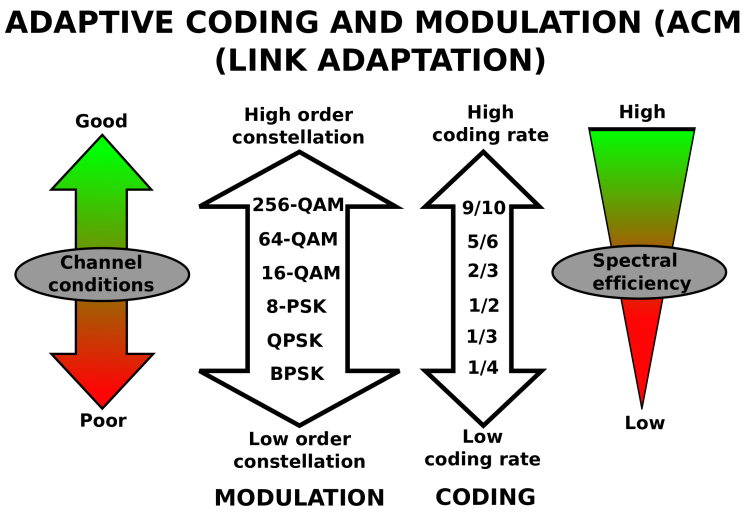


Figura A.4: Diagrama para explicar os principios detrás da adaptación de enlace.

A implementación dalgún tipo de ACM require normalmente algunha

realimentación do receptor ao transmisor. A información retroalimentada pode ser o resultado da decodificación da trama (un asentimento [ACK] se a trama se decodificou correctamente ou un ACK negativo [NAK] se o receptor non puido decodificar a trama), a SNR do enlace, algún indicador da calidade da canle (CQI, Channel Quality Indicator) ou o modo MIMO ou MCS que o emisor debe usar. Nalgúns casos, a adaptación pode realizarse sen retroalimentación se se asume un alto nivel de reciprocidade entre as canles dos enlaces directo e de retorno nunha comunicación bidireccional. Nesta tese, en xeral, consideramos que o receptor retroalimenta algo de información para axudar ao transmisor no proceso de selección do MCS.

Existen diferentes alternativas para realizar a adaptación do enlace. Un enfoque moi común denomínase Outer Loop Link Adaptation (OLLA), que consiste no uso dalgunhas táboas de busca (Look-up tables, LUTs), que mapean os rangos SNR ou un CQI a cada un dos MCS dispoñibles, xunto cunha marxe que se adapta en base aos ACK/NAK que o receptor retroalimenta. No algoritmo clásico de OLLA, a marxe increméntase cando ocorren erros no enlace e redúcese cando as tramas se decodifican correctamente. Este é o enfoque seguido en [16], [76], [78]. Algunhas melloras para o algoritmo clásico inclúense en [21] e [29] para mellorar a velocidade de converxencia do algoritmo.

En sistemas complexos con moitos graos de liberdade é común recorrer a algúns algoritmos de Machine Learning (ML). Por exemplo, en [79] emprégase unha máquina de soporte vectorial (SVM) para seleccionar os parámetros de capa física dun sistema MIMO-OFDM. Nese caso utilízase un enfoque de aprendizaxe supervisado. Máis recentemente, algúns traballos aplican o paradigma ML da aprendizaxe de reforzo, como [37] e [44]. Nesta tese aplícanse algoritmos baseados na aprendizaxe supervisada con redes neuronais e OLLA para a evolución futura dos sistemas de comunicación terrestres e por satélite. Un enfoque radicalmente diferente, que non se explora nesta tese, consiste na formulación da selección MCS como un problema multi-armed bandit, como se fai en [82].

Agora que explicamos o concepto de adaptación de enlace e os tres escenarios abordados na tese, os seus contidos resúmense a continuación. En primeiro lugar, esta tese comeza ofrecendo resultados dunha campaña experimental para validar uns algoritmos de adaptación de enlaces introducidos en [76]. Creouse unha enlace de comunicacións vía satélite bidireccional completo en banda S e, mediante un satélite MEO real que funcionaba en modo transparente, comunicouse un terminal móbil (MT) cunha estación terrestre (GS). Os transceptores implementáronse mediante SDR en dous Universal Software Radio Peripherals (USRPs). Con eles, probouse o rendemento de diferentes algoritmos de adaptación de enlace. Ademais de usar os ACK e

NAK recibidos do outro extremo, os algoritmos empregan diferentes estratexias para incluír a SNR de bucle pechado (a CSI retroalimentada dende o outro extremo) e a SNR de bucle aberto (a CSI medida directamente polo terminal) no proceso de adaptación.

Desde o punto de vista do usuario, a aplicación principal destes resultados é a adaptación automática da velocidade de bits dunha conexión de enlace por satélite móbil que pode beneficiar aos usuarios móbiles de sistemas de comunicacións por satélite. Nas probas de campo, probáronse tanto canles aeronáuticos como terrestres, co MT embarcado na parte superior dun coche e nun vehículo aéreo non tripulado (UAV). Aínda que non se probou, a MT tamén se podía situar nun barco para servir aos usuarios marítimos.

A aplicación de precodificación lineal para sistemas de satélite multifeixe proponse para aumentar a capacidade dos satélites de alto throughput (HTS) e competir mellor coas alternativas terrestres dos servizos de banda ancha. Ademais, só as comunicacións por satélite poden ofrecer conexión a Internet nos mercados marítimo e aeronáutico. Nos últimos anos houbo un intenso traballo de investigación no campo da precodificación para HTS [103], [24], [91], [102] e [100]. Non obstante, hai unha cuestión práctica que non se estudaba na literatura antes: como afecta ao sistema unha información de estado da canle (CSI) moi degradada. É dicir, caracterizamos estatísticamente e xeograficamente os erros na SINR dos usuarios calculadas polo gateway coa CSI imperfecta. A principal causa da degradación do CSI non son os erros de estimación, senón a anulación de moitas das compoñentes do vector de CSI, xa que os receptores de satélite só poden estimar unha pequena parte das canles interferentes dos feixes circundantes.

Os erros na SINR estimados polo gateway son moi problemáticos. Se a SINR é subestimada, o gateway asigna MCS cunha menor eficiencia espectral que o MCS óptimo para a condición da canle dos usuarios. E, aínda peor, se a SINR é sobreestimada, o gateway asigna aos usuarios MCS que non se poden decodificar correctamente, aumentando a taxa de tramas erróneas (Framer Error Rate, FER) e provocando unha caída considerablemente da capacidade do sistema. Nesta tese móstrase como superar este problema cun algoritmo de adaptación de enlace cunha marxe adaptativa por usuario. Así, a marxe adaptativa contrarresta os erros na SINR sen aplicar marxes fixas globais, que evitan os erros de decodificación pero a costa dunha elevada degradación do rendemento no sistema.

Nesta tese tamén se aborda o problema do cálculo da capacidade duns esquemas de modulación novos que se están a propoñer para os sistemas de comunicación de nova xeración, como a modulación espacial (SM) e a modulación espacial xeneralizada (GSM). O cálculo da capacidade ten un interese

práctico para desenvolver algoritmos de adaptación de enlace para futuros sistemas 5G que empregan estes esquemas de modulación. Se o transmisor coñece a capacidade do enlace SM ou GSM, pode adaptar a súa velocidade de transmisión de bits para facer unha mellor explotación da capacidade da canle. Para obter a capacidade destes sistemas proponse un novo enfoque que non se utilizou anteriormente para este propósito. Unha rede neuronal moi sinxela dunha soa capa oculta é adestrada para calcular a capacidade de SM e GSM coa axuda dalgunhas variables de entrada escollidas coidadosamente. Este método é máis preciso que as aproximacións analíticas anteriores da literatura, [49] e [42] e, ademais, implica unha complexidade computacional moito menor.

O cálculo da capacidade de SM ten aplicacións tanto en sistemas de comunicación terrestres coma por satélite. Por un lado, os dispositivos móbiles eficientes enerxeticamente e de baixa complexidade poden beneficiarse deste cálculo, permitindo a implementación de transmisores SM adaptativos. E, por outra banda, estes resultados pódense aplicar aos sistemas satélite móbiles con polarización dobre (DP) que empregan Polarized Modulation (PMod), un caso particular de SM 2×2 onde as dimensións representan diferentes polarizacións ortogonais.

O aproveitamento da dobre polarización nos sistemas de satélite móbiles que funcionan en bandas L e S permite aumentar a capacidade destes sistemas sen ampliar o ancho de banda, grazas á aplicación de técnicas de procesamento de sinal MIMO. Neste contexto, proponse un algoritmo de adaptación de enlace para seleccionar tanto o modo de transmisión MIMO óptimo como o MCS nun sistema de satélite móbil con DP. Ademais, para o caso particular dun sistema DP con PMod, unha nova arquitectura de adaptación proponse con dous codificadores de canle en paralelo con taxas de codificación independentes. Este tipo de adaptación é posible grazas ao método proposto anteriormente para calcular a capacidade de SM e PMod. Desde o punto de vista dun usuario, pódense proporcionar comunicacións de maior throughput a terminais móbiles terrestres, marítimos ou aéreos, e a velocidade da conexión optimízase en tempo real.

Por último, propóñense diferentes técnicas de adaptación de enlaces para seleccionar a taxa de codificación nun sistema SM. A selección da taxa de codificación baseada na capacidade SM calculada por unha rede neuronal permite unha gran mellora do rendemento e a robustez da comunicación en comparación cunha asignación de taxa de codificación fixa. Non obstante, o método que ofrece o mellor rendemento é un método de selección de velocidade de codificación baseado en aprendizaxe profunda (Deep Learning, DL), que se explica no último capítulo da tese. Así, os dispositivos móbiles das futuras redes 5G poden adaptar tamén a súa velocidade de bits de

transmisión cando empregan SM como esquema de modulación. Polo tanto, o mecanismo de adaptación baseado en DL pode beneficiar especialmente aos dispositivos IoT, ofrecéndolles melloras de rendemento pero balanceando a complexidade do transmisor e o consumo de enerxía.

Bibliography

[1] LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures. *ETSI TS 136 213 V8.8.0*, (2009-10).

[2] IEEE standard for air interface for broadband wireless access systems. *IEEE Std 802.16-2012 (Revision of IEEE Std 802.16-2009)*, pages 1 –2542, 17 2012. doi: 10.1109/IEEESTD.2012.6272299.

[3] IEEE standard for information technology–telecommunications and information exchange between systems local and metropolitan area networks–specific requirements part 11: Wireless lan medium access control (MAC) and physical layer (PHY) specifications. *IEEE Std 802.11-2012 (Revision of IEEE Std 802.11-2007)*, pages 1–2793, 29 2012. doi: 10.1109/IEEESTD.2012.6178212.

[4] Digital Video Broadcasting (DVB); Second Generation DVB Interactive Satellite System (DVB-RCS2); Part 1: Overview and System Level specification. *ETSI TS 101 545-1 V1.1.1*, (2012-05).

[5] IEEE Standard for Information technology– Telecommunications and information exchange between systems. Local and metropolitan area networks– Specific requirements–Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications–Amendment 4: Enhancements for Very High Throughput for Operation in Bands below 6 GHz. *IEEE Std 802.11ac-2013*, pages 1–425, Dec 2013. doi: 10.1109/IEEESTD.2013.6687187.

[6] Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications; Part 2: DVB-S2 Extensions (DVB-S2). *ETSI EN 302 307-2 V1.1.1*, (2014-10).

[7] Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications; Part 1: DVB-S2. *ETSI EN 302 307-1 V1.4.1)*, (2014-11).

[8] LTE; evolved universal terrestrial radio access (E-UTRA); physical layer procedures. *ETSI TS 136 213 V14.2.0*, (2017-04).

[9] Satellite component of UMTS (S-UMTS); family SL satellite radio interface. *ETSI TS 102 744*, Oct. 2015.

[10] M. Agiwal, N. Saxena, and A. Roy. Ten commandments of emerging 5g networks. *Wireless Personal Communications*, 98(3):2591–2621, Feb 2018. ISSN 1572-834X. doi: 10.1007/s11277-017-4991-8. URL `https://doi.org/10.1007/s11277-017-4991-8`.

[11] N. Ahad, J. Qadir, and N. Ahsan. Neural networks in wireless networks: Techniques, applications and guidelines. *Journal of Network and Computer Applications*, 68:1 – 27, 2016. ISSN 1084-8045. doi: https://doi.org/10.1016/j.jnca.2016.04.006. URL `http://www.sciencedirect.com/science/article/pii/S1084804516300492`.

[12] S. M. Alamouti. A simple transmit diversity technique for wireless communications. *IEEE Journal on Selected Areas in Communications*, 16(8):1451–1458, Oct 1998. ISSN 0733-8716. doi: 10.1109/49.730453.

[13] S. Andrenacci, S. Chatzinotas, A. Vanelli-Coralli, S. Cioni, A. Ginesi, and B. Ottersten. Exploiting orthogonality in DVB-S2X through timing pre-compensation. In *2016 8th Advanced Satellite Multimedia Systems Conference and the 14th Signal Processing for Space Communications Workshop (ASMS/SPSC)*, pages 1–8, Sept 2016. doi: 10.1109/ASMS-SPSC.2016.7601545.

[14] S. Andrenacci, D. Spano, D. Christopoulos, S. Chatzinotas, J. Krause, and B. Ottersten. Optimized link adaptation for DVB-S2X precoded waveforms based on SNIR estimation. In *2016 50th Asilomar Conference on Signals, Systems and Computers*, pages 502–506, Nov 2016. doi: 10.1109/ACSSC.2016.7869090.

[15] P.-D. Arapoglou, A. Ginesi, S. Cioni, S. Erl, F. Clazzer, S. Andrenacci, and A. Vanelli-Coralli. DVB-S2X-enabled Precoding for High Throughput Satellite Systems. *Int. J. Satell. Commun. Netw.*, 34(3):439–455, May 2016. ISSN 1542-0973. doi: 10.1002/sat.1122. URL `https://doi.org/10.1002/sat.1122`.

[16] J. Arnau, A. Rico-Alvariño, and C. Mosquera. Adaptive transmission techniques for mobile satellite links. In *30th AIAA International Communications Satellite Systems Conference (ICSSC)*, Ottawa, Canada, 2012.

[17] J. Arnau, A. Rico-Alvarino, and C. Mosquera. Adaptive transmission techniques for mobile satellite links. In *Proc. AIAA ICSSC*, Ottawa, Canada, Sept. 2012.

[18] N. Baldo and M. Zorzi. Learning and Adaptation in Cognitive Radios Using Neural Networks. In *2008 5th IEEE Consumer Communications and Networking Conference*, pages 998–1003, Jan 2008. doi: 10.1109/ccnc08.2007.229.

[19] E. Basar. Index modulation techniques for 5G wireless networks. *IEEE Communications Magazine*, 54(7):168–175, July 2016. ISSN 0163-6804. doi: 10.1109/MCOM.2016.7509396.

[20] E. Basar, M. Wen, R. Mesleh, M. D. Renzo, Y. Xiao, and H. Haas. Index Modulation Techniques for Next-Generation Wireless Networks. *IEEE Access*, 5:16693–16746, 2017. ISSN 2169-3536. doi: 10.1109/ACCESS.2017.2737528.

[21] F. Blanquez-Casado, G. Gomez, M. d. C. Aguayo-Torres, and J. T. Entrambasaguas. eolla: an enhanced outer loop link adaptation for cellular networks. *EURASIP Journal on Wireless Communications and Networking*, 2016(1):20, Jan 2016. ISSN 1687-1499. doi: 10.1186/s13638-016-0518-3. URL `https://doi.org/10.1186/s13638-016-0518-3`.

[22] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang. *MIMO-OFDM Wireless Communications with MATLAB*. Wiley Publishing, 2010. ISBN 0470825618, 9780470825617.

[23] D. Christopoulos, P.-D. Arapoglou, and S. Chatzinotas. Linear Precoding in Multibeam SatComs: Practical Constraints. In A. I. of Aeronautics and Astronautics, editors, *31st AIAA International Communications Satellite Systems Conference, International Communications Satellite Systems Conferences (ICSSC)*, 2013.

[24] D. Christopoulos, S. Chatzinotas, and B. Ottersten. Multicast multigroup precoding and user scheduling for frame-based satellite communications. *IEEE Transactions on Wireless Communications*, 14 (9):4695–4707, Sept 2015. ISSN 1536-1276. doi: 10.1109/TWC.2015.2424961.

[25] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, 1991.

[26] T. Cui, F. Lu, V. Sethuraman, A. Goteti, S. P. Rao, and P. Subrahmanya. Throughput optimization in high speed downlink packet access (HSDPA). *IEEE Journal on Wireless Communications*, 10(2): 474–483, 2011.

[27] R. C. Daniels, C. M. Caramanis, and R. W. Heath. Adaptation in convolutionally coded mimo-ofdm wireless systems through su-

pervised learning and snr ordering. *IEEE Transactions on Vehicular Technology*, 59(1):114–126, Jan 2010. ISSN 0018-9545. doi: 10.1109/TVT.2009.2029693.

[28] M. De Sanctis, E. Cianca, G. Araniti, I. Bisio, and R. Prasad. Satellite communications supporting internet of remote things. *IEEE Internet of Things Journal*, 3(1):113–123, Feb 2016. ISSN 2327-4662. doi: 10.1109/JIOT.2015.2487046.

[29] R. A. Delgado, K. Lau, R. Middleton, R. S. Karlsson, T. Wigren, and Y. Sun. Fast convergence outer loop link adaptation with infrequent updates in steady state. In *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, pages 1–5, Sept 2017. doi: 10.1109/VTCFall.2017.8288232.

[30] M. Di Renzo. Spatial modulation based on reconfigurable antennas: A new air interface for the IoT. In *MILCOM 2017 - 2017 IEEE Military Communications Conference (MILCOM)*, pages 495–500, Oct 2017. doi: 10.1109/MILCOM.2017.8170856.

[31] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo. Spatial Modulation for Generalized MIMO: Challenges, Opportunities, and Implementation. *Proceedings of the IEEE*, 102(1):56–103, Jan 2014. ISSN 0018-9219. doi: 10.1109/JPROC.2013.2287851.

[32] Z. Dong, J. Shi, W. Wang, and X. Gao. Machine Learning Based Link Adaptation Method for MIMO System. In *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 1226–1231, Sep. 2018. doi: 10.1109/PIMRC.2018.8580924.

[33] S. Elkawafi, A. Younis, R. Mesleh, A. Abouda, A. Elbarsha, and M. Elmusrati. Spatial Modulation and Spatial Multiplexing Capacity Analysis over 3D mmWave Communications. In *European Wireless 2017; 23th European Wireless Conference*, pages 1–6, May 2017.

[34] M. T. E. A. Elsoufi, X. Ying, W. Jun, and T. Bin. Fletcher-Reeves learning approach for high order MQAM signal modulation recognition. In *2016 7th International Conference on Information and Communication Systems (ICICS)*, pages 74–79, April 2016. doi: 10.1109/IACS.2016.7476089.

[35] Ettus Research, USRP E310. `www.ettus.com/product/details/E310-KIT.`, Accessed 2016-03-30.

[36] A. Fehske, G. Fettweis, J. Malmodin, and G. Biczok. The global footprint of mobile communications: The ecological and economic per-

spective. *IEEE Communications Magazine*, 49(8):55–62, August 2011. ISSN 0163-6804. doi: 10.1109/MCOM.2011.5978416.

[37] P. V. R. Ferreira, R. Paffenroth, A. M. Wyglinski, T. M. Hackett, S. G. Bilén, R. C. Reinhart, and D. J. Mortensen. Multiobjective Reinforcement Learning for Cognitive Satellite Communications Using Deep Neural Network Ensembles. *IEEE Journal on Selected Areas in Communications*, 36(5):1030–1041, May 2018. ISSN 0733-8716. doi: 10.1109/JSAC.2018.2832820.

[38] F. Fontan, M. Vazquez-Castro, C. Cabado, J. Garcia, and E. Kubista. Statistical modeling of the LMS channel. *IEEE Journal on Vehicular Technology*, 50(6):1549–1567, 2001. ISSN 00189545. doi: 10.1109/25.966585.

[39] K.-I. Funahashi. On the approximate realization of continuous mappings by neural networks. *Neural networks*, 2(3):183–192, 1989.

[40] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016. `http://www.deeplearningbook.org`.

[41] A. Guidotti and A. Vanelli-Coralli. Clustering Strategies for Multicast Precoding in Multi-Beam Satellite Systems. *CoRR*, abs/1804.03891, 2018. URL `http://arxiv.org/abs/1804.03891`.

[42] S. Guo, H. Zhang, J. Zhang, and D. Yuan. On the mutual information and constellation design criterion of spatial modulation MIMO systems. In *2014 IEEE International Conference on Communication Systems*, pages 487–491, Nov 2014. doi: 10.1109/ICCS.2014.7024851.

[43] S. Guo, H. Zhang, P. Zhang, D. Wu, and D. Yuan. Generalized 3-D Constellation Design for Spatial Modulation. *IEEE Transactions on Communications*, 65(8):3316–3327, Aug 2017. ISSN 0090-6778. doi: 10.1109/TCOMM.2017.2704578.

[44] T. M. Hackett, S. G. Bilén, P. V. R. Ferreira, A. M. Wyglinski, and R. C. Reinhart. Implementation of a space communications cognitive engine. In *2017 Cognitive Communications for Aerospace Applications Workshop (CCAA)*, pages 1–7, June 2017. doi: 10.1109/CCAAW.2017.8001607.

[45] M. T. Hagan and M. B. Menhaj. Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks*, 5 (6):989–993, Nov 1994. ISSN 1045-9227. doi: 10.1109/72.329697.

[46] J. R. Hampton. *Introduction to MIMO Communications*. Cambridge University Press, New York, NY, USA, 2014. ISBN 1107042836, 9781107042834.

[47] P. Henarejos and A. I. Perez-Neira. Dual Polarized Modulation and Reception for Next Generation Mobile Satellite Communications. *IEEE Transactions on Communications*, 63(10):3803–3812, Oct 2015. ISSN 0090-6778. doi: 10.1109/TCOMM.2015.2461221.

[48] P. Henarejos, A. Perez-Neira, N. Mazzali, and C. Mosquera. Advanced signal processing techniques for fixed and mobile satellite communications. In *2016 8th Advanced Satellite Multimedia Systems Conference and the 14th Signal Processing for Space Communications Workshop (ASMS/SPSC)*, pages 1–8, Spain, Mallorca, Sept 2016. doi: 10.1109/ASMS-SPSC.2016.7601468.

[49] P. Henarejos, A. Perez-Neira, A. Tato, and C. Mosquera. Channel Dependent Mutual Information in Index Modulations. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3261–3265, April 2018. doi: 10.1109/ICASSP.2018.8461424.

[50] K. Hornik. Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 4(2):251 – 257, 1991. ISSN 0893-6080. doi: https://doi.org/10.1016/0893-6080(91)90009-T. URL `http://www.sciencedirect.com/science/article/pii/089360809190009T`.

[51] Y. Jiang, M. K. Varanasi, and J. Li. Performance Analysis of ZF and MMSE Equalizers for MIMO Systems: An In-Depth Study of the High SNR Regime. *IEEE Transactions on Information Theory*, 57(4):2008–2026, April 2011. ISSN 0018-9448. doi: 10.1109/TIT.2011.2112070.

[52] José Miguel García-Rubia, J. M. Riera, P. García-del-Pino, and A. Benarroch. Propagation in the ka band: Experimental characterization for satellite applications. *IEEE Antennas and Propagation Magazine*, 53(2):65–76, April 2011. ISSN 1045-9243. doi: 10.1109/MAP.2011.5949328.

[53] F. Kaltenberger, I. Latif, and R. Knopp. On scalability, robustness and accuracy of physical layer abstraction for large-scale system-level evaluations of LTE networks. In *2013 Asilomar Conference on Signals, Systems and Computers*, pages 1644–1648, Pacific Grove, California, Nov 2013. doi: 10.1109/ACSSC.2013.6810578.

[54] J. Kassab and S. Nagaraj. Adaptive modulation in an OFDM communications system with artificial neural networks. In *2009 International Joint Conference on Neural Networks*, pages 1547–1551, June 2009. doi: 10.1109/IJCNN.2009.5178657.

[55] E. Lagunas, S. Andrenacci, S. Chatzinotas, and B. Ottersten. Cross-Layer Forward Packet Scheduling for Emerging Precoded Broadband

Multibeam Satellite System. In *2018 9th Advanced Satellite Multimedia Systems Conference and the 15th Signal Processing for Space Communications Workshop (ASMS/SPSC)*, pages 1–8, Sep. 2018. doi: 10.1109/ASMS-SPSC.2018.8510717.

[56] Q. Liu. Doppler measurement and compensation in mobile satellite communications systems. In *Military Communications Conference Proceedings, 1999. MILCOM 1999. IEEE*, volume 1, pages 316–320 vol.1, 1999. doi: 10.1109/MILCOM.1999.822695.

[57] R. Lopez-Valcarce and C. Mosquera. Sixth-order statistics-based non-data-aided snr estimation. *IEEE Communications Letters*, 11(4):351–353, April 2007. ISSN 1089-7798. doi: 10.1109/LCOM.2007.348298.

[58] M. Luise and R. Reggiannini. Carrier frequency recovery in all-digital modems for burst-mode transmissions. *IEEE Transactions on Communications*, 43(2/3/4):1169–1178, Feb 1995. ISSN 0090-6778. doi: 10.1109/26.380149.

[59] A. Marseet and F. Sahin. Application of complex-valued convolutional neural network for next generation wireless networks. In *2017 IEEE Western New York Image and Signal Processing Workshop (WNYISPW)*, pages 1–5, Nov 2017. doi: 10.1109/WNYIPW.2017.8356260.

[60] C. Meng, H. Wang, W. Heng, and J. Zhang. Physical layer abstraction algorithm based on RBIR for CDMA EVDO and WLAN. In *2011 International Conference on Wireless Communications and Signal Processing (WCSP)*, pages 1–4, Nov 2011. doi: 10.1109/WCSP.2011.6096776.

[61] R. Mesleh, H. Haas, C. W. Ahn, and S. Yun. Spatial Modulation - A New Low Complexity Spectral Efficiency Enhancing Technique. In *2006 First International Conference on Communications and Networking in China*, pages 1–5, Oct 2006. doi: 10.1109/CHINACOM.2006.344658.

[62] D. Minoli. *Innovations in Satellite Communication and Satellite Technology: The Industry Implications of DVB-S2X, High Throughput Satellites, Ultra HD, M2M, and IP*. 03 2015. doi: 10.1002/9781118984086.

[63] M. Mirmohammadsadeghi, S. S. Hanna, and D. Cabric. Modulation classification using convolutional neural networks and spatial transformer networks. In *2017 51st Asilomar Conference on Signals, Systems, and Computers*, pages 936–939, Oct 2017. doi: 10.1109/ACSSC.2017.8335486.

[64] M. M. A. Moustafa and S. H. A. El-Ramly. Channel estimation and equalization using backpropagation neural networks in OFDM systems. In *2009 IFIP International Conference on Wireless and Optical Communications Networks*, pages 1–4, April 2009. doi: 10.1109/WOCN.2009.5010528.

[65] T. L. Narasimhan and A. Chockalingam. On the Capacity and Performance of Generalized Spatial Modulation. *IEEE Communications Letters*, 20(2):252–255, Feb 2016. ISSN 1089-7798. doi: 10.1109/LCOMM.2015.2497255.

[66] V. Nikolaidis, N. Moraitis, and A. G. Kanatas. Statistical characterization of an urban dual-polarized MIMO LMS channel. *International Journal of Satellite Communications and Networking*, 36(6):474–488, 2018. doi: 10.1002/sat.1253. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/sat.1253`.

[67] M. Oerder and H. Meyr. Digital filter and square timing recovery. *IEEE Transactions on Communications*, 36(5):605–612, May 1988. ISSN 0090-6778. doi: 10.1109/26.1476.

[68] Omnispace. `www.omnispacellc.com/ico-f-2/`, 30 of March 2016.

[69] C. W. Paper. Cisco visual networking index: Global mobile data traffic forecast update, 2017-2022. February 2019. URL `https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-738429.pdf`.

[70] S. Park, R. C. Daniels, and R. W. Heath. Optimizing the target error rate for link adaptation. In *2015 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6, Dec 2015. doi: 10.1109/GLOCOM.2015.7417770.

[71] S. Park, R. C. Daniels, and R. W. Heath. Optimizing the target error rate for link adaptation. In *2015 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6, Dec 2015. doi: 10.1109/GLOCOM.2015.7417770.

[72] S. R and K. D. Modified EESM Link Adaptation Method with Multiple Constellation for Future Wireless Networks. *IJRCCT*, 2(7), 2013. ISSN 2278-5841. URL `http://www.ijrcct.org/index.php/ojs/article/view/278`.

[73] T. Ramírez, C. Mosquera, M. Caus, A. Pastore, M. Navarro, and N. Noels. Message-splitting for interference cancellation in multibeam satellite systems. In *2018 9th Advanced Satellite Multimedia*

Systems Conference and the 15th Signal Processing for Space Communications Workshop (ASMS/SPSC), pages 1–7, Sep. 2018. doi: 10.1109/ASMS-SPSC.2018.8510733.

[74] S. Richard, P. Markland, F. Lanciault, and V. Dupessey. High power dual-polarized, combined Tx/Rx feed for L- and S-band mobile satellite service antennas. In *32nd ESA Antenna Workshop on Antennas for Space Applications, Sesion 15, ESTEC*, 2010.

[75] M. Richharia. *Mobile Satellite Communications: Principles and Trends*. Wiley, 2nd edition, 2014.

[76] A. Rico-Alvariño, J. Arnau, and C. Mosquera. Balancing Closed and Open Loop CSI in Mobile Satellite Link Adaptation. In *Proc. ASMS & SPSC*, pages 226–233, Livorno, Italy, sep 2014.

[77] A. Rico-Alvarino, A. Tato, and C. Mosquera. Robust adaptive coding and modulation scheme for the mobile satellite forward link. In *Signal Processing Advances in Wireless Communications (SPAWC), 2015 IEEE 16th International Workshop on*, pages 530–534, June 2015. doi: 10.1109/SPAWC.2015.7227094.

[78] A. Rico-Alvariño, J. Arnau, and C. Mosquera. Link adaptation in mobile satellite links: Schemes for different degrees of CSI knowledge. *International Journal of Satellite Communications and Networking*, 2016.

[79] A. Rico-Alvariño and R. W. Heath. Learning-Based Adaptive Transmission for Limited Feedback Multiuser MIMO-OFDM. *IEEE Transactions on Wireless Communications*, 13(7):3806–3820, July 2014. ISSN 1536-1276. doi: 10.1109/TWC.2014.2314104.

[80] V. Saxena, B. Cavarec, J. Jaldén, M. Bengtsson, and H. Tullberg. A Learning Approach for Optimal Codebook Selection in Spatial Modulation Systems. In *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, pages 1800–1804, Oct 2018. doi: 10.1109/ACSSC.2018.8645407.

[81] V. Saxena, J. Jaldén, M. Bengtsson, and H. Tullberg. Deep Learning for Frame Error Probability Prediction in BICM-OFDM Systems. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6658–6662, April 2018. doi: 10.1109/ICASSP.2018.8461864.

[82] V. Saxena, J. Jaldén, J. E. Gonzalez, I. Stoica, and H. Tullberg. Constrained Thompson Sampling for Wireless Link Optimization. *arXiv e-prints*, art. arXiv:1902.11102, Feb 2019.

[83] K. Sayana, J. Zhuang, and K. A. Stewart. Link Performance Abstraction based on Mean Mutual Information per Bit (MMIB) of the LLR Channel. In *IEEE 802.16 Broadband Wireless Access Working Group*, 2007.

[84] K. Scharnhorst. Angles in complex vector spaces. *Acta Applicandae Mathematica*, 69(1):95–103, Oct 2001. ISSN 1572-9036. doi: 10.1023/A:1012692601098. URL https://doi.org/10.1023/A:1012692601098.

[85] N. Schorghofer. *Lessons in Scientific Computing: Numerical Mathematics, Computer Technology, and Scientific Discovery*. Routledge, 2018. ISBN 9781138070585. URL https://www.amazon.com/Lessons-Scientific-Computing-Norbert-Schorghofer/dp/1138070580?SubscriptionId=AKIAIOBINVZYXZQZ2U3A&tag=chimbori05-20&linkCode=xm2&camp=2025&creative=165953&creativeASIN=1138070580.

[86] SDR forum. www.sdrforum.org., Accessed 2016-03-30.

[87] G. A. Seber. *Multivariate Observations*. John Wiley & Sons, Inc., 1984.

[88] M. Sellathurai, P. Guinand, and J. Lodge. Space-time coding in mobile satellite communications using dual-polarized channels. *IEEE Transactions on Vehicular Technology*, 55(1):188–199, Jan 2006. ISSN 0018-9545. doi: 10.1109/TVT.2005.861195.

[89] O. Simeone. A Very Brief Introduction to Machine Learning With Applications to Communication Systems. *ArXiv e-prints*, Aug. 2018.

[90] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos. Learning to optimize: Training deep neural networks for wireless resource management. In *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–6, July 2017. doi: 10.1109/SPAWC.2017.8227766.

[91] G. Taricco. Linear Precoding Methods for Multi-Beam Broadband Satellite Systems. In *European Wireless 2014; 20th European Wireless Conference*, pages 1–6, May 2014.

[92] A. Tato and C. Mosquera. Deep Learning Assisted Rate Adaptation in Spatial Modulation Links. In *2019 16th International Symposium on Wireless Communication Systems (ISWCS)*, pages 176–181, Aug. 2019. doi: 10.1109/ISWCS.2019.8877161.

[93] A. Tato, C. Mosquera, and I. Gomez. Link adaptation in mobile satellite links: field trials results. In *2016 8th Advanced Satellite Multimedia Systems Conference and the 14th Signal Processing for Space Communications Workshop (ASMS/SPSC)*, pages 1–8, Sept 2016. doi: 10.1109/ASMS-SPSC.2016.7601547.

[94] A. Tato, S. Andrenacci, S. Chatzinotas, and C. Mosquera. Link Adaptation and Carriers Detection Errors in Multibeam Satellite Systems with Linear Precoding. In *2018 9th Advanced Satellite Multimedia Systems Conference and the 15th Signal Processing for Space Communications Workshop (ASMS/SPSC)*, pages 1–8, Sep. 2018. doi: 10.1109/ASMS-SPSC.2018.8510749.

[95] A. Tato, P. Henarejos, C. Mosquera, and A. Pérez-Neira. Link Adaptation Algorithms for Dual Polarization Mobile Satellite Systems. In P. Pillai, K. Sithamparanathan, G. Giambene, M. Á. Vázquez, and P. D. Mitchell, editors, *Wireless and Satellite Systems*, pages 52–61, Cham, 2018. Springer International Publishing. ISBN 978-3-319-76571-6.

[96] A. Tato, C. Mosquera, P. Henarejos, and A. Pérez-Neira. Practical Implementation of Link Adaptation with Dual Polarized Modulation. In *2018 11th International Symposium on Communication Systems, Networks Digital Signal Processing (CSNDSP)*, pages 1–6, July 2018. doi: 10.1109/CSNDSP.2018.8471820.

[97] A. Tato, S. Andrenacci, E. Lagunas, S. Chatzinotas, and C. Mosquera. Link Adaptation and SINR errors in Practical Multicast Multibeam Satellite Systems with Linear Precoding. *International Journal of Satellite Communications and Networking (under review)*, 2019.

[98] A. Tato, C. Mosquera, P. Henarejos, and A. Pérez-Neira. Neural Network Aided Computation of Mutual Information for Adaptation of Spatial Modulation. *arXiv e-prints*, art. arXiv:1904.10844, Apr 2019.

[99] A. Tato, C. Mosquera, P. Henarejos, and A. Pérez-Neira. Neural Network Aided Computation of Generalized Spatial Modulation Capacity. In *2019 27th European Signal Processing Conference (EUSIPCO)*, pages 1–5, Sep. 2019. doi: 10.23919/EUSIPCO.2019.8903104.

[100] M. Vazquez, B. Shankar, C. Kourogiorgas, P. Arapoglou, V. Icolari, S. Chatzinotas, A. Panagopoulos, and A. Perez-Neira. Precoding, Scheduling and Link Adaptation in Mobile Interactive Multibeam Satellite Systems. *IEEE Journal on Selected Areas in Communications*, 2018.

[101] M. Vazquez, M. R. B. Shankar, C. I. Kourogiorgas, P. Arapoglou, V. Icolari, S. Chatzinotas, A. D. Panagopoulos, and A. I. Pérez-Neira. Precoding, Scheduling, and Link Adaptation in Mobile Interactive Multibeam Satellite Systems. *IEEE Journal on Selected Areas in Communications*, 36(5):971–980, May 2018. ISSN 0733-8716. doi: 10.1109/JSAC.2018.2832778.

[102] M. A. Vazquez, A. Perez-Neira, D. Christopoulos, S. Chatzinotas, B. Ottersten, P. D. Arapoglou, A. Ginesi, and G. Taricco. Precoding in multibeam satellite communications: Present and future challenges. *IEEE Wireless Communications*, 23(6):88–95, December 2016. ISSN 1536-1284. doi: 10.1109/MWC.2016.1500047WC.

[103] O. Vidal, J. Lacan, J. Radzik, E. Alberty, and P. Inigo. Linear Precoding performance analysis in a Broadband satellite system with a 2-color dual-polarization reuse scheme. *31st AIAA International Communications Satellite Systems Conference, International Communications Satellite Systems Conferences (ICSSC), American Institute of Aeronautics and Astronautics*, 2013.

[104] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela. V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel. In *1998 URSI International Symposium on Signals, Systems, and Electronics. Conference Proceedings (Cat. No.98EX167)*, pages 295–300, Sep 1998. doi: 10.1109/ISSSE.1998.738086.

[105] P. Wu and N. Jindal. Coding versus ARQ in Fading Channels: How Reliable Should the PHY Be? *Communications, IEEE Transactions on*, 59(12):3363–3374, December 2011. ISSN 0090-6778. doi: 10.1109/TCOMM.2011.102011.100152.

[106] X. Yan, H. Xiao, C. Wang, K. An, A. T. Chronopoulos, and G. Zheng. Performance Analysis of NOMA-Based Land Mobile Satellite Networks. *IEEE Access*, 6:31327–31339, 2018. ISSN 2169-3536. doi: 10.1109/ACCESS.2018.2844783.

[107] J. Yang. *Multimodulus algorithms for blind equalization.* PhD thesis, University of British Columbia, Vancouver, BC, Canada, 1997.

[108] P. Yang, Y. Xiao, L. Li, Q. Tang, Y. Yu, and S. Li. Link Adaptation for Spatial Modulation With Limited Feedback. *IEEE Transactions on Vehicular Technology*, 61(8):3808–3813, Oct 2012. ISSN 0018-9545. doi: 10.1109/TVT.2012.2207973.

[109] P. Yang, Y. Xiao, B. Zhang, S. Li, M. El-Hajjar, and L. Hanzo. Star-QAM Signaling Constellations for Spatial Modulation. *IEEE Transactions on Vehicular Technology*, 63(8):3741–3749, Oct 2014. ISSN 0018-9545. doi: 10.1109/TVT.2014.2306986.

[110] P. Yang, M. D. Renzo, Y. Xiao, S. Li, and L. Hanzo. Design guidelines for spatial modulation. *IEEE Communications Surveys Tutorials*, 17 (1):6–26, Firstquarter 2015. doi: 10.1109/COMST.2014.2327066.

[111] P. Yang, M. D. Renzo, Y. Xiao, S. Li, and L. Hanzo. Design Guidelines for Spatial Modulation. *IEEE Communications Surveys Tutorials*, 17 (1):6–26, Firstquarter 2015. ISSN 1553-877X. doi: 10.1109/COMST.2014.2327066.

[112] T.-S. Yang and A. Duel-Hallen. Adaptive modulation using outdated samples of another fading channel. In *Proc. WCNC*, volume 1, pages 477–481, Orlando, Florida, USA, Mar. 2002.

[113] A. Younis, D. A. Basnayaka, and H. Haas. Performance analysis for generalised spatial modulation. In *European Wireless 2014; 20th European Wireless Conference*, pages 1–6, May 2014.

[114] G. Zafari, M. Koca, and H. Sari. Dual-Polarized Spatial Modulation Over Correlated Fading Channels. *IEEE Transactions on Communications*, 65(3):1336–1352, March 2017. ISSN 0090-6778. doi: 10.1109/TCOMM.2016.2643664.