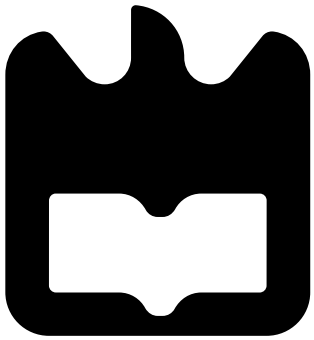
Universidade de Aveiro
Departamento de Electrónica, Telecomunicações e Informática,
2019

Tiago José Esteves Gonçalves Costa

Development of Probabilistic Constellation Shaping algorithms for 60 Gbaud metro networks

Desenvolvimento de algoritmos Probabilistic Constellation Shaping para redes metro a 60 Gbaud



Tiago José Esteves Gonçalves Costa

Development of Probabilistic Constellation Shaping algorithms for 60 Gbaud metro networks

Desenvolvimento de algoritmos Probabilistic Constellation Shaping para redes metro a 60 Gbaud

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requesitos necessários à obtenção do grau de Mestre em Engenharia Eletrónica e Telecomunicações, realizada sob a orientação científica de Professor Doutor António Luís Jesus Teixeira, Professor do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro e do Engenheiro Francisco Manuel Ruivo Rodrigues da empresa PICadvanced.

o júri / the jury

presidente / president

Professor Doutor João Nuno Pimentel da Silva Matos
Professor Catedrático da Universidade de Aveiro

vogais / examiners committee

Professor Doutor António Luís Jesus Teixeira
Professor Associado C/ Agregação da Universidade de Aveiro (orientador)

Doutora Maria do Carmo Raposo de Medeiros
Professora Associada da Universidade de Coimbra

agradecimentos / acknowledgements

É com muito gosto que aproveito esta oportunidade para agradecer a todos os que me ajudaram durante este percurso marcado por altos e baixos.

Aos meus pais, irmã, cunhado e sobrinho, obrigado por serem o apoio fundamental em todos os bons e maus momentos e por nunca terem desistido de mim, permitindo que crescesse com os meus erros. À Patrícia, o maior e especial agradecimento pela paciência e apoio e por ter estado sempre ao meu lado.

Ao Fábio, Pedro, Hugo e Miguel, obrigado por toda a força e motivação que me deram, não me deixando desanimar nos momentos baixos. Um obrigado também a todos os outros colegas e amigos presentes durante este percurso, pois fizeram com que este tempo fosse memorável.

Ao professor Mário Lima, ao Francisco Rodrigues e Ricardo Ferreira da PICAdvanced um obrigado pela forma exemplar como me orientaram durante este percurso.

Este trabalho é financiado pelo Fundo Europeu de Desenvolvimento Regional (FEDER), através do Programa Operacional Regional do Centro (CENTRO 2020) do Portugal 2020 [Projeto HeatIT com o nº 017942 (CENTRO-01-0247-FEDER-017942)]

Resumo

A necessidade de aumentar cada vez mais os ritmos de transmissão fazem das comunicacões óticas a tecnologia a adotar, sendo esta a unica tecnologia que, atualmente, permite transmissão de dados na gama dos Tb/s por canal. Várias novas tecnologias são testadas, sendo que o caminho passa por sitemas de modulação cada vez mais avançados.

Algoritmos de Probabilistic Constellation Shaping foram apresentados como uma possível tecnologia para o futuro, sendo que apresentam melhor desempenho, comparado com um sistema uniforme, principalmente quando associados a técnias de correção de erros como o FEC.

Nesta dissertação são estudados algoritmos de PCS, descrevendo todo processo desde geração, modulação, transmissão e receção de sinais. Foi feito um estudo de simulação usando 60 Gbaud, com bit rates entre os 400 e 600 Gbit/s, e técnicas de modulação QAM.

Esta dissertação visa demonstrar que PCS é um método viável e uma verdadeira alternativa a ser usada em futuras redes metro.

Abstract

The need to increase transmission rates makes optical communications the technology to adopt, being the only one that currently allows data transmission in the range of Tb/s per channel. Several new technologies are being tested, and the path goes through increasingly advanced modulation systems.

Probabilistic Constellation Shaping algorithms have been presented as a possible technology for the future, and they perform better especially when associated with error correction techniques such as FEC.

In this dissertation are studied PCS algorithms, describing the whole process from signal generation, modulation, transmission and reception. A simulation study was made using 60 Gbaud, with bit rates between 400 and 600 Gbit/s, and QAM modulation techniques.

This dissertation aims to demonstrate that PCS is a viable method and a true alternative to use in future metro networks.

Contents

List of Figures

List of Tables

Acronyms

AWGN	Additive White Gaussian Noise
ASK	Amplitude Shift Keying
BER	Bit Error Ratio
CAP	Cut and Paste
CCDM	Constant Composition Distribution Matching
DM	Distribution Matcher
DPSK	Differential Phase Shift Keying
DSP	Digital Signal Processing
ECS	Error Correction Schemes
EDFA	Erbium-doped fiber amplifiers
FEC	Forward Error Correction
FSK	Frequency Shift Keying
HiDM	Hierarchical Distribution Matching
LDPC	Low-density parity-check
LLR	Log-likelihood ratio
LUT	Look up Table
MI	Mutual Information
OH	Overhead
PCS	Probabilistic Constellation Shaping
PRBS	Pseudorandom Binary Sequence
PSK	Phase Shift Keying
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying

SE	System Efficiency
SNR	Signal to Noise Ratio

Chapter 1

Introduction

1.1 Context and motivation

In order to follow the most recent technologic improvements, it is important that communication systems can accommodate higher data rates. The main goal is to achieve spectral efficiency and maximum transmission rate in order to maximize all resources, including fiber infrastructure already in use, and not neglecting the fundamental standards for digital communications in order to maintain current level of reliability. To achieve this, the trend is to develop technologies with higher data rate using modulation formats successively more complex [1].

The integration of telecommunication networks, such as metro networks, and IT infrastructure is continuously growing. Being data centers one of the main blocks of the Internet, with large geographical distance to costumers, it is a huge need to develop interconnecting networks. With very short response times requirements of new technologies such as 5G, data centers tend to be smaller and more distributed in order to provide essential data and functionality closer to the customer, leading to increased data exchange between them. As a consequence, traffic volume is expected to grow much faster in the metro than in core networks. [2]

With the natural migration to high order speeds, it is crucial to have standardized organizations working together, as was achieved with 100G optical interfaces. The usable bandwidth of an optical communication system is effectively limited by the loss profile of the fiber and the erbium-doped fiber amplifiers (EDFAs) placed between every span. It is thus of high practical importance to increase the efficiency of optical fiber systems. [3].

Several improvements can be made in terms of error detection and correction algorithms, complex modulation formats, higher bandwidths or super-channel with advanced spectral shaping and multiplexing technologies. Although it is necessary to increase the complexity of nowadays systems, we can assume that with 400 Gbit/s we can have high quality communications over long distances [4].

Over the past few years, data rate in coherent optics has been increasing, with the launch of 600Gb/s Digital Signal Processing (DSP) algorithms. To achieve this data rate, the electronics and optics are operating at 60-64 GBaud with 64QAM modulation. [5]

It is here that arises the concept of Probabilistic Constaletaion shaping PCS and it has become the de-facto standard to overcome efficiency problems that come with the introduction of higher rates. Throughout the dissertation this concept will be properly scrutinized, focusing

on obtaining Bit error ratio (BER) values for different signal to noise ratio (SNR) over an Additive white Gaussian noise (AWGN) channel.

This technique gathered more interest in September 2016 when Nokia Bell Labs successfully demonstrated working 1 Tbit/s data transmission channels between German cities using 16QAM, 32QAM, and 64QAM and variable bandwidth with a maximum reach of 1500 km [6] .

1.2 Objectives

The main objective of this thesis is to analyze and demonstrate that PCS is a valid path to follow in order to efficiently use 60 Gbaud communications. This dissertation was divided into the following objectives:

- PCS concept presentation and discussion, approaching the most significant algorithms such as Constant Composition Distribution- CCDM and Cut and paste - CAP;
- Development and simulation, using MATLAB software, of CAP algorithm for different QAM modulations;
- Performance analysis of Bit Error Ratio- BER- for different modulations, comparing to uniform distribution systems.

1.3 Structure

- **Chapter 2: High capacity networks in optical communications**
  This chapter begins with an overview of digital communication systems. Begins with a description of types of modulation schemes used in the past and in nowadays communication systems. After the introduction of noise is described and a way to correct erros is introduced with Forward Error Correction - FEC. In the end, performance evaluation methods are introduced.
- **Chapter 3: Probabilistic Constellation Shaping**
  This chapter introduces PCS concept and gives theoretical background behind its concept. Cut-and-Past algorithm is analysed in more detail.
- **Chapter 4: Simulation and Results**
  This chapter will present the simulations of CAP algorithm, through the use of MATLAB software, which allows the simulation of different scenarios using QAM-ary modulation formats.
- **Chapter 5: Conclusions and future work**
  In this last chapter will be presented the main conclusions obtained from the performedwork. It ends with some suggestions for future research.

1.4 Main contributions

The main contributions of this work are:

- Demonstrating that PCS algorithms are a relevant technology for future optical metro networks;
- Performance analysis of BER vs SNR for QAM-ary modulation formats for the various simulation steps.

Chapter 2

High Capacity networks in optical communications

Massive improvements in the areas of high-speed electronics and optical component technologies have been made. From an engineering perspective, advances in highspeed electronics and its spectral efficiencies led to an adaptation of commonly used digital communication and signal processing techniques, such as advanced modulation, coding, and digital equalization. [7]

In this chapter a digital communication system will be discussed. Modulation systems, noise, error correction schemes (ECS) and performance evaluation will be analysed in depth.

Block diagram of a modulation system will be described from transmitter to receiver; Types of noise present in a simulation environmment, mainly additive white gaussian noise - AWGN will be presented; Forward error correction -FEC is introduced, focusing in binary block codem which was used in simulation; lastly, will be shown some criteria able to evaluate performance of a digital communication system.

2.1 Digital Communciation Systems

To design a digital communication system, many variables have to be considered, such as the number of signals to be transmitted and what code to use. At the end, decisions have to be made according with reliability, capacity and budget of the system [8].

The functional block diagram shown in 2.1 illustrates the signal flow and most significant signal processing steps through a typical digital communication system. Upper branch blocks denote signal transformations from the source to the transmitter, while lower branch blocks denote signal transformations from receiver to user, essentially reversing the signal processing steps from the upper brach.

The digital source data is encoded to eliminate as much unnecessary information as possible. The encoded binary data stream is then passed onto the channel encoder, which can take care of some missed data during source encoder, introducing some controlled redundancy into the data stream protecting it from channel induced errors. Also in this block encoded data stream can be grouped in segments of bits, called message or symbol. Finally, the modulator will map the digital symbols into appropriate representation according to which modulation system in use.

The communications channel is the physical/electrical medium through which the signal

goes from the transmitter to the receiver, which can be an optical fiber. Noise is introduced in this channel and will be discussed later.

The receiver will process the channel corrupted output and reverse the transmitters encoding and modulating procedures to extract an estimate of the originally transmitted data sequence. After receiver it is possible to evaluate how well reconstructed transmitted data was and estimate the probability of an error to occur. The probability of error at the receiver depends on many different factors as signal to noise ratio (SNR), modulation type and type of error detection used. [9]

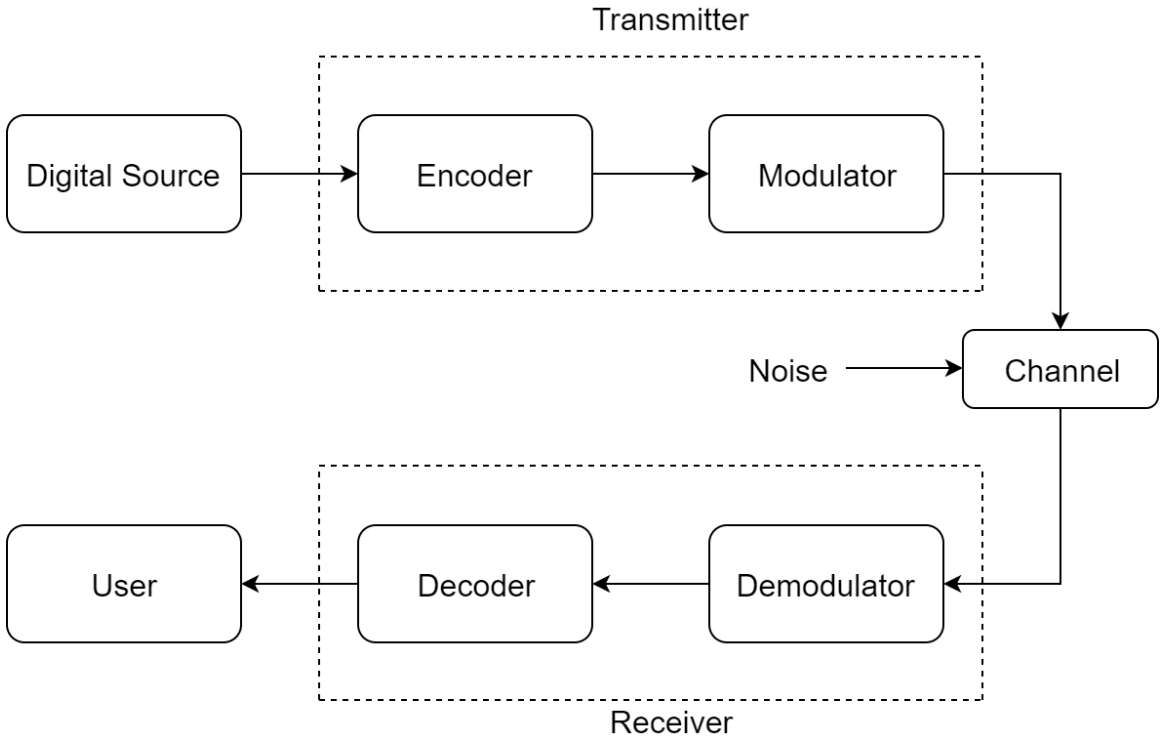


Figure 2.1: Typical digital communication system.

2.2 Digital Modulation techniques

Modulation is the process that promotes the variation of one or more characteristics of one signal (the carrier) according to the information signal to be transmitted. Modulation can be Baseband modulation, also called line coding; or Bandpass modulation, also known as carrier modulation [10].

A signal is baseband or low pass when frequency spectrum extends from (or near) DC up to some finite value, usually less than a few mega hertz, which is not appropriate for propagation through many transmission media and that is the case of system we want to use. Bandpass modulation is the process by which an information signal is transformed into a sinusoidal waveform[11].

A desirable modulation scheme must fulfill three important criteria:

- **Power efficiency:** the ability of a modulation technique to preserve the fidelity of the digital message at low power levels.

- **Bandwidth efficiency:** the ability of a modulation scheme to accommodate data within a limited bandwidth.
- **System complexity:** related with the technical difficulty of the system and associated with the cost.

2.2.1 Bandpass modulation

The baseband signals constitute the modulating signal and the high-frequency carrier signal is a sinusiodal waveform:

$$c(t) = A_c \cos(2\pi f_c t + \Theta_c(t)) \text{ [10]} \tag{2.1}$$

where: A_c represents the amplitude of the carrier;
f_c represents the frequency of the carrier;
θ_c represents the phase of the carrier.

After analysing this formula it is possible to understand that three parameters can be changed: amplitude, frequency and phase. There are three basic digital modulation techniques:

- **Amplitude Shift Keying ASK**;
- **Frequency Shift Keying FSK**;
- **Phase Shift Keying PSK**.

All of them can be represented as M-array modulations, which means that two or more bits are grouped together to form symbols and signals, and one of the possible signals obtained is transmitted. As we are dealing with binary codification $M = 2^n$, where n is an integer. For instance, M-ary Amplitude Shift Keying takes on M different levels. [10]

2.2.1.1 Amplitude Shift Keying - ASK

ASK was one of the earliest forms of digital modulation used in radio communications at the beginning of the nineteenth century. It has the simplest implementation, but is very susceptible to noise and distortion thats why is no longer used in nowadays modulation systems. The advantage of using ASK signals is that it does not require a large amount of bandwidth when transmitting the signal. However, it depends on the strength of the amplitude, which consumes more power. The general form of ASK can be expressed as:

$$s_i(t) = \sqrt{\frac{2E_i(t)}{T}} \cos(\omega_c t + \Phi) \quad 0 \leqslant t \leqslant T \;\; ; \;\; i=1,...,M \tag{2.2}$$

where $\sqrt{\frac{2E_i(t)}{T}}$ represents the amplitude with M discrete values and Φ is a random constant. [12]

In this technique, and assuming binary coding, the baseband signal is multiplied by a carrier frequency, thus, the output signal is null when binary 0 is transmitted and binary 1 is represented with A Watts. At the receiver, the demodulation can be done using a photodetector, which converts the signal from the optical domain to the electrical, resulting in the original transmitted pattern [13].

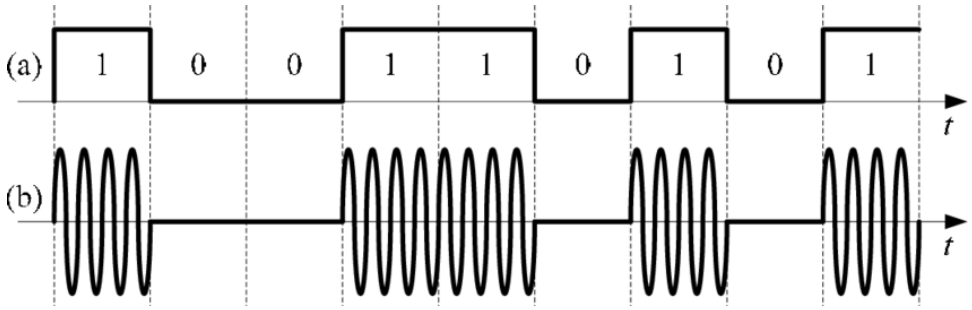


Figure 2.2: Example of ASK modulation format, (a) binary signal, and (b) ASK modulated signal [13].

2.2.1.2 Frequency Shift Keying - FSK

The general form of FSK modulation can be expressed as:

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos(\omega_i t + \Phi) \quad 0 \leqslant t \leqslant T \;\; ; \quad i=1,...,M \tag{2.3}$$

where the frequency component ω_i varies in time with discrete values M and amplitude and phase remains constant. Comparing to ASK it does not require a large amount of power when transmitting the signal. Nonetheless, it requires a large number of bandwidth to transmit the signal. [12]

The illustration of the Frequency Shift Keying modulation principle is shown in Figure 2.3 that represents the binary modulating signal and the modulated signal, which is the result of carrier frequency variation depending on the symbol to be transmitted.

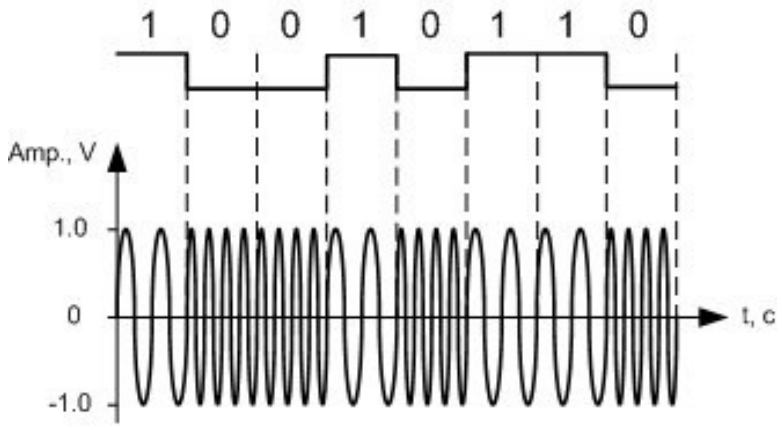


Figure 2.3: Example of FSK modulation format, (up) binary signal, and (down) FSK modulated signal [14].

2.2.1.3 Phase Shift Keying - PSK

PSK modulated signals are normally more immune to noise than ASK and FSK. Systems with M-array phase encoding (e.g. QPSK) assumes that the receiver has the perfect knowledge of the carrier phase. This underlying principle requires the receiver to have a minimal phase estimation error, which is very hard to achieve. A general analytic expression for PSK can be written as [12]:

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos(\omega_c t + \Phi_i(t)) \quad 0 \leqslant t \leqslant T \;\; ; \quad i=1,...,M \tag{2.4}$$

Where the phase component of the signal varies in time :

$$\Phi_i(t) = \frac{2\pi i}{M} \quad i=1,...,M \tag{2.5}$$

Example of BPSK modulation is shown in Figure 2.4. In this example, binary 1 is signed as $\sin(\omega t)$ and binary 0 is signed as $\sin(\omega t + \pi)$.

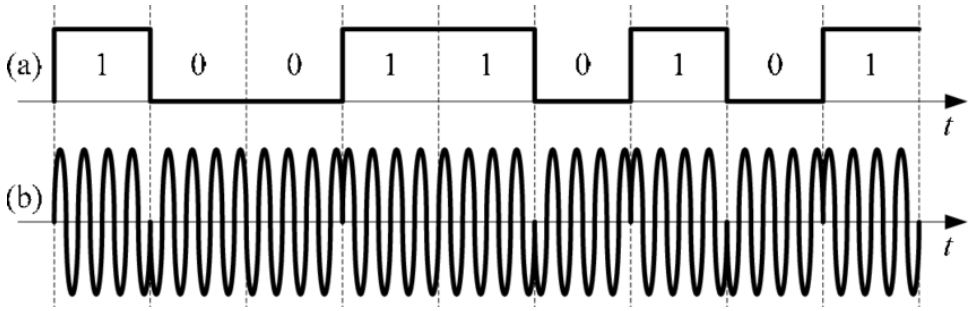


Figure 2.4: Example of PSK modulation format, (a) binary signal, and (b) FSK modulated signal [13].

Due to it's modulator complexity, PSK did not receive much attention in the early days, and Differential PSK (DPSK) was more significant. The term differential comes from symbol's similarity or difference when compared with the preceding symbol resulting in binary 1's or 0's. With this technique the data are first encoded differentially and then modulated onto optical carrier using a phase modulator, which externally changes the optical phase from its original phase to a relatively π phase shift. [13]

2.2.1.4 Quadrature amplitude modulation - QAM

M-QAM is a more generic modulation that includes M-ASK and M-PSK as special cases. In QAM modulation, the messages can be encoded either into amplitude or phase of the carrier. There are a finite number of M alternative symbols that is used to represent the changes in the carrier signal. QAM consists of two signals, I signal (upper brach in figure 2.5) referes to the one In-phase and quadrature phase is known as the Q signal (lower brach in figure 2.5) . The signals are summed,through amplitude modulation, and they form a two channel system. [15] As described in figure 2.5, phase shifters labeled as $-\pi/2$ are responsible for delaying input sinusoid by $-\pi/2$ rad. [16]

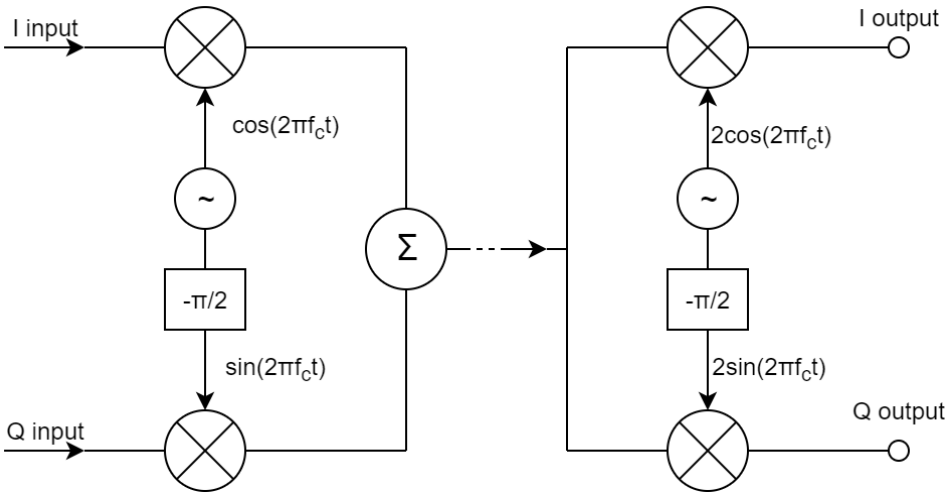


Figure 2.5: QAM modulation and demodulation system.

M-QAM constellations are two-dimensional given by [10]:

$$\begin{aligned} \phi_I(t) &= \sqrt{\tfrac{2}{T_s}}\cos(2\pi f_c t),\ 0 \leqslant t \leqslant T_s \\ \phi_Q(t) &= \sqrt{\tfrac{2}{T_s}}\sin(2\pi f_c t),\ 0 \leqslant t \leqslant T_s \end{aligned} \tag{2.6}$$

The output signal ($s_i(t)$) will be sum of 2.6:

$$s_i(t) = A_{I,i}\sqrt{\frac{2}{T_s}}\cos(2\pi f_c t) + A_{Q,i}\sqrt{\frac{2}{T_s}}\sin(2\pi f_c t) \quad 0 \leqslant t \leqslant T_s\ , i = 1, 2, ..., M \tag{2.7}$$

Alternatively, QAM can be described as:

$$s_i(t) = \sqrt{E_i}\sqrt{\frac{2}{T_s}}\cos(2\pi f_c t + \Theta_i) \tag{2.8}$$

where $E_i = \sqrt{A_{I,i}^2 + A_{Q,i}^2}$ and $\Theta_i = \tan^{-1}(\frac{A_{Q,i}}{A_{I,i}})$. From this equation it is possible to note that QAM is a combination of amplitude (E_i) and phase (Θ_i) modulation.[10]

Output signals of QAM modulation in the form of symbols constitute constelation diagrams represented by points in a both vertical and horizontal grid. In binary digital communication data those points can be the result of multiplications by 2 (2,4,8,16,32,etc.). The most common ones are 16, 64, 128 and 256 QAM [15].

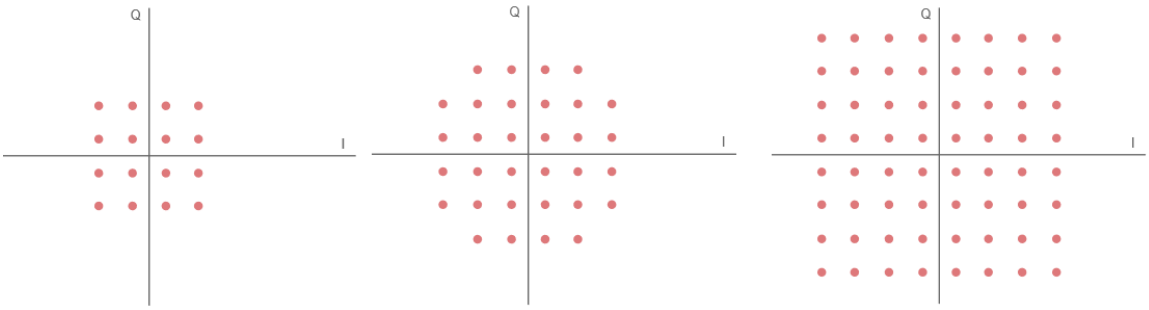


Figure 2.6: Constellation Diagrams for (a) 16-QAM, (b) 32-QAM and (c) 64-QAM .

The number of points present in a constellation represent the number of amplitude and phase combinations possible. To each combination is associated a number of bits, if are availabe N possible states, there are $\log_2 N$ bits. For 64 QAM, each symbol (each constellation point) correspond to a sequence of 6 bits.

Theoretically, higher order QAM modualtion will transfer higher number of bits per symbol. However, it will bring up reliability issues when applied in practice because points will be much more closer to one another and as it is expected to continue with same mean energy of the system as lower order costellations will lead to extra noise and distortion. [15]

2.3 Noise in Digital Systems

Even undesirable, noise is always present in digital communication systems. Any communication system, who sends information through a channel will be affected by noise, so it is always desirable to take it into account when designing a communication model in order to understand which are the system limitations and what measures need to be taken to overcome it.

For instance, in a single digital mobile system, a wide range of noise can be degrading the quality of communication, such as background noise, electronic device noise (e.g. thermal noise and shot noise), acoustic and line echoes, multipath reflections, fading, signal processing noise, among others. To successfully deal with noise it is necessary to use appropriate processing methods to differentiate the signal from the noise.

Based on its frequency spectrum or time characteristics, a noise process can be characterized following different features, such as white noise, narrowband noise, impulsive noise, among others. In the context of this dissertation, Additive White Gaussian Noise (AWGN) will be the most explored due to its simplicity and reliability to implement.[18]

To start, it is important to define white noise as a mathematical random process, that has an impulse autocorrelation and/or power spectral density function. White noise is a set of arbitrary uncorrelated time instant pairs, with a time averaged autocorrelation function, $R(\tau)$ (being τ Dirac Delta), and with a power spectral density function, $G(f)$, constant for all frequencies [19]:

$$R(\tau) = k\delta(\tau) \qquad G(f) = \eta/2 \tag{2.9}$$

It is important to note that white noise has infinite average power, and so it is physically impossible. Nevertheless, those simple mathematical properties make it easy and reliable method for implementation as statistical system analysis. [20]

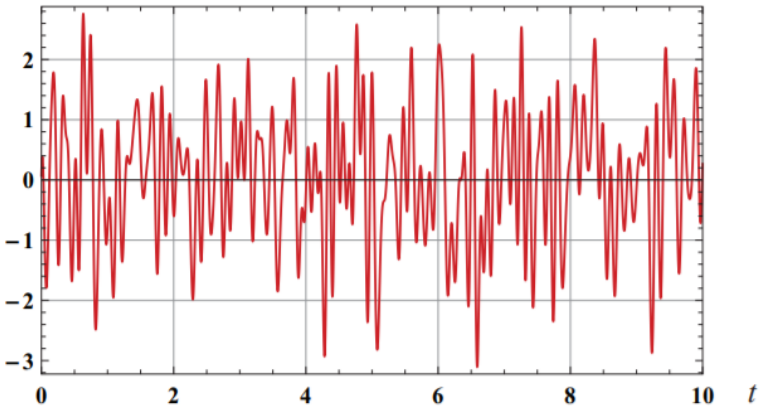


Figure 2.7: Graph of one signal from a white noise random process (yy axis refers to amplitude and xx axis refers to time) [19].

2.3.1 Additive White Gaussian Noise (AWGN)

The Gaussian noise is known to be uncorrelated with itself for any non-zero time offset τ :

$$E[n(t)n(t-\tau)] = \frac{N_0}{2}\delta(\tau) \tag{2.10}$$

and zero mean ($E[n(t)] = 0$). [21]

Being additive means that signal at the receiver is the sum of the transmitted signal plus noise, being the receiver responsible to recover, as best as it can, to the original message. As figure **??** shows, signal $y(t)$ is the result of the original signal $x(t)$ and noise $n(t)$. It is important to note that noise is statistically independent of the transmitted signal. [22]

The sample $y(t)$ at the receiver has a noise term, $n(t)$, contributing to it additively, where $n(t)$ is obtained from the following probability density function (PDF), which specifies a Gaussian distribution: [20]

$$f_W(n) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(n-\mu)^2}{2\sigma^2}} \tag{2.11}$$

Gaussian or Normal distribution is the given designation of a system in which noise can assume either negative, positive or null values, being the probability of near zero values much hihger than close to infinite values. A gaussian distribution graph is shown in figure 2.8 where it is possible to verify the distribution symmetry in relation with its mean (μ). [22]

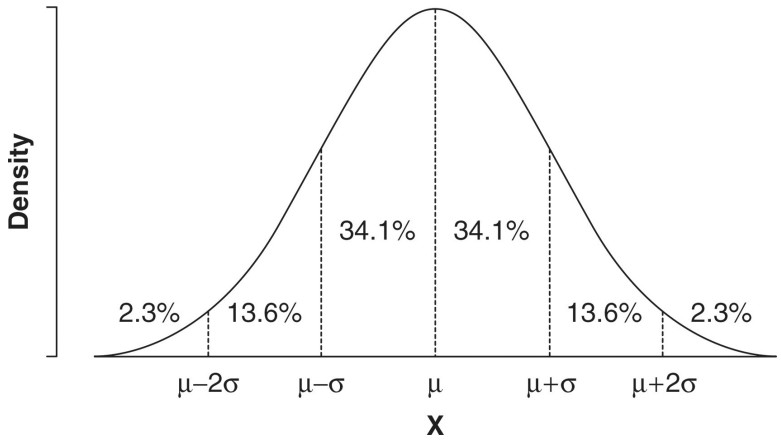


Figure 2.8: Gaussian Distribution model [23].

AWGN principle assumes that the main source of noise is at the receiver, that is independent of the paths over which the signal is being received. This assumption is normally used for most communication systems. [24]

2.4 Forward Error Correction

There are various types of Error correction schemes to support various types of reliability, delay constrained, and data rates in different types of communication systems. They can be loosely grouped in four categories. Figure 2.9 illustrates a broad classification of error correction schemes in order to demonstrate where FEC is categorized. [25]

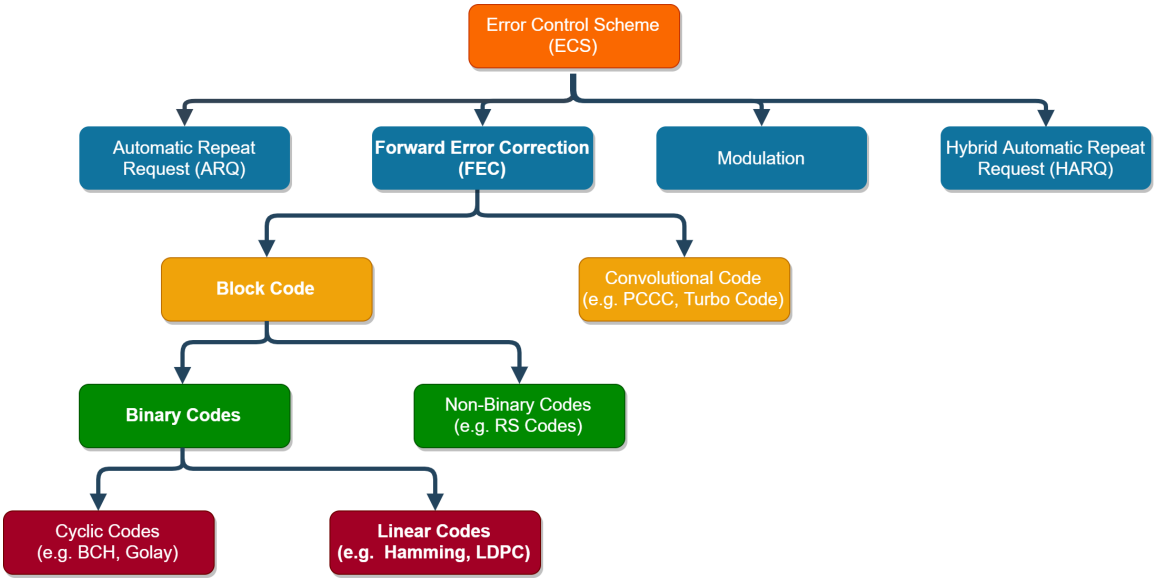


Figure 2.9: Classification of Error Correction Schemes (ECS) for wireless networks [25].

Forward error correction (FEC) or channel coding is a classical approach that is used to enhance the performance of communication link. It is normally used when retransmission is more expensive and delay sensitive. The insertion of FEC on a system grants it the advantage to operate at a significantly lower Signal to Noise Ratio (SNR) than an uncoded system to achieve a certain BER.[25]

As can be seen in figure 2.9, FEC can be implemented using block code (binary or non-binary) or convolutional code. For the interest of simulation process, binary block codes will be introduced.

The principle of FEC is the addition of a certain number of code bits in order to reduce the probability of error during signal transmission. For a signal length of k bits to be encoded, there must be $2k$ separate and orthogonal codes.

The difference in value of the individual codewords is defined by the Hamming distance (d), which gives the minimum number of different bits that codewords differ from eachother. The code capability to detect and correct errors increases with larger Hamming distances. $(d-1)/2$ gives the maximum number of errors that can be corrected and can be detected less than or equal to d errors. [26]

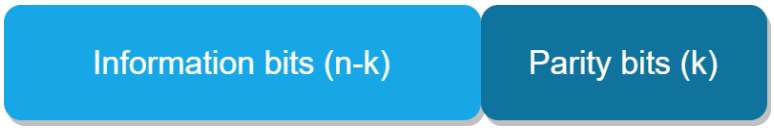


Figure 2.10: Block Code with codeword length n [25].

Focusing on block coding, it is commonly represented by the triple (n, k, t), where n is the length of code word bits, k is the length of information bits in the code word, and t is the error correction capability in term of number of bits that can be corrected. Figure 2.10 illustrates information bits and parity bits in block codes. The code rate (R) of ECC is defined as $R = \frac{k}{n}$. [25]

The parity-check codes use the combination suggested in 2.10 and can be represented as a matrix in 2.11 where each digit is the modulo 2 sum (1 if ordinary sum is odd and 0 if even) of a given set of infomation bits. This matrix represents a set of equations called parity-check equations and their soultions are the set of codewords. This equations provide the number of positions in which the two nearest code words differ. [27]

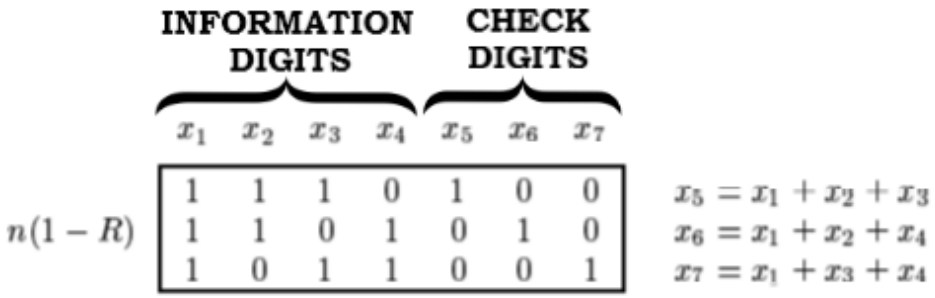


Figure 2.11: Example of parity-check matrix. [27].

Figure 2.12 shows a (n, j, k) matrix representation of LDPC where each column has a j number of 1's and each row contains k 1's and block code length N. The analysis of a low-density code of long block length is difficult because of the large number of code words involved. Hereby, the matrix is divided into j submatrices, each containing a single 1 in each column [27].

1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0
0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0
0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0
0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1
1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0
0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0
0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0
0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0
0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1

Figure 2.12: Example of parity-check matrix.; N=20, j=3, k=4 [27].

Decoding of parity-check codes is not simple to implement, however it can be simplified for a binary symmetric channel, which is a memoryless binary-input binary-output channel. Figure 2.13 shows a parity check tree where an arbitrary digit can be corrected even if its parity-check sets contain more than one transmission error.

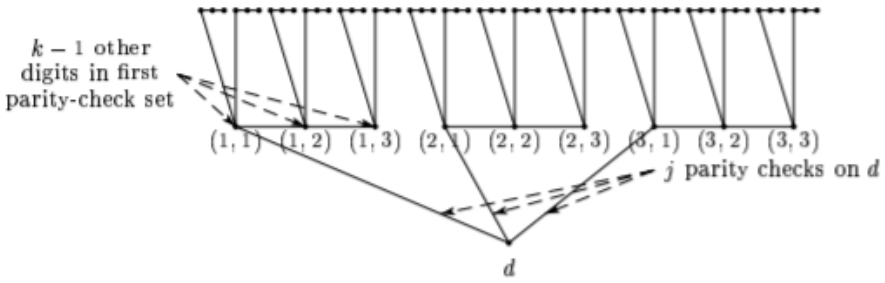


Figure 2.13: Parity-check tree [27].

A random digit (d) or bit is represented by the node at the base of the tree, and each line above this node represents one of the parity-check sets containing that digit. Assuming that both digit d and several other digits in the first tier are transmission errors, then the parity check represented by the lines connecting tiers will allow correction of errors from tiers above reaching to the final correct decoded digit d [27].

The most significant feature of this decoding scheme is that the digit computation per iteration is independent of block length. This is accomplished using log-likelihood ratio beetween transmitted and received bits [27]:

$$\ln\left[\frac{Pr[x_d = 0 \mid \{y\}, S]}{Pr[x_d = 1 \mid \{y\}, S]}\right] = \alpha\beta \tag{2.12}$$

where this gives the ratio beetween the probability of the transmitted digit in position d being 1 and the probability of a transmitted digit in position d being 0, conditional on the set of $\{y\}$ received symbols on the event S that the transmitted digits satisfy the parity-check

equations. The result is given in terms of $\alpha\beta$, being α the signal and β of the ratio. The signal would be positive if the probability of being zero is higher and negative for higher probability of one.

2.5 Performance evaluation of digital systems

Choosing the right performance evaluation model is one of the key factors to design efficient optical systems. Those evaluation models must provide precise measures of dominant system limitations, making them crucial for propagation disturbances reduction and performance improvement. The most common performance measures are BER and eye opening. [28]

2.5.1 Eye Diagram

Eye diagrams are multivalued displays used for measure the quality of high-speed digital signals [29]. Those measures are usually obtained applying multiple data signals from a receiver applied to the vertical input, while the data rate is used to trigger the oscilloscope horizontal sweep. Setting the display for infinite persistence will make remain on screen previous waveforms while subsequent ones are added creating a sample of waveforms for better analysis. [30]

Several system performance measures can be made, directly or indirectly, by analyzing an eye diagram display (figure 2.14) as signal duration or bit period, where the horizontal opening of an eye diagram at the crossing points of the eye is measured; noise can be perceived measuring eye height where the its ideal height will be equal to the eye amplitude, noise will cause the eye to close. The signal to noise ratio of the high speed data signal is also directly indicated by the amount of eye closure.

Jitter is one of the most important measures in high speed digital data and it consists in the time vatiation from the ideal timing of a data-bit event. The time deviations of the rising and falling edges of an eye diagram at the crossing point are measured to compute jitter. [28].

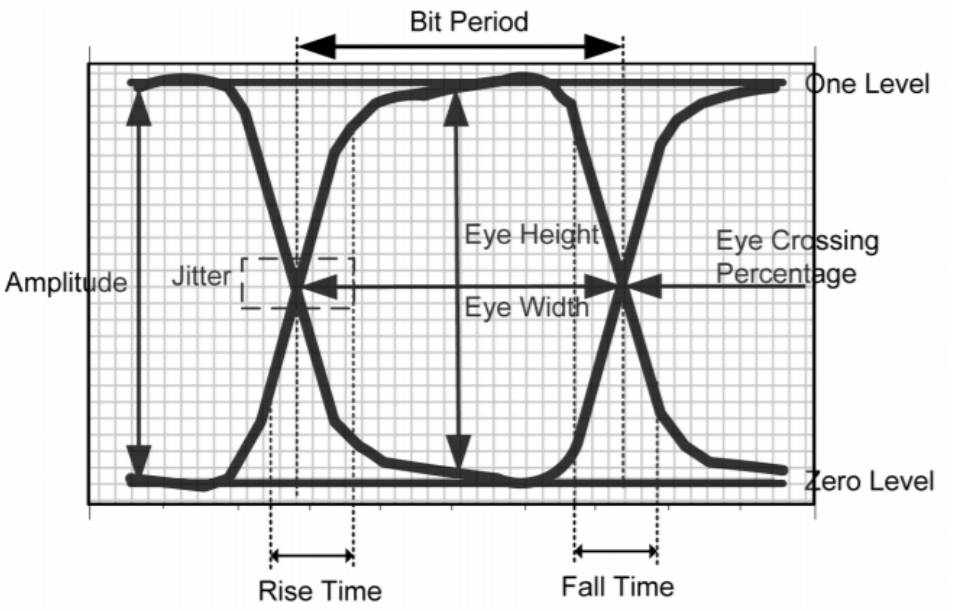


Figure 2.14: Typical Eye Diagram Measurements [31].

2.5.2 Signal to Noise Ratio - SNR

Signal to noise ratio (SNR) measures the ratio of signal power to the noise power, often expressed in decibels: [32]

$$\begin{aligned} SNR &= \frac{Signal\ Power}{Noise\ Power} \\ &= \frac{\frac{1}{T}\sum_{t=1}^{T}[(I_t)^2 + (Q_t)^2]}{\frac{1}{T}\sum_{t=1}^{T}[(n_{I,t})^2 + (n_{Q,t})^2]} \end{aligned} \tag{2.13}$$

where I_t and Q_t represent in-phase and quadrature signal amplitudes of the M-ary modulations, $n_{I,t}$ and $n_{Q,t}$ are the in-phase and quadrature noise amplitudes of the complex noise being considered. Most of the times these estimations can be simplified by measuring signal and noise power as ratio of variances, when both are zero mean processes.

For systems sampled at data rate, $\frac{E_s}{N_0}$, it is possible to obtain signal to noise ratio directly, where E_s is the symbol energy and $\frac{N_0}{2}$ is the noise power spectral density. For such systems (with M-array modulation), symbol energy can be defined as: [32]

$$E_s = \log_2 M E_{bit} \tag{2.14}$$

Average energy can also be obtained in order to evaluate performance, given from the reason between the sum of the energies of all constellation symbols and the number of input points.

In general a high Signal-to-Noise Ratio is good because it means the amount of noise is getting smaller comparing the actual signal. SNR is usually measured using a logarithmic scale, meaning its value is the logarithm of the actual ratio, in order to represent in decibel (dB)[33].

2.5.3 Bit Error Ratio - BER

Bit error ratio (BER) is the ratio between the number of error bits and total number of transmitted bits, unitless expressed. In practice, measuring BER is relatively simple: it's the comparison of input bit sequence with output sequence at the receiver, adding the number of bits that differ between these two sequences.

BER can also be defined in terms of the probability of error and each type of modulation has its own way to calculate it because each type of modulation performs differently in the presence of noise. For example, in the case of QPSK modulation and AWGN channel, the BER as function of the E_b/N_0 is given by:

$$P_{e_{QPSK}} = \frac{1}{2} erfc(\sqrt{\frac{E_b}{N_0}}) \tag{2.15}$$

And for M-QAM,the bit error probability can be defined as :

$$P_{\text{M-QAM}} = 1 - \left[1 - 2\left[1 - \frac{1}{\sqrt{M}}\right] Q\left[\sqrt{\frac{3\frac{E_s}{N_o}}{M-1}}\right]\right]^2 \tag{2.16}$$

The Probability of bit error is directly proportional to the distance between the closest points in the constellation diagram, which suggest that a modulation scheme with a thickly

packed constellation is less energy efficient than a modulation scheme with a more distributed constellation [34].

BER performance against SNR is the popular performance criterion that is used in today's communications systems and the most used throughout this dissertation.

In figure 2.15 it is possible to visualize that BER decreases with increasing SNR and moving to a higher order QAM scheme, it is possible to transmit more bits per symbol. However, in order to maintain the same SNR, as the symbol's constellation is larger, the constellation points must be closer together and are thus more susceptible to noise, resulting in a higher BER. [35]

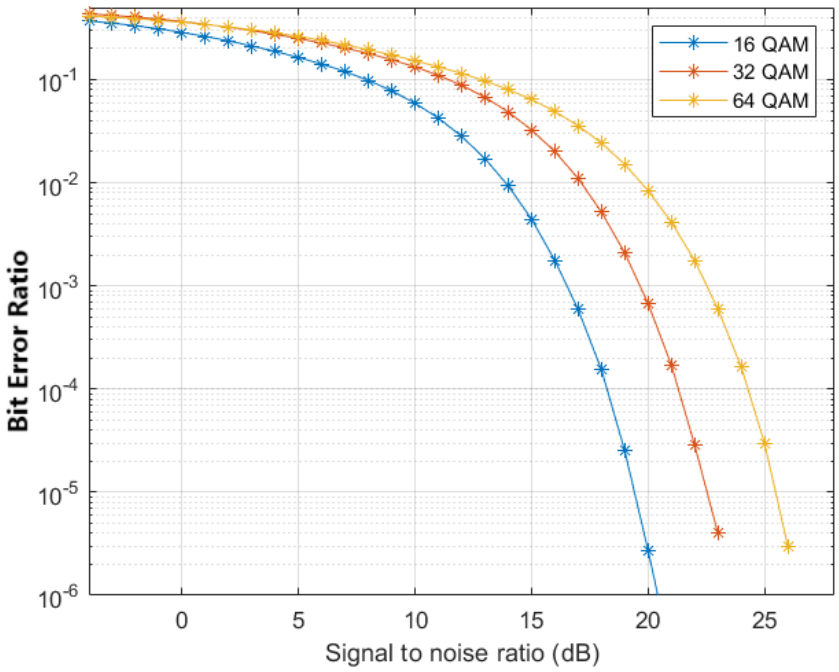


Figure 2.15: 16, 32 and 64-QAM BER vs. SNR comparision

2.5.4 Entropy and Mutual Information

The entropy of a random variable is a function which attempts to characterize the "unpredictability" of a random variable. If a random variable X takes on values in a set $X = \{x_1, x_2, ..., x_n\}$, and is defined by a probability distribution $P(X)$, then the entropy of the random variable will be [36]:

$$H(X) = -\sum_{x \in X} P(x) \log P(x) \tag{2.17}$$

Entropy, expressed in number of bits, returns the average amount of information (in bits) that must be delivered in order to resolve the uncertainty about the outcome of a trial. This is a lower bound on the number of binary digits that should, on the average, be used to encode a message.

If we send fewer binary digits on average, the receiver will have some uncertainty about the outcome described by the message. If more binary digits than required are sent, communications channel capacity is wasted with bits that don't need to be sent. Achieving the entropy lower bound is the optimal result for an encoding system. For uniform 64-QAM, entropy should be equal to 6, being that the optimal number of bits to be transmitted in order to recover to the correct symbol distribution.

Mutual information measures the relation between two random variables that are sampled simultaneously. In particular, it denotes the entropy (or uncertainty) of the random variable X after random variable Y is known: [37]

$$MI(X;Y) = \sum_{x \in X} \sum_{x \in X} P(x,y) \log \frac{P(x,y)}{P(x)P(y)} \tag{2.18}$$

The mutual information (MI) of a 2^{2m}-QAM scheme with uniform signalling can be computed as a function of the signal-to-noise ratio. This mutual information indicates what transmission rates can be realized with low error rates. In Figure 2.16, these rates are represented for m = { 1, 2, 3, 4} [38].

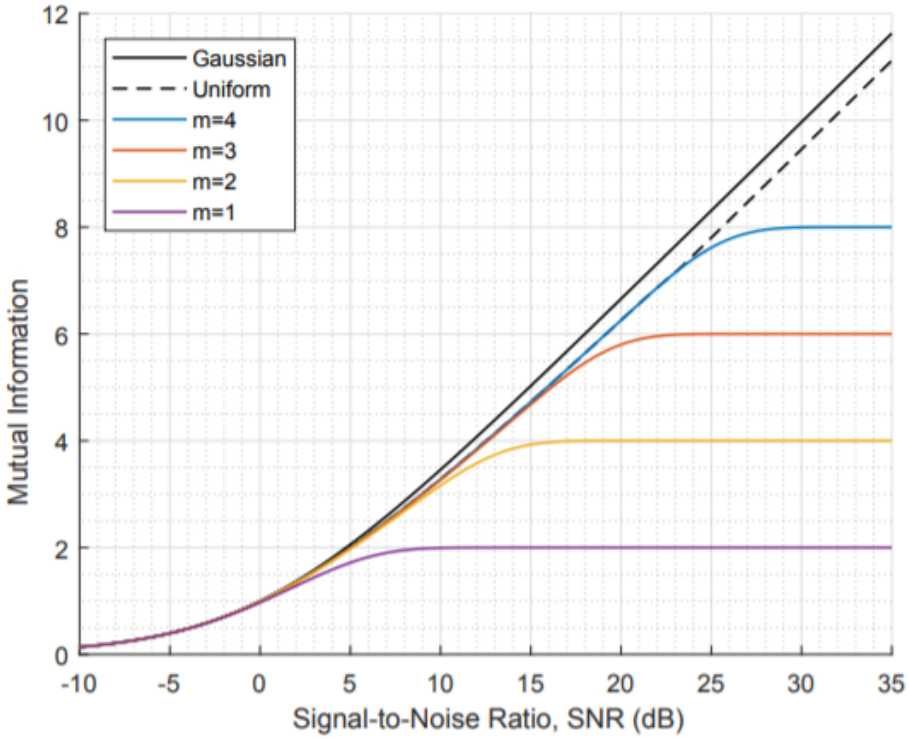


Figure 2.16: Mutual information for Gaussian input (solid), and MI for an uniformly distributed input (dashed), and the MI between channel input and output for 2^{2m}-QAM schemes where m = {1, 2, 3, 4}. [38]

2.5.5 Channel Capacity

In the case of an AWGN interference and according to Shannon's theorem, an ideal bandlimited channel of bandwidth W has a capacity C given by [37]:

$$C = W \log_2(1 + \frac{P}{WN_0})\ bits/s \tag{2.19}$$

where P is the average transmitted power and N_0 is the power-spectral density of the additive noise. We can also define it in terms of system efficiency (SE) for a given bandwidth B [39]:

$$SE = \frac{C}{B}\ \log_2(1 + SNR) \tag{2.20}$$

It is noticeable that if transmission rate R from the source is less than C ($R < C$), then it is theoretically possible to achieve reliable transmission through the channel by appropriate coding.[37] In figure 2.17 it is possible to visualize maximum obtainable channel capacity (in bits) for different QAM orders.

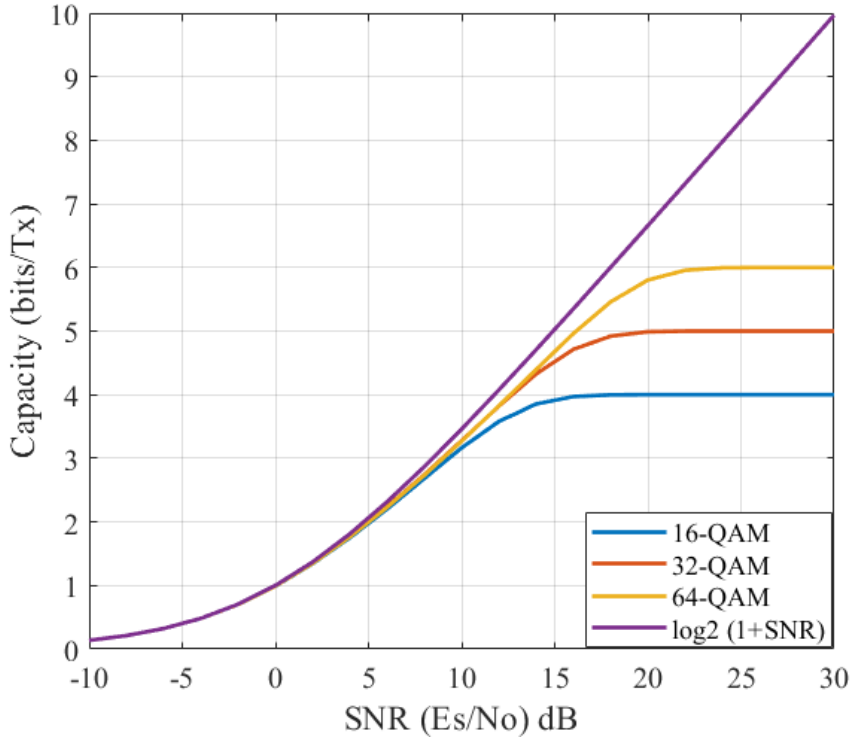


Figure 2.17: Channel capacity C for an AWGN channel and different M-ary constellations.

Chapter 3

Probabilistic Constellation Shaping - PCS

Throughout years, technologic evolution led to an huge increase in data traffic that optical core network was unable to follow. Optical communication systems in use have reduced bandwidth available due to loss profile of the fiber and the amplifiers between every span. The main goals to achieve for the next few years are spectral efficiency and maximum transmission rate in order to maximize all resources, mainly the fiber infrastructure already in use. However, from the joint optimization of these two requirements arises new challenges in terms of signal design. Therefore it is of great interest to introduce digital signal processing (DSP) techniques that are robust against fiber nonlinearities and also offer sensitivity and SE improvements in the linear transmission regime. [3]

It is in this context that arises the concept of signal shaping and it has become the standard to overcome the inherent gap to Shannon capacity of QAM.[41] Shaping can be separated in two types: geometric and probabilistic. In geometric shaping, each symbol probability remains the same and distance between symbols is modified, whereas in probabilistic shaping, the probabilities of each constellation point differs although it remains on an uniform grid.

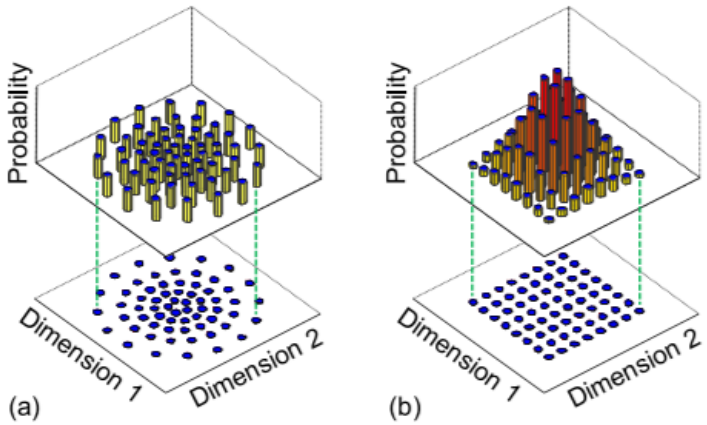


Figure 3.1: (a) Geometric and (b) probabilistic constellation shaping. [42]

Probabilistic shaping offers several advantages over geometric shaping. It gives higher shaping gains and grant rate adaptivity with iterative FEC decoding and demodulation. [3] Due to these advantages, work on this dissertation was restricted to probabilistic shaping.

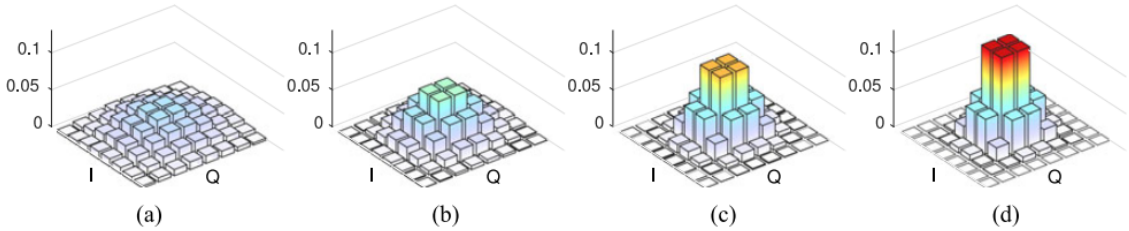


Figure 3.2: Graphical illustration of four different probability distributions for PS-64-QAM. The bars indicate the probability of each modulation symbol. From (a) to (d), the distributions become more shaped and the entropies H(Pi) decrease. [1]

3.1 PCS concept

As stated before, the main idea of probabilistic shaping is the transmission of standard QAM constellations points with different probabilities. Since optimal constellation is the Gaussian constellation itself and this is not implementable in practical communication systems due to its nonlinearities, QAM constellations are used even not optimal in an AWGN channel, however its simplicity and the possibility to directly apply I/Q modulation/demodulation makes it very interesting to use, besides being applied extensively to reach high-capacity and long-reach optical communications [43].

The nonuniform signaling reduces the entropy of the transmitter output, and consequently the average bit rate. However, if points with small energy are chosen more often than points with higher ones, energy savings may compensate for this loss in bit rate. These energy savings are called shaping gain. The maximum gain due to shaping can be achieved only for large constellations with an infinite number of dimensions: [44]

$$G_{s,max} = 10\log_{10}(\frac{\pi e}{6}) \ = 1.53 \ dB \tag{3.1}$$

As proved in [45], the shaping gain of a gaussian probability function, $P(x)$ is $\frac{\pi e}{6} = 1.423$, which converting to dB is 1.53 dB. This proves that the maximum possible shaping gain for any M dimensional constellation is 1.53 dB; and since equality holds when and only when the induced distribution $P(x)$ is Gaussian, and a constellation achieves this shape gain as $M \to \infty$, this argument also shows that the induced distribution of a constellation becomes Gaussian as $M \to \infty$, that is, becomes near optimal.

A schematic of a communication system employing probabilistic shaping is shown in figure 3.3. The information bits are first shaped by a distribution matcher, then encoded by a FEC encoder, before being mapped to symbols and transmitted over the channel. The process is reversed at the receiver, where symbols are de-mapped and FEC code is decoded. The shaped bits are then passed to a distribution dematcher to recover the original information bits. The noise-free shaped bits are then passed to a distribution dematcher to recover the original information bits

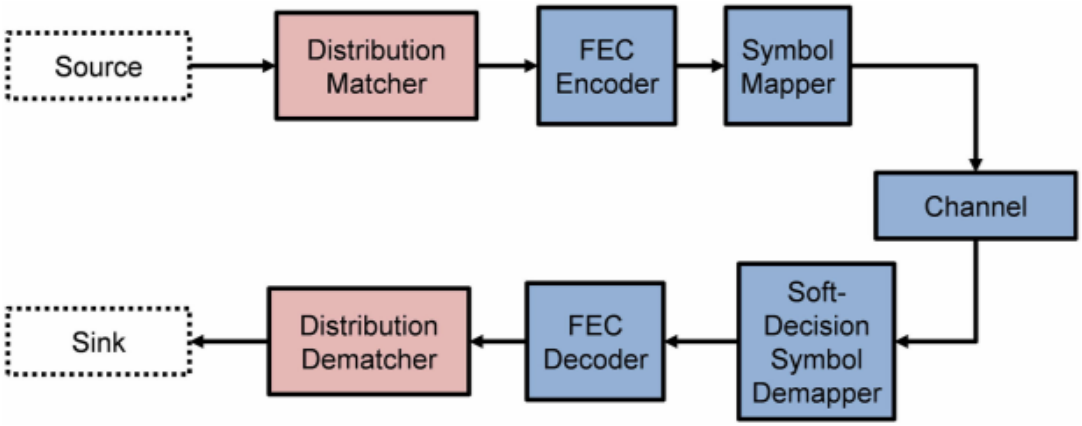


Figure 3.3: Schematic for a communication system employing probabilistic shaping. Blocks are represented in different colours depending on which blocks bits are subject to FEC [46]

The constellation points are normally represented as square QAM, instead of circular; constellations are assumed to be shaped in the amplitude domain only (this could be done in quadrature domain), and therefore they exhibit reflection between real and imaginary axes. This allows for the forward error correction (FEC) code to be applied to the stream of shaped bits, and the uniformly distributed parity bits assigned to the sign bits of the constellation – this configuration is also known as the "reverse concatenation" of FEC and distribution matching [46].

The major advantage of the proposed scheme is that its complexity is similar to most of communication systems used in many optical communications. The demapper only needs to be slightly modified taking in mind the initial probabilities. The receiver do not carry out iterations between demapper and decoder, which greatly simplifies the decoder implementation. The only additional element is the distribution dematcher, which, after all, only adds little complexity.

3.1.1 Maxwell-Boltzmann distribution

Probablistic shaping introduces a non uniform distribution and this is where the concept of Maxwell-Boltzmann distribution is based. A Maxwell-Boltzmann distribution can maximize bit rate for a fixed average energy or minimize average energy for a fixed bit rate, according to the desired optimization. The optimal distribution for a constellation Ω with a point r and energy $\|r\|^2$ can be defined as [47]:

$$p(r) \doteq \frac{\exp(-\lambda\|r\|^2)}{Z(\lambda)} \ , \quad \lambda \geqslant 0 \tag{3.2}$$

where $\lambda \geqslant 0$ is the arrangement between bit rate and average energy and the partition function $Z(\lambda)$ is responsible for normalize the distribution:

$$Z(\lambda) \doteq \sum_{r\in\Omega} \exp(-\lambda \|r\|^2), \lambda \geqslant 0 \tag{3.3}$$

It is important to note that with such a distribution, outer points (points with large energy) are never selected more often than inner points (points with small energy).

An essential property of the Maxwell-Boltzmann distribution is its "separability", where it can be separated into the product of various distributions over the factor constellations. In practice implementation, each factor constellation can have its own nonuniform distribution in order to obtain the optimal nonuniform signalling, which means that the ultimate shaping gain of 1.53 dB (3.1) can be reached in any dimension. [47]

3.1.2 Distribution Matching

A distribution matcher (DM) converts independent Bernoulli($\frac{1}{2}$) distributed input bits into output symbols with a wanted distribution. Inversely, a dematcher recovers the input bits from the output symbols.

A one-to-one fixed-to-fixed distribution matcher with an invertible function f is shown in figure 3.4. The mapping uses a desired distribution P_A by mapping m Bernoulli($\frac{1}{2}$) distributed bits $B^m = B_1, ..., B_m$ to strings $\tilde{A}^n = f(B^m)$ with length n. The output distribution is $P_{\tilde{A}^n}$ and data rate is $\frac{m}{n}[\frac{bits}{output\ \ symbols}]$. [48]

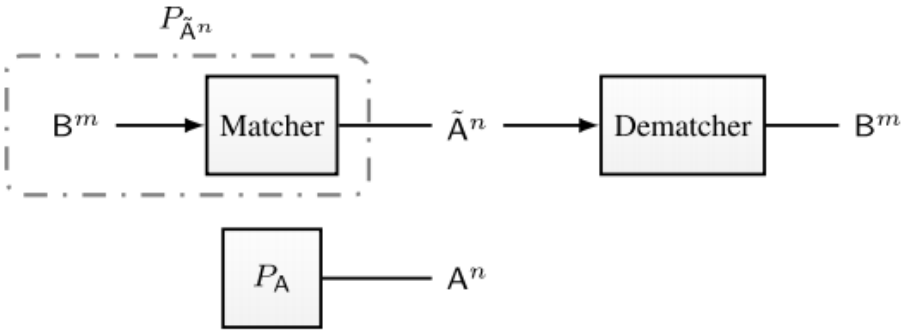


Figure 3.4: One-to-one fiber to fiber distribution matching diagram block [48]

For $N_{Seq} = 2k$ possible DM output sequences where k is the number of input bits given to the DM, with n symbol sequences, the rate loss of the DM can be defined as [46]:

$$R_{loss} = H(\tilde{A}) - \frac{k}{n} \tag{3.4}$$

where $H(\cdot)$ refers to the entropy in bits. The rate loss will be minimal for large n , which means that an infinite-length DM can achieve the target rate $H(\cdot)$. For distribution matching with fixed n, it is intended to make k as large as possible in order to minimize rate loss. [49]

3.2 Probablistic Constellation Shaping Methods

Besides de most common methods used, like Constant Composition Distribution Matching (CCDM) using Maxell Boltzman distribution, there are several new methods in development, like the HiDM and CAP. In this work it is proposed to study the general specification of Cut and Paste (CAP) method in more detail.

3.2.1 Constant Composition Distribution Matching - CCDM

Constant Composition Distribution Matching is as a method for mapping uniformly distributed bit sequences to shaped amplitude sequences and vice-versa. [46] CCDM method has an important role in optical communication because of its simple manipulation of input/output binary codes in Bernoulli($\frac{1}{2}$) distribution, demonstrating considerable nonlinear perturbation tolerance. [50]

Distribution of a vector c of length n is defined as:

$$P_{\tilde{A},c}(a) = \frac{n_a(c)}{n} \tag{3.5}$$

where $n_a(c) =\mid \left\{i : c_i = a\right\} \mid$ is the number of times a appears in c. [48]

The principle to choose $P_{\tilde{A}}$ (optimal aproximation of P_A) is to minimize divergence $\boldsymbol{D}((P_{\tilde{A}} \parallel P_A)$ and can be defined by the following expression:

$$\boldsymbol{D}(P_{\tilde{A}} \parallel P_A) < \log_2(1 + \frac{\mid \mathcal{A} \mid}{\min_{a \in suppP_A} P_A(a)n^2)}) \tag{3.6}$$

where $\mathcal{A}$ represents alphabet size of A variable, $suppP_A$ is part of $\mathcal{A}$ and n is output size of CCDM. [48]

A constant composition code is used with $n_a \approx P_A \cdot n$ wich means that multiple possibilities are possible for n_a. The actual mapping of CCDM encoding function with an output composition defined as C_{typ} can, for example, be carried out via arithmetic coding. The number of input bits of a particular C_{typ} is given by:

$$k = \lfloor \log_2 M(C_{typ}) \rfloor \tag{3.7}$$

where $\lfloor . \rfloor$ is the math representation for rounding down to the closest integer and M refers to multinomial coefficient [49]. The rate loss presented in 3.4 applies for CCDM which can disappear for large n, i.e., at an infinite-length it can achieve the target rate $H(\tilde{A})$ and it is therefore objective to minimize the rate loss reducing k.

CCDM is implemented using a data compression technique called arithmetic coding where the message begins with the representation of encode through a range of values between 0 and 1. Each time a new symbol is represented, i.e. each iteration, the interval needed to represent it is subdivided in smaller ones. As number of transmitting bits increase, interval is becoming narrower, wherein this reduction is settled by transmission probabilities of each symbol [51]

For example, supposing the alphabet is $\{a, e, i, o, u, !\}$, to transmit $'ea'$ sequence, both encoder and decoder start with the range [0, 1], after $'e'$ the encoder narrows its range to a smaller probabibility and even more tight when $'a'$ is transmitted, as illustrated in figure 3.5. [51]

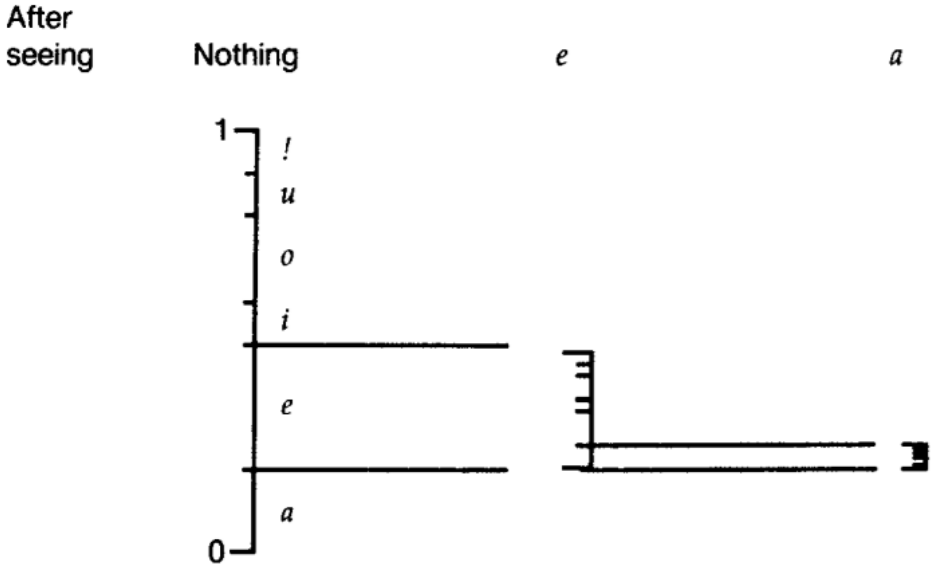


Figure 3.5: Arithmetic Coding Process example [51]

Figure 3.6 shows an example of CCDM, with $m = 2$ input bits and $n = 4$ output bits. Bit transmission '0' and '1' are considered to be equitable events ($P_{\tilde{A}}(0) = P_{\tilde{A}}(1) = 0.5$) as long as 4 equally probable input sequences and 6 equally probable output sequences. Each input sequence is linked with output if the lower border of the output interval is inside the input interval. For instance, '10' can both link with '1001' and '1010', while for '01' only a link to '0110' is possible. At most there are two possible links when the input interval size is less than twice the output interval size. Both can be selected, but in this example lower bound sequences were chosen. C_{typ} is $\{'0011', '0110', '1001', '1100'\}$. Being $\mathcal{T}^n_{P_{\tilde{A}}}$ the subset of bits contained by the symbol sequence, it is not possible to index the whole set unless it is equal to 2^m : [48]

$$2^m \leqslant \mid \mathcal{T}^n_{P_{\tilde{A}}} \mid < 2^{m+1} \tag{3.8}$$

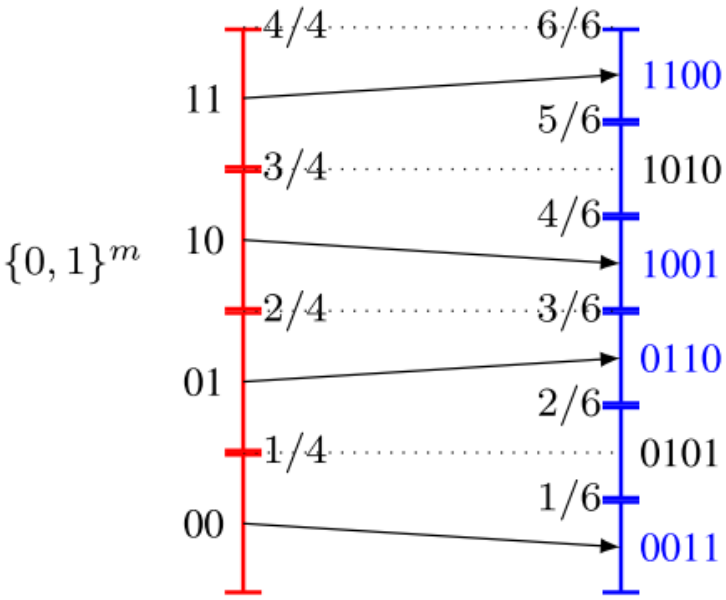


Figure 3.6: Diagram of a constant composition arithmetic encoder [48]

It is now possible to understand that CCDM is a powerful technique to obtain optimal performance as long as output block size n is getting higher values, given that $\boldsymbol{D}(P_{\tilde{A}} \parallel P_A)$ value tends to zero and, consequently, rate loss tend to its minimum value $H(A)$. This makes CCDM the method to choose, however, its need for high precision arithmetic codes and for high number of output bits makes it very difficult to implement. [52]

3.2.2 Cut and Paste - CAP

Considering a communication system with a forward error correction (FEC) encoder and decoder; a bit-to-symbol mapper M_1 and demapper M_1^{-1}, which map the coded bits to constellation points, each of which is typically a complex symbol; and a coherent optical transmitter and receiver (grey box in figure 3.7). CAP shaping only requires the addition of a bit mapper M_2 and demapper M_2^{-1} comparing to a conventional optical modulation scheme. Mapping and demapping is implemented using a "Look up Table" (LUT) with required number of entries. [53]

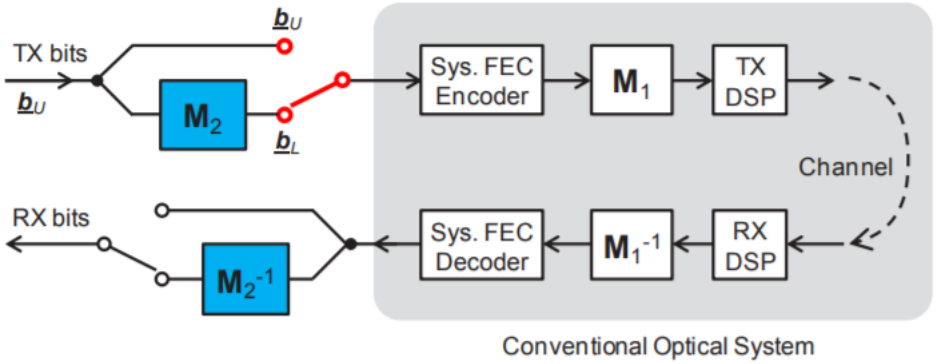


Figure 3.7: CAP shaping system block diagram [53]

At the transmitter, two branches can be selected (red lines in figure 3.7). If the upper branch is chosen, a sequence b_U passes trhough with the conventional bit sequence from transmitted bits. However, in the case of lower branch, that sequence will be first mapped in M_2 mapper wich convertrs that $\underline{b}_U$ sequence into a new $\underline{b}_L$ sequence such that the following bit-to-symbol mapper M_1 will produce a symbol with different transmit energy. M_2 will reorder the transmitted sequence in a way that output symbol at M_1 from $\underline{b}_L$ will have a reverse order energy in contrast with $\underline{b}_U$.

The total energies are compared between the upper and lower branch in each of the $M \times$ D bit input sequences, where D corresponds to the complex dimension of CAP shaping, and the sequence of lower energy is chosen to be transmitted through the fiber.

Due to CAP shaping low-energy symbols are transmitted with more recurrence than high-energy symbols as shown in figure 3.8 and average transmit power is reduced. This reduction comparing to non-shaped signal is called shaping gain. [53]

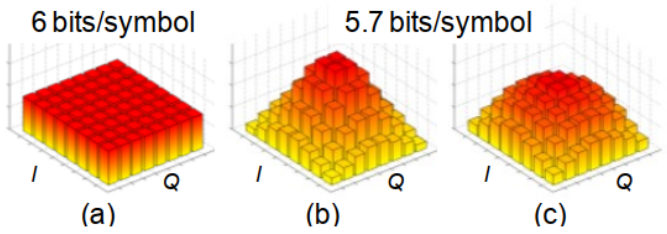


Figure 3.8: Probability distribution of the symbols (a) without shaping, (b) CCDM shaping and (c) CAP shaping (D=4) [53]

With this energy reduction, one-bit shaping overhead (OH) per every D symbols is needed (i.e., per every $6 \times$ D in the case of 64-QAM) which will be used as an indentifier to indicate which branch was chosen by the transmitter and the receiver will recover to the initial bit sequence according to that information. FEC Overhead with code rate R_C is given by [54]:

$$OH = 100\ (\frac{1}{R_C} - 1)\ \% \tag{3.9}$$

OH can be used for FEC coding enhancement that provides an additional coding gain over the baseline coding gain. Likewise this performance improvement, CAP supports flexible-OH which requires one FEC coding for multiple OHs, thus more economic.

While shaping gain in linear channels is independet from the symbol distribution, in nonlinear channels it is vulnerable to different symbol distributions, which introduces different signal-dependent noise characteristics.

The nonlinear interference noise (NLIN) variance depends on the term

$$\frac{(\sum_{i=1,..,D}[(\mid x_i \mid^2 /\mu) - 1])^{2/3}}{D} \tag{3.10}$$

where x_i refers the i-th transmit symbol in the sequence and μ is the average symbol energy. This term denotes that some NLIN contributions are proportional to the deviation of the signal energy to a constant modulus.

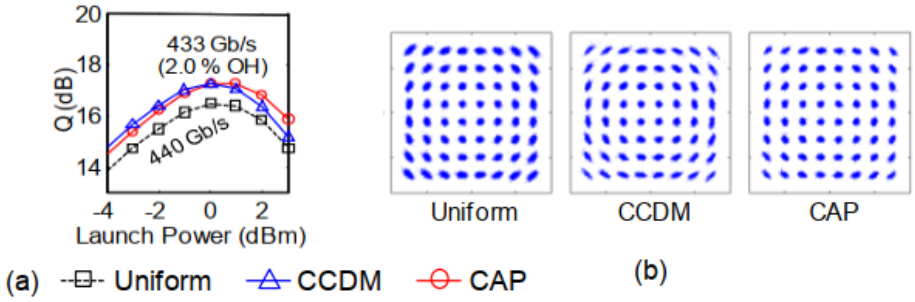


Figure 3.9: (a) Q factor as a function of the launch power, for the OH of 2.0 and (b) received constellation diagrams at a high launch power [53]

In [53] is shown that CAP shaping shows similar performance to CCDM shaping due to the nonlinear gain as predicted by their statistical estimation, which is noticeable in figure 3.9.

Chapter 4

Simulations and Results

The simulation was developed using MATLAB software in order to implement the system. CAP transmitter and receiver algorithms were developed from the beginning, while FEC and CCDM were implemented in this simulation setup using PICAdvanced algorithms. In this section, some details for the system implementation will be specified.

It is important to note that all simulations are made for the highest baud rate of 60 Gbaud per channel. Using a complex CAP dimension of 2, our bit rate for 64-QAM should be 60 $\times$ 2 (due to dual polarization) $\times$6(bits per symbol) $\times$ 5/6 (code rate loss due to FEC soft decision error correction) = 600 Gbit/s. Without FEC the bit rate would be 720 Gbit/s.

Method	Bit Rate (Gbit/s)	OH (%)
Uniform	600	0
CAP P1 (D=2)	550	9
CAP P2 (D=2)	500	20
CAP P3 (D=2)	400	50

Table 4.1: Bit Rates (in Gbit/s) for different methods used with FEC

4.1 Simulation Setup

The first scheme tested is shown in figure 4.1 where no FEC algorithm was used and it is composed by a PRBS (pseudorandom binary sequence), PCS algorithm and its inverse function, QAM mapper and demapper, AWGN and in the end the BER calculation. These blocks will be described throughout this chapter.

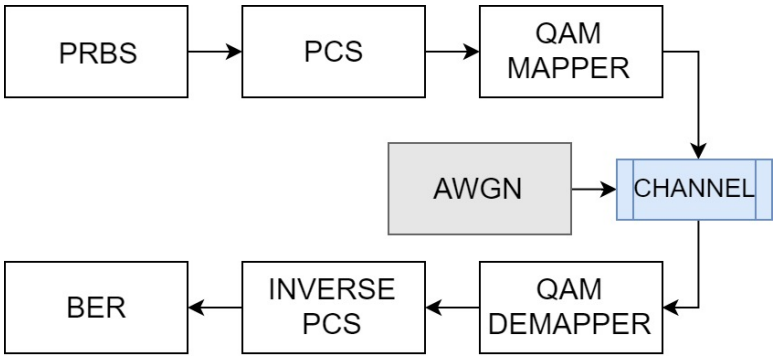


Figure 4.1: First scheme implemented without FEC for BER estimation

4.1.1 Pseudorandom binary sequence

PRBS is a known method to control the operation of digital communications with binary signal sequences whose repetition period is very long and which may be considered as pseudo random sequences. Such a sequence is constituted by a sequence of bits passing at a clock rhythm. [55]

A pseudorandom sequence is a code sequence of 1's and 0's whose autocorrelation properties are similar to those of white noise. The term 'pseudo' comes from its deterministic properties and as a result can't really be called random because after a given number of elements it will start to constantly repeat itself, unlike real random sequences wich are impossible to obtain in simulation domain. [56]

4.1.2 Probabilistic Constellation Shaping - PCS

To implement the probablistic constelation shaping, CAP algorithm with 64-QAM was used with corresponding Mapper M_1 and M_2 represented in figure 3.7 and with Look up table distribution in 4.2 for one quadrant, wich can be easly replicated for the other quadrants.

CAP shaping complex dimension (D) refers to the number of symbols in comparison, in all simulations D is set to be two, which means two symbols, or a sequence of 6×2 (12) bits, are always being compared. In each of the 6×2 input bit sequences, the total energies are compared between the upper and lower branch sequences represented in 4.2 with red dots. If the lower energy value sequence is the one from the lower branch, the symbol corresponding to this value will pass through the fiber.

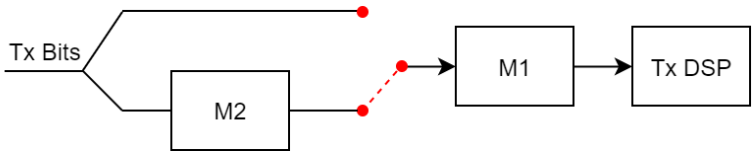


Figure 4.2: Block Diagram for 'CAP P1' - 550 Gbit/s

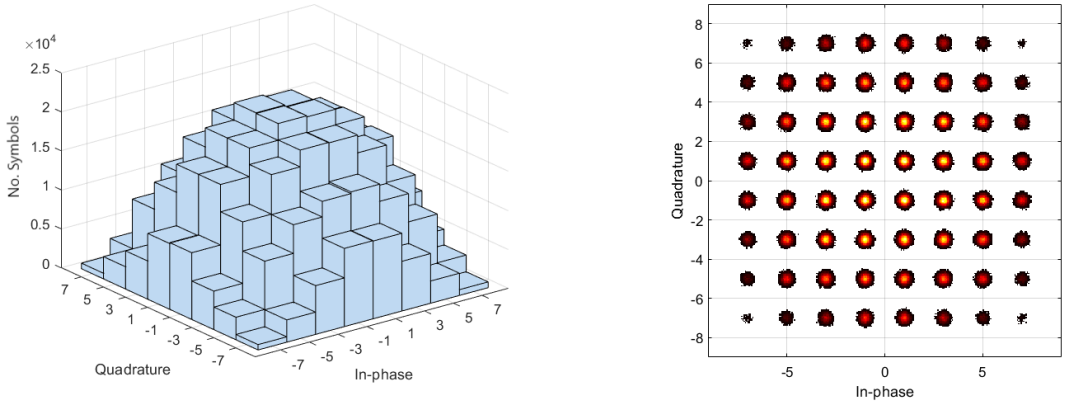
For instance, if the sequence corresponding to 5+5i, with an energy of 50, is selected and 7+7i (with energy of 98) is being compared, the corresponding bit sequence of 5+5i (001111), will be chosen due to its lower energy in this comparison.

The LUT represented in 4.2 was made with the left-most column for memory addressing from the input b_U , then produces an output bit sequence b_L and two energy values, represented in reversed order. Each set of six bits from the input are going to map one symbol with the correspondent energy obtained by the absolute square ($|M(b_x)|^2$) of each symbol, as described in 3.2.2.

Table 4.2: CAP P1 shaping LUT for 64-QAM

b_u	$M_1(b_U)$	$\mid M_1(b_U) \mid^2$	$b_L = M_2(b_U)$	$M_1(b_L)$	$\mid M_1(b_L) \mid^2$
000000	1+1i	2	001100	7+7i	98
000001	3+1i	10	001101	5+7i	74
000010	1+3i	10	001110	7+5i	74
000011	3+3i	18	000110	7+3i	58
000101	5+1i	26	001001	3+7i	58
001010	1+5i	26	001111	5+5i	50
000111	5+3i	34	000100	7+1i	50
001011	3+5i	34	001000	1+7i	50
001000	1+7i	50	001011	3+5i	34
001111	5+5i	50	000111	5+3i	34
000100	7+1i	50	001010	1+5i	26
001001	3+7i	58	000101	5+1i	26
000110	7+3i	58	000011	3+3i	18
001101	5+7i	74	000001	3+1i	10
001110	7+5i	74	000010	1+3i	10
001100	7+7i	98	000000	1+1i	2

In figure 4.3 is shown the simulation result of CAP shapping at the output of the PCS algorithm for acsSNR of 26 dB, obtainded by measuring the total number of symbols and its distributions after the insertion of noise in the form of histogram and constellation diagram. It is possible to observe high density of symbols with lower energy in the middle of the constellation diagram, being outer symbols less likely to happen with much less density.



(a) (b)

Figure 4.3: (a) Histrogram and (b) Constellation for 550 Gbit/s

To implement 'CAP P2', instead of M_2 used in 3.7 were used three more Mappers (M_2 to M_4) in order to implement the modulation scheme, following the same selection criteria - the lower energy symbol will be selected to the next step. This allows us to use the same LUT as for 64-QAM replicated for the new mappers added.

However, this will come with an expense in terms of Overhead (OH)bits, where 1 more bit is necessary to implement it. With this, our bit rate will be affected as now there are 4 information bits and 2 OH bits: $(6 \times D) - 2 \times 60 \times 5/6 = 500$ Gbit/s.

It is relevant to compare this method with a conventional 32-QAM modulation scheme, since its number of bits per symbol is the same, however 'CAP P2' is based on 64QAM modulation format with a LUT described in 4.2.

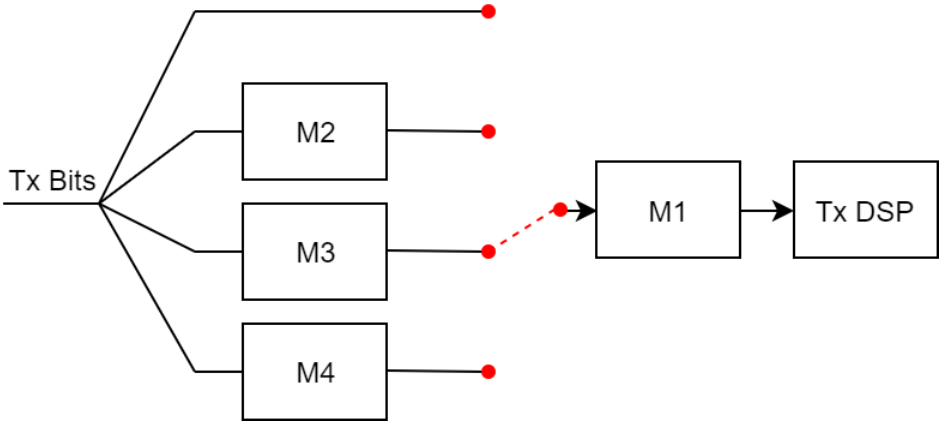


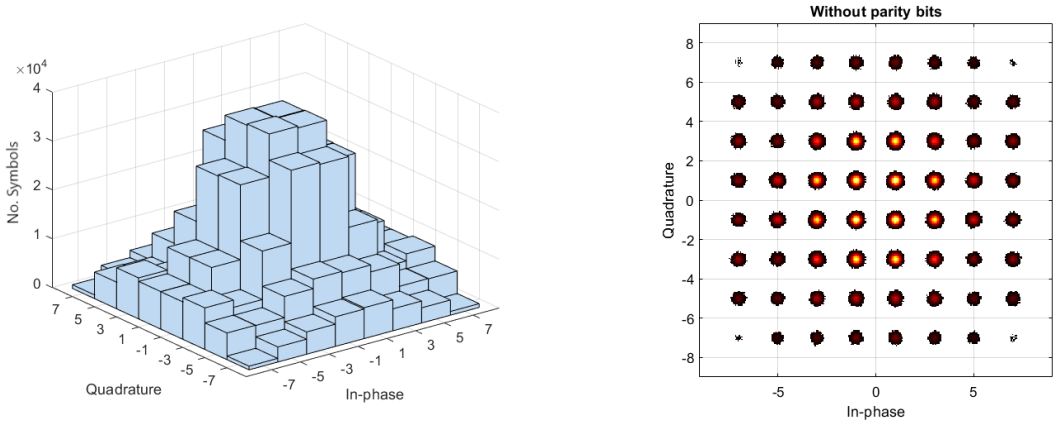
Figure 4.4: Block diagram for 'CAP P2' - 500 Gbit/s

Table 4.3 shows the LUT for 'CAP P2' proposal, with two more branches to choose and respective symbol mapper and energies. Those values follow the same structure as 4.2, where left-most column refers to input bit at upper branch with its symbol mapping and energy obtained by the absolute square ($|M(b_x)|^2$) of each symbol, followed by successively lower branches. Note that all bit to symbol mappers correspond to the same energy level and, by the same logic, all bits have the same symbol mapping - only the order that changes.

Table 4.3: CAP P2 shaping LUT

b_u	M_1 (b_U)	E b_U	$b_L = M_2(b_U)$	M_1 (b_L)	E b_L	$b_{L1} = M_3(b_U)$	M_1 (b_{L1})	E b_{L1}	$b_{L2} = M_4(b_U)$	M_1 (b_{L2})	E b_{L2}
000000	1+1i	2	001100	7+7i	98	001100	7+7i	98	001000	1+7i	50
000001	3+1i	10	001101	5+7i	74	001101	5+7i	74	000100	7+1i	50
000010	1+3i	10	001110	7+5i	74	001110	7+5i	74	001011	3+5i	34
000011	3+3i	18	000110	7+3i	58	000110	7+3i	58	000101	5+1i	26
000101	5+1i	26	001001	3+7i	58	001000	1+7i	50	001010	1+5i	26
001010	1+5i	26	001111	5+5i	50	000100	7+1i	50	000001	3+1i	10
000111	5+3i	34	000100	7+1i	50	001011	3+5i	34	000000	1+1i	2
001011	3+5i	34	001000	1+7i	50	000101	5+1i	26	000010	1+3i	10
001000	1+7i	50	001011	3+5i	34	001010	1+5i	26	000011	3+3i	18
001111	5+5i	50	000111	5+3i	34	000001	3+1i	10	000111	5+3i	34
000100	7+1i	50	001010	1+5i	26	000000	1+1i	2	001111	5+5i	50
001001	3+7i	58	000101	5+1i	26	000010	1+3i	10	001001	3+7i	58
000110	7+3i	58	000011	3+3i	18	000011	3+3i	18	000110	7+3i	58
001101	5+7i	74	000001	3+1i	10	000111	5+3i	34	001110	7+5i	74
001110	7+5i	74	000010	1+3i	10	001111	5+5i	50	001101	5+7i	74
001100	7+7i	98	000000	1+1i	2	001001	3+7i	58	001100	7+7i	98

Figure 4.5 shows the symbol distribution for this method after the insertion of noise through AWGN channel with 26 dB of SNR, where it is possible to see even more concentrated symbols in the centre of the constellation diagram and that those symbols appear in larger number when compared to symbols with lower energy (outer symbols).



(a) (b)

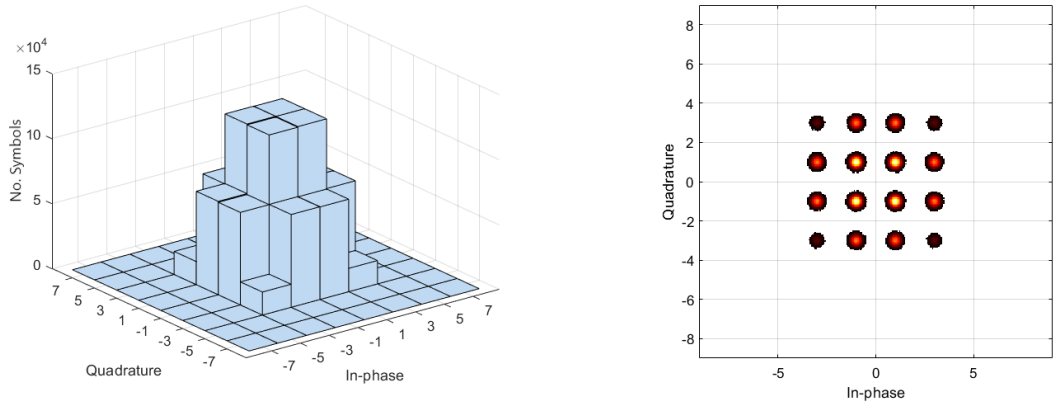
Figure 4.5: (a) Histrogram and (b) Constellation for 500 Gbit/s

'CAP P3' was implemented using a normal 16-QAM format, with the same scheme shown in figure 3.7 and 4.2 using the corresponding LUT for this modulation format, represented in 4.4 for one quadrant, easly replicated for the remaining ones. The bit rate will be 400 Gbit/s, since there are only 4 bits per symbol.

Table 4.4: CAP P3 shaping LUT for 16-QAM

b_u	$M_1(b_U)$	$\mid M_1(b_U) \mid^2$	$b_L = M_2(b_U)$	$M_1(b_L)$	$\mid M_1(b_L) \mid^2$
0000	1+1i	2	0011	3+3i	18
0001	3+1i	10	0010	1+3i	10
0010	1+3i	10	0001	3+1i	10
0011	3+3i	18	0000	1+1i	2

In 4.6 is shown the simulation result of CAP shapping at the output of the PCS algorithm after noise insertion with SNR of 20 dB, obtainded by measuring the total number of output symbols and its distributions in a constellation diagram. Only four symbols per quadrant are represented as result of a conventional 16-QAM modulation format, however it is visible, like the other proposals, an higher concentration of symbols with lower energy.



(a) (b)

Figure 4.6: (a) Histrogram and (b) Constellation for 400 Gbit/s

4.1.3 BER results

BER is represented through a semilogarithmic plot, which enables representation of BER in function of SNR in dB. A sample of 4×10^6 bits were used to get graphs for PCS in order to obtain BER in the same order of magnitude.

A representative graph of SNR versus BER for QAM schemes over a standard digital communication transmission is shown in Figure 4.7 where it is possible to verify that moving to a higher order QAM scheme, more bits per symbol are transmited, however, if the SNR is to remain the same, as the symbol's constellation is larger, the points must be closer together and are thus more susceptible to noise, resulting in a higher BER.

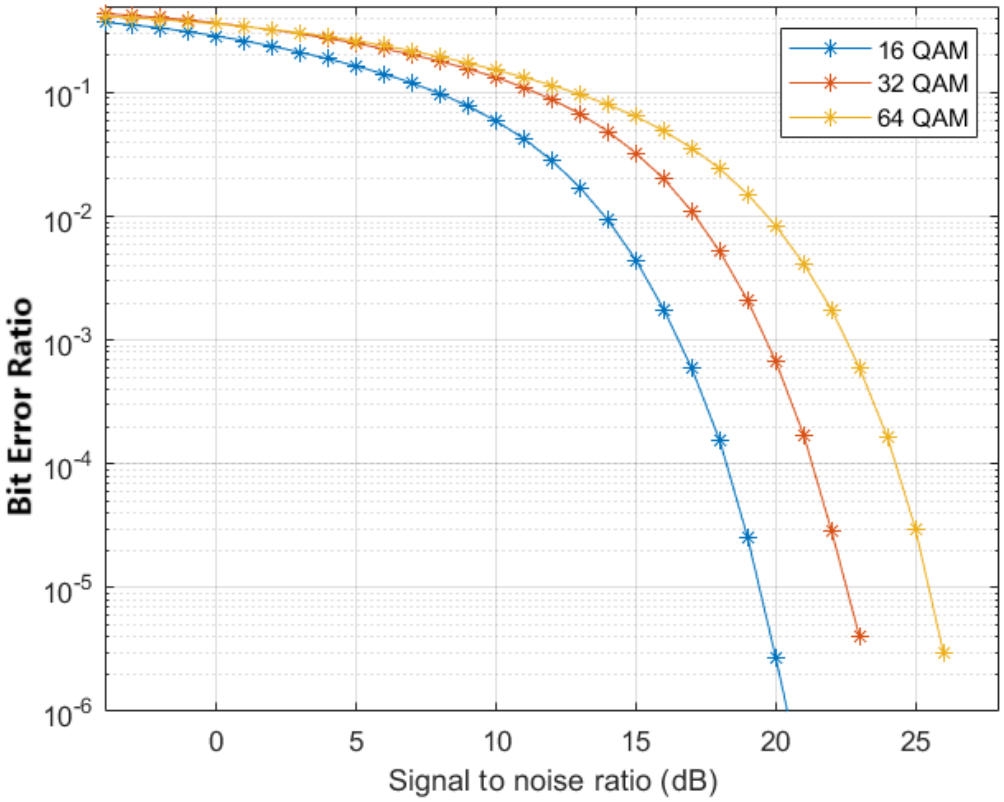


Figure 4.7: BER vs SNR comparing uniformly distributed 16-QAM (blue), 32-QAM (red) and 64-QAM (yellow)

For 20 dB of SNR it is possible to verify BER value of 10^{-6} for 16-QAM. This value will be 10^{-3} for 32-QAM and 10^{-2} for 64-QAM. In terms of gain, there is a gain of approximately 3 dB beetween each order of maginitude for QAM schemes, shown with the SNR values for 10^{-5} BER in the following table:

Modulation format	SNR (dB)
16-QAM	19.8
32-QAM	22.6
64-QAM	25.5

Table 4.5: SNR values for 10^{-5} BER

Figure 4.8 shows BER vs SNR graph for three CAP algorithms proposed. It is important to note that 'CAP P2' does not have an ideal behavior since it is similar to 'CAP P1'. However, for the other considered proposals improvements are verified when compared to an uniform system. Both CAP for 16 and 64-QAM introduce improvements in terms of noise resistance.

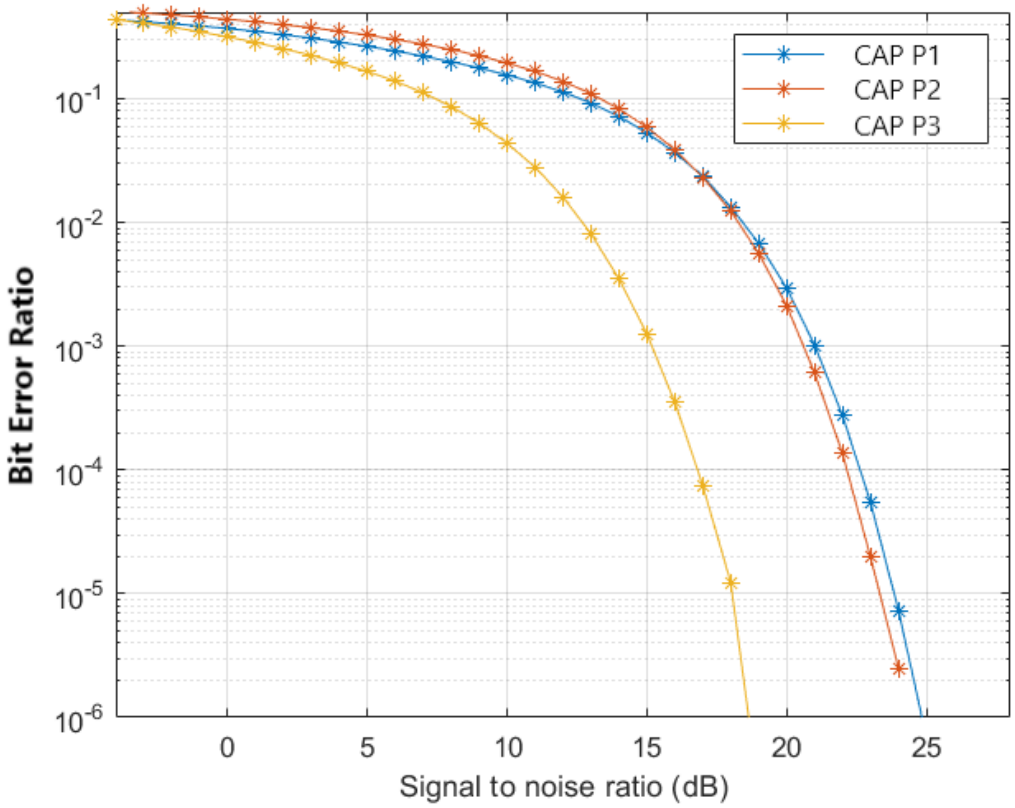
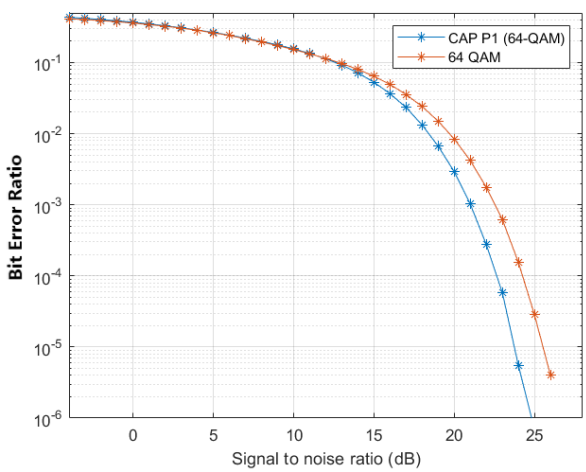


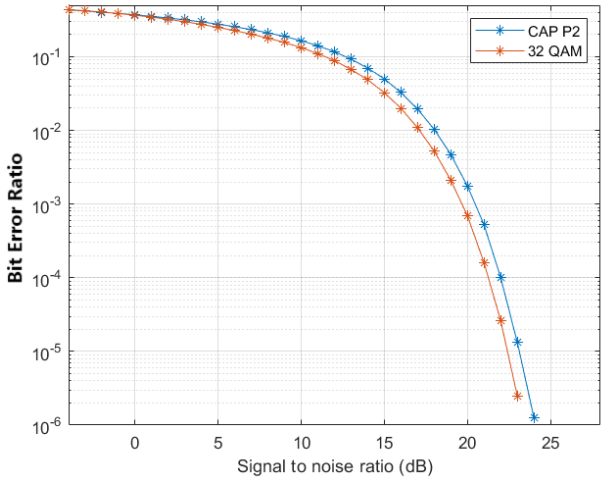
Figure 4.8: BER vs SNR comparing for CAP algorithm

Analysing 4.8 it is possible to note that CAP introduces, approximately, 1 dB gain comparing to uniformly distributed 64-QAM and same goes for 16-QAM. Being the ultimate shaping gain presented in equation 3.1 of 1.53 dB, this 1dB shaping gain can be considered an important result in this simulation.

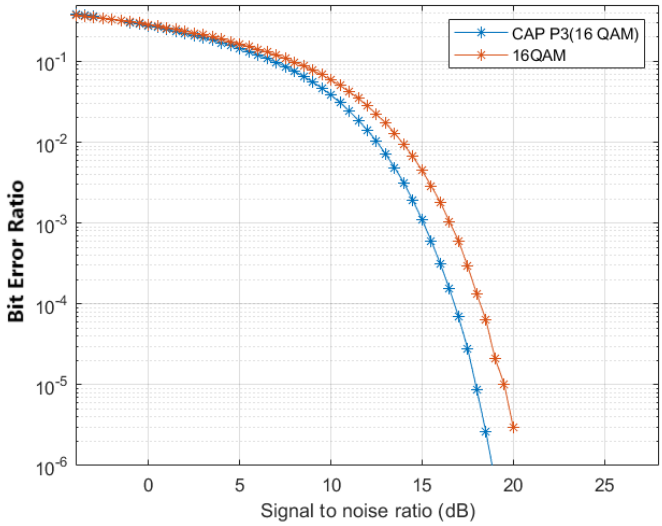
In the case of 'CAP P2' methodology the results are not favorable, it may even be said to have worse results compared to a conventional 32-QAM system. When comparing to a 64-QAM, both present almost the same performance in the graph 4.8 and so, with the decreased transmission rate of 50 Gbit/s that 'CAP P2' presents it does not represent a favorable method for implementation. 'CAP P3' presents similiar beahviour as 'CAP P1' with 1dB shaping gain for 16 QAM.



(a)



(b)



(c)

Figure 4.9: Comparision of CAP algorithm and uniform distribution for (a) 64-QAM, (b) 32-QAM and (c) 16-QAM

4.1.3.1 FEC integration

To implement FEC two new blocks were added to the scheme in 4.1- FEC encoder and its inverse block (decoder) :

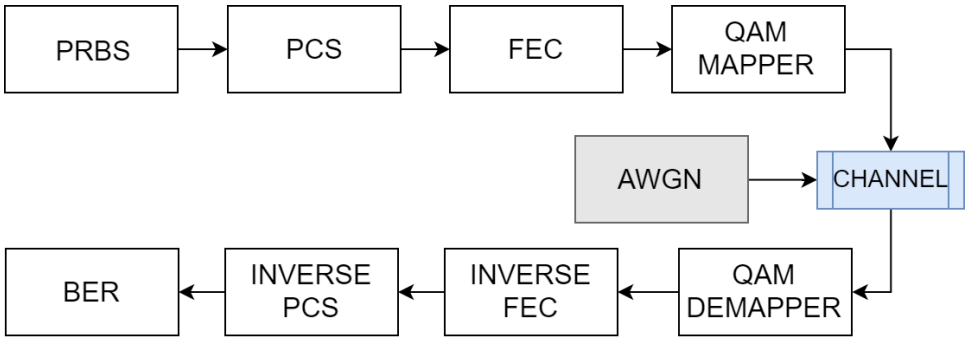


Figure 4.10: Scheme implemented with FEC for BER estimation

FEC integration had to take into account that algorithm had to be repeated a given number of times in order to obtain good results of BER since the number of input bits is reduced here to follow the requirements of FEC implementation - each repetition is called "window" throughout this work.

MATLAB communications toolbox was used to implement FEC, in particular LDPC encoding and deconding algorithms. The Encoder block applies LDPC coding to the binary input message and sets the ParityCheckMatrix property to parity and creates an LDPC encoder System object. The parity input must be specified as described by the ParityCheckMatrix property. In this simulation 'dvbs2ldpc(r)' was used which uses Low-density parity-check codes from DVB-S.2 standard and requires the following [57]:

- code rate 'r' introduced by the DVB-S.2 standard (multiple values possible from 1/4 to 9/10, depending on the number of bits needed) . 5/6 was used in all simulations;
- block length of the code has to be 64800;
- Parity-check matrix specified as a sparse (N – K)-by-N binary-valued matrix. N is the length of the output codeword vector and K is the length of the uncoded message and must be less than N;
- The last (N – K) columns in the parity-check matrix represent parity bits introduced by FEC encoding.

Using the same properties for encoding, the input data using an LDPC code based on the default parity-check matrix is decoded. The 'DecisionMethod' specifies the 'hard decision' to use in Log-likelihood ratio (LLR) where each element is the log-likelihood ratio for a received bit, represented in terms of positive or negative numbers, in which values are more likely to be 0 if the log-likelihood ratio is positive. The first K elements correspond to the information-part of the input message and those are the ones decoder will return.

As FEC algorithm used requires a small number of input values, for each SNR value simulation was repeated at least 10 times ("windows") to have a more decent sample of bits. Also, the granularity of SNR had to be increased since for a certain value, the number of error bits goes dramatically down. Image 4.11 shows a general view of uniform 16, 32 and 64-QAM with FEC algorithm side by side.

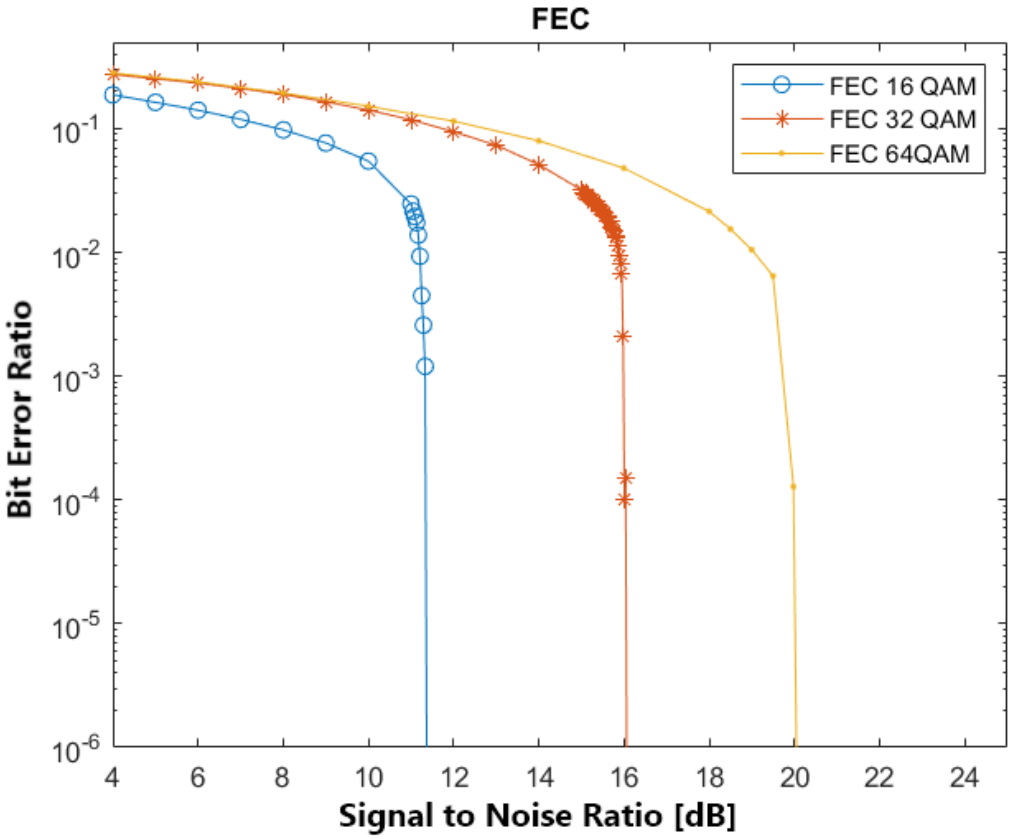
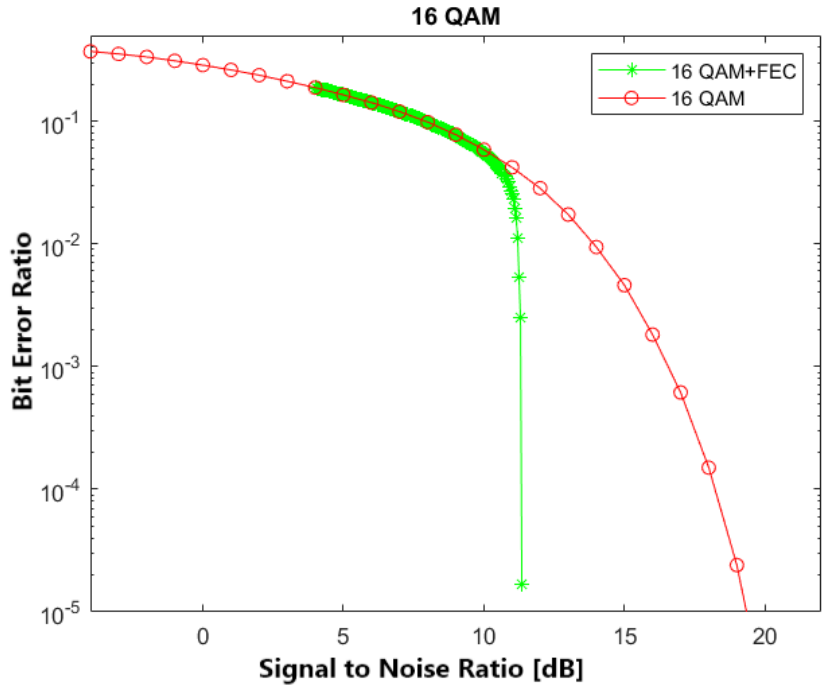
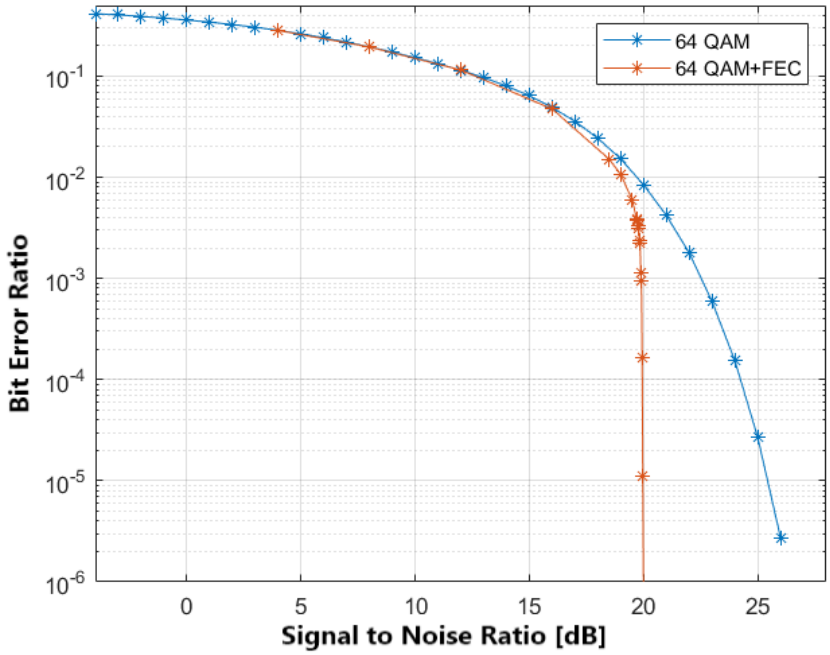


Figure 4.11: BER vs SNR comparing uniform 16-QAM (blue) , 32-QAM (red) and 64-QAM (yellow) with FEC

Figure 4.12 shows a more detailed comparision between an uniform distribution with and without FEC and it is possible to see that FEC gives to 16-QAM 8 dB gain, while for 64-QAM an improvement of 6 dB is visible. For further work, only 'CAP P1' (64 QAM) was tested, since it was proved before (image 4.9c) that 16QAM will present similar behaviour in terms of gain as 64QAM.



(a)



(b)

Figure 4.12: Comparision of uniform distribution with FEC and without FEC for (a) 16-QAM and (b) 64-QAM

Image 4.13 introduces the presence of FEC algorithm in combination with CAP shaping. It is visible an ultimate gain of approximately 3 dB which constitutes an important progress in this simulation reducing the transmission rate by only 50 Gbit/s due to the number of OH bits required for shaping.

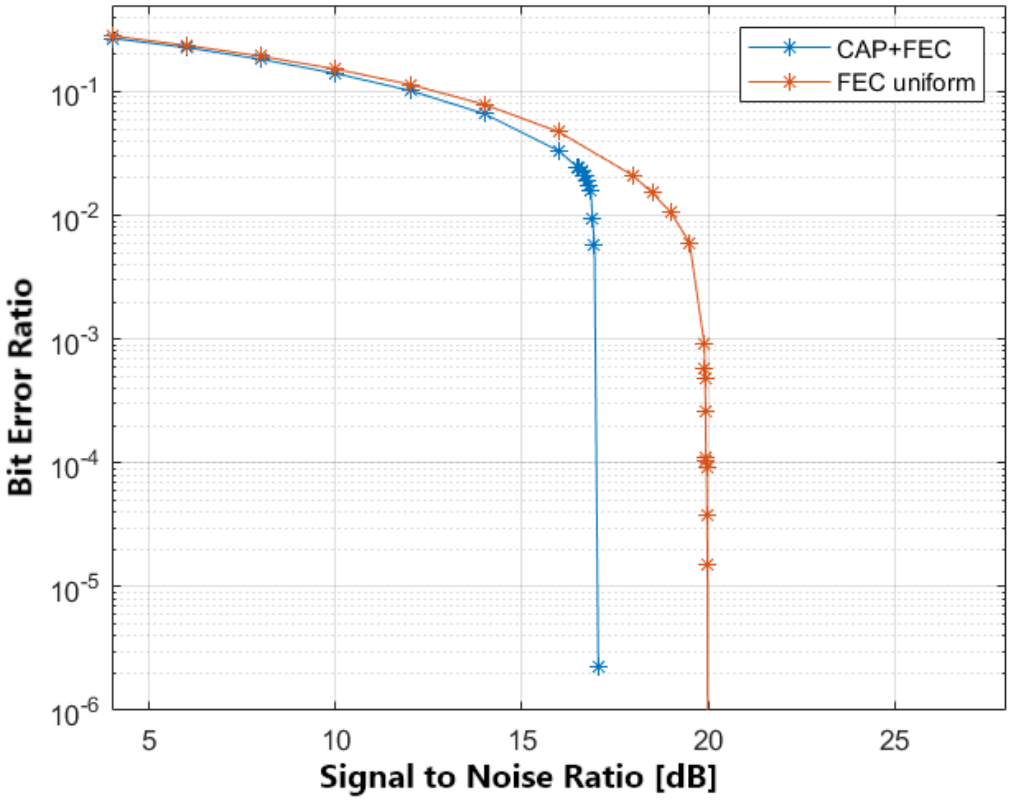
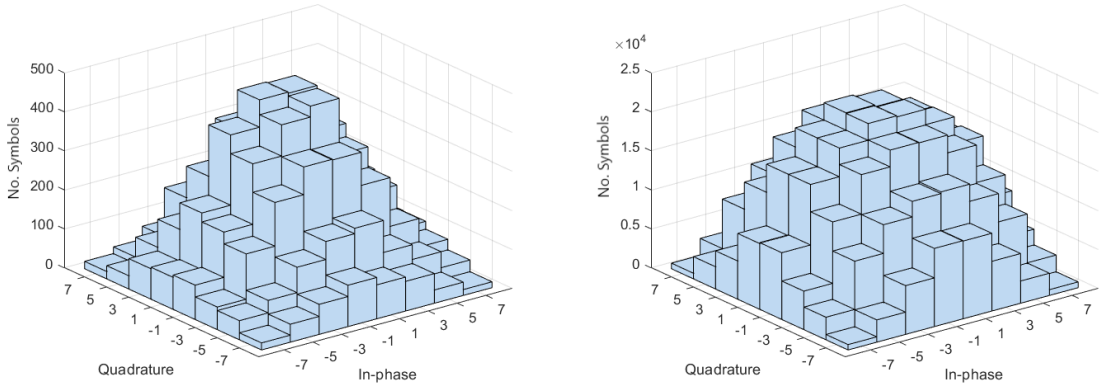


Figure 4.13: BER vs SNR comparing uniform 64-QAM with FEC and 64-QAM with CAP and FEC

4.1.3.2 CCDM comparison

Probabilistic Constellation Shaping techniques analysis end with CCDM comparison, which is implemented by an arithmetic code based on the Maxwell-Boltzmann distribution with $\lambda = 0.03427$ with FEC algorithm. It was defined that 10800 symbols are intended for DM output.

Image 4.14 shows that CCDM achieves a smaller distribution of the outer elements of the constellation that present higher energy, concentrating its distribution on the inner symbols, leading to an average energy of 23, while 'CAP P1' (64 QAM) average energy is 27.



(a) (b)

Figure 4.14: Histrogram for (a) CCDM and (b) for CAP P1 (64 QAM)

CCDM simulation had to take into account that an error in one bit can corrupt all the "window" in simulation, so it was defined that it should have at least 10 error windows. For the acceptable limit SNR values a number of 50000 windows were tested to obtain 7 errors, and with this was possible to obtain 4.15 with BER in the order of 10^{-4} BER. It is vivible that for lower SNR values, BER is constantly 0.5, just showing a decrease for a limited range of SNR (15.4 to 15.61 dB). Comparing with 'CAP P1' (64 QAM) implementation, gain of 1.4 dB is noticeable.

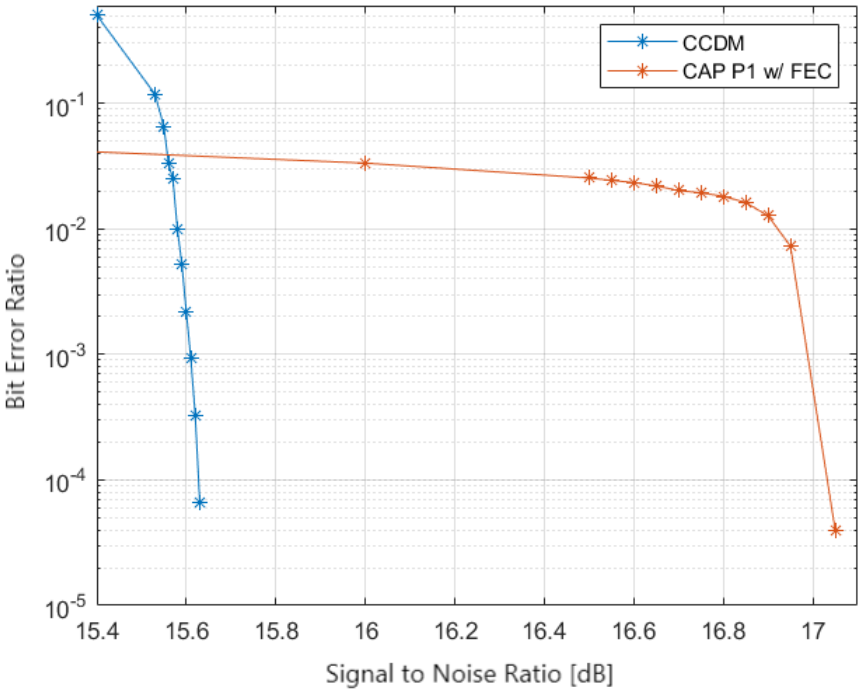


Figure 4.15: BER vs SNR for CCDM implementation and CAP P1

It has been proven that CAP reduces average symbol energy (unitless measure), as supposed, and increases entropy (in bits) which means the need for more bits per symbol to get the correct symbol at the receiver. CAP requires 5.7 bits per symbol while CCDM need 5.6 bits per symbol. Despite the bit rate loss CAP introduces due to OH bits, it also gives greater noise tolerance. When comparing to an uniform 64QAM with FEC distribution there is a difference of 3 dB in SNR to obtain BER in the range of 10^{-4} as shown in 4.13, while for CCDM 4.4 dB gain is visible in 4.15.

As a summary, table 4.6 is presented with the various performance settings evaluated after FEC algorithm has been introduced in the system.

Method	Bit Rate (Gb/s)	OH (%)	Average Energy	Entropy (bits)
16 QAM with FEC	400	50	7.01	3.67
CAP P3 (16 QAM)	400	50	6.70	3.77
64 QAM with FEC	600	0	42.04	5.62
CAP P1 (64 QAM)	550	9	27.39	5.73
CCDM	550	9	23.64	5.59

Table 4.6: Final comparison of methods used.

Chapter 5

Conclusions and Future Work

5.1 Conclusions

Nowadays there is an increasing need to transfer large amounts of data, requiring the maximum of the current metro networks and bringing its' capacity to the limit.

With the constant need to have ever growing bit rates, it is known that optical communications are the only communication technology currently existing that allows high order bit rates per channel. Probabilistic constellation shaping is the method that excels in order to obtain high capacity networks closer to the Shannon Limit.

In this master thesis it was proposed and demonstrated that PCS algorithms, in particular CAP are an important method with advantages in some applications that has to be taken into account for future optical metro networks.

In chapter 2 is given first an overview of some of the most relevant methodologies behind uniform signaling and shaping, as well as error correction and performance evaluation solutions used in digital systems and throughout this work.

In chapter 3 PCS concept is introduced and with it some of their specifications such as Maxwell-Boltzmann distribution and Distribution Matching. CAP and CCDM were discussed with more detail.

In chapter 4, MATLAB software was used in all simulations. Three possible implementations of CAP were tested, however 'CAP P1' was focused for being the best solution, since it has the best Bit rate tested and presents more bits per symbol once it's based in 64-QAM modualtion. When comparing with an uniform distribution it presents 1 dB shaping gain, which is close to 1.53 dB maximum gain achievable for shaping. FEC algorithm gives even more gain and also the average number of bits per symbol is reduced. For the final comparison, CCDM was proposed due to its near optimal performance and it was visible that CAP achieved results close to CCDM, with 1.4 dB difference, and therefore a valid alternative to be implemented.

5.2 Future work

In this master thesis, all simulations were performed using a bit rate from 400 Gbit/s to 600 Gbit/s for 60 Gbaud. With the constant evolution of the demands requirements from the networks, several solutions must be tested in order to improve the transmission performance. With this, the following topics are proposed as future work:

- Complexity evaluation of CAP algorithm, studying its benefits for practical implementation;
- CAP algorithm study for different modulation formats, comparing with QAM and different PCS algorithms;
- Performance study for distinct number representations, such as floating point;
- Experimental analysis of the results obtained in the simulations.

Bibliography

[1] Fred Buchali, Fabian Steiner, Georg Böcherer, Laurent Schmalen, Patrick Schulte, and Wilfried Idler. Rate adaptation and reach increase by probabilistically shaped 64-qam: An experimental demonstration. *Journal of Lightwave Technology*, 34(7):1599–1609, 2016.

[2] Wolfram Lautenschlaeger, Nihel Benzaoui, Fred Buchali, Lars Dembeck, Roman Dischler, Bernd Franz, Ulrich Gebhard, Jens Milbrandt, Yvan Pointurier, Detlef Roesener, et al. Optical ethernet—flexible optical metro networks. *Journal of Lightwave Technology*, 35(12):2346–2357, 2017.

[3] Tobias Fehenberger, Alex Alvarado, Georg Böcherer, and Norbert Hanik. On probabilistic shaping of quadrature amplitude modulation for the nonlinear fiber channel. *Journal of Lightwave Technology*, 34(21):5063–5073, 2016.

[4] Jacklyn D Reis, V Shukla, DR Stauffer, and K Gass. Technology options for 400g implementation. In *Optical Networking Forum (OIF), OIF-Tech-Options-400G-01.0*, 2015.

[5] 600g and above – component requirements. `https://www.neophotonics.com/600g-component-requirements/`. Accessed: 2019-11-05.

[6] Wilfried Idler, Fred Buchali, Laurent Schmalen, Eugen Lach, Ralf-Peter Braun, Georg Böcherer, Patrick Schulte, and Fabian Steiner. Field trial of a 1 tb/s super-channel network using probabilistically shaped constellations. *Journal of Lightwave Technology*, 35(8):1399–1406, 2017.

[7] Peter J Winzer and Renè-Jean Essiambre. Advanced modulation formats for high-capacity optical transport networks. *Journal of Lightwave Technology*, 24(12):4711–4728, 2006.

[8] C Helstrom. The comparison of digital communication systems. *IRE Transactions on Communications Systems*, 8(3):141–150, 1960.

[9] Basab Purkayastha and Kandarpa Sarma. *A Digital Phase Locked Loop based Signal and Symbol Recovery System for Wireless Channel*. 01 2015.

[10] SO Popescu, G Budura, and AS Gontean. Review of psk and qam—digital modulation techniques on fpga. In *2010 International Joint Conference on Computational Cybernetics and Technical Informatics*, pages 327–332. IEEE, 2010.

[11] B Sundareshan. Digital modulation—baseband techniques. *IETE Journal of Education*, 33(1):35–44, 1992.

[12] Bernard Sklar and Fredric J Harris. *Digital communications: fundamentals and applications*, volume 2001. Prentice-hall Englewood Cliffs, NJ, 1988.

[13] Ghafour Amouzad Mahdiraji and Ahmad Fauzi Abas. *Advanced Modulation Formats and Multiplexing Techniques for Optical Telecommunication Systems*. 03 2010.

[14] Fsk - frequency shift keying. `https://www.tmatlantic.com/encyclopedia/index.php?ELEMENT_ID=10422`. Accessed: 2019-09-12.

[15] Nikolaos Voudoukis. Performance analysis, characteristics, and simulation of digital qam. *European Journal of Electrical Engineering and Computer Science*, 1(1), 2017.

[16] Bhagwandas Pannalal Lathi. *Modern Digital and Analog Communication Systems 3e Osece*. Oxford University Press, Inc., 1998.

[17] Manish Trikha, Neha Sharma, Manas Singhal, Ritu Rajan, and Pankaj Bhardwaj. Ber performance comparison between qpsk and 4-qa modulation schemes. *MIT International Journal of Electrical and Instrumentation Engineering*, 3(2):62–66, 2013.

[18] Saeed V Vaseghi. *Advanced digital signal processing and noise reduction*. John Wiley & Sons, 2008.

[19] RM Howard. White noise: A time domain basis. In *2015 International Conference on Noise and Fluctuations (ICNF)*, pages 1–4. IEEE, 2015.

[20] Simon Haykin. *Communication systems*. John Wiley & Sons, 2008.

[21] J Cioffi. Ee379a-digital communication: Signal processing, 2008.

[22] Q. Chaudhari. *Wireless Communications from the Ground Up: Fundamentals of Digital Communication Systems*. CreateSpace Independent Publishing Platform, 2016.

[23] Mike Allen. *The SAGE encyclopedia of communication research methods*. SAGE Publications, 2017.

[24] David Tse and Pramod Viswanath. *Fundamentals of wireless communication*. Cambridge university press, 2005.

[25] Rajan Kadel, Nahina Islam, Khandakar Ahmed, and Sharly J Halder. Opportunities and challenges for error correction scheme for wireless body area network—a survey. *Journal of Sensor and Actuator Networks*, 8(1):1, 2019.

[26] Martin P Clark. *Wireless access networks*. Wiley Online Library, 2000.

[27] Robert Gallager. Low-density parity-check codes. *IRE Transactions on information theory*, 8(1):21–28, 1962.

[28] Amanjot Kaur and Jasbir Singh. Performance evaluation of digital modulation techniques in a wcdma-based radio-overfiber communication system. *International Journal of Advanced Research in Computer Science and Electronics Engineering*, 1(4), 2012.

[29] Dennis Derickson, Christian Hentschel, and Joachim Vobis. *Fiber optic test and measurement*, volume 8. Prentice Hall PTR New Jersey, 1998.

[30] Jeffrey A Jargon, CM Jack Wang, and Paul D Hale. A robust algorithm for eye-diagram analysis. *Journal of Lightwave Technology*, 26(21):3592–3600, 2008.

[31] ON Semiconductor. Understanding data eye diagram methodology for analyzing high speed digital signals. *Application Note.[Online]. Available: http://onsemi. com*, 2014.

[32] Rishad Ahmed Shafik, Md Shahriar Rahman, and AHM Razibul Islam. On the extended relationships among evm, ber and snr as performance metrics. In *2006 International Conference on Electrical and Computer Engineering*, pages 408–411. IEEE, 2006.

[33] Rashmi Suthar, Sunil Joshi, and Navneet Agrawal. Performance analysis of different m-ary modulation techniques in cellular mobile communication. *IP Multimedia Communications A Special Issue from IJCA*, 2009.

[34] Trilochan Patra and Sanjib Sil. Bit error rate performance evaluation of different digital modulation and coding techniques with varying channels. In *2017 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON)*, pages 4–10. IEEE, 2017.

[35] David Coudert, Napoleao Nepomuceno, and Herve Rivano. Wireless backhaul networks: Minimizing energy consumption by power-efficient radio links configuration. 01 2008.

[36] Erik G Learned-Miller. Entropy and mutual information. *Department of Computer Science, University of Massachusetts, Amherst*, 2013.

[37] John G Proakis, Masoud Salehi, Ning Zhou, and Xiaofeng Li. *Communication systems engineering*, volume 2. Prentice Hall New Jersey, 1994.

[38] Yunus Can Gültekin, Wim J van Houtum, Semih Şerbetli, and Frans MJ Willems. Constellation shaping for ieee 802.11. In *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 1–7. IEEE, 2017.

[39] René-Jean Essiambre and Robert W Tkach. Capacity trends and limits of optical communication networks. *Proceedings of the IEEE*, 100(5):1035–1055, 2012.

[40] Thorsten Clevorn, Laurent Schmalen, Peter Vary, and Marc Adrat. On the optimum performance theoretically attainable for scalarly quantized correlated sources. In *Proc. of ISITA*, 2006.

[41] Laurent Schmalen. Probabilistic constellation shaping: challenges and opportunities for forward error correction. In *2018 Optical Fiber Communications Conference and Exposition (OFC)*, pages 1–3. IEEE, 2018.

[42] Junho Cho and Peter J Winzer. Probabilistic constellation shaping for optical fiber communications. *Journal of Lightwave Technology*, 37(6):1590–1607, 2019.

[43] Metodi P Yankov, Darko Zibar, Knud J Larsen, Lars PB Christensen, and Søren Forchhammer. Constellation shaping for fiber-optic channels with qam and high spectral efficiency. *IEEE Photonics Technology Letters*, 26(23):2407–2410, 2014.

[44] Dan Raphaeli and Assaf Gurevitz. An improved pragmatic turbo encoding scheme for high spectral efficiency using constellation shaping. In *IEEE International Conference on Communications, 2003. ICC'03.*, volume 4, pages 2738–2742. IEEE, 2003.

[45] GD Forney and LF Wei. Multidimensional constellations-part i: Introduction, figures of merit, and generalized cross constellations,"ieee j. select. areas commun., vol. 7, pp. 877-892, aug. 1989. si2 ieee transactions on communications, vol. 40, no. 3, march 1992 g. ungerboeck. *A fast encoding method for lattice codes and quantizers," IEEE Trans. Inform. Theory*, 29:820–824, 1983.

[46] David S Millar, Tobias Fehenberger, Toshiaki Koike-Akino, Keisuke Kojima, and Kieran Parsons. Distribution matching for high spectral efficiency optical communication with multiset partitions. *Journal of Lightwave Technology*, 37(2):517–523, 2019.

[47] Frank R Kschischang and Subbarayan Pasupathy. Optimal nonuniform signaling for gaussian channels. *IEEE Transactions on Information Theory*, 39(3):913–929, 1993.

[48] Patrick Schulte and Georg Böcherer. Constant composition distribution matching. *IEEE Transactions on Information Theory*, 62(1):430–434, 2015.

[49] T Fehenberger, DS Millar, T Koike-Akino, K Kojima, and K Parsons. Partition-based distribution matching. *arXiv preprint*, 2018.

[50] Shangxin Lin, Hua Shen, Zhili Lin, Lu Lin, and Ze Dong. Dynamic probabilistic shaping modulation based on fixed-to-fixed symbols projection constant composition distribution matching. In *2017 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, pages 102–107. IEEE, 2017.

[51] Ian H Witten, Radford M Neal, and JG LCEARY. Arithmetic coding for data compression. *Communications of the ACM*, 30(6):520–540, 1987.

[52] Junho Cho and Peter J Winzer. Multi-rate prefix-free code distribution matching. In *2019 Optical Fiber Communications Conference and Exhibition (OFC)*, pages 1–3. IEEE, 2019.

[53] Junho Cho, Sethumadhavan Chandrasekhar, Ronen Dar, and Peter J Winzer. Low-complexity shaping for enhanced nonlinearity tolerance. In *ECOC 2016; 42nd European Conference on Optical Communication*, pages 1–3. VDE, 2016.

[54] Alex Alvarado, David J Ives, Seb J Savory, and Polina Bayvel. On the impact of optimal modulation and fec overhead on future optical networks. *Journal of Lightwave Technology*, 34(9):2339–2352, 2016.

[55] Claude Ailett, Jacques Bigou, and Yves Bretecher. Pseudo-random binary sequence generator, April 29 1975. US Patent 3,881,099.

[56] Vishal Sharma and Richa Sharma. Analysis of spread spectrum in matlab. *Int. J. Sci. Eng. Res*, 5(1):1899–1902, 2014.

[57] MATLAB. *MATLAB Communications Toolbox*. The MathWorks Inc., Natick, Massachusetts, 2019.