



UNIVERSITY OF
CAMBRIDGE

Department of Engineering

High Capacity Digital Coherent Transceivers

Author Name: Mihai Varsandan

Supervisor: Prof. Seb Savory

Date: 27th March 2020

I hereby declare that, except where specifically indicated, the work submitted herin is my own original work.

Signed _____

A handwritten signature in black ink, appearing to read "Mihai Varsandan".

date 27th March 2020



UNIVERSITY OF
CAMBRIDGE

High Capacity Digital Coherent Transceivers

Mihai Varsandan (mv436)
Professor Supervisor: Prof. Seb Savory

27th May 2020

Contents

1 Technical Abstract	3
2 Introduction	5
3 Building the Simulation	13
3.1 Simulating AWGN Channel	14
3.1.1 Bits Input	14
3.1.2 Modulation	15
3.1.3 Upsampling and Downsampling	16
3.1.4 Pulse Shaping	16
3.1.5 AWGN Channel	17
3.1.6 Matched Filter	17
3.1.7 Results	18
3.1.8 Problems Encountered	19
3.2 Adaptive Equaliser	19
3.2.1 Dispersion	19
3.2.2 Adaptive Equaliser Algorithm	20
3.2.3 Results	23
3.2.4 Problems Encountered	24
3.3 Carrier and Phase Recovery	25
3.3.1 Laser Noise	25
3.3.2 Carrier Recovery	26
3.3.3 Phase Recovery	27
3.3.4 Results	28
3.3.5 Optimisation	30
3.3.6 Problems encountered	31
3.4 Non-Linearity	31
3.4.1 SSFT Algorithm	31
3.4.2 Results	33
3.4.3 Problems Encountered	34
3.5 Static Equaliser	34
3.5.1 Static Equaliser Algorithm	34
3.5.2 Results	35
3.5.3 Problems Encountered	36
3.6 Final Results and Conclusions	36
4 Probabilistic Constellation Shaping(PCS)	38
4.1 PAS Architecture	40
4.2 Distribution Matcher	41
4.3 Results and Conclusion	44
5 Final Report Conclusions	47
A Overlap and Save method	48
B Covid-19	49
References	50

1 Technical Abstract

Living in the world driven by economic growth requires engineers to constantly come up with innovations and developments. In the area of optical fibre network, the demand for bandwidth has pushed engineers to achieve serious progress. However, given that the volume of data consumed grew by 25% from last year, it will get harder and harder to sustain this growth. This project aims to show that a technique called *Probability Constellation Shaping(PCS)* will help in meeting this demand for increased bandwidth. This project contains two parts. The first part consists of building an accurate simulation of a real life long-haul optical fibre system. The second one is implementing the PCS technique to this new simulation and see whether an improvement in the capacity has been achieved.

Building the simulation requires having all three main components: *Transmitter*, *Channel* and *Receiver*. At first an *Additive White Gaussian Noise(AWGN)* channel communication has to be implemented and then more and more parts of a more complex system will be added. This model consists of sampling bits '1' or '0' from a Bernoulli distribution with probability 0.5. As bits '1' and '0' cannot be transmitted over a channel it is required to modulate them onto a signal. The modulations used in this system are: *Quadrature Phase-Shift Keying(QPSK)*, *16-Quadrature Amplitude Modulation (QAM)* and *64-QAM*. After the signal is upsampled by a specific factor, as it is required to have digital signal rather than a continuous signal that would have been in real life, the signal is then transformed into a specific shape whose purpose is to minimise the intersymbol interference(ISI).The shape chosen for this transmission is the root raised cosine(RRC). The next step is to send the signal into the channel. There AWGN noise is added to the signal in order to simulate noise added from an amplifier. At the receiver, the signal first goes through a Matched Filter whose job is to identify the RRC template in the noise and maximise the Signal-to-Noise Ratio(SNR). Then the signal is downsampled and passed into the Digital Signal Processing(DSP) block where further subsystems will be added.

After building this simple model, it is required to make it more complex so that it would resemble a real-life model. First, in the channel, dispersion is introduced by using an FIR filter with a certain number of taps. In order to compensate for this dispersion an Equaliser needs to be implemented at the DSP. An adaptive equaliser is introduced whose main purpose is to constantly vary the weights of a filter which can compensate for a varying dispersion. The algorithms used are a Constant Modulus Algorithm(CMA) for the QPSK constellation and Multi-Modulus Algorithm(MMA) for 16-QAM and 64-QAM.All these algorithms aim to change the filter weights so that symbols will lie on circles which form the constellations.

Next the laser noise is introduced. This one acts both at the transmitter due to the transmitting laser and at the receiver due to the local oscillator laser. The noise introduced in the signal is of the form of a Wiener process. Carrier Recovery is used at the DSP to find the right *Intermediate Frequency(IF)* in the system. The Phase Increment algorithm is used to obtain the required frequency. In order to correct for this laser noise a Phase-Recovery stage is implemented at the DSP. This uses a Viterbi-Viterbi Algorithm to determine the phase noise introduced and then compensate for it.

As in real-life transmission there is nonlinear noise which corrupts the signal, it is required to introduce this noise in our simulation as well. A wave propagation in a fibre is described by solving the Nonlinear Schrödinger Equation(NLSE). As this cannot be solved analytically,

a method called Split-Step Fourier Transform(SSFT) method is used to solve the equation numerically. This will also introduce the required dispersion and the nonlinearity in the signal.

The Equaliser is made of two parts, the Static Equaliser and the Adaptive Equaliser. The Static Equaliser is introduced before the Adaptive Equaliser as its purpose is to compensate for the bulk of the dispersion introduced in the signal. It consists of an FIR filter with fixed weights. By implementing this Static Equaliser the first part of this project is finished. Now a simulation that will resemble a real-life transmission has been accomplished. We will now look at implementing the PCS technique in this simulation.

It is important to first find a method of measuring the capacity of the channel which has been created by the simulation. This is done using Monte Carlo method. The capacity of an AWGN channel it is known to have lower bound as described by Shannon. Therefore it has been observed that there is a gap between the capacity of our simulation system and the Shannon Limit. PCS technique aims to reduce this gap by having a non-uniform distribution of the symbols rather than uniform distribution which has been used so far. The non-uniform distribution has to be an optimal distribution and has the form a Maxwell-Boltzmann distribution. In order to implement this distribution a Probability Amplitude Shaping(PAS) architecture is used at the Transmitter. The main component of the PAS architecture is the Distribution Matcher(DM) whose role is to transform uniform distributed bits into non-uniform symbols according to a certain probability. This is done by using a specific DM called Constant Composition Distribution Matching(CCDM). It uses a type class to find a codebook between a pattern of bits and the output symbols and then it encodes it using Arithmetic Coding. The reverse process is used at the output to recover the bits.

Once the PCS technique has been implemented and tested in the simulation, it was observed that indeed the gap between the Shannon limit and the uniform distribution capacity that was found before has been reduced. Therefore it can be said the PCS technique will help to increase the bandwidth of a transmission signal. However it can be observed that there is still room for further improvements. This is due to the fact our system does not compensate for the nonlinearity in signal. Further reduction of this gap could be achieved if some nonlinear equaliser would be used.

2 Introduction

Since the start of the century internet access has transitioned from an inessential service to a basic human necessity. Many of the modern businesses would not have existed provided there would be no internet connection. Internet access has been a pillar which has driven the modern economic growth and created many opportunities. The Internet provides a way for people to obtain widely available educational information, news and even new skills. Because of this, one might say that it could be a way to poor people to escape poverty. It has provided a way for companies to use the internet and the big data it includes, to boost their Research and Development(R&D) so that new products are more tailored and more targeted to specific people. Governments around the world have seen its potential and have been using it extensively in healthcare and education sector, but also in removing the bureaucracy within the departments. Less well appreciated is the boost to human happiness. Online games, widely available entertainment content and socialising have all helped increase our happiness. From messaging apps, which help friends stay in touch and let overseas workers say goodnight to their children back home, to dating apps which help people to meet strangers[1]. The development of smartphones has enhanced the power of the internet as people can now be all the time connected to the internet. In order to observe how people use their smartphones, in the figure below, the fragmentation of different activities that people are using on the smartphone can be observed.

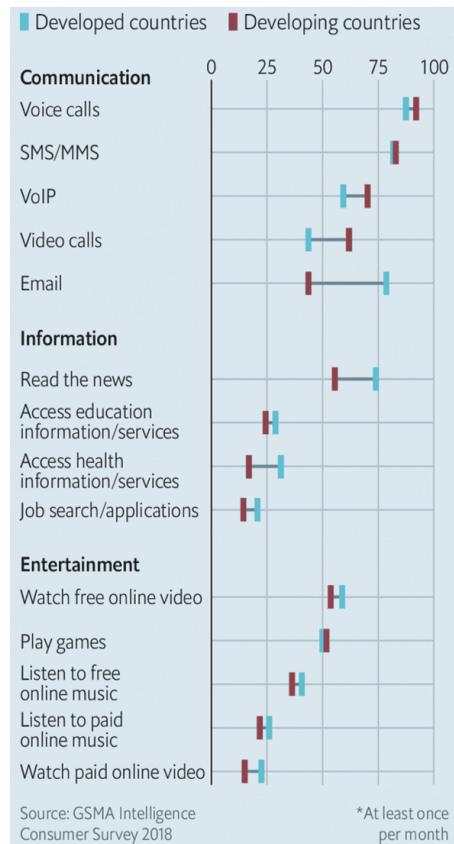


Figure 1: Smartphone Usage by type of activity in at least once a month [2]

As noticed in the figure above, all these activities would require some form of internet access. Therefore in order to meet this necessity, internet access should be widely available to everyone.

The governments and private companies around the world are investing in projects which are focused on bringing the internet to as many people as possible. This can be observed in the figure below:

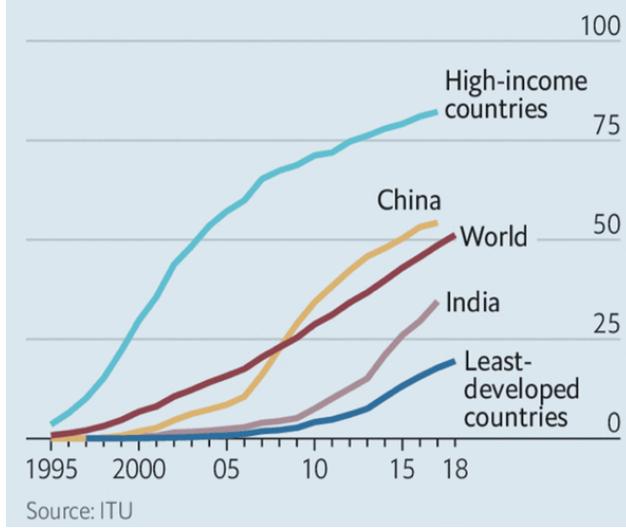


Figure 2: Population using internet % [2]

As more and more people are using the internet the amount of information required to be transmitted is increasing. According to Ofcom [3] in the UK the volume of data used on fixed and mobile connections both grew by around a quarter, with 240GB being used on average each month per fixed broadband connection and 2.9GB in an average month being used on each mobile data connection. But even if the internet access will reach a peak and everybody would have access to internet, the volume of data will continue to increase. This can be observed by taking a look at the situation that is happening right now. We are in the middle of the pandemic and a lot of people have to stay inside the house. Some people work from home using applications like Skype and Zoom to organise work meetings. Others choose to spend their time playing online games or watching streamed series or movies on websites such as Netflix or Amazon Prime. Because of this, the amount of data used has increased significantly which is observed in the *Figure 3* below provided by Cloudflare. As a consequence of this sharp demand increase, services such as Netflix had to reduce the quality of the streaming video in order to cope for this demand while in Spain the mobile operators have asked their users to reduce their data consumption. In the eventuality that this pandemic passes and things will go back to some 'normality', there will be other challenges further ahead which will put a strain on data rate such as emerging new technologies.

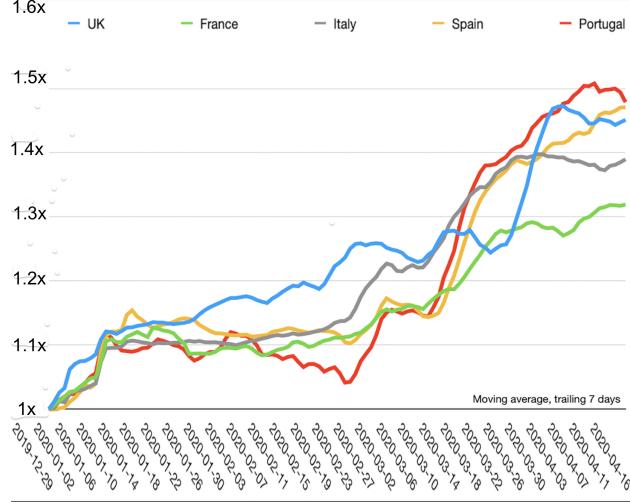


Figure 3: Relative change in internet using moving average and December 29, 2019 as the reference point [4]

Driven by the business growth opportunities and the prospect of enhancing human lives the following technologies: 5G, IoT and Cloud Computing will impact greatly the human life and businesses in the next year. Let's look at how these technologies will increase the demand for bandwidth. First, let's have a look at IoT and 5G which go hand in hand. This technology is ready to be deployed this year in most of the countries worldwide [5]. An IoT network based on 5G connections would allow a million devices to be clustered within a square kilometre, as they might be on a dense factory floor or at a crowded sports arena, far more than the 60,000 or so possible with 4G[5]. Having a 5G system will also enable downloads of about 20Gbps. These new prospects will push the consumers on the trend to consume more and more data. In order to cope with this increase in demand and the latency requirements for the 5G, mobile networks are required to increase the bandwidth available. The second technology which is already in usage but which will become more and more used is the Cloud Computing. This technology is growing at 37% per year with all the major technology companies using it. It has been seen during this pandemic exactly how useful a cloud server is in helping businesses survive when all its employees from around the world are working from home. The cost of this is of course bandwidth which should be widely available for these companies in order to help them survive so that they will continue to help to deliver economic growth. Therefore it is indeed observed that there is a potential problem out there. How to ensure that the increase in bandwidth matches demand? One way is through fiberisation and increasing the data rate within the fibre itself.

First, we are going to have a look at how the data rate in an optical fibre has evolved over the years. For this, we are going to have a look at Nielsen Law which in 1998 predicted that user bandwidth will increase by 50% every year. For a standard Integrated Service Digital Network(ISDN) line, it was observed that empirical data fits the exponential growth curve very closely. This can be observed in the figure below:

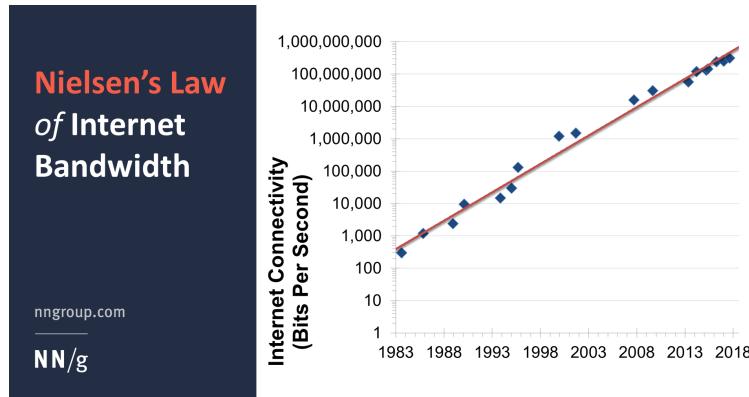


Figure 4: Nielsen's Law [6]

It is worth noticing that in 2019 a speed of 325 Mbps has been reached on an IDSN line. While this law might seem to apply to IDSN line it can certainly be applied to optical fibre speed as well. In the figure below we can observe data rates for optical fibre systems for the past 30 years:

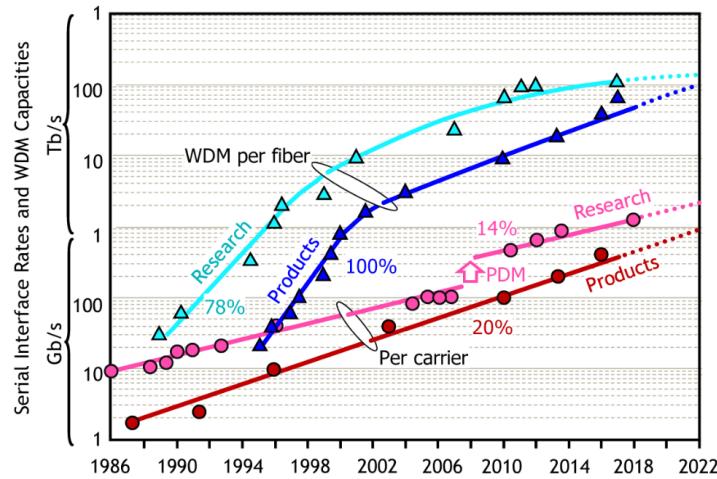


Figure 5: Comparison of products and research records in terms of per-carrier interface rates and Wave Division Multiplexing(WDM) capacities. The discontinuity in single carrier interface records around 2007 is due to the introduction of coherent detection using polarization division multiplexing (PDM).[7]

The exponential growth trend can be observed as well. Looking at the figure, it can also be spotted that there is a factor of 25 - 50 behind concurrent research records, corresponding to a time lag between research and commercialisation of 6 years. In general, the time lag between research and commercialisation can be seen from the figure to have consistently been between about 4 and 8 years[7]. To further understand what this project is about we need to first understand some history into the optical fibre field. Peter J. Winzer whose figure I have used has classified the fibre optic field development into four major periods:

- era of regeneration(1977-1995)
- era of amplified dispersion-managed systems(1995-2008)
- era of amplified coherent systems(2008-present)
- era of space division multiplexing(actively researched since 2008)

The era of regeneration consists in mainly increasing the net bit rate that the transceivers were able to support as the capacity of the early span-by-span regenerated fibre-optic transmission systems depended on this. During this, a technique of sending multiple light beams through a fibre at different frequencies called *Wave Division Multiplexing(WDM)* was invented. It allowed for optical fibre system to have multiple channels each at different frequency through which signal is transmitted at the same time. This invention proved to be later a stepping stone into achieving incredibly high transmission rates. To measure how efficient a WDM system is in transmitting data, spectral efficiency (SE) has been defined as measurement. It is the ratio between the data rate and the channel spacing(frequency difference between the channels). The first optical undersea transmission systems deployed across the Atlantic Ocean (TAT8) and Pacific Ocean (TPC3) in the late 1980s were systems operating at $1.3\text{ }\mu\text{m}$ and carrying 280 Mb/s on each of its three fibre pairs[7].

The second period, the era of amplified dispersion-managed systems started with the invention of practical optical amplifiers. Here, *Erbium-Doped Fibre Amplifier(EDFA)* played a key role in enabling the deployment of cost-efficient WDM systems. The rest of the period is dominated by research into limiting the non-linear effects into the fibre. *Dispersion Management*[8] has been introduced to mitigate this problem. Introducing dual-polarisation enabled us to double the rate of transmission but with it, a big problem has been brought up which is *Polarisation Mode-Dispersion(PMD)*. It caused unpredictable relative delays between the two polarisations carrying an optical signal. As this problem becomes more severe at high bit rates it appeared to be a big problem for bit rate scaling. However, the invention of a manufacturing technique called *GULP method* [9] has allowed us to delay the requirements for transceivers to have PMD compensation capabilities. In 1997 state-of-art optical fibre was a 16 channels at 2.5Gb/s with SE of 0.0125 b/s/Hz.

The third and fourth era both start at the same time with the fact that the speed of CMOS processing and CMOS-integrated analogue-to-digital converters (ADCs) and digital-to-analogue converters (DACs) had caught up with 10-GBaud symbol rates. The new CMOS capabilities enabled maximum-likelihood sequence estimation (MLSE) as well as digital electronic dispersion pre-compensation. This allowed the introduction of digital coherent receivers. Using a free-running local oscillator (LO) laser, they beat the signal roughly to baseband and convert the full optical signal field, i.e., its real and imaginary parts, also called its in-phase (I) and quadrature (Q) components, in both polarisations, to the digital electronics domain, giving them the name “intradyne receivers”. Digital access to the full optical field enables quadrature modulation (I/Q) and polarisation-division multiplexing (PDM) to increase spectral efficiency by a factor of $2\times 2 = 4$, allowing a 40-Gb/s receiver to be based on 10-Gb/s componentry, compatible with the capabilities of CMOS electronics of the time. Access to the full optical field in digital form also opens up the possibility of adaptive digital compensation of chromatic dispersion, PMD, optical filtering effects, and

even distortions from fibre nonlinearities[7] A schematic of a digital coherent receiver can be seen below:

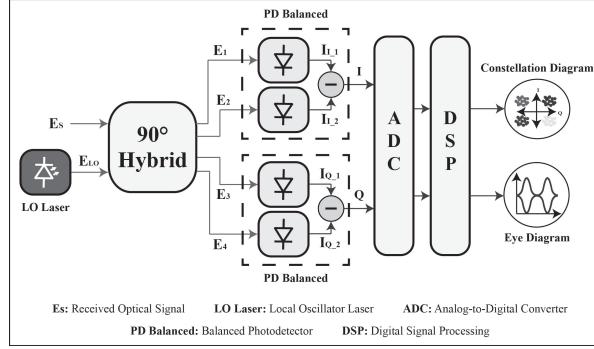


Figure 6: Schematic of Digital Coherent Receiver[10]

Continuous improvement in these domains and introducing *Optical Superchannels* which are optical interfaces consisting of multiple optical carriers that are co-generated, co-propagated, and co-detected, has allowed us to continue this exponential growth and improve the SE [11]. In 2018 the TAT12/13 transatlantic cables had been commissioned. Each cable had 2 fibre pairs and transported a single wavelength at 5 Gb/s per fibre over 5,913 km using erbium-doped fibre amplifiers (EDFAs) as repeaters, for an aggregate cable capacity of 20 Gb/s; the cable's aggregate capacity x distance product was 118 Tb/s.km. The highest-capacity submarine cable, the Pacific Light Cable Network (PLCN) carries an aggregate bidirectional 144 Tb/s across the Pacific, with an overall capacity x distance product across 6 fibre pairs of 3,686 Pb/s.km (counting both transmission directions separately).

This kind of growth has to be continued and even increased if we are to meet the demand for bandwidth required in the future years. However, WDM capacities are quickly approaching their fundamental Shannon limits. Higher-order modulation has been instrumental in the scaling of single-wavelength interface rates and even more critical to the scaling of WDM system capacity as it dictates the achievable SE, and hence the WDM capacity for a given system bandwidth. In the figure below it presents a summary of experimental research records in terms of their SE and their obtained transmission reach.

From *Figure 7* below a trade-off can be observed between SE and reach which follows a logarithmic path with each doubling in transmission reach reducing the SE by 2 bits/s/Hz[7]. For more details check [12].

The maximum capacity that can be reliably communicated per mode over an equivalent additive white Gaussian noise(AWGN) channel can be expressed as:

$$C = 2B \log_2(1 + \text{SNR}) \quad (1)$$

where the logarithmic term captures the maximum possible SE of a single-polarisation complex optical signal (in both in-phase and quadrature components). The preceding factor of 2 accounts for polarisation multiplexing and the final capacity is obtained by multiplying the dual-polarisation SE with the system bandwidth B.

From *Figure 7* it can be observed that there is a special technique which brought the unit cell bit rate closer to Shannon Limit. This technique is called *Probabilistic Constellation*

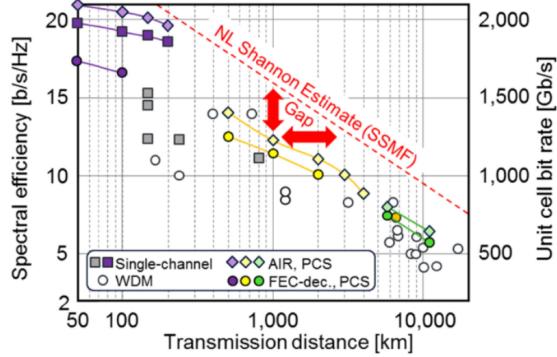


Figure 7: Experimentally achieved record SEs (in two polarizations) versus transmission distance (markers), shown with the Shannon limit estimate on SSMF (dashed).[7]

Shaping(PCS). The main idea about this technique consists of sending higher-energy symbols less frequently than lower energy symbols. Therefore rather than having a uniform distribution amongst all symbols, there is a higher probability of sending low energy symbols. The maximum gain by applying this technique is limited to $\approx 1.53dB$ and it is called *shaping gain*. This was one of the reasons why not until recently this technique has not attracted much popularity as most research focused on improving the *Forward Error Correction(FEC* which has allowed for a *coding gain* that reaches $10dB$ [13]. In the figure below a better visualisation of the shaping gain is observed:

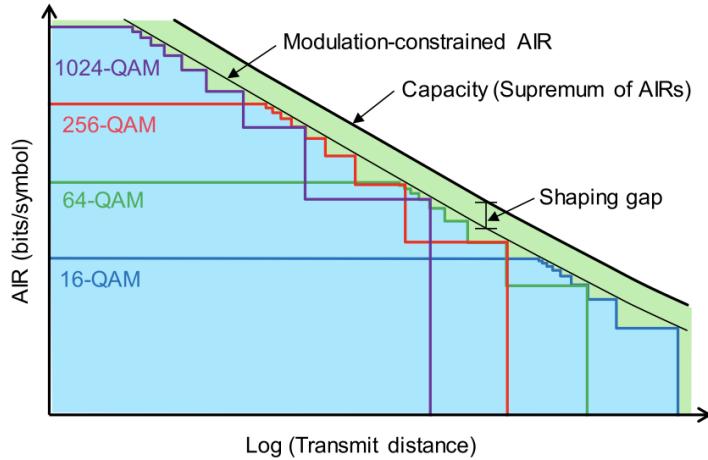


Figure 8: Schematic illustration of the *Achievable Information Rate(AIR)* of the auxiliary AWGN channel modelling an optical fibre channel. Upper solid line: Gaussian signalling (i.e., AWGN capacity), lower solid line: uniform QAMs with arbitrarily rate-adaptable FEC (i.e., modulation-constrained AIR), staircase lines: uniform QAMs with nine different fixed-rate FEC codes (i.e., modulation- and code-constrained AIRs)[13]

The capacity of the auxiliary channel described in the equation above does not represent the fundamental fibre channel capacity, but rather a lower bound of it. The largest recovered SNR, from which AIR can be estimated, of a fibre channel depends on the network scenario and on assumptions about what information is and is not known to the various transponders within the network. This leads to a variety of capacity estimates for the optical fibre channel[13]. Regardless of the sophistication of the optical fibre channel model, it is a general observation that capacity is maximised by a certain optical signal power. Furthermore, as both optical amplifier noise and nonlinear interference noise (NLIN) at optimised optical channel powers are, either exactly or to an excellent approximation for Gaussian signalling, linearly proportional to the transmission reach, the channel capacity decreases logarithmically with transmission distance in the high SNR regime, as illustrated by the upper solid line in *Figure 8*. Achieving the auxiliary AWGN channel capacity implies, at each transmission distance, the use of the optimally chosen variance of a Gaussian-shaped modulation. If we restrict ourselves to uniform square QAM constellations, the modulation-constrained AIR is decreased to below the modulation-unconstrained AIR (i.e., the capacity of the auxiliary channel), as indicated by the lower solid line in *Figure 8*, suffering a loss called the shaping gap due to the non-Gaussianity of the signal. Therefore instead of using a uniform distribution in choosing which constellation to send a probability distribution which would maximise the entropy is required. This distribution is called a Maxwell-Boltzmann distribution which is expressed mathematically as:

$$P_X(x_i) = \frac{1}{\sum_{j=1}^M e^{-\lambda|x_j|^2}} e^{-\lambda|x_i|^2} \quad (2)$$

where λ is a scaling factor and M is the total number of symbols. λ will be a parameter which will be optimised depending on the channel.

The aim of this project is to increase the capacity of a single carrier in an optical fibre transmission. That being the case we will try to replicate the technique of *Probability Constellation Shaping* and show that this shaping gain can be achieved and indeed a higher capacity of a transmission can be achieved. In order to implement this technique a simulation of basic optical fibre transmission needs to be build. Therefore the first part of the project will consist in building an accurate simulation of a long-haul optical fibre transmission system. We will start with implementing the most basic model, an AWGN channel. Then, we will introduce the dispersion and laser noise in the channel and the corresponding algorithms to compensate for them. Afterwards, we will use a numerical method to introduce nonlinearity in the channel. In the second part of the project we will introduce the PCS technique and explain how it can be put in place in our simulation. In the end we are going to run a final simulation with all the subsystems in place and observed whether or not the PCS technique has helped in increasing the capacity of the transmission.

3 Building the Simulation

As I mentioned before it is critically to have a suitable simulation in order to implement the PCS technique. This simulation needs to contain: *Transmitter*, *Channel* and *Receiver*. The first part of the project looks at implementing and simulating an optical fibre system in MATLAB. A block diagram with all the components required for this simulation is presented in the figure below:

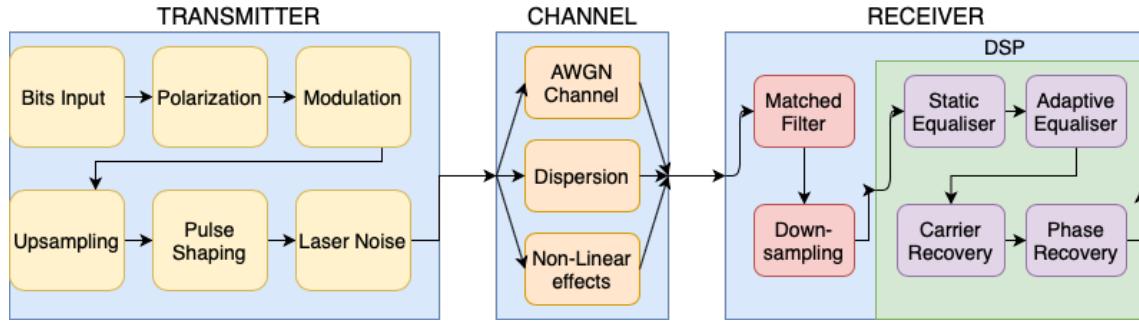


Figure 9: Block Diagram of an Optical Fibre System

First in order to build this simulation some parameters which will determine the nature of this simulation needs to be defined. These are:

Transmission Parameters

$$\text{Symbol Rate} = 28 \text{ Gbd}$$

$$\text{nPol} = 1$$

$$\text{nSyms} = 10^5$$

Laser Parameters

$$\text{roll-off} = 0.25$$

$$F = 8$$

$$\text{Linewidth} = 10^5 \text{ nm}$$

$$\text{PS taps} = 10$$

Channel Parameters

$$\text{Length} = 1000 \text{ km}$$

$$\text{A-spacing} = 100 \text{ km}$$

$$\beta_2 = -22.4 \text{ ps}^2/\text{km}$$

$$\alpha = 0.04605 \text{ km}^{-1}$$

$$\gamma = 1.3 \text{ W}^{-1}\text{km}^{-1}$$

Some clarifications about some of the parameters need to be addressed. n_{Pol} represents the number of polarisations used in this transmission. n_{Syms} is the total number of symbols transmitted. F stands in for the oversampling factor. PS taps denotes the number of taps used for the pulse shaping. A -spacing refers to the amplifier spacing when the *Split Step Fourier Transform (SSFT)* is used. β_2 represents the group velocity while α represents the attenuation of the fibre. γ is the non-linear coefficients which relates to the *Kerr* coefficient n_2 .

Even though this diagram has many components, which all are needed to be implemented to have an accurate representation of a real-life system, in the beginning we will start by modelling a very simple AWGN channel and then add more and more blocks until the full simulation layout will look like the one in *Figure 9*.

3.1 Simulating AWGN Channel

The pillar of the simulation consists of the AWGN basic channel as it is the most basic model of a communication system.. Its block diagram can be observed in the figure below:

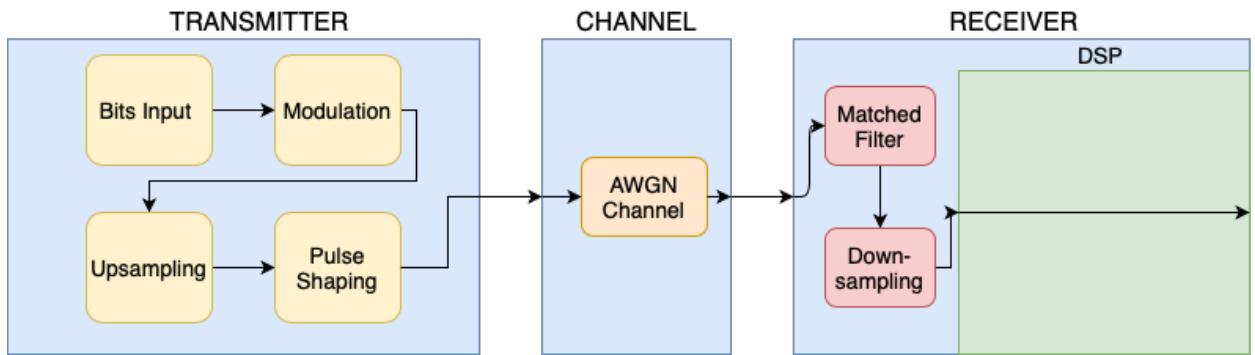


Figure 10: Block Diagram of a basic AWGN channel implementation

As observed this represents a simplified version compared to the diagram in *Figure 9*. Each of these block have been implemented in MATLAB and we will discuss the theory behind each of them.

3.1.1 Bits Input

We will start by first defining M as the number of different constellation used for this simulation. In this project M will take the following values: 4, 16 and 64. We will see in the next section the significance of this number. If a specific value of M is chosen then each symbol will represent $\log_2 M$ number of bits. In this case we have already defined that the total number of symbol for these simulation will be equal to $n_{\text{Syms}} = 10^5$ from which the total number of bits needed for the simulation can be expressed as:

$$n_{\text{Bit}} = n_{\text{Syms}} \times \log_2 M \quad (3)$$

In order to ensure that the probability of having bit 1 is the same as the probability of having bit 0 it is required to sample these bits from a Bernoulli distribution with $p = 0.5$. It

is observed due to *Law of large numbers* that if a large number of bits is sampled from this distribution, at the end of the sampling the probability that bit 1 will appear will be equal to 0.5018 which is very close to the 0.5.

3.1.2 Modulation

As mentioned above there are three values of M which we use: 4, 16 and 64. These correspond to three types of digital modulation: *Quadrature Phase-Shift Keying (QPSK)*, *16-Quadrature Amplitude Modulation (QAM)* and *64-QAM* respectively. This allows us to vary one or more elements of a periodic waveform such as amplitude and phase. To each of the symbols corresponding to these modulation formats, a unique pattern of binary bits are assigned. The structures of these modulations are displayed in the figure below:

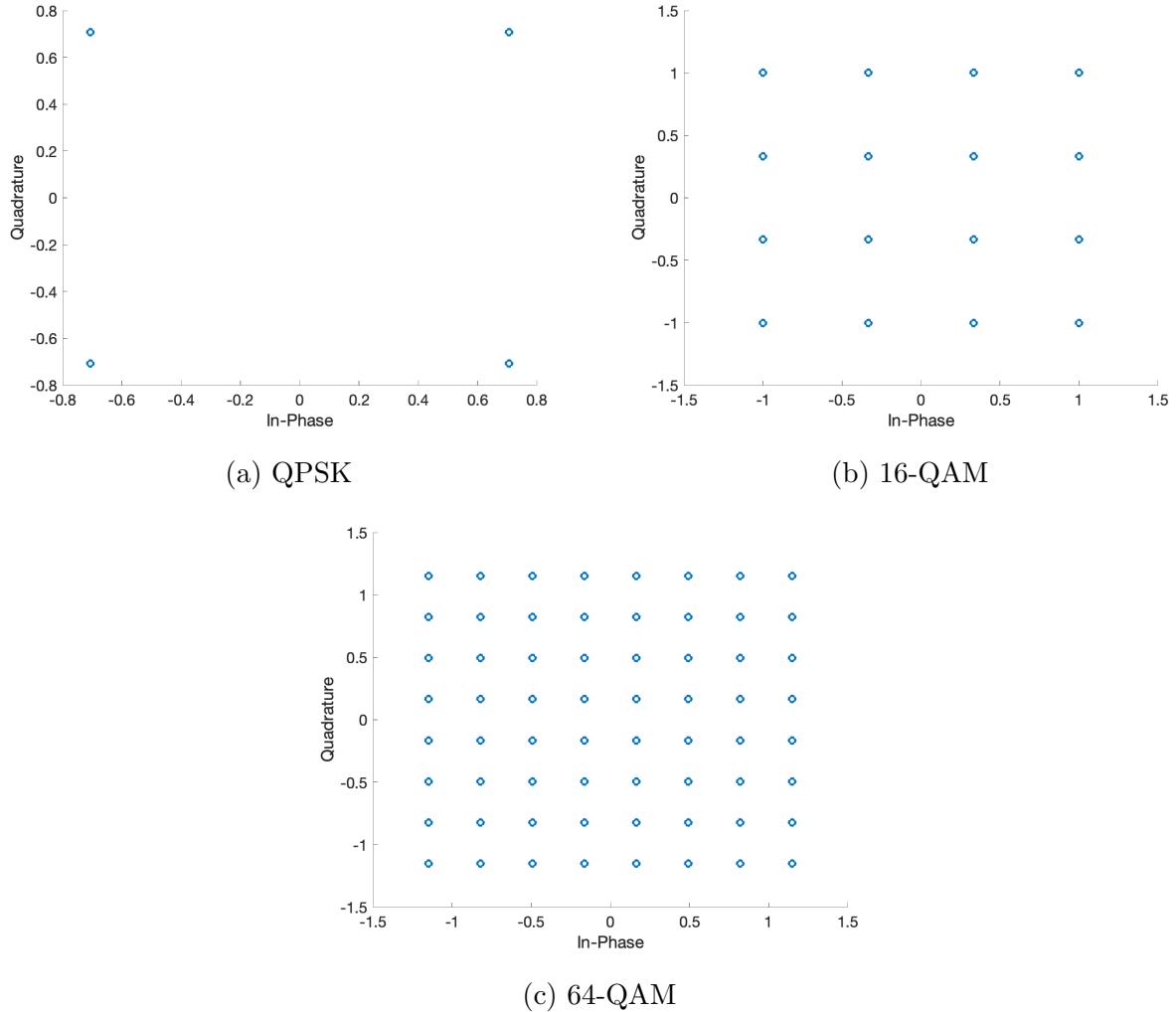


Figure 11: Types of digital modulation techniques

In the figure above *In-Phase* and *Quadrature* refers to the fact that every sinusoid can be expressed as the sum of a sine function (phase zero) and a cosine function ($\pi/2$).

The sine part is called the *In-Phase* component, the cosine part can be called the *Quadrature* component. These can be translated in complex system and therefore the Real part becomes the In-Phase component while the Imaginary part becomes the Quadrature component.

3.1.3 Upsampling and Downsampling

Since the simulation is in digital domain we will work with discrete samples of the signal. Therefore in order have a signal which will resemble the real signal we will have to upsample the signal. That means sampling the signal at a higher rate. At the receiver there is *Analog-to-Digital Converter(ADC)* which will downsample the signal. Downsampling is the reverse operation. In this project an upsampling factor of 8 is chosen which implies that the transmitted signal will have 8 samples per period. The reason why it is upsampled at such a large factor is that when SSFT will be used to model the channel, a high number of samples per symbol will provide a more accurate representation of the channel effects on the signal. Downsampling is used in order to conserve memory and reduce the signal processing time.

3.1.4 Pulse Shaping

It represents the process of changing the waveform of the signal to a specific form that will satisfy *Nyquist Criterion* which will then result in the minimum intersymbol interference(ISI). For this project a *Root-Raised Cosine (RRC) Filter* has been chosen as the shape in which the initial waveform will be transformed to using convolution. Its function is:

$$h(t) = \begin{cases} \frac{1}{T_{\text{samp}}} [1 + \beta(\frac{1}{\pi} - 1)], & t = 0 \\ \frac{\beta}{T_{\text{samp}}\sqrt{2}} [(1 + \frac{2}{\pi}) \sin(\frac{\pi}{4\beta}) + (1 - \frac{2}{\pi}) \cos(\frac{\pi}{4\beta})], & t = \pm \frac{T_{\text{samp}}}{4\beta} \\ \frac{1}{T_{\text{samp}}} \frac{\sin[\pi \frac{t}{T_{\text{samp}}}(1-\beta)] + 4\beta \frac{t}{T_{\text{samp}}} \cos[\pi \frac{t}{T_{\text{samp}}}(1+\beta)]}{\pi \frac{t}{T_{\text{samp}}}[1 - (4\beta \frac{t}{T_{\text{samp}}})^2]}, & \text{otherwise} \end{cases} \quad (4)$$

where β is the roll-off factor which it has been defined above to be 0.25. The shape of this filter is displayed in the figure below:

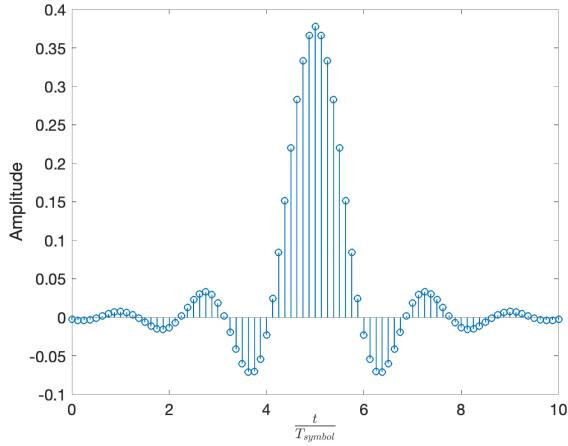


Figure 12: RRC function with 10 taps

The number of taps is chosen to be 10 which implies that the RRC filter spans 10 samples period. RRC has been chosen as the filter because it is required to use it both at the Transmitter(TX) and at the Receiver(RX), after which the shape will become a *Raised Cosine(RC)* which will indeed minimise the ISI.

3.1.5 AWGN Channel

As it has been mentioned above, it represents a basic model used to mimic the effect of noise introduced from external factors on the signal such as amplification noise, when it goes through the channel. Its characteristics are[14]:

- *Additive* because it is added to any noise that might be intrinsic to the information system.
- *White* refers to the idea that it has uniform power across the frequency band for the information system.
- *Gaussian* because it has a normal distribution in the time domain with an average time domain value of zero.

We will define X_Q and X_I the quadrature and in-phase components of signal at the output of the transmitter and respectively Y_Q and Y_I the quadrature and in-phase components of signal at the input of the receiver. The effect of the AWGN will cause the values at the receiver to be:

$$\begin{aligned} Y_Q &= X_Q + N_Q \\ Y_I &= X_I + N_I \end{aligned} \tag{5}$$

where N_Q and N_I are sampled independently from a normal Gaussian distribution.

3.1.6 Matched Filter

In signal processing, a matched filter is obtained by correlating a known delayed signal, or template, with an unknown signal to detect the presence of the template in the unknown signal. This is equivalent to convolving the unknown signal with a conjugated time-reversed version of the template. The matched filter is the optimal linear filter for maximising the signal-to-noise ratio (SNR) in the presence of additive stochastic noise such as AWGN[15]. In the context of this project the template is the RRC filter described in the subsection above. And therefore the output of the matched filter will be given by:

$$Y[n] = \sum_{k=1}^N h[n-k]X[k] \tag{6}$$

where N is total number of samples in the signal, h represents the RRC filter coefficients and X the signal input.

3.1.7 Results

Each of the block in *Figure 10* has been described and implemented in Matlab. In order to test how well this system performs a *Bit Error Rate(BER)* plot is plotted against the *Signal-to-Noise Ratio(SNR)*. If every block was implement correctly we would expect the practical results to follow a theoretical curve which can be found mathematically. The theoretical results for each of the modulations is:

$$\text{BER}_{\text{QPSK}} = \frac{1}{2} \operatorname{erfc} \left(\sqrt{\frac{E_b}{2N_o}} \right) \quad (7)$$

$$\text{BER}_{16\text{-QAM}} = \frac{3}{8} \operatorname{erfc} \left(\sqrt{\frac{E_b}{10N_o}} \right) \quad (8)$$

$$\text{BER}_{64\text{-QAM}} = \frac{7}{24} \operatorname{erfc} \left(\sqrt{\frac{E_b}{42N_o}} \right) \quad (9)$$

where E_b is the energy of the of a bit while N_o represent the amount noise introduced. By varying the amount noise introduced in the simulation a BER curve for the simulation can be plotted as well. The comparison between the practical simulation and the theoretical results is displayed in the figure below:

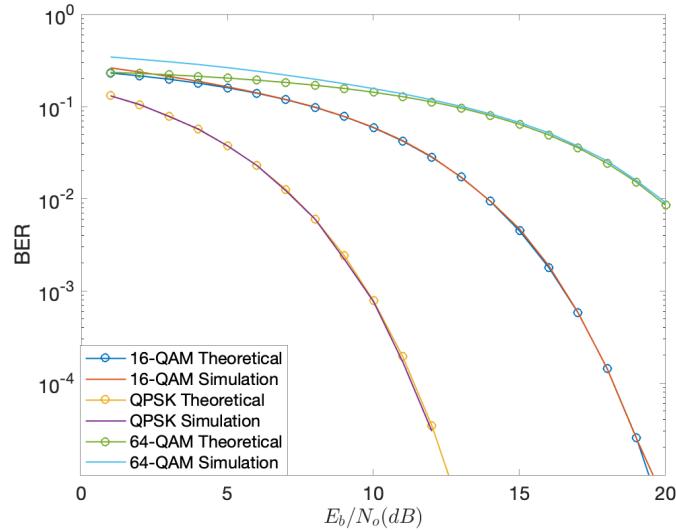


Figure 13: BER plot of Theoretical vs Simulation Results

As observed in the figure above, the simulation implemented matches the theoretical values. Therefore it can be said that this model is a good representation of an AWGN channel.

3.1.8 Problems Encountered

Implementing this basic simulation had not been problem free. The first problem is related to how noise is added to the channel. It was important to notice that noise had to be added both in a real and imaginary part of the transmission. The second problem encountered was normalising the transmission at the receiver to unit energy. Making everything consistent prevents from any scaling errors that could be introduced, for example when the matched filter is applied when the system undergoes the change from a root raised cosine pulse to a raised cosine pulse.

3.2 Adaptive Equaliser

The fibre itself is a dispersive system in which the group speed of light changes with the frequency of the optical wave. In order to make the simulation as accurate as possible the dispersion needs to be introduced in the channel. A receiver tackles this dispersion in two ways. Using a static equaliser and an adaptive equaliser. This section will focus on adaptive equaliser as the static equaliser will be introduced in the last subsection. The block diagram corresponding with adding the dispersion and adaptive equaliser component can be observed in the figure below:

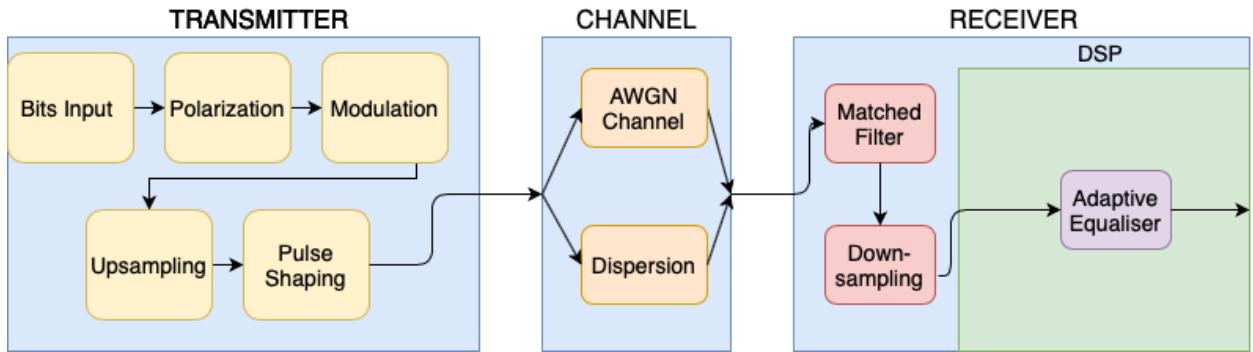


Figure 14: Simulation Block Diagram including the Adaptive Equaliser and Dispersion

As observed, this block diagram gets much closer to *Figure 9* than *Figure 10*. Each additional block will now be further discussed.

3.2.1 Dispersion

Fibre dispersion can be divided in two parts: material dispersion and waveguide dispersion. Material dispersion is due to the frequency dependence of the refractive index of glass. Just like a prism spreads white light into a rainbow of colours, different frequency components travel at different speeds in glass leading to pulse spreading [16]. The second contribution to fiber dispersion dispersion comes from the waveguide effect in which the wave's phase velocity in a fibre depends on the geometry of the fibre.

In order to observe from where the dispersion coefficient β_2 comes from we will first define

the inverse group speed as:

$$\beta_1 = \frac{1}{v_g} = \frac{dk}{dw} \quad (10)$$

where k is the propagation constant. Therefore the dispersion coefficient is defined as:

$$\beta_2 = \frac{d\beta_1}{dw} = \frac{d^2k}{dw^2} \quad (11)$$

Given a slowly varying envelope of the field (caused by modulation of the optical field to transmit data), $A(t, z)$ where t is the time components and z the distance components. Its Fourier Transform will be defined as $\tilde{A}(w, z)$. The Fourier Transform of envelope field at a certain distance L will be equal to:

$$\tilde{A}(w, L) = \tilde{A}(w, 0) \exp\left(-j\frac{\beta_2}{2}w^2L\right) = \tilde{G}(w, L)\tilde{A}(w, 0) \quad (12)$$

It is observed that dispersion can be introduced in the channel via an FIR filter of a certain length or via a numerical method. Initially an FIR filter is implemented in order to be able to test and improve at a higher rate the adaptive equaliser. Using the equation above, filter coefficients are given by taking the inverse Fourier Transform of $\tilde{G}(w, L)$ and discretize. The value of the coefficients are found to be:

$$H[k] = T_{\text{samp}} \sqrt{\frac{1}{-2\pi j\beta_2 L}} \exp\left(\frac{-jk^2T_{\text{samp}}^2}{2\beta_2 L}\right) \quad (13)$$

3.2.2 Adaptive Equaliser Algorithm

To counter for this introduced dispersion and for any other fluctuations in the dispersion due to other conditions such as environmental conditions, an adaptive equaliser is implemented. It consists of a FIR filter,

$$y[n] = \sum_{k=-K}^{K} W[k]x[n-k] \quad (14)$$

whose filter taps are adjusted adaptively based on some criterion and where $2K + 1$ is the number of taps the filter is using. Initially the equaliser employed a time-slot(recurring periodically for time-varying conditions) during which a training signal, known in advance by the receiver would be sent. Then using the *Least-Mean Squared(LMS)* criterion it would adapt its filter taps so that its output would match closely the known reference training signal. However using this type of adaptive equaliser meant that the inclusion of the training signal would sacrifice valuable channel capacity[17]. Therefore it was required another type of adaptive equalisation called *blind equalisation* which does not require any reference training signal. For this kind of equaliser, another algorithm must be used to train the filter taps which is called *Constant Modulus Algorithm(CMA)*.

We begin first by looking at QPSK modulation. The CMA algorithm takes advantage of the fact that the constellations in the QPSK modulation lie on the unit circle so that its criterion for adapting the filter weights will rely on this. In the figure below in can be observed that the QPSK constellation lie on a circle of radius 1.

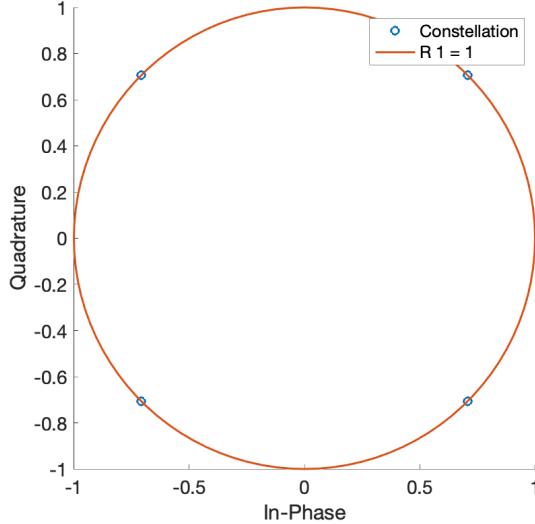


Figure 15: Unit Circle encompassing QPSK constellations

Therefore using this characteristic of the QPSK constellation the CMA algorithm can be implemented in the following way.

Algorithm 1 Constant Modulus Algorithm(CMA)

```

 $K \leftarrow 16$ 
 $\mathbf{W} \leftarrow [\underbrace{0, \dots, 0}_K, 1, \underbrace{0, \dots, 0}_K]$ 
 $\mu \leftarrow 0.002$ 
for  $n \leftarrow 1 : N$  do
     $y[n] \leftarrow \mathbf{x}[n : n + 2K]\mathbf{W}^T$ 
     $\epsilon[n] \leftarrow y[n](1 - |y[n]|^2)$ 
     $\mathbf{W} \leftarrow \mathbf{W} + 2\mu\epsilon[n]\mathbf{x}^*[n : n + 2K]$ 
end for

```

In this algorithm μ represents the incremental step towards the optimum filter weights and ϵ is the error value. It is really important to understand the initialisation of the filter taps. It is observed that the $K + 1$ weight is initialised to 1. This is because of two reasons: the first one is that it is desired to make the filter of unit energy in order not to introduce any amplification in the signal and as such the amplitude of 1 is selected. The second reason is due to the fact in the equation of the FIR filter above the x value corresponds to the $y[n]$ is at $k = 0$ which in this case will be $K + 1$.

It is observed that error criterion represents the distance between the symbol and unit circle. Therefore the algorithm is a form of optimiser which will try to adapt the filter weights such that this error criterion is as small as possible. It is observed that the performance of the filter is dependent on two variables: K and μ . In order to have the best possible performance some form of optimisation is required to find the right values for these parameters.

However if a more complicated modulation technique such as 16-QAM and 64-QAM are chosen (*Figure 11b*, *Figure 11c*) it can be observed how CMA would encounter a problem. Because there is no unique circle on which the constellations would lie on, it is required to move away from CMA to another algorithm. For this purpose a new algorithm is chosen: *Multi-Modulus Algorithm(MMA)*. It relies on the fact that 16-QAM and 64-QAM constellations are placed on multiple circles as observed in the figure below:

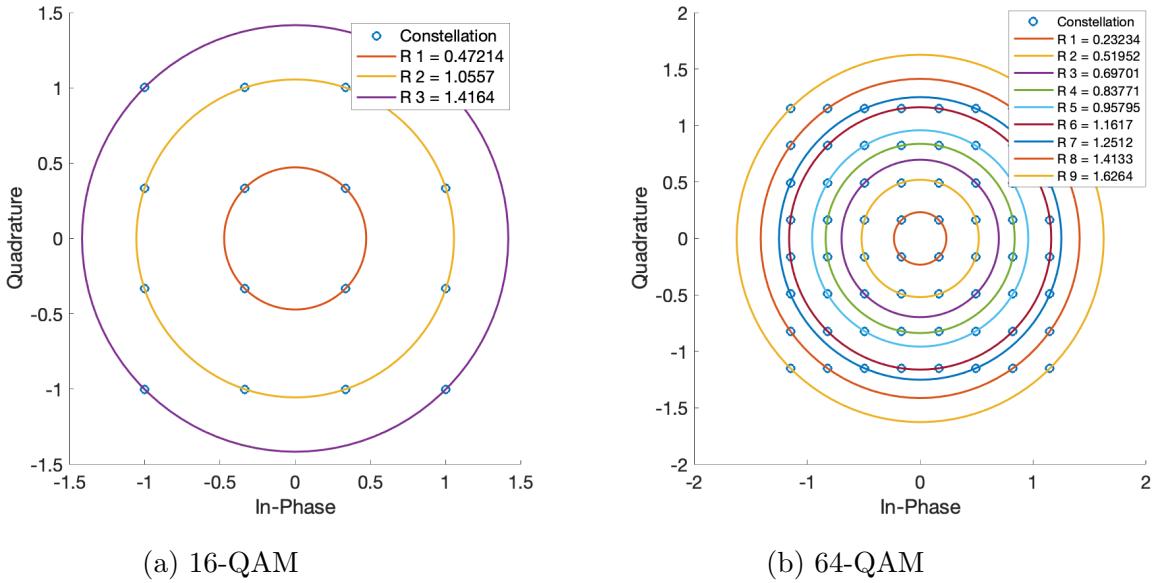


Figure 16: Circles encompassing 16-QAM and 64-QAM constellations

MMA relies on CMA algorithm to place constellations on the circle but instead of having a unit circle, the radius of the circle will be chosen depending on where the output of the adaptive equaliser is placed on the constellation plot. 16-QAM is chosen as an example for the description of MMA because of its simplicity compared to 64-QAM. However it would be observed how easily this algorithm can be modified for 64-QAM.

In the algorithm description below it can be observed that the same initialisation of the filter weights \mathbf{W} is used and the same values of K and μ are used. R_1, R_2 and R_3 represent the radius of the circles that are used in the 16-QAM which can be observed in *Figure 16a*. Therefore it can be noticed that the algorithm will select the radius that is the closest to the current constellation and then apply the CMA criterion with that radius instead of 1 in order to place that constellation on that circle.

As mentioned before the 64-QAM MMA algorithm can be written in the same as way as in the *Algorithm 2* but with the radius that are used to describe the circles in the 64-QAM constellation.

Algorithm 2 Multi-Modulus Algorithm(MMA)

```

 $K \leftarrow 16$ 
 $\mathbf{W} \leftarrow [\underbrace{0, \dots, 0}_K, 1, \underbrace{0, \dots, 0}_K]$ 
 $\mu \leftarrow 0.002$ 
for  $n \leftarrow 1 : N$  do
     $y[n] \leftarrow \mathbf{x}[n : n + 2K]\mathbf{W}^T$ 
    if  $\Re(y[n]) < \frac{\sqrt{R_1} + \sqrt{R_2}}{2}$  then
         $R_R \leftarrow R_1$ 
    else if  $\Re(y[n]) < \frac{\sqrt{R_2} + \sqrt{R_3}}{2}$  then
         $R_R \leftarrow R_2$ 
    else
         $R_R \leftarrow R_3$ 
    end if
    if  $\Im(y[n]) < \frac{\sqrt{R_1} + \sqrt{R_2}}{2}$  then
         $R_I \leftarrow R_1$ 
    else if  $\Im(y[n]) < \frac{\sqrt{R_2} + \sqrt{R_3}}{2}$  then
         $R_I \leftarrow R_2$ 
    else
         $R_I \leftarrow R_3$ 
    end if
     $\epsilon_R \leftarrow \Re(y[n])(R_R - |\Re(y[n])|^2)$ 
     $\epsilon_I \leftarrow \Im(y[n])(R_I - |\Im(y[n])|^2)$ 
     $\epsilon[n] \leftarrow \epsilon_R + j\epsilon I$ 
     $\mathbf{W} \leftarrow \mathbf{W} + 2\mu\epsilon[n]\mathbf{x}^*[n : n + 2K]$ 
end for

```

3.2.3 Results

Both of the blocks Dispersion and Adaptive Equaliser have been described and implemented in MATLAB. The results of the simulation for all the three modulation formats can be observed in the figures below.

As noticed, both the CMA and MMA algorithms correctly manage to compensate for the dispersion introduced in the Channel. In the *Figures 17c, 18c and 19c* it can be observed that, as predicted, the algorithms will try to reduce the error criterion, which indeed happens. Because of this fact, it will take some time before the weights will be optimised so some truncating of the symbols in the beginning is required.

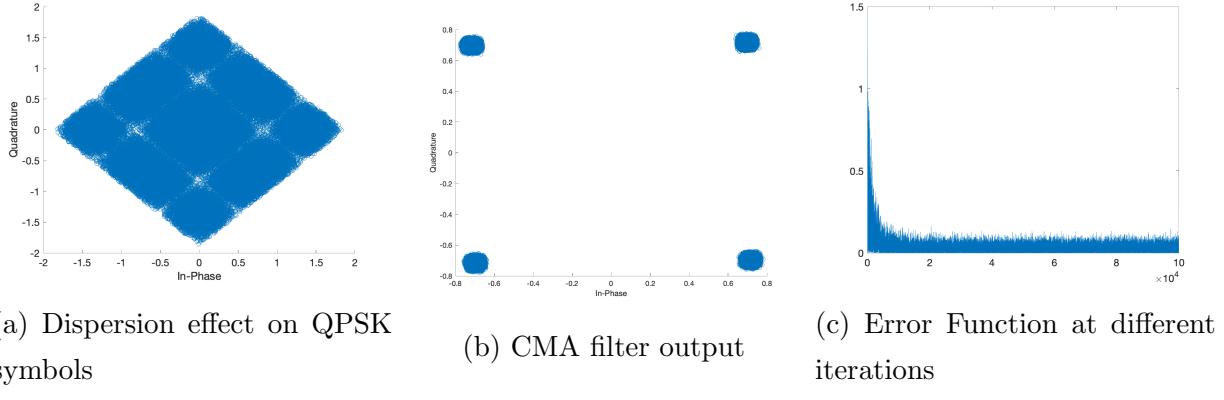


Figure 17: CMA Filter applied to QPSK constellation

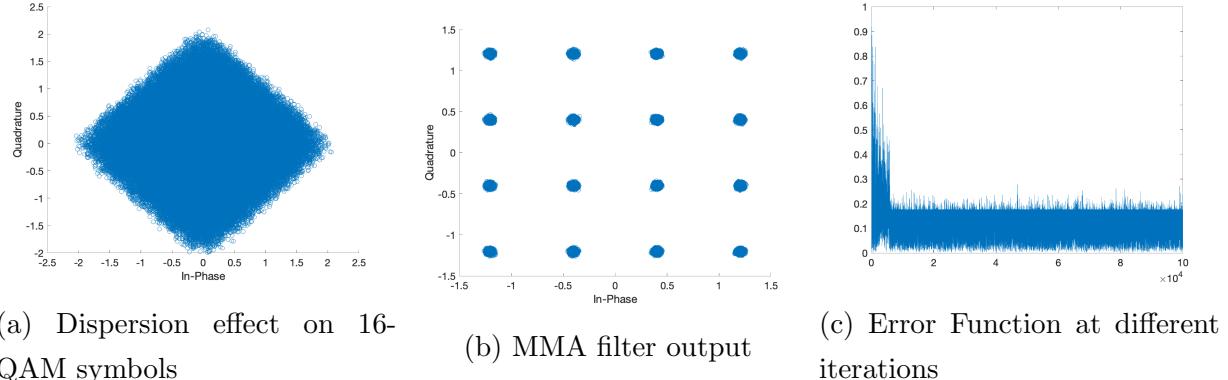


Figure 18: MMA Filter applied to 16-QAM constellation

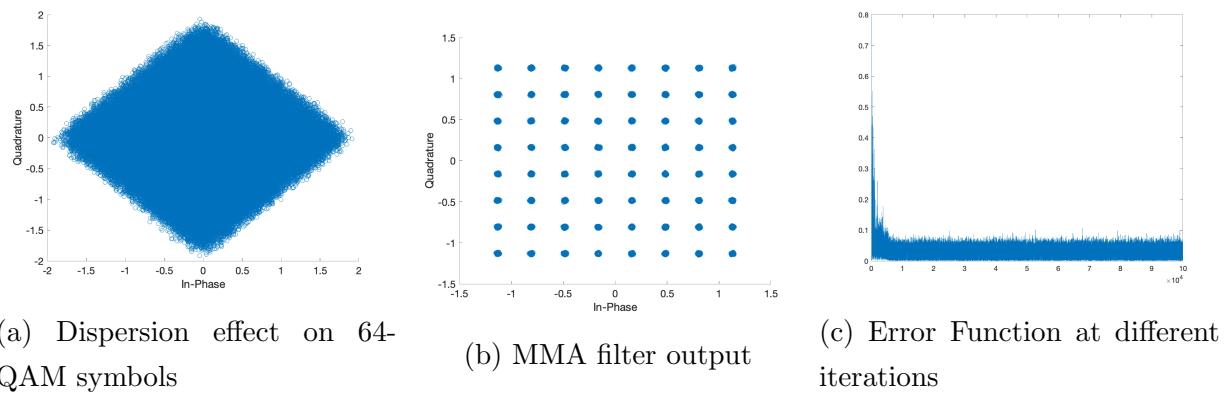


Figure 19: MMA Filter applied to 64-QAM constellation

3.2.4 Problems Encountered

One of the issues encountered and which caused some errors in the simulation was having a suitable length of channel memory and a sensible dispersion. Another problem encountered was the fact that CMA was used with 2 samples per period and, as expected, it failed to move

all constellations on the unit circle. The last obstacle that was overcome was the initialisation of the tap weights for the CMA filter which had to be of unit energy and be considerate of the channel memory.

3.3 Carrier and Phase Recovery

In a real simulation both Carrier and Phase Recovery are essential as phase noise appear due to different reasons. Carrier Recovery is there to compensate for an *Intermediate Frequency(IF)* offset which cause the absolute value of the phase to increase with time and could lead to erroneous phase decisions. Phase Recovery helps the receiver to compensate for the phase errors introduced by the transmitter and the local oscillator(LO) lasers. The output of a single-frequency laser is not strictly monochromatic but rather has frequency deviations that change randomly[16]. This will induce a random change in phase on each symbol. The block diagram of the optical fibre system with these additions will look in the following way:

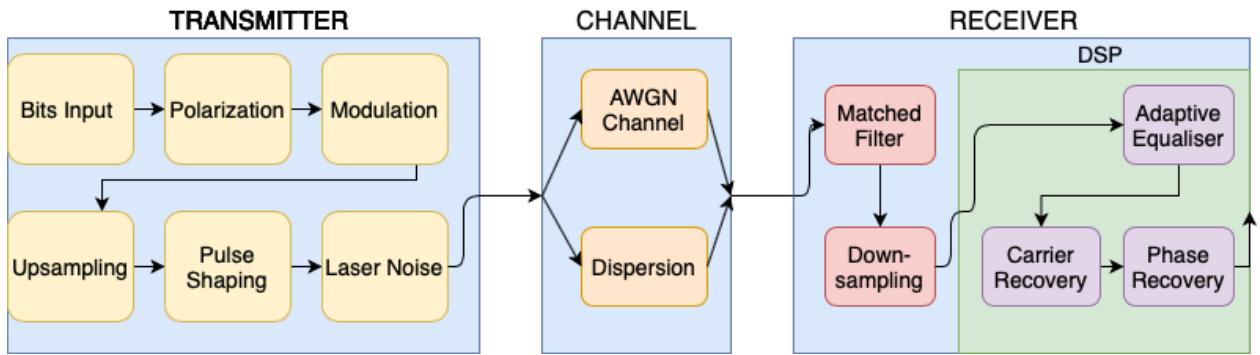


Figure 20: Simulation Block Diagram including the Laser Noise, Carrier Recovery and Phase Recovery

As noticed the Transmitter block diagram now completely matches the one in *Figure 9*. We will now go through each of the new blocks added.

3.3.1 Laser Noise

As mentioned above laser noise is a random process of introducing noise in the phase of the symbols. We will define $\phi(t)$ as the laser phase noise. Therefore the instantaneous frequency deviation can be written as:

$$f_i = -\frac{1}{2\pi} \frac{d\phi}{dt} \quad (15)$$

Because the process is random, the instantaneous frequency deviation is a zero-mean Gaussian noise process with variance σ_f^2 that is equal to:

$$\sigma_f^2 = \frac{\Delta v}{2\pi T_{\text{samp}}} \quad (16)$$

where Δv represents the linewidth of the laser. Integrating *Equation 15* and observing that the integration can be replaced by the rectangle rule, will result in:

$$\phi(t) = \phi(t - T_{\text{samp}}) - 2\pi f_i(t - T_{\text{samp}})T_{\text{samp}} \quad (17)$$

which after discretisation will become:

$$\phi_k = \phi_{k-1} - 2\pi f_{i,k-1}T_{\text{samp}} \quad (18)$$

This equation is equivalent to a Wiener Process. It is observed that the phase of sample k is incremented by $-2\pi T_{\text{samp}} f_{i,k-1}$, where $f_{i,k-1}$ is a value of instantaneous frequency picked from the Gaussian distribution. The phase noise can also be interpreted as a one-dimensional random walk [16]. The phase noise evolution for the transmitted signal can be visualised in the figure below:

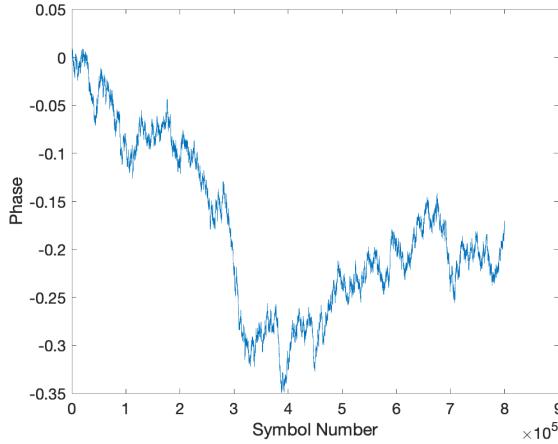


Figure 21: Phase Noise added to the transmission due to Laser Noise

As observed in the figure above the phase noise is random. If the variance is increased, for example due to larger linewidth, the fluctuations will also increase.

3.3.2 Carrier Recovery

As mentioned above, there is this constant offset which causes the phase to increase with time. Carrier Recovery job is to remove this offset. In the absence of laser phase noise the shift between two consecutive symbols y_k, y_{k+1} is:

$$\Delta\theta = 2\pi f_{IF}T_{\text{samp}} \quad (19)$$

The objective of the frequency estimator is to estimate the phase shift $\Delta\theta$ between two consecutive samples. For this, an algorithm called *Phase Increment Algorithm* has been chosen to accomplish the task. The algorithm takes advantage of the fact that:

$$\arg\{(y_k y_{k-1}^*)^M\} = -(2\pi f_{IF}T_{\text{samp}})M \quad (20)$$

or

$$\hat{f}_{IF} = \frac{-1}{2\pi f_{IF} T_{\text{samp}} M} \arg \{(y_k y_{k-1}^*)^M\} \quad (21)$$

where M is the number of unique symbols and \hat{f}_{IF} is the estimated intermediate frequency. However if other phase noises sources such as laser phase noise would be applied to the signal the estimated frequency offset would fluctuate from symbol to symbol. These fluctuations would be minimised if we average over N samples:

$$\hat{f}_{IF} = \frac{-1}{2\pi f_{IF} T_{\text{samp}} M} \arg \left\{ \sum_{k=1}^N (y_k y_{k-1}^*)^M \right\} \quad (22)$$

Therefore the frequency estimate \hat{f}_{IF} will get better as the block size N increases, as long as \hat{f}_{IF} remains constant over the block size. The IF offset is removed by multiplying y_k by $\exp(-j\Delta\theta)$, where $\Delta\theta = 2\pi\hat{f}_{IF}T_{\text{samp}}$.

When this block diagram was created it was desired that this project would be able to simulate a WDM system with a standard 50 GHz frequency spacing. Because it would be extremely hard to simulate the waveform if the true frequency was used, a reference frequency of 0 Hz is used which corresponds to a baseband signal. Therefore for a WDM system, it would have channels at $\dots, -100 \text{ GHz}, -50 \text{ GHz}, 0 \text{ GHz}, 50 \text{ GHz}, 100 \text{ GHz}, \dots$ for a certain number of channels. Due to insufficient time available the WDM system had to be given up on but the carrier frequency was still implemented even though a single carrier with $f_{IF} = 0 \text{ Hz}$ was used. This implies that the offset would be equal to 0 which cause no change in phase and therefore for a single carrier system simulation it is not required to have carrier recovery.

3.3.3 Phase Recovery

The purpose of phase recovery is to remove the noise accumulated in the signal due to Transmitter Laser and Local Oscillator Laser. The idea behind this compensation relies on the fact that if the phase $\Delta\phi_k$ is averaged over many symbol intervals, it is possible to obtain an accurate phase estimate. The algorithm used to compensate for the phase noise is called *Viterbi-Viterbi Algorithm*[16]. It is first looked at QPSK modulation and then more complex modulations will be addressed.

First, the signal is divided into K blocks with each block consisting of N samples, as shown in the figure below. In the block k , $k = 1, 2, \dots, K$, the signal is raised to the M th power and summed over N samples to obtain

$$\sum_{l=(k-1)N+1}^{kN} (\tilde{y}_l)^M \approx A_0 N \exp(-jM\Delta\phi_l) \quad (23)$$

where \tilde{y}_l represents the input symbol into this algorithm, A_0 is the energy of the symbol which for QPSK is constant and M is the number of unique symbols which have energy A_0 . Therefore the phase noise estimate $\Delta\phi_l$ can be found to be

$$\Delta\phi_l = \frac{-1}{M} \arg \left\{ \sum_{l=(k-1)N+1}^{kN} (\tilde{y}_l)^M \right\} \quad (24)$$

The compensated signal at the output of this algorithm is x_l and is found by multiplying the input sample by $\exp(j\Delta\phi_l)$. The block size should be chosen carefully. If N is too small the approximation in the equation above would not hold. If N is too big, the laser phase may drift and $\Delta\phi_l$ may not remain constant within each block [16].

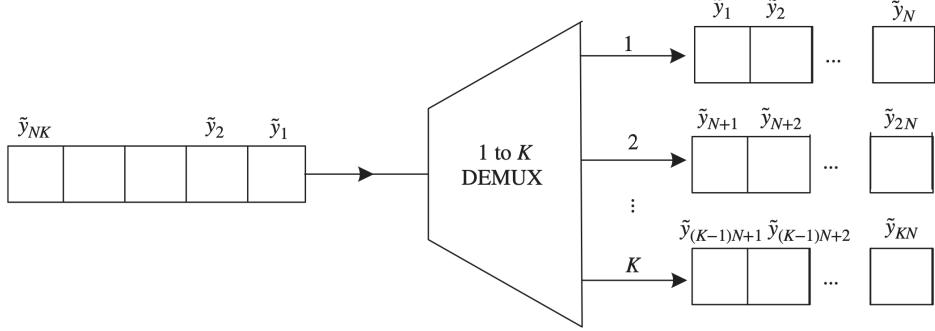


Figure 22: Demultiplexing of the data into K blocks with each block consisting of N samples[16]

It is noticed that if more complex constellations such as 16-QAM and 64-QAM the approximation above fails as A_0 will not be constant due to the fact these constellations have multiple absolute energy values which can be observed in *Figures 16a and 16b*. Therefore to correct for this it is required that before putting the symbols into K blocks only symbols which belong to the same energy value should be selected.

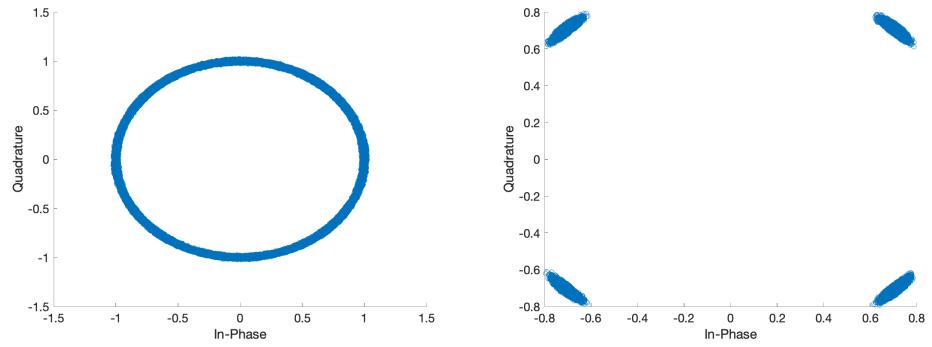
There is yet still another issue with the algorithm which needs to be resolved. The function $\arg()$ in the equation above cannot distinguish between phases that differ by 2π and it returns the results in the interval $[-\pi, \pi]$. If the phase is $\pi + \epsilon$, $\epsilon > 0$, the function $\arg()$ returns a phase of $-\pi + \epsilon$. This is known as *Phase Wrapping* and it could lead to symbol errors. Let the carrier phase prior to the unwrapping for a block k to be $\Delta\tilde{\phi}_k$. The phase wrapping will be corrected if we add an integer multiple of $2\pi/M$ to $\Delta\tilde{\phi}_k$. As a result the phase for the k th block after the phase unwrapping will be:

$$\begin{aligned}\Delta\phi_k &= \Delta\tilde{\phi}_k + m \frac{2\pi}{M}, \quad k = 1, 2, \dots, K \\ m &= \left\lfloor 0.5 + \frac{\Delta\phi_{k-1} - \Delta\tilde{\phi}_k}{2\pi/M} \right\rfloor\end{aligned}\tag{25}$$

Therefore the phase unwrapping needs to be included after the Viterbi-Viterbi Algorithm in order to correctly compensate the phase noise.

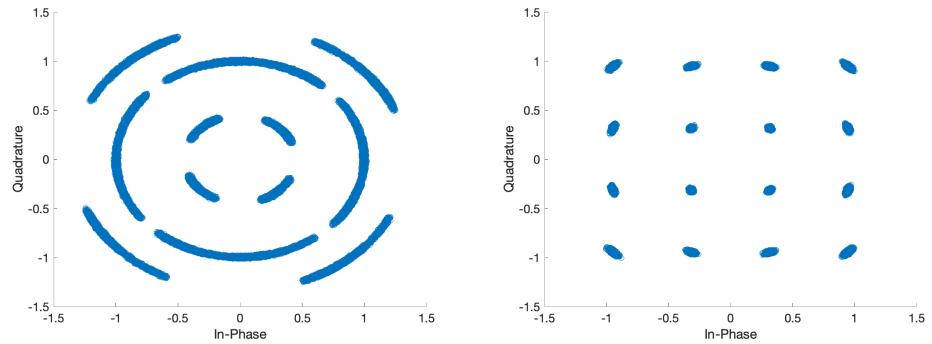
3.3.4 Results

Bringing all these blocks together the simulation results can be observed in the figures below:



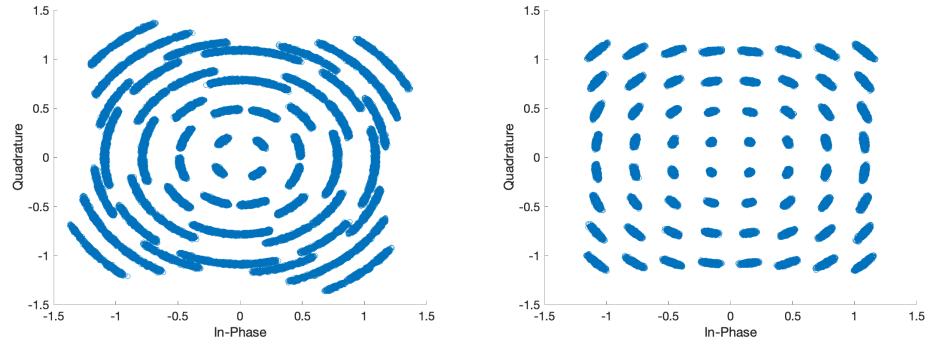
(a) Effect of Laser Noise on QPSK (b) QPSK constellations after passing constellation through Phase Recovery

Figure 23: QPSK Phase Recovery



(a) Effect of Laser Noise on 16-QAM constellation (b) 16-QAM constellations after passing through Phase Recovery

Figure 24: 16-QAM Phase Recovery



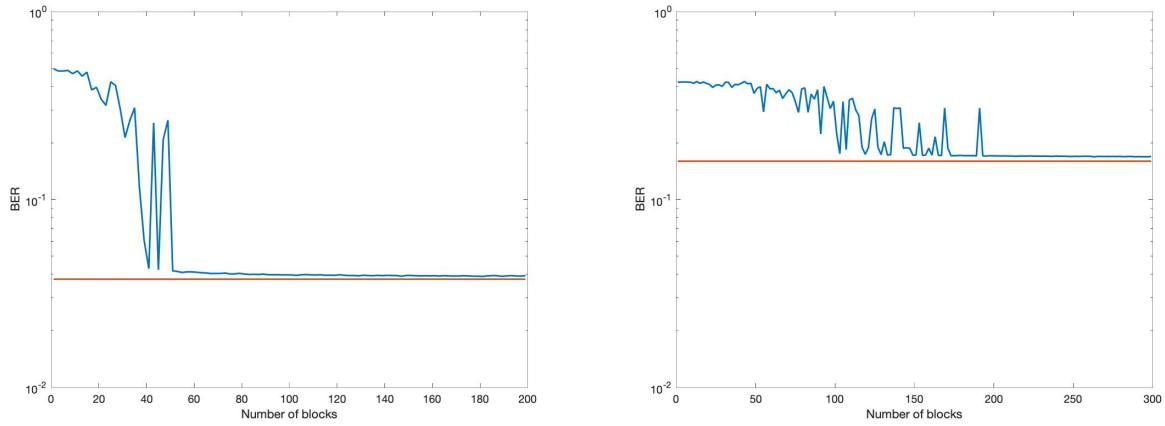
(a) Effect of Laser Noise on 64-QAM constellation (b) 64-QAM constellations after passing through Phase Recovery

Figure 25: 64-QAM Phase Recovery

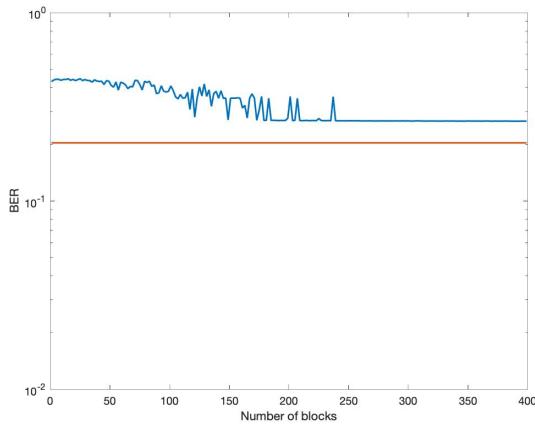
As noticed in the figures above the Phase Recovery algorithms successfully manages to compensate for the phase noise introduced by the lasers.

3.3.5 Optimisation

In order to maximise the efficiency of the Phase Recovery, an optimum value for the block length N needs to be selected. For that, a target BER has been selected and the measured BER has been found as N is varied. For each of the constellation the results are displayed below:



(a) QPSK Phase Compensation Optimisation (b) 16-QAM Phase Compensation Optimisation



(c) 64-QAM Phase Compensation Optimisation

Figure 26: Block Length Optimisation

Using the figures above a suitable value for N for each constellation has been chosen. It is observed that as constellation become more and more complex there is a certain gap in the BER that cannot be reduced. This is due to the fact that for both 16-QAM and 64-QAM only symbols which have a certain absolute value have been selected. To reduce this gap a more complex algorithm for the phase recovery is needed to be implemented.

3.3.6 Problems encountered

Some problems encountered in this section are due to the fact that the symbols in *Figures 23b, 24b and 25b* could be shifted by either $\pi/2$, π and $3\pi/2$. Therefore checks have to be made in each rotation to be able to determine which shift has occurred.

3.4 Non-Linearity

There is one last block which is needed to be added in the channel in order for the simulation to resemble a real-life transmission. This represents the non-linearity in the fibre. The block diagram of the system so far is represented below:

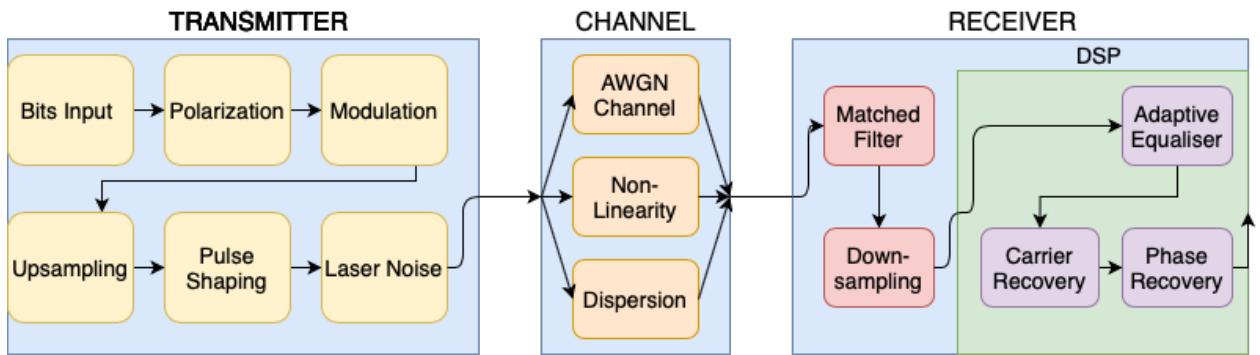


Figure 27: Simulation Block Diagram including the Non-Linearity in the fibre

To be able to introduce the Non-Linearity to a reasonable accuracy that matches a real-life transmission the SSFT method will be used.

3.4.1 SSFT Algorithm

So far the Channel has been modelled as an AWGN channel but with dispersion added by an FIR filter. However the true propagation of pulse through an optical fibre is described by solving the *Nonlinear Schrödinger Equation(NLSE)*:

$$\frac{\partial A(t, z)}{\partial z} = -\frac{\alpha}{2}A(t, z) + j\frac{\beta_2}{2}\frac{\partial^2 A(t, z)}{\partial t^2} - j\gamma|A(t, z)|^2A(t, z) \quad (26)$$

where α is the attenuation, β_2 the dispersion coefficient and γ is the nonlinear coefficient related to the Kerr coefficient n_2 by:

$$\gamma = \frac{n_2 w_0}{c A_{\text{eff}}} \quad (27)$$

where A_{eff} is the effective area of the fibre mode. The Kerr coefficient describes how strong the non-linear effect due to the Kerr effect is. Kerr effect is defined as the phase shift an optical signal undergoes which is proportional to the signal power. This phenomenon is called *Self-Phase Modulation(SPM)*. All these parameters were given values at the beginning of Section 3.

The NLSE equation in 26 is strictly true for a single polarisation signal. For a dual polarisation which, in practice will always be used, the NLSE needs to be modified as the pulse is now represented by a vector $\mathbf{A} = [A_x(t, z), A_y(t, z)]^T$:

$$\frac{\partial \mathbf{A}}{\partial z} = -\frac{\alpha}{2}\mathbf{A} + j\frac{\beta_2}{2}\frac{\partial^2 \mathbf{A}}{\partial t^2} - j\frac{8}{9}\gamma(\mathbf{A}^H \mathbf{A})\mathbf{A} \quad (28)$$

where \mathbf{A}^H represents the Hermitian transpose operator. This equation is known as the *Makanov* equation.

Solving either the NLSE equation in 26 or the Makanov equation in 28 would result in a function that represents how the pulse propagates through the optical fibre. However, both of these equations cannot be solved analytically for some arbitrary inputs. Therefore the SSFT method, which is numerical technique, will be used to solve these equations. For simplicity we will focus on the NLSE equation but this technique can be applied to Makanov equation as well.

The NLSE equation in 26 can be rewritten as:

$$\begin{aligned} \frac{\partial A(t, z)}{\partial z} &= (\hat{D} + \hat{N})A \\ \text{where } \hat{D} &= -\frac{\alpha}{2} + j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2} \\ \hat{N} &= -j\gamma|A(t, z)|^2 \end{aligned} \quad (29)$$

so in moving frame of reference $A(t, z + h)$ can be expressed as:

$$\begin{aligned} A(t, z + h) &= e^{(\hat{D} + \hat{N})h}A(t, z) \\ &= e^{\hat{D}h}e^{\hat{N}h}A(t, z) \\ &\approx F^{-1} \left\{ \underbrace{e^{-(j\beta_2 w^2 + \alpha)h/2}}_{\text{Linear Term}} F \left\{ \underbrace{e^{-jh\gamma|A(t, z)|^2}}_{\text{Nonlinear Term}} A(t, z) \right\} \right\} \end{aligned} \quad (30)$$

where F and F^{-1} represents the Fourier transform and its inverse respectively.

Therefore for a total transmission distance of $L = 1000\text{km}$, an amplifier spacing of 100km and for an initial signal represented by vector \mathbf{S} , the SSFT algorithm to describe the pulse propagation can be written as:

Algorithm 3 Split-Step Fourier Transform(SSFT) Algorithm

$l \leftarrow 1$

```

for  $h \leftarrow 1 : l : L$  do
    Linear Term  $\leftarrow e^{-(j\beta_2 w^2 + \alpha)h/2}$ 
    Nonlinear Term  $\leftarrow e^{-jh\gamma|\mathbf{S}|^2}$ 
     $\mathbf{S} \leftarrow \text{ifft}(\text{Linear Term} \times \text{fft}(\text{Nonlinear Term} \times \mathbf{S}))$ 
    if remainder( $l, 100$ ) = 0 then
         $\mathbf{S} \leftarrow \mathbf{S} \times 100$  Amplify by 20dB
    end if
end for

```

3.4.2 Results

Using the algorithm described above some initial tests have been made. These can be observed in the figure below. In *Figure 28a* β_2 and γ are both set 0 so that only the attenuation and fibre amplification have been able to be observed. In *Figure 28b* α and γ are set to 0 and only a small dispersion is applied. In *Figures 28c and 28d* all the parameters are set to values defined at the beginning of Section 3.

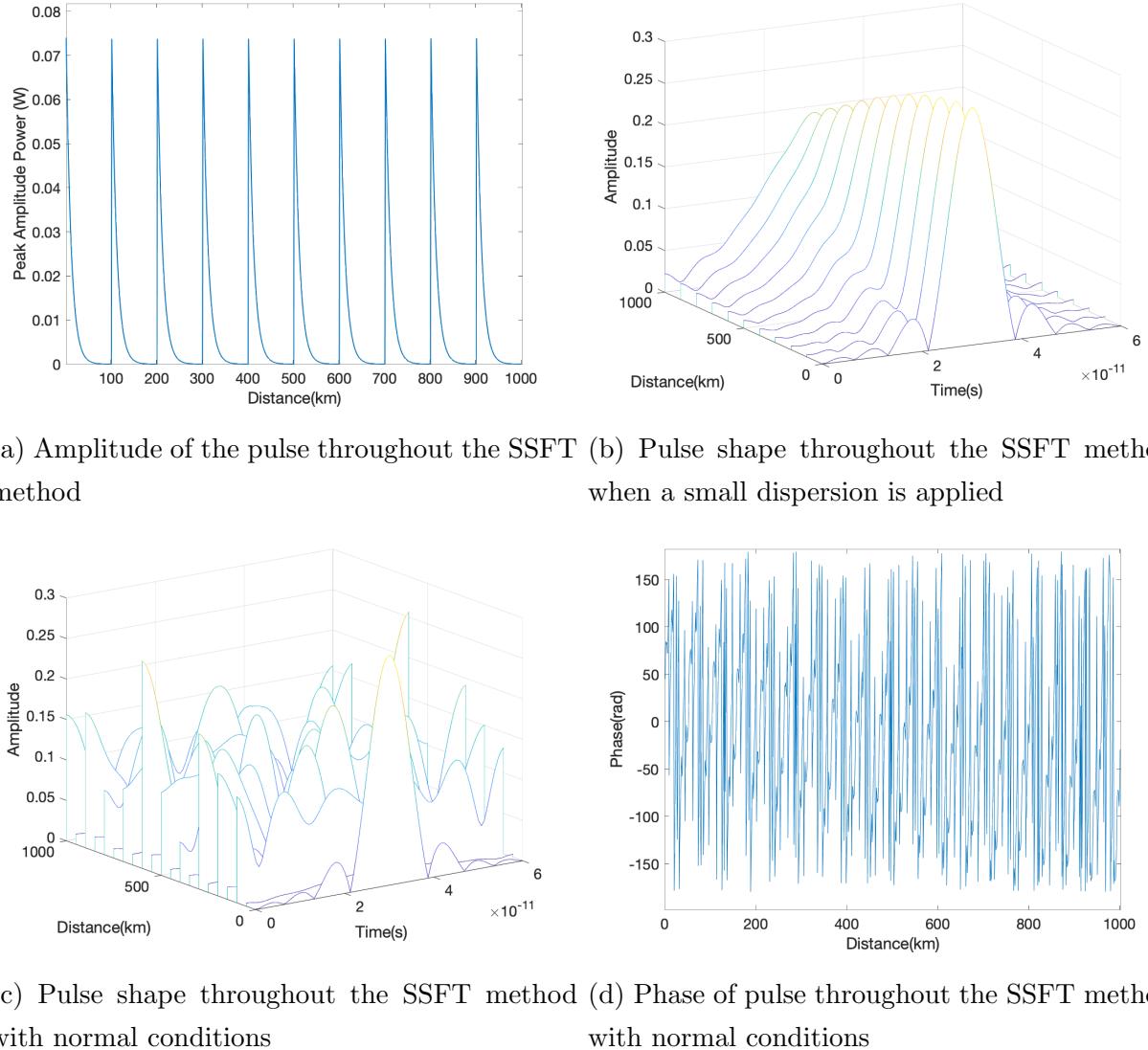


Figure 28: Test results of using SSFT method

In *Figure 28a* it is observed that the amplitude variation behaves as expected of a pulse passing through an optical fibre with some attenuation. In *Figure 28b* is observed that indeed by introducing a small dispersion, the pulse will broaden which is the expected behaviour as the blue components of the pulse will travel faster than the red components of the pulse. Therefore, it can be said that the SSFT method correctly reproduces the expected behaviour

of a signal in optical fibre. In *Figures 28c and 28d* the characteristics of single pulse as it passed through the channel simulation with parameters defined in the beginning is visualised. The simulated channel parameters will be used throughout the rest of the project.

3.4.3 Problems Encountered

The only problem that occurred during the build of this algorithm was the fact that l was chosen too big. This meant the approximation that we made in Equation 30 cannot reproduce accurate results. Therefore it was important to choose a small value of l relative to the total distance travelled.

3.5 Static Equaliser

The last block diagram to be introduced in order to complete our system such that it would resemble the one described in *Figure 9* is the Static Equaliser. The reason why Static Equaliser is the last block introduced is because it was required to have the SSFT method to introduce the dispersion correctly. The role of the static equaliser is to compensate for the bulk dispersion in the fibre.

3.5.1 Static Equaliser Algorithm

It was observed in the previous section that the dispersion was introduced in the SSFT method by the linear term:

$$\exp(-j\beta_2 w^2 L/2) \quad (31)$$

Therefore in order to compensate for the bulk of this dispersion, an FIR filter \mathbf{h} is used of length N . The tap weights will have to equal to $\exp(j\beta_2 w^2 L/2)$ in the frequency domain in order to fully compensate for the dispersion. As a result the tap weights in time domain will be equal to:

$$h[k] = T_{\text{samp}} \sqrt{\frac{1}{2\pi j\beta_2 L}} \exp\left(\frac{jk^2 T_{\text{samp}}^2}{2\beta_2 L}\right), \quad \text{where } -M \leq k \leq M \quad (32)$$

In order to implement this FIR filter digitally and to avoid the finite sampling frequency aliasing it is required to choose the value of M very carefully. We note that if we sample every T_{samp} seconds then aliasing will occur for frequencies which exceed the Nyquist frequency given by $w_k = \frac{\pi}{T_{\text{samp}}}$ and that the impulse response may be considered as a rotating vector whose angular frequency is given by:

$$w = \frac{t}{\beta_2 L} \quad (33)$$

Which implies that the value of M should be:

$$M = \left\lfloor \frac{\pi |\beta_2| L}{T_{\text{samp}}^2} \right\rfloor \quad (34)$$

From the equation above, the total number of taps is $N = 2M + 1$. Because the number of taps will be very large the filter implementation will be more efficient in the frequency

domain. Therefore the algorithm used to implement this filter is named *Overlap and Save* whose full description is explained in the Appendix.

3.5.2 Results

The Static Equaliser is now introduced in the simulation system. Signal with different modulations, QPSK, 16-QAM and 64-QAM, will be sent via the SSFT method. The results right before the Equalisation stage and after the Equalisation stage which include Static Equaliser and Adaptive Equaliser can be observed below.

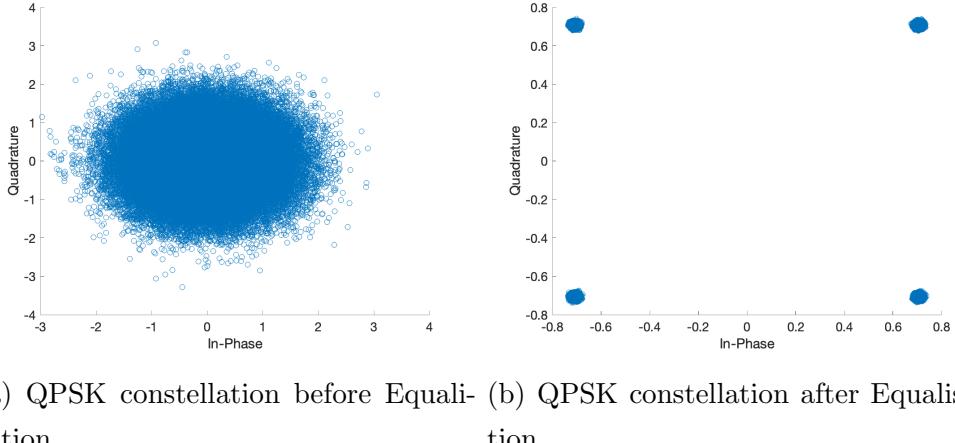


Figure 29: QPSK Equalisation

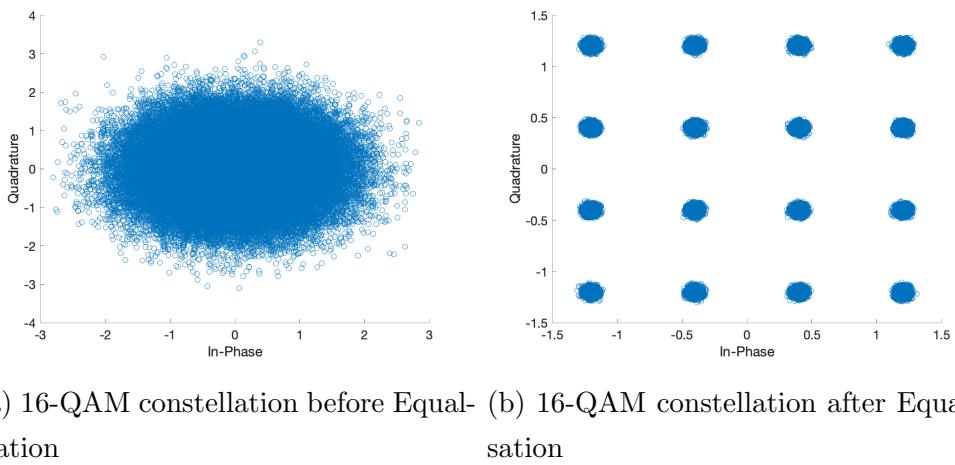
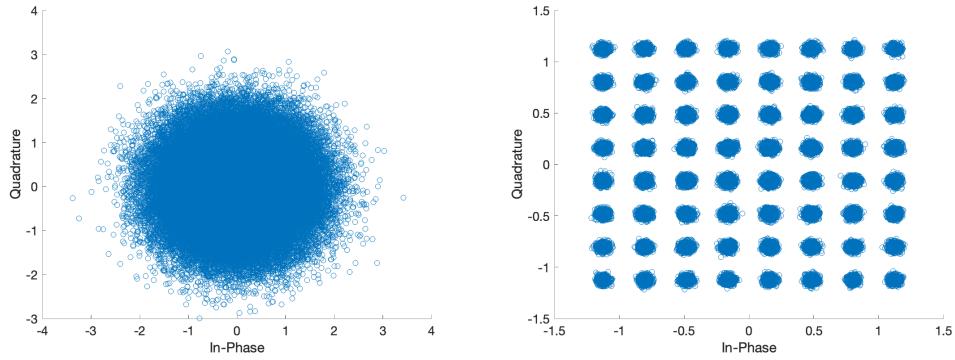


Figure 30: 16-QAM Equalisation



(a) 64-QAM constellation before Equalisation (b) 64-QAM constellation after Equalisation

Figure 31: 64-QAM Equalisation

From the figures above it is observed that both the Static and Adaptive Equaliser successfully manages to remove the dispersion in the channel. It can be observed that there are clear decision regions from which the inference of the right symbol can be made.

3.5.3 Problems Encountered

The main problem which has been causing the main algorithm to fail was the fact that the delay introduced by the FIR filter has not been adjusted correctly. In the main algorithm for static equaliser it is required to remove the end $M + 1$ symbols in order to adjust for the zero padding of the FIR filter. This meant that in the original signal transmitted it is required to remove the $M/2 + 1$ symbols from the beginning and $M/2$ from the end of the symbols sequence.

3.6 Final Results and Conclusions

As by now the simulation system composed of the all the block diagrams in *Figure 9* have been implemented it is now time to see how the full simulation performs. A good metric is to look at the BER curve for this system and compare it to the theoretical AWGN channel BER curve. The result is displayed below for different modulations.

It is observed that there is a difference compared to the AWGN channel simulations that was presented in *Figure 13*. The gap between the theoretical curve and the simulation curve is due to two things. The first one is the fact that it is required to implement more elaborated algorithms than the Viterbi-Viterbi algorithm used in order to fully compensate for the laser noise in more complex constellations . The second contribution to this gap is due to the existence of non-linearity which was introduced in the section above. There are some suitable solutions to addressing this nonlinearity problem. One solution would be implementing a Digital Back Propagation technique which involves passing the received symbols into a virtual fibre with the nonlinear parameter equal to $-\gamma$ such that nonlinearity will be compensated. In this simulation we have not introduced two types of other nonlinear effects

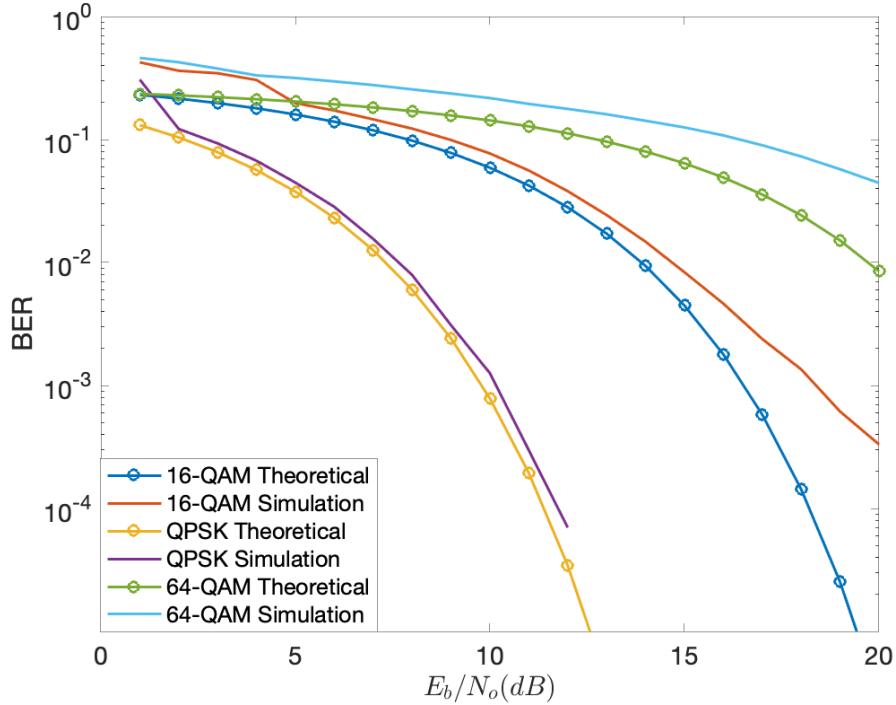


Figure 32: BER plot of Theoretical vs Simulation Results for the overall system

in the channel. These are *Intrachannel Cross-Phase Modulation(IXPM)* and *Intrachannel Four-Wave Mixing(IFWM)*. IXPM causes a phase modulation of a pulse by another in the same channel. The phase modulation caused by IXPM leads to instantaneous frequency change of a pulse. So because in a dispersive fiber different frequency components travel at different speeds so frequency change result in group speed changes. So the first pulse moves faster than the second pulse and it arrives at the earlier stage than the second one leading to a temporal separation longer than the bit interval. IFWM leads to ghost pulses being generated when two or more pulses of the same channel interact with each other. To compensate for these nonlinearities complex special equaliser needs to be implemented at the DSP. Some of the widely used equalisers are: Voltera Equaliser[18], Turbo Equaliser [19] and Kalman Filter [20].However due to time constrains, it was not possible to fully investigate these methods of combating the nonlinearity.

In the next section this basic simulation model will be used to study Probabilistic Constellation Shaping (PCS) technique to improve the capacity of this channel. Therefore from now only 64-QAM modulation simulation will be used.

4 Probabilistic Constellation Shaping(PCS)

A real-life long haul optical channel can be well approximated by an AWGN channel. As with any channel, there is a fundamental limit on how much information you can send reliably. In order to fully understand how the PCS technique will improve the capacity of the channel, some information on the capacity of the AWGN channel needs to be given. The capacity of any given channel represents the highest rate in bits per channel use at which information can be sent with arbitrarily low probability of error. Mathematically, it can be expressed as the maximum of the mutual information between the input and output of the channel, where the maximisation is with respect to the input distribution[21]:

$$C = \max_{p_X(x)} I(X, Y) \quad (35)$$

where $I(X, Y)$ is the mutual information(MI) between X and Y defined as:

$$I(X, Y) = H(Y) - H(Y/X) = H(X) - H(X/Y) \quad (36)$$

where $H(X)$ is the entropy of an ensemble that represents the average Shannon information. Its mathematical expression is:

$$\begin{aligned} H(X) &= - \sum_{x \in X} p(x) \log_2 p(x) \quad \text{for discrete } X \\ H(X) &= - \int_{-\infty}^{\infty} p(x) \log_2 p(x) dx \quad \text{for continuous } X \end{aligned} \quad (37)$$

We start by first denoting X and Y as the complex-valued input and output respectively, into the channel. Therefore assuming an AWGN channel, it be said that:

$$Y = X + N \quad (38)$$

where N is a circularly-symmetric Gaussian distribution with mean 0 and variance σ_n^2 . If it is assumed the X has a Gaussian input distribution with variance σ_x^2 , $p(y)$ and $p(y/x)$ can be expressed as:

$$\begin{aligned} p(y) &= \frac{1}{\pi(\sigma_x^2 + \sigma_n^2)} \exp\left(-\frac{|y-x|^2}{\sigma_x^2 + \sigma_n^2}\right) \\ p(y/x) &= p(x + n/x) = p(n) = \frac{1}{\pi\sigma_n^2} \exp\left(-\frac{|y|^2}{\sigma_n^2}\right) \end{aligned} \quad (39)$$

Therefore if the input signal power is $S = \sigma_x^2$ and $N_0 = \sigma_n^2$ is the noise power, and we use the values above in Equation 37 for the continuous time, the results will be:

$$\begin{aligned} H(Y) &= \frac{1}{2} \log_2 2\pi e(S + N) \\ H(Y/X) &= \frac{1}{2} \log_2 2\pi eN \end{aligned} \quad (40)$$

which will then result in mutual information being equal to:

$$I(X, Y) = \frac{1}{2} \log_2 \left(1 + \underbrace{\frac{S}{N}}_{SNR} \right) \quad (41)$$

The true capacity of the channel will be found if the mutual information will be multiplied by the bandwidth. As showed, a lower bound to the capacity of an AWGN channel has been found. However, the Gaussian constellation has an infinite number of points, which makes it impossible to implement in a communication system. Real-world constellations have a finite number of points, such as M-QAM. Computing the mutual information for these modulations and the simulation system described in Section 3 would require to compute the integral using Monte Carlo methods. As a result, the mean μ and variance σ^2 values of the symbols at the output of the DSP have been measured for each unique symbol in the transmission. Then the mutual information is found expressing the probability of decoded symbol belonging to a certain known unique symbol as a Gaussian. Therefore the mutual information can be expressed as:

$$I(X, Y) = \frac{1}{M} \sum_{i=1}^M \sum_{k=1}^N \exp \left(-\frac{|y[k] - \mu[i]|^2}{\sigma^2[i]} \right) p(x = i) \quad (42)$$

where M is the modulation number and N the total number of symbols in the transmission. The core of the algorithm was taken from [22] and then modified and adapted to meet the system described in Section 3. The results of mutual information for these constellations can be observed below:

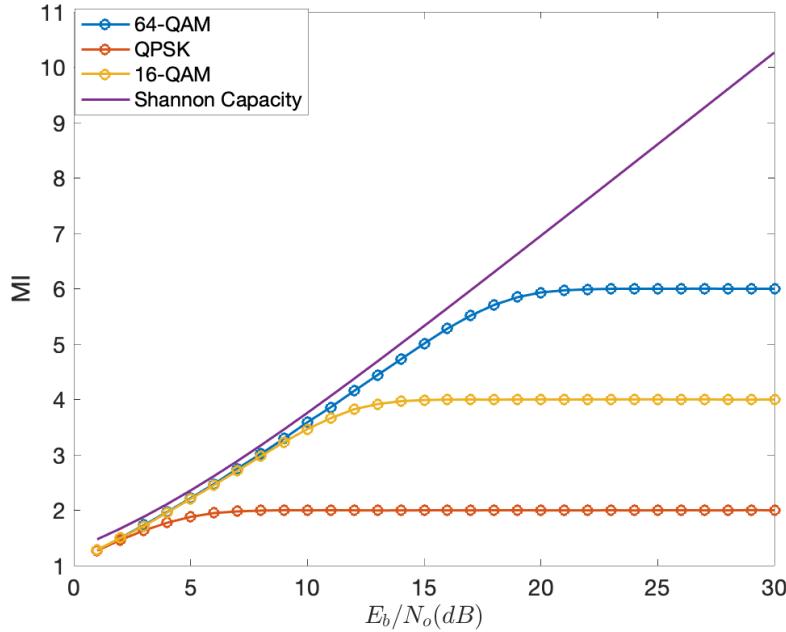


Figure 33: Mutual information of QAM constellations over an AWGN channel

The MIs that can be achieved by QAM constellations are shown in the figure above, where they are compared with the Gaussian constellation. As seen in the figure, at low SNR

the performance of QAM constellations is close to the Gaussian constellation. Then, each constellation saturates at its entropy $H(X_{\text{QPSK}}) = 2$, $H(X_{16\text{-QAM}}) = 4$ and $H(X_{64\text{-QAM}}) = 6$. At medium values of SNR it is observed that there is a fixed gap between QAM constellations and channel capacity. This gap, called shaping gap, is asymptotically equal to $e/6 \approx 1.53$ dB and is related to the sub-optimality of the 2D square shape [23].

Therefore, some modifications of QAM constellation are necessary to fill this gap. In general, there are two different shaping methods that can be applied to a QAM-based communication system in order to reduce this shaping gap. These are:

- Geometric Constellation Shaping(GCS)
- Probabilistic Constellation Shaping(PCS)

Geometric shaping techniques optimise the position of constellation points to maximise a metric (e.g. the MI), while probabilistic shaping changes the probability of transmission of standard QAM constellations[23].

This project focuses on PCS. The technique will be implemented using scheme called Probabilistic Amplitude Shaping(PAS) which takes advantage of the fact that one QAM complex symbol can be mapped by two real Amplitude Shift Keying(ASK) symbols so that uniformly distributed data symbols are transformed into non-uniformly distributed symbols.

4.1 PAS Architecture

As mentioned above the role of the PAS architecture is to change uniform distributed symbols to a pre-selected non-uniform distribution. A diagram explaining the PAS architecture scheme is observed in the figure below:

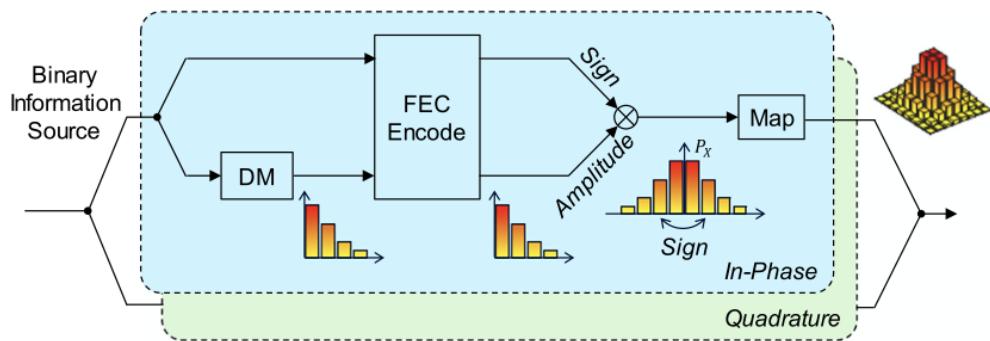


Figure 34: PAS architecture

It is important to notice that the Forward Error Correction Encoding can be applied before or after the PAS scheme. In our project we chose to not use a FEC encoder and decoder as it was desired to look at just the pure signal without adding any redundancy in the signal.

PCS is achieved by this scheme by independently shaping each signal dimension on an M-ary pulse amplitude modulation (PAM) template to construct a probabilistically-shaped (PS) QAM constellation. This is possible since the in-phase and quadrature dimensions of a modulated signal are orthogonal[13].

The key component of the PAS architecture is the Distribution Matcher (DM), which is a block that converts an equiprobable stream of bits into a multi-level sequence with unequal probability. Finding an efficient hardware-implementable DM is a current topic of active research[23].

In order for the DM to perform, we must know what kind of distribution our symbols should have at the output the PAS architecture. We must remember that our goal is to improve the capacity of the channel and therefore the desired distribution should be chosen so that the mutual information is maximised. This probability distribution is called Maxwell-Boltzmann(MB) distribution and it's mathematically expression is:

$$P_X(x) = \frac{\exp(-\lambda x^2)}{\sum_{x' \in \mathcal{X}} \exp(-\lambda x'^2)} \quad (43)$$

where λ is the rate parameters that controls the entropy rate which in order to minimise the shaping gain λ is required to be optimised for each SNR.

When $\lambda = 0$, the MB distribution creates a uniform distribution with the entropy equalling to $H(X) = \log_2(M)$ where M is the modulation number. As $\lambda \rightarrow \infty$, $P_X(x)$ distribution approaches the distribution that chooses the smallest amplitude with probability 1. The resulting entropy is zero.[13]. Therefore λ must be chosen carefully so that the MI will be maximised.

The DM used in the PAS architecture is a special version called Constant-Composition Distribution Matcher (CCDM) which uses modified arithmetic coding that involves multiplications, divisions, and comparisons of real numbers in order to transform the uniform distribution into the MB distribution described above.

4.2 Distribution Matcher

In this section we are going to outline the theory behind the Distribution Matcher and we are especially going to describe Constant-Composition Distribution Matcher(CCDM). The theory and implementation of CCDM has been taken from [24] where there are also more details on the performance of the algorithm.

According to the PAS architecture described above, we need a device that takes as input k_c uniformly distributed independent bits U_1, U_2, \dots, U_k and outputs n_c amplitudes $\tilde{A}_1, \tilde{A}_2, \dots, \tilde{A}_{n_c}$. The device should also have the following properties:

1. The input provides a binary interface to the source coding part of the digital communication system.
2. The input can be recovered from the output, i.e., the mapping is invertible.
3. The output is approximately the output of a *Discrete Memoryless Source*(DMS) \mathbf{P}_A that outputs amplitudes A_1, A_2, \dots, A_k that are iid according to P_A . The modulation described in Section 3 was a DMS with probability $\frac{1}{M}$ where M is the modulation number.

We will start by considering an amplitude sequence a_1, a_2, \dots, a_{n_c} . Let n_α be the number of times that amplitude α occurs in this sequence, i.e, $n_\alpha = |\{i : a_i = \alpha\}|$. Let $P_{\tilde{A}}$ be the empirical distribution of this sequence such that $P_{\tilde{A}}$ will be equal to:

$$P_{\tilde{A}}(\alpha) = \frac{n_\alpha}{n_c}, \quad \alpha \in \mathcal{A} \quad (44)$$

Since each probability $P_{\tilde{A}}(\alpha)$ is an integer multiple of $\frac{1}{n_c}$ the distribution $P_{\tilde{A}}$ is called an n_c -type. The set $\mathcal{T}_{P_{\tilde{A}}}^{n_c}$ contains all the permutations of the sequence a_1, a_2, \dots, a_{n_c} and is called the type class of $P_{\tilde{A}}$. A set of sequences is called a constant composition code, if all the sequences are of the same type.

In order to understand how a CCDM would work lets consider a simple example of 4-ASK constellation $\mathcal{X} = \{\pm 1, \pm 3\}$ and an output amplitude sequence of length $n_c = 4$. The amplitudes are $\mathcal{A} = \{1, 3\}$ and suppose that the desired distribution is:

$$P_A(1) = \frac{3}{4}, \quad P_A(3) = \frac{1}{4} \quad (45)$$

The probabilities are integer multiples of $\frac{1}{4}$, so P_A is a 4- type. The corresponding type class of length will be:

$$\mathcal{T}_{P_A}^4 = \{(1, 1, 1, 3), (1, 1, 3, 1), (1, 3, 1, 1), (3, 1, 1, 1)\} \quad (46)$$

It is observed that in each sequence, amplitudes 1 and 3 occur $n_1 = 3$ and $n_3 = 1$ times, respectively. A CCDM will map a length k_c binary string to sequences in $\mathcal{T}_{P_A}^4$. As $n_c = 4$ we will have:

$$k_c = \log_2 |\mathcal{T}_{P_A}^4| = 2 \quad (47)$$

Therefore the following lookup tables will define the CCDM:

$$\begin{aligned} 00 &\mapsto (1, 1, 1, 3), & 01 &\mapsto (1, 1, 3, 1) \\ 10 &\mapsto (1, 3, 1, 1), & 11 &\mapsto (3, 1, 1, 1) \end{aligned} \quad (48)$$

This mapping maps binary strings to amplitude sequences and thereby provides a binary interface (Property 1). The mapping is one-to-one and therefore invertible (Property 2).

Now it can be seen that instead of the probability distribution described in the example above we would have the MB distribution and instead of the ASK constellation we would have 64-QAM constellation we would be able to use this CCDM technique to create the desired MB distribution over QAM.

Therefore a general construction guidance of the CCDM algorithm can be made as follows:

- knowing what the desired distribution $P_A(\alpha)$ is and the number of symbols n_c we are expecting to transmit which we have defined in at the beginning of Section 3. As a result we can find n_α as:

$$n_\alpha \approx \lfloor n_c \times P_A(\alpha) \rfloor, \quad \alpha \in \mathcal{A}$$

- next it is required to calculate the length k_c needed for the input bits which can be find as:

$$k_c = \lfloor \log_2 |\mathcal{T}_{P(A)}^{n_c}| \rfloor = \left\lfloor \log_2 \frac{n_c}{n_{0.2323}! n_{0.5195}! \dots n_{1.6264}!} \right\rfloor$$

where $n_\alpha = P_A(\alpha) \times n_c, \alpha \in \mathcal{A} = \{0.2323, 0.5195, \dots, 1.6264\}$ which represent the amplitudes of the 64-QAM constellation. The rounding is necessary since the cardinality of the type class $\mathcal{T}_{P(A)}^{n_c}$ may not be a power of 2.

- in the last step we pick a set $\mathcal{C}_{ccdm} \subseteq \mathcal{T}_{P(A)}^{n_c}$ with $|\mathcal{C}_{ccdm}| = 2^{k_c}$ sequences and define a bijective mapping f_{ccdm} from $\{0, 1\}^{k_c}$ to \mathcal{C}_{ccdm} .

It is observed that if n_c is really big which in our case it is $n_c = 10^5$ symbols, the size of the code book \mathcal{C}_{ccdm} grows exponentially so choosing a single value out this lookup table becomes an infeasible operation. Therefore a new method that performs the mapping of f_{ccdm} without storing the whole code book \mathcal{C}_{ccdm} is required.

This new method will consist of using arithmetic coding for indexing sequences efficiently and it is taken from [25]. Our arithmetic encoder associates an interval to each input sequence in $\{0, 1\}_c^k$ and it associates an interval to each output sequence in $\mathcal{T}_{P(A)}^{n_c}$, as can be seen in *Figure 35* below. The size of an interval is equal to the probability of the corresponding sequence according to the input and output model, respectively. For the input model we have described the model in Section 3 as an iid Bernoulli(0.5).

In the figure below it shows input and output intervals with output length $n_c = 4$ and $P(1) = \frac{3}{4}$ and $P(3) = \frac{1}{4}$ as in the example of the ASK constellation above. The intervals on both the input and the output side are $[0, 0.25)$, $[0.25, 0.5)$, $[0.5, 0.75)$ and $[0.75, 1)$. The arithmetic encoder can link an output sequence to an input sequence if the lower border of the output interval is inside the input interval.

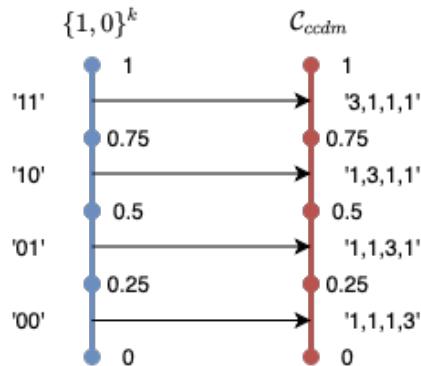


Figure 35: Arithmetic Encoder with $k_c = 2, n_c = 4$ and Amplitudes $\mathcal{A} = \{1, 3\}$ with $P(1) = \frac{3}{4}$ and $P(3) = \frac{1}{4}$

However in a real life scenario we would not know the whole bit sequence and therefore it is required to implement a sequential version of the arithmetic encoder which will have the following structure. In *Figure 36* below this sequential process is applied to an input sequence of bits '10':

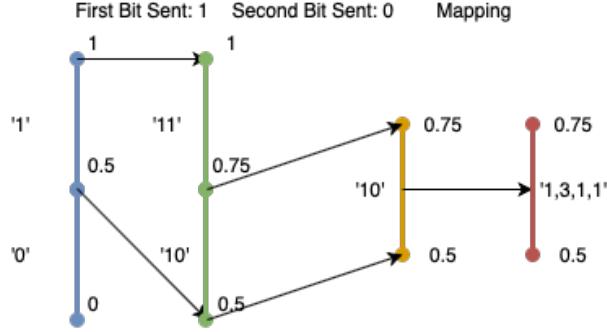


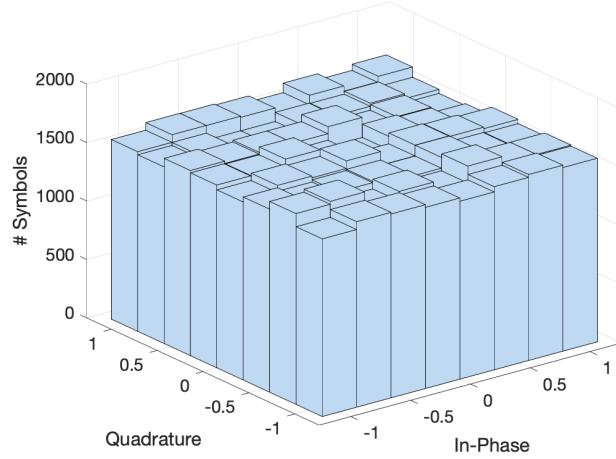
Figure 36: Sequential Arithmetic Encoder with $k_c = 2, n_c = 4$ and Amplitudes $\mathcal{A} = \{1, 3\}$ with $P(1) = \frac{3}{4}$ and $P(3) = \frac{1}{4}$

- As the input model is Bernoulli(0.5) we split the interval into two equally sized intervals and continue with the upper interval in case the first input bit is '1'; otherwise we continue with the lower interval.
- After the next input bit arrives, we repeat the last step and so after k_c input bits we will have an 2^{k_c} interval
- After every refinement of the input interval, the algorithm checks for a sure sequence from the total output sequence. In our case, if we have sent bit '1' first we know that the output intervals can only be either (1, 3, 1, 1) or (3, 1, 1, 1) each one having a certain probability. Every time we extend the sure prefix by a new symbol, we must calculate the probability of the next symbol given the sure prefix. That means we determine the output intervals within the sure interval of the prefix.

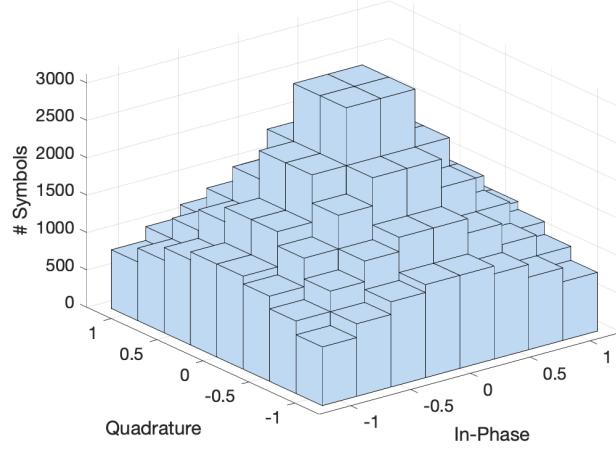
Having now described the algorithm for the DM we will look at the results of introducing this PAS architecture with the DM into the simulation described in Section 3. The core of the DM algorithm has been taken from [26] and been modified and adapted to fit the simulation in Section 3.

4.3 Results and Conclusion

As mentioned, the PAS architecture including the DM have now been introduced into the simulation. First we will be looking at the performance of the DM and observe if indeed a non-uniform distribution over the symbols has been introduced. The results can be observed below:



(a) Uniform Distribution



(b) Non-uniform Distribution

Figure 37: 64-QAM symbols distribution with and without the PCS technique

In *Figure 37b* above it can be seen that lower energy symbols have a higher probability of transmission than higher energy symbols. As a result it can be said that the distribution formed by the symbols will correspond to a Maxwell-Boltzmann Distribution. Therefore it can be said that the PAS architecture has correctly introduced a non-uniform distribution in the 64-QAM symbols.

Since now the desired distribution has been introduced we can now compute the new mutual information of the system using this distribution and observe whether or not it was possible to reduce the shaping gap. The results can be observed below for the 64-QAM modulation:

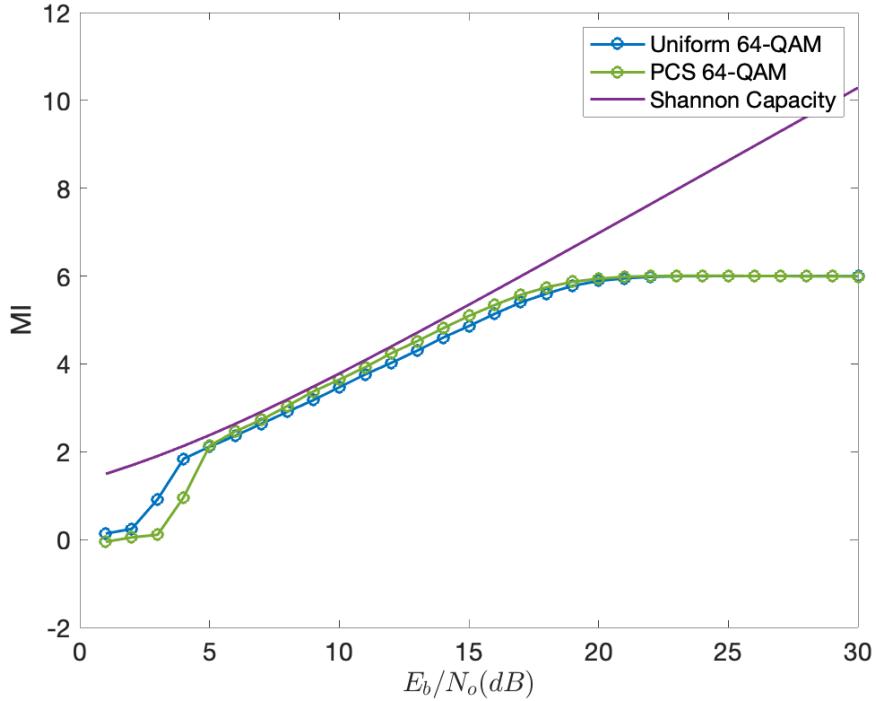


Figure 38: Mutual information of QAM constellations over an AWGN channel

It can be observed that from this figure that the PCS technique has successfully managed to reduce the shaping gap between the uniform distribution 64-QAM and the Shannon Capacity. Therefore it is observed that the MB distribution over the symbols is optimal. However at low values of SNR, it is observed that the PCS technique performs worse than the uniform distribution. This is because the noise added is so big that the amplitudes can be very big, which will increase the non-linearity in the signal. From the figure above it can be observed that even though it has been reduced there is still a gap between the shannon capacity and the non-uniform 64-QAM capacity. This due to the fact that our channel model is not an ideal AWGN channel as it contains some nonlinearity due to the SSFT method. Further reduction in the shaping gap can be achieved by implementing nonlinear equalisers which aim to compensate for these nonlinearities. Since in our model we only introduced the self phase modulation nonlinearity, we would observe that if we were to introduce the rest of nonlinearities in the fibre we would expect that this shaping gap to become bigger, partly due to the fact that uniform 64-QAM capacity will be worse. But we will observe that the improvement in capacity of the non-uniform 64-QAM compared to the uniform 64-QAM will remain the same. Because of this, it can be said that indeed it was possible to improve the rate of transmission of a channel. If we chose an SNR of 14 it will be observed that the capacity of the channel increases by 4.8% and as result we can send 4.8% more data on the channel reliably.

5 Final Report Conclusions

This project has proved how useful the PCS technique is, in improving the line rate of an optical fibre system even by a small margin. Using the optimum power launch and the symbol rate the data throughput can be calculated to be 0.939 Tbit/s while without the PCS the value of data throughput is 0.9038 Tbit/s. This increase in the capacity was demonstrated through a phased approach of building first a simulation of a real long-haul optical fibre transmission and then using the PAS architecture to enable the PCS. If further improvements can be made it to compensate for the nonlinearity and more complex DSP algorithms should be used, a data throughput of 1.0 Tbit/s can be achieved.

The challenges of building the simulation were numerous as it was desired to make as accurate as possible and as a result dispersion, nonlinear noise and phase noise had to be introduced systematically. By far the biggest challenge was building the DSP to correct for these effects. Many lessons were learned about how to account for different delays introduced by filters or how to ensure that the right form of sampling is used. Switching from different digital modulation such as, from QPSK to 16-QAM, has meant that ,most of times, all the algorithms would stop working as it would require either further additional components or to even to change the algorithm completely. However having used QPSK modulation first meant that the initial algorithms were easier to understand and it paved the way for the underlying principle of the more complex algorithms. This project has allowed me to understand the mechanism behind a DSP, to comprehend the different types of noises in the signal that engineers encounter and how difficult it is to compensate for these.

There are two types of future research which can be developed from this project. The first one should aim to improve the simulation by introducing a WDM system and the extra nonlinear noise sources and look into different nonlinear techniques that would help to bring the capacity of the channel even closer to the Shannon Limit. The second is looking into different hardware implementations of the PCS technique which should aim at making it as efficient as possible and with minimum cost.

In the end it can be said that the goal of the project has been accomplished. This report showed that the PCS technique has managed to increase the capacity of the channel. As the world will continue to develop the need for higher data rates will continue to increase and therefore any form of improvements is welcomed and it will help to sustain this growth and improve the lives of billions of people.

A Overlap and Save method

A general FIR filter in order to generate an output $y[n]$ by filtering an input vector $\mathbf{x}^T = (x[n], x[n-1], \dots, x[n-(N-1)])^T$ of N samples by the filter \mathbf{h} can be written as:

$$y[n] = \mathbf{h}^T \mathbf{x} = \sum_{k=0}^{N-1} h[k]x[n-k] \quad (49)$$

requiring $O(N)$ multiplications to find n . Since in our solution N is very big it will be more efficient to implement this method in the frequency domain using Overlap and Save method described below:

1. Append N zeros to the tap weights get an array of length $2N$ which we then transform into the frequency domain (using a $2N$ point FFT). This transformation is done just once as such it can be neglected insofar as the computational cost (since fundamentally this is just an alternative way of representing the tap weights).
2. Take the block of data of length $2N$, e.g. \mathbf{x}_0 . The FFT this block using a $2N$ point FFT. The $2N$ point FFT requires $N \log_2(2N)$ complex multiplications.
3. The taps in the frequency domain from stage (1) are multiplied by the data in the frequency domain from part (2). This requires $2N$ complex multiplications
4. The resulting vector is then transformed back into the time domain using a $2N$ point IFFT, which we can denote \mathbf{x}_0^* , again requiring $N \log_2(2N)$ complex multiplications
5. Discard the first N samples and outputting the last N samples to give $\mathbf{y}_0 = (x^*[N], x^*[N+1], \dots, x^*[2N-1])$
6. Then input the next block of data which overlaps by N samples with the previous block and repeat process 2 – 6 to get $\mathbf{y}_1, \mathbf{y}_2, \dots$.

The complexity of this algorithm will be $O(\log N)$ which is an improvement compared to the time domain implementation of the filter.

B Covid-19

I suffered from Covid-19 flue like symptoms from 17th of March until 27th of March during which I planned to work to finish all the computing for the project. Because of the illness I was had to push back the computing by 10 days. Afterwards I had some serious work to catch up as I was suppose to finish the first draft of the report by beginning of May. This meant that if more time would have been available I would have showed the capacity of the channel vary with the length of the fibre and observed the what are the effects of the PCS technique on this type of plot.

References

- [1] The Economist Author. The second half of humanity is joining the internet. <https://www.economist.com/leaders/2019/06/08/the-second-half-of-humanity-is-joining-the-internet>, 8th June 2019. [Online; accessed 22-Mar-2020].
- [2] The Economist Author. How the pursuit of leisure drives internet use. <https://www.economist.com/briefing/2019/06/08/how-the-pursuit-of-leisure-drives-internet-use>, 8th June 2019. [Online; accessed 22-Mar-2020].
- [3] Ofcom. Communications market report 2019. <https://www.ofcom.org.uk/-/data/assets/pdf-file/0028/155278/communications-market-report-2019.pdf>, 4th July 2019. [Online; accessed 22-Mar-2020].
- [4] John Graham-Cumming, Cloudflare. Internet performance during the covid-19 emergency. <https://blog.cloudflare.com/recent-trends-in-internet-traffic/>, 23rd April 2019. [Online; accessed 23-April-2020].
- [5] The Economist Author. The splinternet of things threatens 5g's potential. <https://www.economist.com/the-world-in/2019/12/25/the-splinternet-of-things-threatens-5gs-potential>, 25 December 2019. [Online; accessed 23-April-2020].
- [6] Jakob Nielsen. Nielsen's law of internet bandwidth. <https://www.nngroup.com/articles/law-of-bandwidth/>, Apr. 4, 1998; Updated Sep. 27, 2019. [Online; accessed 24-April-2020].
- [7] Peter J. Winzer, David T. Neilson, and Andrew R. Chraplyvy. Fiber-optic transmission and networking: the previous 20 and the next 20 years. *Opt. Express*, 26(18):24190–24239, Sep 2018.
- [8] A. R. Chraplyvy, A. H. Gnauck, R. W. Tkach, and R. M. Derosier. 8*10 gb/s transmission through 280 km of dispersion-managed fiber. *IEEE Photonics Technology Letters*, 5(10):1233–1235, 1993.
- [9] NL) Geertman, Robert E. M. (Veldhoven. Method of making a twisted optical fiber with low polarization mode dispersion, November 2000.
- [10] Jesús Alvarez Guerrero and Ferney Amaya-Fernández. Comparison of electronic compensation techniques in access networks with optical phase modulation and coherent detection. *Revista Ingeniería y Competitividad*, 19:87–96, 02 2017.
- [11] G. Gavioli, E. Torrengo, G. Bosco, A. Carena, V. Curri, V. Miot, P. Poggigliini, M. Belmonte, F. Forghieri, C. Muzio, S. Piciaccia, A. Brinciotti, A. L. Porta, C. Lezzi, S. Savory, and S. Abrate. Investigation of the impact of ultra-narrow carrier spacing on the transmission of a 10-carrier 1tb/s superchannel. In *2010 Conference on Optical Fiber Communication (OFC/NFOEC), collocated National Fiber Optic Engineers Conference*, pages 1–3, 2010.

- [12] P. J. Winzer and D. T. Neilson. From scaling disparities to integrated parallelism: A decathlon for a decade. *Journal of Lightwave Technology*, 35(5):1099–1115, 2017.
- [13] J. Cho and P. J. Winzer. Probabilistic constellation shaping for optical fiber communications. *Journal of Lightwave Technology*, 37(6):1590–1607, 2019.
- [14] K. McClaning and T. Vito. *Radio Receiver Design*. EngineeringPro collection. Noble Publishing Corporation, 2000.
- [15] G. Turin. An introduction to matched filters. *IRE Transactions on Information Theory*, 6(3):311–329, 1960.
- [16] Shiva Kumar. *Fiber optic communications : fundamentals and applications / Shiva Kumar and M. Jamal Deen, Department of Electrical and Computer Engineering, McMaster University, Canada*. 2014.
- [17] R. Johnson, P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas. Blind equalization using the constant modulus criterion: a review. *Proceedings of the IEEE*, 86(10):1927–1950, 1998.
- [18] F. P. Guiomar, J. D. Reis, A. L. Teixeira, and A. N. Pinto. Mitigation of intra-channel nonlinearities using a frequency-domain volterra series equalizer. In *2011 37th European Conference and Exhibition on Optical Communication*, pages 1–3, 2011.
- [19] Kyeongyeon Kim, Nargiz Kalantarova, Suleyman Serdar Kozat, and Andrew C Singer. Linear mmse-optimal turbo equalization using context trees. *IEEE transactions on signal processing*, 61(12):3041–3055, 2013.
- [20] O. Golani, M. Feder, and M. Shtaif. Kalman-mlse equalization of nonlinear noise. In *2017 Optical Fiber Communications Conference and Exhibition (OFC)*, pages 1–3, 2017.
- [21] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. Wiley, 2012.
- [22] Fernando Guiomar. Optdsp: Digital signal processing matlab library for the simulation of coherent optical communication systems (v1.1). Sep 2017.
- [23] Dario Pilori. Advanced digital signal processing techniques for high-speed optical communications links, 2019.
- [24] G. Böcherer, F. Steiner, and P. Schulte. Bandwidth efficient and rate-matched low-density parity-check coded modulation. *IEEE Transactions on Communications*, 63(12):4651–4665, 2015.
- [25] P. Schulte and G. Böcherer. Constant composition distribution matching. *IEEE Transactions on Information Theory*, 62(1):430–434, 2016.
- [26] Patrick Schulte, Fabian Steiner, and Georg Böcherer. shapecomm WebDM: Online constant composition distribution matcher. <http://dm.shapecomm.de>, July 2017.