Master's thesis in

Machine Learning Enabled Nonlinear Phase Noise Mitigation for Coherent Optical Systems

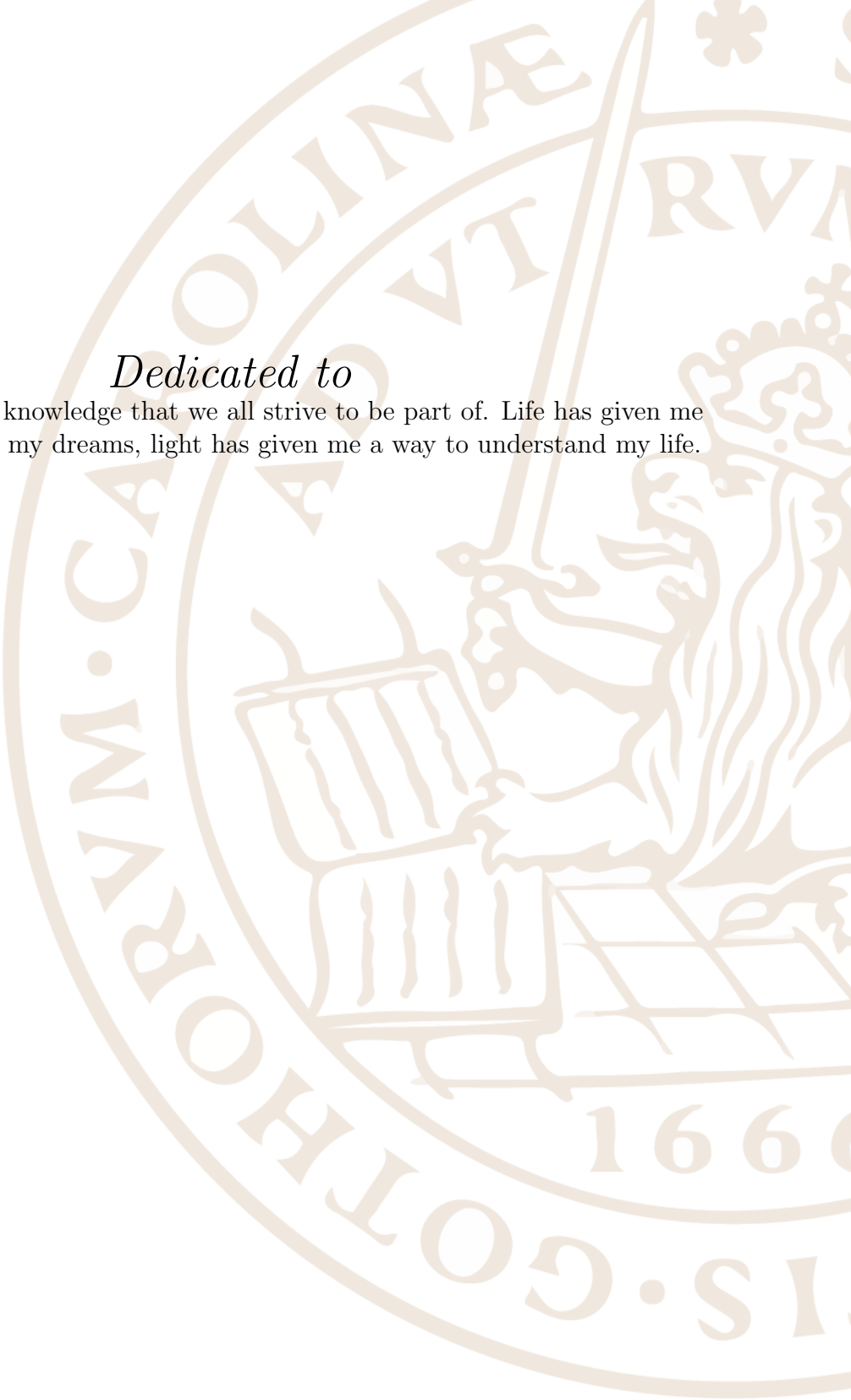
Sergio Steven Gutiérrez Pulgarín

Supervisor: Dr. Cord Arnold **Co-supervisor: Dr. Sergei Popov**

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Photonics

LUND UNIVERSITY

Department of Physics
Atomic Physics Division
Lund, Sweden
May, 2020



Dedicated to

the endless source of knowledge that we all strive to be part of. Life has given me the chance to pursue my dreams, light has given me a way to understand my life.

Machine Learning Enabled Nonlinear Phase Noise Mitigation for Coherent Optical Systems

Sergio Steven Gutiérrez Pulgarín

Submitted for the degree of Master's in Photonics
May 2020

Abstract

In the current development of coherent optical communication systems, nonlinear noise is considered to be the ultimate bottleneck when extending the transmission length. In this report we suggest a nonparamter digital signal processing scheme to extend the transmission length of a fiber link. The processing scheme was enabled by machine learning, implemented to compensate nonlinear noise in a communication system without needing any information about the physical state of the transmission line. It was shown that in the case of nonlinear phase noise in a long-haul fiber system, the proposed processing scheme could extend the transmission length of the fiber link. However, for interchannel noise a clear benefit could not be determined due to limitations in the simulation. It was concluded that for a long-haul fiber link, knowledge of the system could be replaced with learning through an optimal statistical algorithm.

ACKNOWLEDGEMENTS

I want to express my gratitude to my supervisor Cord Arnold for all his constant support throughout my whole master's. In the past two years Cord has given me all the tools to pursue the path I desired with my education at Lund. Special thanks to Cord for encouraging me to contact the Faculty of Optical Communication at *Kungliga Tekniska Högskolan* (KTH Royal Institute of Technology) for my thesis project.

The research carried out in this report was only possible thanks to the guidance and expertise of Sergei Popov. In the months I spend in Stockholm Sergei oversaw and helped me overcome difficulties in my project. I want to thank him for the opportunity to work at KTH University and for facilitating the assistance of the *Kista High-speed Transmission Lab* (Kista HST-Lab) at the *Research Institutes of Sweden* (RISE). This project was not only a a great professional opportunity but it was the experience of a life time.

I would like to mention and express my appreciation to Aleksejs Udalcovs, Xiaodan Pang and Oskars Ozolins from the joint group KTH/RISE for their input throughout my thesis. With their expertise I was able to understand more in depth the objective and results of my project. Thanks to their input I was able to acquire a technical understanding of the simulations implemented.

In a personal note I want to thank Stefan Höst for listening to me through my whole project and giving me valuable insight from an Information Theory point of view.

The physical and technical contribution of Kista High-speed Transmission Lab and KTH Royal Institute of Technology is truly appreciated. Without their support this project could not have reached its goal. To conclude I want to thank *Lund University*, this will always be the start of a new chapter in my life, and I will always be grateful for all the experiences I have had in this journey.

Sergio Steven Gutiérrez

CONTENTS

LIST OF FIGURES

LISTINGS

CHAPTER

ONE

INTRODUCTION

Throughout human history technological innovation has redefined our experience in the world, changing the way we interact with each other and the perception of the distances that separate us. Communication is a basic need for anyone; however, our expectations on what communicating means has transformed through time. Most recently, from the beginning of the new millennium we require fast, long reach and high capacity telecommunication systems, like the telephone network or the internet. This are not simple requirements, consider as an example that just under 250 years ago the United States Postal Services was established to bring a way of sending and receiving information through a robust network throughout North America. Which was a great innovation at the time, but it does not come close to fulfill the requirements that seem natural to us today. Our current communication needs are in some ways more similar to the mailing system than to a phone network in the end of the 19th century. We expect to send and receive packets of any size, to a location as convenient as possible, with reliable service and with little hassle to the user. A phone network from the end of the 19th century was nothing close to this, usually with single phone for large areas. It was connected by operators that required information from the user and had poor coverage with low reliability.

Today we take for granted the technology developed to satisfy our need to communicate over a reliable and robust network. It is common practice for most of us to connect with family members or friends that may be anywhere around the world or request information from any server. Just two decades ago it wasn't possible to depend so reliably on the communications networks due to low capacity, lack of redundancies and inefficient routing schemes that added latency to the network [10]. As powerful as our current networks might be, it is the human way to find higher and more complex requirements for the ways we communicate. When the first consumer deployment of 3G in 2001 was implemented the consumer market was excited about a new service nicknamed 'SMS on steroids' [11]. Today we know that 3G opened the door to a whole set of services, that today are a normal part of our lives.

Current networks can handle heavy loads, but in the future the need of low latency communication (for self-driving cars or remote surgeries), the Internet of Things (IoT), high quality video streaming and other services, will require creative solutions to the current limits of our communication systems. New technology will be

deployed in the network, improving its capacity, but there is also a need to extract as much use from older networks to maintain there current coverage and reduce costs. Coherent Optical telecommunications gives a path forward that improves the network as whole. Improving the current operating network and pushing the limits for the future transparent network [12, 13].

1.1 General Objective & Motivation

In coherent optical communication the effects of dispersion and attenuation can now be compensated for. However, the noise that arises from the interaction between the optical power and the fibers optical Kerr effect limit the transmission length of the system [14]. The noise sources that arise from this interaction are known as nonlinear interference noise. The aim of this thesis is to study the effects of nonlinear noise in a coherent optical communication system and suggest a *nonparameter digital signal processing detection scheme* enabled by machine learning. Two communication links are built, a *long-haul* and a *dense wavelength division multiplexed* system, to detect nonlinear interference noise. Each communication link architecture is used to test if the suggested digital signal processing scheme improves the transmission distance of the signal compered to a traditional scheme.

Bringing together machine learning and optical communications has been on surge for the past decade being implemented in performance monitoring [15, 16], modulation format identification [17] and amplitude/phase noise characterization [18]. This techniques improve the performance of the system and extends the operational capacity. Increasing the transmission length of the optical system allows a further reach simplifying the deployment of the network, and reducing the need of re-transmitters. Given that the most expensive part of deployment is digging and laying optical fiber it is necessary to reduce the deployment complexity and cost. Getting more use from older network segments helps improve coverage as well, resources can be allocated to reach more people instead of having to lay again the same network segments. The current bottleneck in a fiber link is the presence of nonlinear noise [7]. Machine learning could be a way to mitigate nonlinear noise without needing any information on the physical state of the fiber system [19]. This is a relatively novel idea that has been implemented in References [18,19] and others. Our approach implements algorithms that have been improved over time and now enable more efficient and robust predictions.

1.2 Optical Telecommunications

Transporting information requires a protocol that standardizes a measurement of a physical events. Through time there has been many information links used to communicate at a distance, from a smoke signal or a telegraph, up to the telephone. All these have revolutionized our way of communicating, and the reach at which we can expect to get an immediate response. Optical communication is the field interested in studying the implementation of communication systems using light as an information carrier. The beginning of Optical communication by this definition

could be traced back to the Greeks in 1084 B.C where its said that a 500-km-long line of fire beacons was built to convey urgent messages, like the fall of Troy [10]. In the framework we understand optical communication now, a first proposal of a Fiber-Optic Communication was done by Jun-ichi Nishizawa in 1963 at Tohoku University [20]. This type of lightwave system uses optical fiber to contain a light pulse, guiding it to a receiver, where it is detected and decoded into a digital signal containing information like voice message or video.

Today fiber-optic communication is the most prominent technology in long-haul communication. Enabling the vision of a global society, thanks to its effective reduction of immediate communication length. Fiber optics has also had a large implementation in the metro networks in the paste decade, due to improvements in fibers made out of polymers, reducing the cost of deployment [21]. When we place a call or use the internet, we expect a fast, reliable and robust coverage. These requirements set high constraints into the technology development. Not all communication needs can be satisfied with the same technology.

Currently, a combination of electrical and optical systems are used depending on the requirements given, each scheme has benefits an drawbacks. For long haul communications it is necessary to have a system that can be implemented over hundreds of kilometers, with high capacity and low latency, making fiber optics an ideal technology. Today's optical communications systems can fulfill these requirements, but it has taken a lot of effort to develop and perfect a reliable communication link [22].

1.2.1 Brief History of Optical Communication

A more contemporaneous start to *optical communications* would be the invention of the optical semaphore telegraph in the 1790s by C. Chappe [23]. However, it took one more century for a new development in the field, in 1880s A. Bell created the Photophone an optical telephone system that never saw much use [10, 23]. The 1960 marked a true revolution in optical communication thanks to the invention of the *laser* and *glass fiber* [24]. Nobel prize winner for optical fibers C.K. Koa, G. Hockman and J. Nishizawa where among the first to notice the potential of this new technologies [20, 25], they realized that the attenuation in a fiber could be reduced to a usable value of 20 dB/km by decreasing the impurities in the glass fiber. At this time the attenuation was approximately 1000 dB/km in a usual fiber. *Corning Glass Works* in 1970 was able to manufacture the worlds first low-loss single mode fiber operating at 633 nm with an attenuation of 20 dB/km, making it commercially available by 1975 [10, 23]. The technological developments achieved through the history of Optical Communications can been divided into four generations. Each generation is distinguished by its operating wavelength, data rate and transmission distance.

In the starts of 1975 the commercially available communication system was a multimode fiber with a 800 nm GaAs semiconductor laser. The systems main limitations was the high fiber loss, intermode and intramode dispersion [26]. The performance of this system was 45 Mbit/s for a transmission distance of 10 km without repeaters. A set of improvements would be carried out in the next five years that tripled the data rate; thus, naming this first implementation a *first generation* link. [22]

Improvements done in 1980 allowed further reach and higher bit rate operation. By changing the operating wavelength and breakthroughs in glass fiber manufacturing substantial improvements to the technology where archived [22, 27]. Marking the start of the *second generation* link, which had a new InGaAsP semiconductor laser operating at 1330 nm with a low loss -at the time- 1dB/km single mode fiber. Second generation optical communication systems could reach a data rate of 1.7 Gb/s with a transmission distance of 50 km without repeaters [12, 27]. A significant improvement over the first generation; however, the fiber losses where still high and no dispersion management had been implemented yet.

At the beginning of the 1990 a solution was found to mitigate the fiber losses and have the minimum dispersion coefficient around the operating wavelength. The development of dispersion shifted fibers [22] enabled deploying longer fibers without presenting dispersive effects. The new communication system operated with a low-loss 0.2 dB/km single mode fiber at 1550 nm. The two main breakthroughs in this *third generation* system was the shift to an operating wavelength with minimum attenuation, and the improvement in the core-cladding fiber design such that the chromatic dispersion is minimum at the 1550 nm wavelength region. This system achieved a data rate of 2.5 Gbit/s at a transmission distance of 100 km without repeaters [12,27]. A stable phase coherent light source, multiplexing and transparent amplification where the last key technologies needed for a robust communication system.

The last three decades have brought a fleet of improvements, three different types of information multiplexing, the invention of transparent amplification and higher order modulation formats, enabling 109 Tbit/s performance with further reach without repeaters. *Wave division multiplexing* (WDM), *polarization division multiplexing* (PDM) and *time division multiplexing* (TDM) [12, 27, 28] have helped extract as much capacity as possible from different communication systems. The *erbium dope fiber amplifier* (EDFA) gave way to longer transmission range without repeaters. With the *distributed feedback* (DFB) lasers with stable phase coherence, higher order modulation formats could be implemented due to the improved phase stability. *Fourth generation* optical communication system are defined by use of these technologies. Currently, fourth generation systems are still operational, and Telecom operators are interested in extending the life span of this systems as a cost effective update.

1.2.2 Current Outlook in Optical Communications

The lightwave systems currently deployed are referred to as *intensity modulation with direct detection* (IM/DD). An alternative to IM/DD is to transmit information by modulating the frequency, amplitude or phase of the optical carrier. Detection of the transmitted signal is done by a scheme known as *coherent detection*, that allows the extraction of all the signals information. Phase coherence with the optical carrier plays an important role in *higher-order modulation* (HOM) in a fiber link, also known as *coherent lightwave systems*. These systems fall into the field known as *Coherent Optical Communications*, that has been studied extensively since the 1980. Its deployment was slow at the time due to limitations on amplification and cost effective improvements in IM/DD systems and the poor phase stability of the

light sources [13, 29].

Coherent optical receivers follow a heterodyne detection scheme where a *local oscillator* (LO) is required to detect the incoming signal. When the LO and the transmitted signal are mixed, a beat frequency equal to the difference between the two frequencies is created. By adjusting the local oscillator frequency, the beat frequency retrieved could be the signal that was modulated on top of the optical carrier. These detection arrangement makes the receivers require more power and physical components. At the time IM/DD receivers where cheaper to run and manufacture making them more widely deployed. In current networks IM/DD system are reaching their limit [22], this has brought new interest to coherent systems, thanks to the possibility of improving already deployed links and giving space to create new solutions for future demands. Currently, the main objective of coherent optical communications is exploiting Photonics to engineer the next generation transparent network. However, a main limitation is phase noise distorting the signal while it propagates through the fiber, making noise compensation a vital part for a high quality link [30].

1.3 Modulation Formats

A HOM format is a scheme to standardize information for a correct representation of a signal, to extract a bit stream from a physical measurement. In optical communication the amplitude and phase of a DFB *continuous wave* laser is modulated to encode information, this information can later be decoded at the receiver side by heterodyne detection. A single detection of a combination of phase and amplitude components is denominate a *symbol*. A set of M unique symbol can each be mapped to a unique bit sequences m bits long. A graphical representation of the modulation of optical carrier is shown in Figure 1.3.1, which depicts the basic transformations that can be done to the carrier. In digital communication *amplitude* and *phase* are referred to as the *in-phase* and *quadrature* components or IQ components. It is usual to represent the a modulated signal in a phase diagram as shown in Figure 1.3.2, where you map each symbol by its IQ components, also known as a constellation point when displayed in a phase diagram. When the carrier signal is *amplitude modulation* (AM) or *phase modulated* (PM) the transformation under gone in the phase diagram is just a translation or rotation, as shown in the top left of Figure 1.3.2.

With this two basic transformations, an array of modulation formats are possible. A modulation format is in general a scheme to determine a set of equidistant points of unique IQ component that each represents a unique bit sequence, also called a constellation. The most simple case is the *binary phase shift keying* (BPSK) that consists of two unique constellation points. Determining how to classify a detected symbol is simple if it falls to the left of the y-axis the symbol is a 1 or if it falls to the right the symbol is a 0. Thus a simple decision region dividing in half the phase diagram would suffice to accurately classify the detected symbols for BPSK modulation.

An extension to phase modulated BPSK can be achieved with a clear signal, higher resolution on the receiver and increasing the amount of constellation points. Modulation formats like *quadrature phase shift keying* (QPSK) or *eight constellation*

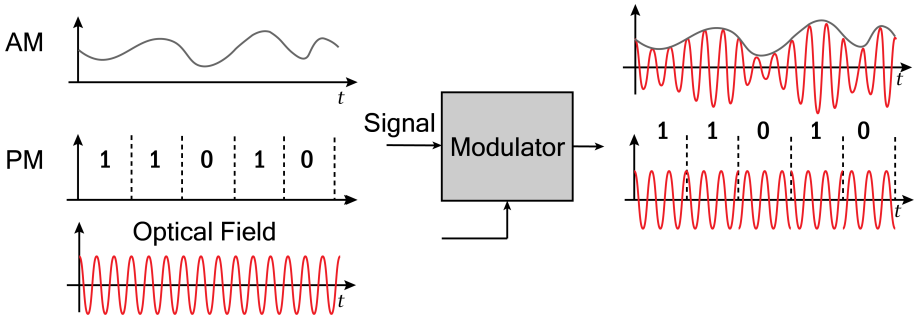


Figure 1.3.1: Amplitude modulation and phase modulation can be implemented to transmit information over a optical field. In this figure a visual representation on how information can be imprinted to a carrier signal is shown.

Image designed inspired by available figures in Reference [1]

point phase shift keying (8PSK) are extension on BPSK as can be seen from the constellation diagram (observe Figure 1.3.2). To transmit more bits per symbol, it is useful to have a compact constellation diagram. Thus, using both AM and PM is a logical next step to modulate the carrier signal.

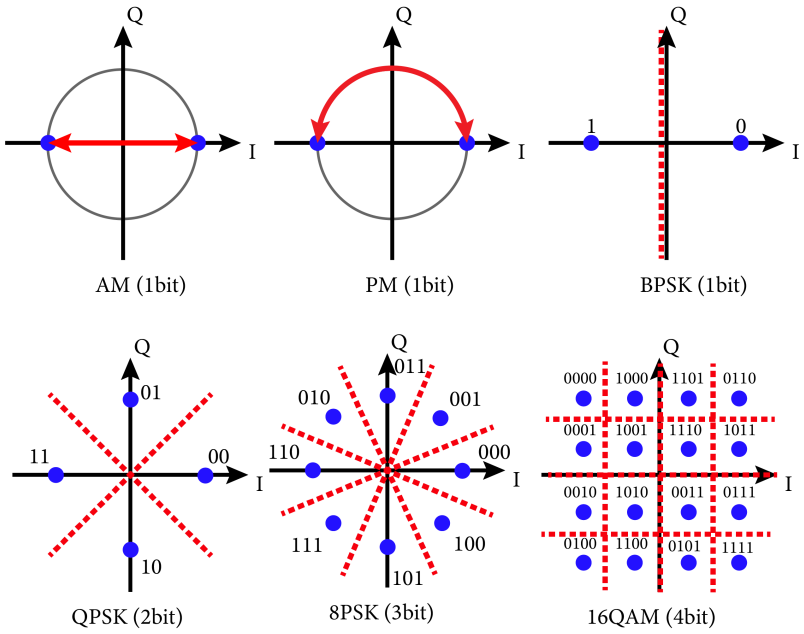


Figure 1.3.2: A phase diagram can be used to represent the changes in the state of detected signal. A HOM implements a more complex constellations architectures to transport more information per symbol. The more constellation points, more information per bit can be transmitted.

Image designed inspired by available figures in Reference [2]

A widely implemented modulation format in modern telecommunications is *quadrature amplitude modulation* (QAM), a technique in which two carriers shifted in phase by 90 degrees are modulated and combined. Due to this phase difference the carriers are in quadrature, giving the modulation format its name [31]. This technique allows a communication system to reach higher data rates within the same bandwidth,

obtaining a higher spectral capacity [32]. The number of bits m needed to decode each symbol, is related to the total number of symbols M (or constellation points) as $M = 2^m$ [33]. A representation of a 16QAM square constellation is shown in the bottom right of Figure 1.3.2, the corresponding constellation point and decode 4 bit message are displayed for each symbol. The quality of the link relies on having little noise, such that the transmitted symbol can be detected with its original phase and intensity. Making noise mitigation a necessary component for robust systems.

1.4 Optical Noise

Photonics gives a framework to describe and engineer a communications link that can have high capacity, low latency and long-distance coverage. A wide range of creative solutions are required to counter impairments that arise from the light-matter interaction between the light pulse and the fiber cable. When noise is present in a channel, it deteriorates the quality of the signal. The signal quality can be quantified by the Optical Signal to Noise Ratio (OSNR). Information is imparted onto a light pulse by modulating its physical properties within a predetermined scale or protocol like a HOM. The main purpose of a communication link is to transmit symbols that can be decoded into bits, the amount of errors encountered in a bit stream is known as *bit error rate* (BER). It is worth mentioning that the *symbol error rate* (SER) does not have to be the same as BER, depending on the modulation format and the *forward error correction* (FEC) there can be significant difference reducing the actual amount of mistakes made [31].

When light propagates through a medium it can suffer a change in speed depending on its wavelength, this effect is known as *chromatic dispersion* (CD) or *group velocity dispersion* (GVD). A pulse of light contains multiple wavelengths rather then a single frequency. Thus, when the pulse acquires CD it stretches in time, affecting the timing of the link, corrupting the information carried. Dispersion is not the only noise source in a fiber link, the inline optical amplifiers that counter the fiber attenuation produces *amplified spontaneous emission* (ASE) noise which is added to the signal as *additive white Gaussian noise* (AWGN), reducing the OSNR. Amplifiers contribute to other types of noise indirectly, like the noise produced by intense light propagating through a medium [29].

The architecture of the detector can also add noise to the received signal. Depending on the resolution of a heterodyne detection or the offset between the incoming signal and the LO. This phase difference is known as the *carrier frequency offset* (CFO). Linear sources of noise arise primarily from the response of the dielectric -the silica optical fiber- to low intensity light propagating through it.

In the regime of low intensity light a transparent medium is represented by its refractive index. The electrical susceptibility of the medium is determines the mediums optical response. The linear or first order response of the electrical susceptibility is related to the refractive index. The higher order or nonlinear response of the susceptibility has a small contribution to the total interaction at low intensity. The higher order effects only have significant contribution in the presence of an intense electromagnetic field, like an intense laser or tightly confined light, as is present in the waveguide geometry of a fiber. Some part of these noise sources can be mitigated

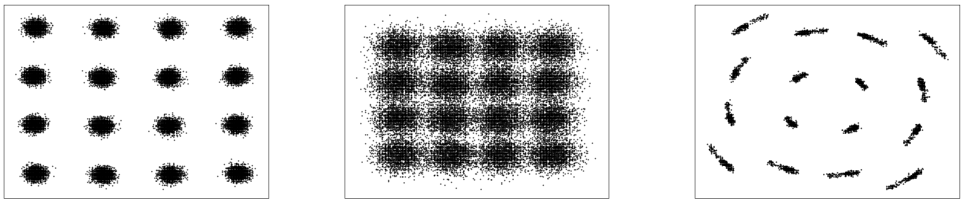
using *digital signal processing* (DSP) or clever exploitation of physical phenomena, extending the throughput of the link [34].

A critical limiting factor is *nonlinear interference noise* (NLIN) produced by the interaction between the signal and the ASE noise via the higher order response of the electrical susceptibility. Compensating for nonlinear effects is possible, using pre-compensation or DSP [14]. However, the main adversity to these techniques is the requirement on knowing the physical sate of the link. Having the physical state of the system is possible in a point-to-point link but it is not practical for a deploying a dynamic optical network.

Nonlinear noise sets the limit for the maximum amount of information transmitted over a fixed distance of an optical communication link [35]. Restriction in optical power help reduce nonlinear noise, but high power is also necessary in HOM to obtain the necessary OSNR, limiting how low the power should be set. Noise mitigation becomes then interesting to retrieve an accurate representation of the link that permits computing decision regions for the transmitted symbols based on the links configuration. However, having a complete physical representation of the link is an impairment in real deployment, since its cumbersome to keep count of system length, type of fiber etc. Nonparameter noise mitigation is a novel solution to mitigate nonlinear noise in a system where the physical attributes are unknown. The detection scheme would learn the most reliable decision regions based on the noisy signal [36].

1.4.1 Optical Noise in a QAM Channel

In a QAM channel NLIN is imparted as a phase shift to each symbol. Since the HOM classifies each symbol based on its phase (Figure 1.3.2), a phase shift could change the classification of the symbol. The reduction of the OSNR due to the noise produced by the linear susceptibility is visualized in Figure 1.4.1, where the different sources of noise previously mentioned are shown for the case of a square 16QAM constellation. Mitigation techniques can improve the quality of the signal in the case of CFO, but improving the OSNR is not possible due to the non-stochastic nature of linear noise.

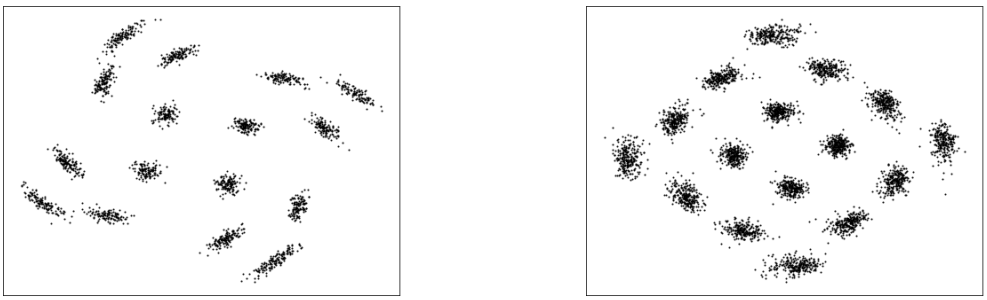


(a) Standard OSNR 25dB (b) Low OSNR 15dB (c) CFO phase noise

Figure 1.4.1: Square 16QAM constellation comparison between different types of linear noise.

In the presence of an intense light source, as is needed in QAM, nonlinear noise distorts the signal and hampers the transmission distance. Nonlinear effects also

alter the constellation of a QAM link, an example of this distortions is shown in Figure 1.4.2. Nonlinear noise has the advantage of being stochastic source of noise, such that a probability distribution could be extracted from the noise [37]. From Figure 1.4.2 it is noticeable that the decision region for each symbol has to be more complex than in the case of Figure 1.4.1.



(a) Gordon-Mollenauer effect (b) Cross phase modulation

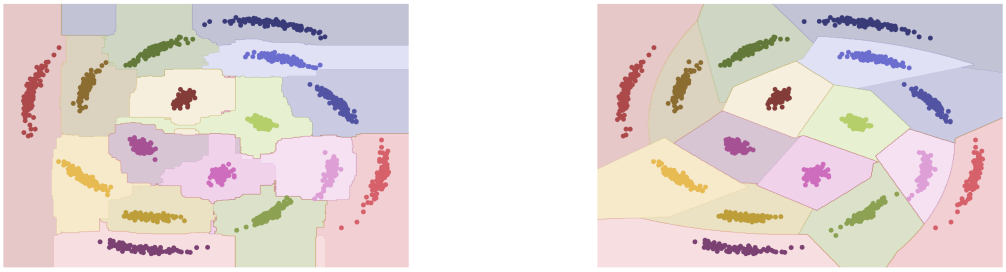
Figure 1.4.2: 16QAM constellation comparison between different types of Nonlinear Noise

Complicated decision regions have been implemented through DSP and nonlinear back propagation [7] showing improvements on a known system. An other approach is being done through *machine learning* which show unique advantages [19, 37] by making accurate predictions with out the need of back propagation or heavy numerical analysis. A more detailed description of the DSP and machine learning will be presented later on in this report.

1.5 Machine Learning

Learning has been a fundamental ability for humans throughout time. Observing our surrounding, using tools and developing technological advancements are only possible due to our ability to remember previous experiences to make prediction of the future. Since 1957 with the development of more powerful computers and more sophisticated software, the notion that a machine could *learn from data* in comparable way as humans *learn from experience* has captivated many [38]. As currently defined Artificial Intelligence (AI) uses a computer to perform cognitive tasks; such as, perception, learning, reasoning or other similar abilities. A branch of AI is *machine learning* (ML), based on the idea that patterns and trends in a given dataset can be learned through some algorithm. From the learned structures a decision or prediction can be made for some other dataset in the system of interest [39].

Machine learning is offering a different perspective to problems that traditionally have been solved with complicated numerical models. The term nonparameter detection schemes comes from the idea that ML can suggest decision regions without needing any information on the current state of the link. Thus, ML can be a tool to work around imperfections in a fiber, or as a scheme to characterize nonlinear



(a) Random Forest (b) Support Vector Machine

Figure 1.5.1: Machine learning determined decision regions for 16QAM constellation with NLIN acquired from 26 segments of 80 kilometer fiber.

noise in a signal. In Figure 1.5.1 two different algorithms are implementedto predict the best decision regions for a noisy constellation. Notice how this two ML algorithms perform differently for the same dataset due to their individual learning approach. Our interest is using this algorithms to extend the transmission distance by improving the prediction capability from a noisy signals.

1.6 Thesis Organization

This thesis is divided into five chapters that cover the physical description of the system, the digital processing algorithms used and the simulation of the system. In the next chapter the physical layer will be discussed, describing the signal detection, modulation and the light-matter interaction undergone by the propagating signal. In Chapter 3 the ML algorithms and an analytically derived nonlinear noise compensating algorithm are introduced. After this the simulation of the two communication links, and the retrieved data will be presented in Chapter 4. To conclude this report the results will be discussed in Chapter 4, with final remarks and conclusion in Chapter 5.

CHAPTER

TWO

PHYSICAL LAYER

In this chapter the physical description and basic concepts implemented through this report will be discussed. The following are the basic assumptions and technology used to design, simulate and analysis the results of this thesis. In each section a different element of the communication link will be discussed, ultimately from this understanding it is perused to investigate nonlinear noise and its fundamental role in the transmission link.

2.1 Emission, Detection and Modulation

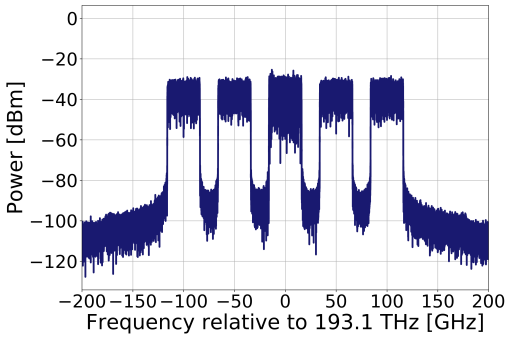
The components of a coherent optical communication system will be given a short description in this section. The basic components for an optical communication system are the *laser*, *modulator* and *detector*. For a coherent system the modulation format is an important design parameter, as was mentioned in the introduction, noise effects affect in different ways a given modulation format. Each component plays a fundamental role in a optical system; the correct selection of this elements can improve the throughput and transmission distance of a link. First let us consider the emission source that enabled coherent communication due to its high phase coherence, the *distributed feedback laser* (DFB) laser.

2.1.1 Light Emitting Source

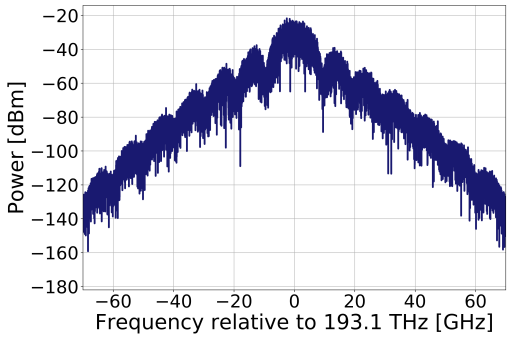
In the beginning, coherent optical communications did not get as much attention as IM/DD systems, because in the 1990 all available light sources presented large phase noise and frequency drift [40], which is not useful for a coherent system. With the invention of the DFB laser new interest sprung in coherent systems. The main advantage of a DFB laser is the ability to emit a single longitudinal mode, which implies high frequency and phase stability.

A DFB laser is either a semiconductor laser or fiber laser where its whole resonator consists of a periodic structure, that acts as a distributed reflector. This reflector can be designed to reflect the desired wavelength range, which is amplified in the gain medium and emitted by the laser [1, 41]. The emitted optical spectrum for a general DFB laser is shown in Figure 2.1.1c, where we can see that a single

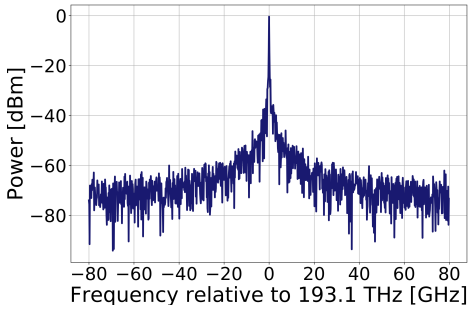
mode is prevalent over all other noise present. The linewidth of a DFB laser is usually bellow 10 MHz [40] and for this implementation a narrow linewidth laser was simulated with a 1MHz linewidth. Wavelength tunability is also possible for a range of 1000 GHz or a wavelength range of 8 nm, allowing the implementation of WDM systems [40,41]. A critical advantage for coherent systems is the use of HOM, which increase the systems capacity for the same bandwidth. Thus, modulating a DFB laser and detecting its physical properties like intensity, amplitude and phase are necessary to achieve higher system capacity.



(a) Nyquist spectral shaping on 5 channels



(b) Single channel QAM modulation



(c) No Modulation

Figure 2.1.1: Optical frequency spectrum for different systems. A DFB laser frequency spectrum is shown in (c) and a set frequency spectrum after modulation are also shown in (a) and (b).

To determine the operating power of a fiber system two noise sources constrain the system. The first noise source arises from the optical fibers refractive index dependence on the transmitted signal power [29]. With a high power signal distortions can be induced if power fluctuations interact with the electrical susceptibility of the fiber. The second noise sources is common amplification effect when a low power signal is amplified AWGN noise is added [42], decreasing the OSNR and transmission length. Optical modulators require a strong carrier signal to operate at high OSNR. This requirements limits the minimum operating power. To determine the optimal operating range a detailed analysis will be discussed throughout this thesis.

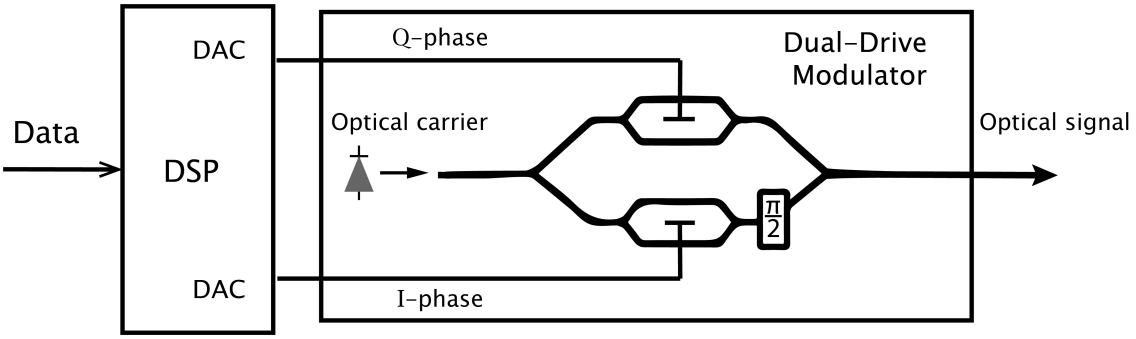
2.1.2 Modulator and Detector

Some of the advantages of optical communication systems rely on the modulation of the carrier signal and on the high sensitivity detection. To transmit more information over the same bandwidth a clever representation of the signal has to be standardized, as is done by HOM. However, it is necessary to modulate the carrier signal physical properties with precision to control reliably the information transmitted. Detection in coherent optical communication requires a scheme that can extract simultaneously multiple physical measurements from a single pulse. These two requirements can be solved with clever implementation and design of well understood optical phenomena.

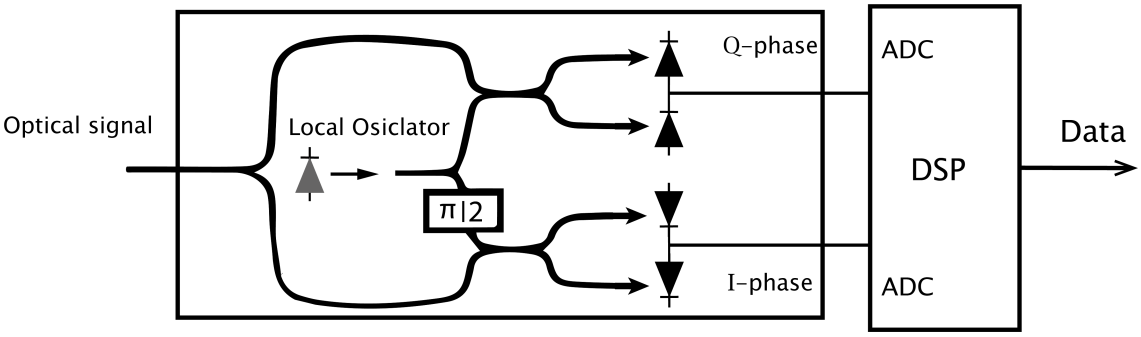
In coherent communication the amplitude and phase of the carrier signal are modulated to transmit information. For an optical system *phase modulation* (PM) can be achieved by passing through a non-centrosymmetric crystal that presents the Pockels effect [4, 29]. This type of crystal have refractive index dependent on the applied voltage allowing PM [4]. If a set of this crystals are used in a Mach-Zehnder configuration *amplitude modulation* (AM) can be achieved. This type of modulators are known as *Mach-Zehnder-type push-pull modulators* or *Mach-Zehnder modulator* (MZM) [4]. To obtain a simultaneous amplitude and phase (In-phase and quadrature components) modulation two MZM in parallel with a $\pi/2$ phase shift can be combined as shown in Figure 2.1.1a, modulating independently the IO components [4]. Any kind of multilevel modulation format can be done with this scheme [3],currently integrated Dual-drive MZM $LiNbO_3$ IQ modulators are available on the market. This is the same modulator architecture that was used for the simulation.

Modulation of the CW optical carrier not only alters the IQ components but it also modifies the optical spectrum of the laser. When the optical carrier undergoes QAM, the optical spectrum of the transmitted signal is altered as seen in Figure 3.1.1b. In Figure 2.1.1 the modulator is controlled by the *digital to analog converters* (DAC) and the driving signal is computed by *digital signal processing* (DSP) [4]. Filtering can be done by the DSP in the digital domain to modify the optical frequency spectrum of the signal and frequency spectrum close to a square pulse [4, 29], as seen in Figure 2.1.1a. The type of filtering implemented is a low-pass filter with a transfer function response as a raised cosine function, aslo know as *raised-cosine filter*. This type of filtering is known as Nyquist filtering and helps reduce the inter-symbol interference and maximize the available bandwidth [43], both of this effects can be observed in Figure2.1.1a. Reducing inter-symbol interference alludes to the clear spacing between channels, and maximizing the bandwidth can also be achieved given that each channel is constrained in a small bandwidth allowing more channels to be add in the side bands.

Coherent detection of the quadrature and in-phase components of the optical signal can be done simultaneously using a *phase-diversity homodyne receiver*. The fundamental idea behind coherent detection is to compute the inner product of the *local oscillator* (LO) phasor and the received signal phasor to determine one of the signals quadratures [4]. A homodyne receiver has LO operating at the same frequency as the carrier signal [29]. However, to detect both IQ components simultaneously two LO signals are needed with a $\pi/2$ offset. This LO architecture,



(a) Dual-drive MZM



(b) Phase-diversity homodyne reciver

Figure 2.1.1: Detection and modulation scheme for a coherent optical communication system.
Image designed inspired by available figures in Reference [3, 4]

as shown in Figure 2.1.1b, is called phase-diversity detection. Polarization diversity detection can be done with another LO that has orthogonal polarization, most commonly a random polarization state laser is used that is then divided using a polarization beam splitter [4]. As for the modulator, an *analog to digital conversion* (ADC) converts the signal to a digital sequence that can be processed by the DSP. Many different DSP schemes exist today to process an optical signal and retrieve the carriers phase [44], determine the timing (clock recovery) [45], LO offset mitigation [4, 44] and other processing can be done to improve the signal quality.

2.1.3 Modulation Format

In the previous section, a short description of a dual-drive MZM was given. In this section, the *multi-level modulation* (or higher order modulation HOM) of the optical carrier will be discussed, which improves the link's spectral efficiency. In the digital domain the transmission of bits is mapped to a set of IQ components that determine the driving voltage of the modulator, producing the encoded optical signal that can be decoded on the receiver side. For the digital transmission, m bits are collected and mapped to a complex symbol chosen from an alphabet

$$s(k) = i_k + jq_k \in \{s_0, s_1, ..., s_{M-1}\}\,,\ M = 2^m \tag{2.1.1}$$

where k represents the symbols interval with length $T_S = mT_B$ and $1/T_B$ is the bit rate [46]. One of the symbols in the alphabet is assigned depending on the bit

combination, that is defined in a *constellation diagram*, where i and q are the IQ component of each constellation point. When m bit alphabet is designed, a M point constellation is defined, the architecture of this constellation can follow a determined standard like PSK, QPSK or QAM (as discussed in Section 1.3. In this project a 4 bit alphabet with a square 16QAM constellation was implemented as is shown in Figure 2.1.1. A detailed explanation on how to determine the complex function that modulates the carrier signal and what electrical driving signal should sent the the MZM can be found in Reference [46]. With this basic principals, our interest is shifted to the interaction between the transmitted optical signal and the optical fiber that guides the signal to the receiver. This interaction deteriorates the quality of the channel by corrupting the phase information imprinted by the modulators.

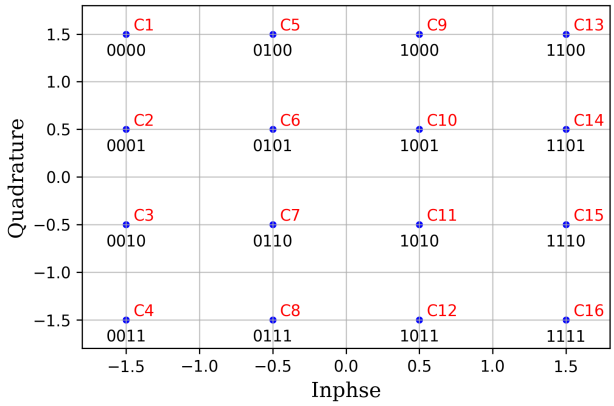


Figure 2.1.1: Four bit per symbol signal represented in a phase diagram of a square 16QAM constellation.

2.2 Electromagnetic Theory of Light

The current theory of light comes from the idea of photons as the smallest unit of radiation composed of electromagnetic energy. As described by quantum mechanics a photon can be represented either by a wave or a particle. For the propagation and nonlinear interaction it is convenient to work in the wave picture, while at the detector you detect individual particles. A photon can be represented by an electromagnetic field that arises from a set of coupled vector fields that are functions of time and position: the *electric field* $\mathbf{E}(\mathbf{r}, t)$ and the *Magnetic field* $\mathbf{H}(\mathbf{r}, t)$. The phenomenological description of the interaction between this fields are known as the *Maxwell's equations* which describe all electromagnetic phenomena [29].

2.2.1 Maxwell Equations in a Medium

In this thesis, the interaction between a modulated optical signal and silica glass fiber is studied to get a better understanding of nonlinear noise that corrupts an optical communication system. For a fundamental understanding, it is necessary to take into account the physical interaction of the medium with the electromagnetic

field. A mediums response in the presence of a electromagnetic field is described by two vector fields: the *electric flux density* $\mathbf{D}(\mathbf{r}, t)$ and the *magnetic flux density* $\mathbf{B}(\mathbf{r}, t)$. For an optical communication system we can assume an optical fiber that is a nonconducting medium and has no free charges or currents, simplifying the interaction. These four fields are related by the Maxwell's equations in a source-free medium [1, 29]:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.2.1a}$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} \tag{2.2.1b}$$

$$\nabla \cdot \mathbf{D} = 0 \tag{2.2.1c}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.2.1d}$$

The electric flux density $\mathbf{D}$ relation to electric vector field $\mathbf{E}$ is dependent on the properties of the medium, that are characterized by the *polarization density* $\mathbf{P}$. For a dielectric medium the polarization density is the sum of the electric dipole moments induced by the electric field [1]. In an analog manner, the magnetic flux density $\mathbf{B}$ relationship with the magnetic field $\mathbf{H}$ is dependent on the *magnetization density* $\mathbf{M}$ that describes the properties of the material. The connection between the vector fields and the flux densities are given by

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \tag{2.2.2a}$$

$$\mathbf{B} = \mu_0 \mathbf{H} + \mathbf{M} \tag{2.2.2b}$$

where ϵ_0 and μ_0 are the vacuum permittivity and permeability respectably. Due to the nonmagnetic nature of silica glass, an optical fiber has $\mathbf{M} = 0$. When operating far from the mediums resonance, a phenomenological relation between $\mathbf{P}$ and $\mathbf{E}$ can be used. This is the case for a optical fiber since we are interested in working in the low-loss wavelength region far from resonance.

In this project, nonlinear effects are the main objective, but it is necessary to have a description of linear dispersion effects as well. For a dispersive medium with a *electric susceptibility* χ the relation between $\mathbf{P}$ an $\mathbf{E}$ is given as [29]:

$$\mathbf{P}(\mathbf{r}, t) = \epsilon_0 \int_{-\infty}^{\infty} \chi(\mathbf{r}, t - t')\mathbf{E}(\mathbf{r}, t')dt' \tag{2.2.3}$$

this response is dynamic, given that the electric field induces the bound electrons to start oscillating, and collectively create the polarization density. Given the oscillatory response of the bound electrons, different harmonics can be reached. The electric susceptibility describes the strength of different harmonic components in respect to the incoming field.

In the simplest case of a *linear*, *nondispersive*, *homogeneous* and *isotropic* dielectric medium the the vectors $\mathbf{P}$ and $\mathbf{E}$ are parallel and proportional for every position and time, such that $\mathbf{P} = \epsilon_0 \chi \mathbf{E}$ [1, 29]. Using the definition of the electric flux density $\mathbf{D}$ the *electric permittivity* of the medium can be computed as

$\epsilon = \epsilon_0(1+\chi)$, which for a nonmagnetic material is connected to the refractive index as $n = \sqrt{\epsilon/\epsilon_0} = \sqrt{1+\chi}$. From now on all materials discussed are *homogeneous* and *isotropic* making the relation between $\mathbf{P}$ and $\mathbf{E}$ independent of position $\mathbf{r}$ and independent from the direction of $\mathbf{E}$. Drawing some insight from the linear interaction response, a nonlinear dielectric media can be understood as an extension of this case. By considering a small change to the polarization density due to higher order oscillatory dependencies of the electric field and the medium.

2.2.2 Nonlinear Dielectric Media

Linear dielectric media gives a framework that can be extended to take into account higher order components of the electric susceptibility interaction with nonlinear orders of the electric field. Let us consider the electric field vector $\mathbf{E}(t)$, only time dependent, as a single harmonic electric field of angular frequency ω and complex amplitude $\mathcal{E}(\omega)$ written as [1]

$$\mathbf{E}(t) = \mathbf{Re}\left\{\mathcal{E}(\omega)e^{i\omega t}\right\} = \frac{1}{2}\left[\mathcal{E}(\omega)e^{i\omega t} + \mathcal{E}^*(\omega)e^{-i\omega t}\right]. \tag{2.2.1}$$

This electric field propagates through a medium as dictated by Maxwell's equations, thus the polarization density describes the mediums response.

Inter atomic or crystalline field forces are much greater than the externally applied optical field, such that nonlinearity is usually weak. Thus, the relation between $\mathbf{E}$ and $\mathbf{P}$ is approximately linear for small $\mathbf{E}$, shifting slightly from linearity as $\mathbf{E}$ increases. For a homogeneous isotropic dielectric medium the polarization density can be expanded as a power series around $\mathbf{E}(t) = 0$, as done in Equation 2.2.2a [1].

Nonlinear Polarization Density Frequency Response

The polarization density for a nonlinear dielectric medium is expressed as the sum of the linear component ($\epsilon_0\chi\boldsymbol{E}$) and the nonlinear component ($\boldsymbol{P}_{NL}$), which describes the medium's response to an incident electric field.

$$\mathbf{P}(t) = \epsilon_0\left(\chi\mathbf{E}(t) + 2\chi^{(2)}\mathbf{E}^2(t) + 4\chi^{(3)}\mathbf{E}^3(t) + \cdots\right) \tag{2.2.2a}$$

$$\mathbf{P} = \epsilon_0\left(\chi\mathbf{E}(t) + \mathbf{P}_{NL}(t)\right), \quad \mathbf{P}_{NL}(t) = \left(2\chi^{(2)}\mathbf{E}^2(t) + 4\chi^{(3)}\mathbf{E}^3(t) + \cdots\right) \tag{2.2.2b}$$

It is enough to take into account the first three terms of the series, linear relation, second order and third order effects to describe the mediums response. The rest of the higher order effects have a very small contribution and can usually be ignored.

It becomes necessary to distinguish between media that have second order or third order effects. A centrosymmetric medium presents dominant nonlinear optical processes described by the terms in the polarization density that are a cubic function of the electric field [47]. In noncentrosymetric media, the polarization density exhibits dependencies on both the squared and cube of the electric field. Throughout this project, our main concern is studying fiber optics made out of centrosymmetric media like fused silica SiO_2, such that nonlinear effects arise from third order nonlinear dependencies of the polarization density.

2.2.3 Third Order Nonlinear Kerr effect

Given this projects interest in centrosymmetric media it is necessary to describe the nonlinear response of the polarization density to an ideal monochromatic electric field like the one given in Equation 2.2.1. Lets consider only the effects of the third order response of the dielectric medium $\mathbf{P}_{NL}(t) = 4\chi^{(3)}\mathbf{E}^3(t)$. Substituting Equation 2.2.1 into the nonlinear polarization density gives Equation 2.2.1, which describes the frequency response that arises.

Nonlinear Polarization Density

The nonlinear polarization density $\boldsymbol{P}_{NL}$ response to a monochromatic field can be decomposed into the individual frequency effects $\mathcal{P}_{NL}(\omega_i)$ that emerge from a third-order nonlinear medium [1].

$$\mathbf{P}_{NL}(t) = \mathbf{Re}\left\{\mathcal{P}_{NL}(\omega)e^{iwt}\right\} + \mathbf{Re}\left\{\mathcal{P}_{NL}(3\omega)e^{i3wt}\right\} \tag{2.2.1a}$$

$$\mathcal{P}_{NL}(\omega) = 3\epsilon_0\chi^{(3)}|\mathcal{E}(\omega)|^2\mathcal{E}(\omega) \tag{2.2.1b}$$

$$\mathcal{P}_{NL}(3\omega) = \epsilon_0\chi^{(3)}\mathcal{E}^3(\omega) \tag{2.2.1c}$$

Two different frequency components ω and 3ω emerge from the nonlinear response of the polarization density. The presences of a 3ω response indicates that *third-harmonic generated* (THG) light can be produced in the nonlinear interaction. However, in this cases the energy conversion efficiency is low for THG, producing a relatively weak 3ω signal from the third order polarization density [1].

As has been mentioned before it is assumed that the polarization density response varies only slightly from the linear response $\epsilon_0\chi\mathbf{E}$ with increasing $\mathbf{E}$. It can be assumed that for a centrosymmetric material an incremental change $\Delta\chi$ would be connected to the nonlinear polarization density [1, 29]. With this insight a similar assumption to the linear case can be made for $\mathcal{P}_{NL}$, relating it to the change in the susceptibility $\Delta\chi$ at a frequency ω by

$$\mathcal{P}_{NL}(\omega) = \epsilon_0\Delta\chi\mathcal{E}(\omega). \tag{2.2.2}$$

Expressing the incremental susceptibility as a change in the refractive index would be useful given the ease of measuring the refractive index in general. Since $n^2 = 1 + \chi$, computing the deferential change would give $2n\Delta n = \Delta\chi$. It is also necessary to define the optical intensity of the electric field given by $I = |\mathcal{E}(w)|^2/2\eta$ where η is the impedance of the dielectric medium. With this and Equation 2.2.2 the incremental refractive index can be written as a function of the optical intensity.

$$\frac{\mathcal{P}_{NL}(\omega)}{\epsilon_0\mathcal{E}(\omega)} = 2n\Delta n \rightarrow \Delta n = \frac{3\chi^{(3)}|\mathcal{E}(\omega)|^2}{\epsilon_0 2n} \tag{2.2.3a}$$

$$\Delta n = \frac{3\eta}{\epsilon_0 n}\chi^{(3)}I \equiv \bar{n}_2 I \tag{2.2.3b}$$

Given the proportionality between the refractive index Δn and the intensity I, the overall refractive index has a linear dependency of the optical intensity. This effect is known as the *Optical Kerr Effect* [48].

Optical Kerr Effect

All material have a linear dependency on the optical intensity for high enough power, known as the Optical Kerr Effect. The effect arises from the anharmonic response of bound electrons to optical fields, resulting in a nonlinear susceptibility. The nonlinear coefficient $\bar{n}_2$ is also defined as the optical Kerr coefficient, but is also commonly known as the nonlinear-index coefficient. [29,49]

$$\bar{n}_2 = \frac{3\eta_0}{n^2\epsilon_0}\chi^{(3)} \qquad \text{Optical Kerr Coefficient} \qquad (2.2.4a)$$

$$n(I) = n + \bar{n}_2 I \qquad \text{Optical Kerr Effect} \qquad (2.2.4b)$$

The optical Kerr effect makes a medium sensitive to power fluctuations. It is expected that some sort of distortion will be manifested in the signal due to its effect. In optical communication the Kerr effect induces nonlinear sources of noise that limit the system. To get a better understanding on the nonlinear effects it is necessary to study the properties of an optical fiber.

2.3 Optical Fiber Overview

In this section the phenomenological description for the optical fiber used to construct the communication links are discussed. To investigate nonlinear noise it is necessary to have a physical description of the interaction between the optical signal and the transporting fiber. First a few comments are made on the usual fiber refractive index profile and the definition of some fundamental properties, like the propagation constant of the medium. To describe the nonlinear interaction an extension will be made to the linear polarization density to include higher order terms. A general overview of the physics will be done, since a more detailed presentation is out of the interest of this thesis.

2.3.1 Fiber Properties

Refractive index fiber profile

An optical fiber in a ray optics frame work is just a clever exploitation of the law of refraction and the critical angle. In Figure 2.3.1 a step index fiber profile is visualized, that has a refractive index n_1 in the core and n_2 in the cladding. The light gathering capacity of an optical fiber is known as the *numerical aperture* (NA) that can be approximated for a step index fiber by

$$\text{NA} = n_1(2\Delta)^{\frac{1}{2}}, \quad \Delta = \frac{n_1 - n_2}{n_1} \qquad (2.3.1)$$

where Δ is the *fractional index*. Increasing the fractional index would improve the amount of light gather by the fiber, but it would also induce a large multipath or *modal dispersion*. Figure 2.3.1 shows a set of propagating rays, each traveling a path of different length. As a result, these rays disperse in time, broadening the pulse.

Current optical communication fibers are design to have a $\Delta < 0.01$ allowing for better modal dispersion performance [29].

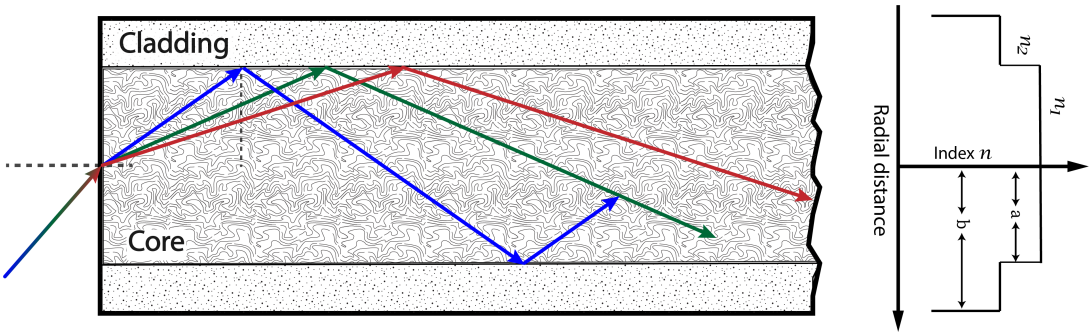


Figure 2.3.1: A step index fiber with its corresponding refractive index profile is shown. Modal dispersion is visualized with a set of propagating rays that enter the fiber at the same time and would exit dispersed in time.

In Figure 2.3.1 the *refractive index profile* of the step index fiber is also shown. The core design plays an important role defining the fibers response to an optical signal. In Section 2.3.3 the refractive index profiles of more interesting fibers will be discussed.

Power loss and Amplification

A critical parameter for an optical system is the attenuation of the fiber. It provides a measurement of the power loss during transmission. For a communication system, having a strong output signal is important to reliably detect the transmitted signal. An optical fiber of length L with an input launch power P_0 has a transmitted power P_T written as

$$P_T = P_0 e^{-\alpha L}, \tag{2.3.2}$$

where α the *attenuation constant* is a measure of all the losses in the fiber [50]. The attenuation constant has units of dB/km and is wavelength dependent, as reported in Reference [51] a fiber with $\Delta = 1.9 \cdot 10^{-3}$ exhibits a fiber-loss of 0.2 dB/km in the wavelength region near 1550 nm. In today's optical communication system the usual operating frequency ω is 193.1 THz because of its low fiber-loss and ease of amplification using *erbium doped fibers*.

Amplification is crucial to counteract the losses in the fiber, extending the reach of the communication system. An ideal amplifier would produce an output power $P_{out} = GP_{in}$ without distortion given an input power P_{in}. In general an amplifier can be described by the *amplification factor* $G(\omega)$ and the *amplification profile* $g(\omega)$ given by $G = e^{gL}$. The amplifiers simulated imparted a minimum amount of noise, having a low *noise figure* of 3 dB. Thus, the dominant sources of noise will arise from the light-matter interaction.

Chromatic dispersion

When electromagnetic waves interacts with a medium's bound electrons, different effects arise through the polarization density, some of which have been covered is

Section 2.2.2. The interaction between the atomic structure and the electromagnetic field is also dependent on the fields frequency ω. This property is known as *chromatic dispersion*. The frequency dependence is manifested through the refractive index $n(\omega)$. Far from the material resonance the refractive index can be approximated by the *Sellmeier equation* [29, 50]

$$n^2(\omega) = 1 + \sum_{j=1}^{m} \frac{B_j \omega_j^2}{\omega_j^2 - \omega^2} \tag{2.3.3}$$

where ω_j is the resonance frequency and B_j is the strength of the jth resonance. In a optical fiber these parameters are found experimentally by fitting the material dispersion curve. Fiber dispersion plays a critical role when building an optical communication system. A pulsed signal frequency components can be found through the Fourier transformation of the signal in the time domain, as is known form signaling theory a pulsed signal contains a range of frequencies. In the same manner, transmitted light pulses are composed of frequencies, in a dispersive medium each frequency would travel at its unique speed $c/n(\omega)$ making the pulse spread in time.

When nonlinear effects are present in the fiber as well, the combination of dispersion and nonlinearity can result in a more complex behavior. This will be discussed in Section 2.3.2. The fiber dispersive effects are dependent on the mediums *propagation constant* β defined as

$$\beta(\omega) = n(\omega)\frac{\omega}{c}. \tag{2.3.4}$$

An approximate response for the complete spectral range can be obtained by doing a Taylor expansion around the pulse center frequency ω_0

$$\beta(\omega) = \beta_0 + \beta_1(\omega - \omega_0) + \frac{1}{2}\beta_2(\omega - \omega_0)^2 + \cdots \tag{2.3.5}$$

where the frequency derivatives of the propagation constant are given by

$$\beta_m = \left(\frac{d^m \beta}{d\omega^m}\right)_{\omega=\omega_0} \quad m = 0, 1, 2..., \tag{2.3.6}$$

and are related to different fundamental properties of the fiber. Due to the pulses spectral composition it propagates with a *group velocity* defined as $\nu_g = (d\beta/d\omega)^{-1}$ [52]. It can be expressed in function of the refractive index as

$$\frac{1}{\nu_g} = \beta_1 = \frac{d}{d\omega} n(\omega)\frac{\omega}{c} = \frac{1}{c}\left(n + \omega\frac{dn}{d\omega}\right) = \frac{n_g}{c} \tag{2.3.7}$$

where the *group refractive index* is $n_g = \left(n + \omega\frac{dn}{d\omega}\right)$. Given the mediums frequency dependence different frequency components undergo different delays. This effect is known as *group velocity dispersion* (GVD) [50], the fibers dispersive effect is described by the *dispersion parameter* D either in terms of wave length D_λ or frequency D_ω differing by a factor of $d\omega/d\lambda = -2\pi c/\lambda^2$.

$$D_\lambda = \frac{d\beta_1}{d\lambda} = -\frac{2\pi c}{\lambda^2}\beta_2 \tag{2.3.8}$$

The units for D_λ are [s/m-Hz] and for D_ω has [ps/nm-km] presenting variation in value given a wavelength. Thus, a pulse spectrum would under go slightly different values of D which would broaden the pulse. This broadening is a higher-order dispersive effects governed by the *dispersion slope* $S = dD/d\lambda$ that can be written in terms of propagation constant β_2 and β_3 as

$$S = \frac{2\pi c}{\lambda^2}\beta_3 + \frac{4\pi c}{\lambda^3}\beta_2. \tag{2.3.9}$$

An optical fibers cladding/core can be designed in such a way that the total dispersion of the fiber can present desirable properties for an communication system. With a clever design, the fibers dispersive response can be flattened to a have a constant value for a spectral range. Another fiber can also be designed to have shifted a dispersion minimum present at frequency with lowest attenuation.

Dispersion in an optical communication link can alter the signals timing, due to the pulse broadening in time different pulses would start to overlap, distorting the signal. Optical compensation can be done using fibers with opposite dispersion parameters to cancel each others effect out. It is more useful to use a highly dispersive fiber to compensate in a short length a large amount of accumulated dispersion. A more detailed description of dispersion compensation is done in Section 2.3.3.

2.3.2 Higher Order Optical Noise

A silica fiber presents a linearly dependent refractive index due to the optical Kerr effect. As is discussed by S. Dris et al. in Reference [53] "*Nonlinear interference noise* (NLIN) due to the Kerr effect presents the most significant bottleneck towards increasing throughput in a flexible WDM optical networks". The objective of this project is to suggest a detection scheme that mitigates NLIN. Through this section a description of the different noise sources of interest will be presented.

First, it is necessary to define quantities that represent the signals resilience. The signals quality is quantified by *optical signal to noise ratio* (OSNR) that is dependent on the signal power P and noise average power N given by

$$\text{OSNR} = 10\log_{10}\left(\frac{P}{N}\right) \tag{2.3.1}$$

in dB units [29]. In a coherent system it is necessary to have a high OSNR at the receiver, usually above 20 dB, for an accurate detection. When noise is present, either linear or nonlinear, the OSNR drops, distorting the signal. As discussed in the introduction, the NLIN distortion is present as a phase noise rotation preserving some information.

In a communication system the quality of the transmission relies on the amount of correctly classified symbols. With a low number of misclassified symbols, correction can be implemented in a higher layer of the communication system to recover or correct bit stream transmitted with FEC [14]. The SER describes the amount of mistakes made in the total transmitted signal as

$$\text{SER} = \frac{\text{no. of incorrect received symbols}}{\text{no. of transmitted symbols}}. \tag{2.3.2}$$

Nonlinear noise in an optical communication system induces a phase shift to the detected signal. If not taken into account, it becomes a fundamental limiting factor to the transmission distance [7]. Having a physical description of the different nonlinear effects, is then useful to alter the expected detection regions to improve the SER.

Three different nonlinear effects will be discussed that fall into two categories, *intra-channel NLIN* or *self-channel interference* and *inter-channel NLIN* or *cross-channel interference*. The intra-channel effects in interest are *self phase modulation* and the *Gordon-Mullenauere effect*. For inter-channel effects *cross phase modulation* will be considered in a WDM system.

Self Phase Modulation

In a fiber absent of dispersion $\beta_2 = 0$ the transmitted pulse undergoes a phase shift due to the intensity dependent refractive index, this modulation is known as *self phase modulation* (SPM). The refractive index intensity dependence induces a power dependence on the propagation constant [29], that can be written as [54]

$$\beta' = \beta + k\bar{n}_2\frac{P}{A_{eff}} = \beta + \gamma P, \quad \gamma = \frac{2\pi\bar{n}_2}{\lambda A_{eff}} \tag{2.3.3}$$

where A_{eff} is the effective area of the guided mode and $k = 2\pi/\lambda$ is the wavenumber. In it self the change to the propagation constant suggest an effect analogues to GVD, where the pulse is chirped and spectral broadening [50]. As is deduced in Reference [54] the total nonlinear phase shift can be computed as the difference between β' and β over the interaction length as

$$\phi_{\text{SPM}} = \int_0^L (\beta' - \beta)dz = \int_0^L \gamma P(z)dz \tag{2.3.4a}$$

$$\phi_{\text{SPM}} < \gamma L_{eff} P_{in}. \tag{2.3.4b}$$

The effective interaction length is defined as $L_{eff} = (1 - e^{-\alpha L})/\alpha$ and the power in the fiber $P(z) = P_{in}e^{-\alpha L}$ [54]. Equation 2.3.4 gives us the amount of phase change over the fiber length L along the propagation axis z. The assumptions made to compute the nonlinear phase shift describes the maximum phase shift possible in a system where a CW driving laser is used with a IM/DD format non-return to zero (NRZ) [29, 42, 54].

It is important to note that when dispersion effects are present the dependencies become more complex and require more sophisticated numerical solutions, making it cumbersome to have a complete model of the system [42, 55]. In multiple references [29, 50, 54] the limit for P_{in} is introduced as a general rule to follow before SPM starts to corrupt the signal. The upper bound is given as $P_{in} < 0.1\alpha/\gamma N_A$, where N_A is the number of amplifiers [54]. For an average fiber system with 10 amplifiers the upper limit is 1.2-3 mW. Making it a critical impermanent for optical system with a high average power. SPM is a NLIN and as concluded in Reference [53] the effects are modulation format independent.

Cross Phase Modulation

In a WDM system the presence of multiple channels in a single fiber span would alter the interaction of the fiber with the transmitted optical signal. If we have an M-channel WDM system with a fixed channel spacing the cross-phase modulation for channel i can be described as

$$\phi^i_{\text{XPM}} = \gamma L_{eff} \left(P^i_{in} + 2 \sum_{j \neq i}^{M} P_j \right) \tag{2.3.5}$$

where we can see that SPM effects are also taken into account for each channel. The XPM effects are important due to the dependence of the channels phase with the bit patterns of all the others [42]. As in the case of SPM a phase shift can be computed to represents the maximum shift acquired in a system

$$\phi^i_{\text{XPM}} < \frac{\gamma}{\alpha}(2M-1)P_{ch} \tag{2.3.6}$$

where P_{ch} is the power per channel. Given the linear relation between the phase shift and the number of channels, there is a limitation to the number of channels or power per channel that can be transmitted.

In a system where GVD effects are present, pattern-dependent phase shifts are converted to power fluctuations, reducing the OSNR at the receiver [42]. A rough estimate for the power restriction is computed in Referene [54] for maximum phase shift $\phi_{XPM} < 1$ rad the upper bound is $P_{ch} < \alpha/\gamma N_A(2M-1)$ before XPM effects arise. In general the power restriction per channel is limited to bellow 1 mW depending on the system, increasing the amount of constraints for the system design. Through this section no mention of the effect of amplification has been mentioned, the interaction between the spontaneous noise produced by amplification and the Kerr effect plays an important role in long-haul transmission links.

Gordon-Mollenauer Effect

In a fiber communication link it is usual to use multiple segments of fiber with optical amplifiers (OA) to extend the systems reach. This type of architecture is known as lump amplification, commonly used in long-haul optical systems with over 10 amplifiers. Noise produced by OA interacts with the fiber through the Kerr effect producing phase fluctuations at the receiver and limiting the transmission distance [42]. This nonlinear source of noise is known as the *Gordon-Mollenauer effect* [55] or simply *nonlinear phase noise* (NLPN).

A transmission line that has N spans of optical fiber followed by its corresponding linear loss amplification is ultimately limited by the presence of *amplified spontaneous emission* (ASE) noise. Due to the refractive index dependence on optical power, fluctuation produced by ASE noise induces a phase shift in the received signal [55]. The nonlinear fluctuation can be approximated as [54]

$$\phi_{\text{NLPN}} \approx N\gamma L_{eff} P_{in}. \tag{2.3.7}$$

Notice that the ASE noise is still a linear source of noise added to the signal as AWGN. However, it also has a second effect by introducing small fluctuations

in the optical power, inducing a phase shift in the transmitted signal due to the Kerr effect. As is explained by P. Gordon and L. Mollenauer in Reference [55] noise produced by the amplifier can be decomposed into a noise field with two degrees of freedom that have the same shape as the pulse. "One in-phase with it, and the other in quadrature. The quadrature component produces an immediate phase shift, and the in-phase one alters the energy of the pulse. The energy change results in an additional nonlinear phase shift by the time the pulse reaches the receiver".

Compound Nonlinear Phase Shift

Three sources of nonlinear interference noise have been presented; self phase modulation, cross phase modulation and nonlinear phase noise. Each of this effects induces a phase shift to the transmitted signal. Represented in a phase diagram as a rotation of the transmitted constellation. Corrupting the detected phase. Each of the NLIN sources has a complex interaction with the dispersive effect, making it necessary to carry out a numerical simulation for a precise model. However, it is not always possible to know the systems parameters before hand, given that in a real system deploying a fiber link requires tolerances.

A maximum phase shift can be expressed for each of the noise source, adding up to a total maximum nonlinear phase shift ϕ_{nl}. A general approximation to quantify the expected phase shift for a signal can be computed as

$$\phi_{\text{nl}} = \phi_{\text{SPM}} + \phi_{\text{XPM}} + \phi_{\text{NLPN}}. \tag{2.3.8}$$

Each effect contributes to the total phase shift induced on the signal. Even though they all arise from the Kerr effect, some can be compensated more efficiently than others. Compensation in the end is what interest us, extracting as much transmission distance or capacity from the system before nonlinear noise corrupts the signal.

Bound Nonlinear Phase Shift

All sources of nonlinear noise; XPM, SPM and NLPN are found to have an upper bound for a fixed system configuration. Thus, an upper bound can be set on the total nonlinear phase noise that is detected at the receiver side.

$$\phi_{nl} < \frac{\gamma}{\alpha}\left(2M + N\right) P_{ch} \tag{2.3.9}$$

Where γ is the nonlinear parameter defined in Equation 2.3.3 and α is the absorption coefficient of the fiber. A reference value can be drawn for a M-channel and N-span link, quantifying the nonlinear phase shift undergone by the signal.

2.3.3 Optical Fiber and Compensation

As presented in the previous section, nonlinear effects are quantified by some fundamental parameters; the nonlinear refractive index, dispersion parameter and the dispersion slope. In Section 2.3.1 there properties were presented. However, it is also necessary to discuss the different types of fiber that can be manufactured and how

they can be used to compensate for fiber dispersion. In this section, the physical description for dispersion compensation is introduced. We conclude by discussing the different fibers used for the simulation and how they where implemented.

Dispersion Compensation in Fibers

All optical fibers manifest dispersive effects, given the refractive index dependence on frequency. In Section 2.3.1 the dispersion coefficient was introduced, concluding that given the different frequency components of a optical pulse, would spread in time due to GVD altering the system's timing [56]. Compensation becomes then necessary to retrieve the original signal. Fortunately, a clever design of the optical core can alter the dispersion parameter D_ω and the dispersion slope S such that a zero dispersion system can be constructed. Consider a link composed of two fibers, a *single mode fiber* (SMF) with a D_1 and a *dispersion compensating fiber* (DCF) with a D_2. The condition for perfect dispersion compensation is

$$D_1(\lambda_n)L_1 + D_2(\lambda_n)L_2 = 0 \tag{2.3.1}$$

where λ_n is the operating wavelength of the nth channel and L_i is the corresponding length of each fiber. It is clear then that the dispersion compensation varies slightly for each channel. Thus, to have perfect compensation it is necessary to also consider the slope of the dispersion coefficient. It would be ideal if the dispersion slope for the DCF could be design to have a value of

$$S_2 = S_1 \frac{D_2}{D_1}. \tag{2.3.2}$$

Unfortunately, manufacturing and fiber design have limits, such that it is not usual to have a perfectly compensated link. When designing a communication system the dispersion scheme is built around the operating wavelength $\lambda = 1550\, nm$ (as has been mentioned before) this wavelength has a low absorption, presents zero dispersion in some fibers and is the current standard in optical communication. Thus, to design a link the fiber is first selected depending on the system's configuration. In practice it is desirable to have a DCF that has a dispersion coefficient and slope that best approximate the values given by Equations 2.3.1 and Equation 2.3.2. A

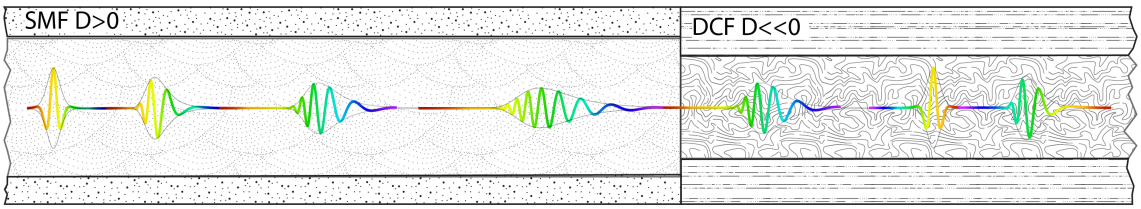


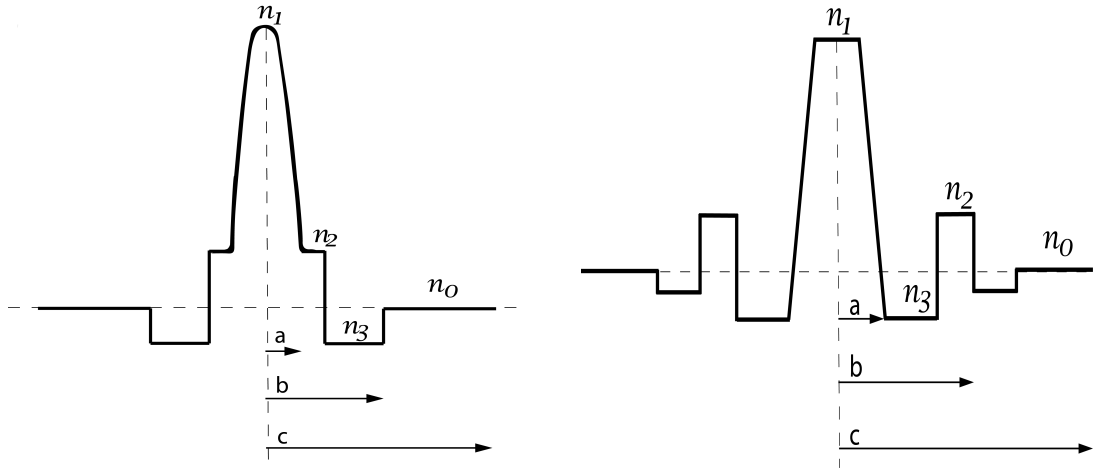
Figure 2.3.1: Visualization of chromatic dispersion of a optical pulse propagating through a SMF. Compensation of pulse was preformed using a highly dispersive DCF.

more accurate visualization of GVD is represented in Figure 2.3.1 where a propagating pulse is spreading in time, compensation is also shown using a short DCF that mitigates the dispersion accumulated in a long SMF. Ignoring losses between fiber

splices notice that the effective core area for a DCF is smaller than SMF. Making compensating fibers have a larger nonlinear refractive index $\bar{n}_2$, thus making NLIN more present in the system.

Manufactured Optical Fibers

The refractive index profile can be designed to alter the optical properties of the fiber. In previous sections the connection between the refractive index and the dispersion coefficient, nonlinear coefficient and other parameters has been discussed. Now we are interested in making a few remarks about the optical fibers that were implemented during the simulation. Each fiber that was used in the simulation followed specifications from a real fibers that have been studied in peer-reviewed articles.



(a) DSF (b) DCF

Figure 2.3.2: Refractive index profile for dispersion shifted or dispersion compensating fibers.

Image designed inspired by available figures in Reference [5, 6]

First, let us consider a single channel system with a operating wavelength of 1550 nm. If only one channel is present then it would be useful that the fiber would have zero dispersion for this wavelength. Multiple fibers exist in the market that present zero dispersion for this wavelength, they are known as *dispersion shifted fiber* (DCF) [5, 57]. To design a optical fiber with a shifted dispersion the refractive index profile has to made in such away that the fibers chromatic dispersion and the waveguide dispersion balance each other out. The general refractive index profile is shown in Figure 2.3.2a, compared to a step index fiber (Figure 2.3.1) four different cladding sections with different refractive index are used, this intricate design achieves modified dispersion effects. However, not all system architectures can use DSF given that when multiple channels are present in the same fiber it is necessary to compensate the dispersion acquired by all channels.

For WDM systems it is usual that all channels present dispersion in the transmission fiber. Instead of implementing a DSF a *nonzero-dispersion shifted fiber*

(NZ-DSF) is deployed that presents a small dispersion value for the center wavelength 1550 *nm*. As was mention in Section 2.3.3, it is not enough to just consider the dispersion parameter, it is necessary to take into account the dispersion slope as well to compensate a multichannel system. To compensate a separate *slope compensating-dispersion compensation fiber* (SC-DCF) or SC-DCF module [6] is used to compensate the acquired dispersion for all channels in a short span of fiber. Figure 2.3.2b shows the refractive index profile for a general DCF, the fiber design is not trivial, but thanks to this clever design, dispersion is not a limiting factor for transmission distance as it once was.

Table 2.3.1: Fiber Characteristics

Physical characteristics for the optical fibers used in the simulations. In the long-haul system the DSF is implemented and a combination of NZ-DSF and SC-DCF is used in the WDM system.

Fiber	$\bar{n}_2$ $[10^{-20}\frac{cm^2}{W}]$	α $[\frac{dB}{km}]$	A_{eff} $[\mu m^2]$	D $[\frac{ps}{nm-km}]$	S $[\frac{ps}{nm^2-km}]$
DSF	2.4	0.21	70	<0.13	0.11
NZ-DSF	2.25	0.2	65	4.2	0.085
SC-DCF	2.8	0.41	16.7	-47	-0.93

All values used for the fiber design in the simulation where taken from References [5, 6, 57, 58]

To simulate the long-haul optical system a DSF with ideal dispersion is considered. Due to our interest in measuring the NLPN after propagating in lump amplification this DSF is a good candidate for what would be deployed today. For the ten channel WDM system two fibers where implemented, a NZ-DSF for transmission and a SC-DCF to compensate for dispersion. This system presents more nonlinear noise given the compensating fibers higher nonlinear coefficient, and both of them have a smaller effective area as well. In Table 2.3.1 the values for all fibers used in the simulation are shown, where all dispersive coefficient values are taken for an operating wavelength of 1550 nm.

CHAPTER

THREE

DIGITAL SIGNAL PROCESSING FOR NONLINEAR PHASE NOISE COMPENSATION

After detecting the transmitted signal, the retrieved information is sent to the *digital signal processing* (DSP) module where the signal can be decoded and electronic noise mitigation can be implemented. In this Chapter two DSP schemes that mitigate nonlinear interference noise will be presented. The first method is a electronic compensation technique that has been analytically derived to perform optimal compensation of nonlinear phase shifts [7,59]. This electronic scheme requires the current state of the system (fiber length, number of amplifiers and detected power) to be able to compensate for the acquired phase shift. Having this information is not always possible in a real system. Thus, a second detection scheme is implemented using machine learning to enable a nonparameter detection scheme that can characterize the nonlinear phase noise as well as other impediments.

3.1 Electronic Compensation of Nonlinear Phase Noise

As was described in Section 2.3.2 the Gordon-Mollenauer effect induces a nonlinear phase shift to the transmitted signal due to the interaction between the ASE noise and the Kerr effect [55]. The analytically derived compensation intendeds to add the opposite amount of phase shift to each IQ parameter. Each detected symbol would acquire a phase shift from the DSP scheme proportional to the optimal compensation coefficient α [59]. This technique can correct accurately the NLPN acquired due to lumped amplification, but it is very sensitive to perturbation. Thus, if any other NLIN is the dominant noise source in the system, the DSP scheme looses its prediction capabilities and will not be able to accurately compensate the constellation.

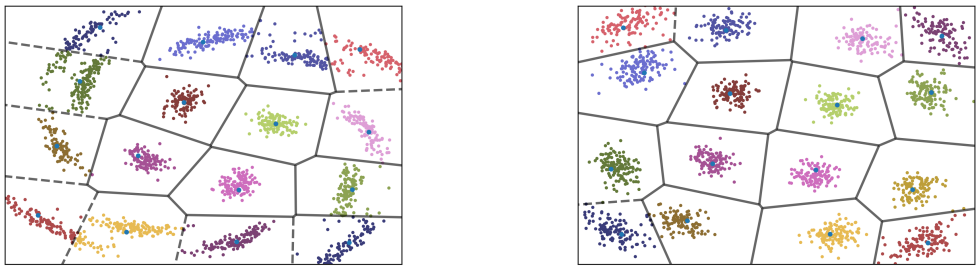
3.1.1 Optimal Nonlinear Noise Compensation

Consider the electric field at the detector side after N spans of lumped amplification as $E_N = E_0 + n_1 + n_2 + \cdots + n_N$, where E_0 is the transmitted signal and n_i is the AWGN of the ith span [55]. The mean nonlinear phase shift acquired is approximately $\langle \phi_{NLPN} \rangle = \gamma N L_{eff} P_N$ as discussed in previous sections [7, 29]. In this system the main impairments are the acquired nonlinear phase shift and the AWGN. Thus, the received electric field can be written as $E_R = E_N e^{-j\phi_{NLPN}}$ containing both sources of noise [7].

To Compensate for the acquired nonlinear phase shift, it is natural to suggest a scheme that rotates the phase as $E_C = E_R e^{-j\alpha P_N}$ where P_N is the output power of the Nth amplifier and α is a *scaling factor* [59]. In Reference [7] the *optimal scaling factor* is analytically deduced to be

. (3.1.1)

The implementation of this NLPN compensating technique on a square 16QAM constellation is shown in Figure. 3.1.1 with decision regions determined using an *expectation maximization* and *Voronoi diagrams*. Notice from this example that the NLPN phase shift present in Figure 3.1.1a causes misclassification of multiple constellation points. It is also important to observe that an imperfect classification is being performed in Figure 3.1.1b, suggesting that when to symbols start to overlap not much is there to do generally.



(a) With NLPN (b) After NLPN compensation

Figure 3.1.1: Square 16 QAM with a nonlinear phase shift of $\langle \phi_{NLPN} \rangle = 0.13$ mrad is shown before and after electronic compensation with optimal scaling factor α, as discussed in Reference [7].

As is reported by K. Ho [7] this compensation scheme can double the transmission distance of a communication system that has NLPN as the dominant noise source. However, it is also acknowledged that other types of nonlinear noise can limit the transmission distance given that only ASE power fluctuations are taken into account. A simple python code can be implemented with this procedure for each detected IQ parameters.

Implemented Algorithm

In this section, a short description is given of the compensation scheme implemented in the simulation to compensate NLPN is discussed. This script is executed by a DSP processing module in the simulation software VPI. The function DSP_ NLPN takes as arguments the IQ parameters detected and request information from the design system through SystemDesign and SystemDesignVolatile. All parameters can be given to the scheme since it has been predefined for the simulation. Each step of the computation is shown in Listing 3.1 where the total output of the function gives the compensated IQ components.

```
def DSP_NLPN(IQ, SystemDesign, SystemDesignVolatile ):
  # Length of one fiber span [m].
  Lspan = SystemDesign[0]
  # Fiber attenuation [dB/m]
  Att_dB_m = SystemDesign[1]
  # Nonlinear coefficient [1/(W*m)]
  Gamm = SystemDesign[2]
  # Received power W
  Ps = SystemDesignVolatile[0]
  # Number of fiber spans
  NSpans = SystemDesignVolatile[1]
  # Fiber loss
  Att = Att_dB_m/4.343
  # Effective fiber length
  if Att > 0:
    Leff = (1-exp(-Att*Lspan))/Att
  else:
    Leff = Lspan
  # Correction factor
  alpha = -Gamm*Leff*(NSpans+1)/2
  # Perform correction
  IQ_C = IQ*exp(-1j * Ps * alpha)
return IQ_C
```

Listing 3.1: Python implementation of an analytically derived compensation scheme to mitigate NLPN. The scheme requires the current state of the system given by user input.

This technique can correctly compensate NLPN present in one channel. A reference can be drawn to this compensation technique to determine the quality of the mitigation predicted by the ML schemes. The next section will cover the ML algorithms used and the implementation suggested to obtain the best decision regions for a constellation with NLIN.

3.2 Implemented Machine Learning Algorithms

The benefit of implementing machine learning to the detection scheme of a communication link is that knowledge of the physical state of the system is not required [18]. This allows the detection scheme to learn about the real impairments of the complete link, computing an accurate decision region for the received constellation. It could

be said that training replaces knowledge. This quality makes it a robust system due to the applicability to different systems configurations without the need of making any previous assumptions.

A drawback of this implementation would be the need to train the scheme when installed or intervened with, requiring a time consuming installation process, if deployed. On the up side, training only has to be done once and it can be performed quickly as well. To initialize the system, a predetermined bit stream would be transmitted as well as being in memory at the receiver side. Since the receiver has the bit stream in memory we get a complete dataset of features and labels. With the dataset a normal training routine can be implemented to any algorithm of interest. Given the natures of the dataset, two classification algorithms *support vector machine* (SVM) and *random forest* (RF) will be used to classify a noisy constellation.

For both of the ML algorithms implemented a simple representation of what training would be can be described as *getting better at a task through practice*. However, uncertainty arises on how to implement an algorithm that exhibits this overall effect. The two principal questions that one encounters are, how does the computer know, if it is getting better, and how does it know how to improve? Several answers have been proposed, each sprung a class of machine learning in it self. Our task falls into a class of algorithms categorized under *Supervised Learning*, or algorithms that can learn from exemplars. A training set of examples is given to the algorithm and based on this it generalizes a response for all possible inputs [39]. A descriptive answer to both of this question will be done for each of the ML algorithms in their corresponding section.

3.2.1 Support Vector Machine

In 1992 V. Vapnik et al. proposed a training algorithm that automatically tunes the capacity of the classification function by maximizing the margin between training examples and class boundary [60]. A critical improvement was done to this algorithm when kernels were implemented to solve the nonlinear separable situation [61]. These two techniques are known now as *support vector machine* (SVM), a ML algorithm that transforms a dataset to a higher space and computes a hyper plain that has the maximum margins separating the plane and the data classes [39].

Analytical Description

Consider an *input space* of $x(i)$ data points containing two different classes that can not be divided by a linear boundary. Applying a transformation like a projection to higher-dimensional space $v(i) = \varphi(x(i))$ projects the data points into to a space known as a *feature space*, where they could be separated. In this feature space, we can define *hyperplain* as all data points $v(i)$ in $\mathbf{v}$ that fulfill $\mathbf{w}^T\mathbf{v} + b = 0$, where $\mathbf{w}$ is a vector perpendicular to the hyper plane. By defining the support vectors as all data point $\mathbf{v}(i)$ that lay on the hypreplane $\mathbf{w}^T\mathbf{v} + b = \pm 1$ the maximum margin b can be computed by maximizing $1/||\mathbf{w}||$ [2, 19, 39]. The kernels implemented where

a Gaussian radial base function, polynomial combination or a hyperbolic tangent mapping, this transformation are a one to one mapping of the given dataset. A visual description of the algorithm is shown in Figure 3.2.1 as a general picture we can keep in mind.

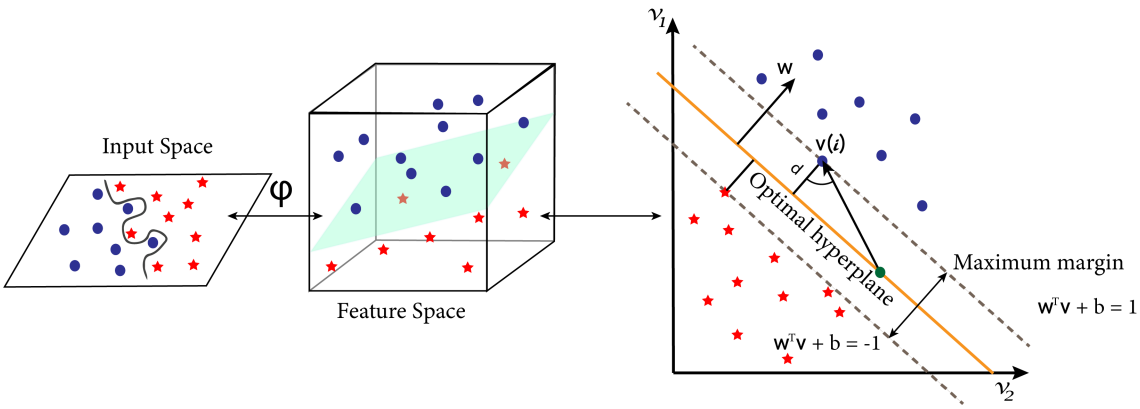


Figure 3.2.1: Visualization of the different steps in SVM to find the optimal nonlinear boundary between a dataset with two classes.

Image designed inspired by available figures in Reference [2]

As a big picture answer the algorithm knows it is getting better if it can decrease the value of $||\mathbf{w}||$, subject that all data points are correctly classified. Or written as an optimization problem as

$$\arg\min_{\mathbf{w},b} \frac{1}{||\mathbf{w}||} \quad \text{subject to } y(l)(\mathbf{w}^T\mathbf{v}(l)+b) \geq 1 \tag{3.2.1}$$

where $y(l)$ is a class label given by ± 1 and $\mathbf{v}(l)$ is the vector of all data points $v(i)$ with label $y(l)$ [60]. This ML algorithm can be trained to compute nonlinear boundaries dividing data classes in a higher-dimensional space. However, to find the optimal solution one must select the correct transformation (or kernel), the weight of each class and the scaling factor b. In the example given a dataset with two classes was described, this was a limitation of SVM in the early 2010. This algorithm was implemented as NLPN compensation scheme in Reference [19] where a 16QAM constellation is processed by four SVM schemes that detect two classes each, the total response of the classifiers would classify all the sixteen constellation points. These algorithms have been now been optimized and improved such that it is capable of determining multiple classes by an extension of the class label $y(l)$, making it faster to implement and compute. Multiple programming languages can implement this classifier in a few lines of code, in this report Python was used to implement all DSP schemes.

Implemented Algorithm

A short description of the python script implemented to computes the optimal parameters for the SVM algorithm is shown in this section. A function DSP_ SVM returns the optimal parameters (kernel, class weight and scaling factor) for the SVM

algorithm given a training dataset. All other functions implemented belong to the *Sklearn* libary in python [62].

```
# Computes optimal parameters for SVM given a set of IQ parameters and the decoded bit sequence(BS)
def DSP_SVM(IQ,BS):
  # Split the IQ parameter and BS into test and train sets
    X_train, X_test, y_train, y_test = train_test_split(IQ, BS, test_size = 0.20)
  # Tuning parameters search grid
  tuned_parameters ={'C': stats.expon(scale=50),
              'gamma': stats.expon(scale=25),
              'kernel': ['linear','rbf', 'polynomial','sigmoid'],
              'class_weight':['balanced', None]}
  # Defines search ledger for SVM  from the different parameter combinations
    Ledger = RandomizedSearchCV(SVC(), tuned_parameters, cv=4)
  # Computes all combinations given by the ledger and returns the best performing parameters
    Ledger.fit(X_train, y_train)
  # Retrieve optimal parameters with the minimum prediction error
    Op_param=Ledger.best_params_
    return Op_param
# Compute optimal parameters for the training set of IQ parameters
Op_param = DSP_SVC(IQ_T,BS_T)
# Define the SVM classifier with the optimal parameters
SVMclass = SVC(kernel=Op_param['kernel'],
              C=Op_param['C'},
              class_weight=Op_param['class_weight'])
# Train the classifier with the training IQ parameters and BS
SVMclass.fit(IQ_T,BS_T)
# Predict the BS of the received IQ parameters
BS_R= SVMclass.predict(IQ_R)
```

Listing 3.2: SVM implementation in python, where the optimal parameter are determined from a training dataset. With the optimal parameters the scheme can classify a new set of IQ parameters and determine the decode bit sequence.

Notice that two different sets of IQ parameters are implemented in this script. To compute the optimal parameters and define the classifier a set of detected IQ parameters for which we know the decode bit sequences is first used as a training set. Given that the IQ parameters are imprinted with the noise characteristics of the system and the receiver knows the decode bit sequence this dataset can be used as an input space. After defining and training our SVM classifier it can be implemented to predict on detected symbols transmitted through the same link.

3.2.2 Random Forest

A *random forest* (RF) or also known as *random decision trees* is an ML algorithm that fits a number of *decision tree classifiers* on different sub-samples of the dataset using averaging over all trees to improve the prediction accuracy and control over-fitting. This algorithm has been developed by incremental incorporation of well understood techniques to optimize its performance and prediction capabilities. By using decision trees, bootstrapping and boosting, three well established techniques,

random forest has been developed to be straight-forward to implement and it can be trained extremely fast [63]. A brief description of the implementation used to characterize NLIN is discussed in this section.

Analytical Description

Let us first discusses the principal element of this ML algorithm, which is a decision tree. Multiple methods exist to grow a decision tree, the method used in this scheme is known as *oblique decision trees* [63]. The first step to grow a tree is generating an initial hyperplane that is not necessarily tailored for the dataset of interest. This d-dimensional hyperplane can be expressed as $H = (h_1, h_2, \cdots, h_d)$. From this hyperplane the algorithm randomly picks an h_i and adds a uniform random variable creating a new hyperplane [64]. For both of these hyperplanes the *sum-minority energy measure* is computed as defined in Reference [64] Theorem 2.1. If the energy difference ΔE between the two hyperplanes is negative, then the energy of the perturbed hyperplane has decreased and it becomes the current split. If the change in energy increases or stay the same then the perturbed hyperplane has a probability of becoming the new hyperplane or being discarded. To determine when to stop perturbing the current split the energy of all perturbations (around 3000 iterations) is kept to select the split that generated the lowest energy [64]. Although generating these trees can be computed very fast a limiting factor is over-fitting.

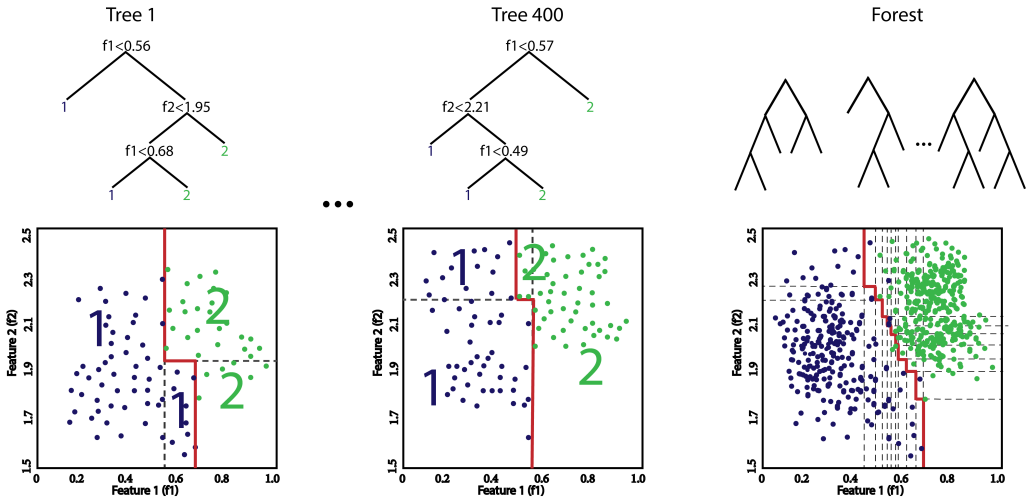


Figure 3.2.1: Two different decision trees from a forest are shown with their computed hyperplane that divides the sub-sampled dataset. When all trees are averaged together a forest is crated giving a decision boundary the is not over-fitted to the original dataset.

Image designed inspired by available figures in Reference [8]

To counter the effects of over-fitting a dataset, two different techniques are implemented that improve its generalization capabilities. The first technique is *bootstraping*, which is a sub-sampling dataset technique to increase the training data [65]. Bootstraping creates datasets with the same statistical properties as the original dataset by random sampling with replacement. The second technique is known as

boosting, where you ensemble a large amount of weak learners to produce a strong learner [66]. These three methods compose the basis of a random forest ML algorithm, each of this techniques effect is depicted in Figure 3.2.1. Implementing random forest can be easily done using multiple programming languages that offer optimized libraries like Python with Sklearn [62].

Implemented Algorithm

Implementing a nonparameter NLIN mitigating algorithm requires training the algorithm in the first instance and then a datasets from the same system can be implemented. Tress are also restrained of growth by delimiting the number of data points (or features), depths of the tree, number of splits and other parameters as shown in Listing 3.3. The function DSP_ RandForest takes a set of IQ parameters and its corresponding bit sequence and searches for the optimal parameters to classify this dataset. With the optimal parameters, the algorithm can be trained to obtain the optimal decision boundaries to predict the bit sequence of a detected symbol.

```
# Optimal parameters for random forest given a set symbol IQ
   parameters and the decoded bit sequence BS
def DSP_RandForest(IQ,BS):
# Split the IQ parameter and BS into test and train sets
  X_train, X_test, y_train, y_test = train_test_split(IQ, BS,
   test_size = 0.2)
# Number of trees in random forest
  n_estimators = [int(x) for x in np.linspace(start = 200, stop =
   500, num = 3)]
# Number of features to consider at every split
  max_features = ['auto', 'sqrt']
# Maximum number of levels in tree
  max_depth = [int(x) for x in np.linspace(5, 20, num = 3)]
  max_depth.append(None)
# Minimum number of samples required to split a node
  min_samples_split = [2, 5, 10]
# Minimum number of samples required at each leaf node
  min_samples_leaf = [1,2]
# Create the random grid
  random_grid = {'n_estimators': n_estimators,
                 'max_features': max_features,
                 'max_depth': max_depth,
                 'min_samples_split': min_samples_split,
                 'min_samples_leaf': min_samples_leaf}
# Defines search grid of input parameters for the RF
    Grid = GridSearchCV(RandomForestClassifier(), random_grid, cv =
    5)
# perform best parameter search using the train set
    Grid.fit(X_train, y_train)
# Retrieve optimal parameters with the minimum prediction error
    Op_param=Grid.best_params_
return Op_param
# Compute the optimal parameters
Op_param= DSP_RandForest(IQ_T,BS_T)
# Define the RF structure and tree construction parameters
  Forest = RandomForestClassifier(max_depth=Op_param['max_depth'],
```

```
                    max_features=Op_param['max_features'],
                    min_samples_leaf=Op_param['min_samples_leaf'],
                    min_samples_split=Op_param['min_samples_split'
    ],
                    n_estimators=Op_param['n_estimators'])
# Train the RF classifier with the training IQ parameters and BS
     Forest.fit(IQ_T,BS_T)
# Predict the BS of the received IQ parameters
       BS_R=Forest.predict(IQ_R)
```

Listing 3.3: Random forest implementation in python where the optimal parameter are determined from a training dataset. With the optimal parameters the scheme can classify a new set of IQ parameters and determine the decode bit sequence.

3.3 Nonparameter Detection Scheme

In the previous sections two different ML algorithms were presented. In this section, both algorithms are implemented to obtain a more powerful and reliable algorithm. Random forest can be operationally very fast but it can present over-fitting which reduces its predictive capability. Support vector machine does not over-fit the training dataset, but it requires a longer time to find the optimal parameters and train the algorithm to perform accurate predictions. Both algorithms have different advantages and disadvantages, making a joint implementation an interesting idea to balance there short comings.

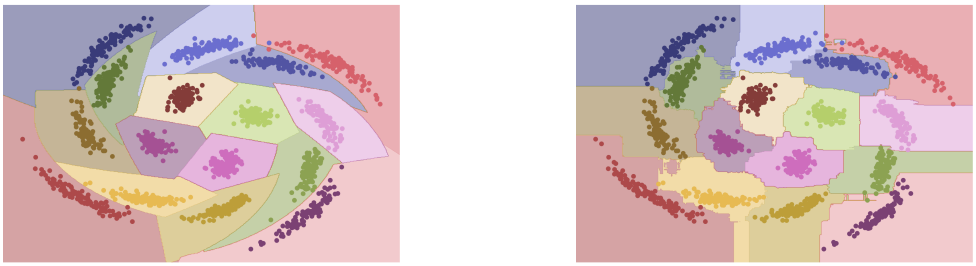
In this thesis a DSP scheme that computes decision regions that characterize the nonlinear noise present in a 16 QAM coherent optical link is suggested. Our intent for this scheme is improving the amount of correctly classified symbols and extending the maximum transmission length. The algorithm computes first the random forest classifier because of its fast execution time. If the SER of the training dataset is greater than zero, support vector machine is also executed to check if a better SER could be achieved. With both ML algorithms the transmission length of the system can be improved for the same transmission link. The structure of the script implemented is shown in Listing 3.4, the execution time of the complete script can take between 2 s to 45 s to compute the optimal decision regions, which is a relatively fast implementation time for the system. Also taking into account that the link in principle could work forever afterwards.

```
def DSP_ML(IQ_T, BS_T):
# Compute optimal parameters for RF given the training set of IQ
    parameters
  Op_param_RF = DSP_RandForest(IQ_T,BS_T)
# Define the RF structure and optimal tree construction parameters
  Forest = RandomForestClassifier(max_depth=Op_param['max_depth'],
                    max_features=Op_param['max_features'],
                    min_samples_leaf=Op_param['min_samples_leaf'],
                    min_samples_split=Op_param['min_samples_split'
    ],
                    n_estimators=Op_param['n_estimators'])
```

```
# Trains the RF classifier given the training IQ parameters and BS
  Forest.fit(IQ_T,BS_T)
#  Predict the BS using RF for the training IQ parameters
  BS_RF = Forest.predict(IQ_T)
# Check the prediction capacity of the RF classifier
  if( SER(BS_RF,BS_T )> 0 ):
  # Compute optimal parameters for SVM given the training set of IQ
     parameters and BS
    Op_param_SVM = DSP_SVM(IQ_T,BS_T)
  # Define the SVM classifier with the optimal parameters
    SVMclass = SVC(kernel = Op_param['kernel'],
              C=Op_param['C'},
              class_weight=Op_param['class_weight'])
  # Trains the SVM classifier with the training IQ parameters and
    BS
    SVMclass.fit(IQ_T,BS_T)
  # Predict the BS using SVM for the received IQ parameters
    BS_SVM = SVMclass.predict(IQ_T)
# If the RF classifier performs no mistake it is returned as the
   optimal classifier
  else:
    return 'RF', Op_param_RF
# The prediction capability for both algorithms is compared
  if(SER(BS_RF,BS_T) > SER(BS_SVM,BS_T)):
  # If SVM performs less mistakes predicting the BS then it is
   returned as the optimal classifier
    return 'SVM', Op_param_SVM
# If RF performs less mistakes predicting the BS then it is
   returned as the optimal classifier
  else:
    return 'RF', Op_param_RF
```

Listing 3.4: Machine learning enabled DSP scheme that returns the best performing ML algorithm and its optimal parameter for the training dataset.

To get a better understanding of the difference between the decision regions determined in Figure 3.1.1 and the predicted regions computed by the ML algorithm, the RF and SVM predicted regions are shown in Figure 3.3.1. Notice how RF (Figure 3.3.1b) tends to overfit the training dataset increasing its SER. Overfitting is observed when the decision region for a constellation point is tightly bound around the constellation point, the system can not produce new predictions, it learns to much structure from the dataset. However, SVM for the same dataset finds optimal regions that do not over lap and extending its prediction capability's, predicting accurately most symbols. From Figure 3.3.1 one can observe the difference on how both algorithms compute the regions by looking at the smoothness of the regions. For this dataset SVM has a lower SER and would be the optimal DSP scheme.



(a) SVM

(b) Random forest

Figure 3.3.1: A constellation diagram for a square 16 QAM with a nonlinear phase shift of $\langle \phi_{NLPN} \rangle = 0.13$ mrad is shown. Two different schemes where used to compute the decision boundaries of this constellation, SVM and random forest.

Two different communication links were simulated to determine the performance of this DSP scheme. In the next chapter the simulation of the different fiber links will be discussed, as well as the measurements retrieved for the analytical compensation scheme (discussed in Section 3.1) and the proposed ML algorithm. In the last chapter the conclusions, improvements and further work will be given.

CHAPTER

FOUR

COMMUNICATION SYSTEM SIMULATION

In the previous chapters the components of an optical communication fiber system has been described, taking special attention to the noise that arises from the nonlinear interaction between the optical fiber and the propagating light. The intent with the simulations carried out was to obtain a measurement of the noise sources described in Section 2.3.2. To observe NLPN it is necessary to have multiple amplification segments. Thus, a long-haul fiber link is an ideal scenario, given that it has more than ten fiber amplification segments where NLPN would be imprinted into the received constellation. The two other nonlinear noise sources are SPM and XPM which are present in systems with multiple channels being transmitted over the same fiber. To measure this effect a WDM scheme is also implemented to detect the interchannel interference imparted to the transmitted constellation. Both of these systems are simulated using *VPItransmissionMaker*^{TM} [9] also known as VPI. This software allows accurate simulations of a multitude of photonic systems including coherent optical links as a plug and play software.

With the simulated links a bit sequence will be transmitted over the system, at the receiver side the detected constellation point will classified and decoded using different DSP schemes. The DSP algorithms that will be implemented have been discussed in Chapter 3. For each fiber link we will compare the proposed nonparameter nonlinear noise characterizing DSP with the DSP computed by VPI automatically. Giving us a point of reference to determine, if there is an improvement in transmission length when using a more robust DSP scheme.

4.1 VPI Detection scheme

Each communication link was intended to give insight into the relationship between a physical variable or system architecture and the SER of the received signal. The components used to build this communication link have been described in Chapter 2. In the present section some additional comments on the simulation parameters and processing scheme will be done. For all architectures the detection and modulation were performed with the same modules in VPI, to measure the difference between using different fiber lengths, launch power and amount of channels transmitted. When the system are simulated, VPI performs a numerical simulation propagation

of the transmitted symbol through the specified link. In the receiver side VPI automatically performs some signal processing to retrieve the IQ parameters of the received symbol. This signal processing is performed automatically by the used VPI module.

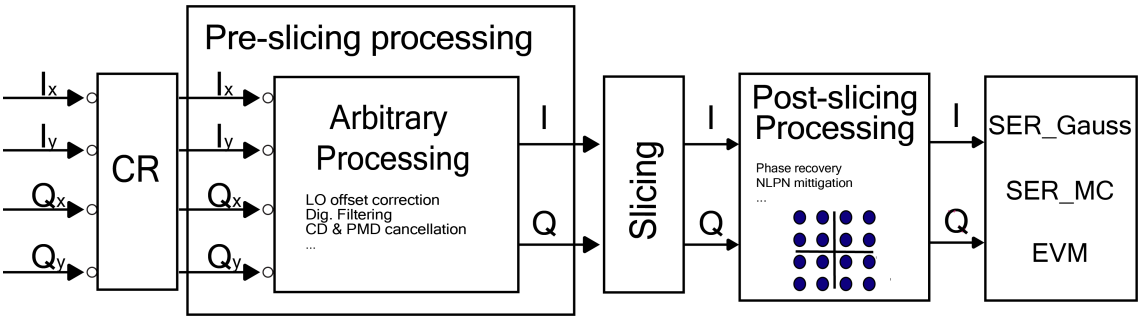


Figure 4.1.1: Block diagram of the signal processing performed by VPI after detecting the IQ parameters of a transmitted symbol.

Image designed inspired by available figures in Reference [9].

Throughout all simulated communication systems, the detection processing scheme follows the same block architecture. After retrieving the IQ parameters with the detection scheme discussed in Section 2.1.2 the VPI module process the symbols as shown in Figure 4.1.1 for both polarization x and y. Each block accepts user inputs to alter its behavior, like selecting different compensation schemes available in VPI. The signal processing scheme implemented by VPI receives the transmitted *IQ* complex samples. With the samples it performs *clock recovery* in order to retrieve the timing information from the detected signal. Selecting one fixed time sample to perform metric calculation is called "slicing", which has the same meaning as "sampling" in conventional BER estimator. Generally, the processing scheme is divided into Pre- and Post-Slicing DSP procedures. As is shown in the block diagram custom DSP procedures can be applied before and after slicing. Impairments in the signal like CD and CFO are time dynamic effects, thus they require being proceed in pre-slicing to have the timing information for each symbol.

Retrieving the IQ phase components is done in post-slicing because we are not considering time dependent phenomena. With the IQ components extracted from VPI the DSP is done in Python as was described in Chapter 3. For each communication link the DSP implemented is used to classify the detected symbol, ultimately lowering the SER of the link. The receiver has a run time simulation of 0.128 μs at symbol rate of 32 $Gbaud/s$ transmitting a total of $2^{12} = 4096$ symbols per frame. This is the total length of the simulation carried out for each frame or a single transmission length (fiber length times fiber segments) of the simulation.

4.2 Coherent Communication System Implementation

High capacity, low latency and other technological advantages like transparent amplifiers and routers have made optical fiber the most implemented technology for

long distance communication. It is also being currently deployed in the access network as well. In this section, a 32 Gbaud 16QAM transmitter was simulated in different architectures to observe the effects of NLPN, SPM and XPM. Each of the communication systems simulated and the data collected for different configuration will be discussed as well.

4.2.1 Gordon-Mollenauer Effect & SPM in a Long-Haul Fiber Link

A system with a single frequency channel, transmitted over lumped amplification is ultimately limited by the NLPN. To measure the effects of this nonlinear noises sources the fiber link shown in Figure 4.2.1 was simulated. This system can be implemented to sweep over launch power, fiber length and amount of fiber spans. With this scheme we can observe the effects of NLPN and SPM on the transmitted constellation and perform DSP to optimize the decision regions to classify each symbol.

Architecture

A long-haul fiber system is composed of multiple segments of fibers spans with their corresponding amplification one after the other. To simulate a long fiber, a loop block is used to propagate the signal over 80 km fiber segments as many times as desired. A set of 64 spans were simulated making the maximum fiber length 5120 km. After each loop the transmitted signal can be detected without altering the system.

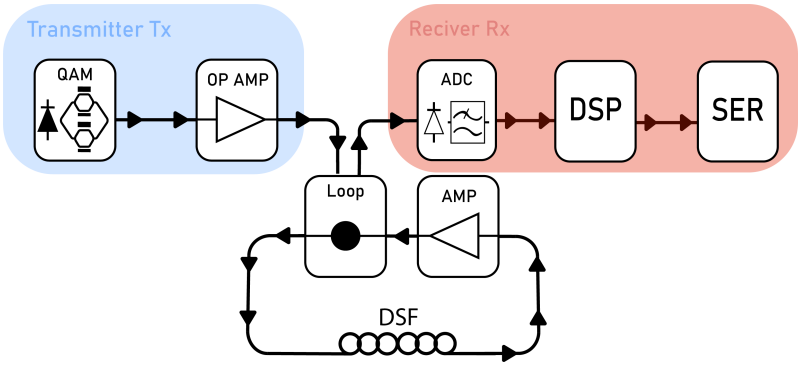


Figure 4.2.1: Architecture of a single channel fiber link. The transmitter is composed of an optical carrier and modulator, that permit independently setting the launch power. The propagating fiber implemented is a DSF with inline amplification to compensate for fiber losses, the amplifiers simulated add small amounts of AWGN. Isolating the noise effects that arise from NLIN. At the receiver side a ADC retrieves the digital IQ parameters and send them to the DSP which improves the signal quality that is quantified by the SER.

Nonlinear effects are prevalent in this system due to the fibers relatively large length and nonlinear coefficient. With this system, we intend to get an understanding on the effects of applying DSP on the received signal. It would be advantageous

if we could perform some digital processing and improve the quality of the signal, resulting in a system that can be transmitted over a longer distance or has a more reliable connection. The detected signal is imprinted with a phase rotation from NLPN and SPM. However, SPM is very weak compered to NLPN, so we will focus primarily in NLPN. As a reference for 30 spans of DSF fiber upper bound channel power (as discussed in Section 2.3.2) $P_{ch} \leq 0.4\ mW$ acquires a phase shift $\phi_{SPM} < 1.3$ mrad, compared to aa NLPN phase shift of $\phi_{NLPN} < 92.6$ mrad. Thus, for this simulation even though SPM is present it is not the dominant noise source in the system.

Transmission Measurement for a Single Channel 128 Gbit/s System

As was discussed in Section 2.3.2 nonlinear effects depend on the fibers length, the amount of fiber segments and the transmission power. A fixed fiber length of 80 km has been chosen for convenience, thus determining the optimal launch power is necessary. An optimal lunch power would be such that it presents high enough OSNR for re-transmission, ultimately increasing the systems reach by improving the signal quality. To decode a HOM format it is necessary to have a clear signal with a high OSNR. This can be achieved by operating at high launch powers. However, if the launch power is to intense the OSNR will decreased do to the nonlinear noise added to the signal. Balancing this trade off is important for the system, given that at the proper operating power a more reliable and further reaching system can be achieved. A power sweep was performed to determine the optimal power and observe the effect on the transmission distance and DSP schemes.

To determine the effect of nonlinear noise on the system a low launch power was first investigated. Even thought the launch power is relatively low $P_{ch} \leq 0.5$ mW the $OSNR \geq 25$ dB can be reliably detected. Given that in this simulation we isolate the NLIN by simulating low noise figure amplifiers adding a small amount AWGN. A *reliable detection* is assumed to be a system that has a SER$> 10^{-5}$, which is the same as a communication link that performs one mistake every hundred thousand symbols. The phase shift acquired in this datasets is $\phi_{SPM} \leq 3.8$ mrad and $\phi_{NLPN} \leq 115.8$ mrad, making it a good region to detect primarily NLPN.

From the simulation we can retrieve the IQ parameters with all the NLIN acquired from transmission, represented by the color orange in all the plots in this chapter. From this dataset all DSP scheme are computed, represented in this section by the color blue for the analytically derived scheme and green for the nonparamter ML algorithm proposed in this report. The vertical bars in all plots in this chapter represent the fiber span at which the system (orange) or the DSP schemes (blue, green) stopped being reliable. The dashed vertical bar for the ML implementation is directly computed from the NLIN dataset, while the dashed-dotted line is determined by an exponential function fitted to the ML data set. This fitting is performed due to an abrupt change in the green scattered points. This effect will be further discussed in the next section. The fitting determines an approximate maximum reliable transmission distance for SER with exponential change, as is in principle anything nonlinear is exponential. However, the change in this ML SER function is faster than a simple exponential fitting. Meaning that the maximum reliable transmission distance could fall in between the two green vertical lines, this

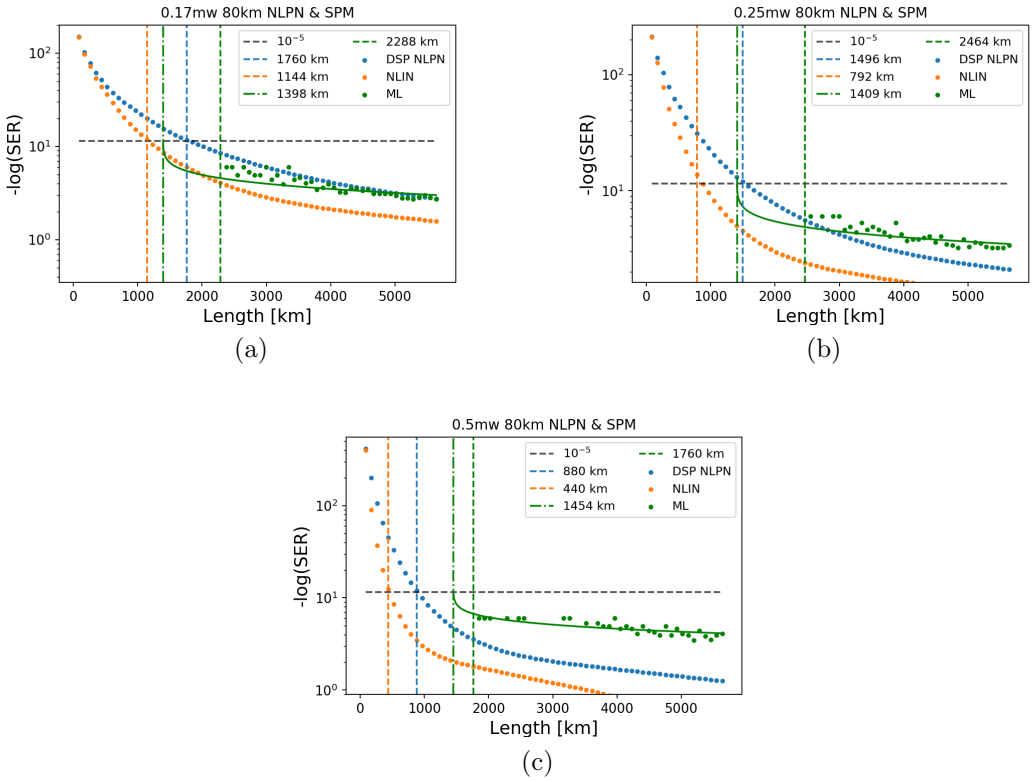


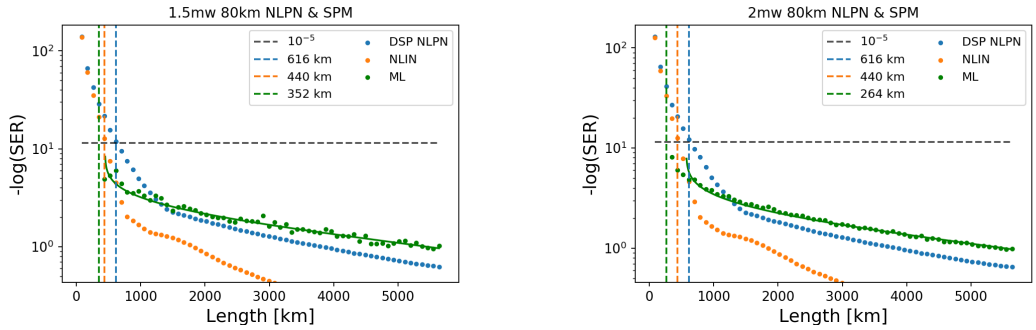
Figure 4.2.2: Single Channel 16QAM long-haul link operating at a low launch power. In these plots a reliable measurement is any SER that falls above the horizontal dotted line, which represents theoretical SER= 10^{-5}. When one of the DSP scheme crosses this line, the processing scheme is not reliable. At this operating power the system present a small amount of nonlinear noise. At higher operating power the ML scheme performs better, as seen by the shift to longer distances of the green vertical lines.

region will be referred to as the *optimal prediction region* (OPR).

From Figure 4.2.2 we can determine that there seems to be a tendency for the ML algorithm to perform better for higher launch power. This is determined by the shift to longer transmission length of the green vertical bars. We can notice that at a operating power of 0.5 mW (Figure 4.2.2a) the OPR at a minimum of 1454 km starts to outperform the analytical DSP scheme.

Given this observation it could be considered that ranking up the launch power would produce a further reaching system. However, it can be seen in Figure 4.2.3 that all DSP schemes quickly start to fail with the increase in NLIN. For a 2 mW launch power, after propagating through 560 km of fiber the signal acquires a phase shift $\phi_{SPM} \leq 15.4$ mrad and $\phi_{NLPN} \leq 108.1$ mrad. Even thought this might seem like is a small phase shift the signal is mixed much faster than for lower powers, making this acquired phase shift to corrupt the signal faster.

Through the power sweep it was determined that the optimal launch power was 1 mW (Figure 4.2.4) due to the overall shift to longer distance in the OPR of the system. The acquired NLIN phase shift at 2288 km is $\phi_{SPM} \leq 7.7$ mrad and



(a) (b)

Figure 4.2.3: Single Channel 16QAM long-haul link operating at a high launch power. At this operating power the system present a large amount of NLIN, critically reducing the performance of the system.

$\phi_{NLPN} \leq 200.7$ mrad. Notice that even thought the acquired phase shift is larger than the one computed for a 2 mW the performance of the system is better.

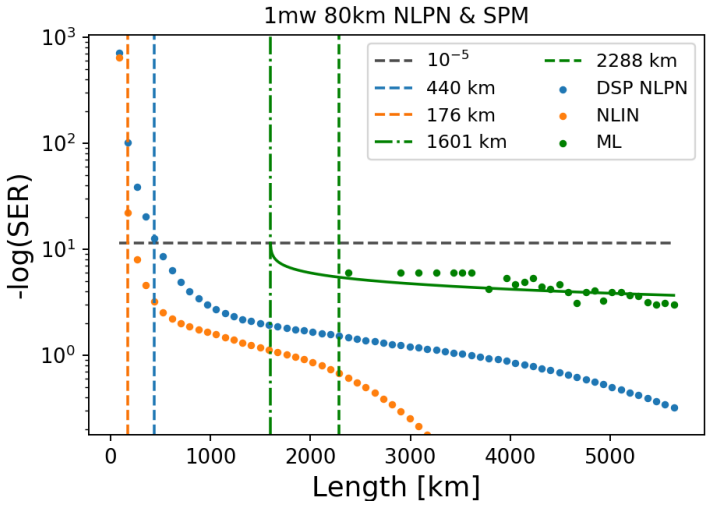


Figure 4.2.4: Single Channel 16QAM long-haul link operating at the optimal launch power of 1 mW. At this operating power the system present the optimal balance between optical power and induce NLIN phase shift. The signal preserves its information allowing it to be accurately decoded at further distances even thought it has acquired a maximum phase shift $\phi_{nl} \leq 239.6$ mrad at a fiber length of 2400 km.

Measurement Discussion

A few remarks should be made about the measured launch powers and their resulting SER. The lowest launch power simulated was 0.17 mW (Figure 4.2.2a). For this system the NLPN is approximately 30 times larger than the SPM noise. Given that this system presents mainly NLPN the analytically derived DSP scheme can compensate most of this noise. With this in mind we intend to compare the performance of the

ML scheme represented by the green scattered dots. For long transmission lengths it can be observed in Figure 4.2.2a that the ML scheme tends to perform equally well as the analytical DSP scheme. Suggesting that the ML scheme can effectively decrease the SER of the system, but it cannot outperform the compensation achieve by the analytical DSP scheme.

In all plots the ML scheme (green scattered dots) abruptly starts making mistakes after a certain fiber length. This abrupt change from the SER is due to a limitation of the simulation performed. A total of 4096 symbols are being simulated per frame, if one mistake is made in a dataset the SER=$24 \cdot 10^{-5}$ automatically failing the reliable detection condition. However, this only means that the fewer mistakes are made the more samples are needed. Simulating for longer times could be a solution to this limitation but it is not practical for the scope of this report, given that increasing the simulation time would increase the computation time for all the project beyond what is available at my disposal. Due to this limitation of the simulation an exponential function was fitted to the SER predicted by the ML algorithm, represented by the solid green plot.

For the simulation performed the maximum transmission length would lay the region between the zero SER predicted ML algorithm (vertical green dashed line), and the length at which the fitted exponential function (vertical green dash-dotted line) intersect the optimal prediction boundary (black horizontal line). From sweeping the power it was determined that the optimal launch power was 1mW as seen in Figure 4.2.4. Due to the shift of the OPR to longer transmission lengths it can be determined that at the very list the ML processing scheme performs comparable to the analytical DSP. Performing as well or better than the analytical DSP scheme is a interesting result, because it suggest that a DSP scheme with no information about the system can resolve NLIN comparable to a scheme that has all the information about the links architecture.

Increasing the launch power results in critically hindering the communication link, ultimately corrupting the prediction of all the DSP schemes and reducing the transmission length. In Figure 4.2.3 it can be observed that all vertical bars or optimal prediction lengths quickly recline to lower transmission lengths. The main take away from this simulation is that the proposed ML scheme at the optimal launch power could increase the transmission length anywhere from 2.5 times (same performance as the analytical DSP) to 9.4 time longer (length predicted by exponential fit), which could motivate a further investigation in this area.

4.2.2 SPM and XPM in a WDM system

Increasing the transmission capacity can be achieved by adding more channels to the system, either by *polarization division multiplexing* or *wavelength division multiplexing*. In this system both of this techniques are implemented to increase the transmission capacity and explore interchannel noise. Every transmitter has two orthogonal polarization channels, a total of five transmitters in the system are used giving a 1.28 Tbit/s link. This system can be used to measure the effect of SPM and XPM, because transmitting multiple channels over a single fiber naturally increase the optical power perceived in the fiber inducing nonlinear noise.

The fiber system implemented induces dispersion at different rates for each communication channel. Transmission is done over NZ-DSF and the dispersion acquired is mitigated for all channels with a SC-DCF fiber. This fiber arrangement presents in total a higher nonlinear coefficient and a smaller effective areas, enhancing nonlinear noise. The high power given by all channels together hinders the transmission length that can be achieved by increasing the NLIN present in the system.

Architecture

An advantage of the modules used in VPI are their flexibility to implement a dual polarization transmitter and alter the operating frequency. The channel spacing between optical carriers was 50 GHz making this system a *dense wavelength division multiplexing* architecture. The propagating fiber implemented was a 80 km NZ-DSF that induces dispersive effect on all channels of the system. Dispersion compensation is done with a 7.14 km SC-DCF that can perform close to an ideal compensation. The maximum transmission length was 10 spans of 87 km fiber, equal to 870 km. It was assumed that all fibers were polarization mode preserving for simplicity. The fiber loss compensation is done with amplifiers with a low noise figure, such that the dominant source of noise is NLIN that is present in the WDM system.

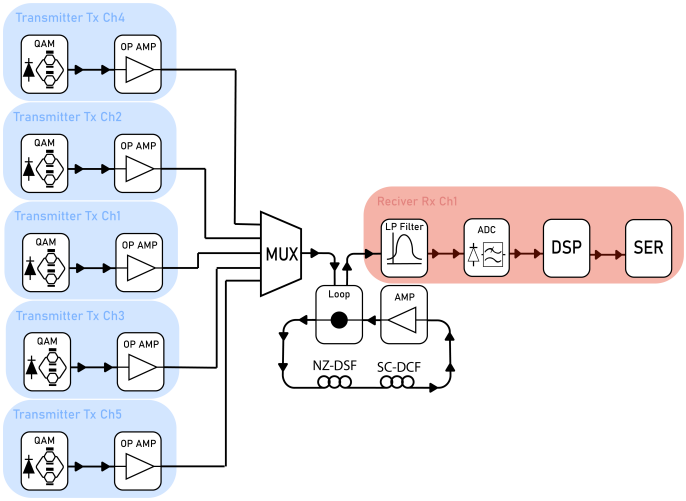


Figure 4.2.1: Architecture of a 5 channel WDM dual polarization fiber link. The transmitter is composed of a dual polarization optical carrier and modulator. The launch power can be independently adjusted as desired. The fiber segment used produce a total signal that presents a small amount of dispersion, that does not critically hinder the system. At the receiver side an ADC retrieves the digital IQ parameters and sends them to the DSP that tries to improve the SER.

All blocks used for the simulation in VPI are the same as in the previous link. Thus once again the amplification performed for each fiber segment is almost ideal, adding a small amount of AWGN to the signal. This is done to isolate the NLIN that is in the system and focus on the nonlinear effects present in the link. In this architecture XPM noise is present for any setup that has a channel power $P_{ch} > 0.07mW$ which would be any relatively realistic implementation.

Transmission Measurement for a $5 \times 256Gbit/s$

The measurements performed were done in the same manner as for the previous long-haul communication channel. To find the optimal launch power a sweep was performed and the SER was measured. In Figure 4.2.2 the SER for the signal with all the NLIN is represented by the blue scattered points. Once again the ML scheme is represented by the green scatter points, with its corresponding exponential fit as the green solid curve.

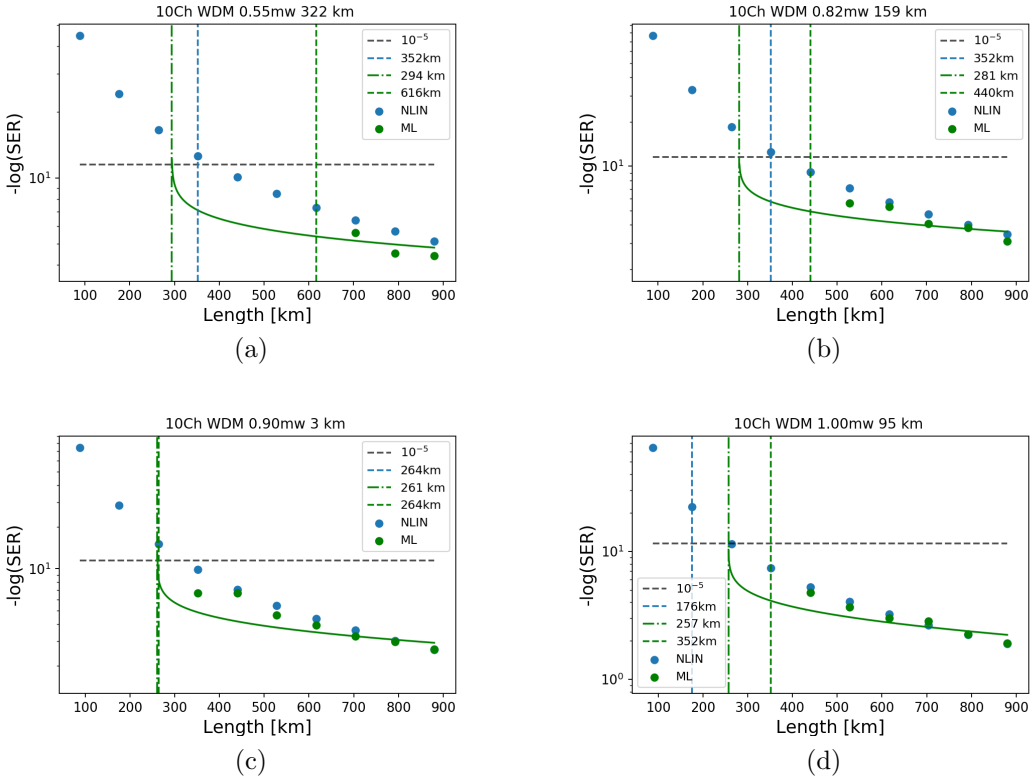


Figure 4.2.2: Ten channel WDM system power sweep measurements. For this system architecture it cannot be determined if the ML algorithm has any effect on the systems performance. In general the ML DSP scheme has a comparable transmission length with the detection scheme implemented by VPI.

From the WDM simulation performed it cannot be determined, if there is a benefit in implementing ML. This result can suggest that XPM is a time dynamic effect which cannot be resolved with the DSP that is being performed.

Measurement Discussion

Extending the transmission length using ML was the overall objective for this simulation. However, the data suggest that a considerable difference cannot be achieved with the DSP scheme implemented. In Figures 4.2.2a a low launch power is tested and it can be determined that some improvement is achieve by the ML scheme. This is determined due to the OPR extending further than the optimal transmission length determined by VPI for the signal with all the NLIN. But since the

optimal transmission length determined by VPI lays inside the OPR it is not clear if there is a benefit to implementing ML.

When the launch power is further increased Figure 4.2.2c-4.2.2d it seems like there is a improvement in implementing ML over VPI. The limitation is that the transmission length is shorter than for previous launch powers, making it a sub-optimal operating power. What this results suggest is that XPM is a noise source dynamic in time and cannot be compensated for in post-slicing processing as mentioned in Section 4.1. What can be said about this simulation is that XPM cannot be compensated with this ML DSP scheme, but it could be used with other technologies to improve the general reliability of the system.

CHAPTER

FIVE

CONCLUSION

To conclude this report, a few closing remarks will be done to give a general scope of the project, its limitations and possible future work that could be done to extend the results found in this thesis. First, let us remember the aim, which was "to study the effects of nonlinear noise in a coherent optical communication system and suggest a *nonparameter digital signal processing detection scheme* enabled by machine learning". The nonparameter scheme suggested used *support vector machine* and *random forest* as a statistical tool to classify symbols detected in a coherent optical communication link. Nonlinear effects were simulated for two different coherent optical communication architectures, a long-haul link and a *wave length division multiplexing* system.

For the long-haul fiber link simulated, the optimal launch power was determined to be 1 mW. It was concluded that for the optimal launch power, the nonparameter digital processing scheme could preform nonlinear phase noise compensation comparable to an analitycal scheme that has all the information of the links architecture. It was determined that the nonparameter scheme could potentially improve the transmission length between 2 to 9 times further. The main observation for this simulation was that for nonlinear phase noise, knowledge of the system can be replaced by training a machine learning algorithm to learn the impairments of the link. A fundamental limitation of the simulation carried out was the limited amount of transmitted symbols. Given that when the *symbol error rate* is low (when few mistakes are made) more symbols are need to accurately measure the error rate (more symbols are need to make a mistake). Given the limitation in simulation run time, it was not possible to find a precises optimal transmission length for the nonparameter scheme. All conclusion had to be drawn from the *optimal prediction region* of the processing scheme.

In the multiplexed system it was determined that no clear benefit could be determined in implementing machine learning to mitigate interchannel noise. Suggesting that *cross phase modulation*, the dominant noise source in the link is a time dynamic effect and can not be resolved with the digital signal processing scheme implemented in this report. However, the joint implementation with other technologies could help improve the reliability of the system. This simulation was also limited by the amount of symbols transmitted. If more symbols could be simulated it could of been

determined if there is clear benefit in implementing the proposed scheme.

The research carried out in this project could be extended in a couple of different ways. To get a clear measurement of the optimal transmission length a simulation with 2^{17} symbols could be carried out. This was not possible in this project because of the increase in computing time throughout the whole implementation. An other relevant aspect that was not pursued was investigating the effect of using different constellation diagrams. In this project only a square 16QAM constellation was implemented. Depending on the constellation, the amount of acquired nonlinear phase noise changes. One last measurement that could be preformed is observing the effect of changing the fiber arrangement in the multiplexed link. Depending on the dispersion management implemented the interaction between the carrier signal optical power and the fibers Kerr effect could be altered. Given a clever fiber arrangement a system with less nonlinear noise could be constructed.

As a final remark, in this project it was observed that some of the nonlinear noise present in a fiber link could be compensated without needing any information on the systems physical state. Suggesting that machine learning as a statistical tool could help reduce the deployment complexity of a network. It can also be mentioned that the proposed processing scheme can learn form the real impairments of the system which is desired. The limitation of the processing scheme lay in its inability to resolve time dynamic noise sources. Personally, the lesson I take away from this project is simple, in some cases knowledge can be substituted by training. Which is not only true for nonlinear phase noise compensation in coherent long-haul fiber system, but for life in general.

BIBLIOGRAPHY

[1] B. Saleh and M. Teich, *Fundamentals of Photonics*. Wiley Series in Pure and Applied Optics, Wiley, 2007.

[2] F. N. Khan, Q. Fan, C. Lu, and A. P. T. Lau, "An optical communication's perspective on machine learning and its applications," *Journal of Lightwave Technology*, vol. 37, no. 2, pp. 493–516, 2019.

[3] K.-P. Ho and H.-W. Cuei, "Generation of arbitrary quadrature signals using one dual-drive modulator," *Journal of lightwave technology*, vol. 23, no. 2, p. 764, 2005.

[4] K. Kikuchi, "Fundamentals of coherent optical fiber communications," *Journal of Lightwave Technology*, vol. 34, no. 1, pp. 157–179, 2015.

[5] T. Kato, E. Sasaoka, and S. Ishikawa, "Dispersion-shifted fiber," June 6 2000. US Patent 6,072,929.

[6] K. Aikawa, S. Shimizu, T. Suzuki, R. Suzuki, M. Nakayama, and K. Himeno, "Dispersion compensating fiber and dispersion compensating fiber module," Aug. 30 2005. US Patent 6,937,805.

[7] K.-P. Ho and J. Kahn, "Electronic compensation technique to mitigate nonlinear phase noise," *Lightwave Technology, Journal of*, vol. 22, pp. 779 – 783, 04 2004.

[8] M. Hanselmann, U. Köthe, M. Kirchner, B. Renard, E. Amstalden van Hove, K. Glunde, R. Heeren, F. Hamprecht, and R. Morgan, "Towards digital staining using imaging mass spectrometry and random forests-technical report," 04 2009.

[9] VPItransmissionMaker, "Vpi component maker™," *Photonic Modules Reference Manual*.

[10] A. A. Huurdeman, *The worldwide history of telecommunications*. John Wiley & Sons, 2003.

[11] A. Arvidsson *et al.*, *Brands: Meaning and value in media culture*. Psychology Press, 2006.

[12] M. Seimetz and R. Freund, "Next generation optical networks based on higher-order modulation formats, coherent receivers and electronic distortion equalization," *Elektrotechnik und Informationstechnik*, vol. 125, pp. 284–289, 08 2008.

[13] M. Marciniak, "Future networks - beyond next generation networking," in *2008 10th Anniversary International Conference on Transparent Optical Networks*, vol. 1, pp. 25–28, June 2008.

[14] S. Kumar, "Analysis of nonlinear phase noise in coherent fiber-optic systems based on phase shift keying," *Journal of Lightwave Technology*, vol. 27, pp. 4722–4733, Nov 2009.

[15] T. S. R. Shen, K. Meng, A. P. T. Lau, and Z. Y. Dong, "Optical performance monitoring using artificial neural network trained with asynchronous amplitude histograms," *IEEE Photonics Technology Letters*, vol. 22, no. 22, pp. 1665–1667, 2010.

[16] J. A. Jargon, X. Wu, and A. E. Willner, "Optical performance monitoring using artificial neural networks trained with eye-diagram parameters," *IEEE Photonics Technology Letters*, vol. 21, no. 1, pp. 54–56, 2008.

[17] F. N. Khan, Y. Zhou, A. P. T. Lau, and C. Lu, "Modulation format identification in heterogeneous fiber-optic networks using artificial neural networks," *Optics express*, vol. 20, no. 11, pp. 12422–12431, 2012.

[18] D. Zibar, L. H. H. de Carvalho, M. Piels, A. Doberstein, J. Diniz, B. Nebendahl, C. Franciscangelis, J. Estaran, H. Haisch, N. G. Gonzalez, *et al.*, "Application of machine learning techniques for amplitude and phase noise characterization," *Journal of Lightwave Technology*, vol. 33, no. 7, pp. 1333–1343, 2015.

[19] M. Li, S. Yu, J. Yang, Z. Chen, Y. Han, and W. Gu, "Nonparameter nonlinear phase noise mitigation by using m-ary support vector machine for coherent optical systems," *IEEE Photonics Journal*, vol. 5, pp. 7800312–7800312, Dec 2013.

[20] J.-i. Nishizawa and K. Suto, "Terahertz wave generation and light amplification using raman effect," *Bhat, KN and DasGupta, Amitava. Physics of semiconductor devices. New Delhi, India: Narosa Publishing House*, p. 27, 2004.

[21] J. B. Gallegos, Doyle; Narimatsu and O. Ariana, *Innovative Business Models for Expanding Fiber-Optic Networks and Closing the Access Gaps*. World Bank Group, 2018.

[22] K. Kikuchi, *Coherent Optical Communications: Historical Perspectives and Future Directions*, pp. 11–49. 06 2010.

[23] M. Arumugam, "Optical fiber communication—an overview," *Pramana*, vol. 57, no. 5-6, pp. 849–869, 2001.

[24] R. Ramaswami, K. Sivarajan, and G. Sasaki, *Optical networks: a practical perspective*. Morgan Kaufmann, 2009.

[25] K. Kao and G. A. Hockham, "Dielectric-fibre surface waveguides for optical frequencies," in *Proceedings of the Institution of Electrical Engineers*, vol. 113, pp. 1151–1158, IET, 1966.

[26] R. Ramaswami, "Optical fiber communication: from transmission to networking," *IEEE Communications Magazine*, vol. 40, pp. 138–147, May 2002.

[27] P. Sharma, S. Pardeshi, R. K. Arora, and M. Singh, "A review of the development in the field of fiber optic communication systems," 2013.

[28] K. Kikuchi, "Coherent optical communications—history, state-of-the-art technologies, and challenges for the future—," in *OECC/ACOFT 2008-Joint Conference of the Opto-Electronics and Communications Conference and the Australian Conference on Optical Fibre Technology*, pp. 1–4, IEEE, 2008.

[29] G. P. Agrawal, *Optical Fibers*. John Wiley & Sons, Ltd, 2011.

[30] C. Behrens, R. I. Killey, S. J. Savory, M. Chen, and P. Bayvel, "Nonlinear distortion in transmission of higher order modulation formats," *IEEE Photonics Technology Letters*, vol. 22, no. 15, pp. 1111–1113, 2010.

[31] S. Benedetto and E. Biglieri, *Principles of digital transmission: with wireless applications*. Springer Science & Business Media, 1999.

[32] W. T. Webb and L. Hanzo, *Modern Quadrature Amplitude Modulation: Principles and applications for fixed and wireless channels: one*. IEEE Press-John Wiley, 1994.

[33] W. H. Paik, S. A. Lery, and J. M. Fox, "Mode selective quadrature amplitude modulation communication system," Nov. 8 1994. US Patent 5,363,408.

[34] G. Agrawal, *Applications of Nonlinear Fiber Optics*. Elsevier Science, 2010.

[35] M. Born, E. Wolf, A. B. Bhatia, P. C. Clemmow, D. Gabor, A. R. Stokes, A. M. Taylor, P. A. Wayman, and W. L. Wilcock, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, 7 ed., 1999.

[36] N. Ekanayake and H. M. V. R. Herath, "Effect of nonlinear phase noise on the performance of m-ary psk signals in optical fiber links," *J. Lightwave Technol.*, vol. 31, pp. 447–454, Feb 2013.

[37] D. Zibar, O. Winther, N. Franceschi, R. Borkowski, A. Caballero, V. Arlunno, M. N. Schmidt, N. G. Gonzales, B. Mao, Y. Ye, K. J. Larsen, and I. T. Monroy, "Nonlinear impairment compensation using expectation maximization for dispersion managed and unmanaged pdm 16-qam transmission," *Opt. Express*, vol. 20, pp. B181–B196, Dec 2012.

[38] A. Kaplan and M. Haenlein, "Siri, siri, in my hand: Who's the fairest in the land? on the interpretations, illustrations, and implications of artificial intelligence," *Business Horizons*, vol. 62, no. 1, pp. 15–25, 2019.

[39] S. Marsland, *Machine learning: an algorithmic perspective*. Chapman and Hall/CRC, 2014.

[40] K. Kikuchi, "Coherent optical communications: Historical perspectives and future directions," in *High Spectral Density Optical Communication Technologies*, pp. 11–49, Springer, 2010.

[41] Y. Kotaki and H. Ishikawa, "Wavelength tunable dfb and dbr lasers for coherent optical fibre communications," *IEE Proceedings J - Optoelectronics*, vol. 138, no. 2, pp. 171–177, 1991.

[42] G. Agrawal, *Applications of nonlinear fiber optics*. Elsevier, 2001.

[43] K. Zou, Y. Zhu, F. Zhang, and Z. Chen, "Spectrally efficient terabit optical transmission with nyquist 64-qam half-cycle subcarrier modulation and direct detection," *Optics letters*, vol. 41, no. 12, pp. 2767–2770, 2016.

[44] M. Kuschnerov, F. N. Hauske, K. Piyawanno, B. Spinnler, M. S. Alfiad, A. Napoli, and B. Lankl, "Dsp for coherent single-carrier receivers," *Journal of lightwave technology*, vol. 27, no. 16, pp. 3614–3622, 2009.

[45] N. Stojanovic, F. N. Hauske, C. Xie, and M. Chen, "Clock recovery in coherent optical receivers," in *Photonic Networks, 12. ITG Symposium*, pp. 1–4, VDE, 2011.

[46] M. Seimetz, "Multi-format transmitters for coherent optical m-psk and m-qam transmission," in *Proceedings of 2005 7th International Conference Transparent Optical Networks, 2005.*, vol. 2, pp. 225–229, IEEE, 2005.

[47] M. Levenson and N. Bloembergen, "Dispersion of the nonlinear optical susceptibility tensor in centrosymmetric media," *Physical Review B*, vol. 10, no. 10, p. 4447, 1974.

[48] R. Stolen and A. Ashkin, "Optical kerr effect in glass waveguide," *Applied Physics Letters*, vol. 22, no. 6, pp. 294–296, 1973.

[49] R. W. Boyd, *Nonlinear optics*. Elsevier, 2003.

[50] G. P. Agrawal, "Nonlinear fiber optics," in *Nonlinear Science at the Dawn of the 21st Century*, Springer, 2000.

[51] T. Miya, Y. Terunuma, T. Hosaka, and T. Miyashita, "Ultimate low-loss single-mode fibre at 1.55 μm," *Electronics Letters*, vol. 15, no. 4, pp. 106–108, 1979.

[52] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

[53] S. Dris, H. Louchet, and A. Richter, "Analysis of nonlinear interference noise in flexible optical networks," in *2017 Opto-Electronics and Communications Conference (OECC) and Photonics Global Conference (PGC)*, pp. 1–3, IEEE, 2017.

[54] N. Le Binh, *Advanced digital optical communications*. CRC Press, 2015.

[55] J. P. Gordon and L. F. Mollenauer, "Phase noise in photonic communications systems using linear amplifiers," *Optics letters*, vol. 15, no. 23, pp. 1351–1353, 1990.

[56] G. Goldfarb and G. Li, "Chromatic dispersion compensation using digital iir filtering with coherent detection," *IEEE Photonics Technology Letters*, vol. 19, no. 13, pp. 969–971, 2007.

[57] K. Kim, R. Stolen, W. Reed, and K. Quoi, "Measurement of the nonlinear index of silica-core and dispersion-shifted fibers," *Optics letters*, vol. 19, no. 4, pp. 257–259, 1994.

[58] Y. Namihira, K. Miyagi, K. Kaneshima, M. Tadakuma, C. Vinegoni, G. Pietra, and K. Kawanami, "A comparison of six techniques for nonlinear coefficient measurements of various single mode optical fibers.," *NIST SPECIAL PUBLICATION SP*, pp. 15–18, 2002.

[59] X. Liu, X. Wei, R. E. Slusher, and C. McKinstrie, "Improving transmission performance in differential phase-shift-keyed systems by use of lumped nonlinear phase-shift compensation," *Optics letters*, vol. 27, no. 18, pp. 1616–1618, 2002.

[60] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the fifth annual workshop on Computational learning theory*, pp. 144–152, 1992.

[61] A. J. Smola and B. Schölkopf, *Learning with kernels*, vol. 4. Citeseer, 1998.

[62] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[63] T. K. Ho, "Random decision forests," in *Proceedings of 3rd international conference on document analysis and recognition*, vol. 1, pp. 278–282, IEEE, 1995.

[64] D. Heath, S. Kasif, and S. Salzberg, "Induction of oblique decision trees," in *IJCAI*, vol. 1993, pp. 1002–1007, 1993.

[65] J. H. Friedman, "Stochastic gradient boosting," *Computational statistics & data analysis*, vol. 38, no. 4, pp. 367–378, 2002.

[66] G. R. Hancock and M. Liu, "Bootstrapping standard errors and data-model fit statistics in structural equation modeling.," 2012.