Impact of Stochastic Transmission Effects on the Design of Digital Back Propagation

vorgelegt von
M.Sc.
Kseniia Goroshko

an der Fakultät IV - Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften
- Dr.-Ing. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr.-Ing. Wolfgang Heinrich
Gutachter: Prof. Dr.-Ing. Klaus Petermann
Gutachter: Prof. Dr.-Ing. Peter Krummrich
Gutachter: Dr.-Ing. André Richter

Tag der wissenschaftlichen Aussprache: 22. August 2019

Berlin 2020

Acknowledgement

This thesis is a result of a 3.5 years research with ICONE project, VPIphotonics and TU Berlin. Dozens of extraordinary people have helped and taught me immensely during this time. First and foremost I wish to thank my supervisors Professor Klaus Petermann, Doctor André Richter and Doctor Hadrien Louchet for their continuous support. I am heartily thankful for their trust, guidance and patience.

Many thanks are sent to the extraordinary VPIphotonics team for their kindness, support, stimulating talks, cookies and hilarious jokes. It was a real pleasure to work with Maria, Vera, Beate, Cristina, Thomas, Kay, Stephanos, Nuno, Igor.

Many hugs and thanks to my peers at the ICONE project, mi casa es su casa, Giuseppe, Asif, Aditya, Jaime, Aleksanders, Oskars, Francesca, Hugo, Simone, Tu, Auro, Faruk, Marti, Hou-Man. I would also like to thank all the experts involved in the project from both the academic side and the industrial side. It was my pleasure to see how Doctor Noreen Akhtar manages the project and to learn from her.

I am grateful for the amazing reception in my two secondments, for the hospitality and excellent possibility to learn: to Professor Popov, Doctor Ozolins, Doctor Pang, Professor Jacobsen, Doctor Iglesias Olmedo from Acreo Swedish ICT and KTH Royal Institute of Technology in Stockholm, and to Professor Juan Diego, Doctor Rizelli, Doctor Corredera Guillén from IO-CSIC in Madrid.

I gratefully acknowledge the EU Marie Curie ITN grant that made my work possible.

I am thankful to Aleksanders and Braulio for checking the first versions of this work and making it more coherent through their remarks.

I am heartily grateful to Pablo whose support and love are so appreciated. Thank you for planting optimism that I will, in fact, finish this work.

Finally, with a lot of love I thank my parents for passing me the best both have, for encouraging me and loving me. Thank you!

Abstract

During the last three decades optical telecommunications industry is constantly pushing for new data transmission capacity limits. With a nowadays paradigm of coherent communication, one of the ways to increase the data rates is closely related to increasing the signal transmission power. That in turns allows reaching higher optical signal-to-noise ratio (OSNR). An intrinsic part of this solution will be proposing a proper management for fibre nonlinearities (NL) that arise with the high optical power.

One of the promising techniques to mitigate NL is called digital back propagation (DBP). The main assumption of DBP is that fibre propagation can be fully described by the Manakov-polarisation mode dispersion (PMD) equations. The most popular DBP implementation is based on the split-step Fourier (SSF) method. Its performance is limited by the accuracy of the actual implementation (signal discretization, step size) and the uncertainty of the transmission link (fibre parameters, amplification map). Fundamental limitations of DBP that are independent from implementation arise from the impact of stochastic propagation impairments on fibre NL, such as amplified spontaneous emission (ASE) noise and polarisation effects.

This thesis will focus on the impact of stochastic propagation effects on DBP performance. The key limiting effects, PMD and polarisation dependent loss (PDL), will be characterised. A supplement to the perturbation model will be proposed to account for PMD analytically and to predict the DBP error. Further, a low-complex and high-efficient DBP modification will be proposed. Finally, the fundamental limits of the conventional and the proposed DBP will be quantified.

Keywords: Nonlinear optics, coherent communications, digital back propagation, polarisation mode dispersion

Zusammenfassung

In den letzten drei Jahrzehnten treibt die optische Datenübertragungsbranche neue Kapazitätsgrenzen voran. Mit dem heutigen Paradigma der kohärenten Kommunikation ist eine der Möglichkeiten, um höhere Datenraten zu erreichen, zugleich eng mit der Erhöhung der Sendeleistung verbunden, wodurch wiederum ein höheres optisches Signal/Rausch-Verhältnis erreicht werden kann. Ein wesentlicher Teil dieser Lösung wird darin bestehen, einen ordnungsgemäßen Umgang mit der Nichtlinearität in Glasfasern, die bei der hohen optischen Leistung auftreten, vorzuschlagen.

Eine der vielversprechenden Techniken zur Minderung von Nichtlinearitäten wird als digitale Rückwärtsausbreitung bezeichnet. Die Hauptannahme der digitalen Rückwärtsausbreitung ist, dass die Wellenausbreitung in Fasern durch die Manakov-Polarisationsmodendispersion Gleichungen vollständig beschrieben werden kann.

Die Leistung der digitalen Rückwärtsausbreitung wird durch die numerische Genauigkeit der tatsächlichen Implementierung (Signaldiskretisierung, Schrittgröße) und die Unbestimmtheit der tatsächlichen Übertragungsverbindung (Faserparameter, Verstärkungskarte) begrenzt. Grundsätzliche Einschränkungen, die von der Implementierung unabhängig sind, ergeben sich aus den Auswirkungen stochastischer Ausbreitungsbeeinträchtigungen auf Fasernichtlinearitäten, wie beispielsweise verstärktes spontanes Emissionsrauschen und Polarisationswirkungen der Doppelbrechung der Faser.

Diese Arbeit konzentriert sich auf die Auswirkungen stochastischer Ausbreitungseffekte auf die Leistung der digitalen Rückwärtsausbreitung. Die wichtigsten einschränkenden Wirkungen, Polarisationsmodendispersion und polarisationsbedingter Verlust werden dabei charakterisiert. Eine Ergänzung des Perturbationsmodells wird vorgeschlagen, um Polarisationsmodendispersion analytisch zu berücksichtigen und den Fehler der digitalen Rückwärtsausbreitung vorherzusagen. Ferner wird eine hocheffiziente Modifikation der digitalen Rückwärtsausbreitung mit niedriger Kapazität vorgeschlagen, um die Nichtlinearität- Polarisationsmodendispersion-Wechselwirkungen zu verringern. Schließlich werden die grundlegenden Grenzen der konventionellen und vorgeschlagenen digitalen Rückwärtsausbreitung quantifiziert.

Schlagwörter: Nichtlineare Optik, kohärente Kommunikation, digitale Rückwärtsausbreitung, Polarisationsmodendispersion

Contents

Chapter 1

Introduction

In this chapter we will introduce the scope of the thesis, in that we will discuss the modern data traffic scenarios and demands, capacity crunch, and techniques to increase optical fibre throughput. We motivate for the importance of the current work and outline its possible impact. The overview of the thesis content can be found in the end of this chapter.

1.1 Overview

In the first two decades of the XXI century we find ourselves in a continuous and accelerating growth in different sectors of society, such as socio-economics, habitat, healthcare, welfare, culture and education, communications and technology. All of these areas show high dependence on global communications and, therefore, on data transfer. Since the first deployment of long-haul optical cables in the middle 80s and the boosting growth of optical networks in the late 90s and early 00s new generations have been growing up with access to the Internet. It is nowadays a fundamental part in our lives.

An easy access to telecommunication and information is strongly incorporated in minds of the so-called Z generation. In fact this generation and the subsequent ones are already setting the global social and political trends and will become world-shapers in 2030-2050s. It is highly probable that the basic demand in the Internet reality will strengthen further and will define the future society development, mindsets and philosophy. Already now, recently emerged applications, such as Internet of Things, automated medical diagnostic, data

analysis and control, autonomous driving, virtual reality (VR) or artificial intelligence (AI) are maturing and to become a technological routine in the near future.

Facing the future, our task as a scientific community today, is to consider these applications with a fundamental perspective and to create a solid basis for the Internet backbone, that is flexible, scalable, and ready to meet the future demand on reliable and qualitative traffic flow.

As of today, optical fibres represent the only medium capable of transporting over longer distances the audacity of data produced by the humanity. Optical links provide necessary bandwidth, and their application ranges from high-capacity Internet backbone networks to data centres and supercomputer parallel interconnects. The traffic especially in core networks is growing from year to year [1], mainly due to massive amounts of data produced by such applications as YouTube, Netflix, Facebook, Instagram, iTunes, Spotify, etc. and is associated with cloud services and content providers, such as Amazon, Google CDN, Akamai, and Level3 Communications.

We are currently living in the third era of fibre communications technology [2]. The first era of regenerated direct detection lasted from late 70s to early 90s, followed by an erbium-doped fibre amplifier (EDFA) wavelength division multiplexing (WDM) era until the beginning of the new millennium, when an era of non-dispersion managed (non-DM) coherent communication emerged. Each time the new technological paradigm enabled a solution for the needs of optical transmission system. The current coherent technology gives a possibility to increase the fibre capacity by providing full access to the optical field information. The spectral efficiency, measure of information transmitted by a bandwidth unit, is increased by encoding more bits per symbol in higher order modulation formats and by utilising digital signal processing for restoring the signal and improving its quality. The innovations in these two areas have a potential to bring the system performance close to the theoretical limit, the Shannon Bound [3]. The possible gap in the future between the demand and the available communication rates is referred to as capacity crunch.

Various estimations on the data traffic demands have been performed. According to Cisco prognoses [1], "annual global Internet protocol (IP) traffic will

reach 3.3 Zettabytes (ZB) by 2021" (one ZB equals to 1000 Exabytes (EB), 10^{21} bytes), showing a compound annual growth rate (CAGR) of 24% from 2016. Particularly in Western Europe the annual IP traffic is predicted to reach 0.45 ZB in 2021 at the CAGR of 22% from 2016. In order to get a very rough estimation of traffic volume transmitted per second (traffic rate), one can divide the annual value by the number of seconds in a year. Thus, in 2021 in Western Europe the Internet average traffic rate of 14.3 TB/s or 114 $Tbit/s$ in all the links combined can be expected.

On top of that the distribution of traffic demand throughout the day should be taken into account. Typically for Western Europe peak hours are between 7 am and 11 pm [4], increasing the average data traffic in downstream by 50%. This means that the network providers in 3 years (from 2018) should be able to meet the demand on rush hour of ≈ 170 $Tbit/s$ downstream data speed (the number refers to the whole network).

In the latest years united efforts of the scientific and research community are focused on finding new technologies to increase the optical networks' transmission rates in order to meet the traffic demand. Two major research paths can be identified: Technologies operating in the linear and nonlinear fibre regimes. By linear transmission the capacity limit of standard single mode fibre (SSMF) O-L-bands (see Table 1.1) is $\sim 100 - 200$ $Tbit/s$ [5]. By transmission in nonlinear regime, the same rates can be potentially achieved without having to utilise all fibre transmission bands. In this thesis we focus on the nonlinear fibre transmission regimes. In such fibre regimes the signal experiences a degradation during propagation, which is attributed to the fundamental nonlinear properties of the silica fibre at high optical powers. The main reason to transmit in nonlinear regimes is a possibility to increase the signal power compared to the noise power. The ratio of the two expressed in dB is traditionally used as a signal quality measure and is referred to as OSNR. While decreasing noise is constrained by fundamental limits, a great potential lies in increasing the optical signal power, thus increasing the OSNR and improving the signal quality at the receiver end.

1.1.1 Transmission techniques roadmap

Various approaches have been proposed for avoiding the capacity crunch. A comprehensive overview of what principally can be done is proposed by Peter J. Winzer, for example in [6]. According to his view, there are five physical dimensions, that can be utilized to increase fibre capacity, and three of them are currently fully utilized: *time*, *quadrature* and *polarisation*. The two remaining physical dimensions are *frequency* and *space*. Examples of expanding the *frequency* dimension would be utilising all fibre transmission O-L-bands, summarized in Table 1.1 Examples of expanding the *space* dimension would be multicore, multimode fibres and their combination.

Table 1.1: Optical fibre transmission windows.

Band	Wavelength range
O	1260 – 1360 nm
E	1360 – 1460 nm
S	1460 – 1530 nm
C	1530 – 1565 nm
L	1565 – 1625 nm
U	1625 – 1675 nm

Additionally to utilising physical dimensions, the nonlinear nature of the fibre can be addressed. Provided there is a way to mitigate the nonlinear distortion in a fibre, the input signal power can be increased thus increasing the OSNR, allowing usage of higher-order modulation formats and providing longer reach distances. A strong research effort was directed to developing mathematical tools and experimental techniques for data transmission over the nonlinear channel in the last decade [7]. Further, innovations in digital signal processing (DSP) and advanced forward-error correction (FEC) [8], have a high potential to bring the performance closer to the Shannon Bound.

Nonetheless, it is worth noting that an OSNR gain would only yield a limited fibre capacity increase due to their logarithmic relationship (see eq. (7.1) in chapter 7), compared to *frequency* and *space* dimensions.

1.1.2 Link scenarios

In order to support the transmission of huge volumes of data, the telecommunications industry should find ways to upgrade the Internet backbone Distortions due to fibre NL can become a significant problem in long-haul networks. In this thesis we will focus on long-haul transmission systems. Nowadays those are represented by the submarine cable network, as in Figure 1.1, and by the long-haul terrestrial links. In metro and regional applications network's cost, simplicity, and power consumption are the dominating concerns.

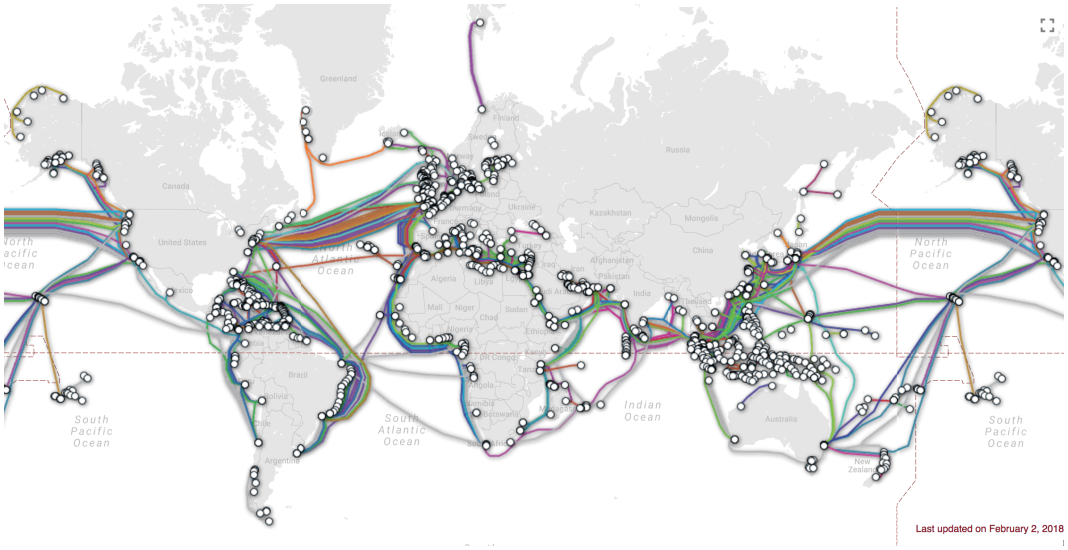


Figure 1.1: Submarine cable map, retrieved from [9].

Undersea communications have historically been one of the most challenging and innovative industries. A great number of post-deadline conference publications are demonstrating the push towards higher capacity and longer transmission distance. Now that the paradigm has shifted to coherent communications, long-haul submarine links are a good candidate do deploy the nonlinearity management, provided the energy efficiency and cost are balanced. The US, Asian, and EU long-haul terrestrial links would as well benefit from the technology, which will allow upgrading existing links and thus increasing the network capacity without costly laying of the new fibres. A comprehensive overview of the commercial application of the NL compensating techniques is presented in [10].

In the current work we concentrate on the intermediate range (from 1000 km

up to 4500 km) links and discuss which propagation effects could appear when deploying one of the promising nonlinearity mitigation techniques, digital back propagation (see Section 2.1).

One of the latest use cases is an undersea communication link Marea (Spanish for "tide"), which was developed and built late 2017 by Telxius, Facebook and Microsoft. It is a 6.600 km link consisting of a bundle of 8 SSMFs, each transmitting 20 Tb/s, and a total capacity of 160 Tb/s. The technologies used include L-Band extension (see Table 1.1), 64-quadrature amplitude modulation (QAM) (modulation format encoding 6 bits per symbol) and above, probabilistic constellation shaping (when low-energy symbols are transmitted more frequently than high-energy ones), hybrid EDFA/Raman amplification (based on the gain from the effect of stimulated Raman scattering [11]), flexible symbol rate, modulation and spacing allocation, and nonlinear compensation [12].

1.1.3 Techniques to reduce NL impact

Generally, the techniques to reduce the impact of Kerr NL can be divided into those that mitigate (or pre-mitigate) the impact on the propagated signal, and those that increase the resistance of the propagated signal towards NL. The first include such techniques as digital back propagation (DBP) [13, 14] and its numerous modifications, techniques based on signal conjugation: optical phase conjugation (OPC) [15] and twin waves [16] (these are illustrated in Figure 1.2); perturbation-based compensation [17], nonlinear equalizers, such as maximum-likelihood sequence estimation (MLSE) and Volterra series nonlinear equalization [18]. The second class of techniques includes advanced modulation formats (geometric [19, 20] and probabilistic constellation shaping for NL [21]), and nonlinear Fourier transform (NFT) [22].

A brief analysis of complexity and flexibility of the methods, their ability to compensate for the NL, including in presence of polarisation mode dispersion (PMD), is presented in Table 1.2. By the **processed band** the received and processed signal is understood, by the **full band** – the co-propagating part of the signal spectrum that is yet not included in NL compensation. Each of the techniques has its benefits and challenges, so that in different scenarios different technique (or combination of techniques) might be applied.

Perturbation and nonlinear equalization methods provide rather limited

compensation gain and require a high computational effort. DBP although also exhibiting a prohibitively high computational complexity, gained attention in the last decade. Numerous studies report a possibility to trade-off high complexity for a reasonable reduction in performance. However, the computational cost of the algorithm still grows exponentially with increasing processed bandwidth. OPC experimentally shows a good full-band performance increase, however, puts strict requirements on the link symmetry and conjugation effectiveness. Twin waves, even with proposed modifications [23], sacrifice up to a quarter of the spectral efficiency, besides, the method allows low link flexibility in terms of dynamic routing.

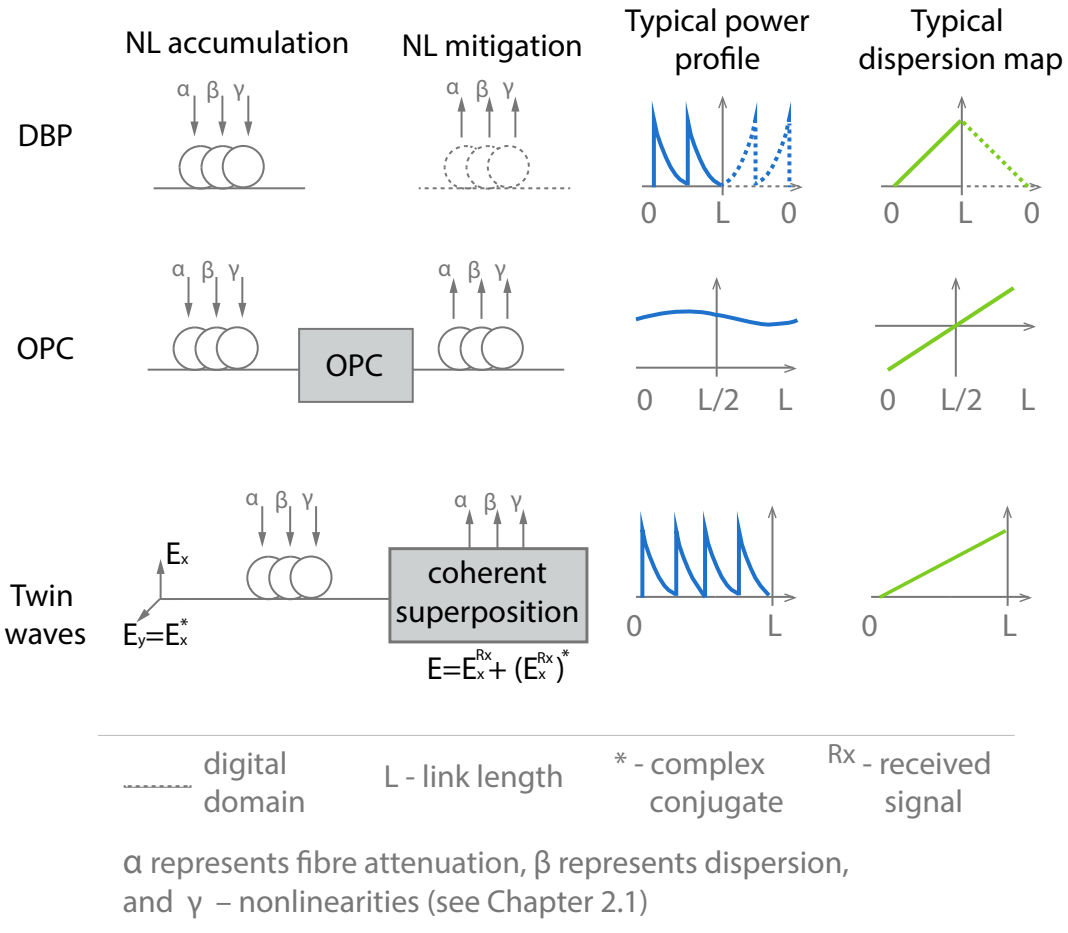


Figure 1.2: Basic illustration of signal conjugation based NL mitigation techniques.

Table 1.2: DSP approaches for NL mitigation.

	Processed band compensation	**Full band compensation**	**Link Flexibility**	**Biggest challenges**	**NL-PMD penalty mitigation**
DBP	$++$	$--$	link is flexible, link parameters should be known	computational complexity, stochastic effects	current thesis, [24]
OPC	$+$	$+$	symmetric dispersion and power maps are required	efficient phase conjugation, stochastic effects	no work found PMD avoidance [25]
Twin Waves	$+$	$+$	symmetric dispersion map is required	loss of 50% of transmission bandwidth, stochastic effects	no work found
Equalization	$-$	$-$	does not put constrains on link	mature technology	no work found
Constellation shaping	$-$	$-$	does not put constrains on link	mature technology	no work found
NFT	$++$	$--$	link is flexible, link parameters should be known no dynamical routing	computational complexity, stochastic effects	no work found

Markers $(++, +, -, --)$ indicate a potential mitigation gain from full compensation, to moderate impact, to typically small, to zero impact

Probabilistic shaping and pulse shaping are effective approaches to increase the signal tolerance to NL. They are reported to improve the received effective signal-to-noise ratio (SNR) on a fraction of decibel (dB) range and can be used complement to other NL mitigation techniques. NFT, a fundamentally different and potentially revolutionary approach, promises high spectral efficiency, but requires significant work to be done in computational complexity reduction, incorporation in add-drop optical networks, and noise treatment.

Regarding the above information, DBP is one of the most promising and mature techniques. Its positive aspects are that it allows for certain link flexibility, compensates for both linear and nonlinear effects, has good intra-channel effects compensation, and with the newly developed modifications it can trade-off computational complexity for the proportion of compensated Kerr NL effects.

1.2 Motivation: fundamental limitations of DBP

1.2.1 Stochastic-nonlinear interactions

During propagation through a fibre a signal is affected by the linear and nonlinear effects, and the stochastic noise, as illustrated in Figure 1.3 (adopted from [26]). These effects are distinct by their nature but are occurring simultaneously during propagation and are not independent. If one would want to simulate the signal propagation in a fibre, accurate results would require a simultaneous simulation of all these effects, as they are influencing each other. From the perspective of Kerr nonlinear effects compensation, regardless of which technique is used, it would mean that linear, nonlinear, and noise effects should be considered simultaneously.

For the NL compensation it is important that both polarisation effects and noise randomly change the signal amplitude profile. As Kerr nonlinearities are proportional to the cube of optical field amplitude, the presence of those effects will affect nonlinearities. An example of how PMD affects the propagating optical field envelope is shown in Figure 1.4. PMD is a linear polarisation effect, which is describing a time delay between the fibre X and Y polarisation components arising during propagation. The strength of the effect can be

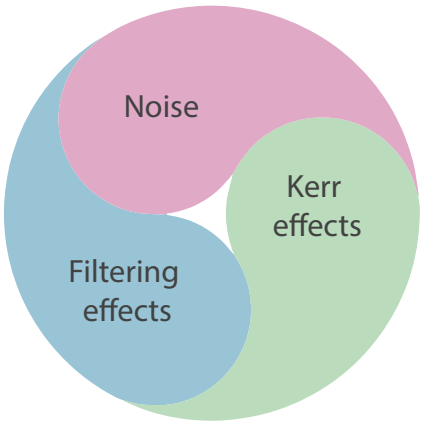
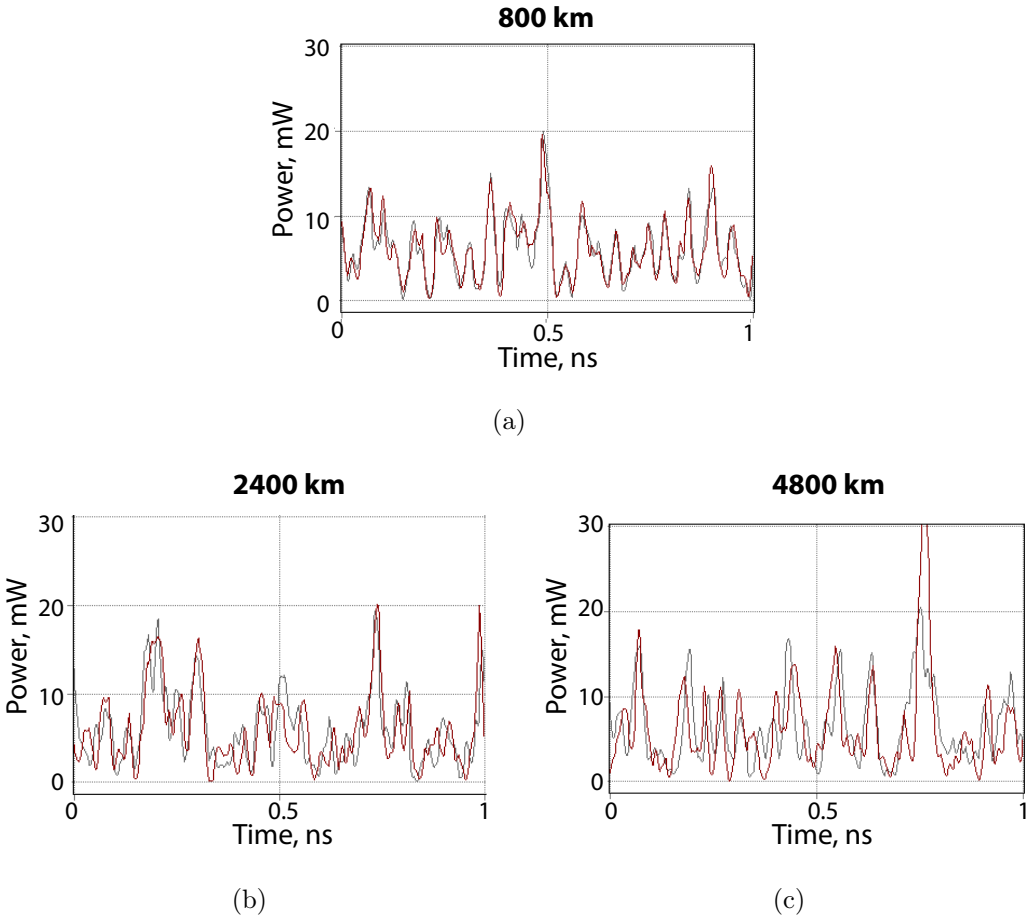


Figure 1.3: Physical phenomena present in the optical transmission system, adopted from [26].



(a)

(b)

(c)

Figure 1.4: Change of the local optical field in a fibre with PMD (red curve) compared to a fibre without PMD (grey curve) at three different distances.

indicated by the polarisation mode dispersion coefficient (PMDc). Most of the modern fibres will have PMDc between 0.04 and 0.1 $ps/\sqrt{km}$, very old ones can have 0.2 $ps/\sqrt{km}$ and higher.

The simulation in Figure 1.4 considers dual polarisation (DP)-16-QAM 28 $GBaud$ signal propagating in a lumped amplified noiseless system with 80 km SSMF spans. The field is presented at three different propagation distances, corresponding to 10, 30, and 50 spans in a fibre without PMD and in a 0.1 $ps/\sqrt{km}$ PMDc fibre. It can be seen that the PMD introduces a significant change in the optical intensity distribution and therefore changes the impact of the nonlinear effects.

1.2.2 Impact of stochastic-nonlinear interactions

A conventional digital Kerr-nonlinearity compensation technique, including DBP, assumes that the signal experienced only deterministic effects, such as Kerr NL or chromatic dispersion (CD). In reality the optical field local amplitudes are affected by stochastic effects as well. And therefore the signal would experience a different impact of the NL. Thus, not the stochastic effects themselves, but the interplay between them and NL becomes a limiting factor of compensation techniques.

The performance of any DBP is further limited by the numerical accuracy of the actual implementation, limited by the signal discretisation, DBP step size, uncertainty of the actual transmission link parameters, such as fibre attenuation, nonlinear index, amplification and dispersion maps. Nonetheless, fundamental limitations of DBP are independent from implementation and arise from the impact of stochastic propagation impairments on fibre nonlinearities, such as ASE noise from EDFAs and fibre birefringence effects, for example, PMD. Not only the performance of DBP, but also of OPC as we showed in [27], the twin waves, perturbation-based compensation, and of NFT will be affected by stochastic effects.

1.2.3 State of the art

The numerical and analytical comparison of impact of the stochastic effects on DBP performance has shown that PMD is of major concern in non-dispersion

managed (non-DM) orthogonal frequency division multiplexing (OFDM) [28], binary phase shift keying (BPSK) [29], quadrature phase shift keying (QPSK) [30] and 16-QAM [31] modulated links. Experimental investigations of dispersion managed (DM) links show that the DBP performance has an equivocal dependence on PMD and link configuration [32]. In non-DM links [33] shows that PMD degrades the filtered DBP performance, though implementation penalties are more significant.

In case of the future deployment of DBP in real-time coherent optical transmission systems, it is worthwhile to make a comprehensive analysis of the stochastic effects', particularly PMD's, impact on the nonlinear transmission. This will allow defining new fundamental limits of DBP. Knowing the principally achievable performance limits will help to resolve the uncertainty of the question: How good the method actually is, compared to how good it could possibly be.

Understanding the impact of stochastic effects on NL compensation will help to propose the modifications in the DBP algorithm, in order to make it resistant to the PMD-induced degradation. This thesis will suggest and investigate one of such DBP modifications. Another practically-implementable approach, besides the one proposed in this thesis, was presented concurrently in [24, 34, 35]. More details and the comparison of the two methods can be found in Section 5.5. Finally, a theoretical approach to mitigate NL-PMD interplay, relying on the fully-known information about PMD along the link, was demonstrated in [29].

1.3 Thesis content

The purpose of this thesis is to provide a comprehensive understanding of the fibre NL-stochastic interactions (NL-ASE, NL-PMD, and NL-PDL), of their impact on DBP, to provide a low-complexity solution to estimate the strength of impact, and to propose a DBP modification greatly reducing the impact of these interactions.

The theoretical background and simulation details are reviewed in Chapter 2. It will focus on key notions, briefly mentioned in the introduction, such as DBP, PMD, ASE. As most of the results in this thesis rely on simulations,

the chapter will also provide a detailed description of modelling concepts, such as split-step Fourier (SSF) or coarse-step model. Different approaches to simulate the stochastic effects will be introduced as well.

In Chapter 3 we will present the statistics of the stochastic propagation effects, such as ASE, PMD and PDL. We will show that the average differential group delay (DGD) evolution along the link is defined by both fibre PMDc and accumulated DGD at the end of the link. In this chapter we will also show how the effects influence DBP performance. Finally, we will present an experiment proving the correlation of DBP gain in linear Q factor and the DGD.

Chapter 4 is dedicated to the perturbation model extended to account for PMD. Different scenarios of DGD evolution will be discussed. It will be shown, that with a constant DGD and span-step DGD, perturbation term calculation can be considerably simplified. We will verify the proposed extension of the perturbation model by comparing the results with SSF simulations. Finally, we will present a new metric describing DBP error due to NL-PMD interaction and will show how it depends on signal power spectral density, symbol rate and modulation format.

The simplification proposed and verified in Chapter 4 builds a basis for the DBP modification. In Chapter 5 we will explicitly investigate the proposed method for typical long-haul transmission systems having different input power, symbol rates and modulation formats, for single- and multichannel. We will discuss practical aspects of the method implementation and its complexity. Finally, we will compare the proposed DBP modification with the method developed in parallel and independently from this work [35].

Some assumptions on a practical implementation of the proposed algorithm are given in Chapter 6. Firstly, we will show how the performance of the proposed DBP is affected when the parameters are not estimated exactly. We will then see how a NL-PMD interaction is pronounced in a realistic DBP implementation. At last, we will focus on one of the methods to estimate DGD for the proposed DBP and demonstrate its accuracy.

In Chapter 7, nonlinearities polarisation mode dispersion (NL-PMD) induced limits are highlighted from the information theoretical point of view. It is shown that for a typical optical transmission systems considered in the thesis it is possible to define PMD-induced fundamental limits on the proportion

of compensated NL and to define maximum achievable spectral efficiency and reach distance.

Finally, the conclusions are drawn in Chapter 8.

Chapter 2

Simulation of stochastic effects and digital back propagation algorithm in fibre transmission links with EDFAs

This chapter will explore the basis for simulating DBP and the stochastic effects in a fibre system. It will cover the DBP principles, EDFA ASE, birefringence, PMD, polarisation dependent gain (PDG) and PDL. Here we will also introduce a notation to describe polarisation effects in a fibre. Finally, we will give the details on how the propagation effects and DBP were simulated in *VPItransmissionMaker Optical Systems*.

2.1 Digital back propagation

2.1.1 Forward propagation effects

The working principle of the DBP is centred at the idea that fibre propagation effects can be described and simulated digitally. The most important propagation effects in SSMF are loss, CD, nonlinear and polarisation effects.

The loss depends on the length of the fibre, L, the attenuation coefficient, α, and the optical angular frequency, ω. The modern SSMFs have α around

0.2 dB/km in the 1550 nm region. It is defined as:

$$\alpha(\omega) = \frac{1}{L} \cdot 10log_{10}\Big(\frac{P_{in}(\omega)}{P_{out}(\omega)}\Big), \tag{2.1}$$

where P_{in} is the input power in the fibre and P_{out} is the output power.

A linearly polarised electrical field, E, propagating along the Z-axis in positive direction in a fibre at frequency ω_0 can be described as [36]:

$$E(z,t) = A(z,t) \cdot e^{i(\beta_0 z - \omega_0 t)}, \tag{2.2}$$

where A is is the slowly varying envelope of the optical field, z is the propagation distance, β_0 is the phase propagation constant at frequency ω_0, that determines the change of phase of an initial light field, t is the time frame moving at ν_g speed. $\nu_g = \frac{\partial\omega}{\partial\beta(\omega)}$ is the group velocity (speed of the optical field envelope propagating at frequency ω). The phase velocity $\nu_{ph} = \frac{\omega}{\beta(\omega)}$ defines the propagation speed of the field phase at a particular frequency ω.

The phase propagation constant β is frequency-dependent and can be expanded into a Taylor series around the centre frequency ω_0 [36]:

$$\beta(\omega) = \beta_0 + \beta_1\Delta\omega + \frac{1}{2!}\beta_2\Delta\omega^2 + \frac{1}{3!}\beta_3\Delta\omega^3 + ..., \tag{2.3}$$

$$\beta_l = \Big(\frac{d^l\beta}{d\omega^l}\Big)_{\omega=\omega_0};\ l = 0, 1, 2... \tag{2.4}$$

where ! denotes the factorial, l is the index, $\Delta\omega = \omega - \omega_0$, $\beta_1 = 1/\nu_g$, β_2 describes the frequency dependence of ν_g. The phenomenon is called chromatic dispersion (CD) or group velocity dispersion (GVD). Dispersion slope β_3 describes frequency dependence of β_2. The CD is typically described with dispersion coefficient D_c in $\frac{ps}{nm\cdot km}$:

$$D_c = -\frac{2\pi c}{\lambda^2}\beta_2, \tag{2.5}$$

where c is the speed of light in vacuum, λ is the wavelength. Typical for SSMF D_c is 16 $\frac{ps}{nm\cdot km}$.

Fibre Kerr nonlinearities arise from a phenomenon called Kerr effect, which presents signal power density dependent refractive index, i.e. signal dependent speed of light. The modified phase propagation constant can be written as:

$$\beta_{nl} = \beta + \gamma P_s, \tag{2.6}$$

where P_s is the signal power, and γ is the nonlinear coefficient:

$$\gamma = \frac{n_2 \omega}{c A_{eff}}, \tag{2.7}$$

n_2 being the nonlinear refractive index, A_{eff} being the fibre effective area. In modern SSMF n_2 is typically $3 \cdot 10^{-20}\ m^2/W$, A_{eff} is 80 μm, γ is 1.37 $\frac{1}{W \cdot km}$.

2.1.2 Manakov equations in time domain

The nonlinear Schrödinger equation (NLSE), derived from Maxwell equations using Equation (2.2) and Equation (2.3), is commonly used to describe the light wave propagation in optical fibre accounting for attenuation, chromatic dispersion, and Kerr nonlinearity while ignoring birefringence (see more in Section 2.3.1) [36]:

$$\frac{\partial A}{\partial z} + \frac{i}{2}\beta_2 \frac{\partial^2 A}{\partial t^2} - \frac{\beta_3}{6}\frac{\partial^3 A}{\partial t^3} + \frac{\alpha}{2}A = i\gamma |A|^2 A, \tag{2.8}$$

where $|A|^2$ is the signal power, and the frame is moving with group velocity.

When birefringence occurs, the two polarisation modes are not independent from each other throughout the propagation. The equations governing evolution of the two polarisation components have been derived in [36] for linearly birefringent fibres (i.e. the polarisation modes' orientation stays constant along the fibre). For the optical field polarised along the birefringence X-axis:

$$\frac{\partial A_x}{\partial z} + \frac{i}{2}\beta_2 \frac{\partial^2 A_x}{\partial t^2} - \frac{\beta_3}{6}\frac{\partial^3 A_x}{\partial t^3} + \frac{\alpha}{2}A_x = i\gamma(|A_x|^2 + \frac{2}{3}|A_y|^2)A_x + \frac{i\gamma}{3}A_x^* A_y^2 e^{-2i\Delta\beta z}, \tag{2.9}$$

where A_x and A_y are the two polarisation components' optical fields, $\Delta\beta$ is the propagation constant difference between the two modes $\Delta\beta = \beta_x - \beta_y$. The solution for Y-polarisation component is obtained by inverting x and y.

In case of a single-mode fibre with birefringence varying faster than the nonlinear interaction an averaging of nonlinear effects over the two polarisation modes occurs, which reduces the strength of the NL. Propagation over such a fibre is described mathematically by the coupled Manakov equations [37–39]:

$$\frac{\partial A_{x,y}}{\partial z} + \frac{i}{2}\beta_2 \frac{\partial^2 A_{x,y}}{\partial t^2} - \frac{\beta_3}{6}\frac{\partial^3 A_{x,y}}{\partial t^3} + \frac{\alpha}{2}A_{x,y} = i\frac{8}{9}\gamma(|A_{x,y}|^2 + |A_{y,x}|^2)A_{x,y}, \tag{2.10}$$

where on the left side the second and third addends account for CD and dispersion slope, fourth addend accounts for attenuation, and on the right side the Kerr nonlinear effects are described.

2.1.3 Digital back propagation fundamentals

Digital back propagation is a DSP method that virtually reconstructs the nonlinear channel in a backward order. For that the backpropagation algorithm solves an inverse NLSE. As a result, the deterministic effects, such as CD, dispersion slope and NL, are compensated after backpropagation.

The Manakov equations (2.10), considering the dispersion slope negligible, can be rewritten as

$$\frac{\partial A_{x,y}}{\partial z} = (\hat{N} + \hat{D} + \hat{P})A_{x,y}, \tag{2.11}$$

where $\hat{N} = i\frac{8}{9}\gamma(\left|A_{x,y}\right|^2 + \left|A_{y,x}\right|^2)$ is the nonlinear operator, $\hat{P} = -\frac{\alpha}{2}$ is the power correction operator, and $\hat{D} = -\frac{i}{2}\beta_2\frac{\partial^2}{\partial t^2}$ is the linear operator. Dispersion slope is typically omitted when the signal is narrow-band and not in the fibre zero GVD region.

The inverse Manakov equations would be:

$$\frac{\partial A_{x,y}}{\partial z} = (-\hat{N} - \hat{D} - \hat{P})A_{x,y}, \tag{2.12}$$

which is also equivalent to sending the signal through a fibre with inverse sign parameters.

The Equation (2.12) can be solved numerically using SSF method [36] as:

$$A_{x,y}(z+h,t) = A_{x,y}(z,t)e^{h(-\hat{N}-\hat{D}-\hat{P})} \approx A_{x,y}(z,t)e^{\frac{h}{2}(-\hat{D})}e^{\int_z^{z+h}(-\hat{N})z'dz'}e^{\frac{h}{2}(-\hat{D})}e^{h(-\hat{P})}, \tag{2.13}$$

where h is the step size. The DBP algorithm flow is illustrated in Figure 2.1. In this DBP implementation the nonlinear compensation is performed in between the dispersion compensation. It is called Wiener-Hammerstein model or symmetrical model. Dispersion can alternatively be compensated fully upfront or after the linear operator.

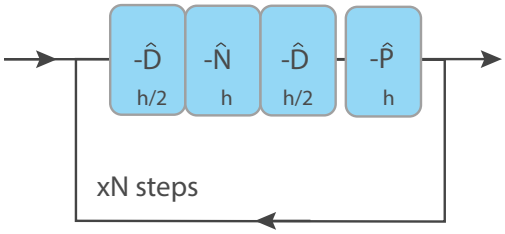


Figure 2.1: DBP block scheme. The received signal power is normalised, backpropagated through all linear and nonlinear steps of DBP, between which the signal power is adjusted.

The linear operator in frequency domain, using the property of a Fourier transform of a second derivative ($\mathfrak{F}\{\frac{d^2k(x)}{dx^2}\} = -(2\pi f)^2 \cdot K(X)$, where $k(x)$ is a function in time domain and $K(X)$ is that function in frequency domain), is written as:

$$\mathfrak{F}\{-\hat{D}\} = \mathfrak{F}\{\frac{i}{2}\beta_2\frac{\partial^2}{\partial t^2} + \frac{\alpha}{2}\} = -\frac{i}{2}\beta_2\Delta\omega^2 + \frac{\alpha}{2}, \tag{2.14}$$

where $\mathfrak{F}$ is the Fourier transformation, f is the optical frequency. Linear operator is applied as:

$$A_{x,y}(z+h,t) = \mathfrak{F}^{-1}\{\mathfrak{F}\{A_{x,y}(z,t)\}e^{-ih\beta_2\Delta\omega^2/2}\}, \tag{2.15}$$

and nonlinearities are compensated in time domain:

$$A_{x,y}(z+h,t) = A_{x,y}(z,t)e^{-i\frac{8}{9}h\gamma\left(\left|A_{x,y}\right|^2+\left|A_{y,x}\right|^2\right)}. \tag{2.16}$$

The signal power is updated in time domain as:

$$A_{x,y}(z+h,t) = A_{x,y}(z,t)e^{(h\alpha/2)}. \tag{2.17}$$

2.1.4 Historical development of DBP

The concepts of backward propagation have been reported as early as 1996 [40], where in optical domain a dispersive medium with a negative nonlinear refractive index is used for dispersion and nonlinearities compensation. In numerical derivations [41] backward propagation was used for reversing femtosecond pulse propagation in an optical fibre. The concept consolidated in 2006 [42, 43]. In the digital domain, first important experimental results are obtained in 2008 with the use of coherent receivers [44, 45]. 20 years after its introduction DBP is considered a promising mathematical tool to address the nonlinearity mitigation, however, only if the high numerical complexity constrain is overcome.

2.1.5 DBP modifications

Numerous DBP modifications have been proposed in the recent years in order to reduce the prohibitively high computational complexity of the standard algorithm. In order to support real-time DBP implementation most of the modifications aim at increasing the DBP step size, h.

One of the earliest proposed modifications involves applying a non-uniform step size, for example, logarithmic as proposed in [46, 47]. The idea is based on the fact that in a link with EDFAs the power in the beginning of a span is the highest and drops exponentially along the span. As nonlinear effects are proportional to the optical power density, it is reasonable to perform shorter DBP steps in the beginning of the span and less frequent steps in the rest of the span. Interestingly, this approach was long used in simulations of the forward propagation to save computational time, before it was also adopted for DBP. Another example of the forward propagation simulation methods that was adopted for DBP is a walk-off method for WDM systems [48], which chooses a step size smaller than a characteristic walk-off length, L_w, a distance after which the relative delay between the edge channels is comparable with the pulse width: $L_w = \tau_p/\Delta\beta_{1-2}$, where τ_p is the pulse duration, $\Delta\beta_{1-2}$ is the difference in propagation constants of the edge channels. In this method, dispersion slope β_3 needs to be considered. A further example is a non-uniform step size modification, maximum phase change (MaxPhCh) algorithm [49], which makes sure that in every step the nonlinear phase rotation does not overshoot a certain set value.

Study [50] suggests that for non-DM optical links optimisation of the linear operator can reduce the number of steps by two, still keeping the performance level. In this work complexity reduction as high as 85% was demonstrated for DM-systems with residual dispersion .

Another commonly used DBP modification is called filtered DBP [51]. It is based on the observation that after each DBP step the high-frequency intensity fluctuations introduce additional noise. The problem can be overcome by introducing a low-pass filter (LPF) after every nonlinear operator. In that way the DBP is optimised for compensating the low-frequency intensity fluctuations. The actual optimization process according to the number of steps is a non-trivial task including fine-tuning the LPF bandwidth, fibre coefficients,

and NL operator position relatively to the linear operator. Generally, the more steps DBP has, the larger LPF bandwidth should be.

A very promising DBP modification was proposed in [52–54], called correlated DBP. Because of the pulse broadening due to dispersion, the self-phase modulation (SPM) impact on one symbol depends not only on that symbol power but also on the neighbouring symbols. When calculating the nonlinear phase shift the correlation between the neighbouring symbols is taken into account by applying a weighted time-domain filter. This approach leads to up to 80% reduction reduction in required back-propagation steps (compared to standard DBP) without performance loss for a single channel DP-QPSK transmission [52].

An interesting approach of *optical* back propagation for real-time implementation was proposed in [55] using dispersion compensating fibre (DCF) for the linear effects compensation and highly nonlinear fibre – for nonlinear.

2.1.6 Digital back propagation used in the thesis simulations

For simulations in this work we always use ideal DBP, i.e. not limited by complexity constraints or receiver bandwidth. Our aim is to separate DBP degradations arising from stochastic propagation effects, such as PMD, ASE and PDL, and therefore we implement the algorithm in its most accurate version, where DBP is performed in optical domain before the receiver (Rx) using the same model as the one for forward propagation, only inverted. The number of steps is comparable.

We use the symmetrical DBP model in simulations depicted in Figure 2.1. The step size is controlled via the MaxPhCh method, where it is chosen adaptively during the propagation along the fibre so that in each step the nonlinear phase of the optical field is changed by no more than a certain value in degrees. According to [49] 0.05° is chosen for all the simulations if not specified otherwise. The step length is calculated as:

$$h \leqslant \frac{MaxPhCh}{\gamma \left|A_p\right|^2}, \tag{2.18}$$

where A_p is the peak value of the input optical field of the step.

2.2 ASE noise in EDFAs

In an EDFA, the silica glass is doped with rare earth Erbium (Er) ions, which adds an additional atom energy level to the silica energy band structure [56]. The fibre can be pumped with an appropriate wavelength laser (980 nm) to excite electrons to the new energy level and thereby to create the so-called electron population inversion, as in Figure 2.2. The lifetime of an electron in the excited state, τ_{sp}, is short, after a few μs it will relax to the metastable lower-energy level without emitting radiation. If such an electron is stimulated with an incident photon, which energy matches the energy gap between the metastable and the ground state, it will relax to the ground state while emitting in the wavelength window around 1550 nm. The nature of the stimulated emission ensures that the new photons have exactly the same properties as the incident ones, thus amplifying the incident light.

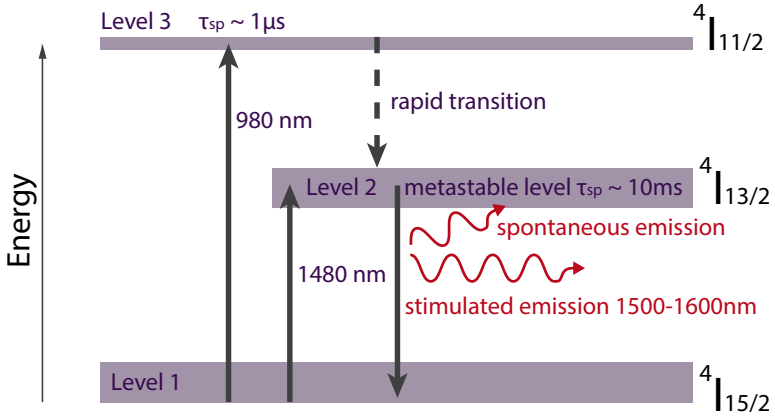


Figure 2.2: Energy levels and principle of EDFA.

However, some of the excited electrons will relax to the ground energy level spontaneously. Such emission is not in-phase with the incident light and can propagate in any direction in space. A part of such photons will be emitted in direction that falls within the numerical aperture of the fibre and will continue propagating with the optical signal. Moreover, this spontaneous emission will be further amplified along the EDFA. Such emission is called amplified spontaneous emission.

The noise performance of amplification can be measured by the noise figure

(NF), a fraction of input and output linear OSNRs in dB [57]:

$$NF = 10log_{10}\frac{OSNR_{lin,in}}{OSNR_{lin,out}}\ [dB], \tag{2.19}$$

where $OSNR$ is expressed in linear units:

$$OSNR_{lin} = \frac{P_s}{P_n}, \tag{2.20}$$

P_s and P_n are the signal and noise power respectively.

2.3 Polarisation effects in SSMF

2.3.1 Birefringence in SSMF

Light propagation in an ideal SSMF can be described with a high level of accuracy, when the initial optical field as well as the fibre characteristics are known. An optical wave can be represented as a linear superposition of two orthogonally polarised HE_{11} modes. When a fibre is ideal these modes cannot be distinguished from each other. However, when the fibre is produced, laid in cable and exploited in real life conditions, the anisotropy gets lost, which results into the two distinct polarisation modes HE_{11} having different fibre propagation constants. This effect is called birefringence.

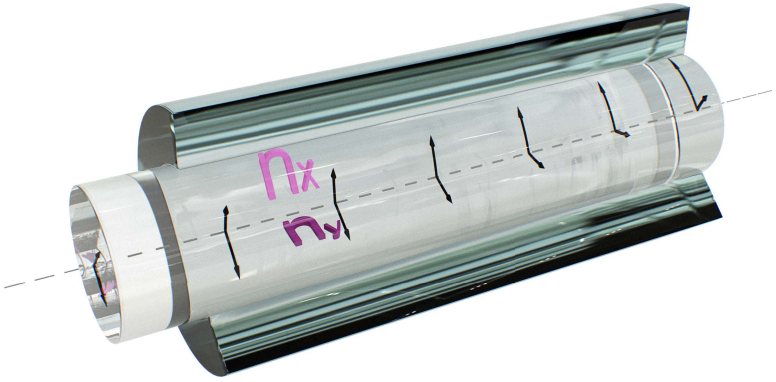


Figure 2.3: Evolution of the fast and the slow birefringence axes along the fibre. $n_x = n_e$, $n_y = n_o$.

A propagation constant consist of a real and imaginary parts and is defined as a sum of attenuation and phase constants: $\alpha + i\beta$ (see Section 2.1.1). Different propagation constants of the modes mean they have each their own

group and phase velocities, attenuation, and refractive indices. The latter are typically denoted as n_o (slow) and n_e (fast) for ordinary and extraordinary refractive indices respectively. The bigger the refractive index is, the slower light is propagating in the according axis, i.e. $n_o > n_e$ by definition. Typical values of $\Delta n_{eff} = n_o - n_e$ are between 10^{-7} and 10^{-5}. The fast and the slow axes in a fibre depend on the source of birefringence and are continuously changing along the fibre, as illustrated in Figure 2.3.

2.3.2 Sources of birefringence in SSMF

Birefringence at the stage of fibre production arises from statistical changes in the manufacturing process, like the soot chemical composition or fibre drawing, which leads to fibre stress in the core or cladding, loss of core alignment, fibre asymmetry, material imperfections and accidental doping. Besides, the mechanical impacts on the fibre during installation, laying into the sewage system or connecting couplers introduces extra birefringence. The mechanical effects include outer pressure, bending and twisting. The material stress caused by temperature change in the exploited fibre is also a known source of birefringence. The above is summarised in Figure 2.4.

Due to these effects a long fibre is characterised by having a random birefringence profile in space, time and frequency dimensions. In order to work with the birefringent properties of a fibre, statistical approach should be applied.

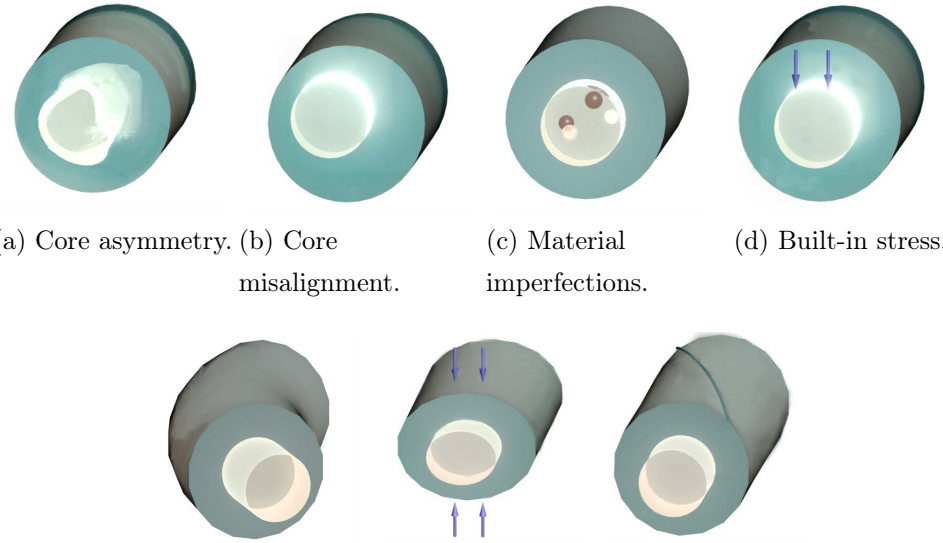
Birefringence giving rise to two distinct propagation modes leads to various polarisation effects in a fibre, such as PMD, PDL and PDG.

2.3.3 PMD

PMD statistics in short-length fibres

In a short-length fibre the slow and fast axes are assumed to stay constant along the fibre [58]. The net effect of a fibre having two propagation modes with different group velocities is then a group delay between the polarisation modes. The delay is known as DGD, is denoted as τ, measured typically in ps, and can be defined as:

$$\tau = \frac{L}{v_o} - \frac{L}{v_e}, \tag{2.21}$$



(a) Core asymmetry. (b) Core misalignment. (c) Material imperfections. (d) Built-in stress.

(e) Bending. (f) External pressure. (g) Twisting.

Figure 2.4: Sources of birefringence in a fibre. Internal: core asymmetry, core misalignment, material imperfections, built-in stress, and external: bending, external pressure, twisting.

where L is the fibre length, υ_o and υ_e are group velocities along the slow and fast axes respectively. Reformulating the latter expression as

$$\tau = PMDc_{slow} \cdot L, \tag{2.22}$$

we can introduce a fibre PMD proportionality coefficient:

$$PMDc_{slow} = \frac{1}{\upsilon_o} - \frac{1}{\upsilon_e} \tag{2.23}$$

in ps/km, that characterizes the DGD arising from fibre birefringence. Experimentally the value can be measured by the difference of slow and fast group refractive indexes: $PMDc_{slow} = \Delta n_{eff}/c$, where c is light speed in vacuum.

The effect of the DGD on a linearly polarised light pulse depends on its polarisation orientation towards the polarisation axes of the fibre. The time-domain evolution of the pulse that is 45° tilted to fast and slow fibre axes can be found in the schematic in Figure 2.5. The figure shows a group of two images at five different time snaps. The two images are depicting the pulse envelope (Figure 2.5b, 2.5d, 2.5f, 2.5h, 2.5j) and its projection on the slow and fast axis

separately (Figure 2.5a, 2.5c, 2.5e, 2.5g, 2.5i). In a general case the pulse can be oriented at an angle to the fast and slow axis. The pulse then splits into two polarisation components that propagate with different group velocities. The mismatch between the two pulse components leads to the change of the pulse polarisation state and shape until the DGD reaches the full pulse duration when the two orthogonally polarised pulses split.

The above description contains a simplification that the fast and slow axes do not rotate along the fibre. In reality they are changing the orientation and while at certain short ranges they can be approximated as constant, at longer ranges rotation should be taken into account. These two PMD regimes in a fibre are called, respectively, short-length and long-length, and the fibre length on the border between the two regimes is referred to as correlation length or coupling length [58]. Imagining a linearly polarised light inserted in a fibre, the correlation length l_c is defined as the distance where the *average* power in that initial polarisation state $P_{\parallel}$ becomes comparable to the *average* power that leaked to the orthogonal polarisation $P_{\perp}$, i.e. when the average power is equalized between the two polarisation modes. The correlation length is strictly defined as [58]:

$$< P_{\parallel}(l_c) > - < P_{\perp}(l_c) >= \frac{1}{e^2} \cdot P_s, \tag{2.24}$$

$< >$ denoting the averaging operator, P_s is the total signal power. According to [59] the correlation length varies dramatically depending on the fibre exploitation conditions from less than 1 m up to more than 1 km with a typical value of 100 m.

DGD statistics in long fibres

For the reasons discussed in Section 2.3.2 a long-length fibre experiences rotation of birefringence axis. This rotation can either contribute to or counteract the accumulation of the group delay. As a result, the average DGD, $< \tau >$, also called PMD, in long fibres accumulates in a random-walk manner and is proportional to the square root of the fibre length (compare with the linear length dependence for short-length fibres) [58]:

$$DGD = \tau, \tag{2.25}$$

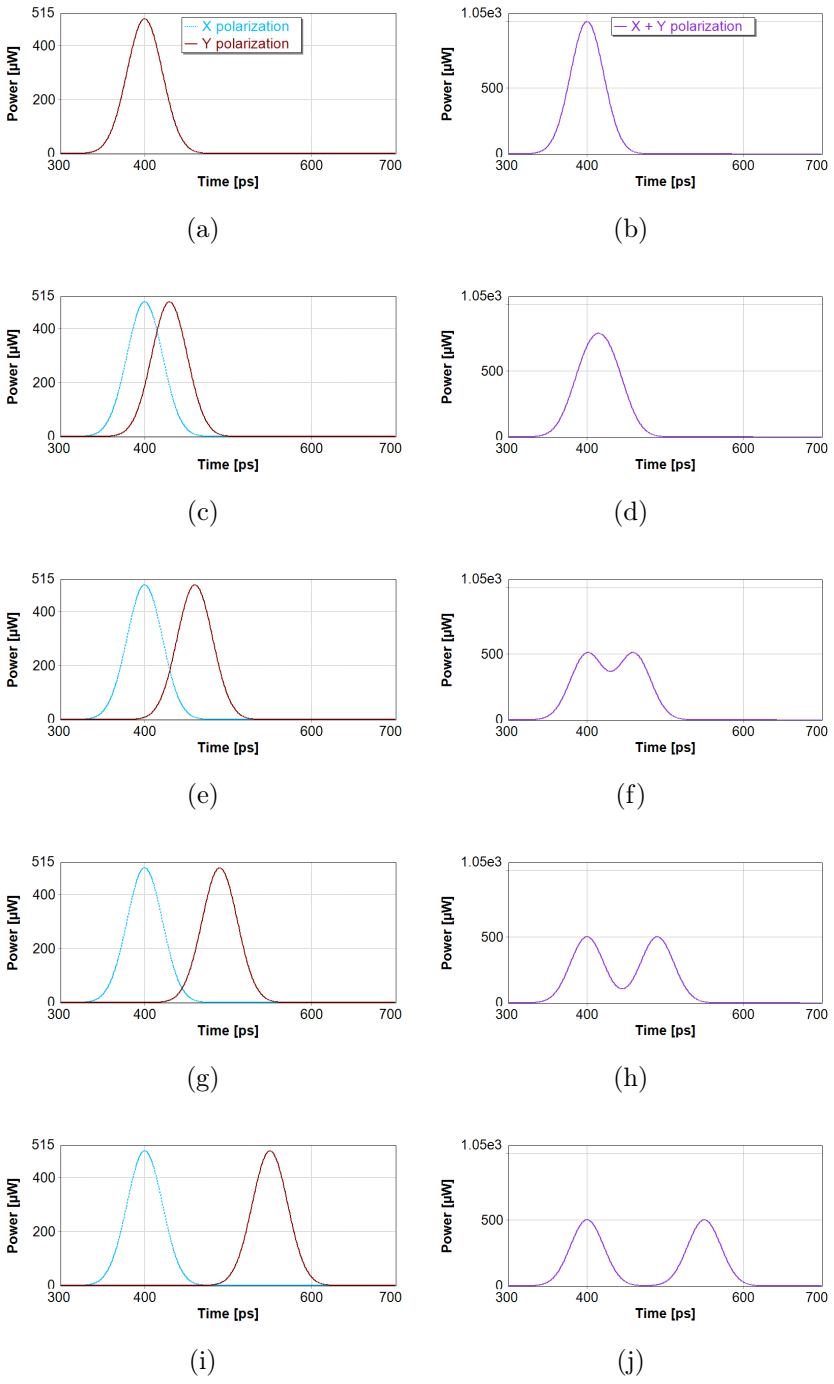


Figure 2.5: Time-domain effect of DGD caused by birefringence on a pulse polarised at 45° to the fast and slow fibre axes. Figures in the second column (b), (d), (f), (h), (j) depict the pulse envelope, and figures in the first column (a), (s), (e), (g), (i) – its projection on the slow and fast axis separately.

$$PMD =< \tau >= PMDc \cdot \sqrt{L}. \tag{2.26}$$

The proportionality coefficient $PMDc$ is now defined in $ps/\sqrt{km}$.

The DGD in a long-length fibre has a Maxwell distribution, defined as:

$$\mathbb{P}(\tau) = \frac{32\tau^2}{\pi^2 < \tau >^3} \cdot exp(-\frac{4\tau^2}{< \tau >^2}/\pi), \tag{2.27}$$

where $\mathbb{P}$ is the occurrence probability density function with

$$\int_0^\infty \mathbb{P}(\tau)d(\tau) \;=\; 1. \tag{2.28}$$

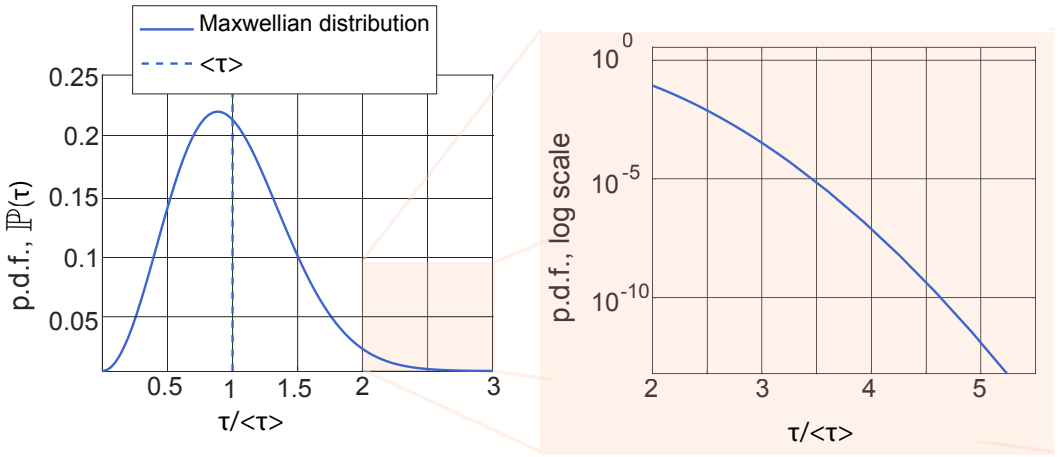


Figure 2.6: Maxwellian distribution of DGD p.d.f. (probability density function) in a long-length fibre.

Figure 2.6 illustrates that in a long-length fibre with a certain PMDc, the output DGD can take a wide range of values. For example, the probability of τ falling into the interval between 0.5 to 1.5 $< \tau >$ is equal to approximately 75%. Regarding the system outage, the most harmful consequence of statistical nature of DGD is that there is a non-zero probability of exceeding the PMD value more than 3-4 times. Such probability is generally defined as:

$$\mathfrak{P}(\tau)\Big|_{\tau \geqslant \left(\mathcal{S}\cdot<\tau>\right)} = \int_{\mathcal{S}\cdot<\tau>}^{\infty} \mathbb{P}(\tau) \cdot d(\tau), \tag{2.29}$$

where $\mathcal{S}$ is the ratio of DGD to PMD, also called safety factor:

$$\mathcal{S} = \tau / < \tau > . \tag{2.30}$$

The calculated probabilities and corresponding duration per year are summarized in Table 2.1.

Table 2.1: The probability of τ being at least $\mathcal{S}$ times larger than the $<\tau>$.

$\mathcal{S}$	$\mathfrak{P}$	**Duration per year**
2.0	1.4e-2	5.4 days
2.5	8.5e-4	7.5 h
3.0	3.8e-5	20 min
3.5	7.5e-7	23 sec
4.0	5.6e-9	less than 1 sec

Probability of DGD being in a certain interval is defined as:

$$\mathfrak{P}(\tau)\Big|_{\tau_1 \leqslant \tau \leqslant \tau_2} = \int_{\tau_1}^{\tau_2} \mathbb{P}(\tau) \cdot d(\tau), \tag{2.31}$$

where τ_1 and τ_2 are the lower and upper interval borders.

2.3.4 PDL and PDG

Attenuation of the propagated light in passive optical components can vary depending on the input polarisation states. PDL is defined as the ratio of the peak-to-minimum transmission intensity given all possible input states of polarisation in logarithmic scale:

$$PDL = 10log_{10}\frac{\Gamma_{max}}{\Gamma_{min}}\ [dB], \tag{2.32}$$

where Γ_{max} and Γ_{min} are the maximum and minimum light transmission coefficients through the fibre or another passive optical component.

PDL of a fibre link has a Rayleigh distribution, which transforms into a Maxwell distribution in presence of PMD [60–62]. The Rayleigh distribution is defined as:

$$\mathbb{P}(PDL) = \frac{PDL}{<PDL>^2}\exp(-\frac{PDL^2}{2<PDL>^2}), \tag{2.33}$$

where PDL is in dB, $<PDL>$ is the average PDL in dB, and $\mathbb{P}$ is the probability density function.

The combination of PMD and PDL effects can have a bigger impact a system than each of those separately. In the presence of PDL the signal states of

polarisation lose their orthogonality during propagation [63], as schematically explained in Figure 2.7.

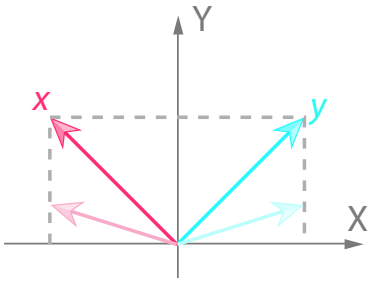


Figure 2.7: Loss of orthogonality between signal's x and y polarisations by PDL (here shown for a relative loss along the Y axis only).

PDG is defined similarly to PDL. It is observed in optical amplifiers (in semiconductor optical amplifiers (SOAs) and EDFAs). PDG is defined as a ratio of maximum G_{max} and minimum G_{min} gain over all possible input polarisations in logarithmic scale:

$$PDG = 10log_{10}\frac{G_{max}}{G_{min}}\ [dB]. \tag{2.34}$$

2.3.5 PSP notation in Stokes space

The principal state of polarisation (PSP) is usually defined in Stokes space as four parameters combined into a Stokes vector [64]:

$$\mathbb{S} = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} S_0 \\ S_0 cos(2\phi) \\ S_0 sin(2\phi) cos(\zeta) \\ S_0 sin(2\phi) sin(\zeta) \end{pmatrix}, \tag{2.35}$$

where S_0 corresponds to light intensity and S_{1-3} describe the polarisation state of the electromagnetic light wave, as in Figure 2.8 and Figure 2.9.

Considering PMD, the PSP axes correspond to the slow and the fast propagation axes. The time difference in propagation along the PSP is called DGD, as introduced in Section 2.3.3.

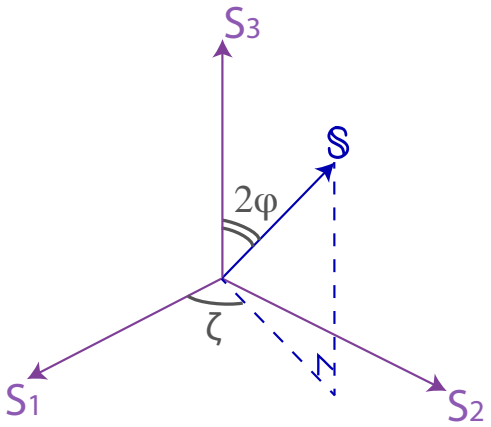


Figure 2.8: Azimuth ζ and elevation 2ϕ visualization on the Poincare sphere.

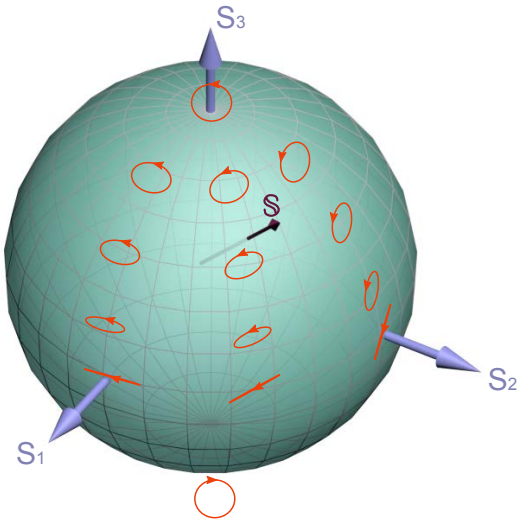


Figure 2.9: Poincare sphere for polarisation state visualization.

2.3.6 Principal States model and Jones notation

The principal state model proposes that a fibre without PDL in the linear regime can be represented as a frequency-dependent channel transfer matrix, called Jones matrix [64, 65]:

$$H(\omega) = e^{\alpha(\omega)} \begin{bmatrix} u_1(\omega) & u_2(\omega) \\ -u_2^*(\omega) & u_1^*(\omega) \end{bmatrix}, \tag{2.36}$$

where $\alpha(\omega)$, $u_1(\omega)$ and $u_2(\omega)$ are complex quantities and $|u_1|^2 + |u_2|^2 = 1$, ω is the signal angular frequency. The complex optical field at the output of the fibre is then:

$$E^{out}(\omega) = H(\omega)E^{in}(\omega), \tag{2.37}$$

where $E^{in,out} = [E_{x,y}^{in,out}, E_{x,y}^{in,out}]^T$.

The principal states model is based on the observation that for every $H(\omega)$ written as in Equation (2.36) at every frequency there is a pair of orthogonal polarisation states, such that if the input signal is aligned with one of those states, the output is invariant to changes in frequency to the first order. Light launched in a fibre in such polarisations will maintain its polarisation at the output to the first order in frequency. These states of polarisation are called principal states of polarisation (PSPs).

In [66, 67] it was shown that Equation (2.36) can describe the fibre channel cross-coupling between both polarisations, including dispersion, $\mathbb{D}(\omega)$, PDL $\mathbb{K}(\omega)$ and PMD $\mathbb{U}(\omega)$ effects:

$$H(\omega) = \mathbb{D}(\omega)\mathbb{U}(\omega)\mathbb{K}(\omega), \tag{2.38}$$

where

$$\mathbb{D}(\omega) = e^{-iL\beta_2\omega^2/2}, \tag{2.39}$$

β_2 being the group velocity dispersion, L fibre length;

$$\mathbb{U}(\omega) = R_1^{-1}(\omega) \begin{pmatrix} e^{i\omega\tau(\omega)/2} & 0 \\ 0 & e^{-i\omega\tau(\omega)/2} \end{pmatrix} R_1(\omega). \tag{2.40}$$

The central matrix describes the group delay between the PSPs, τ, in the so-called Jones notation [64]. $R_1(\omega)$ is a unitary matrix rotating signal x and

y polarisations into X and Y PSP axes in Stokes space (detailed description can be found in Appendix A):

$$R_1(\omega) = \begin{pmatrix} \cos\zeta(\omega) & -e^{-i\phi(\omega)}\sin\zeta(\omega) \\ e^{i\phi(\omega)}\sin\zeta(\omega) & \cos\zeta(\omega) \end{pmatrix}, \tag{2.41}$$

where ζ and 2ϕ are the azimuth and elevation rotation angles (see Figure 2.10).

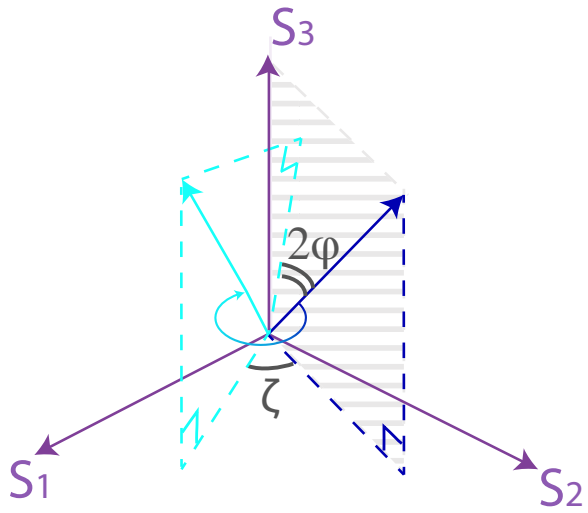


Figure 2.10: Poincare sphere for visualizing a signal state of polarisation.

Finally,

$$\mathbb{K}(\omega) = R_2^{-1}(\omega) \begin{pmatrix} \sqrt{\Gamma_{max}(\omega)} & 0 \\ 0 & \sqrt{\Gamma_{min}(\omega)} \end{pmatrix} R_2(\omega), \tag{2.42}$$

where Γ_{max} and Γ_{min} are the maximum and minimum transmission coefficients and $R_2(\omega)$ is a unitary matrix, similar to Equation (2.41), rotating signal x and y polarisations into PDL eigen modes.

2.3.7 Vectorial notation in Stokes space

To describe the PSP, PDL and the DGD in a fibre, it is convenient to introduce the vectorial notation. The PMD vector is:

$$\mathbb{O}(\omega) = \tau(\omega) \cdot q(\omega), \tag{2.43}$$

where $q(\omega)$ is a unit Stokes vector, aligned with the slow PSP, $\tau(\omega)$ is a scalar, the PMD vector length.

The PDL vector is less commonly used and is defined as:

$$\mathbb{G}(\omega) = \frac{\Gamma_{max}}{\Gamma_{min}} \cdot p(\omega), \tag{2.44}$$

where Γ_{max} and Γ_{min} are the maximum and minimum transmission coefficients (scalar) and $p(\omega)$ is a unit vector, aligned with maximum transmission direction.

Introducing the PMD vector allows describing higher order PMD, when the PMD vector depends on the frequency. (2.43) can be expanded in the Taylor series as in [65]:

$$\mathbb{O}(\omega) = \mathbb{O}(\omega_0 + \Delta\omega) \approx \mathbb{O}(\omega_0) + \frac{\partial\mathbb{O}(\omega_0)}{\partial\omega}\Delta\omega + \frac{1}{2}\frac{\partial(\mathbb{O}(\omega_0))^2}{\partial^2\omega}\Delta\omega^2 + ..., \tag{2.45}$$

where $\Delta\omega = \omega - \omega_0$, and ω_0 is the centre frequency, PMD vector is defined at frequency ω.

$$\mathbb{O}(\omega_0) = \tau(\omega_0) \cdot q(\omega_0) \tag{2.46}$$

is the first-order PMD vector and

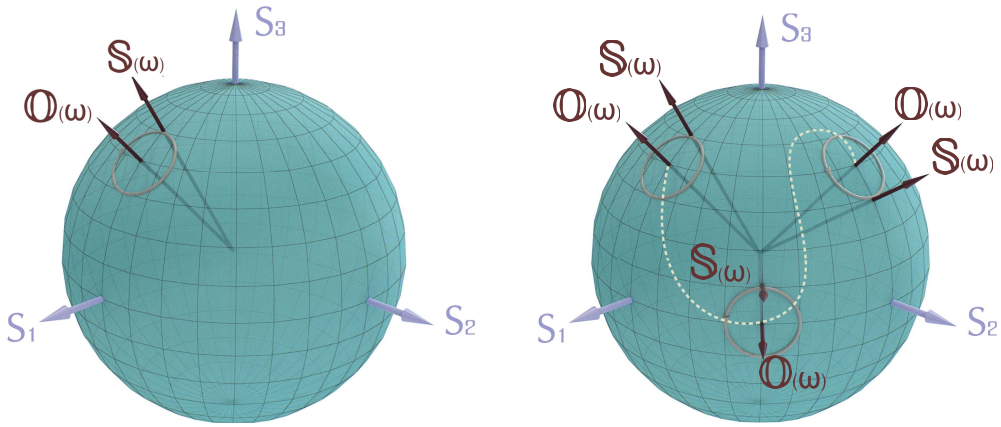
$$\frac{\partial\mathbb{O}(\omega_0)}{\partial\omega} = PCD \cdot q(\omega_0) + PSD \cdot q_\perp(\omega_0) \tag{2.47}$$

is the second order PMD vector, with polarisation-dependent chromatic dispersion (PCD) and polarisation state depolarisation (PSD). $q(\omega_0)$ is the unit Stokes vector pointing in the direction of the PMD vector at the reference frequency and $q_\perp(\omega_0)$ is the orthogonal unit Stokes vector. The second-order PMD magnitude is then described by $\sqrt{PCD^2 + PSD^2}$.

The frequency dependence of signal states of polarisation $\mathbb{S}$ is described as [68]

$$\frac{\partial\mathbb{S}(\omega)}{\partial\omega} = \mathbb{O}(\omega) \times \mathbb{S}(\omega). \tag{2.48}$$

In the Stokes space the PMD vector offers a clear graphical interpretation of the PMD effect on a signal, depicted in Figure 2.11. In case of first-order PMD this results in the signal Stokes vector precession around the PMD vector. When higher order PMD is taken into account, the PMD vector itself starts a complex movement in Stokes space.



(a) First-order PMD. (b) Higher-order PMD.

Figure 2.11: Visual representation of a signal polarisation state change over frequency under effect of PMD.

2.4 Simulation of stochastic effects

2.4.1 ASE noise simulation

The amplifier used in the simulations is an ideal model of an EDFA: It is gain-controlled to match the span loss exactly. Such effects as gain saturation, gain frequency profile, gain power-dependence are not considered. The amplifier model serves for two functions – exact compensation for the power loss and introduction of ASE noise. The ASE noise is modelled by a unpolarised Gaussian distributed noise added independently to both polarisation components. For PMD and PDL simulations, ASE-NL interactions are excluded: EDFAs are noiseless, noise is loaded right before the receiver.

2.4.2 PMD simulation

Coarse-step model

The most common way to simulate optical fibre polarisation resolved signal propagation is to solve a system of two coupled Manakov equations for both orthogonal polarisations using the SSF [69]. The default method to simulate the polarisation effects in a fibre is the coarse-step approximation [70] illustrated in Figure 2.12 (the model is also referred to as waveplate model). The fibre in the model is divided into sections of constant birefringence with a

length chosen randomly from a normal distribution, that are alternated with polarisation scramblers. The length of each section should be longer than the fibre correlation length (see Equation (2.24)). The coupled Manakov equation is solved for each section, which introduces a DGD τ_i. The length of the sections needs to be chosen randomly out of a Gaussian distribution to provide an accurate statistical PMD distribution to the first and second order, close to the one measured experimentally [71]. The multiplication of the signal at different frequencies with a fixed DGD would induces a periodicity in its autocorrelation function, and the PMD characteristics in the frequency domain would therefore repeat in long simulations.

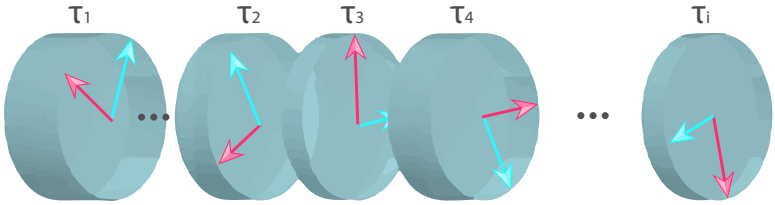


Figure 2.12: Coarse-step PMD simulation method schematic.

If the number of coarse-step sections is finite, only a certain set of DGDs can appear (resulting from the combination of the waveplates' delays). The occurrence ratio of the DGD, especially high values of DGD, should then be weighted with the Treloar formula [72]:

$$\mathbb{P}(\tau, N_{c-s}) = \frac{2\tau}{\pi} \int_0^\infty \sigma \cdot sin(\sigma\tau) \cdot \prod_{i=1}^{N_{c-s}} sinc(\sigma\tau_i) \cdot d\sigma, \tag{2.49}$$

where N_{c-s} is the number of coarse-step sections, τ_i is the DGD of the i-th section, $\prod$ denotes product operator, $sinc$ is the sine cardinal function, and σ is the frequency of the Fourier transform of the DGD probability density function:

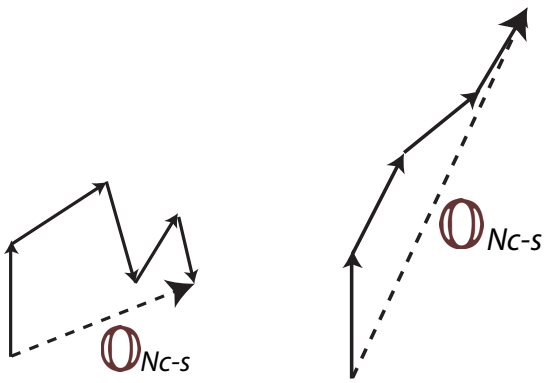
$$\mathfrak{F}\{\mathbb{P}(\tau)\} = \int_0^\infty 4\pi\tau^2 \mathbb{P}(\tau) \cdot e^{i\sigma\tau} d\sigma. \tag{2.50}$$

Importance sampling

In stochastic simulations it can be extremely time-consuming to encounter cases with a high value of DGD or PDL because these are not probable (See

Section 2.3.5). High numerical values, however, have the largest impact on the system performance and can cause system outage. For optical long-haul links typically allowed outage probability is below 10^{-5}, corresponding to 99.999% ('five nines') in system availability [73]. In time it corresponds to 5 minutes per year.

In order to increase the probability of getting high DGD and PSP values in a simulation, importance sampling (IS) can be used. One way of applying IS to simulate PMD is to modify the way the polarisation scrambling is done in the coarse-step model. The angle of the new polarisation can be chosen randomly, or alternatively – aligned, so that the direction of the new PMD vector is close to the previous value [74]. The aligned PMD vectors tend to result in higher DGD values and therefore allow saving time compared to the brute-force random simulations. The idea is illustrated in Figure 2.13.



(a) Unbiased case.

(b) Biased polarisation rotation angle between the PMD sections.

Figure 2.13: Visual representation of the resulting PMD vector $\mathbb{O}_{N_{c-s}}$ in the coarse-step model.

The resulting PMD vector after concatenation of N_{c-s} course-steps sections can be written as:

$$\mathbb{O}_{N_{c-s}} = \Delta\mathbb{O}_{N_{c-s}} + R_{c-s}\mathbb{O}_{N_{c-s}-1}, \tag{2.51}$$

where $\Delta\mathbb{O}_{N_{c-s}}$ is the contribution of the current section, $\mathbb{O}_{N_{c-s}-1}$ is the concatenated PMD vector of the previous sections, and R_{c-s} is the rotation matrix (polarisation scrambler) before the current section. The length of the $\mathbb{O}_{N_{c-s}}$

depends on the alignment of $\Delta\mathbb{O}_{N_{c-s}}$ and $R_{c-s}\mathbb{O}_{N_{c-s}-1}$, or strictly, on the $cos\ \psi$ of the ψ angle between them. In the unbiased case $cos\ \psi$ takes equally probable values between -1 and 1. Using importance sampling technique, it can be biased to have the following distribution:

$$cos\ \psi = 2\xi^{1-\Xi} - 1, \tag{2.52}$$

where ξ is a random number between 0 and 1, and Ξ is the bias parameter between 0 and 1. The closer is Ξ to 1, the closer would PMD vectors of the consecutive sections align.

With IS the probability of the PMD becomes biased and has to be weighted using the weighting factor K [75]:

$$K(cos\ \psi) = \frac{\mathfrak{P}(cos\ \psi)}{\mathfrak{P}^*(cos\ \psi)} = 0.5/\Big(\frac{1}{2(1-\Xi)}\big(\frac{cos\ \psi + 1}{2}\big)^{\frac{\Xi}{1-\Xi}}\Big), \tag{2.53}$$

where $\mathfrak{P}(cos\ \psi)$ is the unbiased, and $\mathfrak{P}^*(cos\ \psi)$ is the biased probability of $cos\ \psi$.

Preselection

When simulating links operating in highly-nonlinear regime, each transmission simulation can take considerable time. It gets even longer with growing signal power, provided the accuracy is fixed. Besides, the probability to encounter a fibre configuration with high value of DGD or PDL is low.

In order to minimise the total simulation time, one can preselect the fibre birefringence profiles (combination of coarse-step section length and polarisation scramblers' angles) that result in high DGD or PDL. The simulation of the fibre polarisation effects can be done disregarding dispersion or Kerr effects. Therefore the time of such estimation is short. The estimation can be done using low-power optical impulses propagating in orthogonal polarisations that will provide the Jones matrix (channel transfer matrix) of the fibre, from which PDL, PDG, PMD, PSP and DGD parameters can be extracted [66, 67]. Finally, when the fibres are preselected, the precise simulations can be performed. In the thesis we combine both IS and preselection for coarse and fine selection of birefringence profiles respectively.

Higher-order PMD

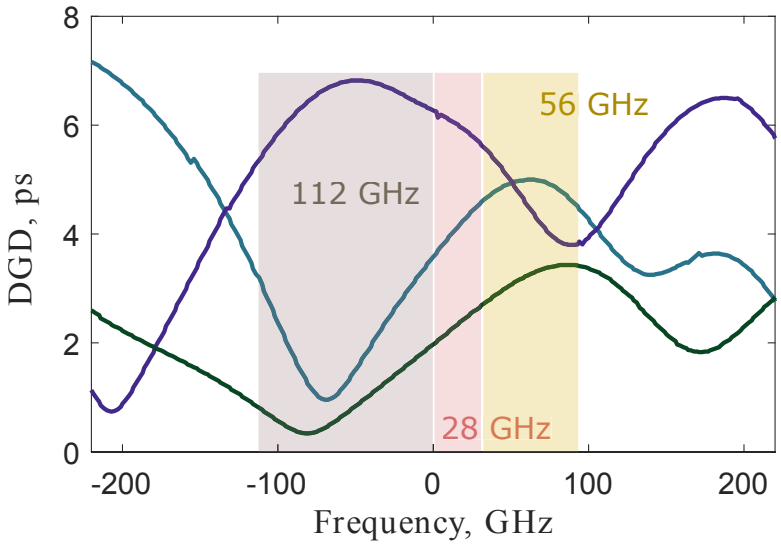


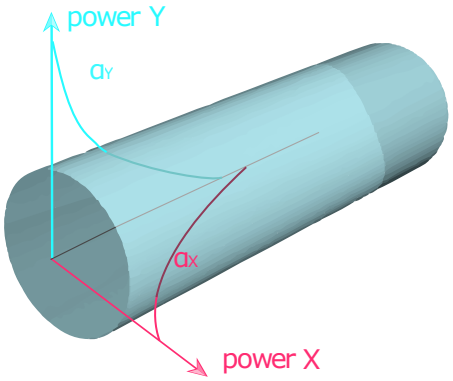
Figure 2.14: Frequency dependence of DGD.

Polarisation effects are wavelength-dependent, but when the signal bandwidth is relatively small, in a coarse approximation, the frequency-independent DGD can be considered. However, it should always be checked if second-order PMD modelling is required. Figure 2.14 illustrates the frequency dependence of DGD spectra overlaid with 28, 56 and 112 GHz signal spectra for three randomly chosen 2400 km SSMF 0.1 $ps\sqrt{km}$ fibre birefringence realizations and central frequency of 193.0985 THz. PMD was simulated using a coarse-step model with an average of 100 m long waveplates.

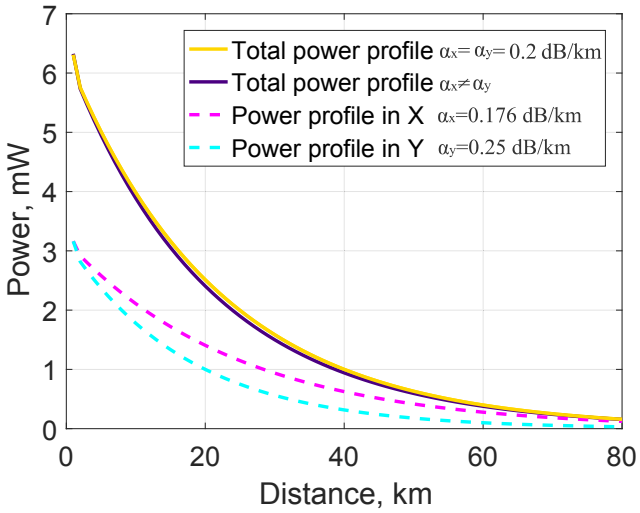
2.4.3 PDL simulation

In order to achieve large PDL or PDG values, similar techniques as for PMD can be used. By performing IS, the angle between the channel and PDL or PDG-device polarisation states can be adjusted [76]. This angle biased to zero would lead to the largest performance penalties. Bearing in mind the two main sources of PDL and PDG, fibre-intrinsic and from the link components, we simulate them accordingly in two ways, as explained in Figure 2.15.

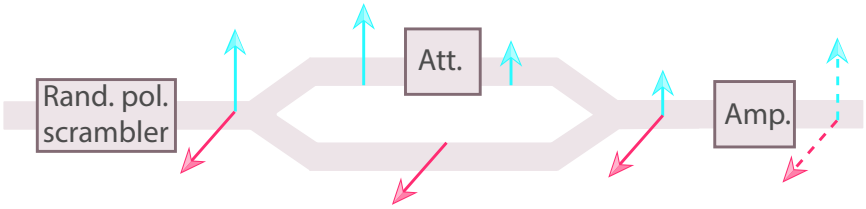
In the first approach, the fibre's slow and fast axes are having different attenuation coefficients, Figure 2.15a and Figure 2.15b. In Figure 2.15b a signal power profile of a 80 km SSMF span is illustrated, the signal input power



(a) First approach: Fibre's propagation axes have different attenuation coefficient.



(b) First approach: Example of power profiles in a span with same (yellow line) or different (purple, pink, blue lines) attenuation coefficients in X and Y axes.



(c) Second approach: PDG in EDFAs.

Figure 2.15: Illustration of intrinsic PDL and component PDG simulation approaches.

is 8 dBm. The case where both polarisations have 0.2 dB/km attenuation coefficient is compared with the case, where $\alpha_x = 0.176$ dB/km and $\alpha_y = 0.25$ dB/km The attenuation coefficients are adjusted so that the total power along the span has the same shape as if the attenuation coefficients were equal. The second approach in Figure 2.15c investigates PDG in optical components, namely in the EDFAs, as a combination of a random polarisation scrambler (implemented as a random polarisation axes rotation) and attenuation of one of the polarisation components. In the result, one of the randomly scattered polarisation components is amplified more strongly than the other.

2.5 Conclusions

In this chapter, we covered the background knowledge on fibre birefringence and polarisation effects, and the DBP algorithm. Further, the basis for simulating polarisation effects and their impact on the DBP was introduced.

Chapter 3

Statistics of stochastic effects and their impact on DBP

This chapter will give an overview of the statistics of the stochastic effects simulated in a fibre and EDFAs, namely, ASE, PMD, and PDL, which were performed in *VPItransmissionMaker Optical Systems*. The impact of these effects on DBP gain will then be discussed. Finally, experimental results on the PMD-induced DBP limit will be presented.

3.1 Statistics of stochastic effects

3.1.1 EDFA ASE noise statistics

Unlike PMD, ASE noise is much less sensitive to practically observed variations in the operation conditions. Provided the temperate and pumping of an EDFA stays constant, the amount of produced ASE noise in short time scales has a low statistical variation.

Figure 3.1 presents the output ASE noise power in dBm of 2400 and 4800 km links with EDFAs between the spans of 80 km (30 and 60 EDFAs per link respectively). The EDFAs ideally compensate for the fibre span loss of 16 dB. EDFA's NF vary from 3 to 8 dB with a step of 1 dB and the power gain is 16 dB (ideal span loss compensation). The signal bandwidth is 28 GHz and the central frequency is 193.0985 THz The results show low deviation from the average. They are in good, although not perfect, conjunction with the

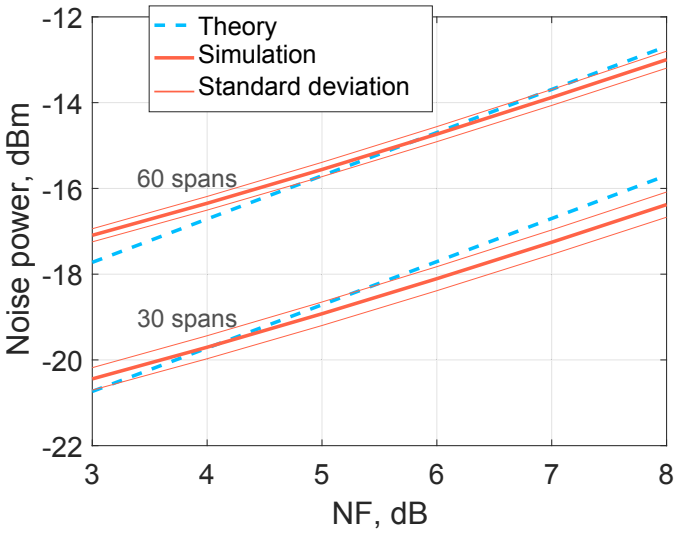


Figure 3.1: Noise power generated by an EDFA. Theoretical vs simulated curves.

theoretical formula giving an estimation on the ASE power [77]:

$$P_{ASE} = NF_{lin} \cdot \mathcal{G}_{lin} \cdot \hbar f \cdot B \cdot N_{spans}, \tag{3.1}$$

where NF_{lin} is the linear scale NF, $\mathcal{G}_{lin}$ is the linear scale amplifier gain, B is the signal bandwidth, $\hbar$ is Plank's constant in $J \cdot s$ and f is central frequency in Hz.

The mismatch of the theoretical and simulated curves is due to theoretical formula assumptions that, firstly, no ASE noise amplification occurs along the link, which could explain why simulated values are higher than theoretical for low NFs. Secondly, for higher NFs, the mismatch could be due to $\mathcal{G}$ being frequency-independent in the theoretical formula, whereas in the simulations a typical frequency-dependent EDFA gain spectrum was used.

3.1.2 PMD statistics

In this section we simulate only PMD independent of any other fibre propagation effects. The PMD statistics is simulated using the coarse-step model. An example of statistical PMD simulation is depicted in Figure 3.2 for a 2400 km system and three different SSMF with PMD coefficient, PMDc, of 0.04, 0.1 and 0.3 $ps/\sqrt{km}$, based on 10^5 fibre birefringence realizations. A birefringence

profile realization describes a specific combination of the parameters of the coarse-step model, like sections' lengths and polarisation scramblers' matrices. The simulation of DGD statistics is shown with bars and is compared to the theoretical Maxwellian distribution shown with solid line. The fibre PMD value is shown with the vertical dashed line. The means of the distributions in Figure 3.2 coincide with the empirical Equation (2.26). The DGD values very much higher or lower than the average PMD have a low occurrence probability. In the example in Figure 3.2 for a 0.3 $ps/\sqrt{km}$ PMDc fibre the probability of DGD being $0-1$ ps is 0.057%, $9-10$ ps is 6.2%, $12-13$ ps is 6.8%, and $34-35$ ps is 0.038%, calculated as:

$$P_{\Delta\tau_1<\Delta\tau<\Delta\tau_2} = \int_{\Delta\tau_1}^{\Delta\tau_2} P(\Delta\tau)\cdot d(\Delta\tau)\cdot 100\%, \tag{3.2}$$

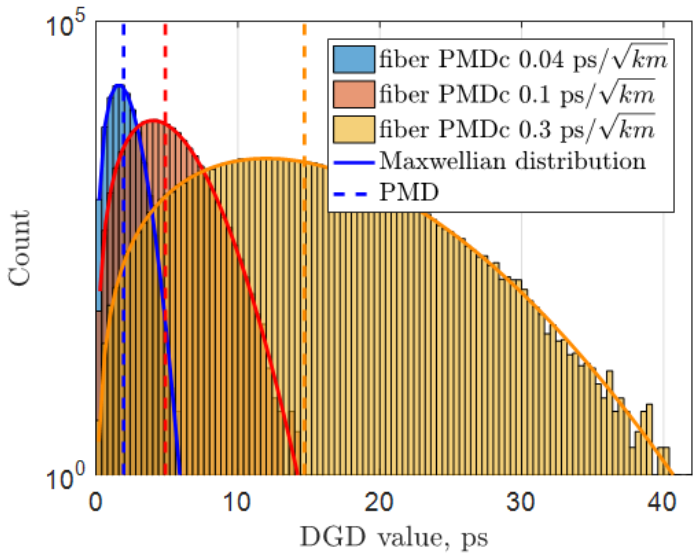


Figure 3.2: Simulated DGD distribution of a 2400 km fibre with PMDc 0.04, 0.1 and 0.3 $ps/\sqrt{km}$.

If one wanted to describe the differential group delay evolution (DGDev) **along** the fibre, one would need to have exact information about the DGD and PSP at each point in the fibre. While the PSP along the fibre are impossible to predict empirically, some speculations can be done about the values of the DGD: Equation (2.26) describes the average DGD value in a fibre of any length.

In Figure 3.3 three DGD values are chosen from the 0.3 $ps/\sqrt{km}$ distribution and 100 fibre realizations are simulated and averaged for each of the three

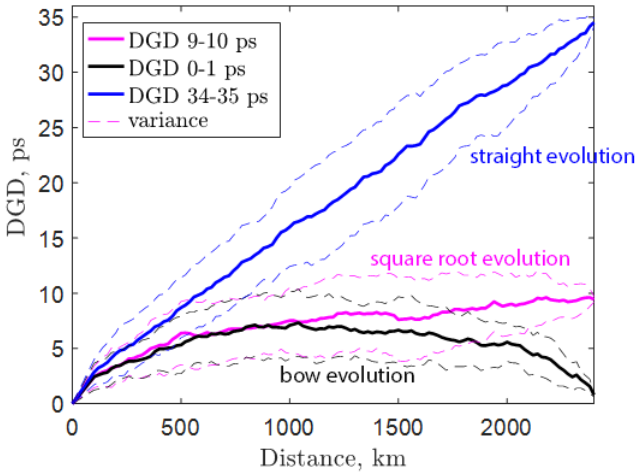
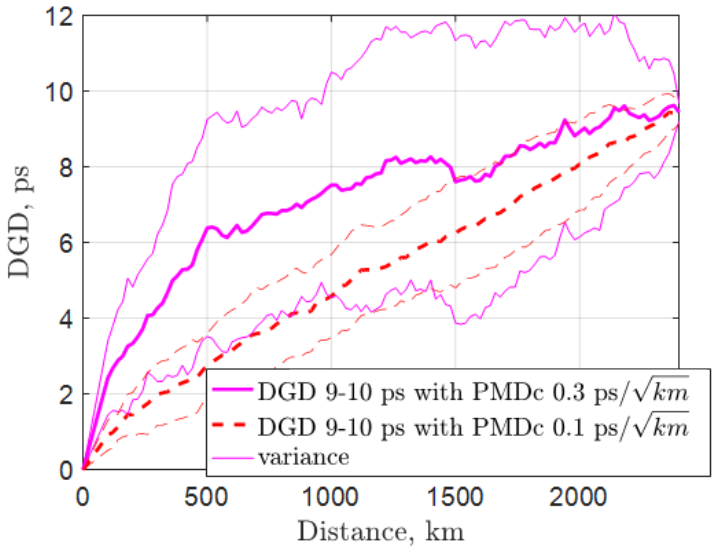


Figure 3.3: Average and standard deviation of DGDev along the link in three probability regions. $0-1$ ps very low probability, $9-10$ ps high probability, $34-35$ ps very low probability. Each case averages 100 fibre realizations with PMDc 0.3 $ps/\sqrt{km}$.

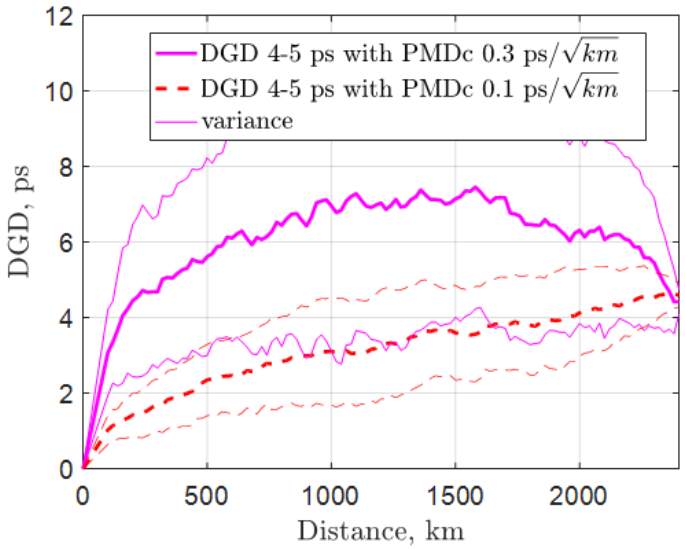
chosen points. The mean DGDev that results in $9-10$ ps DGD in pink colour follows the empirical formula with a square root length dependence. However, in order to achieve improbable high values of $34-35$ ps in blue, DGD follows the linear growth. In order to drop to improbable low values of $0-1$ ps in black, DGDev grows up to a certain point and then reduces again, following a bow shape. We further refer to these cases as square root, straight and bow evolutions.

Figure 3.4 illustrates two different fibres and two intervals from the DGD statistics, based on 100 fibre realizations for each case. As in Figure 3.4a, the evolution follows either the straight line if the final DGD value is improbably high for the fibre or the square root line if the final DGD is close to fibre PMD. In Figure 3.4b the bow and square root evolutions of different fibres but same final DGD are compared. Notice that these evolutions will also have a different probability of occurrence.

Figure 3.3 and 3.4 show that the DGDev is fully characterized by neither the fibre PMDc, nor DGD value alone. The combination of those two values gives more valuable information. The fibre PMDc defines the DGD distribution and the final DGD, in turn, defines, which pattern the DGDev follows. Although

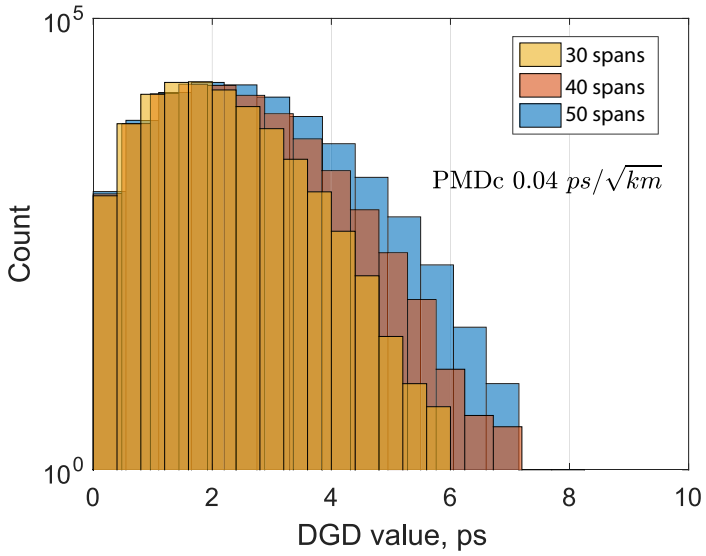


(a) Comparison of square root and straight evolutions.

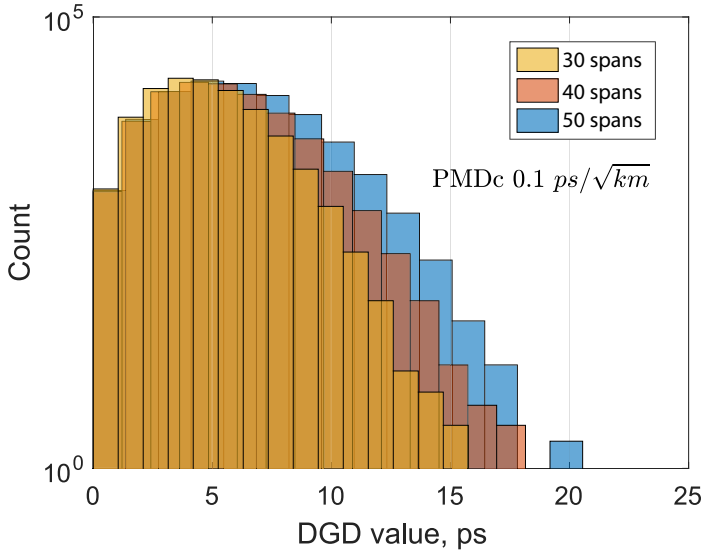


(b) Comparison of square root and bow evolutions.

Figure 3.4: Difference in the average DGDev with a fixed final DGD and variable fibre PMDc in a 2400 km link.



(a) 30, 40, and 50 span links with a fibre PMDc 0.04 $ps/\sqrt{km}$.



(b) 30, 40, and 50 span links with a fibre PMDc 0.1 $ps/\sqrt{km}$.

Figure 3.5: Statistic distribution of DGD depending on the link length.

the actual evolution still remains stochastic, the proposed way of considering PMD minimizes the prediction error.

In the simulations of the PMD impact on nonlinear mitigation techniques such as DBP, the way the DGD evolves along the fibre should be considered. Whether the PMDc is fixed and the investigated DGD values are preselected or whether the PMDc varies and the most probable DGD values are simulated, the evaluated impact of PMD will be different (please find details in Appendix B).

Figure 3.5 shows the statistical distribution of DGD for three different link lengths with 30, 40 and 50 spans, 80 km SSMF each, in two different fibre types, 0.04 and 0.1 $ps/\sqrt{km}$ PMDc. For each of the fibres the three distributions have only a moderate difference. The purpose of the pictures is to demonstrate typical DGD values in practical links.

3.1.3 PDL and PDG statistics

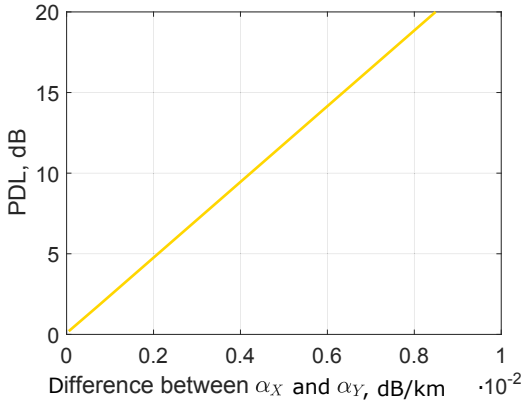
As described in Section 2.4.3, we consider distributed PDL and lumped PDG for simulation. The statistical investigations, performed in other works, conclude that the PDL is following a Maxwellian, a combination of Rayleigh and Maxwellian distributions, or neither of those [60–62], which is consistent with the following presented results.

Distributed PDL

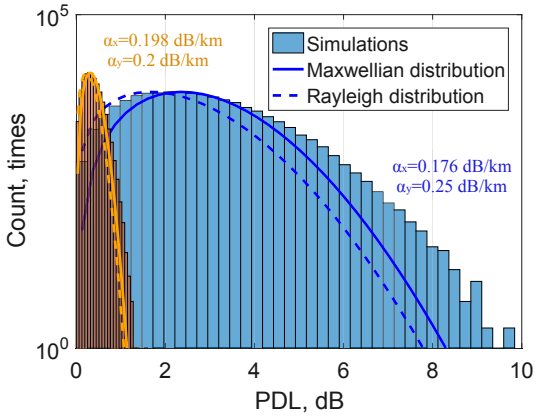
For distributed fibre PDL, statistics is collected for a 2400 km system based on 10^5 fibre realizations. The simulated pairs of α_x and α_y and $\Delta\alpha = \alpha_y - \alpha_x$ are summarized in Table 3.1.

Table 3.1: Pairs of α_x and α_y for simulating distributed PDL while keeping the total power profile unchanged (from the one with $\alpha_x = \alpha_y = 0.2\ dB/km$).

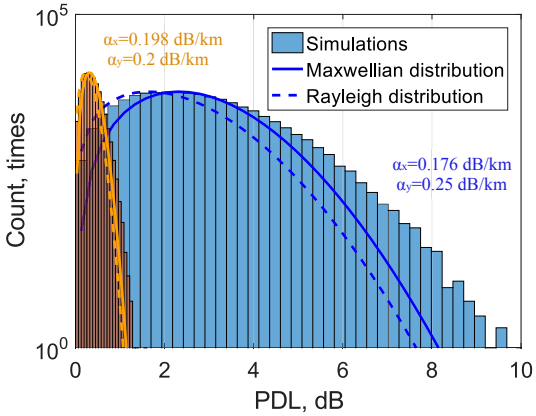
α_x, dB/km	α_y, dB/km	$\Delta\alpha$, dB/km
0.1764	0.2500	0.0736
0.1913	0.2100	0.0187
0.1976	0.2050	0.0074
0.1980	0.2005	0.0025



(a) PDL of a 2400 km link without birefringence.



(b) Statistical distribution of PDL distributed along a 2400 km link without PMD, but with random polarisation rotation.



(c) Statistical distribution of PDL distributed along a 2400 km link with 0.1 $ps/\sqrt{km}$ PMD.

Figure 3.6: Statistic distribution of a distributed PDL depending on PMD and birefringence.

When PMD in the system is negligible, and no birefringence is present, even a slight difference of the slow and fast axes attenuation coefficient causes a considerable PDL, as shown in Figure 3.6a. Note the scaling factor for the X axis. This can be explained by a strong impact of the attenuation coefficient difference not smoothed by polarisation rotation. Generally, distributed PDL should not be investigated independently of polarisation rotation and PMD, as the nature of these effects lies in the same physical phenomena of birefringence [78].

Figure 3.6b and 3.6c show the same system, firstly, without PMD but with random birefringence, and secondly, with both PMD (0.1 $ps/\sqrt{km}$ PMDc) and birefringence. Random polarisation rotation averages out the effect, and the resulting PDL is considerably smaller. Adding the PMD barely changes the resulting PDL statistic.

Lumped PDG

We continue investigating PDL statistics as a lumped gain at EDFAs in Figure 3.7. Here we do not consider PMD in a fibre, and polarisation rotation is ensured by random polarisation scrambling before each EDFA (see Figure 2.15c). Each amplifier has either 0.1 or 0.3 dB insertion loss in one of the polarisations, so that one of the polarisations has 16 dB and the other 16.1 or 16.3 dB gain respectively. The setup ensures that the total power does not change, which is important for investigating Kerr nonlinear effects. Therefore, the X-axis represents link PDL, as defined in Equation (2.32), rather than EDFA PDG. The resulting probability distribution follows neither Maxwellian, nor Rayleigh distributions.

3.2 Impact of stochastic effects on DBP

3.2.1 Simulations

Simulation setup

In order to investigate the impact of stochastic effects on DBP, we use different modulation formats and symbol rates. We consider the transmission of single-channel 28 $GBaud$ DP-QPSK, 56 $GBaud$ DP-QPSK and 28 $GBaud$ DP-16-

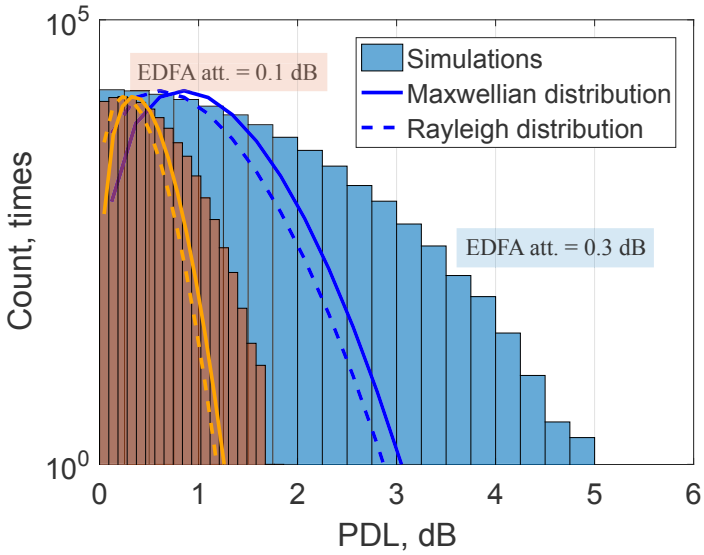


Figure 3.7: Statistical distribution of link PDL caused by lumped PDG in each of the 30 EDFAs.

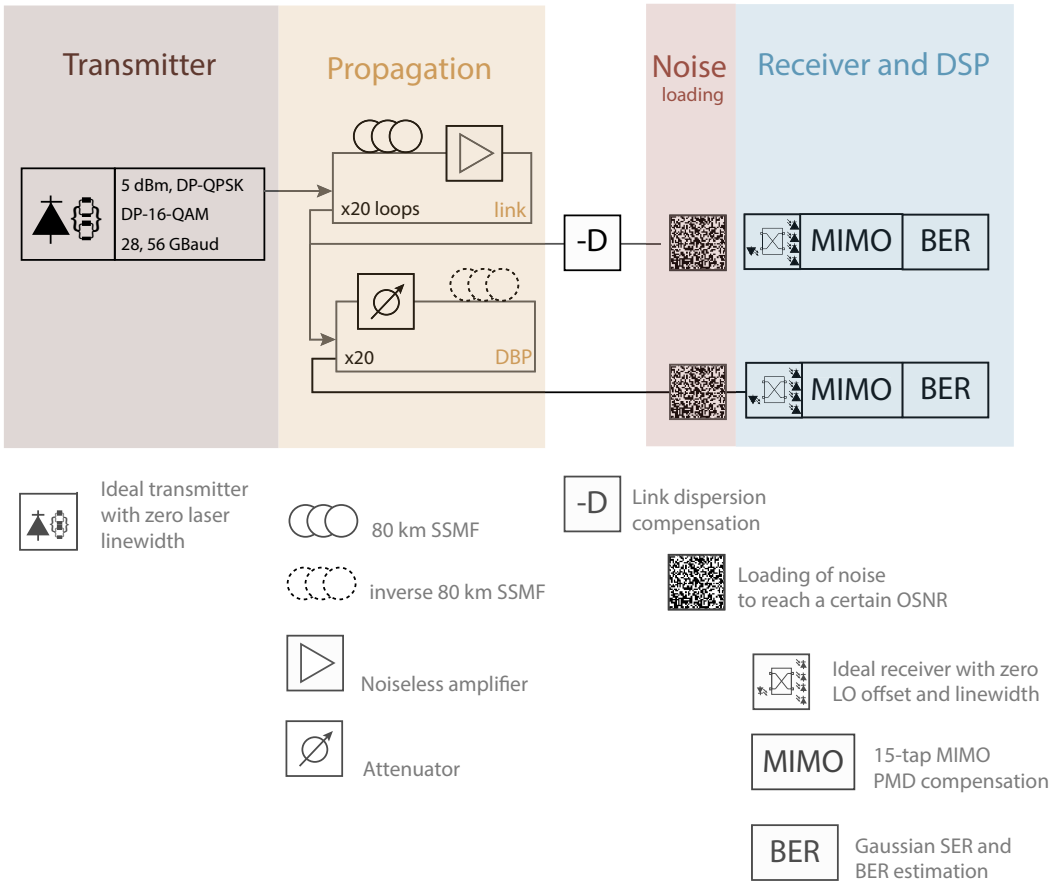


Figure 3.8: Simulation setup.

QAM with non-return-to-zero (NRZ) encoder. The transmitter output signals have 5 dBm power per transmission and are sent into the non-DM transmission link, which is schematically shown in Figure 3.8. The link consists of 20 spans of 80 km SSMF with one EDFA per span to compensate for the fibre loss. The EDFAs are *noiseless*, which means they are not introducing any ASE noise.

The signal is detected by an ideal shot-noise limited coherent receiver: Phase noise of the transmitter laser and local oscillator (LO) are neglected in order to focus on the impact of fibre nonlinearities. The receiver has a filter matched to the signal spectrum (30 or 58 GHz respectively) followed by a DSP unit. In case of PMD- and PDL-impact investigation, noise loading is performed after the DBP to separate the effects and to be able to control the OSNR. DBP is performed using the 0.05° MaxPhCh implementation, using the inverse Manakov equations (see Section 2.1.3). Linear polarisation effects are mitigated using a 15-taps multiple-input multiple-output (MIMO) equalizer, which was optimized using the constant modulus algorithm (CMA) for QPSK and multiple modulus algorithm (MMA) for 16-QAM [79] (also see Appendix C). For each configuration 2^{17} Gray mapped symbols are transmitted.

Impact of ASE on DBP

The impact of EDFA-induced ASE noise has been investigated in [80, 81]. The authors came to the conclusion that the DBP degradation becomes significant in long-haul links ($> 4800\ km$) and high input powers ($> 8\ dBm$). As will be seen in the next sections, PMD and PDL/PDG start inducing DBP degradation at considerably lower distances and powers, thus being the main performance limiting effects.

Impact of PMD on DBP

First we concentrate on a 28 $GBaud$ DP-QPSK signal transmission in fibres with five values of the PMDc (0.1, 0.3, 0.5, 1.0 and 2 $ps/\sqrt{km}$). A hundred different random fibre birefringence profiles are simulated for each configuration.

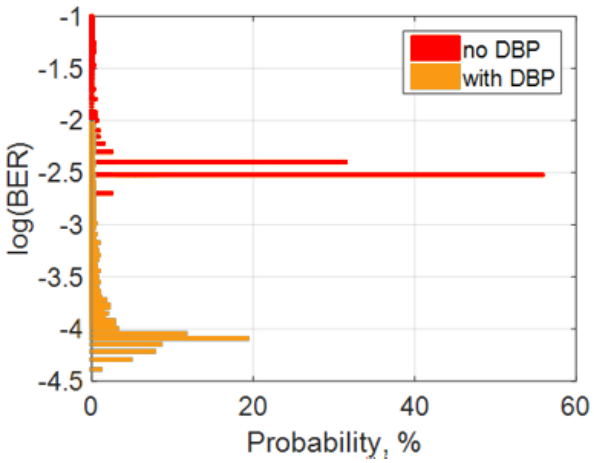
We plot the statistics of the received bit error ratio (BER) in Figure 3.9a. The distribution of the BER ranges from $10^{-4.5}$ up to 10^{-2}. But what does it depend on? Why does in some cases the system performs better than in others?

In order to see that, we plot the same results versus the DGD in Figure 3.9b. It becomes clear now that there is a correlation between DGD and BER. The system with DBP experiences considerable performance degradations for DGD values >30% of the symbol duration T_s. However, note that the degradation occurs at any value of DGD, even <30% of T_s, but its "visibility" is defined by the noise level. In this particular example 16 dB linear OSNR "hides" any PMD degradation below BER 10^{-4}. Compare the performance of exactly the same setup with 50 dB linear OSNR (calculated with respect to the ASE noise not considering nonlinear noise.) in Figure 3.9c. The optical noise is no longer hiding the PMD-induced degradation.

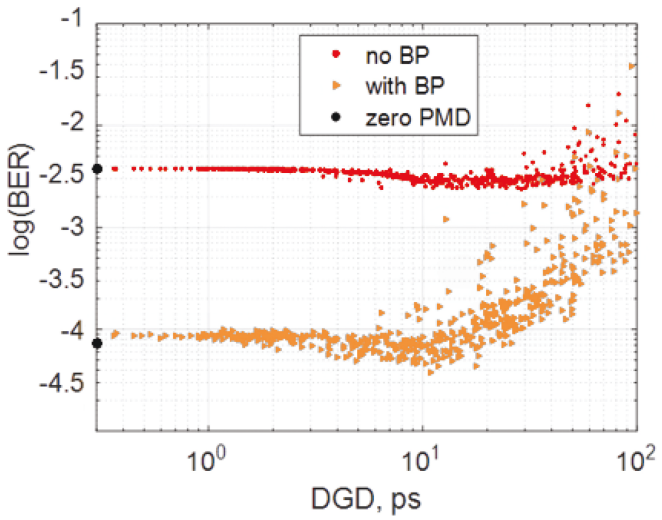
Figure 3.10 and 3.11 show the BER mean and standard deviation of the 28 $GBaud$ DP-QPSK system with and without DBP, compared with results for 56 $GBaud$ DP-QPSK and 28 $GBaud$ DP-16-QAM signal. The linear OSNRs were optimized to reach similar BER values for all considered cases in back-to-back propagation. The results clearly show that BER degradations due to PMD correlate with the DGD of the link.

Figure 3.10 could possibly imply that doubling the symbol rate is worse than doubling the number of bits per symbol. This could be explained by the importance of DGD value relative to the symbol duration. Whereas using higher modulation levels keeps the symbol interval unchanged, increasing the symbol rate proportionally decreases the symbol time interval. However, as will be shown in Section 4.3.4 and Section 5.3.2, the different results are rather due to different OSNRs. The importance of these figures is to show that the DBP performance correlates with DGD.

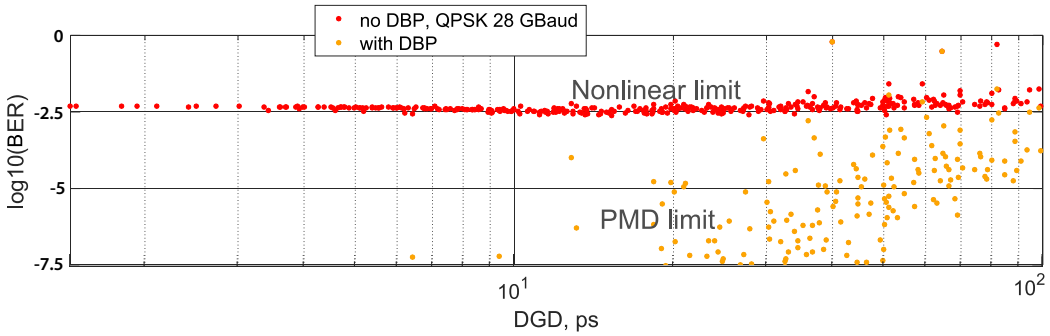
From Figure 3.11 we can conclude that the system performance standard deviation increases with the DGD value. Here the high probability DGD cases for each fibre's PMDc were simulated, which means that the system performance standard deviation grows rather with the fibre PMDc. Recalling the DGD statistical distributions, for example, in Figure 3.4, we can notice that fibres with larger PMDc have bigger variation in DGD evolutions. Therefore, bigger variations of DBP performance will be observed. In Appendix B it is shown that when the PMDc of the fibre is fixed, the results differ considerably, particularly the performance variation does not depend on the value of the DGD.



(a) Satistical distribution of the BER logarythm of a 28 $GBaud$ DP-QPSK 20x80 km system and 16 dB linear OSNR.



(b) BER as a function of DGD. Signal has 16 dB linear OSNR.



(c) Impact of noise on the NL-PMD interaction visibility. Signal has 50 dB linear OSNR.

Figure 3.9: Difference in BER statistics with introduction of DGD as a new performance criterion in a 2400 km system with a 28 $GBaud$ DP-QPSK signal.

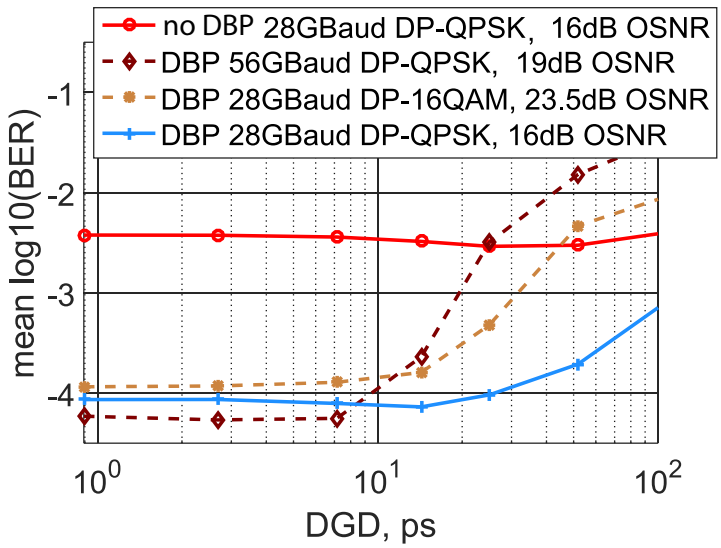


Figure 3.10: Average BER as a function of the link DGD for three investigated systems. OSNR is linear (only ASE noise considered).

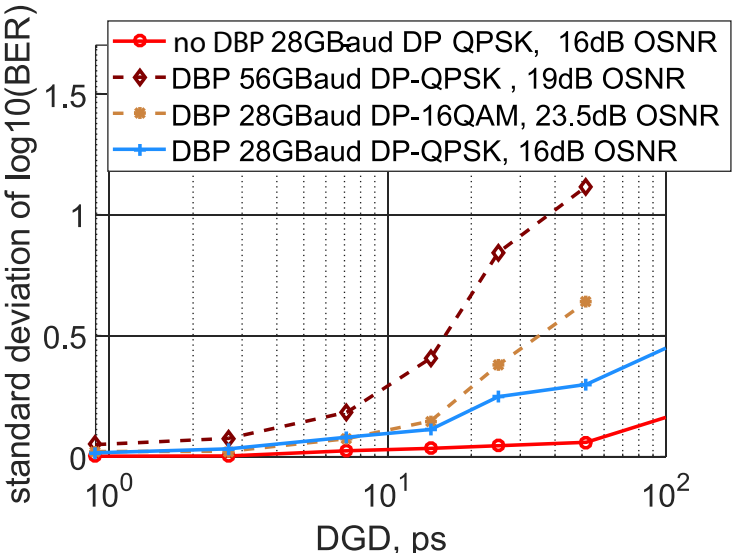


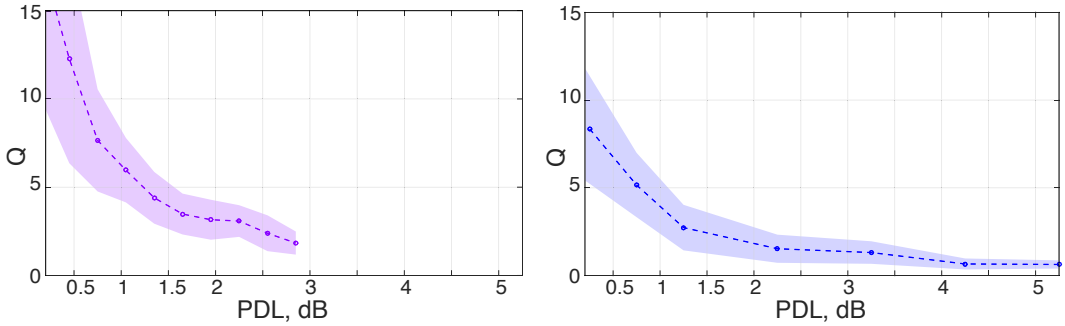
Figure 3.11: Variation of BER as a function of the link DGD for three investigated systems. OSNR is linear (only ASE noise considered).

Impact of PDL on DBP

The impact of PDL on DBP was investigated for two different scenarios, lumped PDG and distributed PDL as specified in Section 2.4.3.

In the lumped PDG case, the gain difference between the X and Y polarisations at each of the EDFAs was 0.05, 0.1, 0.15, 0.2, 0.3, 0.5, and 0.7 dB. These values were chosen to cover typical connector losses. Fifty statistical fibre birefringence realizations for each of the values were simulated.

In the distributed PDL case the simulated difference between the fibre orthogonal modes' attenuation coefficients $\Delta\alpha$ comprised $2.5 \cdot 10^{-3}$, $7.4 \cdot 10^{-3}$, $1.9 \cdot 10^{-2}$, and $7.4 \cdot 10^{-2}$ dB/km along the whole link (see Table 3.1). The actual attenuation coefficients for each axis were selected to keep the total power profile similar to the case when both modes' attenuation coefficients are equal.



(a) DBP degradation by lumped PDG. (b) DBP degradation by distributed PDL.

Figure 3.12: DBP system performance correlation with the PDL at the receiver.

As shown in Figure 3.12, the Q factor is plotted versus the PDL of the link. Q factor and BER are related as:

$$Q = 1/2 \cdot erfc(BER/\sqrt{2}), \tag{3.3}$$

where $erfc\{\}$ is a complementary error function. We switch here from BER to this metric to be able to plot average and standard deviation of the DBP performance in one figure. More information on the Q factor can be found in Appendix D.

The dramatic drop in DBP performance when the PDL is growing can be clearly seen in Figure 3.12. Interestingly, when PDL is up to 0.2 dB, the

performance of the lumped case is considerably better than the performance of the distributed case. This can be explained by the high sensibility of the DBP efficiency to the correctness of the power profile, as we have seen in [82] for Raman amplified links. The standard deviation of performance is, however, higher for the lumped than for the distributed case, which could be connected to the polarisation rotation averaging in the latter case.

3.2.2 Experiment

The experiment of the DBP degradation in presence of PMD was performed in the facilities of "High-speed optical transmission laboratory" of Acreo Swedish ICT (now RISE Acreo). The author is heartily thankful to Oskars Ozolins, Xiaodan Pang, Suhail Al-Awis, Richard Schatz, Anders Djupsjöbacka, Gunnar Jacobsen and Sergei Popov for common working on the experiment, discussions and for granting access to previously created DSP library.

Experimental setup

The experimental setup to characterize the impact of PMD on the performance of DBP is illustrated in Figure 3.13.

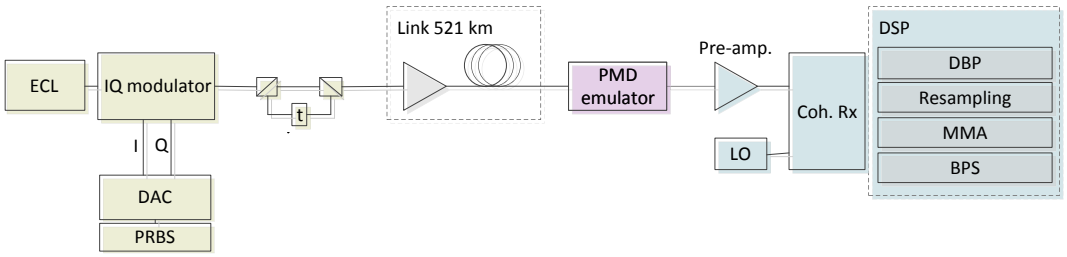


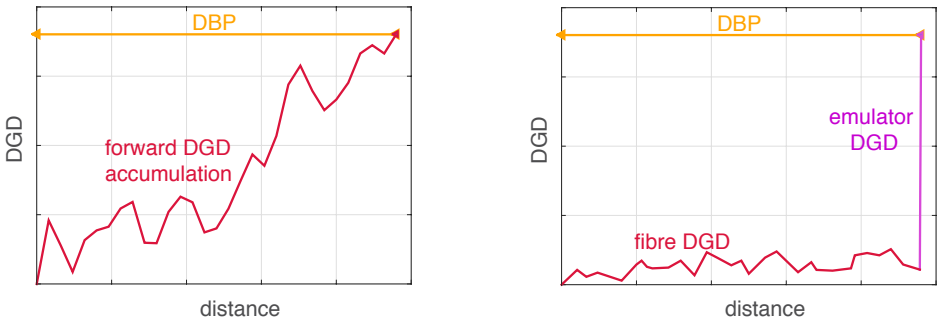
Figure 3.13: Experimental transmission system.

The transmitter consists of an external cavity laser (ECL) with less than 100 kHz linewidth, a pulse pattern generator (PPG), a 2-bit digital to analog converter (DAC) and an optical IQ modulator. The output 28 $GBaud$ 4-pulse-amplitude modulation (PAM) signals are decorrelated and then fed into an optical IQ modulator with a bandwidth of 25 GHz. The output of the IQ modulator is mapped to x- and delayed y-polarisations to produce a single-channel 28 $GBaud$ DP-16-QAM signal, which is sent through a non-DM link. The link consists of 6 spans of 80 to 93 km of SSMF with one lumped EDFA

per span. The DGD of the link is estimated to be 2.8 ± 1.66 ps (mean $\pm$ standard deviation), which corresponds to a PMD coefficient of 0.12 $ps/\sqrt{km}$.

Additional PMD is introduced by one of the two emulators, each one being a passive device formed by random splices of polarisation maintaining fibre [83]. Such emulators have fixed root-mean-square differential group delay (PMD) and varying second-order PMD.

The first PMD emulator has a DGD of 14.4 ± 1.9 ps, and the second one 21.2 ± 2.9 ps standard deviation. Each of the PMD emulators is placed at the end of the link to consider the worst possible scenario for DBP. In such case the error between the actual and back-propagated link is maximised, leading to the highest PMD-induced penalty. Although, such experimental setup excludes the PMD-nonlinear interaction, that are present in practical DBP, it introduces a difference between the forward and backpropagated fields. That difference is the source of the DBP error. The idea is illustrated in Figure 3.14, in conventional DBP the DGD is accumulating along the link, in the experiment it is added in the end of the link in a bulk.



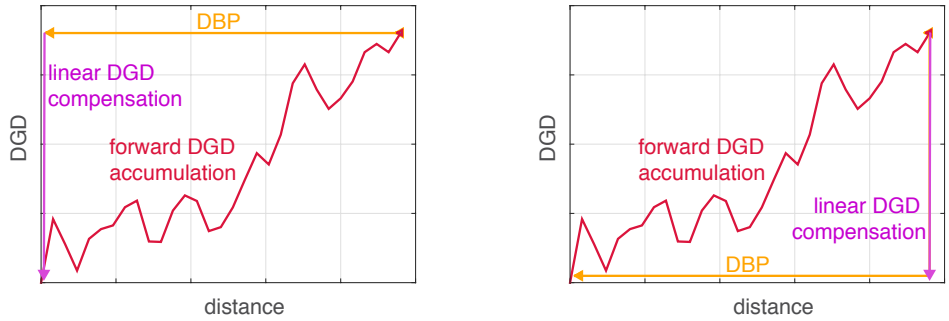
(a) Conventional DBP.

(b) Experiment.

Figure 3.14: Comparison of the PMD emulator and intrinsic fibre PMD impact on DBP.

In comparison to the simulations, performed in this work, the experimental results demonstrate DBP error coming purely from the uncompensated DGD. This error persist even if the linear PMD is compensated after DBP. The simulations will investigate the cases, where DBP error comes from either uncompensated DGD and PMD-nonlinear interaction (linear PMD is compensated after DBP, see Figure 3.15a) or pure PMD-nonlinear interaction (linear

PMD is compensated before DBP, see Figure 3.15b).



(a) Linear PMD is compensated after DBP. (b) Linear PMD is compensated before DBP.

Figure 3.15: Comparison of the impact of two positions for linear PMD compensation on DBP.

The transmission in the experiment is considered for two regimes: linear case (2 dBm transmitter output power) and highly nonlinear case (8 dBm), resulting into 30 and 25.1 dB OSNR before the DSP respectively. At the receiver, the signal is preamplified to 0 dBm and detected by a coherent receiver with 73 GHz bandwidth and with a less than 100 kHz LO linewidth. Digital post-processing includes DBP realized by a 0.05° MaxPhCh algorithm [49] followed by resampling to 2 samples per symbol. MMA equalizer [84] is applied to compensate for linear polarisation effects followed by a blind phase search (BPS) for carrier frequency and phase recovery, and error counting.

Impact of PMD on DBP performance

Figure 3.16 depicts the measured BER dependence on the DGD of the link. Each point corresponds to an average of 50 configurations with low, medium or high DGD. Those are the three transmission cases: without any PMD emulators, with the first emulator, and with the second emulator. The standard deviation of BER is shown within the horizontal coloured areas. The standard deviation of the PMD emulators and the link is depicted with vertical greenish-coloured areas.

In Figure 3.16 the linear link (2 dBm input power), depicted with green diamond marks, shows that linear polarisation effects are fully compensated

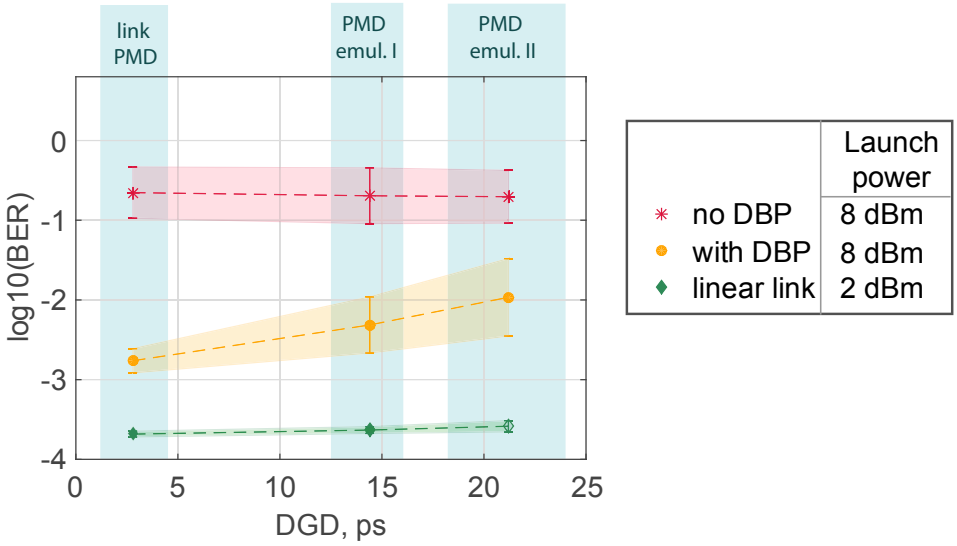


Figure 3.16: Dependence of the experimentally measured BER on the DGD of the link.

by the MMA equalizer. The orange circle marks for the DBP case (8 dBm input power) show that the penalty induced by PMD grows with DGD and can achieve one order of magnitude of BER degradation in highly nonlinear and high-PMD scenarios. It can also be observed that the standard deviation of the DBP performance grows with the PMD emulators' DGD.

3.3 Conclusions

In this chapter we showed the statistics of a typical EDFA ASE noise and PDG, and SSMF PMD and PDL simulations. We demonstrated that the majority of DGD and PDL values in modern fibre links in normal operating conditions will be relatively small. However, there is a non-zero probability of reaching high values, which significantly impacts the performance of DBP.

Using simulations of typical transmission systems, we characterised the PMD-, PDG- and PDL-induced DBP performance degradation. We showed, that the noise level plays a crucial role in the "visibility" of the NL-PMD limit. Furthermore, we presented the experimental demonstration of DBP degradation due to the group delay between x and y signal components.

Chapter 4

Perturbation analysis of PMD

The purpose of this chapter is to show how DGD can be introduced to the perturbation model. The next step is a simplification of the proposed modification to the analytical solution, enabled by an assumption on the DGDev shape. Finally, the chapter presents simulations of optical field propagation using the perturbation model and the SSF, verifying the proposed model and simplifications. First method is implemented in *MathWorks MATLAB*, the second one in *VPItransmissionMaker Optical Systems*.

4.1 Perturbation approach

4.1.1 Perturbation approach foundations

In Section 2.1.2 we showed that the propagation of light over single-mode fibre can be described by a set of the coupled Manakov equations [37–39]:

$$\frac{\partial A_{x,y}}{\partial z} + \frac{i}{2}\beta_2(z)\frac{\partial^2 A_{x,y}}{\partial t^2} - \frac{\beta_3}{6}\frac{\partial^3 A_{x,y}}{\partial t^3} + \frac{\alpha(z)}{2}A_{x,y} = i\frac{8}{9}\gamma(z)(\left|A_{x,y}\right|^2 + \left|A_{y,x}\right|)^2 A_{x,y}. \tag{4.1}$$

The coupled Manakov equations in time domain have an analytical solution for A_x and A_y in a limited set of cases. In all other systems, interesting for simulating currently deployed optical fibres, approximative numerical solutions can be found. It is possible to reduce the computational complexity by solving the equations numerically in the frequency domain. As will be seen further on, loss and chromatic dispersion can be described by a simple attenuation

and phase rotation. The impact of Kerr nonlinearities can be described by a single additional term. The CD slope, β_3, is neglected because its effect is small compared to the chromatic dispersion (unless the propagation is near the zero-dispersion wavelength). By taking the Fourier-transform of $A_{x,y}$ in (4.1) we obtain then:

$$A_{x,y}(\omega, z) = \mathfrak{F}\{A_{x,y}(t, z)\} = \int_{-\infty}^{\infty} A_{x,y}(t, z) exp(-i\omega t) dt. \tag{4.2}$$

Here we simplify the writing of $\Delta\omega = \omega - \omega_0$ to simply ω. Further,

$$\frac{\partial A_{x,y}(\omega, z)}{\partial z} = \frac{i}{2}\omega^2 \beta_2(z) A_{x,y}(\omega, z) - \frac{\alpha(z)}{2} A_{x,y}(\omega, z) + i\frac{8}{9}\gamma(z) \iint \Big(A_{x,y}(\omega_1, z) A^*_{x,y}(\omega_2, z) + A_{y,x}(\omega_1, z) A^*_{y,x}(\omega_2, z) \Big) A_{x,y}(\omega - \omega_1 + \omega_2, z) d\omega_1 d\omega_2, \tag{4.3}$$

where * denotes complex conjugation.

It is convenient to introduce accumulated gain G and dispersion D profiles:

$$G(z) = \int_0^L \Big(-\alpha(z) + \sum_j g_j \delta(z - z_j) \Big) dz, \tag{4.4}$$

g_j is the j-th EDFA amplifier gain and δ is the Dirac function, z_j is the position of j-th EDFA, L is the link length.

$$D(z) = \int_0^L \beta_2(z) dz. \tag{4.5}$$

Rewriting

$$A_{x,y}(\omega, z) = U_{x,y}(\omega, z) \exp\left(\frac{-G(z) + i\omega^2 D(z)}{2}\right), \tag{4.6}$$

so that the $U_{x,y}$ is only affected by NL, from (4.3) we get [85]:

$$\frac{\partial U_{x,y}(\omega, z)}{\partial z} = i \iint \frac{8}{9}\gamma(z) \exp(G(z) - i\Delta\Omega D(z)) \cdot \Big(U_{x,y}(\omega_1, z) U^*_{x,y}(\omega_2, z) + U_{y,x}(\omega_1, z) U^*_{y,x}(\omega_2, z) \Big) U_{x,y}(\omega - \omega_1 + \omega_2, z) d\omega_1 d\omega_2, \tag{4.7}$$

where $\Delta\Omega = (\omega - \omega_1)(\omega - \omega_2)$.

To find the perturbation solution of equation (4.7) let $U_{U_{x,y}}(\omega, z)$ be expanded in Volterra series, where $U_{\mathbf{x}}^{(k)}$ denotes the k^{th} order solution (namely, the k^{th} order Volterra kernel [86] multiplied with the corresponding signal term):

$$U_{U_{x,y}}(\omega, z) \approx \sum_{k=0}^{\infty} U_{U_{x,y}}^{(k)}(\omega, z), \tag{4.8}$$

From equations (4.7) and (4.8) we get

$$\begin{aligned}
&\frac{\partial}{\partial z}(U_{x,y}^{(0)}(\omega, z) + U_{x,y}^{(1)}(\omega, z) + U_{x,y}^{(2)}(\omega, z) + ...) = \\
&\qquad i \iint \frac{8}{9}\gamma(z) \exp(G(z) - i\Delta\Omega D(z)) \cdot \\
&\Big((U_{x,y}^{(0)}(\omega_1, z) + U_{x,y}^{(1)}(\omega_1, z) + ...)(U_{x,y}^{*,(0)}(\omega_2, z) + U_{x,y}^{*,(1)}(\omega_2, z) + U_{x,y}^{*,(2)}(\omega_2, z) + ...) + \\
&(U_{y,x}^{(0)}(\omega_1, z) + U_{y,x}^{(1)}(\omega_1, z) + ...)(U_{y,x}^{*,(0)}(\omega_2, z) + U_{y,x}^{*,(1)}(\omega_2, z) + U_{y,x}^{*,(2)}(\omega_2, z) + ...)\Big) \cdot \\
&(U_{x,y}^{(0)}(\omega - \omega_1 + \omega_2, z) + U_{x,y}^{(1)}(\omega - \omega_1 + \omega_2, z) + U_{x,y}^{(2)}(\omega - \omega_1 + \omega_2, z) + ...) d\omega_1 d\omega_2.
\end{aligned} \tag{4.9}$$

By equating the terms for each addend separately we get [85]:

$$\frac{\partial U_{x,y}^{(0)}(\omega, z)}{\partial z} = 0, \tag{4.10}$$

$$U_{x,y}^{(0)}(\omega, z) = U_{x,y}(\omega, 0), \tag{4.11}$$

and

$$\begin{aligned}
&\frac{\partial U_{x,y}^{(1)}(\omega, z)}{\partial z} = i \iint \frac{8}{9}\gamma(z) \exp(G(z) - i\Delta\Omega D(z)) \cdot \\
&\Big(U_{x,y}^{(0)}(\omega_1, z) U_{x,y}^{*,(0)}(\omega_2, z) + U_{y,x}^{(0)}(\omega_1, z) U_{y,x}^{*,(0)}(\omega_2, z)\Big) U_{x,y}^{(0)}(\omega - \omega_1 + \omega_2, z) d\omega_1 d\omega_2.
\end{aligned} \tag{4.12}$$

Following [85], we further limit the solution to first-order perturbation approximation, so that $U_{x,y}^{(k)}(\omega, z) = 0$ for $k > 1$, assuming low NL and constant

signal spectrum:

$$U_{x,y}(\omega,z) = U^{(0)}_{x,y}+U^{(1)}_{x,y} = U_{x,y}(\omega,0)+i\iint\Big(\int_0^L \frac{8}{9}\gamma(z)\exp(G(z)-i\Delta\Omega D(z))dz\Big)\cdot$$
$$\Big(U_{x,y}(\omega_1,0)U^*_{x,y}(\omega_2,0)+U_{y,x}(\omega_1,0)U^*_{y,x}(\omega_2,0)\Big)U_{x,y}(\omega-\omega_1+\omega_2,0)d\omega_1 d\omega_2. \tag{4.13}$$

We further define a complex link nonlinear transfer function η as:

$$\eta(\Delta\Omega) = \int_0^L \frac{8}{9}\gamma(z)\exp(G(z)-i\Delta\Omega D(z))dz. \tag{4.14}$$

The $S_{x,y}$ term, related to the input optical signal, can be written as:

$$S_{x,y}(\omega,\omega_1,\omega_2) =$$
$$\Big(A_{x,y}(\omega_1,0)A^*_{x,y}(\omega_2,0)+A_{y,x}(\omega_1,0)A^*_{y,x}(\omega_2,0)\Big)A_{x,y}(\omega-\omega_1+\omega_2,0)d\omega_1 d\omega_2. \tag{4.15}$$

Notice that $U_{x,y}(\omega,0) = A_x(\omega,0)$. We then can rewrite (4.13) as

$$U_{x,y}(\omega,z) = U_{x,y}(\omega,0)+i\iint\eta(\Delta\Omega)S_{x,y}(\omega,\omega_1,\omega_2)d\omega_1 d\omega_2. \tag{4.16}$$

Further introducing $\delta_{nl\_x,y}$ as a first-order perturbative term describing the impact of NL:

$$\delta_{nl\_x,y}(\omega,z) = i\iint\eta(\Delta\Omega)S_{x,y}(\omega,\omega_1,\omega_2)d\omega_1 d\omega_2, \tag{4.17}$$

we finally get

$$U_{x,y}(\omega,z) = U_{x,y}(\omega,0)+\delta_{nl\_x,y}(\omega,z). \tag{4.18}$$

Note that $\delta_{nl\_x,y}$ would be different for x and y components.

The propagation of a signal in a fibre, considering Equation (4.6), is then approximated by:

$$A_{x,y}(\omega,z) = \Big(U_{x,y}(\omega,0)+\delta_{nl\_x,y}(\omega,z)\Big)\exp(\frac{-G(z)+i\omega^2 D(z)}{2}) =$$
$$A_{lin\_x,y}(\omega,z)+\delta^{'}_{nl\_x,y}(\omega,z), \tag{4.19}$$

where $A_{lin\_x,y}$ is the linear solution and the nonlinearities are described by a single perturbative term $\delta'_{nl\_x,y}$:

$$\delta'_{nl\_x,y} = \delta_{nl\_x,y} \cdot \exp(\frac{-G(z) + i\omega^2 D(z)}{2}). \tag{4.20}$$

Such signal nonlinearity description can be mapped to a propagation model sketched in Figure 4.1. Fibre propagation effects are calculated for every dz step, and DGD is not yet considered.

Figure 4.1: Sketch of the perturbation model propagation.

4.1.2 Link nonlinear transfer function

The link nonlinear transfer function $\eta(\Delta\Omega)$ can be considerably simplified for certain links. When all the spans in a system are identical, for instance, the function can be analytically integrated over z.

We assume that such a nonlinear transfer function for a SSMF non-DM link of length L with ideal EDFA loss compensation, and α, γ, and β_2 not changing along the span, is:

$$\eta(\Delta\Omega) = \int_0^L \frac{8}{9}\gamma(z)\exp(G(z) - i\Delta\Omega D(z))dz, \tag{4.21}$$

$$G(z) = -\alpha z + g \cdot \delta(z - j \cdot L_{span}),\ D(z) = \beta_2 z,\ \gamma(z) = \gamma, \tag{4.22}$$

where $j = ceil\{z/L_{span}\}$, j is the current span, $ceil\{\}$ is a function rounding towards positive integer.

$$\eta(\Delta\Omega) = \int_0^L \frac{8}{9}\gamma\exp(-\alpha z + g \cdot \delta(z - j \cdot L_{span}) - i\beta_2\Delta\Omega z)dz. \tag{4.23}$$

For a single-span with span length L_{span} and without EDFA amplification in the end:

$$\eta_1(\Delta\Omega) = \int_0^{L_{span}} \frac{8}{9}\gamma \exp(-\alpha - i\beta_2\Delta\Omega)z dz = \frac{8}{9}\gamma \frac{1-\exp((-\alpha - i\beta_2\Delta\Omega)L_{span})}{\alpha + i\beta_2\Delta\Omega}. \quad (4.24)$$

For a multi-span system, where all spans are identical, it can then be written as [85]:

$$\eta_N(\Delta\Omega) = \int_0^{N\cdot L_{span}} \frac{8}{9}\gamma \exp(-\alpha z + g\cdot\delta(z - j\cdot L_{span}) - i\beta_2\Delta\Omega z)dz = \sum_{j=1}^{N}\int_0^{L_{span}} \frac{8}{9}\gamma \exp(-\alpha z - i(\beta_2 z + D_j)\Delta\Omega)dz, \quad (4.25)$$

where N is the number of spans and $D_j = \beta_2(j-1)L_{span}$ is the accumulated CD at the beginning of j^{th} span. Further,

$$\begin{aligned}\eta_N(\Delta\Omega) = \sum_{j=1}^{N}\int_0^{L_{span}} \frac{8}{9}\gamma \exp(-\alpha z - i(\beta_2 z + \beta_2(j-1)L_{span})\Delta\Omega)dz = \\ \eta_1(\Delta\Omega)\sum_{j=1}^{N}\exp(\beta_2(j-1)L_{span}\Delta\Omega) = \\ \eta_1(\Delta\Omega)\frac{\exp(-i\Delta\Omega\beta_2 L_{span}N)-1}{\exp(-i\Delta\Omega\beta_2 L_{span})-1} = \\ \frac{8}{9}\gamma\frac{1-\exp((-\alpha - i\beta_2\Delta\Omega)L_{span})}{\alpha + i\beta_2\Delta\Omega}\frac{\exp(-i\Delta\Omega\beta_2 L_{span}N)-1}{\exp(-i\Delta\Omega\beta_2 L_{span})-1}. \quad (4.26)\end{aligned}$$

The mapping of such a description is presented in Figure 4.2. All effects are calculated at once for the whole link. The only difference from Figure 4.1 is the simplified nonlinear transfer function thanks to assuming a multi-span system consisting of identical spans and having constant dispersion, nonlinear, and attenuation coefficients along the spans.

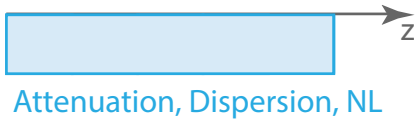


Figure 4.2: Sketch of the perturbation model propagation with simplified nonlinear transfer function.

4.1.3 Graphical representation of the nonlinear transfer function

An example of a link transfer function for a system consisting of 30 spans, 80 km each, with EDFAs exactly compensating for the span loss, is depicted in Figure 4.3.

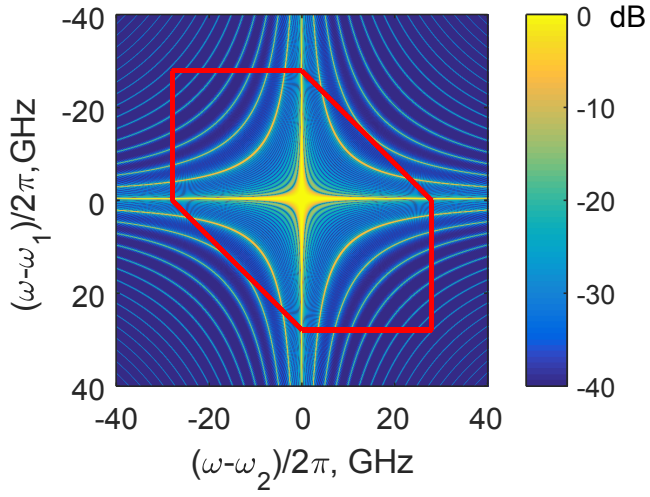


Figure 4.3: Normalised magnitude of the link nonlinear transfer function in dB scale. The red line depicts the integration region of a 28 GHz bandwidth signal.

The $\eta_N(\Delta\Omega)$ function itself is complex and Figure 4.3 shows its normalized magnitude in dB:

$$10 \cdot log_{10}(\frac{|\eta_N(\Delta\Omega)|}{max\{|\eta_N(\Delta\Omega)|\}})\ [dB], \tag{4.27}$$

where $max\{\}$ denotes maximum value. $\eta_N(\Delta\Omega)$ was calculated from Equation (4.26) for SSMF. The red curve depicts the integration region for calculating the $\delta_{nl}(\omega, z)$ for a 28 GHz signal as in Equation (4.17). SSMF parameters are summarised in Table 4.1. $\Delta\Omega$ was chosen in the order of a few kHz to provide good resolution to Figure 4.3.

Table 4.1: SSMF parameters.

Parameter	Value
Reference frequency	193.1 THz
CD	16 $ps/nm/km$
Fibre attenuation coefficient	0.2 dB/km
Nonlinear coefficient	1.37 $1/mW/m$
Core area	80 μm^2
Span length	80 km
Number of spans	30

4.2 DGD in the perturbation model

4.2.1 Introduction of DGD

In this section, we want to introduce a way to take PMD into account in the perturbative model. PMD can be seen as a time delay between the fibre PSP components. We further make a simplification in that we assume signal polarisations stay aligned with the fibre PSP during propagation. This happens, when the signal is launched in the PSPs. Then the time delay between the signal x and y components will account for the DGD.

A group delay τ between the two polarisation states in time domain corresponds to a multiplication with an exponential term $\exp(i\tau\omega)$ in the frequency domain. Only the relative delay is important, no difference is made between positive or negative τ. DGD as a delay of the y component can be described by an exponential term for A_y and A_y^*, as in Equation (4.15):

$$\begin{aligned}
\mathbb{A}_y(\omega_1, 0) &= A_y(\omega_1, 0)\, e^{i\tau(z)\omega_1}, \\
\mathbb{A}_y^*(\omega_2, 0) &= A_y(\omega_1, 0)\, e^{-i\tau(z)\omega_2}, \\
\mathbb{A}_y(\omega - \omega_1 + \omega_2, 0) &= A_y(\omega - \omega_1 + \omega_2, 0)\, e^{i\tau(z)(\omega-\omega_1+\omega_2)},
\end{aligned} \tag{4.28}$$

where $\mathbb{A}_{x,y}$ is the delayed signal $A_{x,y}$ component, $\tau(z)$ is the accumulated DGD of the link at a distance z, a single value in seconds. The key modification is highlighted with *blue* colour. Here $\tau(z)$ is not frequency dependent.

For the S terms we then get the following coefficients:

$$\mathbb{A}_y(\omega_1,0)\cdot\mathbb{A}_y^*(\omega_2,0)\cdot A_x(\omega-\omega_1+\omega_2,0)= A_y(\omega_1,0)\cdot A_y^*(\omega_2,0)\cdot A_x(\omega-\omega_1+\omega_2,0)\, e^{i\tau(z)(\omega_1-\omega_2)}. \quad (4.29)$$

$$A_x(\omega_1,0)\cdot A_x^*(\omega_2,0)\cdot\mathbb{A}_y(\omega-\omega_1+\omega_2,0)= A_x(\omega_1,0)\cdot A_x^*(\omega_2,0)\cdot A_y(\omega-\omega_1+\omega_2,0)\, e^{i\tau(z)(\omega-\omega_1+\omega_2)}, \quad (4.30)$$

$$\mathbb{A}_y(\omega_1,0)\cdot\mathbb{A}_y^*(\omega_2,0)\cdot\mathbb{A}_y(\omega-\omega_1+\omega_2,0)= A_y(\omega_1,0)\cdot A_y^*(\omega_2,0)\cdot A_y(\omega-\omega_1+\omega_2,0)\, e^{i\tau(z)\omega}. \quad (4.31)$$

In Equation (4.28) the delay is applied to one component only for simplicity. It can, however, be split between the x and y components:

$$\begin{aligned}\mathbb{A}_x(\omega_1,z)&=A_x e^{i\tau(z)\omega_1/2},\\ \mathbb{A}_x^*(\omega_2,z)&=A_x e^{-i\tau(z)\omega_2/2},\\ \mathbb{A}_x(\omega-\omega_1+\omega_2,z)&=A_x e^{i\tau(z)(\omega-\omega_1+\omega_2)/2};\end{aligned} \quad (4.32)$$

and

$$\begin{aligned}\mathbb{A}_y(\omega_1,z)&=A_y e^{-i\tau(z)\omega_1/2},\\ \mathbb{A}_y^*(\omega_2,z)&=A_y e^{i\tau(z)\omega_2/2},\\ \mathbb{A}_y(\omega-\omega_1+\omega_2,z)&=A_y e^{-i\tau(z)(\omega-\omega_1+\omega_2)/2}.\end{aligned} \quad (4.33)$$

Further we describe the group delay only in the y component.

The perturbation term for the x signal component for the whole link can be written uniting Equation (4.15), (4.17), (4.23), and (4.29):

$$\delta_{nl\_x}(\omega,z)=i\frac{8}{9}\gamma\iint\Big(\int_0^L\exp(-\alpha z+g\cdot\delta(z-j\cdot L_{span})-i\beta_2\Delta\Omega z)dz\Big)\cdot\Big(A_x(\omega_1,0)A_x^*(\omega_2,0)+ e^{i\tau(z)(\omega_1-\omega_2)}A_y(\omega_1,0)A_y^*(\omega_2,0)\Big)A_x(\omega-\omega_1+\omega_2,0)d\omega_1 d\omega_2 \quad (4.34)$$

It is convenient to split δ_{nl_x} from Equation (4.34) into two addends, where one is independent of DGD:

$$\delta_{nl\_x}(\omega,z) = i\frac{8}{9}\gamma \iint \Big(\int_0^L \exp(-\alpha z + g\cdot\delta(z - j\cdot L_{span}) - i\beta_2\Delta\Omega z) dz \Big)\cdot$$
$$A_x(\omega_1,0)A_x^*(\omega_2,0))\cdot A_x(\omega-\omega_1+\omega_2,0)d\omega_1 d\omega_2 +$$
$$i\frac{8}{9}\gamma \iint \Big(\int_0^L \exp(-\alpha z + g\cdot\delta(z - j\cdot L_{span}) - i\beta_2\Delta\Omega z)\cdot e^{i\tau(z)(\omega_1-\omega_2)} dz \Big)\cdot$$
$$A_y(\omega_1,0)A_y^*(\omega_2,0)A_x(\omega-\omega_1+\omega_2,0)d\omega_1 d\omega_2. \quad (4.35)$$

For a multi-span system consisting of identical spans the first addend can be integrated as in Equation (4.26). The possibility to simplify the second addend depends on the form of the DGD evolution $\tau(z)$. In the general case the evolution follows a random walk-off process and cannot be analytically integrated over z.

For the y signal component both addends depend on DGD:

$$\delta_{nl\_y}(\omega,z) = i\frac{8}{9}\gamma \iint \Big(\int_0^L \exp(-\alpha z + g\cdot\delta(z - j\cdot L_{span}) - i\beta_2\Delta\Omega z)\, e^{i\tau(z)\omega} dz \Big)\cdot$$
$$A_y(\omega_1,0)A_y^*(\omega_2,0))\cdot A_y(\omega-\omega_1+\omega_2,0)d\omega_1 d\omega_2 +$$
$$i\frac{8}{9}\gamma \iint \Big(\int_0^L \exp(-\alpha z + g\cdot\delta(z - j\cdot L_{span}) - i\beta_2\Delta\Omega z)\, e^{i\tau(z)(\omega-\omega_1+\omega_2)} dz \Big)\cdot$$
$$A_x(\omega_1,0)A_x^*(\omega_2,0)A_y(\omega-\omega_1+\omega_2,0)d\omega_1 d\omega_2. \quad (4.36)$$

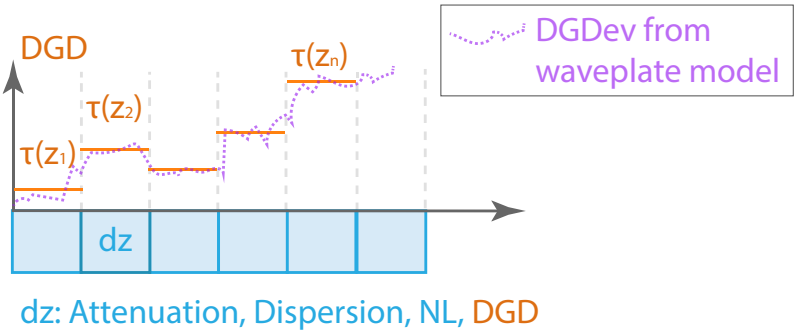


Figure 4.4: Sketch of the perturbation model propagation accounting for DGD.

This general case perturbation model can be mapped to a sketch depicted in Figure 4.4. In comparison to Figure 4.1 nonlinear propagation effects are

still recalculated every dz step, but additionally the DGD is included. For the purpose of modelling, the actual DGD values can be extracted from the coarse-step model simulations. In Figure 4.4 the dz step of the perturbation model is much larger than the one of the waveplate model.

4.2.2 Constant DGD evolution

Equation (4.35) can be considerably simplified if there is a way to integrate η over z. A possible scenario for that could be a one-step $\tau(z)$ evolution: when DGD is added at the beginning of the link and is constant along the whole link ($d\tau(z)/dz = 0$). This case corresponds to a backpropagation of a received signal with uncompensated linear DGD. The perturbative term $\delta_{nl\_x}(\omega, z)$ according to Equation (4.35) is then transformed to:

$$\delta_{nl\_x\_\tau\ const}(\omega, z) = i \iint \eta_N(\Delta\Omega) \cdot (A_x(\omega_1, 0) A_x^*(\omega_2, 0) A_x(\omega - \omega_1 + \omega_2, 0) d\omega_1 d\omega_2 + i\, e^{i\tau(\omega_1 - \omega_2)} \iint \eta_N(\Delta\Omega) \cdot (A_y(\omega_1, 0) A_y^*(\omega_2, 0) A_x(\omega - \omega_1 + \omega_2, 0) d\omega_1 d\omega_2. \quad (4.37)$$

The nonlinear transfer function can be integrated over z as in Equation (4.26). For the y component:

$$\delta_{nl\_y\_\tau\ const}(\omega, z) = i\, e^{i\tau(\omega - \omega_1 + \omega_2)} \iint \eta_N(\Delta\Omega) \cdot (A_x(\omega_1, 0) A_x^*(\omega_2, 0) A_y(\omega - \omega_1 + \omega_2, 0) d\omega_1 d\omega_2 + i\, e^{i\tau\omega} \iint \eta_N(\Delta\Omega) \cdot (A_y(\omega_1, 0) A_y^*(\omega_2, 0) A_y(\omega - \omega_1 + \omega_2, 0) d\omega_1 d\omega_2. \quad (4.38)$$

The corresponding map of the model would be as in Figure 4.5.

4.2.3 DGD evolution as a span-step function

In order to consider a more realistic DGDev, we propose to set $\tau(z)$ to a staircase (or span-step) function, i.e. it stays constant inside each span and rapidly grows between the spans on a τ_{step} value, so that the final DGD is achieved at the last span as in Figure 4.6 and can be written as:

$$\tau(z) = (j-1)\tau_{step},\ j = ceil\{z/L_{span}\}, \quad (4.39)$$

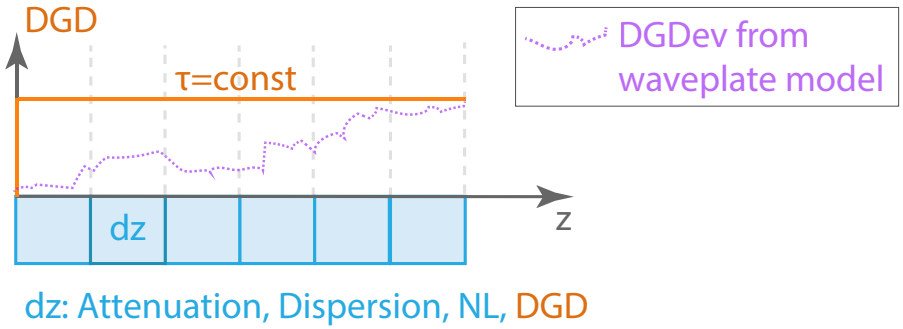


Figure 4.5: Sketch of the perturbation model propagation accounting for constant DGD. The case could describe conventional DBP of a signal with uncompensated DGD.

where j is the current span, $ceil\{\}$ is a function rounding toward positive integer.

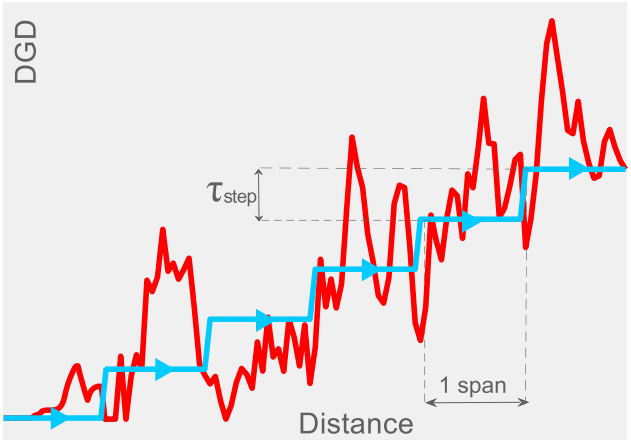


Figure 4.6: Approximation of a real DGDev with a span-step function.

From Equation (4.35) we can then write:

$$\delta_{nl\_x\_\tau\ step}(\omega,z) = i\frac{8}{9}\gamma \iint \Big(\int_0^L \exp(-\alpha z + g\cdot\delta(z-j\cdot L_{span}) - i\beta_2\Delta\Omega z)dz\Big)\cdot$$
$$A_x(\omega_1,0)A_x^*(\omega_2,0))\cdot A_x(\omega-\omega_1+\omega_2,0)d\omega_1 d\omega_2 +$$
$$i\frac{8}{9}\gamma \iint \Big(\int_0^L \exp(-\alpha z + g\cdot\delta(z-j\cdot L_{span}) - i\beta_2\Delta\Omega z)\cdot e^{i(j-1)\tau_{step}(\omega_1-\omega_2)}dz\Big)\cdot$$
$$A_y(\omega_1,0)A_y^*(\omega_2,0)A_x(\omega-\omega_1+\omega_2,0)d\omega_1 d\omega_2. \quad (4.40)$$

We want to simplify the expression for the link nonlinear transfer function η:

$$\eta_{\tau\ step}(\Delta\Omega) = \frac{8}{9}\gamma\Big(\int_0^L \exp(-\alpha z + g\cdot\delta(z - j\cdot L_{span}) - i\beta_2\Delta\Omega z)\cdot e^{i(j-1)\tau_{step}(\omega_1-\omega_2)}\, dz\Big). \tag{4.41}$$

For a single span system the nonlinear transfer function η can be analytically integrated over z using Equation (4.24) (without amplification at the end of the span):

$$\begin{aligned}\eta_{1\_\tau\ step}(\Delta\Omega) = \frac{8}{9}\gamma \int_0^{L_{span}} \exp(-\alpha z - i\beta_2 z\Delta\Omega)\, e^{i\tau_{step}(\omega_1-\omega_2)}\, dz = \\ \frac{8}{9}\gamma\frac{1-\exp((-\alpha - i\beta_2\Delta\Omega)L_{span})}{\alpha + i\beta_2\Delta\Omega}\, e^{i\tau_{step}(\omega_1-\omega_2)} = \\ \eta_1(\Delta\Omega)\, e^{i\tau_{step}(\omega_1-\omega_2)}.\end{aligned} \tag{4.42}$$

For multi-spans with exact span loss EDFA compensation, and all spans being identical, analogously to Equation (4.26) we get:

$$\begin{aligned}\eta_{N\_\tau\ step}(\Delta\Omega) = \sum_{j=1}^{N} \int_0^{L_{span}} \frac{8}{9}\gamma\cdot\exp(-\alpha z - i(\beta_2 z + \beta_2(j-1)L_{span})\Delta\Omega)\cdot \\ e^{i(j-1)\tau_{step}(\omega_1-\omega_2)}\, dz = \\ \eta_1(\Delta\Omega)\sum_{j=1}^{N}\exp(-i\beta_2(j-1)L_{span}\Delta\Omega + i(j-1)\tau_{step}(\omega_1-\omega_2)) = \\ \eta_1(\Delta\Omega)\frac{\exp((-i\Delta\Omega\beta_2 L_{span} + i\tau_{step}(\omega_1-\omega_2))N) - 1}{\exp(-i\Delta\Omega\beta_2 L_{span} + i\tau_{step}(\omega_1-\omega_2)) - 1}.\end{aligned} \tag{4.43}$$

We can then write the expression for δ_{nl} for the x signal component:

$$\begin{aligned}\delta_{nl\_x\_\tau\ step}(\omega, z) = \\ i\iint \eta_N(\Delta\Omega)\cdot A_x(\omega_1,0)A_x^*(\omega_2,0))\cdot A_x(\omega-\omega_1+\omega_2,0)d\omega_1 d\omega_2 + \\ i\iint \eta_1(\Delta\Omega)\frac{\exp((-i\Delta\Omega\beta_2 L_{span} + i\tau_{step}(\omega_1-\omega_2))N) - 1}{\exp(-i\Delta\Omega\beta_2 L_{span} + i\tau_{step}(\omega_1-\omega_2)) - 1}\cdot \\ A_y(\omega_1,0)A_y^*(\omega_2,0)A_x(\omega-\omega_1+\omega_2,0)d\omega_1 d\omega_2.\end{aligned} \tag{4.44}$$

The difference between the nonlinear transfer function η_N without DGD and $\eta_{N\_\tau\ step}$ with staircase DGDev is shown in Figure 4.7b. For convenience the example from Section 4.1.3 Figure 4.3 is repeated in Figure 4.7a. The system is as before consisting of 30 spans SSMF, 80 km each. All the parameters are summarised in Table 4.1.

Figure 4.7b depicts the magnitude of the difference between $\eta_N(\Delta\Omega)$ (as in Equation (4.26)) and $\eta_{N\_\tau\ step}(\Delta\Omega)$ (as in Equation (4.43)). In the latter case DGD accumulates to 25 ps (0.7 symbol periods, Ts). The magnitude of the difference is further normalized to the maximum value of the η_N magnitude and expressed in percent:

$$\frac{\left|\eta_N(\Delta\Omega) - \eta_{N\_\tau\ step}(\Delta\Omega)\right|}{max\{|\eta_N|\}} \cdot 100\%. \tag{4.45}$$

The red line depicts the integration region assuming a uniform signal spectrum of 28 GHz. Later δ_{nl} as in Equation (4.44) can be obtained by integration over this region.

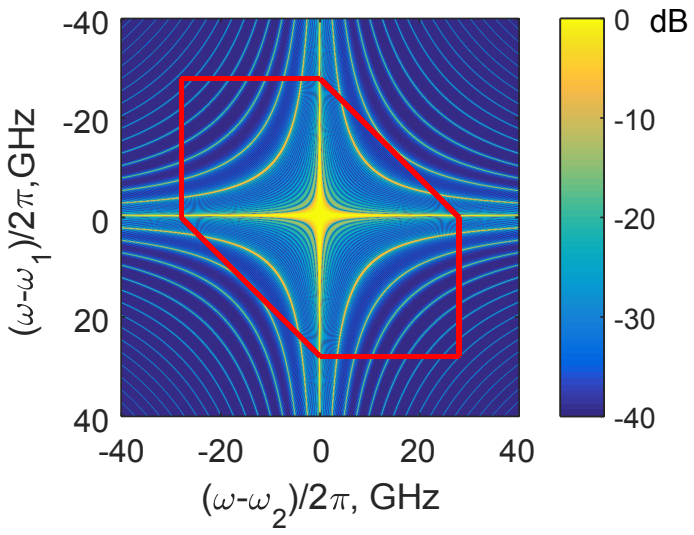
For the y polarisation component the perturbative term with a staircase DGDev looks slightly different:

$$\begin{aligned}
&\delta_{nl_y_\tau\ step}(\omega, z) = \\
&\qquad i\frac{8}{9}\gamma \iint \eta_1(\Delta\Omega) \cdot \frac{\exp((-i\Delta\Omega\beta_2 L + i\tau_{step}\omega)N) - 1}{\exp(-i\Delta\Omega\beta_2 L + i\tau_{step}\omega) - 1} \cdot \\
&\qquad A_y(\omega_1, 0)A_y^*(\omega_2, 0)) \cdot A_y(\omega - \omega_1 + \omega_2, 0)d\omega_1 d\omega_2 + \\
&i\frac{8}{9}\gamma \iint \eta_1(\Delta\Omega) \cdot \frac{\exp((-i\Delta\Omega N\beta_2 L + i\tau_{step}(\omega - \omega_1 + \omega_2))N) - 1}{\exp(-i\Delta\Omega\beta_2 L + i\tau_{step}(\omega - \omega_1 + \omega_2)) - 1} \cdot \\
&\qquad A_x(\omega_1, 0)A_x^*(\omega_2, 0)A_y(\omega - \omega_1 + \omega_2, 0)d\omega_1 d\omega_2.
\end{aligned} \tag{4.46}$$

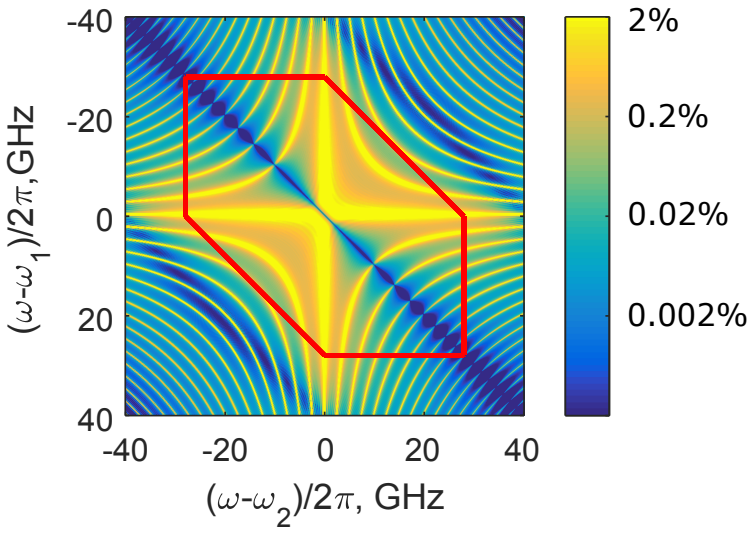
The corresponding map of the model can be found in Figure 4.8.

4.3 Simulation results

In this section we will be using two simulation approaches: perturbation and SSF method. In both methods we will simulate the forward propagation of the signal in a fibre with PMD and then ideally compensate for attenuation,



(a) Magnitude of the link nonlinear transfer function without PMD, $|\eta_N(\Delta\Omega)|$, normalized to its maximum on a dB scale.



(b) Magnitude of the difference between link nonlinear transfer function with and without DGD. Namely, $\left|\eta_N(\Delta\Omega) - \eta_{N\_\tau\ step}(\Delta\Omega)\right|$ normalized to maximum $|\eta_N|$, expressed in % and plotted in logarythmic scale, where $\eta_{N\_\tau\ step}(\Delta\Omega)$ is the magnitude of the link nonlinear transfer function with DGD of 25 ps.

Figure 4.7: Graphical representation of the DGD impact on the nonlinear transfer function.

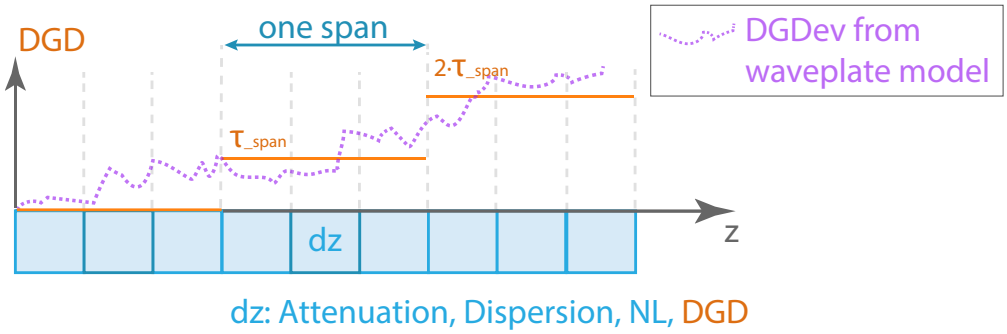


Figure 4.8: Sketch of the perturbation model propagation accounting for span-step DGD.

chromatic dispersion, nonlinearities, and linear PMD. We will then compare the reconstructed optical field with the originally sent. The difference between the two will be purely due to NL-PMD interaction.

4.3.1 Simulation environment

In the perturbation approach signal propagation is simulated by calculating the nonlinear perturbative term $\delta_{nl,\tau}$ with PMD and adding it to the signal optical field, as in Equation (4.19). Attenuation, CD and NL compensation is done by calculating the δ_{nl} without PMD and subtracting it from the propagated optical field. Linear PMD (the group delay itself) is compensated by shifting back the delayed polarisation component.

The perturbation modelling is done in frequency domain, whereas the signal is created and analysed in time domain using fast Fourier transform (FFT) and inverse fast Fourier transform (IFFT). The setup is schematically drawn in Figure 4.9, where one can also imagine an example of the generated and propagated signal.

The transmitter creates a signal similar to a QAM signal by combining random *sinc* pulses and transforming the result into frequency domain. The mean signal power is 8 dBm. The nonlinear perturbative term is calculated for a link consisting of 30 identical SSMF spans of 80 km with noiseless EDFAs with ideal power loss compensation. Fibre parameters can be found in Table 4.1. Other simulation parameters are summarised in Table 4.2, the values are rounded.

In the SSF attenuation, CD, NL are compensated using an ideal DBP (not

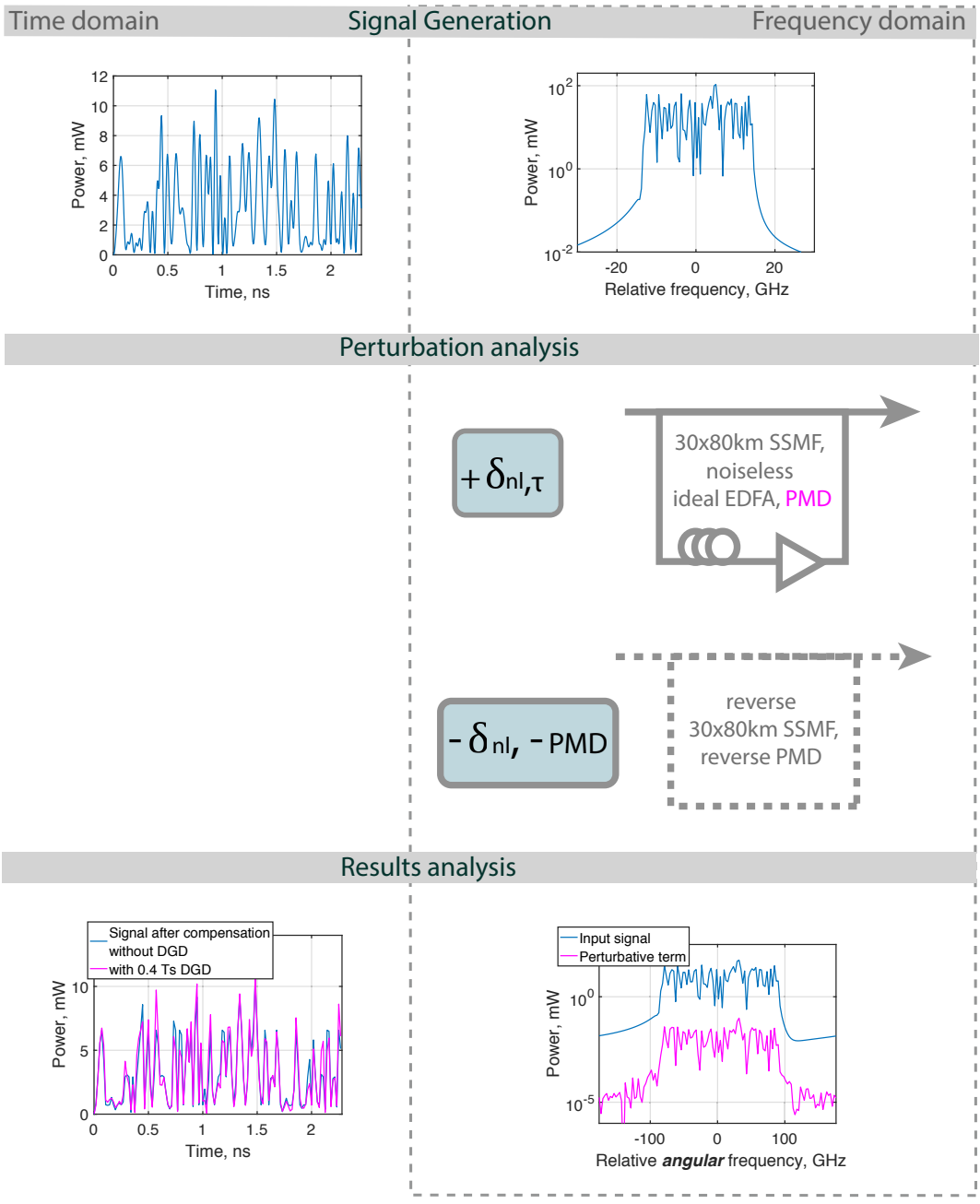


Figure 4.9: Sketch of the simulation setup with examples of the generated and propagated signal.

Table 4.2: Simulation environment parameters:

Parameter	Value	28 GBaud example
Simulation bandwidth, BW	8·symbol rate	224 GHz
Time resolution, dt	8 samples per T_s	4.46 ps
Number of symbols, N_s	64	64
Time window, T	$8 \cdot Ns \cdot dt$	2.28 ns
Frequency resolution, f	$1/T$	0.44 GHz
Signal power	8 dBm	8 dBm

limited by computational complexity), linear PMD is compensated with a MIMO MMA algorithm (after DBP). More information on the simulation parameters can be found in Section 3.2.1. In all the following cases 28 GHz DP-16-QAM signal was used.

4.3.2 Used metric

We introduce a metric to compare reconstructed field with the originally sent one. We will be using the optical fields correlation coefficient ρ as such metric:

$$\rho(A_0, A_r) = \frac{1}{N_s - 1} \sum_{i=1}^{N_s} \frac{A_0(i) - \mu_{A_0}}{\sigma_{A_0}} \frac{A_r(i) - \mu_{A_r}}{\sigma_{A_r}} \cdot 100\%, \tag{4.47}$$

where A_0 is the originally sent field, A_r is the reconstructed field, N_s is the total number of samples, μ_{A_0}, μ_{A_r} and σ_{A_0}, σ_{A_r} are the mean and standard deviation of A_0 and A_r, respectively.

Precisely, we will be looking at the correlation coefficient between the field propagated through the fibre with and without PMD, provided linear and nonlinear effects are compensated. The difference in the two fields is purely due to NL-PMD interactions.

The strength of NL-PMD interaction will reveal itself in the correlation coefficient. The correlation coefficient ρ can be helpful for predicting what part of NL will be compensated in the DBP. ρ deducted from 100% therefore will show the propagated field difference with and without PMD in the link and it will be further referred to as DBP error. In Section 7.1.2 Figure 7.3 an example is given on how the DBP error relates to the part of NL compensated

by DBP.

4.3.3 Model verification

In this section, we verify the proposed DGD extension of the nonlinear perturbation model. This will be performed in several steps. All the possible models and implementations for comparison are summarised in Figure 4.10 and 4.11. Note the mixed order of the columns, which was chosen for visual clarity reasons.

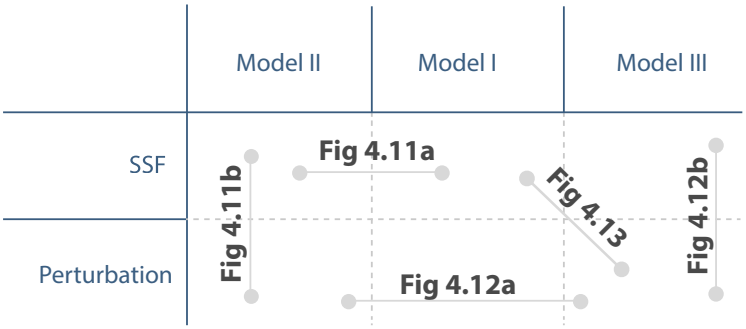


Figure 4.10: Sketch of the SSF simulations and perturbation modelling equivalence.

There are two implementations used: SSF simulation and perturbation model. In the first case for each comparison we ran 1100 simulations, using 11 different fibre PMDcs: 0.04, 0.08, 0.1, 0.2, 0.3, 0.4, 0.5, 0.7, 1.0, 1.2, and 1.5 $ps/\sqrt{km}$, so that the DGDevs follow the square root shape (for each PMDc, 100 random fibre birefringence profiles were simulated). For each case in the perturbation model 550 simulations were performed. The DGDevs were extracted from PMD coarse-step model simulations.

PMD is described in three different ways. In the first one PMD is defined by the waveplate model along with propagation. We refer to this case as Model I. This is the most realistic, but also the most complex PMD simulation. Second case, called further Model II, deals with the fibre split into relatively small (but bigger than coarse-steps) equal-length pieces, for which DGD is predefined. These values are extracted from the waveplate model. In that case no polarisation rotation between the dz steps is considered (signal polarisations are aligned with PSPs). A DGD of one step Model II is an average of all

corresponding DGD values in waveplate model. For example, if the DGD is updated on average every 100 m in the waveplate model, a DGD for the Model II step dz (4 km) is an arithmetic mean of 40 DGD values.

Finally, the third case, Model III, again deals with the fibre split into equal-length pieces, this time having span length, for which DGD is span-step approximated. DGD of a span is calculated by dividing the final accumulated DGD by the number of spans. These three PMD Models are applied along with the SSF or the perturbation term calculation as summarized in Figure 4.11. Model I with waveplate PMD is not considered for perturbation approach.

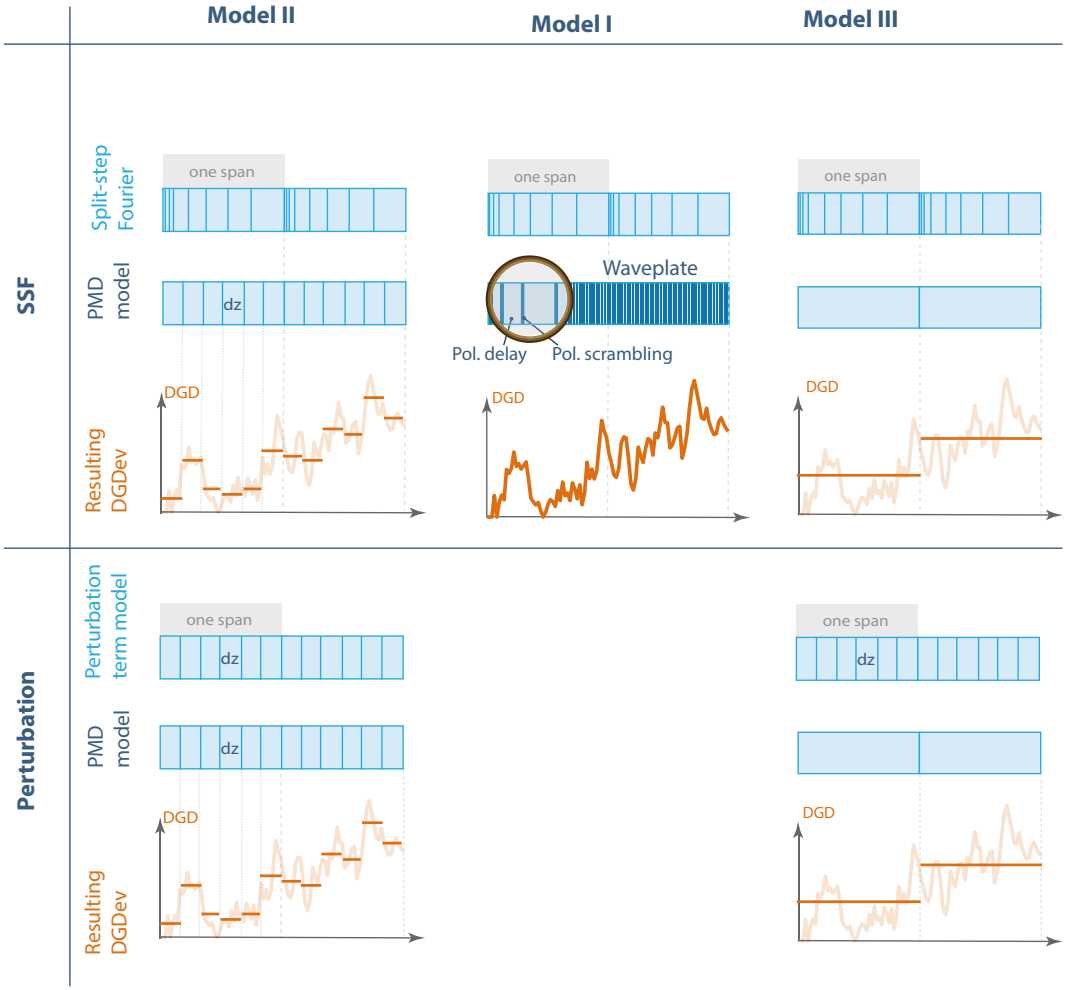


Figure 4.11: Sketch of the PMD models for the SSF and perturbation simulations.

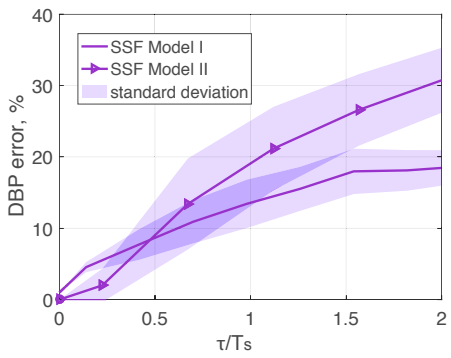
The step sizes for the models are summarised in Table 4.3.

Proceeding with the verification, first of all, we will check, how reasonable

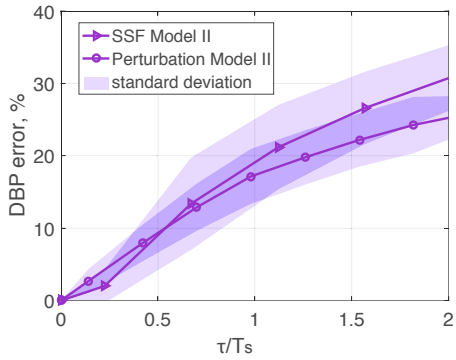
Table 4.3: Step size parameters for SSF and perturbation approach.

		Model II	Model I	Model III
SSF	DBP step	1 cm – 2 m	1 cm – 2 m	1 cm – 2 m
	PMD model step	4 km	100 m	80 km
Perturbation	Pert. term step	4 km		4 km
	PMD model step	4 km		80 km

is it to substitute a waveplate PMD model (Model I) with a fibre with a fixed DGD (Model II) using SSF. As the actual DGD values here are extracted from the waveplate model, the error here is appearing from the DGDev quantization. The resulting correlation coefficient of the originally transmitted and backpropagated optical field, deducted from 100%, is plotted in Figure 4.12a.



(a) Comparison of Model I and II in SSF simulation.



(b) Comparison of PMD Model II implemented in perturbation model and SSF.

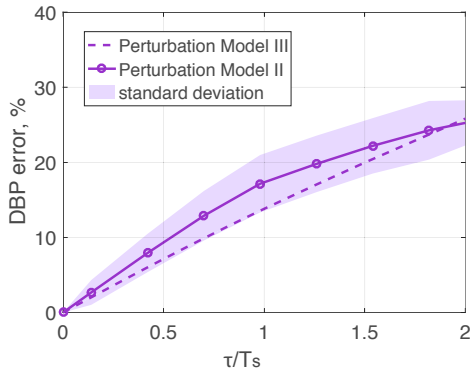
Figure 4.12: From the most accurate SSF simulation to halfway perturbation simplification.

The two cases are not in good agreement, we can conclude that very detailed DGDev information is required for accurate simulations. However, we do not cast this model aside for now, as it will become a bridge to comparing the SSF and perturbation implementations. When the number of DGD updates is decreased even further to once per span, it will surprisingly show much better agreement (see Figure 4.14).

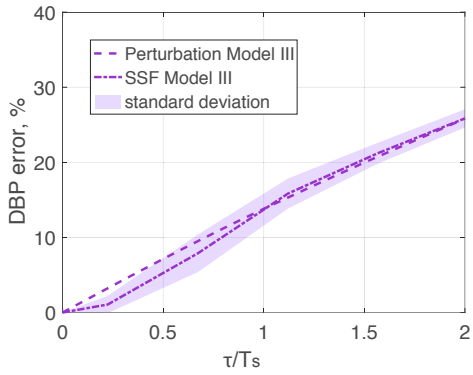
In the next step, we would like to compare the Model II implementation

using SSF and the perturbation model. As seen in Figure 4.12b, the first-order perturbation approach is closer to the SSF simulations Model I, though the approximation is still bad. That means that the error due to propagation step size is smaller than the error due to PMD model step size. Nonetheless, in a hope to trade accuracy for model simplicity, we decrease PMD model step size further as in Model III.

PMD Model III is similar to Model II but it further approximates the DGD evolution with a staircase function. This move allows us an analytical integration of the nonlinear link transfer function in the perturbation model and helps reaching satisfying level of computational complexity. In Figure 4.13a we compare the perturbation implementation of Model II and III, which show good consistency. The dependence between the DBP error and DGD resembles the DGDev shape: span-step approximation case grows linearly with increasing DGD, compared to the square-root growth of the waveplate model case. As a double-check we put together Model III implemented with SSF and in perturbation approach in Figure 4.13b. Due to the DGDev simplifications, both implementations show similar results.



(a) Comparison of PMD Model II and III in perturbative implemention.



(b) Comparison of PMD Model III implemented with SSF and perturbation model.

Figure 4.13: Second halfway simplifications.

The final overall accuracy trade-off for the simplicity can be estimated from Figure 4.14, where the most realistic SSF Model I is compared with the proposed perturbative simplified Model III. The simplification somewhat underestimates the DBP error for PMD below one symbol period and has a more

significant error for DGD higher than one symbol period.

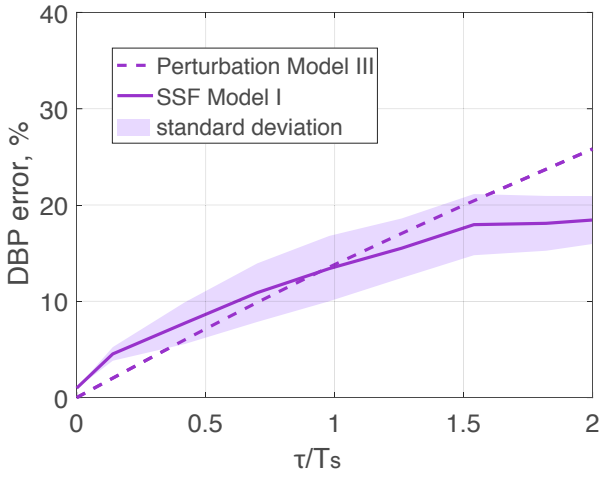


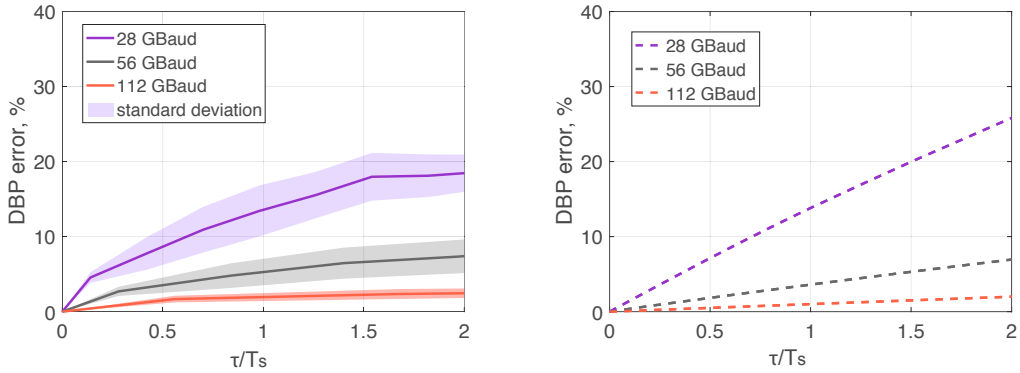
Figure 4.14: Comparison between SSF simulations (most accurate) and proposed simplified perturbation model. PMD Model I is compared here with PMD Model III in perturbative modelling.

4.3.4 Symbol rate investigation

The effect of the *linear* uncompensated DGD, i.e. a group delay between the x and y components, depends on the proportion of DGD to the pulse duration. For example, a DGD of 8 ps would have the same effect on a 28 $GBaud$ signal, as 4 ps on a 56 $GBaud$ signal (symbol periods approximately 36 ps and 18 ps respectively). This does not prove to be true for NL-PMD interactions, as we show further. We present DP-16-QAM 8 dBm signal simulations for three different symbol rates, 28, 56 and 112 $GBaud$ of SSF PMD Model I in Figure 4.15a and perturbative Model III in Figure 4.15b.

The mean DBP error grows with the DGD and decreases with growing symbol rate. However, it does not mean that higher symbol rates are less vulnerable to NL-PMD interactions. The crucial role is played by the power spectral density. While the power was kept constant to 8 dBm the power spectral density decreased two and four times for 56 and 112 $GBaud$ respectively compared to 28 $GBaud$.

To check the assumption we also simulate 28, 56, and 112 $GBaud$ SSF PMD Model I, all having the same power spectral density of 0.225 mW/GHz



(a) PMD Model I simulated as SSF.

(b) PMD Model III realized as perturbation.

Figure 4.15: Investigation of how the symbol rate influences the NL-PMD interaction in signals with same power.

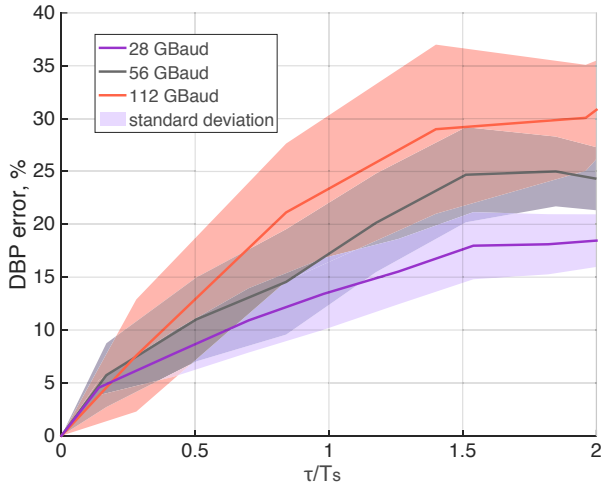


Figure 4.16: Investigation of how the symbol rate influences the NL-PMD interaction in signals with same power spectral density. Simulations are performed using SSF method.

(or 8/28 dBm/GHz) or the same energy per bit. As plotted in Figure 4.16, for DGD up to approximately one third of a symbol period the introduced field error is the same for all symbol rates. The saturation level is, however, different. For DGD bigger than 0.3 Ts higher symbol rates will suffer larger optical field deformation.

The fact that the DBP error for different symbol rates overlap with each other neither in case of constant power, nor in case of constant spectral efficiency (nor when the X axis is expressed in DGD in ps), shows that it has a complex dependence on both, the signal bandwidth and power. Comparing Figure 4.15a and 4.16, we conclude that power spectral density has a bigger impact on DBP error than signal symbol rate.

4.3.5 Modulation format investigation

In order to see if the modulation format makes a difference for the NL-PMD interaction, we simulate DP-16-, 64- and 256-QAM signals of 28 $GBaud$, 8 dBm, using PMD Model I and plot the results in Figure 4.17.

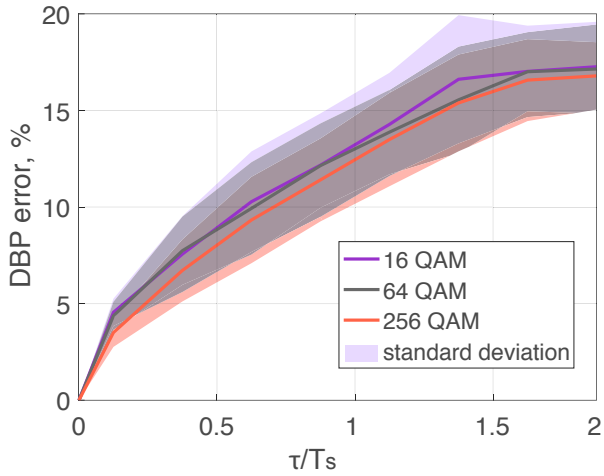


Figure 4.17: Investigation of how the QAM signal modulation level influences the NL-PMD interaction.

The DBP error due to PMD is independent of the modulation format. Higher-order modulation formats, however, are more vulnerable to the errors than the lower-order ones. When the impact on BER, symbol-error rate (SER), Q factor or similar metrics is estimated, the different modulation formats will

show clearly different performance in a link with the same accumulated DGD.

4.4 Conclusions

We showed how the perturbation model can be extended to account for DGD. We suggested an analytical simplification for the model under the assumption that the DGD follows a staircase step evolution along the link. We verified the model by comparing the optical field propagation results with the SSF simulations. The two implementations showed a fairly good match regarding the numerous simplifications and assumptions.

We introduced a metric, DBP error, defined as cross-correlation of the back-propagated field with and without DGD deducted from 100% (100% results for equal fields). The investigations showed that the DBP error has a complex dependence on both, signal bandwidth and power spectral density, the latter having a larger effect. Further we showed that the DBP error is modulation format independent.

Chapter 5

Mitigation of PMD-induced degradation

This chapter describes a novel PMD-tolerant DBP modification, based on the PMD model from Section 4.2.3. All the necessary steps are described, and the complexity is discussed. The performance of the method is tested in a number of typical long-haul optical transmission systems of different length, fibre PMDc and different symbol rates, modulation formats, and signal power levels. Multichannel implementation is demonstrated. The chapter is concluded by comparing the scope and performance of the proposed method with alternatives found in the literature.

5.1 Mitigation basis

Even though PMD is stochastic by nature, it can be described with a certain accuracy by measurable link characteristics such as the PMDc, total DGD, fibre PSP (Section 3.1.2). Though the knowledge of such metrics is not sufficient to describe the exact evolution of the local birefringence along the link, it can be used to limit the mismatch between the actual and virtual back-propagated links.

As was shown in Section 3.1.2, the knowledge of fibre PMDc and DGD can give a good approximation of the evolution of the DGD along the link. The cases with improbable high DGD are of particular interest as they cause the most damage on the DBP gain. Such cases are likely to be best approximated

by a straight-line evolution. In parallel, in Section 4.3.3 we showed that the NL-PMD interaction can be described in a simple and fairly accurate way if the stair-case DGDev is assumed.

Combining these two observations we come up with an idea to approximate the DGDev in the link while performing DBP. Just as the DBP algorithm digitally inverts dispersion and nonlinearities, the group delay between PSP components can be inverted too. Assuming accumulated DGD, τ, is known at the receiver, we suggest correcting for the delay during DBP one time per span on a fixed value $\tau_{step} = \tau/N$, where N is the number of spans (comparable with Model III in Section 4.3.3).

In forward propagation the delay was "applied" in the fibre PSPs rather than in signal logical x and y components, therefore we suggest that the PSP approximation is also needed. Further, we assume the whole link PSP can be approximated by the final PSP measured at the receiver. This approach empirically proved to give the best results among the other approaches to randomly chose PSP and keep it constant along the link or to change PSP randomly while backpropagation.

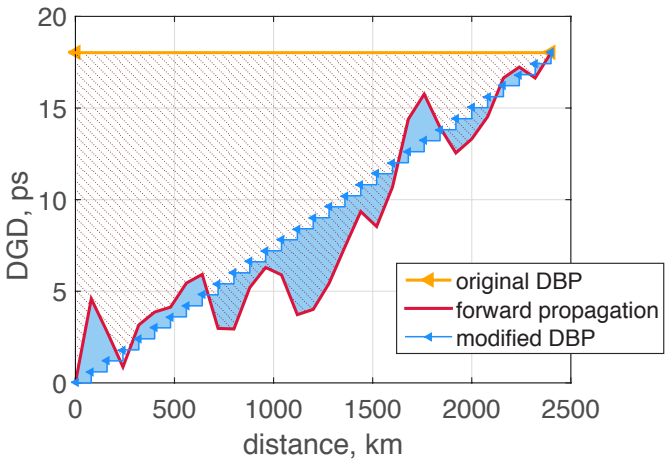


Figure 5.1: Block scheme of DGDev in forward link, conventional DBP (DGD is ignored) and proposed DBP modification (DGDev is approximated with a span-step function).

Figure 5.1 illustrates the proposed modification of the DBP algorithm. The red line depicts the forward link DGDev (in fibre PSP), the orange line shows that original DBP does not account for DGD and the blue triangle-marked

curve shows the proposed approximation of the DGDev in the link. Intuitively the DBP error can be connected with the difference between the actual PMD evolution and the one used in DBP. As clearly visible, the approximation comes closer to the actual DGDev, and therefore DBP has a potential to perform better.

5.2 DBP modification description

The steps of the modified DBP algorithm are described in Figure 5.2. The signal propagates from the left to the right.

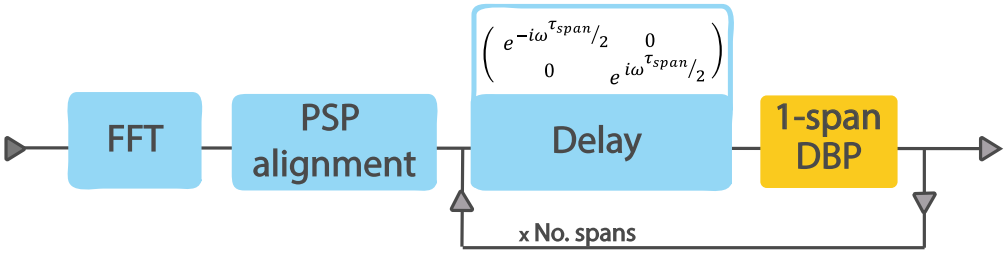


Figure 5.2: Proposed DBP modification block scheme. Orange DBP block depicts conventional DBP, τ_{span} is the DGD that corresponds to one step of the DGDev approximation function.

First, the signal needs to be converted to frequency domain through FFT. Then the signal in X and Y receiver axes, $l_{Rx,x}$ and $l_{Rx,y}$, needs to be rotated to the PSP, $l_{PSP,slow}$ and $l_{PSP,fast}$:

$$\begin{bmatrix} l_{PSP,slow} \\ l_{PSP,fast} \end{bmatrix} = \begin{pmatrix} \cos\zeta(\omega) & -e^{-i\phi(\omega)}\sin\zeta(\omega) \\ e^{i\phi(\omega)}\sin\zeta(\omega) & \cos\zeta(\omega) \end{pmatrix} \begin{bmatrix} l_{Rx,x} \\ l_{Rx,y} \end{bmatrix}, \tag{5.1}$$

here ζ and 2ϕ are the azimuth and elevation rotation angles on the Poincare sphere (see Figure 2.10). Details on how the rotation is performed can be found in Appendix A.

Next the delay between the x and y signal components, A_x and A_y, is adjusted by an exponential term $\exp(i\omega\tau_{span})$ in frequency domain (see Section 4.2.1):

$$\begin{bmatrix} \mathbb{A}_x \\ \mathbb{A}_y \end{bmatrix} = \begin{bmatrix} e^{-i\omega\tau_{span}/2} & 0 \\ 0 & e^{i\omega\tau_{span}/2} \end{bmatrix} \begin{bmatrix} A_x \\ A_y \end{bmatrix}, \tag{5.2}$$

Here, half of the delay is applied to the x component, half to the y component. The whole delay might also be applied to one component:

$$\begin{bmatrix} \mathbb{A}_x \\ \mathbb{A}_y \end{bmatrix} = \begin{bmatrix} e^{-i\omega\tau_{span}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_x \\ A_y \end{bmatrix}, \tag{5.3}$$

The conventional backpropagation of one span is then performed. Before backpropagating the next span the DGD is adjusted again.

The method described above will be further referred to as modified digital back propagation (DBPmod), proposed DBP or modified DBP.

5.3 Method performance

5.3.1 Simulation link scenarios

In this section we describe the types of systems that have been considered for numerical simulations. An 8 dBm launch power, DP signal is propagated over 30 non-DM spans, each consisting of an 80 km SSMF (see Table 4.1) and one noiseless EDFA (to exclude nonlinear signal-noise interactions) ideally compensating for 16 dB link loss. For all the cases the power of 8 dBm is chosen as being the optimum for the considered link with DBP compensation (i.e. giving the lowest SER for a system with 28 $GBaud$ DP-16-QAM signal propagating through a link with negligible PMD). The complete setup is summarized in Figure 5.3.

Simulations consider Gray coded modulations of 16- and 64-QAM, symbol rates of 28 $GBaud$ and 112 $GBaud$, and fibre PMDc of 0.1 and 0.04 $ps/\sqrt{km}$. Particularly, we simulate 16-QAM 28 $GBaud$, 64-QAM 28 $GBaud$, 16-QAM 112 $GBaud$ signals over currently deployed fibres with a typical 0.1 $ps/\sqrt{km}$ PMDc, and 16-QAM 28 $GBaud$ and 16-QAM 112 $GBaud$ signal over modern fibres with a typical 0.04 $ps/\sqrt{km}$ PMDc. The transmitter and local oscillator linewidths are set to zero. PMD is ideally compensated by measuring the Jones matrix of a fibre-link realization in the linear regime (with a few kHz resolution) and by applying its inverse in front of the shot-noise limited receiver with a matched filter and with a bandwidth twice as large as the one of the signal. Chromatic dispersion and nonlinearities are compensated with a SSF-

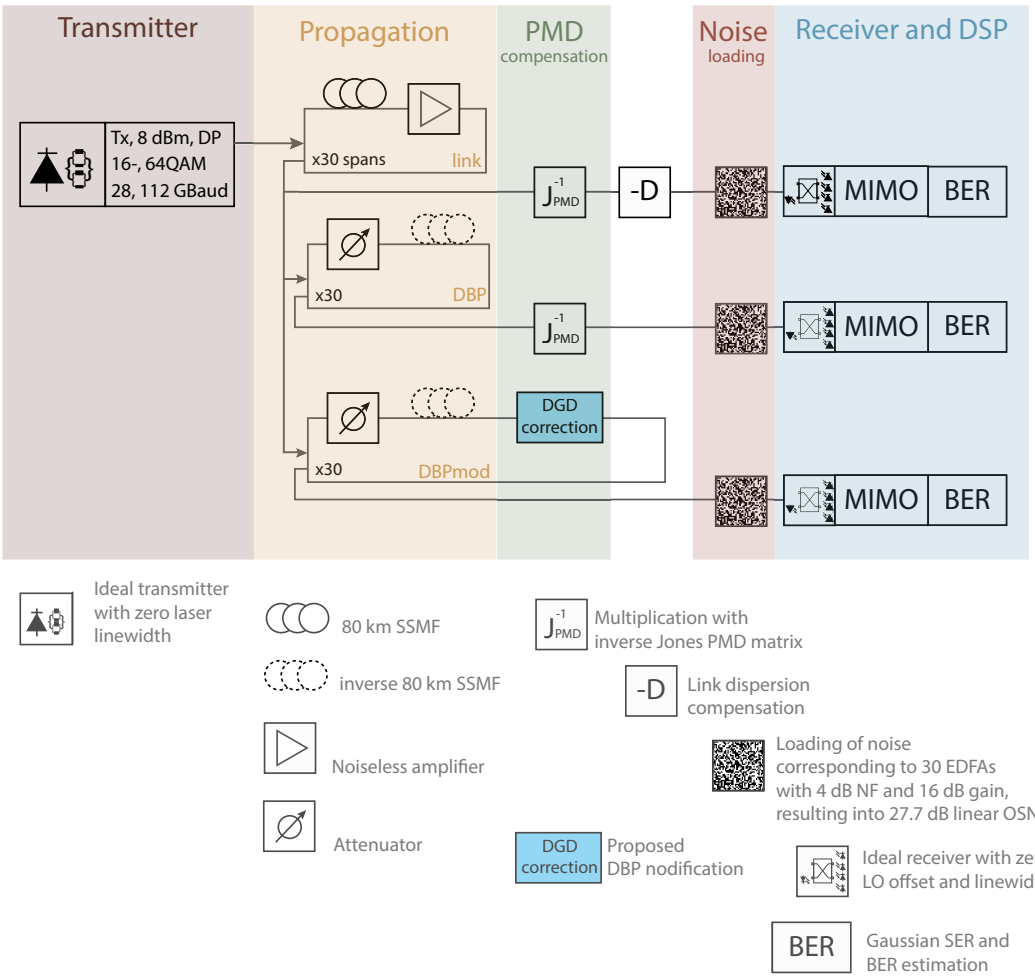


Figure 5.3: Simulation setup.

based DBP where the step size is controlled by a 0.05 ° MaxPhCh algorithm [49]. The number of steps in such implementation is comparable with the one used in SSF model for the signal forward propagation. The amount of loaded noise corresponds to 4 dB noise figure EDFAs with 16 dB gain: -19.7 dBm, so that the resulting linear OSNR is 27.7 dB. The noise is split equally between the two polarisations. Resampling to 1 sample per symbol and compensation of residual polarisation effects in the DSP is performed by a 15 taps MMA MIMO equalizer (see details in Appendix C). The number of taps is chosen to give the best performance by direct search.

5.3.2 Simulation results

Single channel

In Figures 5.4-5.7 each point corresponds to an averaged result of 20 to 55 different fibre birefringence profiles for a given interval of the DGD. Note that these profiles were preselected from a pool of 10^5 birefringence profiles, and probabilities of certain DGD intervals were estimated based on these statistics using Equation (2.31).

Linear Q factor was calculated from BER as in Appendix D. For convenience, we summarize in Table 5.1 some of the Q and corresponding BER values, values are rounded.

Table 5.1: table

BER	Q linear	Q, dB
$2 \cdot 10^{-2}$	2	3.1
$1 \cdot 10^{-3}$	3	4.9
$3 \cdot 10^{-5}$	4	6
$2 \cdot 10^{-7}$	5	7
$9 \cdot 10^{-10}$	6	7.8
$1 \cdot 10^{-12}$	7	8.4

The red dots mark the linear Q factor without DBP, orange circles – with conventional DBP (as in Section 2.1.3) and blue triangles – with the proposed DBP (as in Section 5.2). The standard deviation with respect to the mean

value is shown with the thin lines. The probability of the DGD to appear in certain intervals is indicated in the figures with vertical coloured stripes, the expectation $E(Q)$ of the linear Q factor (probability weighted sum of simulated results for a particular case) is printed to the right of the curves.

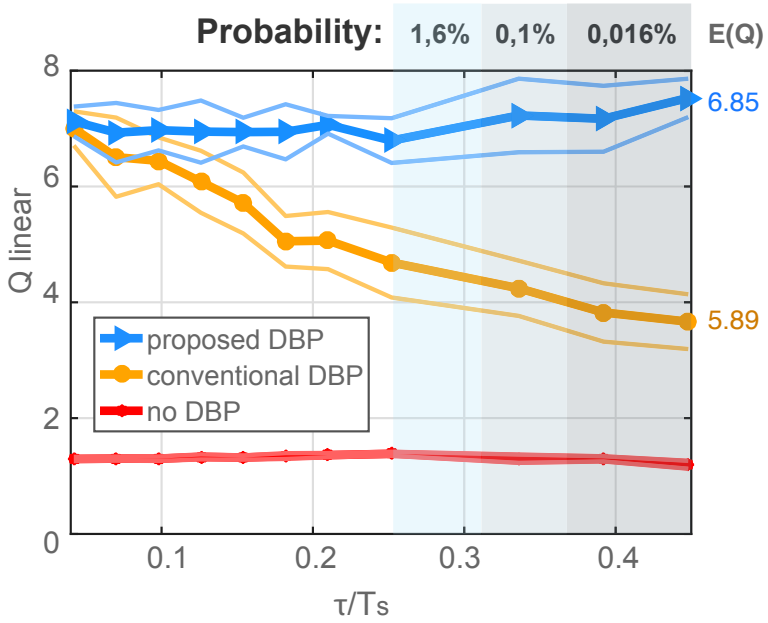


Figure 5.4: 0.1 $ps/\sqrt{km}$ PMDc fibre, DP-16-QAM, 8 dBm, 28 $GBaud$, 27.7 dB linear OSNR. T_s is the symbol period. "Probability" indicates the probability of accumulated DGD to appear in a certain interval, according to Equation (2.31) (shown with coloured vertical columns). $E(Q)$ is the expectation of the linear Q factor, i.e. probability weighted sum of the results for a particular simulation case).

It is visible in Figure 5.4 that growing DGD, τ , degrades the DBP performance. The proposed algorithm modification keeps the performance stable at the level when DGD is low. For the considered system, the expectation of the linear Q factor, $E(Q)$, can be increased by up to 1.

Similar tendency is observed for higher data rates as in Figure 5.5. In faded colours we additionally plot the 16-QAM 112 $GBaud$ system with four times lower noise loading, -25.7 dBm, resulting in 33.7 dB linear OSNR. Practically such OSNR can only be realised in considerably shorter links (for example, 8 spans with NF 4 dB), but in this case it is used to illustrate the following point: We found that with increased OSNR, and thus lower impact of

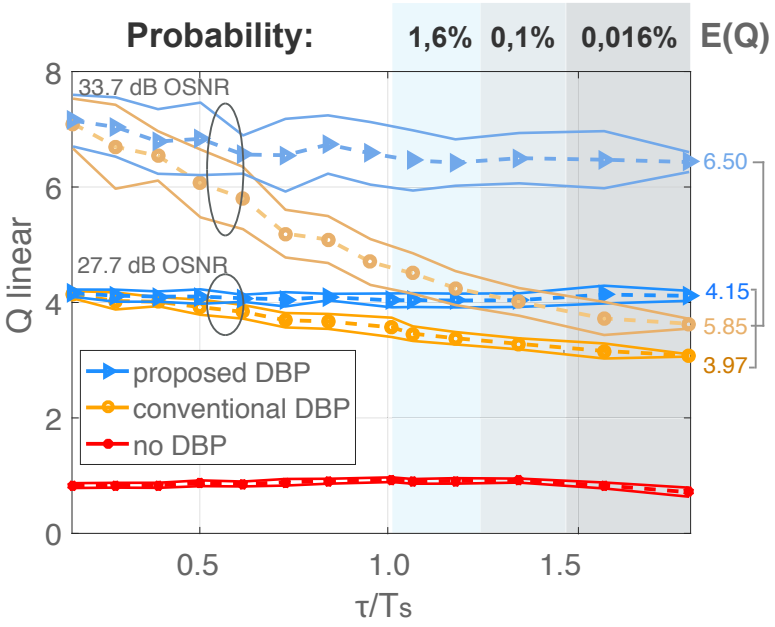


Figure 5.5: 0.1 $ps/\sqrt{km}$ PMDc fibre, DP-16-QAM, 8 dBm, 112 $GBaud$, at linear OSNRs of 27.7 dB and 33.7 dB.

noise, PMD-induced degradations become more significant and consequently, the benefit of the proposed DBP modification increases. Although the results for 33.7 dB linear OSNR in Figure 5.5 might look similar to 27.7 dB linear OSNR in Figure 5.4, the similarity is random and does not hold for other investigated cases. Note, that the fibre configurations are the same in two cases (only the signals are different), which explains the resemblance of performance variation. Also note the different scales of the two X axes (which, however, become identical if results are plotted against DGD value in ps).

Similarly to QPSK and 16-QAM, DBP of any higher-order modulated signal is affected by PMD, as, for example, shown in Figure 5.6 for 28 $GBaud$ DP-64-QAM 8 dBm signal.

Figure 5.7 shows the performance of DP-16-QAM 28 and 112 $GBaud$ signal transmission in a fibre with 0.04 $ps/\sqrt{km}$ PMDc (note different X axes, the ticks correspond to the same value of DGD in ps but to a different part of the symbol period). Depending on the data rate the limiting factor is either noise or PMD.

Generally, when identifying the system limiting factor, two situations are possible. If the system has a large OSNR margin, the PMD will significantly

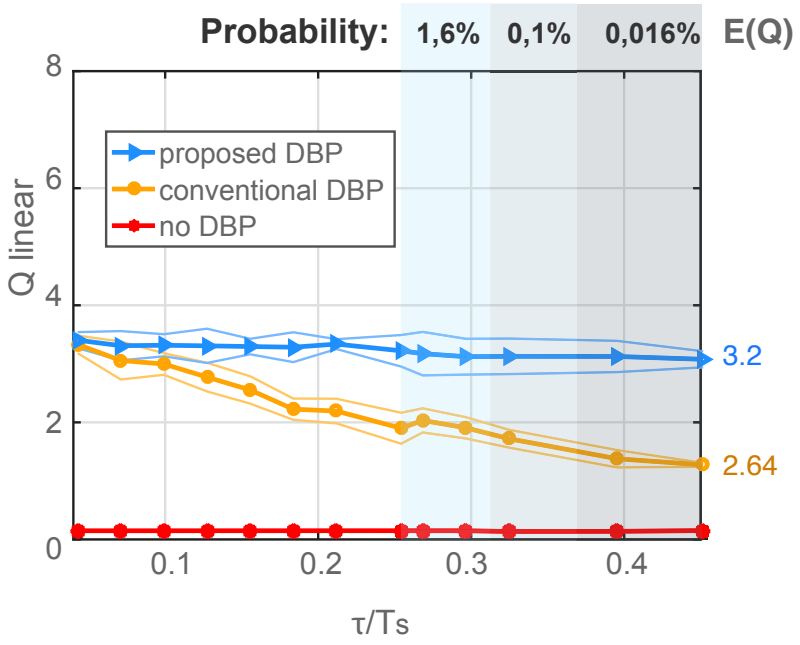


Figure 5.6: 0.1 $ps/\sqrt{km}$ PMDc fibre, DP-64-QAM, 28 $GBaud$, 27.7 dB linear OSNR.

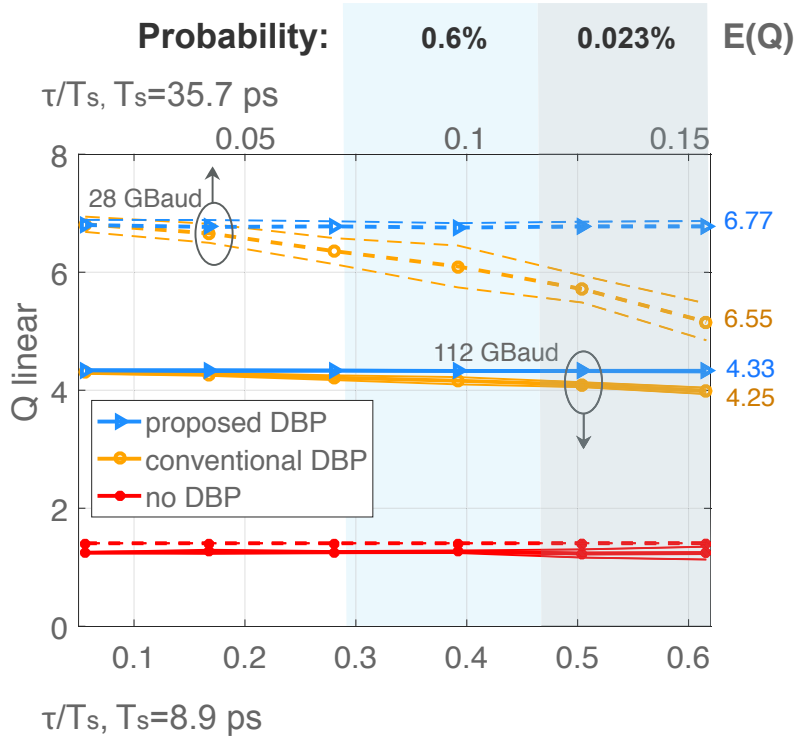


Figure 5.7: 0.04 $ps/\sqrt{km}$ PMDc fibre, DP-16-QAM, 8 dBm, 27.7 dB linear OSNR, upper X axis for 28 and lower – for 112 $GBaud$ cases.

decrease the performance of DBP. However, if the system is already operating at the highest tolerable noise levels, the PMD impact will be only marginal.

Gains achieved for single channel DBPmod in high-DGD links

Hereby we investigate in detail the extreme cases of high DGD, 12 to 17 ps with 0.016% probability for 0.1 $ps/\sqrt{km}$ PMDc fibres, and 5 to 6 ps with 0.023% probability for 0.04 $ps/\sqrt{km}$ PMDc fibres. Table 5.2 shows the linear gain of the proposed DBP modification. In other words, it shows the difference between the linear Q factor of the conventional and the proposed DBP. The cases when conventional DBP performance would be considered as outage are printed in bold. The system outage is defined by Q factor lower than 2.67 (linear), corresponding to $3.8 \cdot 10^{-3}$ BER [87]. If the value is not printed, the DBP and the DBPmod both failed to provide the minimum required performance (due to PMD).

Table 5.2: DBPmod gains (regarding linear Q) compared to conventional DBP for typical transmission systems. Those cases, where BER after conventional DBP will be higher than $3.8 \cdot 10^{-3}$, are highlighted in **bold**. Those are the cases, where DBPmod is required to satisfy the today's hard-decision FEC performance criteria.

	30 spans				50 spans	
	16-QAM		64-QAM		16-QAM	
PMDc, ps/$\sqrt{\text{km}}$	0.04	0.1	0.04	0.1	0.04	0.1
Signal						
28 $GBaud$, 8 dBm	1.66	3.5	**0.86**	**1.72**	1.76	2.88
112 $GBaud$, 8 dBm	0.26	**2.86**			**0.28**	
112 $GBaud$, 14 dBm	**3.42**				**2.18**	

The table considers four different links, differing by length and the fibre PMDc: links of 30x80 km and 50x80 km, and 0.1 and 0.04 $ps/\sqrt{km}$ PMDc fibres. Two signal symbol rates are considered, 28 and 112 $GBaud$ along with two modulation formats, 16-QAM and 64-QAM. Further, different powers are chosen for 112 $GBaud$ signal in order to match either the optical power or power spectral density of the 28 $GBaud$ signal. For 30 loops for cases of

8 dBm and 14 dBm signal powers the linear OSNRs are 27.7 dB and 33.7 dB respectively, for 50 loops – 25.5 dB and 31.5 dB respectively.

The gain of DBPmod depends on how pronounced the PMD effect on DBP is. In Table 5.2 we see that the gains for links with higher PMDc are always larger. Also the cases with higher OSNR will show higher gains.

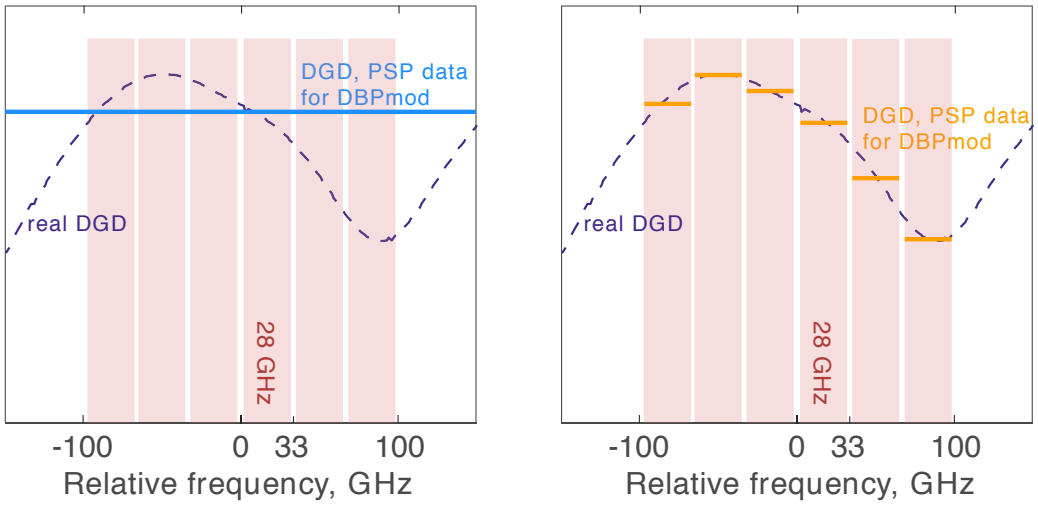
For some systems, such as 30 loops deploying 28 $GBaud$ 64-QAM or 50 loops deploying 112 $GBaud$ 16-QAM signal, DBPmod allows avoiding system outage in the simulated cases. Further, for 0.1 $ps/\sqrt{km}$ systems at high DGD regimes, typical linear gains of approximately 2 can be expected.

Multichannel

In order to test the implementation of the modified DBP for the WDM propagation, we consider the single channel setup in Figure 3.8 with slight modifications. Firstly, the signal now consists of six 28 $GBaud$ 16-QAM channels, placed on a grid of 33 GHz with a central frequency of 193.1 THz. Each channel has a raised-cosine pulse shaping with 0.01 roll-off factor. Each channel has an optimum 2 dBm optical power (defined by the power that results into lowest BER for the 4^{th} channel after conventional DBP and without PMD in the link).

The algorithm performance is estimated on the 4^{th} channel (chosen randomly) at frequency 193.1 $THz + 16.5$ GHz. The conventional DBP is performed on the whole optical field, so that both inter- and intrachannel effects are compensated. The receiver has a matched filter (raised-cosine with 0.01 roll-off factor) at the frequency of the 4^{th} channel. The DBP is still performed full-field as it is done before the receiver. Practical implementation of such setup is not considered here.

There are two possibilities considered for the DBPmod. First one, called further frequency-independent, performs DGD correction with the data (PSP and DGD) measured at the central frequency. Second option, called respectively frequency-dependent, uses data measured for each of the 6 channels' central frequencies. The principles are illustrated in Figure 5.8. The results of the transmission simulation of the 4^{th} channel are plotted in Figure 5.9. Note that the three curves converge to one point at zero DGD. However, points with close-to-zero DGD have low probability and were not specially simulated and



(a) Frequency-independent DBPmod implementation.

(b) Frequency-dependent DBPmod implementation.

Figure 5.8: Two ways to implement proposed DBP for a multichannel transmission.

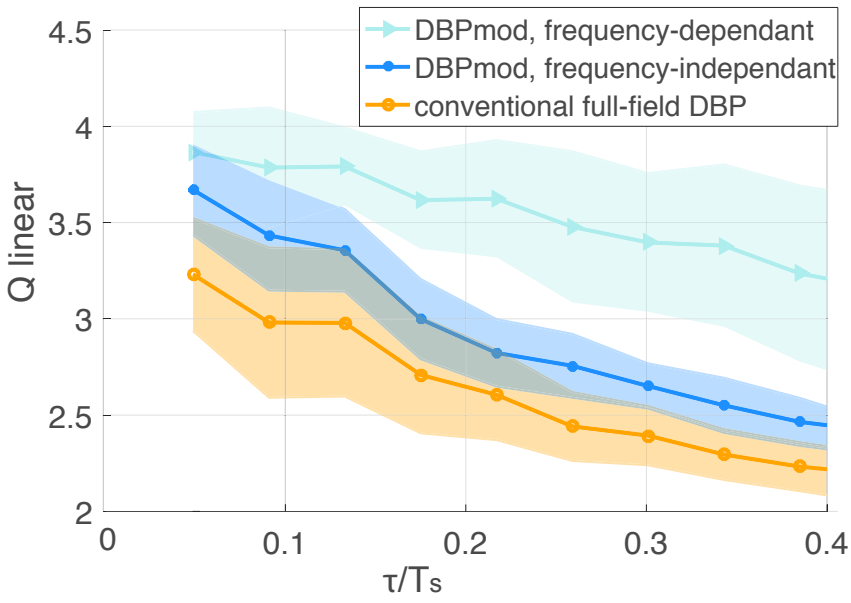


Figure 5.9: Performance of the 4th channel of a six channel transmission, 0.1 $ps/\sqrt{km}$ PMDc fibre, 28 $GBaud$ DP-16-QAM, 2 dBm, 27.7 dB linear OSNR.

are therefore not plotted.

The conventional DBP performance of one 28 $GBaud$ channel in a multi-channel 198 $GBaud$ transmission declines faster than the DBP performance of a single-channel with bandwidth of 112 $GBaud$ (see Figure 5.5). The strength of decline is rather similar to the single channel transmission of 28 GHz bandwidth, as in Figure 5.4. The proposed DBP in the frequency-independent implementation provides some gain. However, the frequency dependence is crucial for considerable performance gain. Especially in the high DGD regimes, effect on each channel's performance strongly depends on the frequency and varies significantly from channel to channel. In order to fully compensate for the PMD effect, denser DGD and PSP measurement discretization might be required.

5.4 Practical aspects of modified DBP implementation

5.4.1 PSP and DGD measurement

In order to perform the steps described above it is necessary to know two physical values, DGD and the PSP vector components. We found at least two approaches in the literature to extract the values: to analyse the response of the adaptive MIMO filter [66, 67] or to introduce additional polarised training sequences and perform a polarisation analysis [88–90]. We will further comment on both methods one after another.

Monitoring adaptive MIMO filter taps

According to Equation (2.38) derived in Section 2.3.6 the linear propagation characteristics of a fibre can be represented as a frequency-dependent channel transfer matrix, called Jones matrix [64], consisting of dispersion $\mathbb{D}(\omega)$, PDL $\mathbb{K}(\omega)$ and PMD $\mathbb{U}(\omega)$ elements:

$$H(\omega) = \mathbb{D}(\omega)\mathbb{U}(\omega)\mathbb{K}(\omega). \tag{5.4}$$

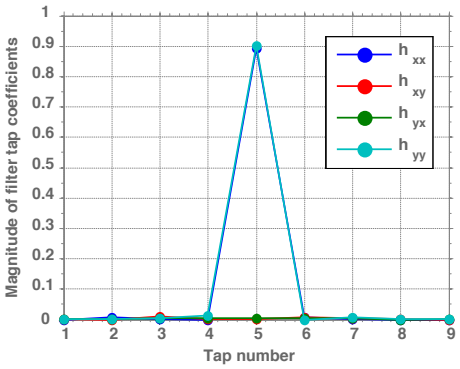
Once the MIMO algorithm converges a matrix is produced as:

$$M(\omega) \approx H^{-1}(\omega) = \left(DFT\left(\begin{bmatrix} h_{xx}(m) & h_{xy}(m) \\ h_{yx}(m) & h_{yy}(m) \end{bmatrix} \right) \right)^{-1}, \tag{5.5}$$

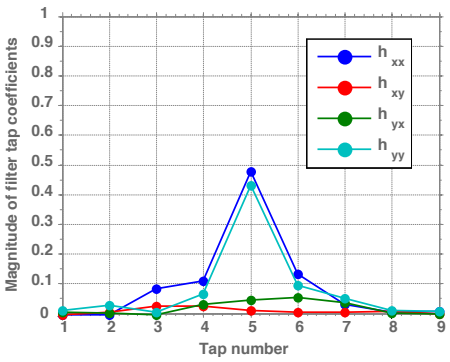
$$h_p(m) = [h_{p,0}(m), h_{p,1}(m), ..., h_{p,N_{taps}-1}(m)]^T, \tag{5.6}$$

where DFT denotes discrete Fourier transform (implemented, for example, employing FFT), h_p are the tap coefficients for the m^{th} signal sample, and p takes values of xx, xy, yx, yy. Thus, each signal frequency sample has four weighting coefficients (xx, xy, yx, yy) for each of the N_{taps} taps. More information can be found in Appendix C.

An example of how the filter coefficients look like is given in Figure 5.10 for low and high DGD. While by low DGDs there is little mixing between the x and y signal components and neighbouring taps, at high DGDs, neighbouring taps and mixing in polarisations need to be considered.



(a) MIMO taps to compensate low DGD.



(b) MIMO taps to compensate high DGD.

Figure 5.10: Example of MIMO filter coefficients.

Finally, following [66, 67] and Section 2.3.6 Equation (2.38), and assuming there is no PDL, $M(\omega)$ provides a basis to extract DGD and PSP:

$$M(\omega + \Delta\omega)M(\omega)^{-1} = \mathbb{D}(\omega)R_1^{-1}(\omega) \begin{bmatrix} e^{i\Delta\omega\tau/2} & 0 \\ 0 & e^{-i\Delta\omega\tau/2} \end{bmatrix} R_1(\omega), \tag{5.7}$$

where $\mathbb{D}(\omega)$ is a term defining the chromatic dispersion matrix. R_1 is a rotation matrix as in Section 2.3.6 Equation (2.41) and τ is the DGD. The

DGD at frequency ω can then be estimated with the help of eigenvalues of matrix $M(\omega + \Delta\omega)M(\omega)^{-1}$, ρ_1 and ρ_2:

$$\tau = \left| \frac{\arg(\frac{\rho_1}{\rho_2})}{\Delta\omega} \right|. \tag{5.8}$$

The PSP components at frequency ω correspond to the eigenvectors of matrix $M(\omega + \Delta\omega)M(\omega)^{-1}$, $t_{slow,fast}$:

$$t_{slow} = [s_x, s_y], \tag{5.9}$$

the slow Stokes PSP vector $\mathbb{S}_{slow}$ in Stokes space (see Section 2.3.5) is then:

$$\mathbb{S}_{slow} = [1, s_x s_x^* - s_y s_y^*, s_x s_y^* + s_y s_x^*, i(s_x s_y^* - s_y s_x^*)]. \tag{5.10}$$

Data-aided monitoring

The measurement of PSP and DGD was demonstrated experimentally in [88] by measuring the difference between fibre output polarisation states of two random input polarisations with an electro-optic polarimeter.

Another data-aided PSP and DGD measurement method proposed in [89] suggests sending a monochromatic linearly polarised optical field in a fibre and receiving the X and Y components separately at different receiver polarisations. The interference map allows retrieving the DGD by its amplitude. The angle between the PSP and the polarisation beam splitter (PBS) axis is retrieved from the phase distribution of the interference map.

Finally, the MIMO filter taps can be considered for DGD and PSP extraction in the same way as described in the previous section but with a predefined sequence. As in [90], a constant-amplitude zero-autocorrelation (CAZAC) sequence might be used, especially effective with a guard interval of 25% of the training sequence length. With as low as 16 training symbols and receiver-side sampling of 2 samples per symbol it was shown that DGD can be estimated with a standard deviation of 7.7 ps at mean 25 ps.

5.4.2 Computational complexity

The computational complexity of the method is determined by the effort to perform a single signal rotation to a new coordinate system, a single complex multiplication per span for the DGD compensation, and calculation or measurement of DGD and PSP.

If the DGD and PSP are measured through the MIMO taps analysis, the complexity can be summarized as in Table 5.3 with the help of Big-O notation [91,92]. For the Big-O complexities chart see Appendix E. In the Table N_{taps} is the number of MIMO taps, n refers to the calculation precision, N_{fft} is the FFT size, N_{spans} is the number of spans. The highest complexity of the algorithm is $O(n^2)$. For the DGD and PSP values high precision is not required, which results in an extreme low-complexity of the algorithm.

5.5 Comparison with other existing methods

Generally, the impact of PMD on DBP can be compensated in a straightforward way if the information about the DGD and polarisation states along the fibre is fully available as, for example, demonstrated in simulation [29]. In that case the signal polarisation components during DBP should be rotated and delayed "backward" at each step following the local birefringence profile of the fibre. The technique was simulated in [29] by tracking the polarisation state with Jones matrices on a span-by-span basis. Such approach, however, is extremely challenging to realize in practice, as the measurement data are not available.

A feasible approach, besides the one proposed in this thesis and developed parallel to it, was presented recently [24]. Both are based on the approximation of the DGDev along the link and are using this approximation while performing DBP. In the method presented here (see [27], Chapter 5) the DGD evolution is predicted based on a measurement of DGD and PSP at the receiver at the central channels' frequencies. In the other method [24] DGDs are measured at the receiver for each of the transmission channels', and an inverse channel PMD Jones matrix is constructed iteratively, optimizing PMD sections' polarisation scramblers and retardation plates.

The two methods use a different approach for predicting the PSP evolution.

Table 5.3: Complexity of the proposed DBP modification.
Mult stands for multiplication, impl. – for implementation.

		Operation	**Times**	**Complexity**	**Notes**
Calculations from MIMO taps	Preparations for one sample:	DFT	4	$O(nlog(n))$	radix 2 impl. $n = N_{taps}$
		matrix inversion	$2/3N_{taps}$	$O(n^3)$	$n = 2$
		eigenvectors	$2/3N_{taps}$	$O(n^3)$	$n = 2$
	DGD	6 scalar mult	1	$O(n^2)$	Using Gauss' complex mult algorithm $n = N_{digits}$
		1 scalar division	1	$O(n^2)$	$n = N_{digits}$
	PSP	4 scalar mult	1	$O(n^2)$	$n = N_{digits}$
	Periodicity factor			reduces complexity by the factor of necessary periodicity of renewing DGD and PSP values	
DBP modification	τ step	1 scalar division	1	$O(n^2)$	$n = N_{digits}$
	Axes rotation	8 scalar mult	1	$O(n^2)$	$n = N_{digits}$
	DGD correction	2 scalar mult	$N_{fft} \cdot N_{spans}$	$O(n^2)$	$n = N_{digits}$

While in this work estimated PSPs are assumed constant along the link, in the other case the PSP evolution is constructed in a way that it "results" into the measured DGD. However, such evolution does not coincide with the actual PSP evolution and can have a very different shape.

Regarding the implementation constrains, in this thesis the PSP should be measured at the receiver, for example, by adding a training sequence to the data header [88–90] that is analysed by a polarimeter or by extracting the Stokes vector from the MIMO digital filter taps [66, 67]. In the other proposed method, DGD also has to be estimated but not the PSP, those are reconstructed, and the construction of the Jones Matrix puts computational load on the system.

A recent development of the method [24] has tried to reduce the computational complexity [34,35]. It proposed to extract the link PMD Jones matrices for each of the transmitted frequencies (corresponding to Fourier frequencies of the inverse Jones matrix) from the time-domain complex-valued finite impulse response (FIR) filter. Next step, in order to get the partial PMD Jones matrices corresponding to the part of the link, the link PMD Jones matrix for each frequency is expanded into a Taylor series around the central frequency. It then goes through a cascade of approximations. As the result, a partial correction matrix is obtained. It is applied in two steps: partially during DBP and partially after DBP. Finally, after DBP a correction matrix is applied to compensate for the tight reference to the central frequency, which is important for large bandwidths. Additionally, the following equalizer feeds the estimated Jones matrix of residual PMD at a specific frequency back to the DBP in a feedback loop. The frequency interval for the PMD Jones matrices calculation is estimated at frequencies defined by the equalizer taps and sampling rate. Note, for the description in [34, 35], the method for f_0 central frequency behaves as conventional DBP with PMD compensated already after DBP. This, however, might only be the descriptive, not the implementation issue.

The PMD is corrected in two steps in [34, 35]. We would like to introduce a point for discussion about the way the first PMD correction term is applied. The part that is applied after DBP is the same for all frequencies and is namely the link Jones matrix at the central frequency. The part applied during DBP accounts for the frequency dependence of the partial Jones matrices. The

idea of compensating for PMD between the series of DBP steps implies that the optical field is corrected accordingly. Even if the correction terms have a common multiplier, it cannot be "taken out of the brackets", because it will affect the optical field and, therefore, DBP performance. Although not proving the concept, we imply that the method performance could possibly be improved if the correction terms are slightly modified. Using the notation in [34,35] Figure 1, the first PMD operator before DBP would become $J_1^{new}(f) = J(f_0)^{1/N_{PMD}} \cdot \exp(\frac{(f-f_0)J^{-1}(f_0)J^{'}(f_0)}{N_{PMD}})$ and the second PMD operator after DBP $J_2^{new}(f) = J_c(f)$. Here N_{PMD} is the number of PMD corrections, J is the inverse Jones matrix. Such approach would also solve inconsistency of the method for the central frequency f_0. As described in [34,35], the algorithm at f_0 would not compensate for the PMD along the link, but would perform all the compensation after DBP.

Another interesting point for discussion would be if the the partial PMD Jones matrix can be constructed without reference to the central frequency. As the total Jones matrix is measured by the equalizer at every frequency f, the partial matrix could be constructed as $J_{part}(f) = J(f)^{1/N_{PMD}}$, so that when the operator is applied N_{PMD} times, it results into $J(f)$, the estimated link PMD Jones matrix. In light of these ideas we suggest that the methods proposed in this thesis can also be further simplified. An open question remains if such approach would reduce the complexity and if so, to what extend.

The major difference of the two methods, the one discussed above and presented in this thesis, is the way to approximate the reverse PMD evolution. While the earlier work of the group [24, 34] differed only in the PSP prediction evolution, the later work [35] differs in both DGD and PSP prediction. Because the Jones matrix is no longer decomposed to rotation and time delay components, the whole matrix gets approximated, without regard of what happens to its components. Further, the approximation is tightly bonded to the central frequency.

The consequence of the approximation process is the methods' complexities. The one in this thesis involves channel transfer matrix measurement at central frequency and its decomposition, and PMD correction terms application. The other one requires channel transfer matrix measurement for a range of transmission frequencies, PMD correction terms calculations and applications.

The latter method requires considerable computational resources, challenging to realise in real-time DSP.

We finally notice that the manner of approximating PMD evolution in both cases allows performing DBP at the receiver only.

The two methods comparison is summarized in Table 5.4.

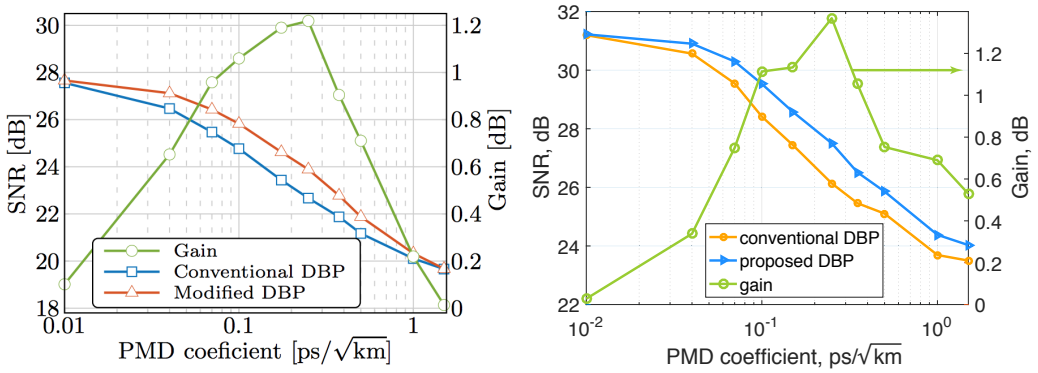
Table 5.4: Comparison of DBP proposed in this thesis and in [35].

	DBP modification in this thesis	**DBP modification in [35]**
Source of DGD	Inverse channel Jones matrix	Constructed partial Jones matrices
Source of PSP	Inverse channel Jones matrix	Constructed partial Jones matrices
Values extracted from FIR filter taps	DGD, PSP	–
Computational steps	Find eigenvalue and eigenvector of the inverse channel Jones matrix for central frequency	Construct an inverse PMD Jones matrix for central frequency, expand it in Taylor series and simplify to obtain two partial correction terms.
Application of DGD correction term	Multiplication of signal samples with an exponential term once per span	Multiplication of signal samples with the correction matrix once per span, multiplication with two correction matrices after DBP
Performance 50 GBaud channel (See Section 5.5.1)	Similar	
Performance 350 GBaud channel (See Section 5.5.2)	Further investigation is required	

5.5.1 Single-channel performance comparison

In this section we compare the performance of the DBP modification presented in this thesis and the one described in [35]. The comparison will be made for Figure 5a from [35], plotted for the same transmission system.

The investigated system consists of 10x100km SSMF spans, one per span amplifier with 4.5 dB NF, ideally compensating for loss, fibre PMDc takes values 0.01, 0.04, 0.07, 0.1, 0.15, 0.25, 0.35, 0.5, 1.0, 1.5 $ps/\sqrt{km}$ (see Appendix B and Figure B.1a particularly). For each case we perform 100 simulations of random fibre birefringence. The signal is DP-16-QAM 50 $GBaud$ shaped using a root-raised cosine with roll-off factor 0.01. For each fibre PMDc the optimum power is chosen (resulting in lowest BER), which ranges from 14 to 8 dBm. PMD correction is performed once per span, i.e. 10 times during DBP. The SNR value was calculated from the measured error vector magnitude (EVM) (described in Appendix F). The simulation results are plotted in Figure 5.11 for DBPmod presented in this thesis (Figure 5.11b) and proposed in [35] (retrieved in Figure 5.11a).



(a) Results from Fig. 5. (a) [35].

(b) DBP proposed in this thesis.

Figure 5.11: Optimum SNR and SNR gain versus the fibre PMDc at the optimum input power. Each point is an average of results for 100 fibre realisations. DP-16-QAM signal 50 $GBaud$.

The results for the modified DBP gain are to a great extent similar. The zero PMD case as a starting point for the curves has an approximately 3 dB difference in SNR. This might be due to better adjusted matched filtering and time synchronisation parameters. We also perform DBP in the optical

domain, and compensate for PMD ideally using the inverse Jones matrix, which might give better results. Nonetheless, the gain values are only slightly better than the ones in [35], which is probably also a statistical issue (specific fibre birefringence realisations). The maximum gain is achieved at 0.25 $ps/\sqrt{km}$ PMDc in both cases and is 1.37 dB for DBP proposed in this thesis, and 1.22 dB in [35].

The comparison above proves that for the single-channel and the systems with variable fibre PMDc (see Appendix B), both methods show similar results.

5.5.2 Multi-channel performance comparison

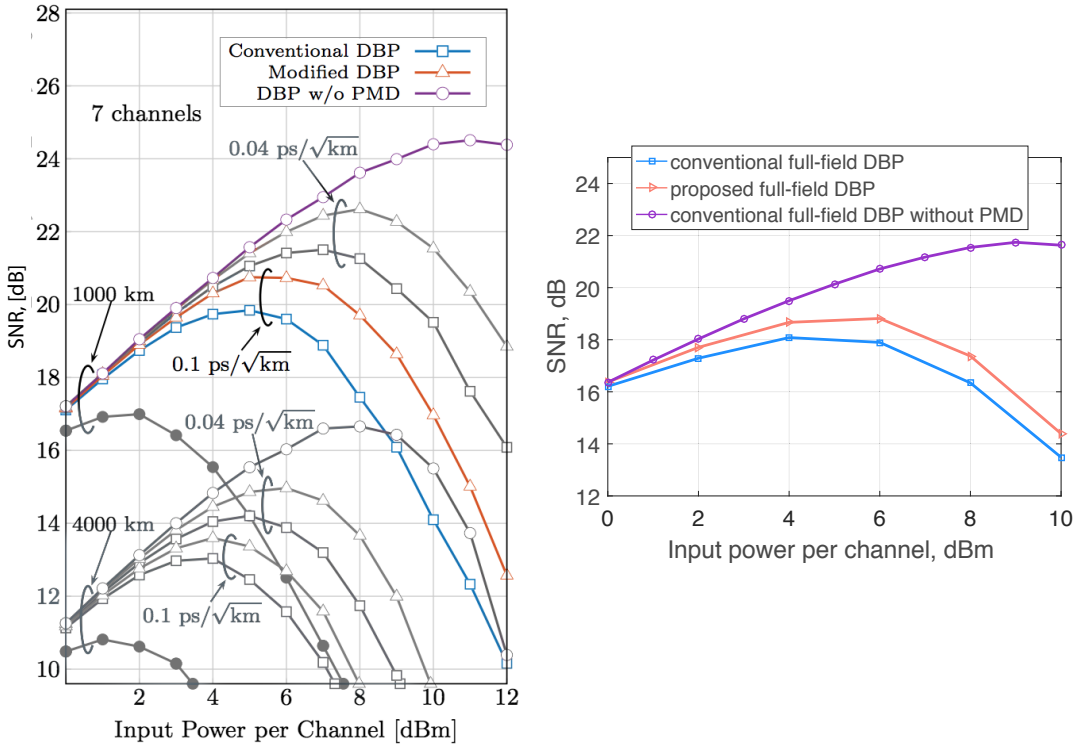
Here we will compare the performance of the multichannel DBP modification presented in this thesis and the one described in [35]. The comparison will be made for Fig. 3. (b) from [35], plotted for the same transmission system.

The investigated system consists of 10x100 km SSMF spans, one per span amplifier with NF 4.5 dB, ideally compensating for loss, fibre PMDc takes value of 0.1 $ps/\sqrt{km}$. For each case we perform 100 simulations of random fibre birefringence. The signal consists of 7 DP-16-QAM 50 $GBaud$ signals, each shaped using a root-raised cosine with roll-off factor 0.01. The 4^{th} channel is at the central frequency of 193.1 THz. Channel spacing is 50.1 GHz. Optical input power per channel ranges between 0 and 10 dBm with a step of 2 dBm. PMD correction is performed once per span, i.e. 10 times during DBP. DBP is performed on the full 350.7 GHz bandwidth, which might not be practically implementable but serves here for the method performance evaluation.

The simulation results in Figure 5.12b are compared to the respective data from [35] Figure 5.12a. The performance of all three cases, propagation without PMD, with DBP, and with DBPmod, shows lower SNRs. This might be the reason for the lower DBPmod gains: 0.92 dB for 6 dBm and 1 dB for 8 dBm for this work compared to 1.1 dB for 6 dBm and approximately 2.1 dB for 8 dBm for [35]. More investigation at similar SNR levels is required.

5.6 Conclusions

In this chapter a modification of the conventional DBP based on a single measurement of the DGD and PSP at the receiver was proposed. The im-



(a) Results from Fig. 3. (b) [35]. The original picture was modified in that the irrelevant curves were bleached, and the legend was added.

(b) DBP proposed in this thesis.

Figure 5.12: Average SNR of the 4$^{\text{th}}$ channel of 7-channels transmission depending on the input power. The channel is DP-16-QAM 50 *GBaud* propagating over a 1000 *km* link.

plementation process, including multichannel, and the complexity of the new algorithm were discussed. The new algorithm keeps the performance at "no PMD" levels, which was verified by simulations on a number of typical transmission systems. In high DGD regimes typical gains of the linear Q factor of approximately 2 can be expected (corresponds to 5.8 dB). The impact of the noise on the "visibility" of the NL-PMD limit was confirmed.

The proposed algorithm was compared with a DBP modification suggested recently and developed independently from this work [35]. The algorithm proposed in this thesis shows similar performance improvement and has lower complexity.

In the method proposed in this thesis we assume exact knowledge of DGD and PSP, which in reality might be measured with error. The next chapter will investigate this question. Note, however, that method [35] will be affected by the accuracy of the channel transfer matrix measurement as well.

Chapter 6

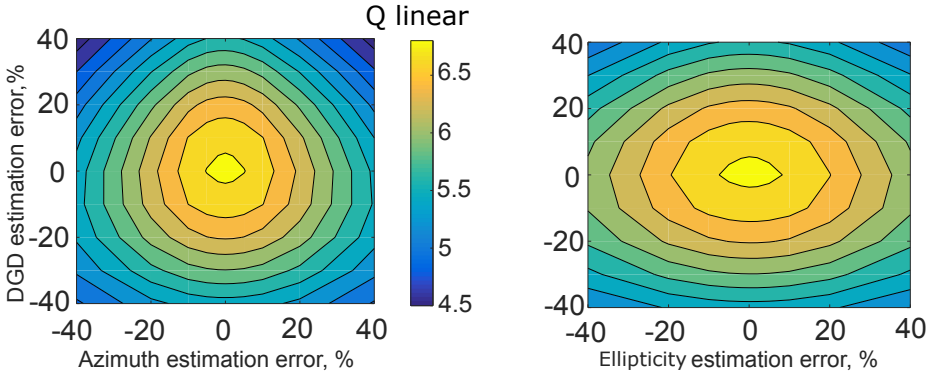
Required accuracy for modified DBP

In this section we investigate what accuracy of the measured DGD and PSP is required for the effective performance of the modified DBP. In the simulations in Section 5.3 we assumed the exact knowledge of the values. We will also show how the DBP implementation affects the performance of the proposed method.

6.1 Measurement accuracy

In order to investigate how tolerant the DBPmod is to the measurement inaccuracy, we artificially introduce a mismatch in DGD and PSP parameters. Both experience 2 to 40% under- and overestimation with a 5% step. For example, 25 % overestimation means that the measured value is 0.25 times bigger than the real one. The PSP estimation error is considered in azimuth and ellipticity on the Poincare sphere.

Figure 6.1 shows Q factor estimation of a 28 $GBaud$ 16-QAM 0.1 $ps/\sqrt{km}$ PMDc 2400 km system for a high DGD interval of $13 - 17$ ps or $0.36 - 0.47$ symbol periods (see Figure 5.4 blue curve). Inaccuracies of up to 2% do not change the performance, inaccuracies of up to 10% decrease the linear Q factor by less than 0.2. The method is most tolerant to estimation error of PSP ellipticity, followed by DGD and then PSP azimuth.



(a) PSP estimation error in azimuth.

(b) PSP estimation error in ellipticity.

Figure 6.1: Average Q factor depending on the DGD and PSP estimation error for a transmission system as in Figure 5.4.

6.2 DBP implementation accuracy

As shown in Figure 5.5 and 5.7, the significance of the PMD impact on the DBP performance depends on the system noise level. We investigate the penalty that occurs in a DBP with realistic number of steps per span.

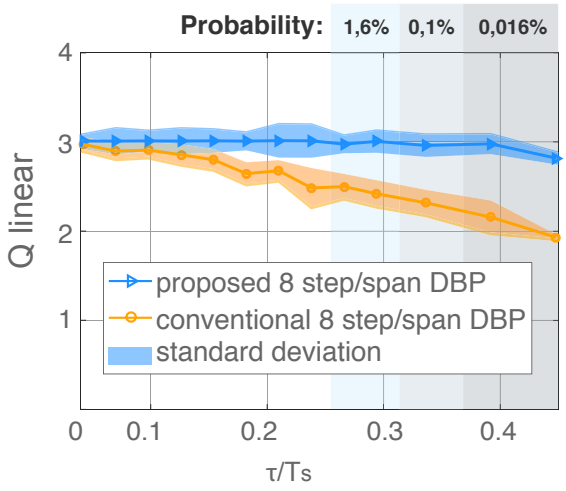


Figure 6.2: Original and modified DBP performance with average 8 steps per span and no noise loading, transmission system as in Figure 5.4. DGD is adjusted 1 time per span in proposed DBP.

DBPs in Figure 5.4 and 6.2 differ by the number of DBP steps and noise loading. While in Figure 5.4 DBP with hundreds of steps (tolerated MaxPhCh

of 0.05°) is investigated, the DBP in Figure 6.2 uses an average of 8 steps per span (tolerated MaxPhCh of 5°). Additionally, the DBP in Figure 6.2 has neither noise loading before the Rx (as in the setup scheme in Figure 5.3), nor EDFA-added ASE noise. This is to allow the system having reasonable OSNR (OSNR is different from case in Figure 5.4). Would the noise be simulated, the system would show outage. In case of non-ideal DBP the uncompensated nonlinear noise defines the system limits, although the PMD effect is still present.

6.3 MIMO DGD measurement accuracy

6.3.1 MIMO filter requirements

We investigate here one of the methods to measure the DGD by analysing the digital MIMO coefficients (see Appendix C). To ensure sufficient MIMO accuracy the signal has to be oversampled and the MIMO needs to have a sufficiently large number of taps. An oversampling factor of S_r samples per symbol leads to a MIMO resolution of:

$$r = \frac{T_s}{S_r}, \tag{6.1}$$

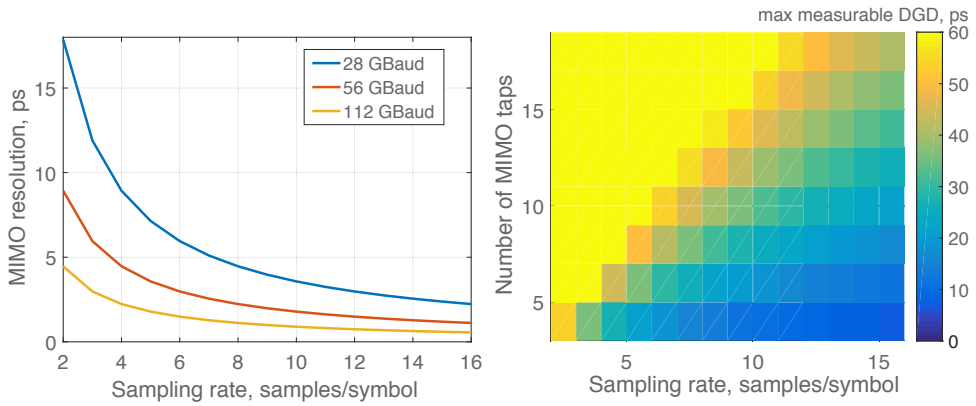
in time domain, where T_s is the symbol duration and r in *seconds* defines the minimal available accuracy of DGD measurement. The MIMO memory (in time domain), or the maximum DGD that can be measured, is defined by the number of taps N_{taps}:

$$\Delta\tau_{max} = N_{taps} \cdot r \overset{!}{>} \Delta\tau \tag{6.2}$$

Higher sampling rates and number of taps are beneficial for the DGD and PSP estimation. There are minimum MIMO taps requirements for specific transmission systems. For example, a 112 $GBaud$ signal, sampled with 8 samples per symbol will allow a MIMO resolution of 1.1 ps, according to Equation (6.1). This, however, results into an unrealistically high receiver required sampling rate. A 28 $GBaud$ signal would require an oversampling factor of 32 samples per symbol for the same accuracy. In currently deployed fibres of length 2400 – 4000 km with a PMDc= 0.1 $ps/\sqrt{km}$ the comprised

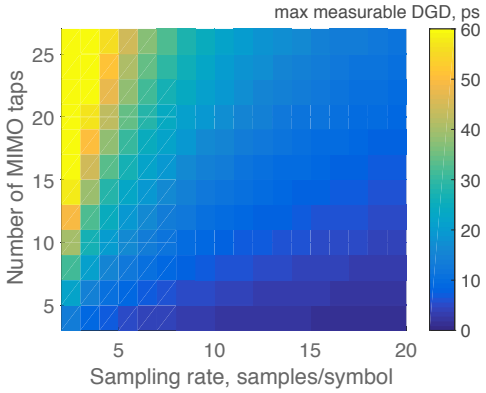
DGD ranges from 0 to approximately 15 ps and in modern fibre with PMDc 0.04 $ps/\sqrt{km}$ – from 0 to 6 ps (see Figure 3.5). In order to cover these ranges the MIMO filters need to have a channel memory that is long enough. Having a resolution of 1.1 ps will require 15 taps or 7 taps respectively. These requirements are summarised in Figure 6.3.

Note that the estimation of DGD and PSP in a multichannel system still holds on the Equation (6.1) and (6.2).



(a) MIMO taps resolution depending on sampling rate and signal symbol rate.

(b) Maximum measurable DGD in case of 28 $GBaud$ signal.



(c) Maximum measurable DGD in case of 112 $GBaud$ signal.

Figure 6.3: Influence of DBP and MIMO implementation on the modified DBP performance.

6.3.2 Emulation of DGD estimation

In order to check how accurate the DGD can be extracted from the MIMO taps we emulate a 28 $GBaud$ signal with 8 samples per symbol sampling. Following the waveplate model (Section 2.4.2), PMD was emulated using 40 plates per link and random polarisation scrambling between them. The section's DGD is chosen from a normal distribution with a mean of $< \Delta\tau > /N_{c-s}$ and standard deviation of 20%. The DGD is estimated from the constructed PMD Jones matrix. The MIMO FFT size is 1024 samples, the number of MIMO taps is swept from 3 to 65; the MIMO resolution according to Equation (6.1) is then 4.4 ps. The estimated deviation comes exclusively from the MIMO architecture, no receiver noise and no additional propagation effects are assumed.

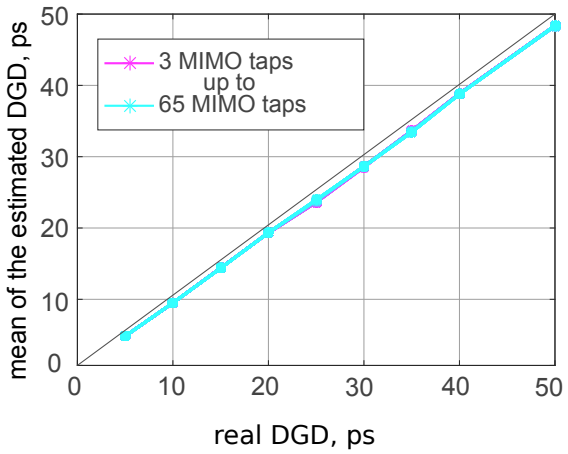
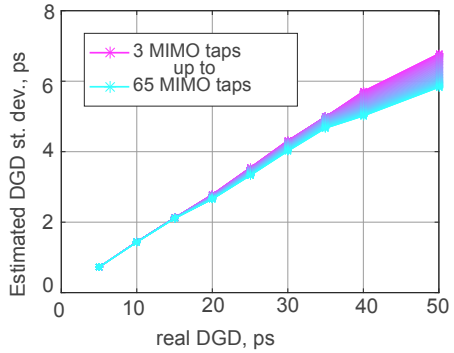
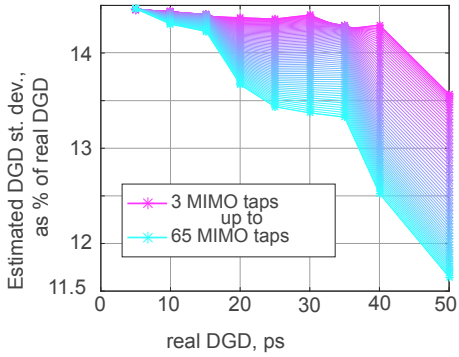


Figure 6.4: Mean of the DGD estimated digitally from MIMO taps. The number of taps is from 3 to 65.

In Figure 6.4 and 6.5 each of the 10 resulting points represents an average of 500 coarse step model realisations. In Figure 6.4 the DGD estimation bias was corrected (see [67]), which was slightly underestimated and which explains why the estimated DGD is always slightly smaller than the real DGD. The results show that when the bias is taken into account the mean of the estimated DGD can be predicted fairly well even for low number of taps. In Figure 6.5a-Figure 6.5b the standard deviation grows steadily with the actual DGD value, however, having more MIMO taps is not beneficial for low actual DGD values. Figure 6.5b shows the results in Figure 6.5a expressed as percentage of the real

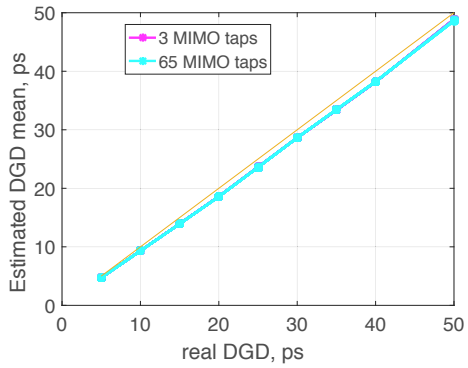


(a) Standard deviation of the DGD estimated digitally from MIMO taps.

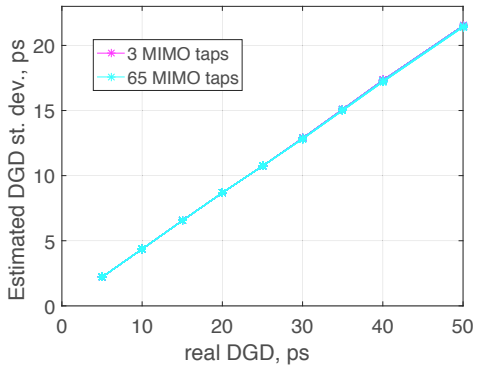


(b) Standard deviation of the DGD estimated digitally from MIMO taps, expressed in percent of the actual DGD value.

Figure 6.5: DGD estimation accuracy for a 28 *GBaud* signal 8 samples per symbol.



(a) Mean of the DGD estimated digitally from the MIMO taps. Number of taps is sweptranges from 3 to 65.



(b) Standard deviation of the DGD estimated digitally from MIMO taps

Figure 6.6: DGD estimation accuracy for a 28 *GBaud* signal 2 samples per symbol.

DGD, which shows that the average error up to 14.5% can be expected.

The calculation above is, however, performed for an impractically high oversampling rate. Recalculated for a signal with 2 samples per symbol as in Figure 6.6, MIMO taps analysis still gives a good estimation of the mean DGD. The standard deviation of approximately 44% of the actual DGD value, on the contrary is quite high. In practice, invention of other methods or upgrades of existing methods are required for more accurate estimation of DGD, as discussed in Section 5.4.1.

6.4 Conclusions

In this chapter we investigated how the proposed DBP algorithm could be implemented in practice. Deviation of DGD or PSP estimation error up to 15% decreases the linear Q factor by less than 0.2. Current DBP implementation limits by far overweight any NL-PMD interactions.

Further, new solutions for accurate DGD and PSP estimation should be developed. With the existing methods, the proposed DBP will have a certain loss in performance due to inaccuracy in the measurement of these parameters.

Chapter 7

PMD-induced fundamental limits of DBP

This chapter suggest a method on how to define the fundamental limits of a long-haul lumped amplified system's spectral efficiency (SE) and reach distance (RD) that employs DBP and is limited by NL-PMD interactions. The limits are derived and the method is demonstrated for an exemplary system for both the conventional DBP and the proposed DBPmod.

7.1 Shannon Capacity theory

The channel capacity defines the maximum rate of error-free information transmission through the channel. According to the Shannon-Hartley theorem [3], the information rate R in $bits/s$ is less or equal to the channel capacity C in $bits/s$, so that there is an upper bound on the maximum possible transmitted error-free information in a channel with a fixed bandwidth B and added white Gaussian noise:

$$R \leqslant C = B \log_2(1 + \frac{P_s}{P_n}). \tag{7.1}$$

where P_s and P_n are the signal input power and noise power, respectively. Equation (7.1) predicts the maximum rate at which information can be transmitted over a memoryless communication channel of a specified bandwidth in the presence of additive white Gaussian noise (AWGN) . Equation (7.1) can

be rewritten to set the upper bound on the channel spectral efficiency S_e in $bits/s/Hz$:

$$S_e = \frac{C}{B} = \log_2(1 + \frac{P_s}{P_n}). \tag{7.2}$$

In a nonlinear channel the sources of noise are ASE-noise with power P_{ASE}, receiver and transmitter noise, and nonlinear noise with power P_{NL}. Nonlinear noise is not white Gaussian, but as shown in the Gaussian noise (GN) model it is assumed to be so [93] (in fact, according to the enhanced GN (EGN) model, for the systems considered in this thesis, nonlinear noise possesses statistical properties of Gaussian noise after propagating approximately 2 spans [94]). Considering the channel only, independent of the receiver and transmitter we get:

$$S_e = \log_2(1 + \frac{P_s}{P_{ASE} + P_{NL}}). \tag{7.3}$$

While the ASE noise power can be estimated as in Section 3.1.1 Equation (3.1), we assume that the NL compensation techniques are compensating for the P_{NL} with a coefficient κ. If κ is 0, all NL are ideally compensated, if κ is 1 none are compensated. Choosing different values for κ we can see how the spectral efficiency evolves relative to P_{NL} and P_{ASE}:

$$S_e = \log_2(1 + \frac{P_s}{P_{ASE} + \kappa P_{NL}}). \tag{7.4}$$

7.1.1 Investigated system and metrics

We take as example a single channel 28 $GBaud$ system with 30 SSMF spans and 4 dB NF EDFAs with ideal power loss compensation (16 dB gain). The signal input power P_s varies from -10 to 10 dBm with 1 dB step. κ varies from 0 to 1 with 0.05 step. P_{ASE} is estimated from Equation (3.1), P_{NL} is derived from the simplified GN model Section VI.C of [93]. We then calculate the S_e as in Equation (7.4), changing the NL compensation coefficient. The optimum power value is always selected with lowest BER as a criterion. The results are plotted in Figure 7.1.

Alternatively, the system maximum reach distance can be estimated. The ROSNR at the receiver is chosen as a reach distance estimation criterion. It is

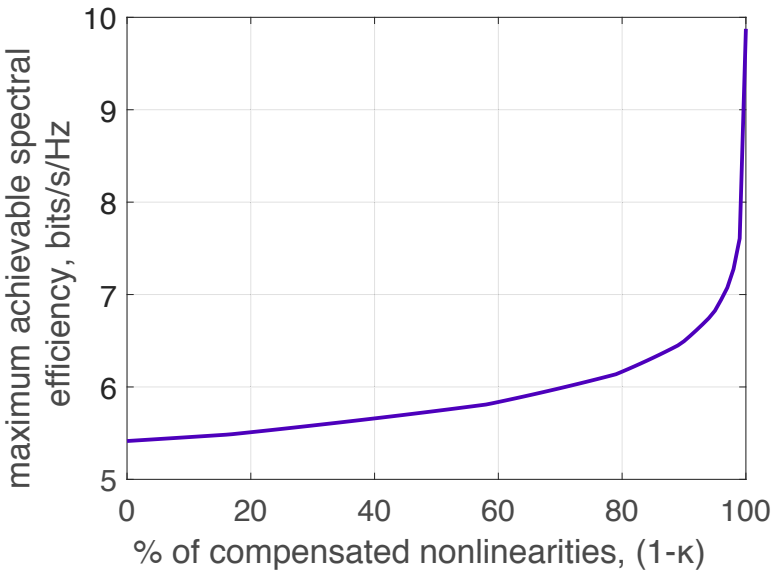


Figure 7.1: Maximum SE of a 30 spans, EDFA 4 dB NF amplified system, 28 $GBaud$ signal.

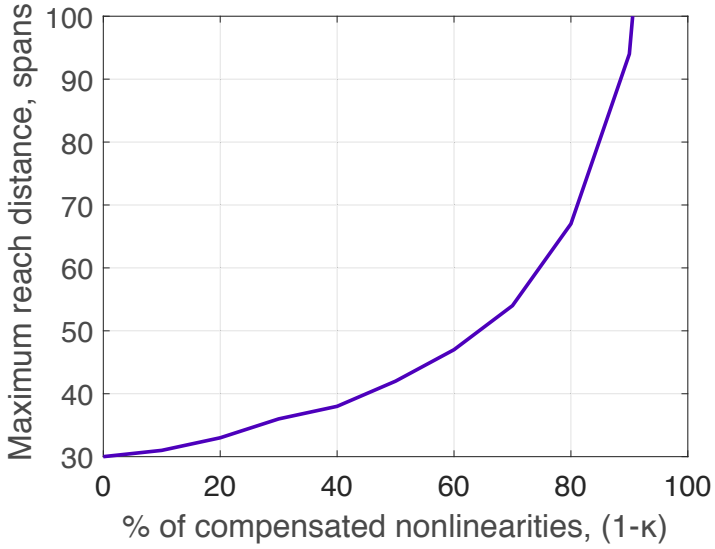


Figure 7.2: Maximum reach distance of a EDFA 4 dB NF amplified system, 28 $GBaud$ signal. Reach distance is defined by required optical signal-to-noise ratio (ROSNR) of 16.2 dB (modulation-format independent).

defined as the OSNR of a reference system (modulation-format independent) at the end of 30^{th} span with zero PMD, optimum input power of -2 dBm, and no NL compensation. This OSNR equals 16.2 dB for DP-16-QAM-modulated signal. Choosing the optimum input power in each case and sweeping the part of the nonlinear effects that would be compensated we get Figure 7.2.

Note that in both Figures the SE or the reach distance do not grow to infinity but have a finite endpoint, which is not plotted to keep the scale reasonable. Principally, if all the NL are compensated, and no other sources of noise are present, the power can continue increasing until the fibre melting point (approx. 33 dBm for SSMF). However, as was noted in Section 3.2.1 the nonlinearities amplified spontaneous emission (NL-ASE) noise interactions start limiting DBP performance at distances of order of 60 spans.

7.1.2 Estimation of how much NL can be compensated

In Section 4.3.2 we introduced a metric to estimate the impact of PMD on DBP. This was the cross-correlation of the optical field propagated forward and then backward with and without PMD. In this section we want to connect this metric with the proportion of NL that can be compensated. Further, we can use it to estimate the SE and maximum RD as in Figure 7.1 and 7.2. Clearly, the more the NL-PMD interactions affect the optical field, the worse standard DBP would perform, and the lower % of NL will be compensated (higher κ).

The procedure to estimate the proportion of compensated NL starts with propagating the signal forward and backward with and without PMD (without ASE). The signal EVM is then estimated. Further EVM is converted to SNR [95] and OSNR [26] (see Appendix F). Finally, knowing OSNR and signal power, the total noise is estimated, ASE noise is deducted which leaves us with pure NL noise. We estimate this for a system with PMD and without PMD by comparing the residual P_{NL} and therefore find out which part of NL is not compensated due to PMD.

We present the simulation results for the DP-16-QAM 8 dBm 28 $GBaud$ signal, propagating in a 0.1 $ps/\sqrt{km}$ fibre PMDc system, consisting of 30x80 km spans, each with one noiseless EDFA ideally compensating for span loss (full setup in Figure 5.3 in Section 5.3.1). The dependence between the cross-

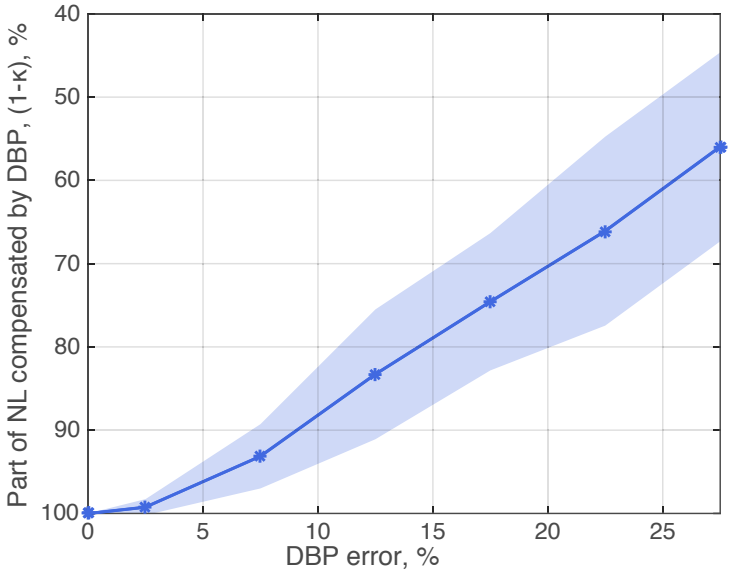
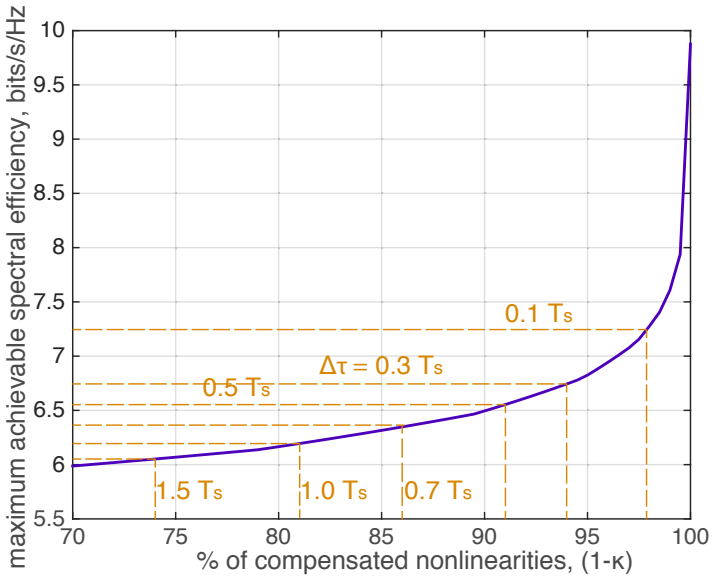
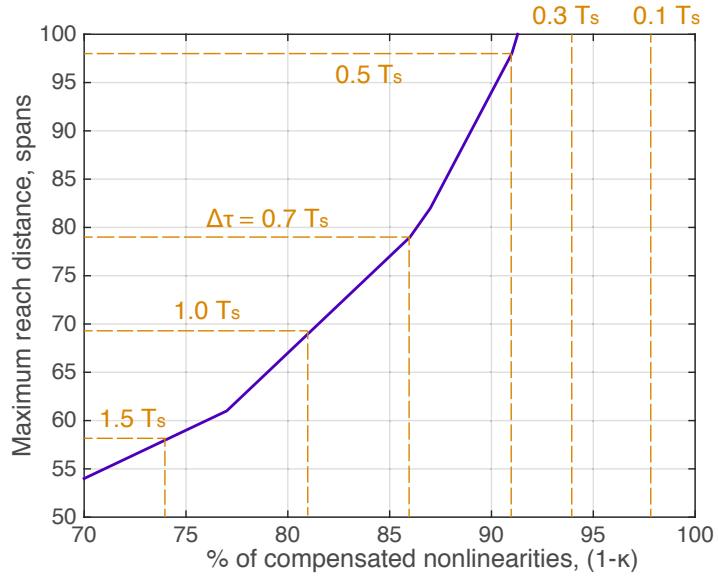


Figure 7.3: Correlation between DBP error (defined in Section 4.3.2) and proportion of compensated NL for a DP-16-QAM 8 dBm 28 $GBaud$ signal is a 30x80 km link with 0.1 $ps/\sqrt{km}$ fibre PMDc.

correlation metric and percent of compensated NL is evident.



(a) Spectral efficiency PMD-induced limitation on DBP.



(b) Reach distance PMD-induced limitation on DBP.

Figure 7.5: PMD-induced limitations of conventional DBP for a DP-16-QAM 28 *GBaud* signal with variable input power propagating in a 30x80 *km* SSMF link with variable fibre PMDc. Each span has a 4 *dB* NF EDFA.

7.1.3 Conventional DBP PMD-induced limits

The connection explored in the previous section allows to leap from the DGD value to the SE and RD. The transition is schematically pictured in Figure 7.4

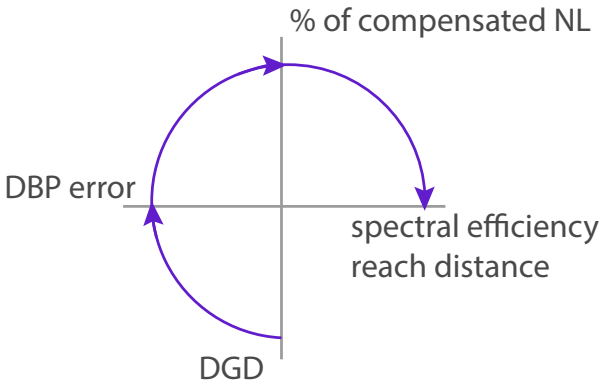


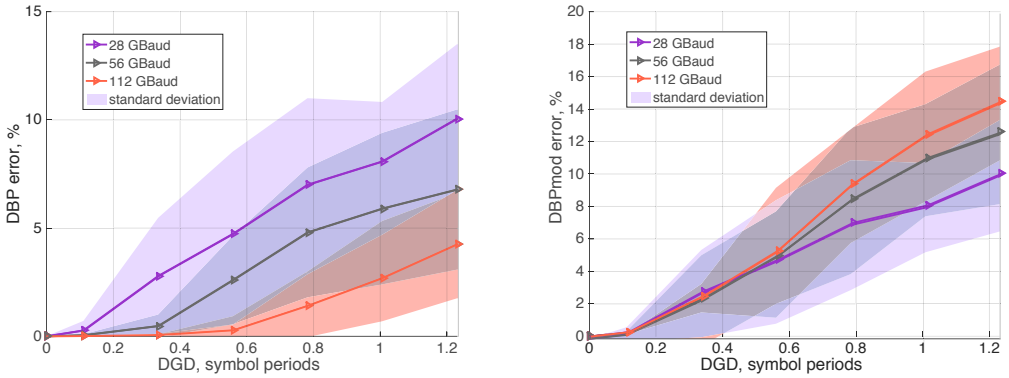
Figure 7.4: Scheme of transition to connect the DGD with SE and RD.

We start from the simulated dependence of the correlation metric (DBP error) on the DGD (ex. see Figure 4.15a in Section 4.3.4). Then by a simple value matching we move from DBP error to percentage of compensated NL (see, for example, Figure 7.3) and finally we get the desired Shannon metrics (see Figure 7.1, Figure 7.2). The results for conventional DBP are summarised in Figure 7.5.

Both figures show the average lower limit that PMD sets on the NL compensation. For instance, when the link has 0.1 symbol periods DGD, 98% of NL will be compensated by conventional DBP. For the considered single-channel system it allows to increase SE from 5.4 $bits/s/Hz$ (without DBP) to 7.3 $bits/s/Hz$ (with conventional DBP). When the DGD grows up to 0.5 symbol periods, 91% of NL are compensated and the SE grows up to 6.5 $bits/s/Hz$.

Reaching longer distances is a less challenging task, even by DGD of 1.5 symbol periods and 74% of compensated NL, the system RD doubles. Note that even though ASE-noise keeps growing, here we assume that we can keep increasing the optical power and compensate the required % of NL (NL-ASE interactions are not considered). Summing up, DBP is potentially a good candidate technique for increasing RD but will not be a game-changer for increasing SE. Adding to that, feasible amounts of PMD that are reached in modern fibres, do not seem to affect DBP ability to increase system length.

7.1.4 Proposed DBP PMD-induced limits



(a) Optical field error due to PMD for signals having same 8 dBm power.

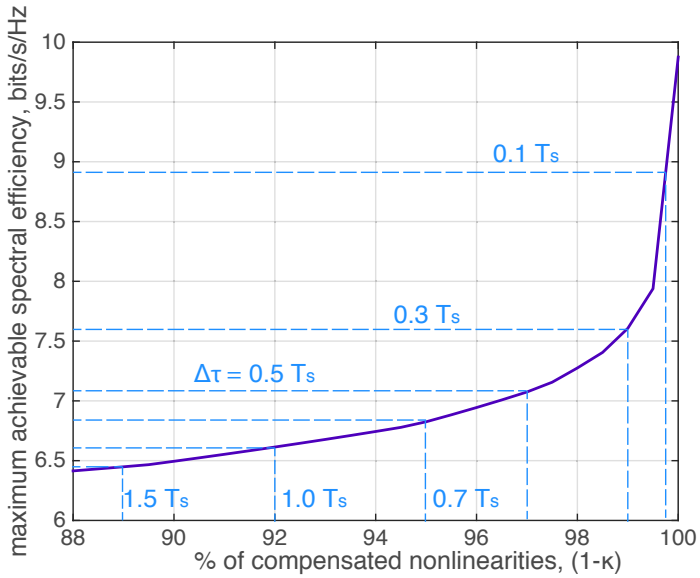
(b) Optical field error due to PMD for signals having same power spectral density of 0.223 mW/GHz.

Figure 7.6: Investigation of how the signal symbol rates influence the PMD-nonlinear interaction in the proposed DBP modification.

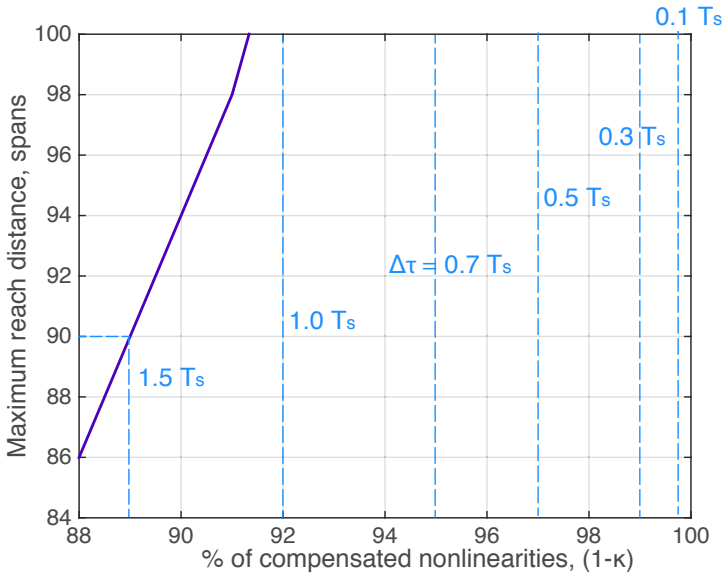
In the previous section we investigated how much standard DBP implementation would underperform in presence of PMD. Here we present the same investigation with the DBP proposed in Chapter 5.

For the procedure we still lack the cross-correlation metric estimation, which is described in Section 4.3.3 as SSF PMD Model I. All simulation parameters are summarized in Section 4.3. Filling the gap we perform the simulations for DBPmod, results are presented in Figure 7.6. Figures compare between different baud rates, 28, 56, and 112 $GBaud$, having the same power of 8 dBm or the same power spectral density of 0.223 mW/GHz. As compared to the conventional DBP (see Figure 4.15 in Section 4.3.4), the field error is around twice as small.

Finally, the SE and RD can be estimated, as in Figure 7.7. Proposed DBP has a considerably better capability to compensate for NL. At DGD of 0.5 symbol periods it removes 97% of NL compared with 91% by conventional DBP. At a DGD of one full symbol period the difference is 11%. Figure 7.7b shows that feasible amounts of PMD will not affect DBP RD increase potential. Even a DGD of 1.5 symbol periods will allow 89% of compensated NL and will theoretically triple the system length.



(a) Spectral efficiency PMD-induced limitation on modified DBP.



(b) Reach distance PMD-induced limitation on modified DBP.

Figure 7.7: PMD-induced limitations of proposed DBP for a DP-16-QAM 28 *GBaud* signal with variable input power propagating in a 30x80 *km* SSMF link with variable fibre PMDc. Each span has a 4 *dB* NF EDFA.

Note that the results presented in Figure 7.5 and in Figure 7.7 are dependent on the receiver DSP. Further, here we assumed perfect estimation of DGD and PSP for the DBPmod.

7.2 Conclusions

In this chapter we demonstrated that for a long-haul optical transmission system it is possible to define the fundamental limits of conventional and proposed DBP (Section 5.2). DBP is a very useful technique to increase the RD. Practically observed DGDs will not have any significant effect on it. Rather, systems will be limited by NL-ASE interactions at long distances. However, it is by far more challenging to use DBP to increase system SE because of the NL-PMD interactions.

Chapter 8

Conclusions

The main objective of this thesis was to provide a comprehensive understanding of how stochastic effects in a fibre interact with NL and how they affect DBP. The analysis was performed for non-DM EDFA amplified long-haul links.

Firstly, we identified the stochastic effects in a fibre and showed their statistical distributions. We pointed out that PDL and PMD are by far the major limiting effect. We showed that the DBP performance can be greatly degraded in presence of PMD and that it can be well characterised by the DGD value at the receiver. The DGDs that have improbable but high values are the ones responsible for the largest system degradation and outage. In such cases DGDev grows linearly with distance, in comparison to the most probable DGDevs that follow the distance square root rule.

Another important conclusion was that the simulation results of PMD-induced DBP degradation are dependent on the way PMD was simulated. Fibre PMDcs can be swept, then the most probable DGD cases are simulated. Or it can be fixed, then the pre-selection of fibre birefringence profiles that lead to improbable DGD values is required. While the first way is easier to perform and is dominating in current research, the second one reflects real systems in a better way.

In order to develop an analytical model to estimate the PMD-induced DBP error, we considered the perturbation model. We showed how under certain assumptions PMD can be accounted for. Furthermore, we suggested and verified a simplification in order to get an analytical solution for the perturbation term. This can be done when the DGDev is approximated by a staircase span-step

function. We compared the proposed supplement with the SSF simulations and found a good consistency level. Using the supplemented perturbation model and SSF simulations we showed that the DBP error is independent on the modulation format, but is affected by both signal bandwidth and signal power density.

Using the facts that the best approximation for DGDevs with high final DGD is a straight line, which can even further be digitised to a staircase span-step function, we suggested a modification of conventional DBP. It is based on the knowledge of DGD and PSP for the central signal frequency at the receiver. The DGDev is then approximated during DBP to minimise the mismatch with the forward propagation. We discussed the low-complexity of the method and rather high tolerance to measurement errors of the data feed. We also pointed out that new methods for accurate DGD and PSP estimation are required. The proposed algorithm nearly fully mitigates the PMD-induced performance degradation and "returns" it to the no-PMD-levels.

Throughout the thesis it was shown that the linear noise remains the largest system limitation factor. It by far overweights NL-PMD interaction effects. So that the NL-PMD management is necessary to increase the already acceptably high OSNR. If the system has low OSNR, the NL-PMD effect will not be visible.

Finally, we identified the NL-PMD induced fundamental limits of DBP. For a long-haul optical transmission system it is possible to tell which percentage of NL on average can be compensated, provided DGD reached a certain value. We showed, that in modern fibre systems PMD will not affect the ability of DBP to increase the reach distance, but will greatly limit the gain of spectral efficiency.

Appendix A

Signal rotation to a new coordinate system

The signal coordinates rotation to a new coordinate system, for example, to the PSP can be performed in frequency domain. The known Stokes parameters are used to construct the Mueller and then the Jones rotation matrices. The rotation should be done in Jones calculus, rather than in Mueller notation, as it works with the electric field, rather than with the light intensity.

We assume that the Stokes parameters of the fast polarisation axis are known:

$$\mathbb{S}_{fast} = \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix}, \tag{A.1}$$

In the Mueller rotation matrix the first column, m_1, is constructed from the fast PSP Stokes parameters $\mathbb{S}_{fast}$. The second column m_2 specifies the PSD vector. In case of first order PMD, the PSD vector is not specified and the second column Mueller matrix is constructed from $\mathbb{S}_{fast}$:

- if $S_1 = S_3 = 0$ then $m_2 = (0, 1, 0)$,
- otherwise, $m_2 = (0, -S_3, S_2)$.

This ensures the second column vector is orthogonal to $\mathbb{S}_{fast}$. The third or-

thogonal vector $m_3 = m_1 \times m_2$ and is specified as

$$m_3 = \begin{pmatrix} m_{1,2}m_{2,3} - m_{1,3}m_{2,2} \\ m_{1,3}m_{2,1} - m_{1,1}m_{2,3} \\ m_{1,1}m_{2,2} - m_{1,2}m_{2,1} \end{pmatrix}. \tag{A.2}$$

Finally, Mueller rotation matrix can be constructed:

$$\mathbb{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & S_1 & m_{2,1} & m_{3,1} \\ 0 & S_2 & m_{2,2} & m_{3,2} \\ 0 & S_3 & m_{2,3} & m_{3,3} \end{bmatrix}. \tag{A.3}$$

Now the Jones rotation Matrix can be constructed as in [96]:

$$H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}, \tag{A.4}$$

$$H_{kl} = |H_{kl}| \cdot e^{i\psi_{kl}}, \tag{A.5}$$

$$H_{11} = \sqrt{(\mathbb{M}_{00} + \mathbb{M}_{01} + \mathbb{M}_{10} + \mathbb{M}_{11})/2}, \tag{A.6}$$

$$H_{12} = \sqrt{(\mathbb{M}_{00} - \mathbb{M}_{01} + \mathbb{M}_{10} - \mathbb{M}_{11})/2}, \tag{A.7}$$

$$H_{21} = \sqrt{(\mathbb{M}_{00} + \mathbb{M}_{01} - \mathbb{M}_{10} - \mathbb{M}_{11})/2}, \tag{A.8}$$

$$H_{22} = \sqrt{(\mathbb{M}_{00} - \mathbb{M}_{01} - \mathbb{M}_{10} + \mathbb{M}_{11})/2}. \tag{A.9}$$

$$\cos(\psi_{11} - \psi_{12}) = \frac{\mathbb{M}_{02} + \mathbb{M}_{12}}{\sqrt{(\mathbb{M}_{00} + \mathbb{M}_{10})^2 - (\mathbb{M}_{01} + \mathbb{M}_{11})^2}}, \tag{A.10}$$

$$\sin(\psi_{11} - \psi_{12}) = \frac{\mathbb{M}_{03} + \mathbb{M}_{13}}{\sqrt{(\mathbb{M}_{00} + \mathbb{M}_{10})^2 - (\mathbb{M}_{01} + \mathbb{M}_{11})^2}}, \tag{A.11}$$

$$\cos(\psi_{21} - \psi_{11}) = \frac{\mathbb{M}_{20} + \mathbb{M}_{21}}{\sqrt{(M_{00} + M_{01})^2 - (\mathbb{M}_{10} + \mathbb{M}_{11})^2}}, \tag{A.12}$$

$$\sin(\psi_{21} - \psi_{11}) = \frac{\mathbb{M}_{30} + \mathbb{M}_{31}}{\sqrt{(\mathbb{M}_{00} + \mathbb{M}_{01})^2 - (\mathbb{M}_{10} + \mathbb{M}_{11})^2}}, \tag{A.13}$$

$$\cos(\psi_{22} - \psi_{11}) = \frac{\mathbb{M}_{22} + \mathbb{M}_{33}}{\sqrt{(\mathbb{M}_{00} + \mathbb{M}_{11})^2 - (\mathbb{M}_{10} + \mathbb{M}_{01})^2}}, \tag{A.14}$$

$$\sin(\psi_{22} - \psi_{11}) = \frac{\mathbb{M}_{32} + \mathbb{M}_{23}}{\sqrt{(\mathbb{M}_{00} + \mathbb{M}_{11})^2 - (\mathbb{M}_{10} + \mathbb{M}_{01})^2}}. \tag{A.15}$$

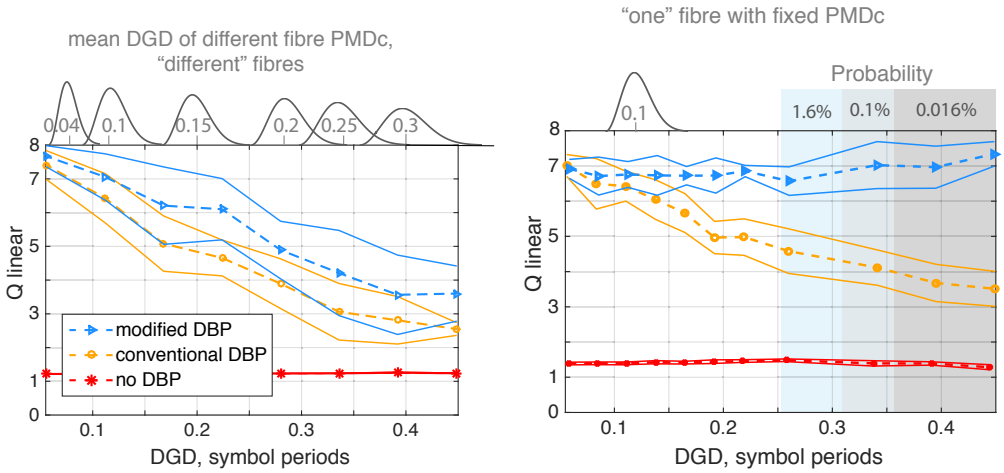
It is principally impossible to extract all phases from the Mueller matrix. The differences between the phases $\psi_{11}, .., \psi_{22}$ are important. Practically $\psi_{11} = 0$ can be assumed. Finally, when the H matrix is constructed, the signal multiplied with it will be rotated to the new coordinate system.

Appendix B

Impact of PMD statistics on the simulation results

In order to demonstrate how the choice of PMD statistics influences the simulation results of the PMD impact on DBP, we simulated the system described in Section 5.3.2 using two different approaches. The first one and most commonly used one is to vary the fibre PMDc and to simulate the most probable DGD cases for each fibre. The second approach keeps the fibre PMDc fixed and uses pre-selection to manually choose low-probable DGD values from one distribution. In the simulations the fibre PMDc runs through the values 0.04, 0.1, 0.15, 0.2, 0.25, 0.3 $ps/\sqrt{km}$ in the first case and is fixed to 0.1 $ps/\sqrt{km}$ in the second. Here we present the results for a 28 $GBaud$ 16-QAM modulated signal.

The results in Figure B.1 show that depending on the PMD statistics, both the DBP and the modified DBP (see Chapter 5) will be affected differently. The orange line marked with circles illustrates the conventional DBP method. The curves have different starting points. While in the first case the low DGD scenarios follow the square root evolution, in the second case the DGD has a bow evolution, similar to Figure 3.3 in Section 3.1.2. Although both of the evolutions finish at the same value, the one that accumulated higher DGD during propagation will affect the DBP more (covering a larger area under the line, as in Figure 3.4b). Note also that the curves in Figure B.1 have different slopes. At high DGD values the first case evolves like a square root and the second case as a straight line, similar to Figure 3.4a). The average straight



(a) PMDc is changing.

(b) PMDc is fixed.

Figure B.1: Comparison between two different approaches of PMD simulation.

line evolution now has a lower cumulative DGD and therefore shows better performance. The variance in performance is slightly larger in the first case as the higher PMDc implying a larger variety of DGD evolutions.

The DBPmod method performance, illustrated with blue triangle marked curve, differs significantly depending on the way PMD is simulated. While in the second case in Figure B.1b performance is stable and is equivalent to low DGD levels, in the first case in Figure B.1a its efficiency drops, however, still performing better than conventional DBP . The performance of the DBPmod cannot be better than the low DGD case for a particular PMDc. The growing PMDc in Figure B.1a therefore pulls the modified DBP gain down. We also remark here, that the DBPmod performance in Figure B.1a can be improved if a square-root DGD distribution is used for approximation (instead of linear distribution).

Appendix C

Basic principles of MIMO equalizer

In optical fibre coherent transmission the name MIMO refers to the equaliser that reconstructs the sent logical x_r and y_r signal polarisations from the received components x_{Rx} and y_{Rx} (which constitute a mixture of x_r and y_r). Functionally, a MIMO reconstructs the inverse Jones matrix H of a channel. The general principle of the equalizer is depicted in Figure C.1.

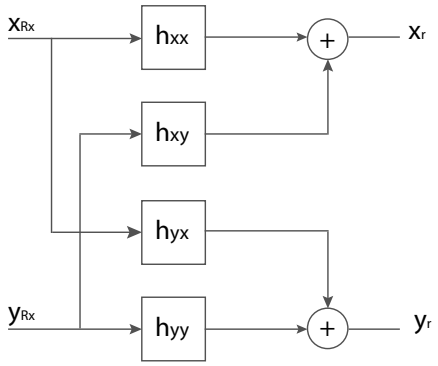


Figure C.1: Working principle of a MIMO equalizer.

The reconstructed signal is described as [79, 97]:

$$\begin{aligned} x_r(m) &= h_{xx}(m) \cdot x_{Rx,\,block}(m) + h_{xy}(m) \cdot y_{Rx,\,block}(m), \\ y_r(m) &= h_{yx}(m) \cdot x_{Rx,\,block}(m) + h_{yy}(m) \cdot y_{Rx,\,block}(m), \end{aligned} \tag{C.1}$$

where $x_{Rx,\,block}$ and $y_{Rx,\,block}$ are the sliding blocks of N_{taps} samples, to which

the filter is applied:

$$\begin{aligned} x_{Rx,\,block}(m) &= [x_{Rx}(m), x_{Rx}(m-1), ...x_{Rx}(m-N_{taps})], \\ y_{Rx,\,block}(m) &= [y_{Rx}(m), y_{Rx}(m-1), ...y_{Rx}(m-N_{taps})], \end{aligned} \tag{C.2}$$

h_p (p takes values of xx, xy, yx, yy) are the components of the inverse Jones matrix H, namely vectors of length N_{taps} containing filter taps weights:

$$h_p(m) = [h_{p,0}(m), h_{p,1}(m), ..., h_{p,N_{taps}-1}(m)]^T, \tag{C.3}$$

m is the signal sample.

The MIMO coefficients h_p are adapted to minimize the so-called cost function, determining the distance between the reconstructed point and the constellation amplitude circle:

$$\begin{aligned} \epsilon_x(m) &= \Lambda^2 - |x_r(m)|^2, \\ \epsilon_y(m) &= \Lambda^2 - |y_r(m)|^2, \end{aligned} \tag{C.4}$$

For example, for a QPSK signal $\Lambda = 1$ (CMA), for a 16-QAM signal $\Lambda = [1/\sqrt{5}, 1, 3/\sqrt{5}]$ (MMA). The equalizer weights are updated as following:

$$\begin{aligned} h_{xx} &= h_{xx} + \mu\epsilon_x x_{Rx} x_r^*, \\ h_{xy} &= h_{xy} + \mu\epsilon_x y_{Rx} x_r^*, \\ h_{yx} &= h_{yx} + \mu\epsilon_y x_{Rx} y_r^*, \\ h_{yy} &= h_{yy} + \mu\epsilon_y y_{Rx} y_r^*, \end{aligned} \tag{C.5}$$

where μ is a convergence factor.

Appendix D

Q factor metric

The Quality factor or Q factor is a metric that is useful for estimating the performance of an optical link. In this thesis we chose it to be able to plot the mean along with the standard deviation of the performance results in one figure.

As defined in [87], the Q factor and BER relate as

$$Q = 1/2 \cdot erfc(BER/\sqrt{2}), \tag{D.1}$$

where $erfc\{\}$ is a complementary error function.

The relation between the two metrics is plotted in Figure D.1.

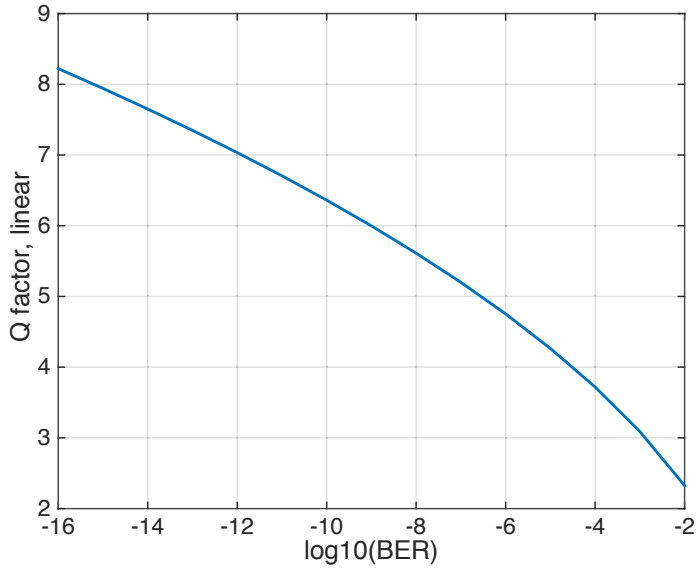


Figure D.1: Relationship between Q factor and BER.

Appendix E

Big-O notation

Big-O notation is a metric used to describe the complexity of algorithms and functions. It describes the behaviour of a function when its argument tends towards a particular value or infinity. Namely, it answers the question of how fast the number of operations grows depending on the number of elements, n, (also called problem size). This growth can be constant, $O(1)$, linear $O(n)$, logarithmic $O(log(n))$, exponential $O(\chi^n)$, polynomial $O(n^\chi)$, where χ is a constant. The term of a function with the highest growth rate defines its Big-O. Some examples of the Big-O notations are summarized in Figure E.1

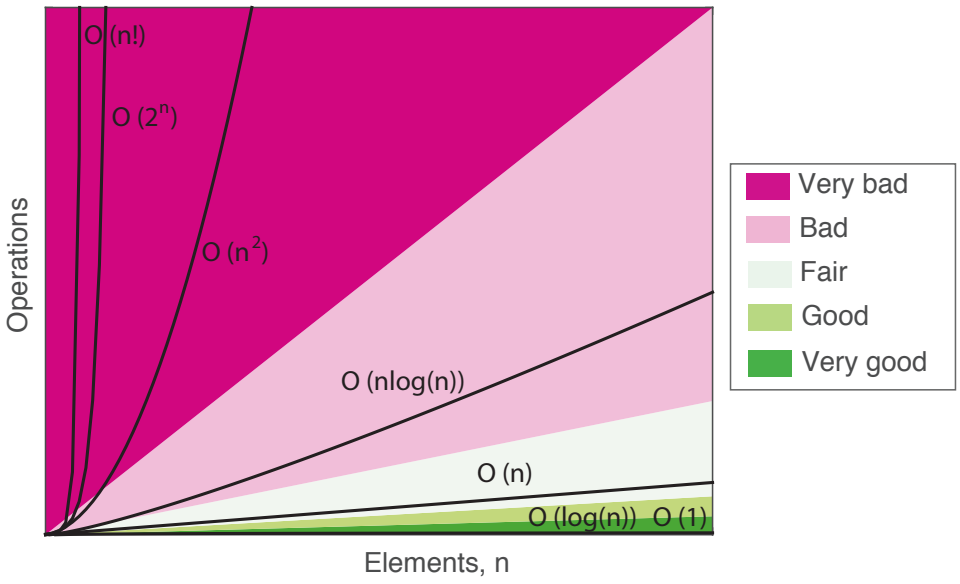


Figure E.1: Big-O Complexity Types Comparison. Adopted from [98].

Appendix F

Relation between EVM, SNR, and OSNR

The metric EVM is a measure of a received signal's constellation diagram quality. It compares the ideal and the received demodulated signals and can be calculated as a relative constellation root-mean-square error:

$$EVM = \sqrt{\frac{\sum_{i=1}^{N_c} |b_i - a_i|^2}{\sum_{i=1}^{N_c} |a_i|^2}}, \tag{F.1}$$

where N_c is the number of constellation points (transmitted symbols) and i is the symbol count, a_i is the vector of reference constellation points, b_i is the vector of received constellation points.

According to derivations in [95], there is an approximate analytical connection between EVM and SNR, provided the noise has Gaussian distribution, and number of received constellation points is much larger than the number of unique symbols in the constellation:

$$SNR \approx \frac{1}{EVM^2}. \tag{F.2}$$

The SNR itself, independent of the receiver parameters and recalculated for a given bandwidth (commonly 0.1 nm), can be further matched to the OSNR, as in [26]:

$$OSNR = SNR \cdot \frac{B}{12.5 \cdot 10^9} \cdot \frac{N_{pol}}{2}, \tag{F.3}$$

where B is the signal bandwidth and N_{pol} is the number of signal polarisations.

Bibliography

[1] "Cisco Visual Networking Index: Forecast and Methodology, 2016-2021 - Cisco," visited on 26 March 2018. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html 2

[2] A. R. Chraplyvy, "Roadmap of optical communications: History," *Journal of Optics*, vol. 18, no. 6, pp. 5–6, 2016. 2

[3] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948. 2, 123

[4] "Fathom Business Solutions | Responsive Design," visited on 26 March 2018. [Online]. Available: http://www.fathomgroup.co.uk/articles/responsive_design.html 3

[5] P. Sillard, "New fibers for ultra-high capacity transport," *Optical Fiber Technology*, vol. 17, no. 5, pp. 495–502, 2011. 3

[6] P. J. Winzer, "Scaling Optical Fiber Networks: Challenges and Solutions," *Optics & Photonics News*, vol. 26, no. 3, pp. 28–35, 2015. 4

[7] N. Alic, M. Karlsson, A. Mecozzi, and C. Sethumadhavan, "Nonlinearity Mitigation for Coherent Transmission Systems [Feature issue]," *Optics Express*, vol. 25, no. 4, 2017. 4

[8] F. R. Kschischang, "Roadmap of optical communications: Forward error correction," *Journal of Optics*, vol. 18, no. 6, pp. 23–24, 2016. 4

[9] "Submarine cable map," retrieved on 2 February 2018. Used by permission from TeleGeography. [Online]. Available: www.telegeography.com 5

[10] D. Rafique, "Fiber nonlinearity compensation: Commercial applications and complexity analysis," *Journal of Lightwave Technology*, vol. 34, no. 2, pp. 544–553, 2016. 5

[11] J. Bromage, "Raman amplification for fiber communication systems," in *OFC 2003 Optical Fiber Communications Conference, 2003.*, 2003, p. 156 vol.1. 6

[12] S. Grubb, "Facebook: The Drive Towards Flexible Optical Networking," in *OFC/NFOEC Optical Fiber Communication Conference*, 2017, visited on 5 May 2018. [Online]. Available: http://on2020.org/files/SG-Facebook-ON202016-9.pdf 6

[13] R. Maher, T. Xu, L. Galdino, M. Sato, A. Alvarado, K. Shi, S. J. Savory, B. C. Thomsen, R. I. Killey, and P. Bayvel, "Spectrally Shaped DP-16QAM Super-Channel Transmission with Multi-Channel Digital Back-Propagation," *Scientific Reports*, vol. 5, 2015. 6

[14] S. B. Amado, F. P. Guiomar, N. J. Muga, R. M. Ferreira, J. D. Reis, S. M. Rossi, A. Chiuchiarelli, J. R. Oliveira, A. L. Teixeira, and A. N. Pinto, "Low Complexity Advanced DBP Algorithms for Ultra-Long-Haul 400 G Transmission Systems," *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1793–1799, 2016. 6

[15] S. Yoshima, Y. Sun, K. R. H. Bottrill, F. Parmigiani, P. Petropoulos, and D. J. Richardson, "Nonlinearity mitigation through optical phase conjugation in a deployed fibre link with full bandwidth utilization," in *European Conference on Optical Communication (ECOC)*, no. We2.6.3, 2015. 6

[16] X. Liu, A. R. Chraplyvy, P. J. Winzer, R. W. Tkach, and S. Chandrasekhar, "Phase-conjugated twin waves for communication beyond the Kerr nonlinearity limit," *Nature Photonics*, vol. 7, no. 7, pp. 560–568, 2013. 6

[17] Z. Tao, L. Dou, W. Yan, L. Li, T. Hoshida, and J. C. Rasmussen, "Multiplier-free intrachannel nonlinearity compensating algorithm operating at symbol rate," *Journal of Lightwave Technology*, vol. 29, no. 17, pp. 2570–2576, 2011. 6

[18] F. P. Guiomar and A. N. Pinto, "Simplified volterra series nonlinear equalizer for polarization-multiplexed coherent optical systems," *Journal of Lightwave Technology*, vol. 31, no. 23, pp. 3879–3891, 2013. 6

[19] D. S. Millar, T. Koike-Akino, S. Ö. Arik, K. Kojima, K. Parsons, T. Yoshida, and T. Sugihara, "High-dimensional modulation for coherent optical communications systems," *Optics Express*, vol. 22, no. 7, pp. 8798–8812, 2014. 6

[20] A. D. Shiner, M. Reimer, A. Borowiec, S. O. Gharan, J. Gaudette, P. Mehta, D. Charlton, K. Roberts, and M. O'Sullivan, "Demonstration of an 8-dimensional modulation format with reduced inter-channel nonlinearities in a polarization multiplexed coherent system," *Optics Express*, vol. 22, no. 17, p. 20366, 2014. 6

[21] A. Ghazisaeidi, I. Fernandez De Jauregui Ruiz, R. Rios-Muller, L. Schmalen, P. Tran, P. Brindel, A. Carbo Meseguer, Q. Hu, F. Buchali, G. Charlet, and J. Renaudier, "Advanced C+L-Band Transoceanic Transmission Systems Based on Probabilistically Shaped PDM-64QAM," *Journal of Lightwave Technology*, vol. 35, no. 7, pp. 1291–1299, 2017. 6

[22] S. Hari, F. Kschischang, and M. Yousefi, "Multi-eigenvalue communication via the nonlinear Fourier transform," in *27th Biennial Symposium on Communications (QBSC)*, 2014, pp. 92–95. 6

[23] Y. Yu, W. Wang, P. D. Townsend, and J. Zhao, "Modified phase-conjugate twin wave schemes for spectral efficiency enhancement," in *European Conference on Optical Communication (ECOC)*, no. We.2.6.5, 2015. 7

[24] C. B. Czegledi, G. Liga, D. Lavery, M. Karlsson, E. Agrell, S. J. Savory, and P. Bayvel, "Polarization-Mode Dispersion Aware Digital Backpropagation," in *European Conference on Optical Communication (ECOC)*, no. Th.2.P2.SC3.25, 2016. 8, 12, 104, 106, 107

[25] A. D. Ellis, M. E. McCarthy, M. A. Z. Al-Khateeb, and S. Sygletos, "Capacity limits of systems employing multiple optical phase conjugators," *Optics Express*, vol. 23, no. 16, p. 20381, 2015. 8

[26] R. J. Essiambre, G. Kramer, P. Winzer, G. Foschini, and B. Goebel, "Capacity Limits of Optical Fiber Networks," *Journal of Lightwave Technology*, vol. 28, no. 4, 2010. 9, 10, 126, 147

[27] K. Goroshko, H. Louchet, and A. Richter, "Overcoming performance limitations of digital back propagation due to polarization mode dispersion," in *18th International Conference on Transparent Optical Networks (ICTON)*, no. Mo.B1.4, 2016. 11, 104

[28] G. Gao, X. Chen, and W. Shieh, "Influence of PMD on fiber nonlinearity compensation using digital back propagation," *Optics Express*, vol. 20, no. 13, p. 14406, 2012. 12

[29] F. Yaman and G. Li, "Nonlinear impairment compensation for polarization-division multiplexed WDM transmission using digital backward propagation," *IEEE Photonics Journal*, vol. 2, no. 5, pp. 816–832, 2010. 12, 104

[30] K. Goroshko, H. Louchet, and A. Richter, "Fundamental Limitations of Digital Back Propagation due to Polarization Mode Dispersion," in *Asia Communications and Photonics Conference (ACP)*, no. ASu3F.5, 2015. 12

[31] G. Liga, T. Xu, A. Alvarado, R. I. Killey, and P. Bayvel, "On the performance of multichannel digital backpropagation in high-capacity long-haul optical transmission," *Optics Express*, vol. 22, no. 24, 2014. 12

[32] T. Tanimura, S. Oda, T. Hoshida, L. Li, Z. Tao, and J. C. Rasmussen, "Experimental Characterization of Nonlinearity Mitigation by Digital Back Propagation and Nonlinear Polarization Crosstalk Canceller under High PMD condition," in *OFC/NFOEC Optical Fiber Communication Conference*, no. JWA020, 2011. 12

[33] I. Fernandez de Jauregui Ruiz, A. Ghazisaeidi, G. Charlet, and P. Tran, "Impact of Polarization Mode Dispersion on Digital Nonlinear Compensation Algorithms in Dispersion Unmanaged Systems," *OFC/NFOEC Optical Fiber Communication Conference*, no. Th3D.3, 2016. 12

[34] C. B. Czegledi, G. Liga, D. Lavery, E. Magnus Karlsson, e. J. Savory, and P. Bayvel, "Modified digital backpropagation accounting for polarization-mode dispersion," in *OFC/NFOEC Optical Fiber Communication Conference*, no. W1G.6, 2017. 12, 106, 107

[35] C. B. Czegledi, G. Liga, D. Lavery, M. Karlsson, E. Agrell, S. J. Savory, and P. Bayvel, "Digital backpropagation accounting for polarization-mode dispersion," *Optics Express*, vol. 25, no. 3, p. 1903, 2017. 12, 13, 106, 107, 109, 110, 111, 112, 113, 171, 172

[36] G. Agrawal, *Nonlinear fiber optics, 5th ed.* Academic Press, 2013. 16, 17, 18

[37] P. K. A. Wai and C. R. Menyuk, "Polarization decorrelation in optical fibers with randomly varying birefringence," *Optics Letters*, vol. 19, no. 19, p. 1517, 1994. 17, 63

[38] D. Marcuse, C. R. Menyuk, and P. K. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *Journal of Lightwave Technology*, vol. 15, no. 9, pp. 1735–1745, 1997. 17, 63

[39] S. V. Manakov, "On the theory of two-dimensional stationary self-focusing of electromagnetic waves," *Soviet Physics JETP*, vol. 38, no. 2, pp. 248–253, 1974. 17, 63

[40] C. Paré, A. Villeneuve, P.-A. A. Bélanger, and N. J. Doran, "Compensating for dispersion and the nonlinear Kerr effect without phase conjugation," *Optics Letters*, vol. 21, no. 7, pp. 459–461, 1996. 19

[41] M. Tsang, D. Psaltis, and F. G. Omenetto, "Reverse propagation of femtosecond pulses in optical fibers." *Optics letters*, vol. 28, no. 20, pp. 1873–1875, 2003. 19

[42] R. J. Essiambre, P. J. Winzer, X. Q. Wang, W. Lee, C. A. White, and E. C. Burrows, "Electronic predistortion and fiber nonlinearity," *IEEE Photonics Technology Letters*, vol. 18, no. 17, pp. 1804–1806, 2006. 19

[43] K. Roberts, C. Li, L. Strawczynski, M. O'Sullivan, and I. Hardcastle, "Electronic precompensation of optical nonlinearity," *IEEE Photonics Technology Letters*, vol. 18, no. 2, pp. 403–405, 2006. 19

[44] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *Journal of Lightwave Technology*, vol. 26, no. 20, pp. 3416–3425, 2008. 19

[45] X. Li, X. Chen, G. Goldfarb, E. Mateo, I. Kim, F. Yaman, and G. Li, "Electronic post-compensation of WDM transmission impairments using coherent detection and digital signal processing," *Optics Express*, vol. 16, no. 2, p. 880, 2008. 19

[46] C.-Y. Lin, R. Asif, M. Holtmannspoetter, and B. Schmauss, "Step-size selection for split-step based nonlinear compensation with coherent detection in 112-Gb/s 16-QAM transmission," *Chinese Optics Letters*, vol. 10, no. 2, 2012. 20

[47] R. Asif, C.-Y. Lin, and B. Schmauss, "Efficient compensation of chromatic dispersion and nonlinearities using logarithmic digital backward propagation in N-channel DWDM 1.12Tbit/s DP-QPSK transmission," *Journal of Modern Optics*, vol. 59, no. 2, pp. 95–101, 2012. 20

[48] E. F. Mateo, X. Zhou, and G. Li, "Improved digital backward propagation for the compensation of inter-channel nonlinear effects in polarization-multiplexed WDM systems," *Optics Express*, vol. 19, no. 2, p. 570, 2011. 20

[49] O. V. Sinkin, R. Holzlöhner, J. Zweck, and C. R. Menyuk, "Optimization of the split-step Fourier method in modeling optical-fiber communications systems," *Journal of Lightwave Technology*, vol. 21, no. 1, pp. 61–68, 2003. 20, 21, 60, 94

[50] A. Napoli, Z. Maalej, V. A. J. M. Sleiffer, M. Kuschnerov, D. Rafique, E. Timmers, B. Spinnler, T. Rahman, L. D. Coelho, and N. Hanik, "Reduced complexity digital back-propagation methods for optical communication systems," *Journal of Lightwave Technology*, vol. 32, no. 7, pp. 1351–1362, 2014. 20

[51] L. B. Du and A. J. Lowery, "Improved single channel backpropagation for intra-channel fiber nonlinearity compensation in long-haul optical communication systems." *Optics Express*, vol. 18, no. 16, pp. 17 075–17 088, 2010. 20

[52] D. Rafique, M. Mussolin, M. Forzati, J. Mårtensson, M. N. Chugtai, and A. D. Ellis, "Compensation of intra-channel nonlinear fibre impairments using simplified digital back-propagation algorithm," *Optics Express*, vol. 19, no. 10, p. 9453, 2011. 21

[53] L. Li, Z. Tao, L. Dou, W. Yan, S. Oda, T. Tanimura, T. Hoshida, and J. C. Rasmussen, "Implementation Efficient Nonlinear Equalizer Based on Correlated Digital Backpropagation," in *OFC/NFOEC Optical Fiber Communication Conference*, no. OWW3, 2011. 21

[54] M. Secondini, S. Rommel, F. Fresi, E. Forestieri, G. Meloni, and L. Poti, "Coherent 100G nonlinear compensation with single-step digital backpropagation," in *International Conference on Optical Network Design and Modeling, (ONDM)*, no. Tu4.3O, 2015. 21

[55] S. Kumar and D. Yang, "Optical backpropagation for fiber-optic communications using highly nonlinear fibers," *Optics Letters*, vol. 36, no. 7, p. 1038, 2011. 21

[56] C. Giles and E. Desurvire, "Modeling erbium-doped fiber amplifiers," *Journal of Lightwave Technology*, vol. 9, no. 2, pp. 271–283, 1991. 22

[57] E. Desurvire, *Erbium Doped Fiber Amplifiers: Principles and Applications.* John Wiley & Sons, 1994. 23

[58] J. Strand, "Polarization mode dispersion," in *Optical Fiber Telecommunications.* Academic Press, 2002, vol. IV-B, pp. 752–861. 24, 26

[59] C. D. Poole and J. Nagel, "Polarization Effects in Lightwave Systems," in *Optical Fiber Telecommunications.* Academic Press, 2012, vol. III-A, pp. 114–161. 26

[60] L. E. Nelson, C. Antonelli, A. Mecozzi, M. Birk, P. Magill, A. Schex, and L. Rapp, "Statistics of polarization dependent loss in an installed long-haul WDM system," *Optics Express*, vol. 19, no. 7, pp. 6790–6796, 2011. 29, 49

[61] J. Jiang, D. Richards, S. Oliva, P. Green, and R. Hui, "In-situ monitoring of PMD and PDL in a traffic-carrying transatlantic fiber-optic system," *OFC/NFOEC Optical Fiber Communication Conference*, vol. 2, no. NMB1, 2009. 29, 49

[62] A. Mecozzi and M. Shtaif, "The statistics of polarization-dependent loss in optical communication systems," *IEEE Photonics Technology Letters*, vol. 14, no. 3, pp. 313–315, 2002. 29, 49

[63] Z. Wang, C. Xie, and X. Ren, "PMD and PDL impairments in polarization division multiplexing signals with direct detection," *Optics Express*, vol. 17, no. 10, pp. 7993–8004, 2009. 30

[64] J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," *Proceedings of the National Academy of Sciences*, vol. 97, no. 9, pp. 4541–4550, 2000. 30, 32, 101

[65] C. Poole and R. Wagner, "Phenomenological approach to polarisation dispersion in long single-mode fibres," *Electronics Letters*, vol. 22, no. 19, p. 1029, 1986. 32, 34

[66] F. Md. Saifuddin, "Optical Performance Monitoring in Digital Coherent Receivers," Ph.D. dissertation, The University of Tokyo, 2011, visited on 12 May 2018. [Online]. Available: http://repository-old.dl.itc.u-tokyo.ac.jp/dspace/bitstream/2261/52233/1/37087287.pdf 32, 38, 101, 102, 106

[67] F. N. Hauske, J. C. Geyer, M. Kuschnerov, K. Piyawanno, and T. Duthel, "Optical Performance Monitoring from FIR Filter Coefficients in Coherent Receivers," in *OFC/NFOEC Optical Fiber Communication Conference*, 2008, pp. 3623–3631. 32, 38, 101, 102, 106, 119

[68] I. T. Lima, R. Khosravani, P. Ebrahimi, E. Ibragimov, C. R. Menyuk, and A. E. Willner, "Comparison of polarization mode dispersion emulators," *Journal of Lightwave Technology*, vol. 19, no. 12, pp. 1872–1881, 2001. 34

[69] G. Agrawal, "Fiber-Optic Communication Systems." Wiley, 1997, pp. 62–63. 35

[70] P. K. A. Wai, C. R. Menyuk, and H. H. Chen, "Stability of solitons in randomly varying birefringent fibers," *Optics Letters*, vol. 16, no. 16, p. 1231, 1991. 35

[71] A. O. Dal Forno, A. Paradisi, R. Passy, and J. P. Von Der Weid, "Experimental and theoretical modeling of polarization-mode dispersion in single-mode fibers," *IEEE Photonics Technology Letters*, vol. 12, no. 3, pp. 296–298, 2000. 36

[72] M. Karlsson, "Probability density functions of the differential group delay in optical fiber communication systems," *Journal of Lightwave Technology*, vol. 19, no. 3, pp. 324–331, 2001. 36

[73] P. J. Winzer and G. J. Foschini, "MIMO capacities and outage probabilities in spatially multiplexed optical transport systems," *Optics Express*, vol. 19, no. 17, p. 16680, 2011. 37

[74] G. Biondini, W. L. Kath, and C. R. Menyuk, "Importance sampling for polarization-mode dispersion: Techniques and applications," *Journal of Lightwave Technology*, vol. 22, no. 4, pp. 1201–1215, 2004. 37

[75] M. Winter, "Effect of polarization-mode dispersion on transmission schemes employing orthogonal polarizations." Studieanarbeit, Technische Universität Berlin, 2003. 38

[76] P. Lu, L. Chen, and X. Bao, "Statistical distribution of polarization-dependent loss in the presence of polarization-mode dispersion in single-mode fibers," *IEEE Photonics Technology Letters*, vol. 13, no. 5, pp. 451–453, 2001. 39

[77] N. S. Bergano, "Undersea Communication Systems," in *Optical Fiber Telecommunications*. Academic Press, 2002, vol. IV-B, pp. 154–197. 44

[78] P. Lu, S. Mihailov, L. Chen, and X. Bao, "Importance Sampling for the Combination of Polarization Mode Dispersion and Polarization Dependent Loss," in *Optical Fiber Communication Conference (OFC)*, 2003. 51

[79] S. J. Savory, "Digital coherent optical receivers: Algorithms and subsystems," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 5, pp. 1164–1179, 2010. 53, 141

[80] D. Rafique and A. D. Ellis, "The impact of signal-ASE four-wave mixing in coherent transmission systems," *OFC/NFOEC Optical Fiber Communication Conference*, pp. 3449–3454, 2011. 53

[81] D. Semrau, D. Lavery, L. Galdino, R. I. Killey, and P. Bayvel, "The Impact of Transceiver Noise on Digital Nonlinearity Compensation," *Journal of Lightwave Technology*, vol. 36, no. 3, pp. 695–702, 2018. 53

[82] S. M. Bilal, K. Goroshko, H. Louchet, I. Koltchanov, and A. Richter, "Nonlinearities tolerant modulation format enabled Tb/s superchannel transmission over 420 km of unrepeated Raman amplified link," *Optical Fiber Technology*, vol. 36, pp. 306–311, 2017. 58

[83] V. Musara, L. Wu, and A. W. Leitch, "A fixed differential group delay but varying Second-Order polarization mode dispersion emulator," in *International Conference on Broadband Communications, Informatics and Biomedical Applications (BroadCom)*, 2008, pp. 468–472. 59

[84] H. Louchet, K. Kuzmin, and A. Richter, "Improved DSP algorithms for coherent 16-QAM transmission," in *European Conference on Optical Communication (ECOC)*, no. Tu.1.E.6, 2008. 60

[85] H. Louchet, "Top-Down Analysis of High-Capacity Fiber-Optic Transmission," Ph.D. dissertation, Technische Universität Berlin, 2006, visited on 15 May 2018. [Online]. Available: http://dx.doi.org/10.14279/depositonce-1445 64, 65, 68

[86] W. Rugh, *Nonlinear System Theory The Volterra/Wiener Approach*. Johns Hopkins University Press, 1981. 65

[87] ITU-T, "Rec. G.976 (05/14) Test methods applicable to optical fibre submarine cable systems," 2014, visited on 26 March 2018. [Online]. Available: https://www.itu.int/rec/T-REC-G.976/en 98, 143

[88] C. D. Poole, N. S. Bergano, R. E. Wagner, and H. J. Schulte, "Polarization Dispersion and Principal States in a 147-km Undersea Lightwave Cable," *Journal of Lightwave Technology*, vol. 6, no. 7, pp. 1185–1190, 1988. 101, 103, 106

[89] K. H. Han, W. J. Lee, and J. S. Ko, "Improved technique for tracking of principal states of polarization and OSNR monitoring," in *Optical Fiber Communication Conference (OFC)*, vol. 2, no. FF4, 2004, p. 3. 101, 103, 106

[90] F. Pittal, F. N. Hauske, Y. Ye, N. G. Gonzalez, and I. T. Monroy, "Fast and robust CD and DGD estimation based on data-aided channel estimation," in *International Conference on Transparent Optical Networks (ICTON)*, no. We.D1.5, 2011. 101, 103, 106

[91] I. Parberry, "Problems on Algorithms," *ACM SIGACT News*, vol. 26, no. 2, pp. 50–56, 1995. 104

[92] J. M. Borwein and P. B. Borwein, *Pi and the AGM–A Study in Analytic Number Theory and Computational Complexity.* Wiley, 1987. 104

[93] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications, Chapter IV.C," *Journal of Lightwave Technology*, vol. 32, no. 4, pp. 694–721, 2014. 124

[94] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Optics Express*, vol. 22, no. 13, 2014. 124

[95] R. A. Shafik, M. S. Rahman, and A. H. Islam, "On the extended relationships among EVM, BER and SNR as performance metrics," in *International Conference on Electrical and Computer Engineering (ICECE)*, 2007, pp. 408–411. 126, 147

[96] A. Thetford, "Introduction to Matrix Methods in Optics," *Optica Acta: International Journal of Optics*, vol. 23, no. 3, pp. 255–256, 1976. 136

[97] K. Kikuchi, "Polarization-demultiplexing algorithm in the digital coherent receiver," in *2008 Digest of the IEEE/LEOS Summer Topical Meetings*, 2008, pp. 101–102. 141

[98] E. Rowell, "Know Thy Complexities!" 2013, visited on 3 May 2018. [Online]. Available: http://bigocheatsheet.com/ 145

Acronyms

AI artificial intelligence. 2

ASE amplified spontaneous emission. i, 11–13, 15, 21, 34, 41, 42, 51–53, 58, 118, 120, 123, 166

BER bit error ratio. 51–53, 57, 58, 83, 90, 94, 95, 137

BPS blind phase search. 57

BPSK binary phase shift keying. 11

CAGR compound annual growth rate. 3

CAZAC constant-amplitude zero-autocorrelation. 99

CD chromatic dispersion. 11, 15, 16, 18, 60, 64, 65, 74

CMA constant modulus algorithm. 51, 136

DAC digital to analog converter. 55

dB decibel. 9

DBP digital back propagation. i, ix, 6–9, 11–13, 15, 19–21, 41, 47, 50–52, 54–58, 69, 74, 76, 80, 81, 83–88, 90, 92, 94–97, 99–103, 105–107, 109–112, 115, 117, 120–125, 127, 128, 133, 134

DBPmod modified digital back propagation. 88, 94, 95, 105, 107, 109, 117, 125, 134

DCF dispersion compensating fibre. 21

DCN distributed content network. 5

DGD differential group delay. 13, 25–28, 31, 32, 34–37, 42–47, 51–54, 56–59, 66–69, 71–73, 77, 79–81, 83–88, 90, 92, 94–101, 103, 107, 109–115, 121, 123, 125, 127, 128, 133, 134, 163

DGDev differential group delay evolution. 43–45, 59, 69–71, 73, 77, 79, 80, 86, 87, 101, 127, 128

DM dispersion managed. 11, 20, 63

DP dual polarisation. 11, 21, 50–52, 56, 74, 81, 83, 88, 91–93, 96, 105–107, 120–122, 124

DSP digital signal processing. 4, 18, 51, 90, 125

EB Exabytes. 3

ECL external cavity laser. 55

EDFA erbium-doped fibre amplifier. 2, 6, 11, 15, 20, 22, 29, 34, 38, 39, 41, 42, 49–51, 54, 56, 58, 60, 63, 64, 71, 74, 88, 90, 118–120, 122, 124, 127, 165, 168

EVM error vector magnitude. 105, 120, 141

FEC forward-error correction. 4, 94

FFT fast Fourier transform. 74, 86, 99, 113, 166

FIR finite impulse response. 101

GN Gaussian noise. 118

GVD group velocity dispersion. 16, 18, 30

IFFT inverse fast Fourier transform. 74

IP Internet protocol. 3

IS importance sampling. 35–38

LO local oscillator. 51, 57

LPF low-pass filter. 20

MaxPhCh maximum phase change. 20, 21, 51, 57, 88, 110, 111

MIMO multiple-input multiple-output. 51, 74, 90, 97–101, 111–114, 135, 136

MLSE maximum-likelihood sequence estimation. 6

MMA multiple modulus algorithm. 51, 57, 58, 74, 90, 136

NF noise figure. 22, 41, 42, 92, 105, 118, 119, 122, 124

NFT nonlinear Fourier transform. 6, 9, 11

NL nonlinearities. i, 5–9, 11–13, 17, 18, 20, 34, 60–62, 74, 76, 81–83, 118, 120, 121, 123, 125, 127, 128

NL-ASE nonlinearities amplified spontaneous emission. 120

NL-PMD nonlinearities polarisation mode dispersion. 13, 58, 86, 107, 115, 117, 120, 125, 128

NLSE nonlinear Schrödinger equation. 17, 18

non-DM non-dispersion managed. 2, 11, 12, 20, 50, 56, 88, 127

NRZ non-return-to-zero. 50

OFDM orthogonal frequency division multiplexing. 11

OPC optical phase conjugation. 6, 7, 11

OSNR optical signal-to-noise ratio. i, 3, 4, 22, 51–54, 57, 90–94, 96, 107, 118, 120, 128, 141

PAM pulse-amplitude modulation. 55

PBS polarisation beam splitter. 99

PCD polarisation-dependent chromatic dispersion. 33

PDG polarisation dependent gain. 15, 24, 29, 37

PDL polarisation dependent loss. i, 13, 15, 21, 24, 29–32, 34–39, 41, 47–49, 51, 54, 55, 58, 97, 98, 127

PMD polarisation mode dispersion. i, ix, 7–13, 15, 21, 24, 25, 27, 29–38, 41–44, 47–49, 51, 52, 55–58, 62, 65, 68, 72, 74, 76–86, 88, 91, 92, 94–97, 99, 101–103, 105, 107, 109–113, 118, 120, 122–125, 127–129, 133, 134

PMDc polarisation mode dispersion coefficient. 9, 11, 13, 27, 36, 42–47, 49, 51, 54, 77, 85, 88, 92–94, 96, 105–107, 111, 120–122, 124, 127, 133, 134

PPG pulse pattern generator. 55

PSD polarisation state depolarisation. 33, 129

PSP principal state of polarisation. 30–32, 35, 37, 43, 66, 85, 86, 95–101, 103, 107, 109, 110, 112, 115, 128, 129, 162

QAM quadrature amplitude modulation. 6, 11, 50–52, 56, 74, 81, 83, 88, 91–96, 105–107, 109, 120–122, 124, 133, 136

QPSK quadrature phase shift keying. 11, 21, 50–52, 92, 136

RD reach distance. 117, 120, 121, 123, 125

ROSNR required optical signal-to-noise ratio. 118, 119

Rx receiver. 21, 103

SE spectral efficiency. 117–121, 123, 125

SER symbol-error rate. 83

SNR signal-to-noise ratio. 9, 105, 106, 120, 141

SOAs semiconductor optical amplifiers. 29

SPM self-phase modulation. 21

SSF split-step Fourier. i, 12, 13, 18, 34, 59, 74, 77–82, 84, 88, 125, 127, 128

SSMF standard single mode fibre. 3, 6, 11, 15–17, 23, 37, 38, 42, 47, 50, 56, 58, 63–65, 71, 74, 88, 105, 107, 122, 124

VR virtual reality. 2

WDM wavelength division multiplexing. 2, 20

ZB Zettabytes. 3

List of symbols

!	Factorial	16
2ζ	Azimuth of the rotation angle	33
$< >$	Averaging operator	26
$< \tau >$	Polarisation mode dispersion	24
$\| \ \|$	Absolute value	76
α	Attenuation coefficient	16
α_x	Attenuation coefficient in X axis	49
α_y	Attenuation coefficient in Y axis	49
β	Phase propagation constant	16
β_0	Phase propagation constant at central frequency	16
β_2	Group velocity dispersion parameter	16
β_3	Dispersion slope parameter	16
β_{nl}	Nonlinear phase propagation constant	16
$\Delta\Omega$	Term related to signal frequency	66
Δn_{eff}	Difference between ordinary and extraordinary fibre refractive indices	24
δ	Dirac function	64
$\Delta\alpha_y$	Difference in attenuation coefficients in X and Y axes	49

$\Delta\beta$	Propagation constant difference between X and Y polarisation compoments 17
$\delta_{nl\_x,y}$	First-order perturbative term 66
ϵ	Cost function 144
η	Link nonlinear transfer function 66
γ	Fibre nonlinear coefficient 17
Γ_{max}	Maximum light transmission coefficient 29
Γ_{min}	Minimum light transmission coefficient 29
$\hat{D}$	Linear operator in DBP 18
$\hat{N}$	Nonlinear operator in DBP 18
$\hat{P}$	Power correction operator in DBP 18
$\hbar$	Plank's constant 44
$\int$	Integral 29
κ	NL compensation coefficient 126
Λ	Set of possible amplitudes for a specific modulation format . 144
λ	Wavelength 16
$\mathbb{A}_x,\ \mathbb{A}_y$	Delayed signal A_x and A_y components 70
$\mathbb{D}$	Dispersion component of channel transfer matrix 32
$\mathbb{G}$	PDL vector 34
$\mathbb{K}$	PDL component of channel transfer matrix 32
$\mathbb{M}$	Mueller rotation matrix 138
$\mathbb{O}$	PMD vector 34
$\mathbb{P}$	Probability density function 28