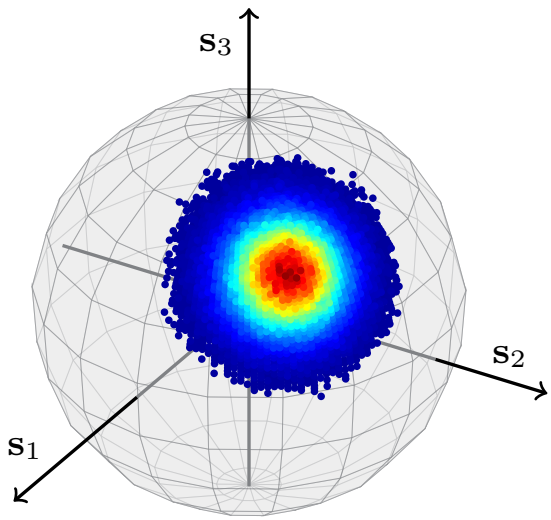
CHALMERS



Modeling and Compensation of Polarization Effects in Fiber-Optic Communication Systems

CRISTIAN B. CZEGLEDI

Department of Electrical Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2018

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Modeling and Compensation of Polarization Effects in Fiber-Optic Communication Systems

Cristian B. Czegledi

CHALMERS
UNIVERSITY OF TECHNOLOGY

Communication Systems Group
Department of Electrical Engineering
Chalmers University of Technology
Gothenburg, Sweden, 2018

Modeling and Compensation of Polarization Effects in Fiber-Optic Communication Systems

Cristian B. Czegledi
ISBN 978-91-7597-702-7

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie Nr 4383
ISSN 0346-718X

This thesis has been prepared using LaTeX and PSTricks.

Communication Systems Group
Department of Electrical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg, Sweden
Phone: +46 (0)31 772 1000
www.chalmers.se

Front cover illustration:
Histogram of a random-walk step of the state of polarization illustrated on the Poincaré sphere (see [Paper A] for details).

Printed by Chalmers Reproservice
Gothenburg, Sweden, February 2018

Abstract

Optical communication systems that exploit the orthogonality between two polarizations of light convey information over optical fibers by modulating data over the two polarizations. In an idealized scenario, the two polarizations propagate through the fiber without interfering. However, this is not the case for practical fibers, which suffer from various imperfections that lead to polarization-related interference between the two polarizations. This thesis is concerned with polarization effects that arise in communication systems over optical fibers. In particular, we consider modeling and compensation of such effects, and their impact on and improvement of nonlinearity mitigation algorithms.

The impact of an impairment on the performance of a transmission system can be understood via a channel model, which should describe the behavior of the channel as accurately as possible. A theoretical framework is introduced to model the stochastic nature of the state of polarization during transmission. The model generalizes the one-dimensional carrier phase noise random walk to higher dimensions, modeling the phase noise and state of polarization drift jointly as rotations of the electric field and it has been successfully verified using experimental data. Thereafter, the model is extended to account for polarization-mode dispersion and its temporal random fluctuations. Such models will be increasingly important in simulating and optimizing future systems, where sophisticated digital signal processing will be natural parts.

The typical digital signal processing solution to mitigate phase noise and drift of the state of polarization consists of two separate blocks that track each phenomenon independently and have been developed without taking into account mathematical models describing the impairments. Based on the proposed model for the state of polarization, we study a blind tracking algorithm to compensate for these impairments. The algorithm dynamically recovers the carrier phase and state of polarization jointly for an arbitrary modulation format. Simulation results show the effectiveness of the proposed algorithm, having a fast convergence rate and an excellent tolerance to phase and polarization noise.

The optical fiber is a nonlinear medium with respect to the intensity of the incident light. This effect leads to nonlinear interference as the intensity of light increases, which made nonlinear interference mitigation techniques to be an intensively studied topic. Typically, these techniques do not take into account polarization-mode dispersion, which becomes detrimental as the nonlinear effects interact with polarization-mode dispersion. We study digital-domain nonlinear interference mitigation algorithms that take into account polarization-mode dispersion by i) reversing the polarization effects concurrently with reversing the nonlinear effects and by ii) mitigating only the polarization-insensitive nonlinear contributions. These algorithms will be increasingly important in future optical systems capable of performing large bandwidth nonlinear interference mitigation, where even small amounts of polarization-mode dispersion become a limiting factor.

Keywords: Channel model, model-based, polarization demultiplexing, polarization drift, polarization-mode dispersion, PMD, DBP, nonlinear compensation, backpropagation.

List of Publications

This thesis is based on the following publications:

(A) C. B. Czegledi, M. Karlsson, E. Agrell, and P. Johannisson, "Polarization drift channel model for coherent fibre-optic systems," *Nature Scientific Reports*, vol. 6, art. no. 21217, Feb. 2016.

(B) C. B. Czegledi, M. Karlsson, P. Johannisson, and E. Agrell, "Temporal stochastic channel model for absolute polarization state and polarization-mode dispersion," in *Proc. Optical Fiber Communication Conference (OFC)*, Los Angeles, CA, Mar. 2017, p. Th3F.2.

(C) C. B. Czegledi, E. Agrell, M. Karlsson, and P. Johannisson, "Modulation format independent joint polarization and phase tracking for optical communications," *IEEE/OSA Journal of Lightwave Technology*, vol. 34, no. 14, pp. 3354–3364, July 2016.

(D) C. B. Czegledi, G. Liga, D. Lavery, M. Karlsson, E. Agrell, S. J. Savory, and P. Bayvel, "Polarization-mode dispersion aware digital backpropagation," in *Proc. European Conference on Optical Communication (ECOC)*, Düsseldorf, Germany, Sept. 2016, pp. 1091–1093.

(E) G. Liga, C. B. Czegledi, and P. Bayvel, "A PMD-adaptive DBP receiver based on SNR optimization," *to appear in Proc. Optical Fiber Communication Conference (OFC)*, San Diego, CA, Mar. 2018.

(F) 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, pp. 1903–1915, Jan. 2017.

(G) C. B. Czegledi and R. Dar, "Volterra series digital backpropagation accounting for PMD," in *Proc. European Conference on Optical Communication (ECOC)*, Gothenburg, Sweden, Sept. 2017, p. W.1.D.1.

Publications by the author not included in the thesis:

(a) C. B. Czegledi, M. R. Khanzadi, and E. Agrell, "Bandlimited power-efficient signaling and pulse design for intensity modulation," *IEEE Transactions on Communications*, vol. 62, no. 9, pp. 3274–3284, Sept. 2014.

(b) C. B. Czegledi, M. R. Khanzadi, and E. Agrell, "Bandlimited power-efficient signaling for intensity modulation," in *Proc. European Conference on Optical Communication (ECOC)*, Cannes, France, Sept. 2014, p. P.3.7.

(c) C. B. Czegledi, E. Agrell, and M. Karlsson, "Symbol-by-symbol joint polarization and phase tracking in coherent receivers," in *Proc. Optical Fiber Communication Conference (OFC)*, Los Angeles, CA, Mar. 2015, p. W1E.3.

(d) M. Karlsson, C. B. Czegledi, and E. Agrell, "Coherent transmission channels as 4d rotations," (invited paper) in *Proc. Signal Processing in Photonic Communication (SPPCom)*, Boston, MA, Jul. 2015, p. SpM3E.2.

(e) C. B. Czegledi, "Modeling and Tracking of Stochastic Polarization Drifts in Fiber-Optic Systems," Gothenburg, Sweden, Licentiate Thesis, Feb. 2016.

(f) G. Liga, C. B. Czegledi, T. Xu, E. Agrell, R. I. Killey, and P. Bayvel, "Ultra-wideband nonlinearity compensation performance in the presence of PMD," in *Proc. European Conference on Optical Communication (ECOC)*, Düsseldorf, Germany, Sept. 2016, pp. 794–796.

(g) L. Galdino, D. Semrau, D. Lavery, G. Saavedra, C. B. Czegledi, E. Agrell, R. I. Killey, and P. Bayvel, "On the limits of digital back-propagation in the presence of transceiver noise," *Optics Express*, vol. 25, no. 4, pp. 4564–4578, Jan. 2017.

(h) C. B. Czegledi, G. Liga, D. Lavery, M. Karlsson, E. Agrell, S. J. Savory, and P. Bayvel, "Modified digital backpropagation accounting for polarization-mode dispersion," in *Proc. Optical Fiber Communication Conference (OFC)*, Los Angeles, CA, Mar. 2017, p. W1G.6.

(i) R. Dar and C. B. Czegledi, "Backward propagation with compensation of some nonlinear effects of polarization mode dispersion," Patent Application Ref. 820681-US-NP (o/r 990.0866), filed 15 Jun. 2017.

Acknowledgments

This thesis is not only the end result of my hours spent at the keyboard, but the outcome of four incredible and joyful years supported by a continuous collaboration and encouragement from many people that I wish to thank.

My first big thank you goes to my main supervisor Erik Agrell for his support and guidance throughout these years. Working with you is such a pleasure! The way you do research, your critical thinking, and the confidence you put in your students are things I really appreciated and learned from. You introduced me to the realm of preciseness; a resource[1] I will benefit from regardless of what the future will be like. I am also grateful to my co-supervisor Magnus Karlsson for his patience with all my questions and misunderstandings about photonics. I feel very lucky that I have been supervised by you and I enjoyed all our discussions about rotation matrices and differential equations[2]. I would also like to thank Pontus Johannisson for his unlimited constructive feedback, pushing me to revise and improve my drafts.

I would like to extend my thanks to friends and colleagues at Comsys for creating an exciting work place and for the good times we had in the past years. These four years were amazing! Special thanks go to Àlex, Andreas, Árni, Christopher, Christian, Erik, Fredrik, Gabo, Henk, Jesper, Johan, Kathi, Keerthi, Li, Markus, Misha, Naga, Nil, Rahul[3], Rajet, Reza, Sven, Themis, and Wanlu. You guys rock! Many thanks to Erik Ström for his dedication in creating an inspiring research group. I would also like to thank every member of FORCE for providing a unique and diverse research environment.

I am forever indebted to the Optical Network Group at University College London, where I had the privilege to collaborate with many talented people. I specially thank Polina Bayvel for giving me the opportunity to spend four months there, Domaniç Lavery for introducing me to the digital backpropagation, and Gabriele Liga for being a good friend and collaborator, and for showing me all the good restaurants in London.

I would like to acknowledge Peter Winzer for giving me the opportunity to spend four months at Nokia Bell Labs, NJ. I would also like to thank Ronen Dar for his support and for teaching me about nonlinearities. Many thanks to René–Jean Essiambre for all the interesting non-work-related discussions.

My gratitude goes to my parents, who raised me with an appreciation for science and supported me in my pursuits. Last but not least, I want to thank Alexandra for her constant support and motivation. After dedicating myself to this, now it is time to dedicate myself to you.

Cristian B. Czegledi

Göteborg, February 2018

[1] This resource should be used with care; thanks to it I am called the "grammar police" in the group.

[2] I am still looking forward to see you eat Riccati equations for lunch though.

[3] For the uninformed, it is pronounced "Reeeeehul", or simply "Sassy Devassy".

Financial support

This work was funded in part by the Swedish Research Council (VR) under grants #2010-4236, #2012-5280, and #2013-5271, by the Swedish Foundation for Strategic Research (SSF) under grant #RE07-0026, and by Vinnova under grant #2017-05228. I would also like to thank Ericsson's Research Foundation, the Optical Networks Group at University College London, and Nokia Bell Labs for partially sponsoring my research visits. The simulations were performed in part on resources at the Chalmers Centre for Computational Science and Engineering (C3SE) provided by the Swedish National Infrastructure for Computing (SNIC).

Acronyms

4D	four-dimensional
ASE	amplified spontaneous emission
AWGN	additive white Gaussian noise
BPS	blind phase search
CD	chromatic dispersion
CMA	constant modulus algorithm
DBP	digital backpropagation
DGD	differential group delay
DOF	degree of freedom
DSP	digital signal processing
MMA	multiple modulus algorithm
NLSE	nonlinear Schrödinger equation
pdf	probability density function
PM	polarization-multiplexed
PMD	polarization-mode dispersion
QAM	quadrature-amplitude modulation
QPSK	quadrature phase-shift keying
SNR	signal-to-noise ratio
SOP	state of polarization
SSFM	split-step Fourier method

Contents

Part I

Overview

CHAPTER 1

Background

Digital communication systems have an important role in today's modern society and have changed the way we connect to the world. Historically, information has been conveyed through many different media ranging from cave paintings, rock-carved petroglyphs, and smoke signals to the modern, digital communication network that we have today, known as the *Internet*. In today's digital era, it is easier than ever to access information, which leads to a rapid growth in popularity of social media, online gaming, and broadcast multimedia systems. Moreover, various emerging technologies, such as traffic safety, Internet-of-Things, and virtual reality, are about to materialize, putting more pressure on the Internet service providers to support high-speed Internet connections and motivates the need for the design of faster communication systems. The long-term trend is summarized in Table 1.1, showing that the total Internet traffic has experienced a remarkable growth in the last two decades according to CISCO's statistics and predictions [1].

This remarkable growth is supported by fiber-optic communications, which is the fastest form of communication technology available today and comprise the backbone of the Internet. The optical technology transports information using light in the near-infrared spectrum over distances varying from a few meters in data-centers to thousands of kilometers over transoceanic links.

The optical revolution started in 1966 with Kao and Hockham's [2] vision of using silica glass as a medium for guided transmission at optical frequencies. Later on in 1970 followed the first demonstration of < 20 dB/km optical fiber loss [3], starting the fiber-optic-communications era. The commercial deployment of optical cables started in the 1980s when they revolutionized the communication networks and became the

Table 1.1. Global Internet traffic: past and forecast [1].

Year	Global Internet Traffic
1992	100 GB per day
1997	100 GB per hour
2002	100 GB per second
2007	2000 GB per second
2016	26600 GB per second
2021	105800 GB per second

predominant transmission medium in telecommunication links. The first transoceanic link was installed in 1988 connecting USA and Europe [4]. The invention of optical amplifiers [5,6] enabled wavelength-division multiplexing initiating a massive investment in system development. This technical revolution increased the data rates of commercial optical systems from approximately 1 Gb/s in the mid-1980s to 1 Tb/s by 2000 [7].

Although communication over fibers offers high data rates, historically, the employed transmission schemes utilized the available spectrum rather inefficiently. As an example, the digital data was represented as the presence or absence of light, known as on–off keying. As the request for higher data rates increased tremendously in recent times, this spectrally inefficient transmission method was replaced by more sophisticated methods that modulate the data over multiple dimensions of the optical field. However, the improved spectral efficiency comes at the cost of expensive hardware, which is still prohibitive for shorter-reach applications where on–off keying is widely used. In conjunction with the coherent receiver, digital signal processing (DSP) tools made it possible to achieve transmission schemes with high-spectral efficiency that convey data through multiple levels of the amplitude and phase of the transmitted light [8]. Coherent detection was first envisioned as early as 1980s for attenuation-limited single-span transmissions, where the receiver sensitivity is severely limited [9]. Research on coherent systems decayed with the advent of efficient optical amplifiers in the beginning of 1990s, and then revived in the 2000s followed by the first demonstration in 2008 [10].

Another dimension that improves the throughput of optical systems even further is the polarization of light. Information is conveyed in two orthogonal polarizations of light. These systems are known as being polarization-multiplexed (PM) and they can double the throughput compared to a single-polarization transmission. In order to benefit from the native four-fold dimensionality of the optical field, consisting of the in-phase and quadrature components[1] in two polarizations, PM transmission with four degrees of freedom (DOFs) has been constructed for optical channels [11]. PM-quadrature phase-shift

[1] The polar representation of the optical field in each polarization, i.e., its amplitude and phase, can be described in Cartesian coordinates as the sum of two orthogonal in-phase and quadrature components.

keying (PM-QPSK) commercially introduced in 2008 for 40 Gb/s per channel transmission, and then adopted for 100 Gb/s per channel and 10 Tb/s per fiber transmission in 2010, has now been widely deployed and reached maturity. Recently, 200 Gb/s per channel transceivers have been made commercially available based on PM-16-quadrature-amplitude modulation (PM-16-QAM) and it is expected that in the near future, higher-order PM-M-QAM modulation formats will become a necessity for higher data rates. However, the improved spectral efficiency comes at the cost of a reduced tolerance to impairments, such as additive noise, nonlinearities, or laser phase noise. Therefore, future higher-order modulation formats require more powerful DSP that can mitigate these impairments accurately.

Although higher-order modulation formats do improve the spectral efficiency, besides the fact that they require sophisticated engineering and DSP technologies to overcome practical and fundamental obstacles, the returns in terms of throughput diminish as the size of the constellation increases. Therefore, the pursuit for higher data rates has mobilized in recent years the research front to explore another dimension and that is *space*. The idea of multiple spatial channels, referred to as space-division multiplexing, was first approached in 1979 [12] by using arrays of thin single-core fibers, so called fiber bundles or multi-element fibers, that share the same coating. However, this approach does not offer big integration advantages. An option that is more integrated is to incorporate several cores into the cross-section of a single glass strand, refereed to as multicore fiber [13]. Another alternative enabling spatial multiplexing is using individual modes of a multimode fiber, where each mode is considered to be a separate spatial channel. The current target of spatial division multiplexing aims to combine multiple approaches in order to achieve higher levels of spatial channels multiplicity and diversity. A more detailed tutorial review on spatial division multiplexing techniques can be found in [14] and references therein.

The propagation in optical fibers, within some parameter range (see Section 3.1), is governed by the nonlinear Schrödinger equation (NLSE), regardless of the transmission scheme or fiber type. The NLSE is a partial differential equation with respect to time and propagation distance. Although many attempts, an exact input–output relationship has not been explicitly found yet. As hinted by its name, the NLSE is nonlinear in respect to the power of the field. Therefore, communication techniques developed for (linear) wireless communications do not achieve optimality or simply do not work in the power regimes where the nonlinear effects become significant. This has encouraged the research community to focus its efforts also on nonlinearity mitigation techniques. These can be categorized mainly into two groups: analog in the optical domain and digital in the electrical domain. The first category includes large effective area fibers [15] that allow higher signal launch powers before reaching the nonlinear regime, optical phase conjugation using twin waves [16], or mid-span optical phase conjugation [17]. Various digital techniques are available in the literature to mitigate fiber nonlinearities, including perturbation-based precompensation [18], Kalman equalization [19], least-squares equal-

ization [20], and digital backpropagation (DBP) [21–24]. More on nonlinear interference mitigation techniques can be found in [25–27] and references therein.

Regardless of the transmission scheme or fiber type, before commercial deployment, newly-designed DSP algorithms must be verified and tuned. This can be done practically in controlled lab environments or field-trial experiments, or numerically using simulations. Simulations offer a greater flexibility than experiments and can be used to predict the behavior of a system and quantify performance. Nonetheless, simulations rely on mathematical channel models, which, in order to achieve conclusive results, should describe propagation through fibers accurately.

In this thesis, we are investigating channel modeling and DSP algorithms for fiber-optic communication systems. We first propose a channel model for the temporal drift of the state of polarization (SOP). In order to assess the accuracy of the model, we validate it by comparing it to measured data over a 127-km long buried fiber link. Thereafter, we extend the channel model to account for polarization-mode dispersion (PMD) and its temporal drift. Based on the channel model for the SOP, we derive a DSP algorithm that can track ultra-fast joint phase noise and drifts of the SOP. The last part of the thesis is concerned with nonlinearity mitigation DSP algorithms in the presence of PMD. PMD reduces the effectiveness of nonlinearity mitigation algorithms and we are proposing four different variants of algorithms that account for PMD.

1.1 Thesis Organization

This thesis is a support of candidature for the doctor of philosophy degree. The thesis documents the progress made by the candidate over a period of approximately four years. This thesis is formatted as a collection of papers, where Part I serves as an introduction to Part II consisting of the appended papers.

The remainder of the thesis is structured as follows. Chapter 2 presents a general overview of communications over optical fibers by discussing the fiber-optic channel and impairments and that arise in fiber-optic communication systems and their mitigation. In Chapter 3, the background information for the considered mathematical representation of optical signals is introduced, and continuous-time and discrete-time channel models for the fiber-optic channel present in the literature are reviewed. This chapter provides the preliminaries for Papers A–G regarding the adopted channel models therein. Chapter 4 serves as background for Papers C–G by providing a summary of various DSP algorithms for nonlinearity mitigation, phase-noise compensation, and polarization-tracking schemes. Chapter 5 provides a short description of the appended contributions in Part II.

1.2 Notation

The following notation conventions are used throughout the thesis: column vectors are denoted by bold lower case (e.g., $\mathbf{u}$) and matrices by bold upper case (e.g., $\mathbf{U}$), except a

few specific cases, for literature consistency reasons, denoted by lowercase Greek letters such as the Pauli matrices $\boldsymbol{\sigma}_i$, the basis matrices $\boldsymbol{\rho}_i$, $\boldsymbol{\lambda}_i$, and the electric field Jones vector $\mathbf{E}$. Transposition is written as $\mathbf{u}^{\mathrm{T}}$, conjugation as $\mathbf{u}^*$, and conjugate (Hermitian) transpose as $\mathbf{u}^{\mathrm{H}}$. The $n \times n$ identity matrix is written as $\mathbf{I}_n$ and the expectation operator as $\mathbb{E}[\cdot]$. The dot operation $\boldsymbol{\alpha} \cdot \vec{\boldsymbol{\sigma}}$ should be interpreted as a linear combination of the three matrices forming the tensor $\vec{\boldsymbol{\sigma}} = (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3)$. Multiplication of a matrix with the tensor $\vec{\boldsymbol{\sigma}}$ results in a tensor with element-wise multiplications, e.g., $\mathbf{U}\vec{\boldsymbol{\sigma}} = (\mathbf{U}\boldsymbol{\sigma}_1, \mathbf{U}\boldsymbol{\sigma}_2, \mathbf{U}\boldsymbol{\sigma}_3)$. The absolute value is denoted by $|\cdot|$ and the Euclidean norm by $\|\cdot\|$. The imaginary unit is denoted $i = \sqrt{-1}$.

Notational Inconsistencies

We warn the reader about the following inconsistencies in the notation across the appended papers in Part II and the thesis overview given in Part I:

- The phase noise is modeled as $e^{i\phi}$ in Paper A, whereas $e^{-i\phi}$ is used in the thesis overview and Paper C.
- The variable L is used to denote the length of the fiber link throughout the entire thesis, except in Paper F where it is a design parameter in the multipoint differentiation (F.14).
- The number of fiber spans is denoted with N_{sp} in the thesis overview, opposed to N_{s} used in Paper F.
- The variable $\mathbf{T}(z, f)$ is used to denote the accumulated PMD at distance z and frequency f in the thesis overview, whereas in Paper F it is used as $\mathbf{T}_k^{\mathrm{eq}}[n]$ to denote the nth tap of the impulse response of the equalizer at discrete-time k, and in Paper B is used as $\mathbf{T}(k, f)$ to denote the accumulated PMD over the entire link at discrete-time k and frequency f.
- In the thesis overview, the parameter N_{PMD} denotes the number of birefringent sections used to model PMD, whereas in Papers D, E, and F it denotes the number of birefringent sections used by the proposed algorithms to compensate for PMD in the digital backwards propagation.
- In Papers D and F, N_{DBP} denotes the total number of DBP steps per link, whereas in Paper E it is used to denote the number of DBP steps between two PMD steps.
- The variable $\mathbf{J}_k$ is used to denote drifts of the SOP in the overview of the thesis and Papers A and C, whereas $\mathbf{J}_k(f)$ is used to denote the inverse of the accumulated PMD over the link in Paper F. On the other hand, in Paper G, $\mathbf{J}(\omega, z)$ denotes the accumulated PMD at distance z and angular frequency ω.
- The mean DGD of the PMD retardation plates is denoted with τ_p in the thesis overview and Paper B, wheres the notation $\Delta\tau_p$ was adopted in Papers D and F.

CHAPTER 2

Fiber-Optic Communication Systems

The inherent difference between optical communication systems and microwave systems is in the carrier frequency range they operate at. The optical carrier frequencies are around 193 THz, whereas microwaves system use carrier frequencies ranging from 1 GHz up to recently 100 GHz, called millimeter waves. This high optical carrier frequency permits the use of much larger modulated bandwidths compared to microwave systems, and pushed the world development of lightwave systems.

Lightwave systems can be categorized into two groups: guided and unguided systems. In the unguided systems, as the name suggests, the transmitted beam of light is not confined in a medium but it is spread in free space, like in microwave links. Unguided systems, commonly called wireless optical communication systems, are an emerging technology and can be broadly categorized into two types of applications: broadcast [28–30] or point-to-point [31–33]. The aim of the former is to provide indoor wireless communication, whereas the latter aims to bridge the gap between the end-user and the high-capacity fiber infrastructure (the so-called "last mile" problem) or to provide high-speed intersatellite communication links. In the case of guided systems, the light remains confined in a wave guide, which is usually an optical fiber. This thesis focuses on fiber-optic communication systems only.

Since their development in the 1970s, fiber-optical communication links have been deployed for various applications. These can be largely classified into two categories, long-haul and short-haul, depending on the distance they cover. Short-haul links cover distances up to 100 km and serve intracity links, access networks, consumer electronics, and data-center applications. Such links are typically cost constrained and, in order to keep the costs down, noncoherent detection is employed, i.e., only the intensity of light is

used to convey information, disregarding the phase. On the other hand, long-haul fiber links cover transoceanic distances and, although very expensive to install and operate, the cost of such links is shared among many users, thus lowering the cost constraints. Historically, the fiber optical communication technology was driven by these long-haul links accelerated by the pursuit of the operators for higher throughput. Such links exploit coherent transceivers that modulate both the phase and amplitude of the optical field, doubling[1] the throughput, and employ periodical amplification compensating for the fiber loss.

Fiber-optic communications rely on the total internal reflection phenomenon to guide a beam of light trough a cylindrical core made of silica glass; a phenomenon known since 1854 [34]. This core is surrounded by a cladding that confines the light into the core and whose refractive index is smaller than the one of the core. The light can propagate through the fiber in different optical modes, which are solutions to Maxwell's equations [35, Ch. 2] satisfying the boundary conditions. Each solution of the wave equation is called a propagation mode. An optical fiber can be engineered to support a different number of propagation modes depending on its physical properties, such as refractive index or core area. Single-mode fibers support only one mode, known as the fundamental mode, and it is used primarily used in long-haul coherent links. Fibers with a larger core diameter than single-mode fibers allow more modes to propagate. These fibers allow the use of low-cost optoelectronics at the expense of increased interference levels, where the different modes have different transit times, thus leading to modal dispersion, and suffer from inter-modal crosstalk. This is called a multi-mode fiber and it is used mainly in short-range links such as access networks, consumer electronics, and within data centers. As previously mentioned in Chapter 1, multi-mode fibers are currently considered to improve the throughput per fiber cable by exploiting the different modes as separate spatial channels.

Although optical fibers benefit from a much lower attenuation compared to the previously used copper wires, the signal still has to be amplified periodically in order to reach long transmission distances. This amplification process is not ideal and corrupts the signal with additive noise. Besides the additive noise, the transmission may suffer from various impairments, both stochastic and deterministic, which affect the transmitted data stronger as the link length increases. In this thesis, we are mainly interested in long-haul transmission, where the data transfer is strongly affected by the channel. This chapter presents a brief description, without going into the mathematical details, of various impairments present in communications through fibers (Section 2.1), thereafter the structure of a conventional transceiver is discussed (Section 2.2).

[1] Or quadrupling since traditional noncoherent detection typically exploits single-polarization signalling.

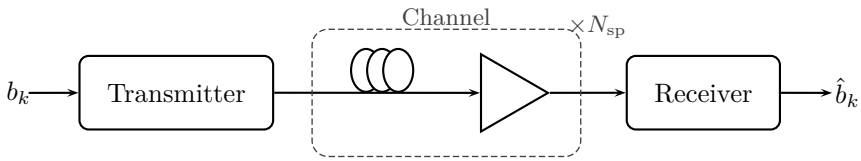


Figure 2.1. A fiber-optic link with N_{sp} spans, each consisting of a fiber span and an optical amplifier.

2.1 The Fiber-Optic Channel

Figure 2.1 shows a generic fiber-optic link. The transmitter maps the bits b_k to an optical waveform, which propagates through the channel to the receiver. The receiver outputs an estimate of the transmitted bits $\hat{b}_k$ based on the received optical waveform.

2.1.1 Signal Attenuation and Additive Noise

The optical signal attenuates as it propagates through the fiber and as the receiver requires a minimum signal power, fiber losses became a limiting factor of reaching long transmission distances. Optical fibers with practical attenuation coefficients became available in the 1970s, when silica fibers were introduced [36]. Silica fibers have a wavelength-dependent loss spectrum, exhibiting a minimum of around 0.2 dB/km in the 1550 nm region and a secondary minimum below 0.5 dB/km around 1300 nm. Although these values are close to the fundamental limit of about 0.16 dB/km for silica fibers [35, Sec. 2.5.1], they are still high enough to prohibit long-haul transmission. For example, considering a fiber link from Göteborg, Sweden, to Cluj-Napoca, Romania, of approximately 2000 km with 0.2 dB/km attenuation, the accumulated loss becomes 400 dB, corresponding to a power loss of 10^{40}, which leads to a practically undetectable signal at the receiver [37]. Therefore, improvements that allow transmission over thousands of kilometers were sought after by other means. Practical optical amplifiers became available in the 1990s, which allowed transmission over very long distances by periodical amplification (illustrated in Fig. 2.1) in the optical domain [5, 6]. For long-haul communications, the span length is typically in the 50–120 km range, after which the signal is amplified. This setup is called lumped amplification and is widely used in current installed systems. On the other hand, distributed amplification has gained recent research interest due to its lower noise figure and wider amplification bandwidth, where the amplification is carried out throughout the entire link [38].

Optical amplifiers add noise to the transmitted signal through the generation of amplified spontaneous emission (ASE), thus degrading the signal-to-noise ratio (SNR). Fig. 2.2 shows the power profile of the signal and the accumulated noise during propagation. As can be seen, the signal suffers from attenuation in each span, after which the signal is amplified, thus increasing the noise level [39].

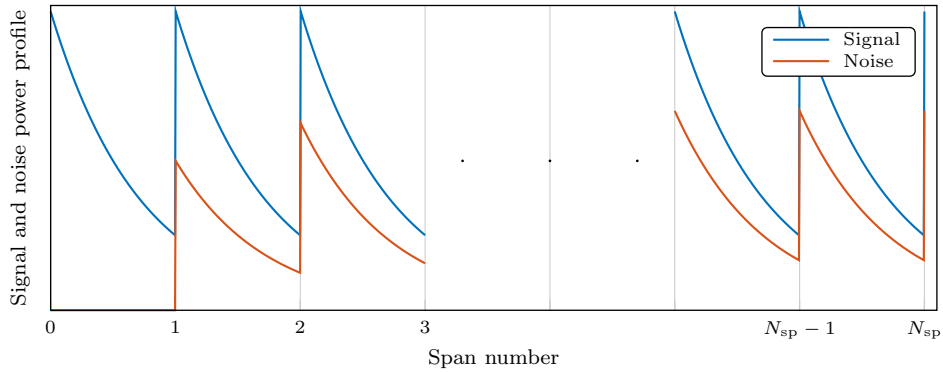


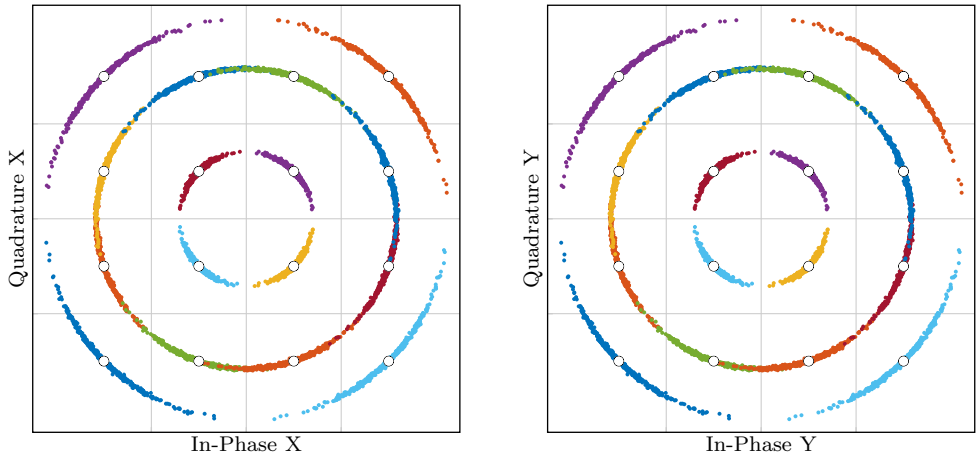
Figure 2.2. The power profile of the signal and noise as a function of the span number is shown.

2.1.2 Carrier-Frequency Offset and Phase Noise

Coherent systems benefit from an improved spectral efficiency by modulating the phase and the amplitude of the optical field, in both polarizations. In order to have access to both phase and amplitude, coherent receivers mix the incoming modulated wave with a continuous-wave optical field, serving as a reference and converting the incoming pass-band field to baseband. The continuous wave at the receiver can be obtained from a separate laser at the receiver, or from the incoming field transmitted as a pilot tone on the orthogonal polarization to the data, or on a separate frequency or spacial channel. Typically, the first option where the continuous wave at the receiver is obtained from a laser at the receiver, often called local oscillator, is preferred because it does not decrease the spectral efficiency.

The frequencies of the transmitter laser and local oscillator can be chosen to be the same or different resulting in two different detection techniques. They are known as homodyne and heterodyne receivers, respectively [40]. In the case of heterodyne detection, the incoming optical signal is first down-converted from the carrier frequency (191–195 THz) to an intermediate frequency that is larger than the bandwidth of the signal and then processed digitally. On the other hand, in the case of homodyne detection, the intermediate frequency is (ideally) zero. The scenario when the reference field is obtained from a pilot is classified as a homodyne technique and is referred to as self-homodyne detection [41].

The local oscillator and the transmitter laser are not synchronized and have frequency and phase fluctuations, resulting in a frequency offset between the two and phase noise, thus creating the need for carrier-frequency offset and carrier-phase noise synchronization. The homodyne architecture requires a demanding optical loop synchronization circuit to phase lock the local oscillator. On the other hand, heterodyne detection does not require the demanding optical phase-locked loop, but has the disadvantage of a 3 dB lower



(a) **(b)**

Figure 2.3. PM-16-QAM constellation affected by phase noise. The different colors represent the 16 points of the constellation in one polarization, whereas the filled dots are the transmitted constellation points. Both X (a) and Y (b) polarizations are affected by the same rotation. The grid lines represent the Voronoi regions of a detector designed for additive white Gaussian noise (AWGN), which, in this case, will lead to catastrophic errors.

sensitivity compared to homodyne detection and requires a doubled processing bandwidth [40]. Similarly to heterodyne detection, intradyne detection [42] is an architecture that down-converts the incoming optical signal to an intermediate frequency, but opposed to heterodyne detection, the intermediate frequency is smaller than the symbol rate with typical values in the range 0–5 GHz, keeping the processing bandwidth low. Compared to homodyne, intradyne detection has the advantage of not requiring an optical phase-locked loop and compensates for the frequency offset and phase noise digitally. On the other hand, self-homodyne detection significantly reduces the phase- and frequency-tracking requirements, since the local oscillator and the transmitter laser are the same and, therefore synchronized. However, the relative phase of the two changes during propagation, but a much lower rate compared to intradyne detection.

Ideally, the spectral shape of a laser is a delta function at the carrier frequency. However, practical lasers have a broader spectrum, which is the source of the phase noise [43]. The phase noise of an optical link can be quantified by the linewidth sum $\Delta\nu$ of the transmitter and receiver lasers in relation to the symbol period T. The laser linewidth relates to the inverse of the coherence time of a laser, which is a measure to characterize the time duration over which the laser frequency is maintained. The coherence time is desired to be much larger than the symbol period, and often the phase-noise sensitivity of a system

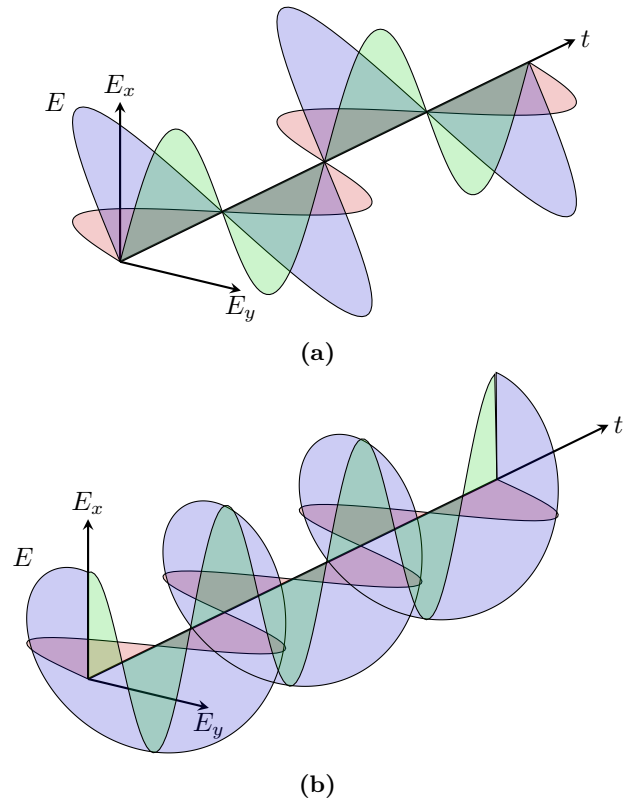


Figure 2.4. The E_x (green) and E_y (red) electric-field components of a plane electromagnetic wave as a function of time t, which combine and form the aggregate field E (blue). In the figure on top, linear polarization is shown where the field E oscillates in the $-45°$ plane, whereas the bottom figure shows circular polarization (obtained by shifting E_x by $\pi/2$) where the field E describes a circle as time progresses.

is quantified by the dimensionless measure $\Delta\nu T$.

Phase noise leads to a random rotation of the received constellation. Since both polarizations are modulated/demodulated based on the same transmitter/receiver lasers, PM constellations are rotated with a common angle in both polarizations. This random rotation induces catastrophic errors in phase-modulated transmission if it is not compensated for. Fig. 2.3 shows the effect of phase noise on a PM-16-QAM constellation. Phase noise is further discussed in Chapters 3 and 4, where modeling and different compensation techniques are reviewed.

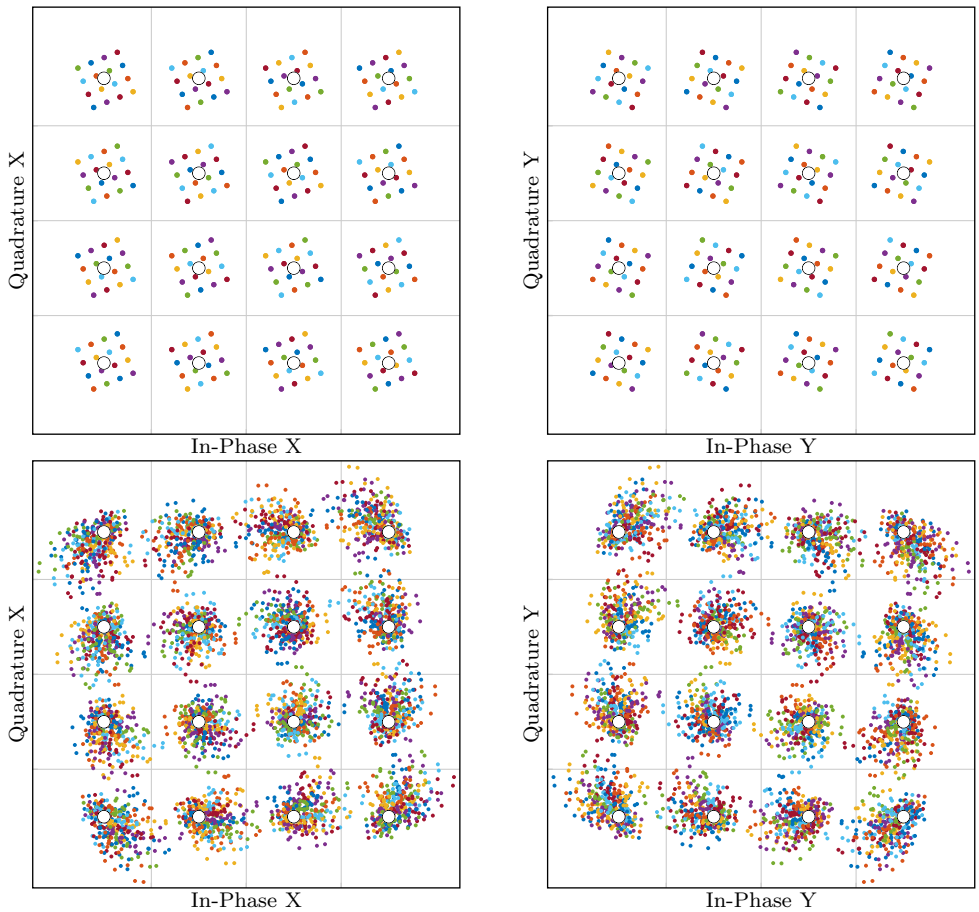


Figure 2.5. PM-16-QAM constellation affected by SOP drift. PM-16-QAM consists of 256 points obtained by all possible combinations between the X and Y polarization. The different colors represent the 16 possible combinations of a point with the complement polarization, whereas the filled dots are the transmitted constellation points. The top row illustrates the effect of a static SOP rotation, where each point is split into 16 points. In the bottom row, the effects of a dynamic SOP drift are shown, where different points of the same color represent different time instances. The grid lines represent the Voronoi regions of a detector designed for AWGN, which, in this case, will lead to catastrophic errors.

2.1.3 Polarization State Drift

Light is an oscillating electromagnetic field, whose propagation can be described, similarly to radio waves, via a sinusoidal wave in the electric field. The orientation of the oscillation of this field is called state of polarization (SOP). The SOP can take on different types. Fig. 2.4(a) shows an example of a linearly polarized lightwave, whose electromagnetic wave E is confined to oscillate only in the $-45°$ plane. The E field can be decomposed into two orthogonal components, commonly referred as E_x and E_y. In the example shown in Fig. 2.4(a), the E_x and E_y components are sinusoidals of the same amplitude and phase. Combinations of E_x and E_y that do not have the same phase or/and amplitude will result in a different behavior of the total field E. In general, the field E will not oscillate in a single plane, but will rotate at the optical frequency. Fig. 2.4(b) illustrates the case called circular polarization where E_x and E_y have the same amplitude but a $\pi/2$ phase difference; thus, the wave E will describe a circle as time progresses. The SOP of light can be fully described by three DOFs corresponding to the amplitudes of the E_x and E_y components and the relative phase difference between them.

Coherent fiber-optic communications make use of the orthogonality between the E_x and E_y components, so called the X and Y polarizations, which can be independently modulated in phase and amplitude. Ideally, light propagating through a straight fiber will preserve the SOP in which it was launched. However, this is not the case for real fibers, which have asymmetrical cores due to the manufacturing process and due to the applied mechanical/thermal stress during and after installation. These imperfections break the orthogonality between E_x and E_y, resulting in a power exchange between the four components of the electric field, a phenomenon called *SOP drift*. The geometry of fibers varies both along the fiber and in time due to the diversity of the environment passed by the cables and temperature/mechanical changes; thus, the drift of the SOP is considered to be *random*.

The SOP drift can be seen as a four-dimensional (4D) rotation of the 4D PM constellation, and it is illustrated in Fig. 2.5. In order to successfully convey data through such a channel, a synchronization block is required to undo these effects. The fact that the drift is stochastic restricts the compensation to be done at the receiver since the behavior of the phenomenon depends on the fiber and cannot be predicted. Different approaches on how to mathematically model SOP drifts are discussed in Chapter 3, whereas compensation techniques are examined in Chapter 4.

2.1.4 Chromatic Dispersion

The optical fiber is a dispersive medium where the group velocity is different for different frequency components; a phenomenon that leads to group velocity dispersion or *chromatic dispersion* (CD). Communication systems are affected by CD since the waveform generated by the transmitter and launched into the fiber spreads over different frequencies/wavelengths. These different frequency components travel through the optical fiber

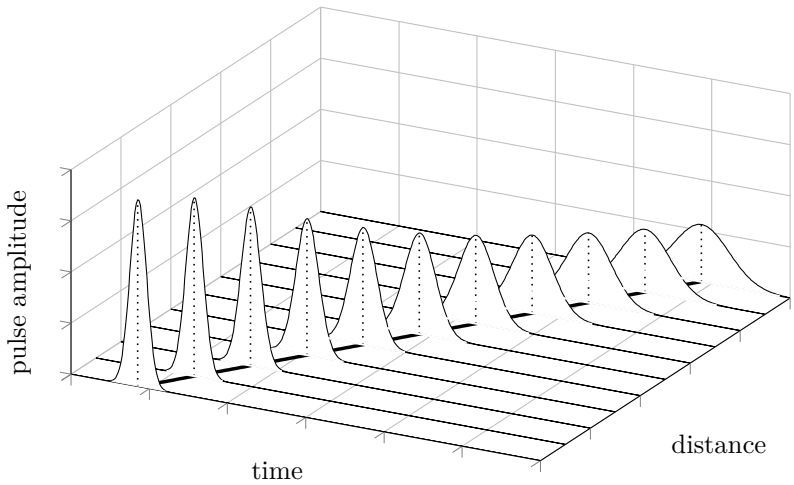


Figure 2.6. The evolution in the time domain of a Gaussian pulse affected by CD. The pulse broadens as it propagates, thus causing interference to the neighboring pulses. The time axis represents the absolute time; therefore, as the pulse propagates, the center of the propagating pulse (indicated by the dotted, vertical lines) shifts according to the oblique thick line.

at different speeds causing pulse broadening in the time domain [44, Sec. 1.2.3]. In other words, different spectral components launched into the fiber at the same time have different transit times and arrive at the receiver at different times. This can be seen as an all-pass filter that causes a frequency-dependent phase shift in the frequency domain without changing the amplitude of the spectrum.

The time of arrival difference between the fast and slow spectral components, i.e., the pulse broadening, depends on the length of the link and on the dispersion parameter, which is specific for each fiber. In principle, CD is a deterministic effect and can be fully compensated for if the length of the transmission and the dispersion parameter are known. Traditionally, CD has been compensated for in the optical domain by inserting a dispersion-compensating fiber [44, Sec. 1.2.3] or a fiber Bragg grating [45] in each span, compensating for the accumulated dispersion in that span. Recently, digital compensation attracted attention in the form of a DSP block at the receiver/transmitter that compensates/precompensates for the entire link at once [46].

Fig. 2.6 illustrates the evolution in the time domain of a pulse affected by CD. It should be noted that both polarizations are affected identically by CD regardless of the SOP. Although the figure is plotted versus absolute time, in general, pulses and signals propagating through optical fibers are analyzed relative to the retarded time, which is the time with refernece at the center of the propagating signal indicated by the dotted, vertical lines in Fig. 2.6.

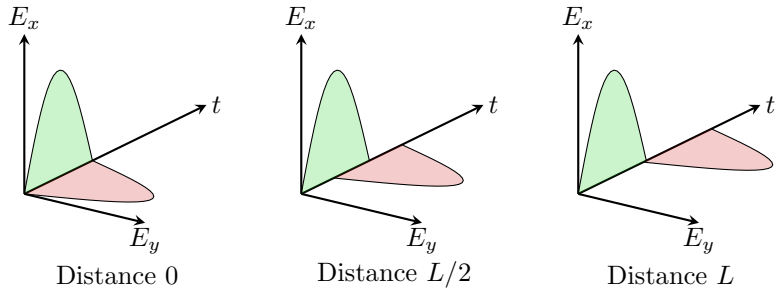


Figure 2.7. Example of (first-order) PMD effects in the time domain, where two pulses with equal power in the X and Y polarizations propagate through a short fiber of length L. The pulses are shown at three different propagated distances and the time separation between them increases with the distance.

2.1.5 Polarization-Mode Dispersion

Practical fibers exhibit imperfections in their shape of the core along the fiber due to the manufacturing process and due to the applied mechanical/thermal stress during and after installation. These imperfections break the cylindrical symmetry of the fiber, leading to a phenomenon called *birefringence*. Birefringence causes dispersion, where the two orthogonal polarizations, which normally travel at the same speed, travel at different speed. This phenomenon is called PMD [47]. Fig. 2.7 shows the effect of (first-order) PMD over a short fiber in the time domain on two pulses launched with equal power in the two orthogonal polarizations. During propagation, the two pulses have different velocities acquiring a time separation between them, which is called *differential group delay* (DGD). For simplicity, in this example the delay occurred between the X and Y axes, which is not always the case. The delay can occur between any pair of orthogonal axes, called *birefringence axes*.

Similarly to CD, this effect is deterministic for short fibers and the DGD grows linearly with the fiber length. However, this is not the case for long fibers where PMD is no longer additive. The birefringence axes have a random orientation along the fiber and their orientation changes randomly with time. The acquired PMD during transmission highly depends on the installation details of the fiber, such as bends, thermal variations, etc.; therefore PMD behaves differently on a fiber spool than on a deployed fiber.

2.1.6 Nonlinear Optical Effects

Nonlinear effects occur when the response of a medium is a nonlinear function of the applied electric and magnetic field, and the optical fibers are no exception. Optical fiber nonliniarities can be classified into two types: stimulated scattering (Raman and Brillouin) and intensity-dependent nonlinear effects. Stimulated scattering effects lead to intensity-dependent gain or loss, and both can be regarded as scattering of a photon to

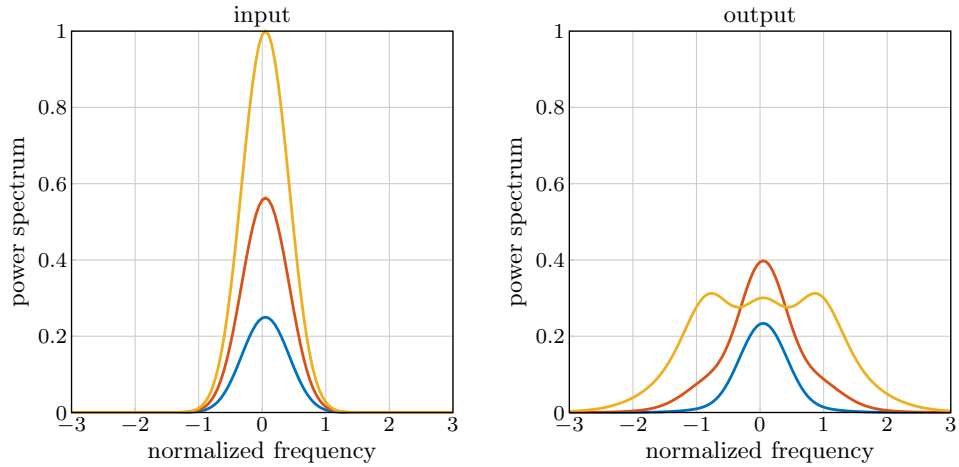


Figure 2.8. The effects of Kerr nonlinearities after propagation in the frequency domain on a Gaussian pulse at different input powers. As the input power is increased (a), the output spectrum broadens (b).

a lower energy photon, i.e., energy transfer occurs from shorter to longer wavelengths, leading to a loss of power at the incident frequency. Although the two scattering phenomena are quite similar, a few things differentiates them. Brillouin scattering [44, Ch. 9] gives rise to a narrow-band (<100 MHz) wave shifted by ∼10 GHz and traveling in the opposite direction. On the other hand, Raman scattering [44, Ch. 8] can occur in both directions and the frequency shift is much larger (∼12 THz). Both these phenomena are detrimental for communication applications, stimulated Raman scattering in particular, but they can also be constructively used in, for example, amplification schemes [38, 48] or fiber lasers [49].

The other nonlinear effect, also known as the Kerr effect, arises due to the nonlinear behavior of the refractive index with respect to the light intensity and produces an intensity-dependent phase shift of the optical field and spectrum broadening. Albeit silica glass is not a highly nonlinear medium, since the section of a fiber core is relatively small, the intensity of the light becomes significant and changes the refractive index. In general, this phenomenon is more detrimental to communication applications than scattering effects and it becomes significant for long fiber lengths. In particular, it is the origin of self-phase modulation [44, Ch. 4], cross-phase modulation [44, Ch. 7], and four-wave mixing effects [44, Ch. 10]. These effects can however be partially compensated for using various digital or optical techniques. The effects of Kerr nonlinearity in the frequency domain can be seen in Fig. 2.8. Modeling and compensation techniques of Kerr nonlinearities are further discussed in Chapters 3 and 4.

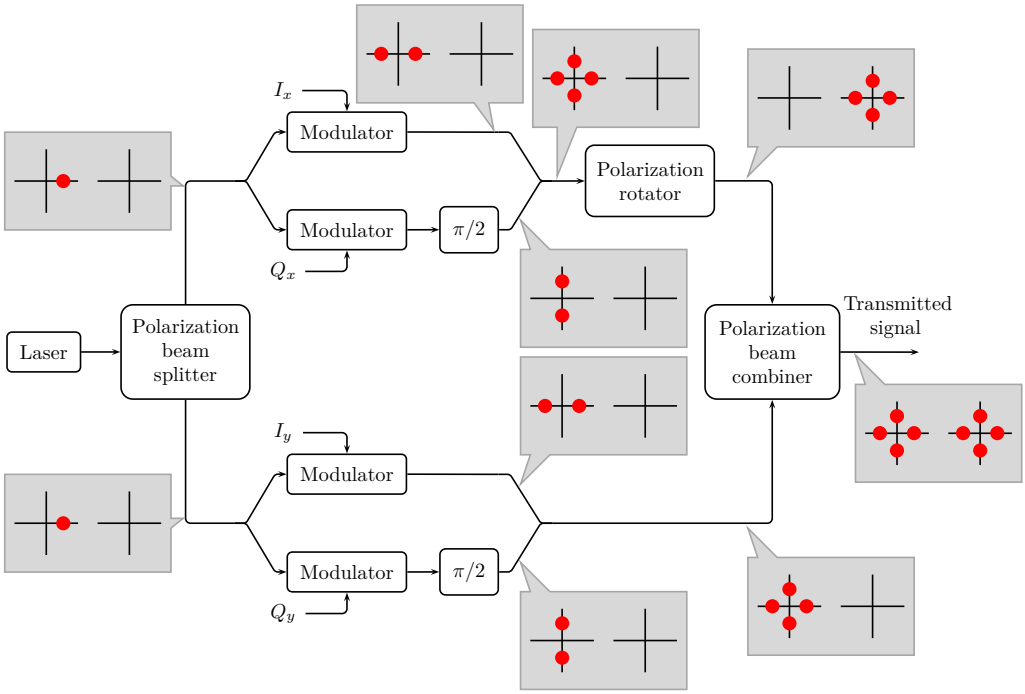


Figure 2.9. Schematic of an optical coherent transmitter [53]. The shown constellations exemplify the generation of PM-QPSK, where the left (right) side corresponds to the X (Y) polarization.

2.2 Coherent Transceivers

Coherent transmission systems are the default solution for optical core networks due to their superior performance offered compared to intensity-modulated direct-detection based alternatives. Nevertheless, direct detection is often used in shorter-reach applications, due to power consumption and cost constraints associated with digital coherent receivers [50–52].

2.2.1 Transmitter

The common solution to modulate the coherent optical signal is based on Mach–Zehnder modulators, which independently modulate each dimension of the optical field. Fig. 2.9 shows an overview schematic of a coherent transmitter [53]. A coherent light source, such as a laser, is split into two arms by a polarization beam splitter. Each arm corresponds to the X and Y polarization. Thereafter, each arm is again split into two arms, corresponding to the in-phase and quadrature components. Each arm is modulated by a Mach–Zehnder modulator, and then recombined after a $\pi/2$ phase shift has been applied

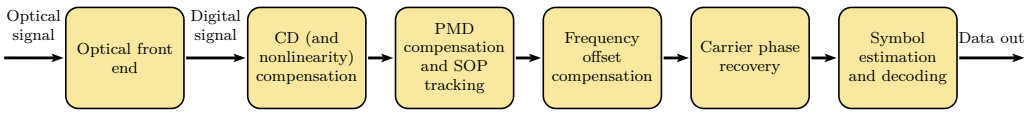


Figure 2.10. Simplified schematic of a coherent receiver.

to one of the arms, thus forming the phase- and amplitude-modulated signal. One of the two such phase- and amplitude-modulated signals is polarization rotated such that it becomes Y-polarized from X-polarized. Finally, the two X- and Y-polarized signals are recombined by a polarization beam combiner and then launched into the fiber.

2.2.2 Receiver

A simplified schematic[2] of a coherent receiver and the flow from the optical signal to data are illustrated in Fig. 2.10 [54]. The shown blocks mitigate the impairments discussed in the earlier sections.

- Optical front end. This block linearly maps the received optical field to electric signals, which are then sampled using analog-to-digital converters to obtain the corresponding digital signals. The sampling rate is typically twice the symbol rate, satisfying the Nyquist-Shannon sampling theorem, but higher sampling rates can however be used. The incoming optical signal is mixed with the local oscillator using two 90° hybrids [55, Sec. 3.1.3], one for each polarization, and four electric signals are output corresponding to the two quadratures in the two polarizations. Phase noise, discussed in Section 2.1.2, as well as a frequency offset, are generated at this stage as a result of the phase and frequency difference between the transmitter laser and the local oscillator.

- CD compensation. This is a static equalization stage usually realized with finite-impulse response filters [54], although infinite-impulse response filters are also an option [56]. Two identical filters are used independently for the two polarizations, typically implemented in the frequency domain due to a more efficient implementation. However, it has been shown that time-domain implementations are more power-efficient for short-haul links [57, 58]. At this stage, matched filtering can be applied by convolving the filters with the desired pulse shape [54] and, for long distances, compensation of nonlinear transmission effects can be employed [24, 59].

- PMD compensation and SOP tracking. This block compensates for time-varying impairments using a bank of finite-impulse response filters in the time domain.

[2] Several blocks, such as the analog-to-digital conversion, de-skew and orthonormalization, are skipped since they are not relevant to this thesis.

These filters have a shorter impulse response than the ones used for CD compensation, and are dynamically updated such that they adapt to the varying channel. The update of the filters is often done in the Jones space using the constant modulus algorithm (CMA) or the multiple modulus algorithm (MMA) (both discussed in Chapter 4) or in the Stokes space [60, 61]. Typically, the sampling rate is down-converted to the symbol rate at the output of this block.

- Frequency offset compensation. The compensation for residual carrier frequency can be done in the time domain using fourth-order methods that take advantage of the four-fold rotational symmetry of PM-QAM constellations [39]. Nevertheless, methods in the frequency domain have been proposed [62], which center the peak of the spectrum around the zero frequency.

- Carrier phase recovery. Similar to frequency estimation, the carrier phase can be estimated using fourth-order methods [63] or methods that scan over a set of possible test phases in order to minimize an error function [64]. Blind algorithms, which do not know the transmitted data, suffer from the four-fold ambiguity of PM-QAM constellations. In order to resolve this issue, and also minimize the impact of cycle slips, differential coding [55, Sec. 2.6.1] is employed, which however induces an increased bit error rate. On the other hand, data-aided algorithms do not require differential coding, but reduce the spectral efficiency by inserting pilots. Phase-noise compensation techniques are examined in Chapter 4.

- Symbol detection and decoding. At this stage, the signal is ready to be decoded into bits. Forward error correction is applied using either soft-decision or hard-decision decoding [65]. The latter is more common due to its weaker complexity at the cost of lower performance, but soft decision has gained more interest recently [66].

The description above considers impairment-compensation only at the receiver. However, DSP can be applied also at the transmitter in order to optimize the overall performance. Moreover, future receivers may combine different DSP blocks in order to improve performance [67]. For example, the algorithm performing joint-polarization phase-noise estimation and symbol detection presented in [68] significantly improves upon the laser linewidth tolerance compared to performing independent phase-noise estimation and symbol detection.

CHAPTER 3

Channel Models

A channel model is an analytical description of the signal propagation through a medium relating the output to the input. This model can include, besides the impairments that arise in the propagation medium, elements of the transmitter and receiver. Channel models can be used in simulations to test and verify transmission schemes and their accompanying DSP. Moreover, improved signal processing tools can be derived from an accurate channel model that account optimally for the corresponding impairment. Channel models can also be used to derive fundamental limits imposed by the transmission medium, such as capacity bounds. In general, simulations offer a greater flexibility than experiments and can be used to predict the behavior of a system before setting up time-consuming experiments.

This chapter describes mathematically the propagation of the electric field in optical fibers and the modeling of various impairments that arise due to propagation effects or nonideal hardware. Section 3.1 describes waveform propagation equations through optical fibers. In Section 3.2, we review how information is mapped from data symbols to waveforms to be transmitted and then describe the reverse operation performed at the receiver, which demaps waveforms to data symbols. We continue by introducing discrete-time channel modeling for the ASE-noise channel. Thereafter, we describe channel modeling of phase noise, SOP drift, and PMD in Sections 3.3 and 3.4. Lastly, approximate channel models for fiber propagation are discussed in Sections 3.5 and 3.6.

The different channel models discussed below are described in two variants: with and without including modulation and demodulation. When modulation and demodulation is included, the underlying channel model characterizes discrete-time symbol propagation (Sections 3.2, 3.3, and 3.6.2), whereas continuous-waveform propagation is modeled

without including modulation and demodulation (Sections 3.1, 3.4, 3.5, and 3.6.1).

3.1 Fiber-Propagation Models

All electromagnetic phenomena, including fiber-optic signal propagation, are modeled by Maxwell's equations. These equations are however troublesome to evaluate, but can be simplified under the following approximations [44, Ch. 2]

- the medium is non-magnetic and has no free charges, which is always true for optical fibers.
- the nonlinearity is small, i.e., can be treated as a perturbation, which is justified since the nonlinear changes in the refractive index are $< 10^{-6}$ in practice.
- the fiber is weakly guiding, i.e., the index difference of the fiber waveguide is small, which it is (approximately 0.1% in single-mode fibers).
- the slowly-varying envelope approximation, i.e., the modulation bandwidth is small compared to the carrier frequency of light, which for communication applications usually holds. One direct implication of this approximation is that only unidirectional propagation is assumed.
- the optical field maintains its polarization along the fiber length. This is not the case, except for polarization-maintaining fibers, but the approximation often works well in practice and it will be relaxed in Sections 3.1.3 and 3.1.4.
- higher-order dispersion is negligible, which holds for most practical cases, but has to be taken into account for high bandwidths (> 1 THz) or when the lowest-order dispersion is small ($\beta_2 \approx 0$, see below). However, third-order dispersion can be relatively easily incorporated when required [44, Sec. 3.3].

Based on these approximations, simplified propagation equations of the optical signal can be derived. First, we will discuss the NLSE in Section 3.1.1, which describes signal propagation in single polarization, and then generalize it to dual polarization in Section 3.1.2. Thereafter, the propagation equations are extended to account for polarization effects in Sections 3.1.3 and 3.1.4. A tutorial by Menyuk and Marks can be found in [69], where these propagation equations are derived starting from Maxwell's equations, and their range of validity is discussed.

3.1.1 The Nonlinear Schrödinger Equation

The NLSE is a partial differential equation defining the input–output relationship for optical baseband[1] signals. Accounting for signal attenuation, CD, and Kerr nonlinear

[1] Due to the slowly-varying approximation above.

effects, it can be written as

$$\frac{\partial E_x(z,t)}{\partial z} = \underbrace{-\frac{\alpha}{2}E_x(z,t)}_{\text{attenuation}} \underbrace{-i\frac{\beta_2}{2}\frac{\partial^2 E_x(z,t)}{\partial t^2}}_{\text{dispersion}} + \underbrace{i\gamma E_x(z,t)|E_x(z,t)|^2}_{\text{nonlinearity}}, \tag{3.1}$$

where $E_x(z,t)$ is a complex-baseband signal denoting the (arbitrarily chosen and without loss of generality) X field component and $i = \sqrt{-1}$ is the imaginary part. The signal $E_x(z,t)$ is a function of distance $0 \leq z \leq L$, where L is the total length of the fiber, and continuous time $t \in \mathbb{R}$. The in-phase and quadrature components are combined as the real and imaginary parts of $E_x(z,t)$. Note that t does not denote the absolute time, but the retarded time, that is, the time relative to the center of the propagating signal (see Fig. 2.6). The fiber is parameterized by three coefficients: γ is the nonlinear coefficient, β_2 is the group velocity dispersion coefficient, and α is the attenuation coefficient.

Despite many attempts, so far no exact analytical solution for the general NLSE has been found. Nonetheless, solutions to (3.1) can be found in special cases, e.g., without losses, it can be solved exactly by applying the inverse scattering theory [70], which recently emerged in the literature as the nonlinear Fourier transform [71]. In the following, we examine the effects of attenuation, dispersion, and nonlinearity.

Signal Attenuation

If we neglect the dispersion and nonlinearity, i.e., by setting $\beta_2 = 0$ and $\gamma = 0$ in (3.1), equation (3.1) can be solved and its solution becomes

$$E_x(z,t) = E_x(0,t)\exp(-\alpha z/2), \tag{3.2}$$

which indicates that the signal experiences an exponential decay due to the fiber loss.

Chromatic Dispersion

The effects of CD can be studied by ignoring the nonlinearity, i.e., setting $\gamma = 0$. In this case, the closed-form solution of (3.1) in the frequency domain is

$$\tilde{E}_x(z,f) = \tilde{E}_x(0,f)\exp(i2\beta_2\pi^2 f^2 z)\exp(-\alpha z/2), \tag{3.3}$$

where $\tilde{E}_x(z,f)$ is the Fourier Transform of $E_x(z,t)$. As can be seen, CD behaves as an all-pass filter without changing the amplitude of the spectrum, but only the phase. The phase has a quadratic frequency dependency and leads to pulse broadening in the time domain (see Fig. 2.6), in turn leading to inter-symbol interference.

Kerr Nonlinearity

In the absence of CD, i.e., $\beta_2 = 0$, the solution of (3.1) can be written as

$$E_x(z,t) = E_x(0,t)\exp(i\gamma|E_x(0,t)|^2 L_{\text{eff}}(z))\exp(-\alpha z/2), \tag{3.4}$$

where

$$L_{\text{eff}}(z) = \int_0^z e^{-\alpha z'}\mathrm{d}z' = \frac{1-\exp(-\alpha z)}{\alpha} \tag{3.5}$$

is called the effective (nonlinear) length, which indicates the nonlinear interaction region and $L_{\text{eff}}(z) \to z$ as $\alpha \to 0$. For a lossless fiber with $\alpha = 0$, (3.4) simplifies to

$$E_x(z,t) = E_x(0,t)\exp(i\gamma|E_x(0,t)|^2 z). \tag{3.6}$$

As it appears from (3.4) and (3.6), the nonlinear effects do not change the signal amplitude in the time domain, but introduce an intensity dependent phase shift. As a result, the spectrum of the signal broadens and the broadening depends on the signal intensity (see Fig. 2.8). Note however that (3.4) is the exact solution of (3.1), including all three effects, if the input field $E_x(0,t)$ is constant.

Amplification Noise

As earlier discussed in Section 2.1, long-haul transmissions engage various optical amplification schemes to compensate for the fiber loss, which in turn introduce noise. The noise is introduced through the process of spontaneous emission, hence the name of ASE noise, and compared to other noise sources, such as thermal noise from electrical components, the ASE noise dominates and the others can be neglected [37, Sec. IX-A].

The NLSE can be extended to account for amplification and noise by inserting a real gain profile function $g(z)$ and a complex-valued stochastic process $n_x(z,t)$ in (3.1) resulting in

$$\frac{\partial E_x(z,t)}{\partial z} = -\frac{\alpha - g(z)}{2}E_x(z,t) - i\frac{\beta_2}{2}\frac{\partial^2 E_x(z,t)}{\partial t^2} + i\gamma E_x(z,t)|E_x(z,t)|^2 + n_x(z,t), \tag{3.7}$$

which is referred to as the stochastic NLSE [72]. The term $n_x(z,t)$ satisfying [73]

$$\mathbb{E}[n_x(z,t)] = 0, \tag{3.8}$$

accounts for the ASE noise in the X polarization. In [74] it was shown that ASE can be accurately modeled as an additive Gaussian random variable. This allows us to model $n_x(z,t)$ as a circularly symmetric complex white Gaussian noise. Such a process is fully characterized by its mean (3.8) and its autocorrelation function $\mathbb{E}[n_x(z,t)n_x^*(z',t')]$, which depends on the amplification scheme. Next we discuss two common amplification schemes and their implications on $g(z)$ and $n_x(z,t)$.

Lumped Amplification Most systems employ lumped amplification, where the total accumulated loss over one span is compensated using an erbium-doped fiber amplifier at the end of each fiber span. In this case, the gain profile becomes $g(z) = \alpha L_{\mathrm{sp}} \sum_{n=1}^{N_{\mathrm{sp}}} \delta(z - nL_{\mathrm{sp}})$, where L_{sp} is the span length and $\delta(\cdot)$ is the Dirac delta function. The autocorrelation function of $n_x(z,t)$ is in this case [73]

$$\mathbb{E}[n_x(z,t)n_x^*(z',t')] = n_{\mathrm{sp}}(G-1)h\nu\delta(t-t')\delta(z-z')\sum_{n=1}^{N_{\mathrm{sp}}}\delta(z-nL_{\mathrm{sp}}), \tag{3.9}$$

where $G = \exp(\alpha L_{\mathrm{sp}})$ is the gain of the amplifier and here we assume that it compensates for the exact span loss, $h \approx 6.626{\cdot}10^{-34}$ Js is Planck's constant, $\nu \approx 193$ THz is the optical carrier frequency, and $n_{\mathrm{sp}} = \frac{GF_n-1}{2(G-1)}$ is the spontaneous emission factor [35, Eq. (7.2.15)], whereas F_n is the noise figure of the amplifier. Typical values for F_n range from 3 dB (fundamental lower bound) to 6 dB. In (3.9), the first two delta functions ensure that the noise realizations are independent with respect to time and distance, whereas the summation accounts for all N_{sp} amplifiers of the link and the delta function within ensures that the noise is lumped, that is, occurs at the end of each span.

Distributed Amplification In contrast to lumped amplification, distributed amplification continuously compensates for the fiber loss using the fiber itself by exploiting the nonlinear phenomenon of stimulated Raman scattering, hence the name of Raman amplification. Here we consider an ideal Raman amplification, where the signal power level is considered to remain equal to the launch power throughout the entire fiber[2]. In this case, the gain profile becomes $g(z) = \alpha$ and the noise autocorrelation is [37]

$$\mathbb{E}[n_x(z,t)n_x^*(z',t')] = \alpha K_{\mathrm{T}} h\nu\delta(t-t')\delta(z-z'), \tag{3.10}$$

where K_{T} is the photon occupancy factor and it is approximately 1.13 for Raman amplification of fiber-optic communication systems at room temperature.

It should be noted that, according to (3.9) or (3.10), the noise term $n_x(z,t)$ is white yielding an $E_x(z,t)$ with infinite power. However, this is not the case in practice where $n_x(z,t)$ is bandlimited since the amplifiers have a finite bandwidth and are sometimes followed by bandpass filters to limit the noise bandwidth. This discrepancy is typically solved in numerical and analytical analyses by considering an ideal square filter after each amplifier in the case of lumped amplification, or just before the receiver for distributed amplification. The bandwidth of the filter can vary from to the bandwidth of the signal up to higher values accounting for spectral broadening.

[2]However, this is not the case in practice, where the power level decays with distance and may increase towards the end of the span in the case of a bidirectional pumping scheme.

3.1.2 The Coupled-Mode Nonlinear Schrödinger Equations

The NLSE and stochastic NLSE above can be used to model single-polarization propagation through optical fibers. In the case of PM transmission, the propagation of each field component is governed by the coupled-mode NLSEs [75]

$$\begin{aligned}\frac{\partial E_x(z,t)}{\partial z} &= -\frac{\alpha - g(z)}{2}E_x(z,t) - i\frac{\beta_2}{2}\frac{\partial^2 E_x(z,t)}{\partial t^2} \\ &\quad + i\gamma E_x(z,t)\Big(|E_x(z,t)|^2 + \frac{2}{3}|E_y(z,t)|^2\Big) + i\gamma\frac{1}{3}E_y(z,t)^2 E_x^*(z,t) + n_x(z,t),\end{aligned} \tag{3.11}$$

$$\begin{aligned}\frac{\partial E_y(z,t)}{\partial z} &= -\frac{\alpha - g(z)}{2}E_y(z,t) - i\frac{\beta_2}{2}\frac{\partial^2 E_y(z,t)}{\partial t^2} \\ &\quad + i\gamma E_y(z,t)\Big(|E_y(z,t)|^2 + \frac{2}{3}|E_x(z,t)|^2\Big) + i\gamma\frac{1}{3}E_x(z,t)^2 E_y^*(z,t) + n_y(z,t),\end{aligned} \tag{3.12}$$

where $E_y(z,t)$ is the complex-baseband signal denoting the Y field component and $n_y(z,t)$ accounts for the ASE noise in the Y polarization and is independent of and identically distributed as $n_x(z,t)$. We notice that the two polarizations experience different nonlinear phase shifts. Equations (3.11)–(3.12) can be compactly rewritten in a vector form as [76]

$$\begin{aligned}\frac{\partial \mathbf{E}(z,t)}{\partial z} &= -\frac{\alpha - g(z)}{2}\mathbf{E}(z,t) - i\frac{\beta_2}{2}\frac{\partial^2 \mathbf{E}(z,t)}{\partial t^2} \\ &\quad + i\gamma\Big(\mathbf{E}(z,t)\|\mathbf{E}(z,t)\|^2 - \frac{1}{3}\big(\mathbf{E}(z,t)^{\mathrm{H}}\boldsymbol{\sigma}_3\mathbf{E}(z,t)\big)\boldsymbol{\sigma}_3\mathbf{E}(z,t)\Big) + \mathbf{n}(z,t),\end{aligned} \tag{3.13}$$

where $\mathbf{E}(z,t) = [E_x(z,t), E_y(z,t)]^{\mathrm{T}}$ is the so-called *Jones vector*, $\boldsymbol{\sigma}_3$ is one Pauli spin matrix defined later in (3.30), and $\mathbf{n}(z,t) = [n_x(z,t), n_y(z,t)]^{\mathrm{T}}$.

3.1.3 The Manakov Equation

The propagation equations above are valid for circularly symmetric ideal optical fibers that do not experience any birefringence. Practical fibers are not ideal and have nonnegligible birefringence that changes along the link, leading to PMD. This effect, however, changes the nonlinear response of the fiber and has to be incorporated in (3.11)–(3.13). By averaging over the rapidly and randomly changing polarization state due to PMD, the Manakov equation is obtained [77, 78]

$$\frac{\partial \mathbf{E}(z,t)}{\partial z} = -\frac{\alpha - g(z)}{2}\mathbf{E}(z,t) - i\frac{\beta_2}{2}\frac{\partial^2 \mathbf{E}(z,t)}{\partial t^2} + i\gamma\frac{8}{9}\mathbf{E}(z,t)\|\mathbf{E}(z,t)\|^2 + \mathbf{n}(z,t). \tag{3.14}$$

It can be noted that averaging over the polarization state has two consequences: i) the nonlinear coefficient is multiplied with the 8/9 factor, and ii) opposed to the coupled-

mode NLSEs, both polarizations experience the same nonlinear phase shift equal to the energy of $\mathbf{E}(z,t)$.

3.1.4 The Manakov–PMD Equation

The propagation equations presented earlier do not account for PMD. However, some scenarios require the inclusion of PMD explicitly. For example, as we will see in Papers D–G, nonlinear compensation techniques are very sensitive to PMD and require PMD to be incorporated in the propagation equation in order to acquire meaningful results. The Manakov equation (3.14) is obtained by averaging over the polarization rotations, which are assumed to be fast, but PMD is not taken into account. PMD is taken into account in the Manakov–PMD equation [76, 79] given by [69, Eq. (68)]

$$\begin{aligned}\frac{\partial \mathbf{E}(z,t)}{\partial z} =& -\frac{\alpha - g(z)}{2}\mathbf{E}(z,t) + \Delta\beta_1 \mathbf{\Sigma}(z)\frac{\partial \mathbf{E}(z,t)}{\partial t} - i\frac{\beta_2}{2}\frac{\partial^2 \mathbf{E}(z,t)}{\partial t^2} \\ &+ i\gamma\frac{8}{9}\mathbf{E}(z,t)\|\mathbf{E}(z,t)\|^2 + \mathbf{n}(z,t),\end{aligned} \tag{3.15}$$

where, in addition to (3.14), we have the second term on the right-hand side accounting for PMD. Note that this is the simplified version of the Manakov–PMD equation [69, Eq. (63)] without including the nonlinear PMD term since it is negligible for fiber-optic communication applications [69,76]. The coefficient $\Delta\beta_1 = (\beta_{1x} - \beta_{1y})/2$ is the group delay[3] between the two polarizations with propagation constants β_{1x} and β_{1y}, respectively, and it can be related to the PMD coefficient of the fiber D_{PMD} as $\Delta\beta_1 = \frac{D_{\text{PMD}}}{2\sqrt{2L_c}}$ [76, Eq. (28)], where L_c is the fiber correlation length. The fiber correlation length L_c is defined as the length over which two polarization components remain correlated; typical values range from 10 to 100 m. The PMD coefficient D_{PMD} has typical values of $0.04 - 0.1\,\text{ps}/\sqrt{\text{km}}$ for modern fibers and can go up to several $\text{ps}/\sqrt{\text{km}}$ for old fibers. The matrix $\mathbf{\Sigma}(z)$ describes the linear evolution of PMD along the fiber length. Similarly to CD, the effects of PMD can be studied in the frequency domain by ignoring the other impairments. The solution to (3.15) if $\alpha = 0$, $\beta_2 = 0$, and $\gamma = 0$ becomes

$$\tilde{\mathbf{E}}(z,f) = \mathbf{T}(z,f)\tilde{\mathbf{E}}(0,f), \tag{3.16}$$

where $\tilde{\mathbf{E}}(z,f)$ is the Fourier Transform of $\mathbf{E}(z,f)$ and $\mathbf{T}(z,f) \in \mathbb{C}^{2\times 2}$ is a unitary matrix modeling PMD and satisfies [76, Eq. (22)]

$$\frac{\partial \mathbf{T}(z,f)}{\partial z} = i\Delta\beta_1 f \mathbf{\Sigma}(z)\mathbf{T}(z,f), \tag{3.17}$$

$$\mathbf{T}(0,f) = \mathbf{I}_2, \tag{3.18}$$

[3] We assume that $\Delta\beta_1$ does not change along the fiber, which may not be the case in practical fibers, but it is a good approximation.

where $\mathbf{I}_n$ is the $n \times n$ identity matrix.

The matrix $\mathbf{T}(z, f)$ is typically modeled using the coarse-step method where the continuous variation of the birefringence along the fiber is replaced by a concatenation of equal-length sections with a given mean birefringence and random polarization coupling. The axis of the polarization coupling varies from section to section. This method is further discussed in Section 3.4. The discussion above and in Section 3.4 considers a static PMD, which however, is not the case in practical fibers. PMD does evolve with time due to temperature variations and mechanical vibrations, changing the local birefringence of the fiber and leading to stochastic temporal PMD fluctuations. In Paper B we propose a model for the temporal stochastic evolution of PMD and absolute polarization state.

3.2 Linear Modulation, Matched Filtering, and Sampling

The transmitted signal $\mathbf{E}(0, t)$ into the transmission medium is obtained by linearly modulating the information symbols $\mathbf{u}_k \in \mathbb{C}^2$ as

$$\mathbf{E}(0, t) = \sum_k \mathbf{u}_k p(t - kT) \tag{3.19}$$

using a real-valued pulse shape $p(t)$, where T is the symbol (baud) interval and $k \in \mathbb{Z}$ is the discrete-time index. The pulse shape is chosen such that its time shifts form an orthonormal basis

$$\int_{-\infty}^{\infty} p(t - kT)p(t - k'T)\mathrm{d}t = \begin{cases} 1, & k = k' \\ 0, & \text{otherwise} \end{cases}, \tag{3.20}$$

for any $k, k' \in \mathbb{Z}$. As can be seen, we consider the transmitted waveform $\mathbf{E}(0, t)$ to consist of a single (wavelength) channel only. However, the expression above can be straightforwardly modified to account for multiple channels.

The discrete transmitted symbols $\mathbf{u}_k$ are drawn from a finite constellation $\mathscr{C} = \{\mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_M\}$. The average energy per symbol is the average of $\|\mathbf{u}_k\|^2$ and, in the case of equiprobable and independent symbols, it is equal to

$$E_s = \frac{1}{M} \sum_{k=1}^{M} \|\mathbf{c}_k\|^2. \tag{3.21}$$

The received discrete symbols at distance L are obtained from the received electric field $\mathbf{E}(L, t)$ as

$$\mathbf{r}_k = \int_{-\infty}^{\infty} \mathbf{E}(L, t)p(t - kT)\mathrm{d}t, \tag{3.22}$$

by matched filtering and sampling.

Having established the modulation and demodulation operations, we can now proceed

by establishing discrete-time channel models that relate output data (constellation) symbols to input data symbols. We start with a discrete-time channel model for ASE noise only, which we further extend to phase noise and SOP drift in Section 3.3.

In the absence of all impairments but ASE noise, the transmission can be accurately modeled as an AWGN channel. The received symbols after demodulation can be related to the transmitted symbols by

$$\mathbf{r}_k = \mathbf{u}_k + \mathbf{n}_k. \tag{3.23}$$

The term $\mathbf{n}_k = [n_{x,k}, n_{y,k}]^{\mathrm{T}} \in \mathbb{C}^2$ with

$$n_{x,k} = \int_{-\infty}^{\infty} \left(\int_0^L n_x(z,t)\mathrm{d}z \right) p(t-kT)\mathrm{d}t, \tag{3.24}$$

and similarly for $n_{y,k}$, models the additive ASE noise accumulated over the link. It can be represented by two independent complex circular zero-mean Gaussian random variables with variance $N_0/2$ per real dimension, i.e., $\mathbb{E}[\mathbf{n}_k \mathbf{n}_k^{\mathrm{H}}] = N_0 \mathbf{I}_2$ [11], where

$$N_0 = \begin{cases} n_{\mathrm{sp}}(G-1)h\nu N_{\mathrm{sp}}, & \text{for lumped amplification} \\ \alpha K_{\mathrm{T}} h\nu L, & \text{for (ideal) distributed amplification} \end{cases}. \tag{3.25}$$

Note that the variance N_0 is not scaled with the bandwidth due to the normalization in (3.20). However, in general, by passing the accumulated ASE noise through a unit-gain filter $h(t)$ with bandwidth $\mathcal{W}$, i.e., $n_x^{\mathcal{W}} = \int_{-\infty}^{\infty} (\int_0^L n_x(z,t)\mathrm{d}z) h(t)\mathrm{d}t$, its variance becomes $\mathbb{E}[n_x^{\mathcal{W}} (n_x^{\mathcal{W}})^{\mathrm{H}}] = N_0 \mathcal{W}$.

3.3 Phase Noise and Polarization State Drift

In this section, we discuss channel modeling of phase noise and rotations of the SOP. We chose to associate these two impairments since they can be conveniently modeled as 4D rotations of the field/symbols. Moreover, this material serves as an introduction to the second part of the thesis containing the appended papers, where the Papers A and C are based on this joint formulation. Nonetheless, this is not a requirement and, as we will see later, the two impairments can be easily decoupled. For completeness, ASE noise from (3.23) is considered too.

Note however that, as already mentioned in Chapter 2, these two impairments emerge due to different phenomena and arise at different parts of the communication system. Phase noise is added before and after fiber propagation at the transmitter and receiver, respectively, due to nonideal fiber lasers. On the other hand, the SOP experiences a random drift during fiber propagation. Nonetheless, under the assumption of linear propagation supposed in this section, the two impairments can be modeled jointly.

So far we considered the Jones formalism, which relies on 2×1 complex Jones vectors to represent the electric field and describe fiber propagation. In the following we introduce

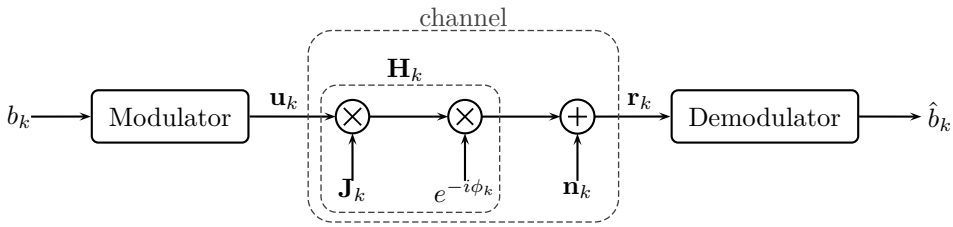


Figure 3.1. Block diagram of a system considering a transmitter, channel, and a receiver. The channel consists of phase noise, SOP drift, and AWGN.

two complementary formalisms, namely, the 4D real and Stokes descriptions. As we will see, these offer a richer or restricted signal space, respectively, which can be useful in some situations. It should be noted that a big part of this section can be found in [80], where the 4D real formalism was properly defined. In what comes next, we will introduce the mathematical description of the these two phenomena first, and then continue with channel modeling.

3.3.1 Mathematical Representation

Jones Description

The propagation of the symbols (including modulation and demodulation) in the presence of phase noise, SOP drift, and ASE noise can be described as

$$\mathbf{r}_k = \mathbf{H}_k \mathbf{u}_k + \mathbf{n}_k, \tag{3.26}$$

where $\mathbf{H}_k \in \mathbb{C}^{2\times 2}$ is a unitary matrix, the so-called *Jones matrix*, which preserves the input power during propagation and models the phase-noise and SOP-drift effects. Fig. 3.1 shows a block diagram of the model described above. This model is a 2×2 multiple-input multiple-output system, where the inputs and outputs are the two different polarizations of light. Such multiple-input multiple-output schemes are widely deployed in communication over radio frequencies, where multiple inputs and outputs are achieved by using several different antennas [81].

The matrix $\mathbf{H}_k$ belongs to the unitary group of degree two, denoted by $U(2)$, such that

$$\mathbf{H}_k^{\mathrm{H}} \mathbf{H}_k = \mathbf{H}_k \mathbf{H}_k^{\mathrm{H}} = \mathbf{I}_2, \tag{3.27}$$

$$|\det \mathbf{H}_k| = 1. \tag{3.28}$$

In general, complex 2×2 matrices have eight DOFs, i.e., the real and imaginary parts of the four elements, whereas the matrices in $U(2)$, after applying (3.27), have only *four* DOFs[4]. Such matrices can be expressed by the matrix function $\mathbf{H}_k = H(\phi_k, \boldsymbol{\alpha}_k)$ using

[4]Note that the constraint (3.28) is already covered by (3.27); therefore it does not reduce the DOFs.

the *matrix exponential* parameterized by four variables: ϕ_k modeling the phase noise and $\boldsymbol{\alpha}_k = (\alpha_{1,k}, \alpha_{2,k}, \alpha_{3,k})$ modeling the SOP drift, according to [82, 83]

$$H(\phi, \boldsymbol{\alpha}) = \exp(-i(\boldsymbol{\alpha} \cdot \vec{\boldsymbol{\sigma}} + \phi \mathbf{I}_2)), \tag{3.29}$$

for any $\phi \in \mathbb{R}$ and $\boldsymbol{\alpha} \in \mathbb{R}^3$, where $\vec{\boldsymbol{\sigma}} = (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3)$ is a tensor of the Pauli spin matrices [83]

$$\boldsymbol{\sigma}_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \boldsymbol{\sigma}_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \boldsymbol{\sigma}_3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{3.30}$$

This notation of the Pauli spin matrices $\boldsymbol{\sigma}_i$ complies with the definition of the Stokes vector (3.58), but it is different from the notation introduced by Frigo [82] and used in [76].

The Pauli matrices are commonly used in the polarization literature and in addition to the three matrices stated above, the identity matrix is often denoted by $\boldsymbol{\sigma}_0$. These matrices are Hermitian and have zero trace, except $\boldsymbol{\sigma}_0$. They also satisfy the identities

$$\boldsymbol{\sigma}_0 = \boldsymbol{\sigma}_i^2 \tag{3.31}$$

for $i = 1, 2, 3$, and

$$\boldsymbol{\sigma}_1 \boldsymbol{\sigma}_2 = -\boldsymbol{\sigma}_2 \boldsymbol{\sigma}_1 = i\boldsymbol{\sigma}_3. \tag{3.32}$$

The Pauli matrices are linearly independent; therefore any complex 2×2 matrix can be written as

$$\mathbf{C} = \sum_{i=0}^{3} c_i \boldsymbol{\sigma}_i, \tag{3.33}$$

where $c_i \in \mathbb{C}$ and can be calculated as

$$c_i = \frac{\mathrm{Tr}(\mathbf{C}\boldsymbol{\sigma}_i)}{2}. \tag{3.34}$$

The matrix exponential in (3.29) is defined as [84, p. 165]

$$\exp \mathbf{A} = \sum_{n=0}^{\infty} \frac{1}{n!} \mathbf{A}^n, \tag{3.35}$$

which converges for any $\mathbf{A} \in \mathbb{C}^{m \times m}$.

The two phenomena modeled by $\mathbf{H}_k$ can be separated by factoring[5] out the phase

[5] This factorization is possible only for $\phi \mathbf{I}_2$ since $\exp(\mathbf{X} + \mathbf{Y}) = \exp \mathbf{X} \exp \mathbf{Y} = \exp \mathbf{Y} \exp \mathbf{X}$ holds only if $\mathbf{XY} = \mathbf{YX}$ for any $\mathbf{X}, \mathbf{Y} \in \mathbb{C}^{n \times n}$.

noise ϕ_k such that

$$\mathbf{H}_k = e^{-i\phi_k}\mathbf{J}_k, \tag{3.36}$$

where $\mathbf{J}_k$ models strictly the drift of the SOP. In this case, (3.26) can be rewritten as

$$\mathbf{r}_k = e^{-i\phi_k}\mathbf{J}_k\mathbf{u}_k + \mathbf{n}_k. \tag{3.37}$$

The modeling of ϕ_k and $\mathbf{J}_k$ (or $\mathbf{H}_k$) and their temporal evolution is discussed in Section 3.3.2.

The matrix $\mathbf{J}_k$ is a unitary matrix that belongs to the special unitary group of degree two, denoted by $SU(2)$, and can be described by the matrix function

$$J(\boldsymbol{\alpha}) = \exp(-i\boldsymbol{\alpha}\cdot\vec{\boldsymbol{\sigma}}) \tag{3.38}$$

as $\mathbf{J}_k = J(\boldsymbol{\alpha}_k)$. This group of matrices is a subgroup of $U(2)$ and satisfies

$$\det \mathbf{J}_k = 1, \tag{3.39}$$

opposed to (3.28). The constraint (3.39) reduces the DOFs of such matrices to only *three*, i.e., the elements of $\boldsymbol{\alpha}$.

The functions $H(\phi,\boldsymbol{\alpha})$ and $J(\boldsymbol{\alpha})$ can be rewritten after expanding (3.29) and (3.38) into the Taylor series (3.35) and using $(\boldsymbol{\alpha}\cdot\vec{\boldsymbol{\sigma}})^2 = \theta^2\mathbf{I}_2$ as

$$H(\phi,\boldsymbol{\alpha}) = e^{-i\phi}(\mathbf{I}_2\cos\theta - i\mathbf{a}\cdot\vec{\boldsymbol{\sigma}}\sin\theta), \tag{3.40}$$

$$J(\boldsymbol{\alpha}) = \mathbf{I}_2\cos\theta - i\mathbf{a}\cdot\vec{\boldsymbol{\sigma}}\sin\theta, \tag{3.41}$$

where

$$\boldsymbol{\alpha} = \theta\mathbf{a}. \tag{3.42}$$

In (3.42), $\boldsymbol{\alpha}$ is represented as the product of its length $\theta = \|\boldsymbol{\alpha}\|$ in the interval $[0,\pi)$ and the unit vector $\mathbf{a} = (a_1, a_2, a_3)$, which represents its direction on the unit sphere.

Since $H(\phi,\boldsymbol{\alpha})$ and $J(\boldsymbol{\alpha})$ are unitary, their inverses can be found by the conjugate transpose operation or by negating ϕ and $\boldsymbol{\alpha}$, since

$$H(\phi,\boldsymbol{\alpha})^{-1} = H(\phi,\boldsymbol{\alpha})^{\mathrm{H}} = H(-\phi,-\boldsymbol{\alpha}), \tag{3.43}$$

$$J(\boldsymbol{\alpha})^{-1} = J(\boldsymbol{\alpha})^{\mathrm{H}} = J(-\boldsymbol{\alpha}). \tag{3.44}$$

4D Real Description

In the 4D formalism, the electric field is represented using 4D real vectors obtained by expressing any Jones vector $\mathbf{z} = [z_1, z_2]^{\mathrm{T}} \in \mathbb{C}^2$ as

$$\mathbf{v}_{\mathbf{z}} = \begin{pmatrix} \mathcal{R}e(z_1) \\ Im(z_1) \\ \mathcal{R}e(z_2) \\ Im(z_2) \end{pmatrix}. \tag{3.45}$$

The analog to the propagation model (3.26) of the field is described by [80, 85, 86]

$$\mathbf{v}_{\mathbf{r}_k} = \mathbf{R}_k \mathbf{v}_{\mathbf{u}_k} + \mathbf{v}_{\mathbf{n}_k}, \tag{3.46}$$

where $\mathbf{v}_{\mathbf{u}_k}$ ($\mathbf{v}_{\mathbf{r}_k}$) is the transmitted (received) symbol, $\mathbf{v}_{\mathbf{n}_k}$ is the ASE noise, and $\mathbf{R}_k$ is a 4×4 real orthogonal matrix (corresponding to $\mathbf{H}_k$) modeling both the phase noise and the SOP drift such that

$$\mathbf{v}_{\mathbf{H}_k \mathbf{u}_k} = \mathbf{R}_k \mathbf{v}_{\mathbf{u}_k}, \tag{3.47}$$

where $\mathbf{v}_{\mathbf{H}_k \mathbf{u}_k}$ is obtained by applying (3.45) to $\mathbf{H}_k \mathbf{u}_k$.

The matrix $\mathbf{R}_k$ belongs to the special orthogonal group of 4×4 matrices, denoted $SO(4)$, and satisfies

$$\mathbf{R}_k^{\mathrm{T}} \mathbf{R}_k = \mathbf{R}_k \mathbf{R}_k^{\mathrm{T}} = \mathbf{I}_4, \tag{3.48}$$

$$\det \mathbf{R}_k = 1. \tag{3.49}$$

This group is also called the rotation group since it consists of all possible 4D rotation matrices and is a subset of the orthogonal group, where the determinant (3.49) can be either 1 or -1.

After constraining any real 4×4 matrix with 16 DOFs to satisfy (3.48) and (3.49), it can be shown that a matrix $\mathbf{R}_k$ that belongs to $SO(4)$ has *six* DOFs. Therefore, 4×4 real orthogonal matrices $\mathbf{R}_k$ are able to span over a richer space than the Jones matrices $\mathbf{H}_k$ can (four DOFs). Jones matrices can model all physically realizable phenomena of wave propagation and the extra two DOFs of $\mathbf{R}_k$ are not possible for photon propagation [80].

The matrix $\mathbf{R}_k$ can be expressed as $\mathbf{R}_k = R(\phi_k, \boldsymbol{\alpha}_k)$ using the matrix function [80]

$$R(\phi, \boldsymbol{\alpha}) = \exp((\phi, 0, 0) \cdot \vec{\boldsymbol{\lambda}} - \boldsymbol{\alpha} \cdot \vec{\boldsymbol{\rho}}), \tag{3.50}$$

$$= (\mathbf{I}_4 \cos\phi - \boldsymbol{\lambda}_1 \sin\phi)(\mathbf{I}_4 \cos\theta - \mathbf{a} \cdot \vec{\boldsymbol{\rho}} \sin\theta), \tag{3.51}$$

and its inverse is

$$\mathbf{R}_k^{-1} = \mathbf{R}_k^{\mathrm{T}} = R(-\phi_k, -\boldsymbol{\alpha}_k). \tag{3.52}$$

The tensors $\vec{\boldsymbol{\rho}} = (\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \boldsymbol{\rho}_3)$, $\vec{\boldsymbol{\lambda}} = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \boldsymbol{\lambda}_3)$ with [80, Eqs. (20)–(25)]

$$\boldsymbol{\rho}_1 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \ \boldsymbol{\rho}_2 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \ \boldsymbol{\rho}_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad (3.53)$$

$$\boldsymbol{\lambda}_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \ \boldsymbol{\lambda}_2 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \ \boldsymbol{\lambda}_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad (3.54)$$

form the basis for any 4D skew-symmetric real matrix and satisfy [80]

$$\boldsymbol{\rho}_i^2 = \boldsymbol{\lambda}_i^2 = -\mathbf{I}_4 \text{ for } i = 1, 2, 3, \quad (3.55)$$

$$\boldsymbol{\rho}_1\boldsymbol{\rho}_2\boldsymbol{\rho}_3 = -\boldsymbol{\rho}_2\boldsymbol{\rho}_1\boldsymbol{\rho}_3 = \mathbf{I}_4, \quad (3.56)$$

$$\boldsymbol{\rho}_1\boldsymbol{\rho}_2\boldsymbol{\rho}_3 = -\boldsymbol{\lambda}_2\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3 = \mathbf{I}_4. \quad (3.57)$$

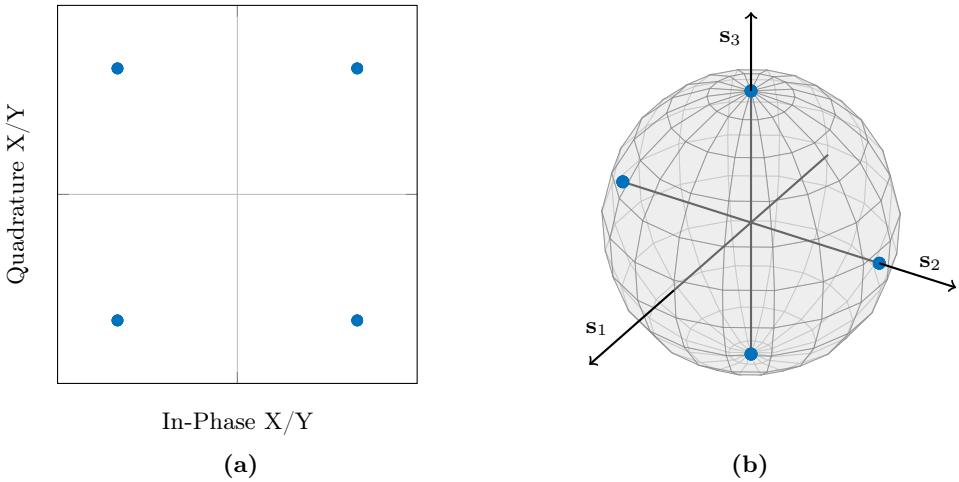
Note that in (3.50), two DOFs modeled by the scalars corresponding to $\boldsymbol{\lambda}_2, \boldsymbol{\lambda}_3$ are shown for pedagogical reasons, but have been deliberately set to zero such that a one-to-one mapping between Jones matrices $\mathbf{H}_k$ and real 4×4 matrices $\mathbf{R}_k$ is possible. However, any 4D real rotation matrix belonging to $SO(4)$ can be expressed using (3.50) and parameterizing all six DOFs.

Stokes Description

Another option for vector representation of optical signals is using Stokes vectors [87]. In general, Stokes vectors are 4D, but for fully polarized light and no polarization-dependent losses (which is the case for this thesis), the Stokes vectors can be reduced to be three-dimensional. In this case, the equivalent Stokes vector of a Jones vector $\mathbf{z}$ is [88, Eq. (2.5.26)]

$$\mathbf{s}_\mathbf{z} = \mathbf{z}^\mathrm{H}\vec{\boldsymbol{\sigma}}\mathbf{z} = \begin{pmatrix} |z_1|^2 - |z_2|^2 \\ 2\mathcal{R}e(z_1 z_2^*) \\ -2\mathit{Im}(z_1 z_2^*) \end{pmatrix}, \quad (3.58)$$

for any $\mathbf{z} = [z_1, z_2]^\mathrm{T} \in \mathbb{C}^2$, where the ith component of $\mathbf{s}_\mathbf{z}$ is given by $\mathbf{z}^\mathrm{H}\boldsymbol{\sigma}_i\mathbf{z}$. The absent fourth component is the optical power $\mathbf{z}^\mathrm{H}\mathbf{z}$. In can be noted from (3.58) that applying a common phase rotation to both z_1 and z_2 will not change $\mathbf{s}_\mathbf{z}$. E.g., both $\mathbf{z}_1 = [1+i, 1+i]^\mathrm{T}$



(a) **(b)**

Figure 3.2. PM-QPSK constellation shown in the Jones/4D space (a) and Stokes space (b).

and $\mathbf{z}_2 = [-1-i, -1-i]^{\mathrm{T}}$ will result in $\mathbf{s}_{\mathbf{z}} = [0, 4, 0]^{\mathrm{T}}$, where a rotation by π was applied to $\mathbf{z}_1$ to obtain $\mathbf{z}_2$. In order to convert a 4D real vector to a Stokes vector, first, the Jones vector should be obtained by reversing the operation in (3.45), and then the Stokes vector can be obtained by (3.58).

The Stokes vectors have the advantage that they are optically observable quantities proportional to the field intensity, and are often represented as points on a sphere, called the *Poincaré sphere*. The Poincaré sphere offers a good visual representation of the SOPs, where each SOP represents a point on this sphere and the drift of the SOP can be visualized as a rotation of the sphere. Fig. 3.2 illustrates PM-QPSK in the Jones/4D real space and in the Stokes space (plotted on the Poincaré sphere). In the former case, the constellation has four points in each polarization, thus resulting in a total of 16 points (all possible combinations between the two polarizations). On the other hand, in the Stokes space, the constellation has only four points since this representation is invariant to common phase shifts.

The analogous Stokes propagation model of (3.37) or (3.46) models only the SOP drift without phase noise and can be written as

$$\mathbf{s}_{\mathbf{r}_k} = \mathbf{M}_k \mathbf{s}_{\mathbf{u}_k} + \mathbf{s}_{\mathbf{n}_k}, \tag{3.59}$$

where $\mathbf{s}_{\mathbf{u}_k}$ ($\mathbf{s}_{\mathbf{r}_k}$) is the transmitted (received) Stokes vector and can be obtained by applying (3.58) to $\mathbf{u}_k$ ($\mathbf{r}_k$) and $\mathbf{s}_{\mathbf{n}_k}$ is the noise term. The matrix $\mathbf{M}_k$ is a real 3×3 orthogonal matrix, referred as *Mueller matrix*, which models only the SOP drift equivalently to $\mathbf{J}_k$ in (3.37) such that

$$\mathbf{s}_{\mathbf{J}_k \mathbf{u}_k} = \mathbf{s}_{\mathbf{H}_k \mathbf{u}_k} = \mathbf{M}_k \mathbf{s}_{\mathbf{u}_k}, \tag{3.60}$$

where $\mathbf{s}_{\mathbf{J}_k\mathbf{u}_k}$ ($\mathbf{s}_{\mathbf{H}_k\mathbf{u}_k}$) is obtained by applying (3.58) to $\mathbf{J}_k\mathbf{u}_k$ ($\mathbf{H}_k\mathbf{u}_k$). The noise component $\mathbf{s}_{\mathbf{n}_k}$ consists of three terms and cannot be obtained by (3.58). It can be identified by equating terms in two expressions for $\mathbf{s}_{\mathbf{r}_k}$, where one is obtained by applying (3.58) to both sides of (3.26) and then applying (3.60), and the other is (3.59). Thus,

$$\mathbf{s}_{\mathbf{n}_k} = (\mathbf{H}_k\mathbf{u}_k)^{\mathrm{H}}\vec{\boldsymbol{\sigma}}\mathbf{n}_k + \mathbf{n}_k^{\mathrm{H}}\vec{\boldsymbol{\sigma}}\mathbf{H}_k\mathbf{u}_k + \mathbf{n}_k^{\mathrm{H}}\vec{\boldsymbol{\sigma}}\mathbf{n}_k, \tag{3.61}$$

$$= (e^{-i\phi_k}\mathbf{J}_k\mathbf{u}_k)^{\mathrm{H}}\vec{\boldsymbol{\sigma}}\mathbf{n}_k + \mathbf{n}_k^{\mathrm{H}}\vec{\boldsymbol{\sigma}}e^{-i\phi_k}\mathbf{J}_k\mathbf{u}_k + \mathbf{n}_k^{\mathrm{H}}\vec{\boldsymbol{\sigma}}\mathbf{n}_k, \tag{3.62}$$

where the first two terms represent the signal–noise interaction and the last one the noise–noise interaction. As can be noted, $\mathbf{s}_{\mathbf{n}_k}$ is signal-dependent and (3.59) is not an additive-noise model, opposed to (3.26), (3.37), and (3.46).

The matrix $\mathbf{M}_k$ belongs to the special orthogonal group of 3×3 matrices $SO(3)$ satisfying

$$\mathbf{M}_k^{\mathrm{T}}\mathbf{M}_k = \mathbf{M}_k\mathbf{M}_k^{\mathrm{T}} = \mathbf{I}_3, \tag{3.63}$$

$$\det \mathbf{M}_k = 1, \tag{3.64}$$

and the polarization transformation modeled by it can be seen as a rotation of the Poincaré sphere. A generic 3×3 real matrix has nine DOFs, whereas after constraining $\mathbf{M}_k$ to fulfill (3.63)–(3.64), it can be shown that it has only three DOFs. Such matrices can be expressed as $\mathbf{M}_k = M(\boldsymbol{\alpha}_k)$ using the matrix function [83]

$$M(\boldsymbol{\alpha}) = \exp(2\mathcal{K}(\boldsymbol{\alpha})) = \exp(2\theta\mathcal{K}(\mathbf{a})), \tag{3.65}$$

$$= \mathbf{I}_3 + \sin(2\theta)\mathcal{K}(\mathbf{a}) + (1 - \cos(2\theta))\mathcal{K}(\mathbf{a})^2, \tag{3.66}$$

where

$$\mathcal{K}(\mathbf{a}) = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}. \tag{3.67}$$

The transformation in (3.65) can be viewed in the axis-angle rotation description as a rotation around the unit vector $\mathbf{a}$ by an angle 2θ. The inverse of $M(\boldsymbol{\alpha})$ can be obtained from

$$M(\boldsymbol{\alpha})^{-1} = M(\boldsymbol{\alpha})^{\mathrm{T}} = M(-\boldsymbol{\alpha}). \tag{3.68}$$

3.3.2 Channel Modeling

After establishing the mathematical representation of phase noise and rotations of the SOP above, we will now discuss options to model ϕ_k, $\mathbf{J}_k$, $\mathbf{R}_k$, and $\mathbf{M}_k$ such that the behavior of real optical fibers is achieved. First, we will discuss phase noise and then SOP drift. We will however, without loss of generality, restrict the discussion below to Mueller

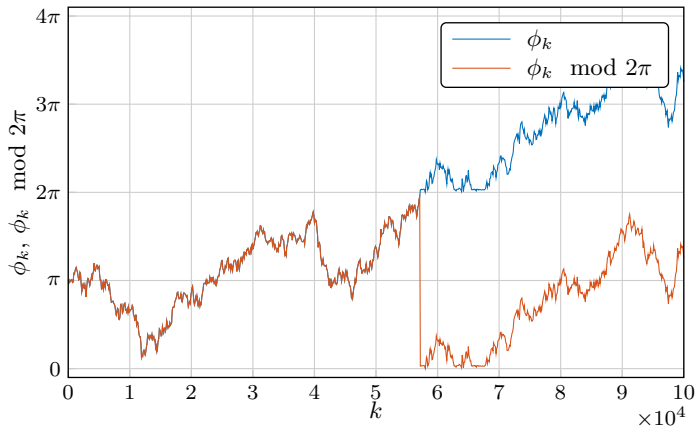


Figure 3.3. An example of phase noise random walk for $\sigma_\nu^2 = 2.25 \cdot 10^{-4}$, where mod is the modulo operation.

matrices $\mathbf{M}_k$ as regards to SOP drifts. Nonetheless, these can be straightforwardly mapped to Jones or 4D matrices.

Phase Noise

The phase noise in (3.26), (3.37), or (3.46) is modeled as a Wiener process [64, 89]

$$\phi_k = \dot{\phi}_k + \phi_{k-1}, \tag{3.69}$$

where $\dot{\phi}_k$ is the *innovation* of the phase noise. The innovation $\dot{\phi}_k$ is a real Gaussian random variable drawn independently at each time instance k as

$$\dot{\phi}_k \sim \mathcal{N}(0, \sigma_\nu^2), \tag{3.70}$$

where the variance $\sigma_\nu^2 = 2\pi\Delta\nu T$, and $\Delta\nu$ is the sum of the linewidths of the transmitter and receiver lasers.

The acquired phase noise at time k is the summation of the innovations $\dot{\phi}_1, \ldots, \dot{\phi}_k$ and the initial phase ϕ_0. Since the $\dot{\phi}_1, \ldots, \dot{\phi}_k$ terms are Gaussian, ϕ_k becomes a Gaussian-distributed random variable with mean ϕ_0 and variance $k\sigma_\nu^2$. Due to the periodicity with period 2π of the function $e^{-i\phi_k}$, the phase angle ϕ_k can be bounded to the interval $[0, 2\pi)$ by applying the modulo 2π operation. In this case, the probability density function (pdf) of ϕ_k becomes a wrapped Gaussian distribution. Fig. 3.3 shows an example of a phase noise random walk and the corresponding wrapped phase noise. The evolution of the wrapped phase noise pdf is shown in [Paper A, Fig. 1]. The initial phase difference ϕ_0 between the two free-running lasers has equal probability for every value; therefore it is common to model ϕ_0 as a random variable uniformly distributed in the interval $[0, 2\pi)$.

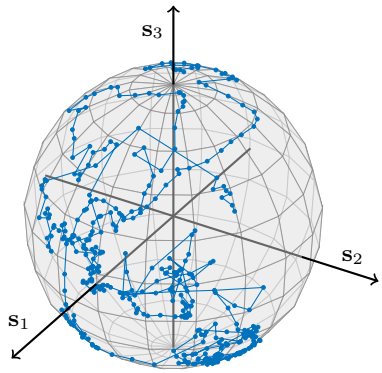


Figure 3.4. The trajectory of a measured SOP drift is plotted. Each dot represents a measurement taken ~ 2.2 h apart.

Polarization State Drift

The change in the SOP between the input and output is a dynamic process that changes over time irrespectively of the input SOP. Plots of the Stokes vectors on the Poincaré sphere are often used as good visual representations of the phenomenon. Fig. 3.4 illustrates the trajectory of the received Stokes vectors for a fixed input over time. The plotted data was obtained by measuring a 127-km long buried fiber for 36 days at every ~ 2.2 h (shown by the dots in the figure). The technicalities of the measurement setup and postprocessing have been published elsewhere [90]. As can be seen, the SOP has a random behavior, taking steps of various sizes in no preferred direction.

Several variations of polarization drift models have been proposed and used in the literature, most commonly in the context of equalization design. These models can be distinguished by the dynamism and randomness of their behavior in time. In the next sections, different variations will be reviewed.

Static The simplest and most straightforward solution is to generate a Jones/Mueller/4D matrix randomly, which thereafter is kept constant [91–94]. This assumption is reasonable if the considered time scale is small compared to time scale of the SOP drift. As can be seen in Fig. 3.4, in the case of this particular fiber the SOP does not change significantly between two consecutive measurements taken ~ 2.2 h apart. Therefore, if, e.g., 10^4 symbols at 28 Gbaud are considered, these correspond to 0.35 µs, which is a too short time to have significant changes of the SOP. Of course this depends on the installation specifics of the fiber. Aerial fibers will change at a much faster pace than the more stable buried fibers [95].

The Jones/Mueller/4D matrix can be generated using several approaches by either choosing $\boldsymbol{\alpha}$ according to a certain pdf or randomly selecting the elements of the matrix. The latter approach does not ensure unitarity/orthogonality of the matrices, and there-

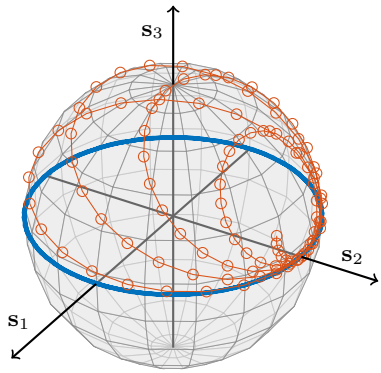


Figure 3.5. SOP trajectories obtained by modeling $\alpha_{3,k} = \omega T k$ (blue curve in the $\mathbf{s}_1$-$\mathbf{s}_2$ plane), and $\alpha_{2,k} = \omega_1 T k$ and $\alpha_{3,k} = \omega_2 T k$ (red curve with round markers).

fore it must be followed by a normalization process. Different approaches will sample the space of possible unitary/orthogonal matrices according to different distributions.

In [96], a method to generate $\boldsymbol{\alpha}$ is described such that the space of all possible Mueller matrices is uniformly sampled. In order to generate such a matrix, the axis $\mathbf{a}$, cf. (3.42), must be uniformly distributed over the unit sphere and the pdf of the angle $\theta \in [0, \pi/2)$ must be $(1 - \cos\theta)/\pi$ [97]. However, we present an alternative to generate the axis and the angle simultaneously [96]. The vector $\boldsymbol{\alpha} = \theta\mathbf{a}$ is formed from the unit vector $(\cos\theta, a_1 \sin\theta, a_2 \sin\theta, a_3 \sin\theta)^{\mathrm{T}} = \mathbf{g}/\|\mathbf{g}\|$, where $\mathbf{g} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_4)$, which will satisfy the conditions for both axis $\mathbf{a}$ and angle θ. In this case, all the possible SOP are equally likely after multiplying such a matrix with a constant Stokes vector. The same method to generate $\boldsymbol{\alpha}$ can be used to form Jones (3.38) or 4D matrices (3.50).

Dynamic Deterministic One option to emulate dynamic SOP drift is to model the elements of $\boldsymbol{\alpha}$ as frequency components. For example, in [61, 98, 99], SOP drift was simulated by modeling $\alpha_{3,k} = \omega T k$ and setting $\alpha_{1,k} = \alpha_{2,k} = 0$. Another practice is to only set $\alpha_{1,k} = 0$ and vary $\alpha_{2,k}$ and $\alpha_{3,k}$ at different frequencies, i.e., $\alpha_{2,k} = \omega_1 T k$ and $\alpha_{3,k} = \omega_2 T k$ [54, 100, 101]. Fig. 3.5 illustrates the trajectories obtained by these two methods. As can be seen, these methods have a cyclic/quasi-cyclic deterministic behavior and cover only a subset of all possible SOPs on the Poincaré sphere.

Dynamic Stochastic As can be seen in Fig. 3.4, the SOP has a random trajectory and cannot be accurately modeled with a deterministic behavior such as those discussed in the previous section and shown in Fig. 3.5. A attempt to emulate stochastic polarization drifts for the first time in the polarization literature was made in our prestudy [102] for

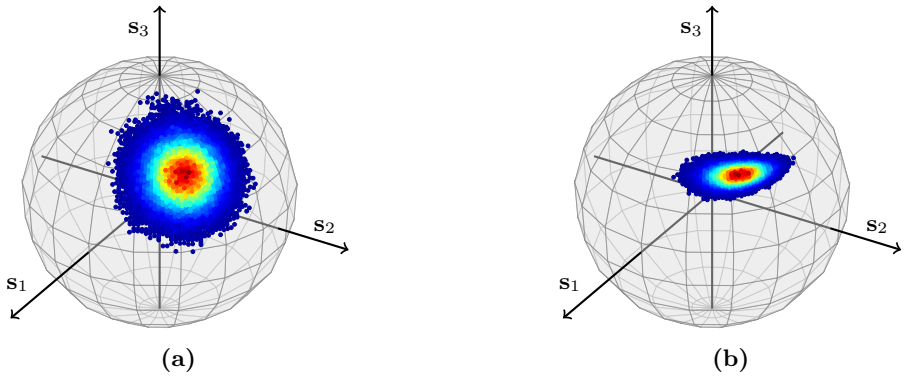


Figure 3.6. The histograms of $M(\dot{\boldsymbol{\alpha}}_k + \boldsymbol{\alpha}_{k-1})\mathbf{s}_{\mathbf{u}_k}$ for different $\boldsymbol{\alpha}_{k-1}$ and a fixed $\mathbf{s}_{\mathbf{u}_k}$ are shown. The highest density is represented by dark red and the lowest by dark blue, the outer part of the histogram. The figure on the left (a) was obtained for $\boldsymbol{\alpha}_{k-1} = [0, 0, 0]^{\mathrm{T}}$, whereas the figure on the right (b) was obtained for $\boldsymbol{\alpha}_{k-1} = [0, 0, 3\pi/4]^{\mathrm{T}}$. In both cases $\sigma_p^2 = 0.007$.

Paper A. In [102], we modeled[6] $\mathbf{M}_k$ in (3.59) as

$$\mathbf{M}_k = M(\boldsymbol{\alpha}_k), \tag{3.71}$$

where $\boldsymbol{\alpha}_k$ follows, analogously to the phase noise, a Wiener process

$$\boldsymbol{\alpha}_k = \dot{\boldsymbol{\alpha}}_k + \boldsymbol{\alpha}_{k-1}, \tag{3.72}$$

and $\dot{\boldsymbol{\alpha}}_k$ is the *innovation* of the SOP drift, cf. (3.69). The innovation parameters are random and drawn independently from a zero-mean real Gaussian distribution at each time instance k

$$\dot{\boldsymbol{\alpha}}_k \sim \mathcal{N}(\mathbf{0}, \sigma_p^2 \mathbf{I}_3), \tag{3.73}$$

where $\sigma_p^2 = 2\pi\Delta p T$. We refer to Δp as the *polarization linewidth*, which quantifies the speed of the SOP drift, analogous to the linewidth describing the phase noise, cf. (3.70).

However, this model is not a stationary random process. If we consider the innovation at k given the previous state, i.e., $\mathbf{M}_k\mathbf{M}_{k-1}^{-1}$, its pdf depends not only on $\dot{\boldsymbol{\alpha}}_k$ in (3.72), but on $\boldsymbol{\alpha}_{k-1}$ as well. Therefore, the statistics of $M(\boldsymbol{\alpha}_k)$ vary with $\boldsymbol{\alpha}_{k-1}$ and time k. On the contrary, the statistics of a stationary random process should depend only on the statistics of the innovation $\dot{\boldsymbol{\alpha}}_k$. A (visual) comparison between two histograms of $\mathbf{M}_k\mathbf{s}_{\mathbf{u}_k}$ obtained for different $\boldsymbol{\alpha}_{k-1}$ but the same σ_p^2 and $\mathbf{s}_{\mathbf{u}_k}$ is shown in Fig. 3.6. As can be seen, the shape of the histograms is different and confirm the nonstationary property of the random process $\mathbf{M}_k\mathbf{s}_{\mathbf{u}_k}$. Moreover, for $\boldsymbol{\alpha}_{k-1} = [0, 0, 3\pi/4]^{\mathrm{T}}$ Fig. 3.6(b), in contrast

[6]The model proposed in [102] is described here in the Stokes space. However, the same model applies to the other two descriptions as well.

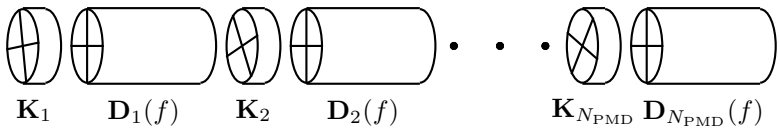


Figure 3.7. PMD model of a long fiber as a concatenation of N_{PMD} frequency-dependent birefringent sections $\mathbf{D}_n(f)$ coupled with polarization scramblers $\mathbf{K}_n$.

with the other case $\boldsymbol{\alpha}_{k-1} = [0,0,0]^{\mathrm{T}}$ Fig. 3.6(a), the histogram is not isotropic, and by isotropic we mean that all the possible orientations of the changes of the Stokes vector are equally likely.

In conclusion, this model is neither stationary nor isotropic. Our goal in Paper A is to find a model that satisfies both these conditions. The main difference between the model in [102] and the one in Paper A consists in the updating method of the matrix $\mathbf{M}_k$, which is done as a multiplication of matrices [Paper A, Eq. (A.15)]

$$\mathbf{M}_k = M(\dot{\boldsymbol{\alpha}}_k)\mathbf{M}_{k-1}, \tag{3.74}$$

where $\dot{\boldsymbol{\alpha}}_k$ is modeled as (3.73). This model has been experimentally validated and its details together with the analogous Jones and 4D descriptions can be found in Paper A.

3.4 Polarization-Mode Dispersion

When trying to understand the effects of PMD, it is often good to understand first its effects over a short fiber segment. To this end, we assume that the segment is short enough such that any PMD effects are constant over its length. Any fiber can be modeled as a concatenation of such fiber segments. In the short-fiber-segment regime, perturbations such as mechanical and thermal stress, bends, and twists break the cylindrical symmetry of the core and lead to nonequal effective indices of refraction β_{1x}, β_{1y} for the slow and fast propagation modes. This phenomenon is exemplified in Fig. 2.7, where two pulses launched in the two orthogonal polarizations acquire a time separation. This time separation is called DGD, denoted with τ, and arises along the so-called slow and fast (birefringence) axes of the fiber, and it is usually expressed in picoseconds.

While for short fiber segments, PMD exhibits a predictable behavior due to the additive nature of DGD, this is not the case for long fiber links of hundreds or thousands of km. Long links can be viewed as a concatenation of many short sections where the birefringence axes of each section change randomly from section to section. This phenomenon is often called polarization coupling, and it originates from splices, optical components, variations in the fiber drawing process, and from the intentional fiber spinning during drawing [103].

A common model to emulate PMD in long fibers is the waveplate model, which models

the entire fiber as a concatenation of many birefringent fiber sections with random coupling. This model is shown in Fig. 3.7, where the short sections $\mathbf{K}_n$, for $n = 1, \ldots, N_{\text{PMD}}$, scramble randomly the birefringence axes at the beginning of each birefringent section $\mathbf{D}_n(f)$, which in turn introduce DGD. The birefringence sections $\mathbf{D}_n(f)$ are frequency dependent, whereas polarization scramblers $\mathbf{K}_n$ are not.

Simulations are often performed in the Jones space, perhaps due to the mapping between Jones matrices and practical components that practitioners often make. The transfer function of the process shown in Fig. 3.7 is given by the product of the different sections as

$$\mathbf{T}(L, f) = \mathbf{D}_{N_{\text{PMD}}}(f)\mathbf{K}_{N_{\text{PMD}}} \cdots \mathbf{D}_2(f)\mathbf{K}_2\mathbf{D}_1(f)\mathbf{K}_1, \tag{3.75}$$

where the matrices $\mathbf{K}_n$ and $\mathbf{D}_n(f)$ are 2×2 complex Jones matrices and the output of the fiber is related to the input as

$$\tilde{\mathbf{E}}(L, f) = \mathbf{T}(L, f)\tilde{\mathbf{E}}(0, f). \tag{3.76}$$

Each birefringent section is a first-order PMD element and is modeled as

$$\mathbf{D}_n(f) = \begin{pmatrix} e^{i\pi f\tau_n} & 0 \\ 0 & e^{-i\pi f\tau_n} \end{pmatrix}, \tag{3.77}$$

introducing a τ_n delay between the two polarizations. The delays τ_n can be chosen to be equal $\tau_n = \tau_p$ for all sections, or randomly and independently for each section from a Gaussian distribution $\tau_n \sim \mathcal{N}(\tau_p, (\tau_p/5)^2)$ [104], where the $1/5$ coefficient was heuristically found to mimic well experimental data. The value of τ_p is chosen according to the links length, PMD coefficient, and number of sections N_{PMD}, which we discuss below. Choosing τ_n differently for each section avoids spectral periodicity with period $2/\tau_n$ of $\mathbf{D}_n(f)$, and consequently of $\mathbf{T}(L, f)$.

The polarization scramblers $\mathbf{K}_n$ can be modeled using (3.38) and by choosing the parameters $\boldsymbol{\alpha}$ randomly. Several options have been considered in the literature, some already mentioned in Section 3.3.2, in the context of SOP drift. One option is to chose uniformly $\alpha_3 \in [0, 2\pi)$ and letting $\alpha_1 = \alpha_2 = 0$. Another practice is to only set $\alpha_1 = 0$ and chose α_2 and α_3 uniformly in the $[0, 2\pi)$ interval. None of these methods will however generate a uniformly distributed polarization scrambling matrix $\mathbf{K}_n$. This can be achieved using the method described in the case of static SOP drift discussed earlier.

The accumulated DGD over the link can be obtained from the frequency derivative of the transfer matrix [83]

$$\tau(f) = \frac{1}{\pi}\sqrt{\det\left(\frac{\partial \mathbf{T}(L, f)}{\partial f}\right)}, \tag{3.78}$$

which, in the limit of an infinite number of sections, is Maxwellian distributed, i.e., the pdf of $\tau(f)$ approaches $p(\tau) = \sqrt{2/\pi}\tau^2 \exp(-\tau^2/(2a^2))/a^3$ as $N_{\text{PMD}} \to \infty$ for any f,

where $a = \tau_p\sqrt{N_{\mathrm{PMD}}/3}$. Valid for the Maxwellian distribution, the first and second moments of $\tau(f)$ follow the relation $\mathbb{E}[\tau(f)] = \sqrt{8\mathbb{E}[\tau^2(f)]/(3\pi)}$. Note that the constant $1/\pi$ in (3.78) replaces the more common constant 2, such as in [83, Eq. (5.7)], since the differentiation is with respect to frequency f as opposed to angular frequency ω in the latter. For a given N_{PMD}, the average DGD is given by [105]

$$\mathbb{E}[\tau(f)] = \sqrt{\frac{8N_{\mathrm{PMD}}}{3\pi}}\tau_p. \tag{3.79}$$

Based on this equation and on the links details, i.e., length and PMD coefficient D_{PMD}, the parameter τ_p can be chosen. For example, a 1000-km link with $D_{\mathrm{PMD}} = 0.1$ ps/$\sqrt{\mathrm{km}}$ can be emulated by letting $N_{\mathrm{PMD}} = 50$ and $\tau_p = 0.4854$ ps.

One direct implication of letting $N_{\mathrm{PMD}} \to \infty$ is that the random variable $\tau(f)$ becomes unbounded since the support of the Maxwellian distribution extends to infinity. However, in practical simulations, the number of sections is finite and $\tau(f)$ becomes upperbounded to the sum of DGDs introduced by each birefringent section in (3.75). For example, for sections with DGD equal to τ_p, the upper bound becomes $N_{\mathrm{PMD}}\tau_p$. This is more in line with real fiber transmissions where an infinite delay is unpractical. However, the question of how many sections is suitable for simulations is not easy to answer. A few studies [104, 106, 107] have tried to analyze this problem and it seems there is a general agreement that a too low number of sections is not suitable. This results in a distribution of the DGD that mismatches the Maxwellian distribution to a high degree and is inadequate. On the other hand, emulators with 15 sections appear to be adequate in [106], whereas 50 sections were used in [107]. Nevertheless, as N_{PMD} grows large, the statistics of $\tau(f)$ approach a truncated version of the Maxwellian distribution and the truncation occurs at $N_{\mathrm{PMD}}\tau_p$.

The PMD model presented above is static and does not have any temporal variations. However, this not the case for real fibers, which have been measured to have time-varying PMD [90, 95, 108–110]. Time variations can be introduced in the model above by letting the polarization scramblers $\mathbf{K}_n$ change with time.

In the classic PMD literature, all polarization scramblers are assumed to be time varying. This assumption however has been questioned by the so-called hinge PMD model [111, 112], where most of the polarization scramblers are considered to be frozen in time and only a few, called hinges, are time varying. The latter could be amplifier sites, fiber segments exposed to temperature or mechanical variations in servicing huts, railroad bridges, etc. The major statistical implication of the hinge model suggests that, given the fronzen (in time) sections between the hinges, the DGD at different spectral components do not have the same statistics, which is the case for the classical PMD model. In the classical PMD model, the DGD is the same at all frequencies for each frozen section $\mathbf{D}_n(f)$, i.e., $\tau(f) = \tau_n$ for all f given n. In the case of the hinge model, the DGD of the frozen sections is not constant but random with respect to frequency and follows a (truncated) Maxwellian distribution. The concatenation of such frozen sections

leads to DGD that has different statistics with respect to frequency, which depend on the particular DGD realization of the frozen sections. For example, let's examine the DGD at two frequencies f_1 and f_2 of a simple link made up by two frozen section $\mathbf{D}_1(f)$ and $\mathbf{D}_2(f)$. The first section has a DGD of 1 ps and 5 ps at f_1 and f_2, respectively, whereas the second section has a DGD of 3 ps and 2 ps at f_1 and f_2, respectively. Therefore after concatenating the two sections, the entire link can experience, in the worst case scenario, a total DGD of 4 ps and 7 ps at f_1 and f_2, respectively, making the statistics of the DGD frequency dependent. However, this analysis holds when conditioned on the frozen sections, i.e., they are fixed. When the entire ensemble of fibers is considered, with all the possible realizations of the frozen sections, the DGD has the same statistics at all frequencies. In Paper B, we study a channel model for the temporal variations of PMD and absolute SOP within the hinge-model framework and compare it to field-trial data.

3.5 The Split-Step Fourier Method

The propagation equations discussed in Section 3.1 do not have general solutions except for some specific cases. Therefore, various numerical methods are used in simulations or to get insights into the nonlinear propagation phenomenon. The split-step Fourier method (SSFM) is an efficient numerical tool[7] to solve nonlinear partial differential equations, which was originally proposed in [114], and it is commonly used to solve the NLSE or Manakov equation. Finite difference methods [115] is another class of numerical tools that can be used to solve the propagation equations, where the derivatives are approximated with difference equations. In general, the SSFM is more commonly used due to its faster evaluation at the same accuracy. The speed advantage can be associated mainly with the use of the fast Fourier transform algorithm. A more detailed discussion about the two approaches can be found in [44, Sec. 2.4] and references therein.

The assumption behind the SSFM is that the linear and nonlinear parts can be decoupled and act independently over a small enough spatial step size. In order to facilitate this assumption, the Manakov-PMD equation[8] (3.15), ignoring the ASE noise, can be reformulated as

$$\frac{\partial \mathbf{E}(z,t)}{\partial z} = (\widehat{D} + \widehat{N})\mathbf{E}(z,t), \tag{3.80}$$

where $\widehat{D}$ is the linear operator accounting for losses, amplification, CD, and PMD, whereas $\widehat{N}$ is the nonlinear operator accounting for the Kerr nonlinearity. These are

[7] In general, the SSFM is considered to be a numerical tool, rather than a channel model. However, these two are not mutually exclusive and moreover, the SSFM has been used in [113] as a channel model to establish an upper bound on the capacity of the fiber-optical channel.

[8] Here we focus on the Manakov-PMD equation due its generality, but the SSFM can be applied to all propagation equations presented in Section 3.1.

given by

$$\widehat{D} = -\frac{\alpha - g(z)}{2} + \Delta\beta_1 \mathbf{\Sigma}(z)\frac{\partial}{\partial t} - i\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}, \tag{3.81}$$

$$\widehat{N} = i\gamma\frac{8}{9}\|\mathbf{E}(z,t)\|^2. \tag{3.82}$$

The propagation is performed recursively over spatial steps Δz_i under the assumption that the linear and nonlinear effects act independently if the steps are small enough. Although not required by the method itself, it is common to discretize the total link in and equal number of $N_{\text{span}-\text{SSF}} \in \mathbb{N}$ steps per span such that $\sum_{i=1}^{N_{\text{sp}} \times N_{\text{span}-\text{SSF}}} \Delta z_i = L$, where $N_{\text{sp}} \times N_{\text{span}-\text{SSF}}$ is the total number of discrete spatial steps[9]. The propagation over one spatial step can be approximated in a two-phase procedure by solving for the linear part first (by setting $\widehat{N} = 0$) and then for the nonlinear part (by setting $\widehat{D} = 0$), or vice versa. This becomes mathematically

$$\mathbf{E}(z + \Delta z_i, t) \approx \exp(\widehat{D}\Delta z_i)\exp(\widehat{N}\Delta z_i)\mathbf{E}(z,t). \tag{3.83}$$

This equation can be implemented as follows:

- the nonlinear step is applied first in the time domain as (cf. (3.6))

$$\mathbf{E}(z + \Delta z_i, t) \approx \exp\Big(i\frac{8}{9}\gamma\|\mathbf{E}(z,t)\|^2\Delta z_i\Big)\mathbf{E}(z,t) \tag{3.84}$$

- in the second step, the linear part is applied in the frequency domain

$$\tilde{\mathbf{E}}(z + \Delta z_i, f) \approx \exp\big(i2\beta_2\pi^2 f^2 \Delta z_i\big)\exp\big(-(\alpha - g(z))\Delta z_i/2\big)\mathbf{T}(\Delta z_i, f)\tilde{\mathbf{E}}(z,f), \tag{3.85}$$

  where the first three terms on the right-hand side account for CD, attenuation and gain, and PMD, respectively.

The steps above are repeated $N_{\text{sp}} \times N_{\text{span}-\text{SSF}}$ times such that the propagation through the entire link is performed.

In the description above, we included the attenuation and gain in the linear step. Nevertheless, this can be performed during the nonlinear step and it has been shown that it leads to a better accuracy [116]. In this case, (3.84) and (3.85) become

$$\mathbf{E}(z + \Delta z_i, t) \approx \exp\big(-(\alpha - g(z))\Delta z_i/2\big)\exp\Big(i\frac{8}{9}\gamma\|\mathbf{E}(z,t)\|^2 L_{\text{eff}}(\Delta z_i)\Big)\mathbf{E}(z,t), \tag{3.86}$$

$$\tilde{\mathbf{E}}(z + \Delta z_i, f) \approx \exp\big(i2\beta_2\pi^2 f^2\Delta z_i\big)\mathbf{T}(\Delta z_i, f)\tilde{\mathbf{E}}(z,f), \tag{3.87}$$

respectively.

[9] This spatial distribution facilities the noise insertion at the end of each span in the case of lumped amplification.

However, the procedure above is concerned only with propagation effects and does not add amplification noise. Noise is typically added in the time domain at spatial points where amplification occurs. This occurs at the end of each span in the case of lumped amplification and at every step in the case of distributed amplification (see Section 3.1.1).

The formulation above of the SSFM is known as the asymmetric method. The naming origin will become evident after we will present the symmetric method. The accuracy of the asymmetric method can be improved by reformulating (3.83) as

$$\mathbf{E}(z+\Delta z_i,t) \approx \exp(\widehat{D}\Delta z_i/2)\exp(\widehat{N}\Delta z_i)\exp(\widehat{D}\Delta z_i/2)\mathbf{E}(z,t); \tag{3.88}$$

known as the symmetric method, where the nonlinear step is included in the middle of the linear step rather than at the boundary. This method has an error that is of third order in the step size Δz_i, compared to second order achieved by the asymmetric method. Nonetheless, the two methods converge as Δz_i becomes small.

Although the algorithms above are straightforward to implement, the selection of the step size Δz_i is crucial and governs the desired complexity–accuracy trade-off [117]. The most trivial approach is to set all steps equal; a method that is also highly computationally inefficient. Another method is to chose a logarithmic distribution of steps such that spurious four-wave mixing is avoided [118]. Other methods [116, 117, 119] optimize the step size at every step such that, e.g., the maximum nonlinear phase shift is kept at a given maximum or the local relative error is kept within some limits.

Another important accuracy aspect related to simulations based on the SSFM is the time discretization. The propagation equations above are presented in continuous time, but any numerical simulation is performed in the discrete-time domain. Linear numerical simulations bridge losslessly the continuous-time and discrete-time domains by satisfying the Nyquist–Shannon sampling theorem. This however is not the case for the nonlinear equations presented here, which require higher sampling rates as the power levels, and implicitly the nonlinear effects, increase [120]. This effect can be partially explained by the bandwidth expansion, but so far there are no available specific guidelines on how these two relate.

3.6 Perturbative Channel Models

The propagation models presented in Section 3.1 are difficult to analyze using information-theoretic tools since they are formulated in continuous time and the input–output relation is not explicitly given, but is rather implicit in the differential equation. In order to overcome this complication, the research community has resorted to various methods to obtain an explicit input–output relationship for the fiber optical channel. Several common approaches are reviewed in the tutorial [121], from which we will consider a few examples that are based on perturbation theory in the following.

Perturbation theory is a class of methods aimed for finding approximate solutions for

(typically) unsolvable problems. The solution obtained with such methods is expressed in terms of power series truncated at the desired accuracy. The fact that the NLSE (and its variations) does not have a general solution yet and that the nonlinear interference is weak make the NLSE a good candidate for perturbation theory. Typically, these methods are applied to the propagation equation modeling strictly propagation effects ignoring ASE noise, which is added separately at the receiver; hence signal–noise interactions are ignored. This is a reasonable assumption as long as the nonlinear signal–signal interactions are dominant.

Perturbation theory analysis can be broadly classified into two categories: frequency-domain [122, 123] and time-domain [124, 125]. In turn, these can be further categorized into two groups depending on how nonlinearity is modeled: deterministic and random nonlinearity [121]. In the following, we will focus first on the deterministic frequency-domain nonlinearity modeling based on Volterra series and then on random time-domain nonlinearity modeling.

3.6.1 The Volterra Series Model

The Volterra series model is a generic tool used to model nonlinear behavior [126]. Peddanarappagari and Brandt–Pearce have applied this tool to solve the NLSE in [122] accounting for up to the fifth-order Volterra kernels. In [123] has been shown that the $2n+1$ order Volterra series solution in [122] coincides with the solution obtained based on the order n regular perturbation method. Later on in [127] the equivalency between the Volterra series model and the time-domain model proposed in [124] has been established.

The output field $\tilde{\mathbf{E}}(z,f)$ can be related to the input $\tilde{\mathbf{E}}(0,f)$ using the third-order truncated Volterra series as

$$\begin{aligned}\tilde{\mathbf{E}}(z,f) \approx & H_1(z,f)\mathbf{T}(z,f)\tilde{\mathbf{E}}(0,f)+ \\ & i\gamma\frac{8}{9}(2\pi)^2\int_0^z H_1(z-z',f)\mathbf{T}(z-z',f)\iint \left(H_1(z',f_2)\mathbf{T}(z',f_2)\tilde{\mathbf{E}}(0,f_2)\right)^{\mathrm{H}} \\ & H_1(z',f_1)\mathbf{T}(z',f_1)\tilde{\mathbf{E}}(0,f_1)H_1(z',f_3)\mathbf{T}(z',f_3)\tilde{\mathbf{E}}(0,f_3)\ \mathrm{d}f_1\ \mathrm{d}f_2\ \mathrm{d}z', \end{aligned} \tag{3.89}$$

where

$$H_1(z,f) = v(z)\exp(i2\beta_2\pi^2 f^2 z) \tag{3.90}$$

is the linear first-order Volterra series kernel, modeling the fiber attenuation and CD, and $f_3 = f - f_1 + f_2$. In (3.90), the function

$$v(z) = \exp\left(-\frac{\alpha z - \int_0^z g(z')\mathrm{d}z'}{2}\right) \tag{3.91}$$

accounts for the power profile of the link. E.g., it is equal to 1 in the case of perfectly distributed amplification and $v(z) = \exp(-\alpha(z \mod L_{\mathrm{sp}})/2)$ in the case of lumped amplification.

In the absence of PMD, i.e., $\mathbf{T}(z, f) = \mathbf{I}_2$ for any z and f, the integral with respect to the distance z' can be analytically solved and (3.89) becomes

$$\begin{aligned}\tilde{\mathbf{E}}(N_{\mathrm{sp}}L_{\mathrm{sp}}, f) \approx & H_1(N_{\mathrm{sp}}L_{\mathrm{sp}}, f)\mathbf{T}(z, f)\tilde{\mathbf{E}}(0, f) + \\ & i\gamma\frac{8}{9}(2\pi)^2 \iint H_3(N_{\mathrm{sp}}, L_{\mathrm{sp}}, f, f_1, f_2)\tilde{\mathbf{E}}(0, f_2)^{\mathrm{H}}\tilde{\mathbf{E}}(0, f_1)\tilde{\mathbf{E}}(0, f_3)\ \mathrm{d}f_1\ \mathrm{d}f_2,\end{aligned} \tag{3.92}$$

where we constrained the propagated distance to be an integer number of spans $N_{\mathrm{sp}}L_{\mathrm{sp}}$ and the amplification scheme to be either lumped or ideal distributed. The quantity $H_3(N_{\mathrm{sp}}, L_{\mathrm{sp}}, f, f_1, f_2)$ is the third-order Volterra series kernel and it reads

$$\begin{aligned}H_3(N_{\mathrm{sp}}, L_{\mathrm{sp}}, f, f_1, f_2) = & \exp(-\alpha_{\mathrm{eff}}L_{\mathrm{sp}}/2)\exp(i2\beta_2\pi^2 f^2 N_{\mathrm{sp}}L_{\mathrm{sp}}) \\ & \frac{1 - \exp\left(i\beta_2(2\pi)^2(f_1 - f_2)(f_1 - f)L_{\mathrm{sp}}N_{\mathrm{sp}}\right)}{1 - \exp\left(i\beta_2(2\pi)^2(f_1 - f_2)(f_1 - f)L_{\mathrm{sp}}\right)} \\ & \frac{1 - \exp\left(-L_{\mathrm{sp}}(\alpha_{\mathrm{eff}} - i\beta_2(2\pi)^2(f_1 - f_2)(f_1 - f))\right)}{\alpha_{\mathrm{eff}} - i\beta_2(2\pi)^2(f_1 - f_2)(f_1 - f)},\end{aligned} \tag{3.93}$$

where α_{eff} is the effective attenuation of the fiber, i.e., the combined effects of fiber loss and amplification scheme. In the case of lumped amplification, it is equal to the fiber attenuation, $\alpha_{\mathrm{eff}} = \alpha$, whereas $\alpha_{\mathrm{eff}} = 0$ in the case of ideal distributed amplification.

The double frequency integral in (3.89) is known to have divergence problems when the input is large. The authors in [128] proposed a modified version of (3.89) to alleviate this issue as

$$\mathbf{E}(z, t) \approx \begin{cases}\mathbf{E}_{\mathrm{L}}(z, t) + \mathbf{E}_{\mathrm{NL}}(z, t) & |\mathbf{E}_{\mathrm{NL}}(z, t)| > |\mathbf{E}_{\mathrm{L}}(z, t)|, \\ \mathbf{E}_{\mathrm{L}}(z, t)\exp\left(\frac{\mathbf{E}_{\mathrm{NL}}(z,t)}{\mathbf{E}_{\mathrm{L}}(z,t)}\right) & \text{otherwise},\end{cases} \tag{3.94}$$

where $\mathbf{E}_{\mathrm{L}}(z, t), \mathbf{E}_{\mathrm{NL}}(z, t)$ are the inverse Fourier transformations of the first and second terms of the right-hand side of (3.89), respectively. Note that the division and $\exp(\cdot)$ operations above are performed element-wise over the two elements of $\mathbf{E}$.

3.6.2 The Gaussian Noise Model

The model above succeeds to produce a simplified input–output relationship over the propagation equations in Section 3.1. However, the evaluation of the double frequency integration in (3.89) can be involved, in particular for multichannel scenarios. By assum-

ing that the nonlinear interference does not arise from data-modulated signals, but rather from a stochastic Gaussian process (hence the name), a further simplified expression for fiber propagation can be obtained. This is the so-called Gaussian Noise model and it was initially derived in 1993 [129] and then rediscovered later [130,131]. As hinted by its name, the nonlinear interference is modeled as an additive Gaussian random variable

$$\mathbf{r}_k = \mathbf{u}_k + \mathbf{w}_k \sqrt{P_{\mathrm{NLI}}}, \tag{3.95}$$

where $\mathbf{w}_k$ is zero-mean, complex, and circular-symmetric with $\mathbb{E}[\mathbf{w}_k \mathbf{w}_k^{\mathrm{H}}] = \mathbf{I}_2$. The scalar P_{NLI} is the power of the nonlinear interference, which has different expressions depending on the amplification type. In the case of lumped amplification, it becomes [132, Eqs. (16),(36)]

$$P_{\mathrm{NLI}} = N_{\mathrm{sp}}^{1+\epsilon} \frac{8}{27} \gamma^2 P^3 T^2 L_{\mathrm{eff}} (L_{\mathrm{sp}})^2 \frac{\alpha \mathrm{asinh}\left(\frac{\pi^2 |\beta_2| (N_{\mathrm{ch}}/T)^2}{2\alpha}\right)}{\pi |\beta_2|}, \tag{3.96}$$

whereas for ideal distributed amplification is [132, Eq. (44)]

$$P_{\mathrm{NLI}} = \frac{16}{27} \gamma^2 P^3 T^2 L \frac{\mathrm{asinh}\left(\frac{\pi^2 |\beta_2| L (N_{\mathrm{ch}}/T)^2}{3}\right)}{\pi |\beta_2|}, \tag{3.97}$$

where [132, Eq. (40)]

$$\epsilon = \frac{3}{10} \log_e \left(1 + \frac{6}{L_{\mathrm{sp}}} \frac{1}{\alpha \mathrm{asinh}\left(\frac{\pi^2 |\beta_2| (N_{\mathrm{ch}}/T)^2}{2\alpha}\right)}\right), \tag{3.98}$$

$\mathrm{asinh}(\cdot)$ is the hyperbolic arcsin function, P is the average power per channel, and N_{ch} is the number of wavelength-division multiplexed channels. The formulation above is a simplification of the Gaussian Noise model that assumes that the spectrum of each channel has a perfect rectangular shape of width $1/T$ and that the channel spacing is the same as the baud rate $1/T$.

As can be seen, the model above is independent of the modulation format. However, it has been shown that the nonlinear interference does depend on the modulation format and that this model overestimates the nonlinear interference [125]. Later on, the Enhanced Gaussian Noise model has been proposed [133], aiming to correct these inaccuracies.

CHAPTER 4

Digital Signal Processing

In the present chapter are discussed estimation and compensation algorithms of nonlinear effects, phase noise, SOP drift, and PMD presented in Chapter 3. First, nonlinearity mitigation based on DBP is discussed in Section 4.1. In Section 4.2, an algorithm for phase tracking is described. Thereafter, in Section 4.3, SOP and PMD recovery algorithms are presented. In Section 4.4, a tracking algorithm that accounts jointly for both phase noise and SOP drift is presented. This chapter reviews the required background knowledge required for: i) Paper C, wherein we propose a tracking algorithm (summarized in Section 4.4.2) that jointly recovers the phase and the SOP, and for ii) Papers D–G where we propose modified DBP techniques that account for PMD in the backwards propagation.

4.1 Digital Backpropagation

DBP is a digital domain technique used to compensate for fiber nonlinearity. In the absence of stochastic effects, such as ASE noise and PMD, the propagation equation can be exactly[1] reversed with inverted channel parameters $(-\beta_2, -\gamma, -\alpha)$ such that $\mathbf{E}(0,t)$ is obtained from $\mathbf{E}(L,t)$. While its main scope is to compensate for nonlinearities, DBP compensates also for CD. DBP can be used as a predistorter [22, 134], where the backwards propagation is preapplied at the transmitter before propagation. It can also be performed at the receiver [21,135] for postcompensation, or a combination of both [136,137]. In [137, 138], it is showed that splitting the compensation between the transmitter and

[1] Within the accuracy boundaries imposed by the simplified propagation equation and the numerical solver.

receiver reduces the nonlinear signal–noise interactions and leads to improved performance.

Although DBP promises many theoretical benefits, many factors contribute to its performance. We summarize a few aspects that degrade the potential gains in the following:

- Uncompensated stochastic effects. Typically the propagation equations are inverted based on a zero-forcing approach ignoring stochastic effects and only deterministic signal–signal nonlinear interaction are compensated for by DBP. By not accounting for stochastic effects, such as nonlinear signal–noise interactions [139] and PMD [140,141], which lead to mismatched signal–signal nonlinear interactions, the performance is degraded. However, these fundamental limitations have been considered and accounted for in modified DBP algorithms [59, 142, 143], including Papers D–G in this thesis. These modified versions do improve the performance compared to the conventional approach, but they are not optimal and finding the optimal algorithm is still an open problem.

- Limited nonlinearity compensation bandwidth. The principle of DBP for a multichannel transmission is not different from single channel, except for the larger bandwidth of the backpropagated signal. However, in a network scenario channels are added and dropped along the optical path and these are not available at the transmitter or receiver performing DBP. Applying DBP over a smaller bandwidth compared to the overall transmitted optical bandwidth severely reduces the potential gains [144].

- Limited DSP accuracy due to complexity restrictions. Simulations employing DBP have the advantage of floating-point accuracy and can perform the backpropagation over sufficiently many spatial steps such that the performance is maximized. However, these are currently highly prohibitive for a real-time implementation and impose strong upper bounds on the achievable gains [145,146]. This restriction relates also to the previous aspect on limited nonlinearity compensation bandwidth, where, even in a point-to-point transmission scenario with access to all transmitted channels, currently available hardware cannot perform large bandwidth DBP.

The numerical channel inversion implied by DBP can be performed based on various techniques. The SSFM discussed in Section 3.5 is arguably the most popular method. However, DBP can be also be performed using perturbation theory in the frequency domain [147,148] or in the time domain [149,150], which provide complexity gains at the expense of being less accurate compared to the SSFM.

This thesis is concerned with DBP algorithms that account for PMD. The nonlinear interference generated in the forward propagation depends on the PMD evolution along the link (cf. the second term of the right-hand side of (3.89)). However, the receiver or transmitter does not have access to this information when performing nonlinearity mitigation. Therefore, conventional algorithms ignore PMD and simply assume $\mathbf{T}(f, z) =$

$\mathbf{I}_2$ for all f and z when performing predistortion or backwards propagation, and PMD is compensated for in a lumped fashion by compensation for the entire link at once at the receiver. Since PMD is a linear effect, compensating for the entire link at once is adequate in the linear (power) regime. On the other hand, in the nonlinear regime, where nonlinearity mitigation is desired, compensating for the entire PMD at once impacts the performance and can, in extreme scenarios, completely cancel the gains of nonlinearity mitigation.

In Papers D–G, we study modified nonlinearity mitigation algorithms that take into account PMD. In Papers D, E, and F, the backpropagation is performed using the SSFM and PMD sections are inserted periodically in the backwards propagation to compensate for PMD in a distributed fashion as it naturally occurs in the forward propagation. As we do not have access to the PMD evolution along the link, the backwards-PMD sections are chosen such that once concatenated, they equal the inverse of the total PMD in the forward direction[2]. To fulfill this constraint, the sections are selected based on

- Paper D: the Nelder–Mead simplex optimization algorithm [151]. In this case, the sections will be different from each other and the solution of the optimization algorithm highly depends on the initialization of the algorithm.
- Paper E: as in Paper D, the backwards-PMD sections are initially chosen using the Nelder–Mead simplex optimization algorithm. However, this optimization has infinitely many solutions and does not guarantee the convergence to the true fiber solution, which would guarantee optimal performance. Therefore, we extend the algorithm by further optimizing the backwards-PMD sections using the estimated SNR as an objective function.
- Paper F: taking the generalized nth root operation of the matrix $\mathbf{T}(f, L)$ that models the PMD of the entire link, where n is the number of PMD sections inserted in the backwards direction. In this case all the n PMD sections are the same.

In Paper G, we take a different approach by studying a frequency-domain perturbation backpropagation algorithm that is based on Section 3.6.1. In this case, instead of compensating for PMD in a distributed fashion by inserting PMD sections in the backwards propagation, we identify and perform the integration in the second term of the right-hand side of (3.89) on only (partially) PMD-insensitive frequency domains of f_1, f_2, f_3, where PMD cancels out. In this case, we do not need to know to total PMD of the link given by $\mathbf{T}(f, L)$.

While the algorithms presented in Papers D–G have the same scope, i.e., to improve the performance of nonlinearity mitigation in the presence of PMD, they are different and can be used in different scenarios. In principle, the algorithms in Papers D and E are the only ones that can fully overcome the penalties imposed by PMD on DBP performance for any PMD realization in the forward propagation. In Paper D, for a fixed given

[2]We can access the total accumulated PMD from the channel equalizer.

forward PMD evolution, the solution of the algorithm for the PMD sections in the backward propagation is random due to the random initialization stage of the optimization algorithm. The performance can be maximized by running the optimization algorithm sufficiently many times with different initializations and choosing the one that maximizes the performance metric, such as SNR. Presumably, the solution maximizing the performance is the closest to the PMD evolution in the forward propagation. However, running many parallel instances of the optimization algorithm can be prohibitively complex. Paper E aims to bypass this obstacle by improving the selection of the backwards PMD sections obtained from the optimization algorithm by further optimizing them such that the SNR is maximized. On the other hand, the algorithm in Paper E outputs a fixed sequence of PMD sections given to the total PMD. Comparing this method to Paper D, it performs better on average, but it will not achieve the performance of DBP without PMD in the link, unless PMD evolves linearly in the forward propagation, for which it is optimal. Lastly, the algorithm in Paper F does not reverse the PMD effects in the reverse propagation. Although we do not have a direct comparison, it appears that its performance is inferior compared to the other three. However, it has the advantage that it does not require any knowledge about PMD, except for the PMD coefficient and fiber length, and it reduces the required computational complexity. Nevertheless, the approaches in Papers D, E, or F can be combined with the approach in Paper G for an improved performance.

4.2 Phase Noise Tracking

We consider the blind phase search (BPS) algorithm [64] for phase noise tracking, which is arguably one of the most popular algorithms for phase noise compensation due its universality and good performance. The BPS algorithm estimates the phase noise ϕ_k and compensates for it in a feedforward fashion in each polarization separately[3].

Assuming that the received signal is sampled at the symbol rate and it is only affected by phase noise and additive noise, the received symbols can be modeled as a special case of (3.37)

$$\mathbf{r}_k = e^{-i\phi_k}\mathbf{u}_k + \mathbf{n}_k, \tag{4.1}$$

and can seen in Fig. 2.3.

The BPS algorithm for QAM constellations can be summarized by the following steps:

1. The received signal is rotated by P_{BPS} test phases

$$z_{k,b} = r_k e^{i\phi_b}, \tag{4.2}$$

[3]It is possible to phase track jointly the two polarizations. However, in this case, the relative phase offset between the two polarization must be compensated in advance.

where r_k is one element of $\mathbf{r}_k$ and

$$\phi_b = \frac{b}{P_{\mathrm{BPS}}} \cdot \frac{\pi}{2}, \text{ for } b = 0, \ldots, P_{\mathrm{BPS}} - 1. \tag{4.3}$$

Note that the range of ϕ_b is $[0, \pi/2)$ due to the $\pi/2$ rotational symmetry of QAM constellations. For other constellations, this range must be adjusted accordingly.

2. The rotated symbols $z_{k,b}$ are fed into a decision circuit to calculate the squared Euclidean distance to the closest constellation point

$$d_{k,b}{}^2 = |z_{k,b} - \hat{z}_{k,b}|^2, \tag{4.4}$$

where $\hat{z}_{k,b}$ is the point of the constellation whose position is the closest to $z_{k,b}$.

3. The impact of the additive noise $\mathbf{n}_k$ is reduced by a moving sum over $2N_{\mathrm{BPS}} + 1$ distances

$$s_{k,b} = \sum_{n=-N_{\mathrm{BPS}}}^{N_{\mathrm{BPS}}} d_{k-n,b}{}^2. \tag{4.5}$$

4. The optimal phase angle is determined by the minimum sum of distances, and the detected symbol can be selected from $\hat{z}_{k,b}$ based on the index b of the minimum sum $s_{k,b}$ as

$$\hat{b} = \arg\min_b s_{k,b}, \tag{4.6}$$

$$\hat{z}_{k,b} = \hat{z}_{k,\hat{b}}. \tag{4.7}$$

The $\pi/2$ symmetry of QAM constellations leads to a four-fold phase ambiguity. For example, if the phase noise passes to the second quadrant, say, $\phi_k = 3\pi/4$, the algorithm will misinterpret the phase angle as being in the first quadrant $\hat{\phi}_k = 3\pi/4 - \pi/2 = \pi/4$. This will lead to catastrophic errors and can be avoided using differential coding [55, Sec. 2.6.1]. Moreover, the calculated phase angles by the algorithm are in the interval $[0, \pi/2)$, but the phase evolves and extends over this interval leading to a performance penalty. Therefore, to account for this event, a phase unwrapper is needed [152], thus introducing feedback in the algorithm. The purpose of the phase unwrapper, in this case, is to add multiples of $\pi/2$ to the estimated phase such that the maximum phase difference between two adjacent phase estimates is less than $\pi/4$.

4.3 SOP Drift and PMD Compensation

The drift of SOP and PMD are typically compensated jointly using a bank of four finite-impulse response filters. The filters can be updated based on various criteria such that

the time-varying nature of SOP and PMD is tracked. In the sections below, we present the most common option to update the filters based on the gradient descent method. First, the CMA for SOP tracking is presented, followed by its generalization the MMA, and then we discuss how to extend these algorithms to compensate for PMD.

4.3.1 Constant Modulus Algorithm

The CMA was initially developed for two-dimensional QPSK signals [153] and then applied to 4D PM-QPSK optical constellations to recover the SOP [54]. In the presence of SOP drift and additive noise, the considered signal model is

$$\mathbf{r}_k = \mathbf{J}_k \mathbf{u}_k + \mathbf{n}_k, \tag{4.8}$$

obtained by neglecting the phase noise in (3.37).

The CMA reverses the channel effects using a previous estimate of $\mathbf{J}_k^{-1}$

$$\mathbf{r}'_k = \hat{\mathbf{J}}_{k-1} \mathbf{r}_k, \tag{4.9}$$

which can be expanded into

$$\begin{pmatrix} r'_{x,k} \\ r'_{y,k} \end{pmatrix} = \begin{pmatrix} \hat{J}_{xx,k-1} & \hat{J}_{xy,k-1} \\ \hat{J}_{yx,k-1} & \hat{J}_{yy,k-1} \end{pmatrix} \begin{pmatrix} r_{x,k} \\ r_{y,k} \end{pmatrix}. \tag{4.10}$$

The purpose of the CMA is to minimize the magnitude of the error functions

$$\epsilon_x = \left| r'_{x,k} \right|^2 - E_s, \tag{4.11}$$

$$\epsilon_y = \left| r'_{y,k} \right|^2 - E_s, \tag{4.12}$$

which minimize the distance of the algorithm's output $r'_{x/y,k}$ to the circle of radius equal to the QPSK symbol energy. The update of $\hat{\mathbf{J}}_{k-1}$ is often done using the gradient descent method [154, p. 466] resulting in the update rules

$$\hat{J}_{xx,k} = \hat{J}_{xx,k-1} - \mu_{\text{CMA}} \epsilon_x r'_{x,k} (r_{x,k})^*, \tag{4.13}$$

$$\hat{J}_{xy,k} = \hat{J}_{xy,k-1} - \mu_{\text{CMA}} \epsilon_x r'_{x,k} (r_{y,k})^*, \tag{4.14}$$

$$\hat{J}_{yx,k} = \hat{J}_{yx,k-1} - \mu_{\text{CMA}} \epsilon_y r'_{y,k} (r_{x,k})^*, \tag{4.15}$$

$$\hat{J}_{yy,k} = \hat{J}_{yy,k-1} - \mu_{\text{CMA}} \epsilon_y r'_{y,k} (r_{y,k})^*, \tag{4.16}$$

where μ_{CMA} is a positive tracking step size parameter.

As suggested by its name, the algorithm was designed for constellations with constant modulus, such as phase-shift keying. Applying the algorithm to constellations with multiple modulus, such as 16-QAM, degrades the performance significantly. However,

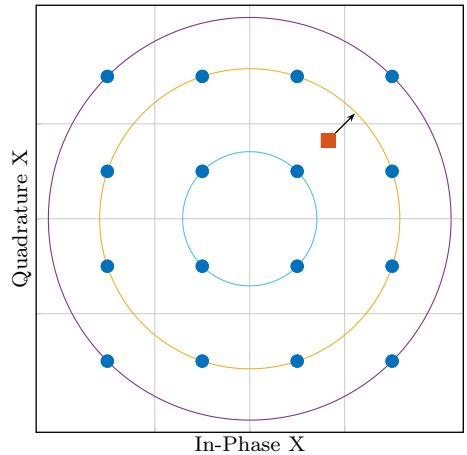


Figure 4.1. The principle of MMA illustrated for 16-QAM. The algorithm updates $\hat{\mathbf{J}}_{k-1}$ such that the error between the received symbol (red square) and its nearest constellation ring is minimized.

the CMA has been extended to multiple-modulus constellations and it is presented in the next section.

4.3.2 Multiple Modulus Algorithm

The MMA is based on the same principle as the CMA discussed earlier, except that the error function is calculated between the algorithm's output and the nearest constellation radius [155]. Due to the radius-directed error, the algorithm is known in the literature also as the radius-directed algorithm/equalizer [156].

The MMA makes first a decision on the constellation ring to which the received symbol most likely belongs, and then adapts $\hat{\mathbf{J}}_{k-1}$ such that this distance is minimized using the same updating rules as the CMA (4.13)–(4.16). The modified error functions (4.11)–(4.12) are

$$\epsilon_x = \left|r'_{x,k}\right|^2 - a_{x,k}^2, \tag{4.17}$$

$$\epsilon_y = \left|r'_{y,k}\right|^2 - a_{y,k}^2, \tag{4.18}$$

where $a_{x,k}$ ($a_{y,k}$) is the nearest radius to $r'_{x,k}$ ($r'_{y,k}$). It should be noted that the CMA is a special case of the MMA applied to a one-ring constellation. Fig. 4.1 shows an example of the algorithm's principle. In this example, only the X polarization is shown, however, the same operation is carried out in the Y polarization.

As can be seen in (4.11)–(4.12) and (4.17)–(4.18), the error functions of both the CMA and MMA are immune to the phase of the received signal. Therefore, the algorithms are

not affected by phase variations and can be applied before frequency-offset and phase-noise compensation (Section 4.2). However, an extra phase shift may be inserted when updating $\hat{\mathbf{J}}_k$ in (4.13)–(4.16). Compared to the laser phase noise, the drift of this phase shift is slow, but it can introduce a relative phase offset between the two polarizations, which must be tracked for 4D constellations.

The CMA and MMA stated above are presented as single-tap equalizers and can track only drifts of the SOP but not PMD. However, once a certain link length is exceeded, PMD becomes detrimental and must be compensated for. This can be achieved by extending $\hat{\mathbf{J}}_k$ to a bank of four finite-impulse response filters. In this case, the elements of $\hat{\mathbf{J}}_k$ and $\mathbf{r}_k$, i.e., $\hat{J}_{xx,k}, \hat{J}_{xy,k}, \hat{J}_{yx,k}, \hat{J}_{yy,k}, r_{x,k},$ and $r_{y,k}$, become vectors and $\mathbf{r}'_k$ is calculated in (4.9) as a weighted sum of a block of received symbols $\mathbf{r}_k$. By extending the algorithms to the multi-tap version, it is possible to perform, besides SOP and PMD tracking, adaptive-channel equalization such as residual-dispersion compensation, intersymbol-interference mitigation, and approximate a matched filter.

4.4 Joint Phase Noise and SOP Compensation

In general, the phase and SOP noise are compensated independently as described earlier. However, since these two phenomena can be jointly modeled as rotations of the electric field (see Section 3.3.1), it is possible to jointly compensate for them. In this section, we discuss such joint compensation, first using the Kabsch algorithm [100], and then we briefly describe the proposed algorithm in Paper C. It should be however noted that these algorithms do not compensate for PMD, which becomes detrimental for long-haul links. Similarly to [157], these algorithms can be complemented with a separate equalization stage that compensates for PMD only, after which the SOP is corrected by the algorithms below.

4.4.1 Kabsch Algorithm

The Kabsch algorithm [100] addresses jointly the phase and SOP tracking by estimating $\mathbf{R}_k$ in the 4D space (see 4D Real Description in Section 3.3.1). The tracking is carried out over blocks of N_{Kab} PM symbols

$$\mathbf{V}_{\mathbf{u}_l} = [\mathbf{v}_{\mathbf{u}_k}, \mathbf{v}_{\mathbf{u}_{k+1}}, \ldots, \mathbf{v}_{\mathbf{u}_{k+N_{\text{Kab}}-1}}], \tag{4.19}$$

under the assumption that $\mathbf{R}_k$ does not change (significantly) over the block l.

The algorithm can be summarized by the following steps:

1. The received block of symbols $\mathbf{V}_{\mathbf{r}_l}$ is derotated using the inverse of a previous estimate of $\mathbf{R}_l^{-1}$

$$\mathbf{V}'_{\mathbf{r}_l} = \hat{\mathbf{R}}_{l-1}\mathbf{V}_{\mathbf{r}_l}. \tag{4.20}$$

2. The estimated transmitted block of symbols $\mathbf{V}_{\hat{\mathbf{u}}_l}$ is obtained by the minimum sum of Euclidean distances between $\mathbf{V}'_{\mathbf{r}_l}$ and a block of constellation points.

3. The matrix $\hat{\mathbf{R}}_l$ is updated as

$$\hat{\mathbf{R}}_l = \mathbf{U}_l \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & s \end{pmatrix} \mathbf{W}_l^{\mathrm{T}}, \tag{4.21}$$

where $s = \mathrm{sign}(\det(\mathbf{U}_l \mathbf{W}_l^{\mathrm{T}}))$, and $\mathbf{U}_l$ and $\mathbf{W}_l$ are the left- and right-singular vectors of $\mathbf{C}_l = \mathbf{V}_{\hat{\mathbf{u}}_l} \mathbf{V}_{\mathbf{r}_l}^{\mathrm{T}}$, i.e., $\mathbf{U}_l \Sigma_l \mathbf{W}_l^{\mathrm{T}} = \mathbf{C}_l$ is the singular-value decomposition [158, p. 35] of $\mathbf{C}_l$.

The Kabsch algorithm can be applied to arbitrary constellations by only changing the decision stage in step 2. This algorithm, as the BPS, suffers from ambiguities, and therefore differential coding must be applied to ensure reliability.

4.4.2 Proposed Algorithm

The proposed algorithm in Paper C recovers the carrier phase and SOP for arbitrary modulation formats using a non-data-aided decision-directed architecture. Similarly to the Kabsch algorithm, the proposed algorithm operates jointly on both polarizations, but in a symbol-by-symbol fashion without averaging blocks. The algorithm has been developed based on the channel model proposed in Paper A and is the first model-based SOP tracking algorithm. Model-based algorithms have a restricted flexibility, and therefore fewer DOFs to adjust, resulting in a more efficient impairment cancellation, rather than scanning over a larger domain in order to find the optimal setup. At similar or better performance, the proposed algorithm offers a good trade-off between complexity and performance compared to state-of-the-art algorithms, regardless of the modulation format. The details of the algorithm can be found in Paper C.

CHAPTER 5

Contributions and Future Work

This chapter summarizes the contributions of the appended papers and discusses potential future research connected to the topics addressed in this thesis.

5.1 Paper A

"**Polarization Drift Channel Model for Coherent Fibre-Optic Systems**"

In this paper, we propose a theoretical framework to model the dynamical changes of the SOP in coherent fiber-optic systems, based on a generalization of the one-dimensional phase-noise random walk to higher dimensions, accounting for the random polarization drift. The model is stated in the Jones, Stokes, and real 4D formalisms and can be easily combined with other transmission impairments to form a complete channel model. The proposed polarization drift model is the first of its kind and will likely be useful in many areas of photonics where stochastic polarization fluctuation is an issue. Based on this model, better polarization tracking algorithms can be found and more accurate simulations that reflect fiber behavior closely can be performed in order to quantify system performance.

Contributions: CBC designed and analyzed the model, carried out simulations, and wrote the paper. MK and EA formulated the problem and contributed to the analysis. PJ provided mathematical expertise and interpretation of the results. All authors reviewed and revised the paper.

Context: Section 3.3.

5.2 Paper B

"**Temporal Stochastic Channel Model for Absolute Polarization State and Polarization-Mode Dispersion**"

In Paper B, we extend our channel model proposed in Paper A to account for PMD and its temporal variation. The model is in the discrete-time domain and models the temporal drift of the absolute polarization state and PMD. The autocorrelation function of the Jones matrix in frequency and time is derived and validated with experimental data. The model can be used in simulations to test and develop DSP for coherent receivers, such as, polarization-tracking or nonlinearity mitigation, where PMD is an issue.

Contributions: CBC proposed, designed, and analyzed the model, carried out simulations, and wrote the paper. MK, EA, and PJ contributed to the analysis and provided mathematical expertise. All authors reviewed and revised the paper.

Context: Section 3.4.

5.3 Paper C

"**Modulation Format Independent Joint Polarization and Phase Tracking for Coherent Receivers**"

Based on the model proposed in Paper A, we propose an algorithm to recover jointly the carrier phase and SOP for arbitrary modulation formats. The algorithm uses a non-data aided, decision-directed architecture, hence zero overhead, and operates jointly on both polarizations. The performance and complexity of the algorithm is investigated by comparing it with state-of-the-art algorithms for different modulation formats. The proposed algorithm performs similarly or better than state-of-the-art algorithms presented in Sections 4.2–4.4.1 and provides a good trade-off between complexity and performance regardless of the modulation format. High performance and fast convergence rate, for any modulation format at low complexity, make the algorithm a strong candidate for future elastic optical systems, where the modulation format can be changed dynamically during transmission to accommodate for various channel and network conditions.

Contributions: CBC analyzed the algorithm, carried out simulations, and wrote the paper. EA proposed the algorithm and contributed to the analysis. MK and PJ provided mathematical expertise and interpretation of the results. All authors reviewed and revised the paper.

Context: Section 4.4.

5.4 Paper D

"**Polarization-Mode Dispersion Aware Digital Backpropagation**"

In this paper, we study a modified SSFM-based DBP algorithm that accounts for PMD. Based on the accumulated PMD at the receiver, the algorithm distributively compensates for PMD in the reverse propagation and outperforms the conventional approach by up to 2.1 dB SNR gains. The PMD sections in the reverse propagation are selected based on the Nelder–Mead simplex optimization algorithm such that concatenated they equal the inverse of the PMD accumulated over the link. We compare the proposed algorithm with the conventional algorithm for various PMD coefficients and investigate how the initialization of the optimization algorithm impacts the performance.

Contributions: CBC designed and analyzed the algorithm, carried out simulations, and wrote the paper. DL formulated the problem and contributed to the analysis. GL, MK, and SJS provided mathematical expertise and interpretation of the results. EA and PB contributed to the analysis and interpretation of the results. All authors reviewed and revised the paper.

Context: Section 4.1.

5.5 Paper E

"**A PMD-adaptive DBP Receiver Based on SNR Optimization**"

In this paper, the algorithm in Paper D is extended such that it further optimizes the backwards-PMD sections in order to maximize the estimated SNR. After the initial selection of the PMD sections in the reverse propagation based on the Nelder–Mead simplex optimization algorithm, the sections are continuously updated using the same optimization algorithm such that the SNR is maximized. In the studied cases based on up to 2000 iterations, the SNR is monotonically increasing with the number of iterations. However, a further study is required to assess the convexity of the problem and potential singularities.

Contributions: GL designed and analyzed the algorithm, carried out simulations, and wrote the paper. CBC contributed to the design of the algorithm, provided mathematical expertise, interpretation of the results, and contributed to the analysis. PB contributed to the analysis and interpretation of the results. All authors reviewed and revised the paper.

Context: Section 4.1.

5.6 Paper F

"**Digital Backpropagation Accounting for Polarization-Mode Dispersion**"

In this paper, we study the same problem as in Papers D and E of a modified SSFM-based DBP algorithm that takes into account PMD. Instead of selecting the reverse

PMD sections based on an optimization algorithm, we choose them analytically based on a generalized nth-root operation of the matrix that models the accumulated PMD of the entire link. We show that this matrix can be accurately recovered from the channel equalizers at negligible penalty. Accounting for nonlinear polarization-related interactions in the modified DBP algorithm, we obtain average SNR gains over conventional DBP of 1.1 dB for transmission over 1000 km for both 1-channel and 7-channel full-field backpropagation. We also examine the effects of the stochastic nature of PMD on the performance and how the performance changes as a function of the number of reverse PMD sections.

Contributions: CBC analyzed the algorithm, carried out simulations, and wrote the paper. CBC and MK designed the algorithm. GL, DL, MK, and SJS provided mathematical expertise and interpretation of the results. EA and PB contributed to the analysis and interpretation of the results. All authors reviewed and revised the paper.

Context: Section 4.1.

5.7 Paper G

"**Volterra Series Digital Backpropagation Accounting for PMD**"

In this paper, we study a perturbative DBP algorithm based on Volterra series that accounts for PMD by only considering nonlinear terms that are (partially) PMD-insensitive. This restriction leads to both performance enhancement and substantial complexity reduction in the high-PMD regime. The discrimination between PMD-sensitive and PMD-insensitive terms is based on an assumption that approximates the spectral autocorrelation function of the PMD-Jones matrix as a step function. This approximation allows us to distinguish between different types of four-wave mixing products and to carefully select the integration domain in (3.89). Considering a 1000-km link with strong PMD, $D_{\mathrm{PMD}} = 0.5\ \mathrm{ps}/\sqrt{\mathrm{km}}$, the proposed algorithm provides 0.4 dB SNR improvement compared to conventional Volterra series DBP, at a reduced complexity by 42%.

Contributions: CBC designed and analyzed the algorithm, carried out simulations, and wrote the paper. RD formulated the problem and the solution method, provided mathematical expertise, and contributed to the analysis and interpretation of the results. Both CBC and RD reviewed and revised the paper.

Context: Section 4.1.

5.8 Future Work

An interesting direction for future work is to adapt the channel models presented in Papers A and B for multi-mode dispersion. Conceptually, the phenomena of PMD and modal dispersion are very similar. In fact, the two orthogonal polarizations are two different modes. However, the difference in modeling these two phenomena consists in

the number of modes to model and in the coupling strength between the modes. In the case of multi-mode dispersion, different modes interact differently with the other modes; therefore leading to a nonuniform coupling matrix. Moreover, the temporal dynamics of modal dispersion have many different sources that vary at different time scales and have to be accounted for [159], similarly to how phase noise varies at a much faster pace than the SOP drift.

The algorithm in Paper C is limited to SOP and phase-noise tracking and does not compensate for PMD. However, long-haul links suffer from PMD, which becomes detrimental once a certain link length is exceeded. The overcome this, similarly to [157], the algorithm in Paper C can be coupled with a separate equalization stage that compensates for PMD only, after which the SOP is tracked with the proposed algorithm. In this case, the tracking of PMD and absolute SOP are decoupled, and, since the two phenomena drift at different time scales, they can be tracked at different rates, leading to complexity savings.

A possible interesting research direction regarding DBP in the presence of PMD would be to investigate an optimal approach to this problem. This can be achieved, e.g., based on the factor-graphs framework, which was used in [59] to derive an improved DBP that accounts for signal–noise interactions. Such a contribution would be useful for i) in establishing the fundamental limits and ii) comparing existing algorithms (like the ones present in this thesis) with optimal detection. Another interesting direction would be to apply the concepts presented in Papers D–G on DBP for multi-mode systems. When performing nonlinearity mitigation in multi-mode systems, the performance is degraded if multi-mode dispersion is not properly accounted for. Due to the similarities between PMD and multi-mode dispersion, DBP algorithms can be modified as in Papers D–G to account for multi-mode dispersion.

Bibliography

[1] CISCO, "The Zettabyte Era: Trends and Analysis, white paper," https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.html, Jun. 2017, accessed: 2017-10-30.

[2] K. C. Kao and G. A. Hockham, "Dielectric–fibre surface waveguides for optical frequencies," *Proceedings of the Institution of Electrical Engineers*, vol. 113, no. 7, pp. 1151–1158, Jul. 1966.

[3] F. P. Kapron, D. B. Keck, and R. D. Maurer, "Radiation losses in glass optical waveguides," *Applied Physics Letters*, vol. 17, no. 10, pp. 423–425, Nov. 1970.

[4] S. Abbott, "Review of 20 years of undersea optical fiber transmission system development and deployment since TAT-8," in *Proc. of European Conference on Optical Communication (ECOC)*, Brussels, Belgium, Sept. 2008, p. Mo.4.E.1.

[5] R. Mears, L. Reekie, I. Jauncey, and D. Payne, "Low-noise erbium-doped fibre amplifier operating at 1.54 μm," *Electronics Letters*, vol. 23, no. 19, pp. 1026–1028, Sept. 1987.

[6] E. Desurvire, J. R. Simpson, and P. C. Becker, "High-gain erbium-doped traveling-wave fiber amplifier," *Optics Letters*, vol. 12, no. 11, pp. 888–890, Nov. 1987.

[7] R. W. Tkach, "Scaling optical communications for the next decade and beyond," *Bell Labs Technical Journal*, vol. 14, no. 4, pp. 3–10, Feb. 2010.

[8] P. J. Winzer and R.-J. Essiambre, "Advanced optical modulation formats," *Proceedings of the IEEE*, vol. 94, no. 5, pp. 952–985, Jun. 2006.

[9] R. A. Linke and A. H. Gnauck, "High-capacity coherent lightwave systems," *Journal of Lightwave Technology*, vol. 6, no. 11, pp. 1750–1769, Nov. 1988.

[10] H. Sun, K.-T. Wu, and K. Roberts, "Real-time measurements of a 40 Gb/s coherent system," *Optics Express*, vol. 16, no. 2, pp. 873–879, Jan. 2008.

[11] E. Agrell and M. Karlsson, "Power-efficient modulation formats in coherent transmission systems," *Journal of Lightwave Technology*, vol. 27, no. 22, pp. 5115–5126, Nov. 2009.

[12] S. Iano, T. Sato, S. Sentsui, T. Kuroha, and Y. Nishimura, "Multicore optical fiber," in *Proc. of Optical Fiber Communication Conference (OFC)*, Washington, D.C., Mar. 1979, p. WB1.

[13] J. Sakaguchi, Y. Awaji, N. Wada, A. Kanno, T. Kawanishi, T. Hayashi, T. Taru, T. Kobayashi, and M. Watanabe, "109-Tb/s (7x97x172-Gb/s SDM/WDM/PDM) QPSK transmission through 16.8-km homogeneous multi-core fiber," in *Proc. of Optical Fiber Communication Conference (OFC)*, Los Angeles, CA, Mar. 2011, p. PDPB.6.

[14] D. J. Richardson, "New optical fibres for high–capacity optical communications," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2062, Jan. 2016.

[15] Y. Yamamoto, Y. Kawaguchi, and M. Hirano, "Low-loss and low-nonlinearity pure-silica-core fiber for C- and L-band broadband transmission," *Journal of Lightwave Technology*, vol. 34, no. 2, pp. 321–326, Jan. 2016.

[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, May 2013.

[17] S. L. Jansen, D. van den Borne, P. M. Krummrich, S. Sälter, G.-D. Khoe, and H. de Waardt, "Long-haul DWDM transmission systems employing optical phase conjugation," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 12, no. 4, pp. 505–520, Jul. 2006.

[18] Y. Gao, J. C. Cartledge, A. S. Karar, S. S.-H. Yam, M. O'Sullivan, C. Laperle, A. Borowiec, and K. Roberts, "Reducing the complexity of perturbation based nonlinearity pre-compensation using symmetric EDC and pulse shaping," *Optics Express*, vol. 22, no. 2, pp. 1209–1219, Jan. 2014.

[19] M. Secondini and E. Forestieri, "On XPM mitigation in WDM fiber-optic systems," *IEEE Photonics Technology Letters*, vol. 26, no. 22, pp. 2252–2255, Nov. 2014.

[20] R. Dar, O. Geller, M. Feder, A. Mecozzi, and M. Shtaif, "Mitigation of inter-channel nonlinear interference in WDM systems," in *Proc. of European Conference on Optical Communication (ECOC)*, Cannes, France, Sept 2014, p. P.5.6.

[21] R.-J. Essiambre and P. J. Winzer, "Fibre nonlinearities in electronically predistorted transmission," in *Proc. of European Conference on Optical Communication (ECOC)*, vol. 2, Glasgow, UK, Sept. 2005, pp. 191–192.

[22] 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, Sept. 2006.

[23] E. Ip, "Nonlinear compensation using backpropagation for polarization-multiplexed transmission," *Journal of Lightwave Technology*, vol. 28, no. 6, pp. 939–951, Mar. 2010.

[24] 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, Oct. 2008.

[25] R. Dar and P. J. Winzer, "Nonlinear interference mitigation: methods and potential gain," *Journal of Lightwave Technology*, vol. 35, no. 4, pp. 903–930, Feb. 2017.

[26] M. Karlsson, N. Alic, S. Chandrsekhar, and A. Mecozzi, "Feature issue introduction: nonlinearity mitigation for coherent transmission systems," *Optics Express*, vol. 25, no. 4, pp. 4552–4553, Feb. 2017.

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

[28] J. R. Barry, *Wireless Infrared Communications*. New York, NY: Springer, 1994.

[29] R. Otte, L. P. de Jong, and A. H. van Roermund, *Low-Power Wireless Infrared Communications*. New York, NY: Springer, 1999.

[30] S. Hranilovic, *Wireless Optical Communication Systems*. Boston, MA: Springer, 2005.

[31] H. Willebrand and B. S. Ghuman, *Free Space Optics: Enabling Optical Connectivity in Today's Networks*. Indianapolis, IN: SAMS publishing, 2002.

[32] O. Bouchet, H. Sizun, C. Boisrobert, and F. De Fornel, *Free-Space Optics: Propagation and Communication*. London, UK: ISTE, 2006.

[33] S. Arnon, J. R. Barry, G. Karagiannidis, R. Schober, and M. Uysal, *Advanced Optical Wireless Communication Systems*. Cambridge, UK: Cambridge University Press, 2012.

[34] J. Tyndall, "On some phenomena connected with the motion of liquids," *Proceedings of the Royal Institution of Great Britain*, vol. 1, p. 446, May 1854.

[35] G. P. Agrawal, *Fiber-Optic Communication Systems*, 4th ed. Hoboken, NJ: John Wiley & Sons, 2010.

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

[37] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *Journal of Lightwave Technology*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[38] M. N. Islam, "Raman amplifiers for telecommunications," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 8, no. 3, pp. 548–559, May 2002.

[39] A. Leven, N. Kaneda, U.-V. Koc, and Y.-K. Chen, "Frequency estimation in intradyne reception," *IEEE Photonics Technology Letters*, vol. 19, no. 6, pp. 366–368, Mar. 2007.

[40] E. Ip, A. P. T. Lau, D. J. F. Barros, and J. M. Kahn, "Coherent detection in optical fiber systems," *Optics Express*, vol. 16, no. 2, pp. 753–791, Jan. 2008.

[41] T. Miyazaki and F. Kubota, "PSK self-homodyne detection using a pilot carrier for multibit/symbol transmission with inverse-RZ signal," *IEEE Photonics Technology Letters*, vol. 17, no. 6, pp. 1334–1336, Jun. 2005.

[42] F. Derr, "Coherent optical QPSK intradyne system: concept and digital receiver realization," *Journal of Lightwave Technology*, vol. 10, no. 9, pp. 1290–1296, Sept. 1992.

[43] K. Petermann, *Laser Diode Modulation and Noise*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1988.

[44] G. P. Agrawal, *Nonlinear Fiber Optics*, 4th ed. Burlington, MA: Academic Press, 2007.

[45] R. Kashyap, *Fiber Bragg Gratings*. San Diego, CA: Academic Press, 1999.

[46] S. J. Savory, G. Gavioli, R. I. Killey, and P. Bayvel, "Electronic compensation of chromatic dispersion using a digital coherent receiver," *Optics Express*, vol. 15, no. 5, pp. 2120–2126, Mar. 2007.

[47] H. Kogelnik, R. M. Jopson, and L. E. Nelson, "Polarization-mode dispersion," in *Optical Fiber Telecommunications IV-B: Systems and Impairments*, I. Kaminow and T. Li, Eds. San Diego, CA: Academic Press, 2002, ch. 15, pp. 725–861.

[48] C. Headley and G. Agrawal, *Raman Amplification in Fiber Optical Communication Systems*. Boston, MA: Academic Press, 2005.

[49] S. P. Smith, F. Zarinetchi, and S. Ezekiel, "Narrow-linewidth stimulated brillouin fiber laser and applications," *Optics letters*, vol. 16, no. 6, pp. 393–395, Mar. 1991.

[50] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proceedings of the IEEE*, vol. 85, no. 2, pp. 265–298, Feb. 1997.

[51] A. Larsson, "Advances in VCSELs for communication and sensing," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 17, no. 6, pp. 1552–1567, Nov.–Dec. 2011.

[52] C. B. Czegledi, M. R. Khanzadi, and E. Agrell, "Bandlimited power-efficient signaling and pulse design for intensity modulation," *IEEE Transactions on Communications*, vol. 62, no. 9, pp. 3274–3284, Sept. 2014.

[53] A. M. Joshi, S. Datta, and A. Crawford, "Next-gen communications fiber: multilevel modulation formats push capacities beyond 100 Gbit/s," *Laser Focus World*, vol. 48, no. 2, pp. 58–63, Feb. 2012.

[54] 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, Sept.–Oct. 2010.

[55] M. Seimetz, *High-Order Modulation for Optical Fiber Transmission.* Heidelberg, Germany: Springer, 2009.

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

[57] A. Sheikh, C. Fougstedt, P. Johannisson, A. Graell i Amat, and P. Larsson-Edefors, "Dispersion compensation filter design optimized for robustness and power efficiency," in *Proc. of Signal Processing in Photonic Communication (SPPCom)*, Boston, MA, Jul. 2015, p. SpT3D.2.

[58] C. Fougstedt, A. Sheikh, P. Johannisson, A. Graell i Amat, and P. Larsson-Edefors, "Power-efficient time-domain dispersion compensation using optimized FIR filter implementation," in *Proc. of Signal Processing in Photonic Communication (SPPCom)*, Boston, MA, Jul. 2015, p. SpT3D.3.

[59] N. V. Irukulapati, H. Wymeersch, P. Johannisson, and E. Agrell, "Stochastic digital backpropagation," *IEEE Transactions on Communications*, vol. 62, no. 11, pp. 3956–3968, Nov. 2014.

[60] M. Visintin, G. Bosco, P. Poggiolini, and F. Forghieri, "Adaptive digital equalization in optical coherent receivers with Stokes-space update algorithm," *Journal of Lightwave Technology*, vol. 32, no. 24, pp. 4759–4767, Dec. 2014.

[61] N. J. Muga and A. N. Pinto, "Adaptive 3-D Stokes space-based polarization demultiplexing algorithm," *Journal of Lightwave Technology*, vol. 32, no. 19, pp. 3290–3298, Oct. 2014.

[62] M. Selmi, Y. Jaouen, and P. Ciblat, "Accurate digital frequency offset estimator for coherent PolMux QAM transmission systems," in *Proc. of European Conference on Optical Communication (ECOC)*, Vienna, Austria, Sept. 2009, p. P3.08.

[63] A. Viterbi and A. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission," *IEEE Transactions on Information Theory*, vol. 29, no. 4, pp. 543–551, Jul. 1983.

[64] T. Pfau, S. Hoffmann, and R. Noé, "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for M-QAM constellations," *Journal of Lightwave Technology*, vol. 27, no. 8, pp. 989–999, Apr. 2009.

[65] B. P. Smith and F. R. Kschischang, "Future prospects for FEC in fiber-optic communications," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 5, pp. 1245–1257, Sept. 2010.

[66] L. Schmalen, A. J. de Lind van Wijngaarden, and S. ten Brink, "Forward error correction in optical core and optical access networks," *Bell Labs Technical Journal*, vol. 18, no. 3, pp. 39–66, Dec. 2013.

[67] D. A. Morero, M. A. Castrillón, A. Aguirre, M. R. Hueda, and O. E. Agazzi, "Design trade-offs and challenges in practical coherent optical transceiver implementations," *Journal of Lightwave Technology*, vol. 34, no. 1, pp. 121–136, Jan. 2016.

[68] A. F. Alfredsson, R. Krishnan, and E. Agrell, "Joint-polarization phase-noise estimation and symbol detection for optical coherent receivers," *Journal of Lightwave Technology*, vol. 34, no. 18, pp. 4394–4405, Sept. 2016.

[69] C. R. Menyuk and B. S. Marks, "Interaction of polarization mode dispersion and nonlinearity in optical fiber transmission systems," *Journal of Lightwave Technology*, vol. 24, no. 7, pp. 2806–2826, Jul. 2006.

[70] A. Shabat and V. Zakharov, "Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media," *Soviet physics JETP*, vol. 34, no. 1, p. 62, Jan. 1972.

[71] M. I. Yousefi and F. R. Kschischang, "Information transmission using the nonlinear fourier transform, part i: Mathematical tools," *IEEE Transactions on Information Theory*, vol. 60, no. 7, pp. 4312–4328, Jul. 2014.

[72] A. Mecozzi, "Limits to long-haul coherent transmission set by the Kerr nonlinearity and noise of the in-line amplifiers," *Journal of Lightwave Technology*, vol. 12, no. 11, pp. 1993–2000, Nov. 1994.

[73] H. A. Haus, "Quantum noise in a solitonlike repeater system," *Journal of the Optical Society of America B*, vol. 8, no. 5, pp. 1122–1126, May 1991.

[74] J. P. Gordon, W. H. Louisell, and L. R. Walker, "Quantum fluctuations and noise in parametric processes. II," *Physical Review*, vol. 129, no. 1, pp. 481–485, Jan. 1963.

[75] C. R. Menyuk, "Nonlinear pulse propagation in birefringent optical fibers," *IEEE Journal of Quantum electronics*, vol. 23, no. 2, pp. 174–176, Feb. 1987.

[76] D. Marcuse, C. R. Manyuk, and P. K. A. 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. 1735–1746, Sept. 1997.

[77] 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, pp. 1231–1233, Aug. 1991.

[78] ——, "Effects of randomly varying birefringence on soliton interactions in optical fibers," *Optics Letters*, vol. 16, no. 22, pp. 1735–1737, Nov. 1991.

[79] P. K. A. Wai and C. R. Menyuk, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence," *Journal of Lightwave Technology*, vol. 14, no. 2, pp. 148–157, Feb. 1996.

[80] M. Karlsson, "Four-dimensional rotations in coherent optical communications," *Journal of Lightwave Technology*, vol. 32, no. 6, pp. 1246–1257, Mar. 2014.

[81] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, Jun. 2003.

[82] N. J. Frigo, "A generalized geometrical representation of coupled mode theory," *IEEE Journal of Quantum Electronics*, vol. 22, no. 11, pp. 2131–2140, Nov. 1986.

[83] J. P. Gordon and H. Kogelnik, "PMD fundamentals: polarization mode dispersion in optical fibers," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 9, pp. 4541–4550, Apr. 2000.

[84] R. Bellman, *Introduction to Matrix Analysis*. New York, NY: McGraw-Hill, 1960.

[85] S. Betti, F. Curti, G. De Marchis, and E. Iannone, "A novel multilevel coherent optical system: 4-quadrature signaling," *Journal of Lightwave Technology*, vol. 9, no. 4, pp. 514–523, Apr. 1991.

[86] R. Cusani, E. Iannone, A. M. Salonico, and M. Todaro, "An efficient multilevel coherent optical system: M-4Q-QAM," *Journal of Lightwave Technology*, vol. 10, no. 6, pp. 777–786, Jun. 1992.

[87] G. G. Stokes, "On the composition and resolution of streams of polarized light from different sources," *Transactions of the Cambridge Philosophical Society*, vol. 9, pp. 399–416, Jan. 1851.

[88] J. N. Damask, *Polarization Optics in Telecommunications*. New York, NY: Springer, 2005.

[89] M. Tur, B. Moslehi, and J. Goodman, "Theory of laser phase noise in recirculating fiber-optic delay lines," *Journal of Lightwave Technology*, vol. 3, no. 1, pp. 20–31, Feb. 1985.

[90] M. Karlsson, J. Brentel, and P. A. Andrekson, "Long-term measurement of PMD and polarization drift in installed fibers," *Journal of Lightwave Technology*, vol. 18, no. 7, pp. 941–951, Jul. 2000.

[91] K. Kikuchi, "Polarization-demultiplexing algorithm in the digital coherent receiver," in *Digest of the IEEE/LEOS Summer Topical Meetings*, Acapulco, Mexico, Jul. 2008, p. MC22.

[92] L. Liu, Z. Tao, W. Yan, S. Oda, T. Hoshida, and J. C. Rasmussen, "Initial tap setup of constant modulus algorithm for polarization de-multiplexing in optical coherent receivers," in *Proc. of Optical Fiber Communication Conference (OFC)*, San Diego, CA, Mar. 2009, p. OMT2.

[93] I. Roudas, A. Vgenis, C. S. Petrou, D. Toumpakaris, J. Hurley, M. Sauer, J. Downie, Y. Mauro, and S. Raghavan, "Optimal polarization demultiplexing for coherent optical communications systems," *Journal of Lightwave Technology*, vol. 28, no. 7, pp. 1121–1134, Apr. 2010.

[94] P. Johannisson, H. Wymeersch, M. Sjödin, A. S. Tan, E. Agrell, P. A. Andrekson, and M. Karlsson, "Convergence comparison of the CMA and ICA for blind polarization demultiplexing," *Journal of Optical Communications and Networking*, vol. 3, no. 6, pp. 493–501, Jun. 2011.

[95] K. Ogaki, M. Nakada, Y. Nagao, and K. Nishijima, "Fluctuation differences in the principal states of polarization in aerial and buried cables," in *Proc. of Optical Fiber Communication Conference (OFC)*, Atlanta, GA, Mar. 2003, p. MF13.

[96] M. Karlsson, C. B. Czegledi, and E. Agrell, "Coherent transmission channels as 4d rotations," in *Proc. of Signal Processing in Photonic Communication (SPPCom)*, Boston, MA, Jul. 2015, p. SpM3E.2.

[97] H. Rummler, "On the distribution of rotation angles how great is the mean rotation angle of a random rotation?" *The Mathematical Intelligencer*, vol. 24, no. 4, pp. 6–11, Sept. 2002.

[98] P. Johannisson, M. Sjödin, M. Karlsson, H. Wymeersch, E. Agrell, and P. A. Andrekson, "Modified constant modulus algorithm for polarization-switched QPSK," *Optics Express*, vol. 19, no. 8, pp. 7734–7741, Apr. 2011.

[99] S. J. Savory, "Digital filters for coherent optical receivers," *Optics Express*, vol. 16, no. 2, pp. 804–817, Jan. 2008.

[100] H. Louchet, K. Kuzmin, and A. Richter, "Joint carrier-phase and polarization rotation recovery for arbitrary signal constellations," *IEEE Photonics Technology Letters*, vol. 26, no. 9, pp. 922–924, May 2014.

[101] F. Heismann and K. L. Tokuda, "Polarization-independent electro-optic depolarizer," *Optics Letters*, vol. 20, no. 9, pp. 1008–1010, May 1995.

[102] C. B. Czegledi, E. Agrell, and M. Karlsson, "Symbol-by-symbol joint polarization and phase tracking in coherent receivers," in *Proc. of Optical Fiber Communication Conference (OFC)*, Los Angeles, CA, Mar. 2015, p. W1E.3.

[103] A. Judy, "Improved PMD stability in optical fibers and cables," in *Proc. of the 43rd International Wire & Cable Symposium*, Atlanta, GA, Nov. 1994, pp. 658–664.

[104] C. H. Prola, J. A. Pereira da Silva, A. O. Dal Forno, R. Passy, J. P. von der Weid, and N. Gisin, "PMD emulators and signal distortion in 2.48-Gb/s IM-DD lightwave systems," *IEEE Photonics Technology Letters*, vol. 9, no. 6, pp. 842–844, Jun. 1997.

[105] C. Poole and D. Favin, "Polarization-mode dispersion measurements based on transmission spectra through a polarizer," *Journal of Lightwave Technology*, vol. 12, no. 6, pp. 917–929, Jun. 1994.

[106] 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, Dec. 2001.

[107] 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, Mar. 2000.

[108] P. Krummrich, E.-D. Schmidt, W. Weiershausen, and A. Mattheus, "Field trial results on statistics of fast polarization changes in long haul WDM transmission systems," in *Proc. of Optical Fiber Communication Conference (OFC)*, Anaheim, CA, Mar. 2005, p. OThT6.

[109] P. Krummirich and K. Kotten, "Extremely fast (microsecond timescale) polarization changes in high speed long haul WDM transmission systems," in *Proc. of Optical Fiber Communication Conference (OFC)*, Los Angeles, CA, Feb. 2004, p. FI3.

[110] Y. Akasaka, X. Wang, A. Lee, M. Davy, and T. Naito, "PMD measurement of 160-km buried fiber with low DGD," in *Proc. of Optical Fiber Communication Conference (OFC)*, San Diego, CA, Feb. 2008, p. NThE.3.

[111] M. Brodsky, M. Boroditsky, P. Magill, N. Frigo, and M. Tur, "A "hinge" model for the temporal dynamics of polarization mode dispersion," in *Proc. of the 17th Annual Meeting of the IEEE Lasers and Electro-Optics Society (LEOS)*, Rio Grande, Puerto Rico, Nov. 2004, p. MJ5.

[112] M. Brodsky, N. J. Frigo, M. Boroditsky, and M. Tur, "Polarization mode dispersion of installed fibers," *Journal of Lightwave Technology*, vol. 24, no. 12, pp. 4584–4599, Dec. 2006.

[113] G. Kramer, M. I. Yousefi, and F. R. Kschischang, "Upper bound on the capacity of a cascade of nonlinear and noisy channels," in *Proc. of IEEE Information Theory Workshop*, Jerusalem, Israel, Apr.–May 2015.

[114] R. H. Hardin and F. Tappert, "Applications of the split-step Fourier method to the numerical solution of nonlinear and variable coefficient wave equation," *SIAM Review Chronicle*, vol. 15, no. 2, p. 423, Apr. 1973.

[115] T. R. Taha and M. J. Ablowitz, "Analytical and numerical aspects of certain nonlinear evolution equations. II. Numerical, nonlinear Schrödinger equation," *Journal of Computational Physics*, vol. 55, no. 2, pp. 203–230, Aug. 1984.

[116] J. Shao, X. Liang, and S. Kumar, "Comparison of split-step Fourier schemes for simulating fiber optic communication systems," *IEEE Photonics Journal*, vol. 6, no. 4, Aug. 2014.

[117] 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, Jan. 2003.

[118] G. Bosco, A. Carena, V. Curri, R. Gaudino, P. Poggiolini, and S. Benedetto, "Suppression of spurious tones induced by the split-step method in fiber systems simulation," *IEEE Photonics Technology Letters*, vol. 12, no. 5, pp. 489–491, May 2000.

[119] Q. Zhang and M. I. Hayee, "Symmetrized split-step Fourier scheme to control global simulation accuracy in fiber-optic communication systems," *Journal of Lightwave Technology*, vol. 26, no. 2, pp. 302–316, Jan. 2008.

[120] E. Agrell, "Capacity bounds in optical communications," in *Proc. of European Conference on Optical Communication (ECOC)*, Göteborg, Sweden, Sept. 2017, p. M.2.D.1.

[121] E. Agrell, G. Durisi, and P. Johannisson, "Information-theory-friendly models for fiber-optic channels: A primer," in *Proc. of IEEE Information Theory Workshop*, Jerusalem, Israel, Apr.–May 2015.

[122] K. V. Peddanarappagari and M. Brandt-Pearce, "Volterra series transfer function of single-mode fibers," *Journal of Lightwave Technology*, vol. 15, no. 12, pp. 2232–2241, Dec. 1997.

[123] A. Vannucci and A. Bononi, "Statistical characterization of the Jones matrix of long fibers affected by polarization mode dispersion (PMD)," *Journal of Lightwave Technology*, vol. 20, no. 5, pp. 811–821, May 2002.

[124] A. Mecozzi and R.-J. Essiambre, "Nonlinear Shannon limit in pseudolinear coherent systems," *Journal of Lightwave Technology*, vol. 30, no. 12, pp. 2011–2024, Jun. 2012.

[125] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of nonlinear noise in long, dispersion–uncompensated fiber links," *Optics Express*, vol. 21, no. 22, pp. 25 685–25 699, Oct. 2013.

[126] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*. New York, NY: Wiley, 1980.

[127] R. Dar, "Analytical and semi-analytical models for nonlinear transmission," in *Proc. of European Conference on Optical Communication (ECOC)*, Düsseldorf, Germany, Sept. 2016, pp. 491–493.

[128] B. Xu and M. Brandt-Pearce, "Modified Volterra series transfer function method," *IEEE Photonics Technology Letters*, vol. 14, no. 1, pp. 47–49, Jan. 2002.

[129] A. Splett, C. Kurtzke, and K. Petermann, "Ultimate transmission capacity of amplified optical fiber communication systems taking into account fiber nonlinearities," in *Proc. of European Conference on Optical Communication (ECOC)*, Montreux, Switzerland, Sept. 1993, pp. 41–44.

[130] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, "Modeling of the impact of nonlinear propagation effects in uncompensated optical coherent transmission links," *Journal of Lightwave Technology*, vol. 30, no. 10, pp. 1524–1539, May 2012.

[131] P. Johannisson and E. Agrell, "Modeling of nonlinear signal distortion in fiber-optic networks," *Journal of Lightwave Technology*, vol. 32, no. 23, pp. 4544–4552, Dec. 2014.

[132] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications," *Journal of Lightwave Technology*, vol. 32, no. 4, pp. 694–721, Feb. 2014.

[133] 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, pp. 16 335–16 362, Jun. 2014.

[134] 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, Jan. 2006.

[135] E. M. Ip and J. M. Kahn, "Fiber impairment compensation using coherent detection and digital signal processing," *Journal of Lightwave Technology*, vol. 28, no. 4, pp. 502–519, Feb. 2010.

[136] A. J. Lowery, "Fiber nonlinearity pre- and post-compensation for long-haul optical links using OFDM," *Optics Express*, vol. 15, no. 20, pp. 12 965–12 970, Sept. 2007.

[137] D. Lavery, D. Ives, G. Liga, A. Alvarado, S. J. Savory, and P. Bayvel, "The benefit of split nonlinearity compensation for optical fiber communications," *IEEE Photonics Technology Letters*, vol. 28, no. 17, pp. 1803–1806, Sept. 2016.

[138] 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, pp. 20 381–20 393, Jul. 2015.

[139] D. Rafique and A. D. Ellis, "Impact of signal–ASE four-wave mixing on the effectiveness of digital back-propagation in 112 Gb/s PM-QPSK systems," *Optics Express*, vol. 19, no. 4, pp. 3449–3454, Feb. 2011.

[140] 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, Oct. 2010.

[141] G. Gao, X. Chen, and W. Shieh, "Influence of PMD on fiber nonlinearity compensation using digital back propagation," *Optics Express*, vol. 20, no. 13, pp. 14 406–14 418, Jun. 2012.

[142] K. Goroshko, H. Louchet, and A. Richter, "Fundamental limitations of digital back propagation due to polarization mode dispersion," in *Proc. of Asia Communications and Photonics Conference (ACP)*, Hong Kong, China, Nov. 2015, p. ASu3F.5.

[143] C. B. Czegledi, G. Liga, D. Lavery, M. Karlsson, E. Agrell, S. J. Savory, and P. Bayvel, "Modified digital backpropagation accounting for polarization-mode dispersion," in *Proc. of Optical Fiber Communication Conference (OFC)*, Los Angeles, CA, Mar. 2017, p. W1G.6.

[144] J. C. Cartledge, F. P. Guiomar, F. R. Kschischang, G. Liga, and M. P. Yankov, "Digital signal processing for fiber nonlinearities," *Optics Express*, vol. 25, no. 3, pp. 1916–1936, Jan. 2017.

[145] C. Fougstedt, M. Mazur, L. Svensson, H. Eliasson, M. Karlsson, and P. Larsson-Edefors, "Time-domain digital back propagation: algorithm and finite-precision implementation aspects," in *Proc. of Optical Fiber Communication Conference (OFC)*, Los Angeles, CA, Mar. 2017, p. W1G.4.

[146] L. Galdino, D. Semrau, D. Lavery, G. Saavedra, C. B. Czegledi, E. Agrell, R. I. Killey, and P. Bayvel, "On the limits of digital back-propagation in the presence of transceiver noise," *Optics Express*, vol. 25, no. 4, pp. 4564–4578, Jan. 2017.

[147] F. P. Guiomar, J. D. Reis, A. L. Teixeira, and A. N. Pinto, "Mitigation of intra-channel nonlinearities using a frequency-domain Volterra series equalizer," *Optics Express*, vol. 20, no. 2, pp. 1360–1369, Jan. 2012.

[148] L. Liu, L. Li, Y. Huang, K. Cui, Q. Xiong, F. N. Hauske, C. Xie, and Y. Cai, "Intrachannel nonlinearity compensation by inverse Volterra series transfer function," *Journal of Lightwave Technology*, vol. 30, no. 3, pp. 310–316, Feb. 2012.

[149] W. Yan, Z. Tao, L. Dou, L. Li, S. Oda, T. Tanimura, T. Hoshida, and J. C. Rasmussen, "Low complexity digital perturbation back-propagation," in *Proc. of European Conference on Optical Communication (ECOC)*, Geneva, Switzerland, Sept. 2011, p. Tu.3.A.2.

[150] L. Dou, Z. Tao, L. Li, W. Yan, T. Tanimura, T. Hoshida, and J. C. Rasmussen, "A low complexity pre-distortion method for intra-channel nonlinearity," in *Proc. of Optical Fiber Communication Conference (OFC)*, Los Angeles, CA, Mar. 2011, p. OThF.5.

[151] J. A. Nelder and R. Mead, "A simplex method for function minimization," *The Computer Journal*, vol. 7, no. 4, pp. 308–313, Jan. 1965.

[152] E. Ip and J. M. Kahn, "Feedforward carrier recovery for coherent optical communications," *Journal of Lightwave Technology*, vol. 25, no. 9, pp. 2675–2692, Sept. 2007.

[153] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Transactions on Communications*, vol. 28, no. 11, pp. 1867–1875, Nov. 1980.

[154] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY: Cambridge University Press, 2004.

[155] H. Louchet, K. Kuzmin, and A. Richter, "Improved DSP algorithms for coherent 16-QAM transmission," in *Proc. of European Conference on Optical Communication (ECOC)*, Brussels, Belgium, Sept. 2008, p. Tu.1.E.6.

[156] M. J. Ready and R. P. Gooch, "Blind equalization based on radius directed adaptation," in *Proc. of International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Albuquerque, NM, Apr. 1990, pp. 1699–1702.

[157] F. Buchali, H. Bülow, K. Schuh, and W. Idler, "4D-CMA: enabling separation of channel compensation and polarization demultiplex," in *Proc. of Optical Fiber Communication Conference (OFC)*, Los Angeles, CA, Mar. 2015, p. Th2A.15.

[158] A. J. Laub, *Matrix Analysis for Scientists & Engineers*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2005.

[159] K. Choutagunta and J. M. Kahn, "Dynamic channel modeling for mode-division multiplexing," *Journal of Lightwave Technology*, vol. 35, no. 12, pp. 2451–2463, Jun. 2017.

Part II

Papers

PAPER A

Polarization Drift Channel Model for Coherent Fibre-Optic Systems

Cristian B. Czegledi, Magnus Karlsson, Erik Agrell, Pontus Johannisson

Nature Scientific Reports,
vol. 6, art. no. 21217, Feb. 2016

PAPER B

Temporal Stochastic Channel Model for Absolute Polarization State and Polarization-Mode Dispersion

Cristian B. Czegledi, Magnus Karlsson, Pontus Johannisson, Erik Agrell

Proc. of Optical Fiber Communication Conference (OFC),
Los Angeles, CA, Mar. 2017, p. Th3F.2

PAPER C

Modulation Format Independent Joint Polarization and Phase Tracking for Coherent Receivers

Cristian B. Czegledi, Erik Agrell, Magnus Karlsson, Pontus Johannisson

Journal of Lightwave Technology,
vol. 34, no. 14, pp. 3354–3364, July 2016

PAPER D

Polarization-Mode Dispersion Aware Digital Backpropagation

Cristian B. Czegledi, Gabriele Liga, Domaniç Lavery, Magnus Karlsson,
Erik Agrell, Seb J. Savory, Polina Bayvel

Proc. of European Conference on Optical Communication (ECOC),
Düsseldorf, Germany, Sept. 2016, pp. 1091–1093

PAPER E

A PMD-adaptive DBP Receiver Based on SNR Optimization

Gabriele Liga, Cristian B. Czegledi, Polina Bayvel

to appear in Proc. Optical Fiber Communication Conference (OFC),
San Diego, CA, Mar. 2018

PAPER F

Digital backpropagation accounting for polarization-mode dispersion

Cristian B. Czegledi, Gabriele Liga, Domaniç Lavery, Magnus Karlsson,
Erik Agrell, Seb J. Savory, Polina Bayvel

Optics Express,
vol. 25, no. 3, pp. 1903–1915, Jan. 2017

PAPER G

Volterra Series Digital Backpropagation Accounting for PMD

Cristian B. Czegledi and Ronen Dar

Proc. of European Conference on Optical Communication (ECOC),
Gothenburg, Sweden, Sept. 2017, p. W.1.D.1

CHALMERS UNIVERSITY OF TECHNOLOGY
SE-412 96 Gothenburg, Sweden
Telephone: +46 (0)31 772 10 00
www.chalmers.se

Popular Scientific Summary

Data transmission through optical fibers is the fastest form of digital communication available today and comprises the backbone of the Internet network. Every time a web page is accessed, an image is Instagrammed, or a YouTube video is watched, the data will travel most likely through an optical fiber at some point on the way to the user.

Even though optical fibers can provide very high-speed communications, the transmitted signal is disturbed along the way by various sources of noise. If the noise is beyond some acceptable levels, the information intended by the transmitter cannot be "understood" by the receiver, therefore leading to a failed communication attempt. In order to avoid these situations, the transmitter and receiver are designed such that they are able to distinguish between noise and useful information.

Before designing the transmitter and receiver such that they can tolerate more noise, the nature of the noise has to be understood and modeled mathematically. These mathematical models, often called channel models, have to reflect the behavior of the noise accurately. On the contrary, designing the transmitter and receiver based on inaccurate channel models leads to suboptimal performance.

In this thesis, we are concerned with modeling and mitigation of noise related to polarization effects. We first develop channel models for polarization effects that occur during propagation. These models can be used in computer-based simulations to reproduce polarization effects that occur in a fiber-optic communication system. Simulations offer a greater flexibility than experiments and can be used to predict the behavior of a system before setting up time-consuming experiments. Furthermore, we propose various methods that improve the tolerance to polarization effects of fiber optical communication systems, leading to an increased transmission speed and improved energy efficiency. The interested reader is referred to page i of the thesis for a technical abstract summarizing the contributions.