

A Statistical Treatment of Cross-Polarization Modulation in DWDM Systems & its Application

vorgelegt von
Diplom-Ingenieur Marcus Winter aus Potsdam

von der Fakultät IV - Elektrotechnik und Informatik - der Technischen Universität Berlin zur
Erlangung des akademischen Grades DOKTOR DER INGENIEURWISSENSCHAFTEN (DR.-ING.)

genehmigte Dissertation

PROMOTIONSAUSSCHUSS

| | |
|--------------|--------------------------------|
| Vorsitzender | Prof. Dr.-Ing. Heino Henke |
| Berichter | Prof. Dr.-Ing. Klaus Petermann |
| Berichter | Prof. Dr.-Ing. Peter Krummrich |

| | |
|---------------------------------------|---------------|
| Tag der wissenschaftlichen Aussprache | 26. Juni 2010 |
|---------------------------------------|---------------|

Berlin 2010
D83

Das, wobei unsere Berechnungen versagen, nennen wir Zufall.

– Albert Einstein

PREFACE

The present work is the culmination of an almost three-and-a-half year research effort on a topic that initially seemed (at least to me) not much more than a curious nonlinear side effect of multi-wavelength data transmission. Previous research merely indicated that compensation of PMD may be affected, and the project initially started out in this direction. This changed with the renaissance of coherent reception, which allowed powerful compensation of transmission-related distortions, but also permitted a cheap doubling of the data rate by making use of both orthogonal eigenmodes of the fiber to transmit independent data. Enter 100 Gbps Ethernet and the desire to keep the ITU 50 GHz channel spacing and there seemed to be no way around polarization-division multiplex. Suddenly, polarization effects which have not yet played a major role in direct or differential detection systems became relevant and cross-polarization modulation became interesting to a much larger audience. In this way, I consider myself lucky, as it made it easier to publish my work.

I am afraid that the following chapters contain much more mathematics than I initially intended them to have. However, when I started out there was no single notation in use – people dealing with PMD had their ways of writing the propagation equations (often Dirac’s bra-ket notation) and people dealing with nonlinearities had theirs. Finding myself at the intersection of both areas, I decided to use the best of both worlds and come up with a single way of writing things. This made it necessary to more or less start from scratch, which resulted in rather extensive chapters dealing with the basics. My desire to write everything down as formally correct as possible certainly made one or the other passage quite lengthy. However, most of the equations given herein serve a purpose related to the topic.

After reading this text, the reader should have a profound understanding of the effect of cross-polarization modulation, how it can be modeled, and how this model can be quantified. Furthermore, and perhaps most interesting, he or she will see how cross-polarization modulation affects optical communication. For some simple exemplary systems limits on the launch power are derived, but I believe the greatest benefit of this work is to show how such limits can be determined for quite arbitrary systems. At the expense of accuracy many simplifying assumptions were made to arrive at expressions that can be easier understood and from which conclusions can be drawn without the need for numerics. However, comparison to computer simulations show that the results are remarkably accurate. I would have liked to confirm those results in a laboratory, but the relevant boundary conditions were unfavorable. I must therefore rely on the accuracy of the (self-programmed) simulations for confirmation.

Large parts of Chapter 3 have been previously published in an article in the *Journal of Lightwave Technology* [Winter 2009b], and portions of Chapters 3 and 5 have been presented at various conferences and workshops. Most of the papers and slides are available on my web site at

<http://www.marcuswinter.de/research>

I would like to thank Prof. Dr.-Ing. Klaus Petermann for his continuing support, both technical and administrative, throughout these years. His notes on the first draft of this thesis – often not more than a few words followed by a question mark – have regularly resulted in rewriting of complete passages, including the addition of yet more equations. I would also like to thank my colleagues Christian Weber, Johannes Fischer, and Stefan Warm for listening to my musings and contemplations on and off the topic (and at least pretending to be interested) and also Dr.-Ing. Christian-Alexander Bunge for fruitful discussions within the first year. Furthermore, I owe thanks to Nokia Siemens Networks for funding fundamental research without an immediate return on investment and in particular Dario Setti for his support throughout the project. Prof. Peter Krummrich, initially on the project team at Nokia Siemens Networks and later as referee of this thesis, also deserves my gratitude. I also would like to thank the anonymous reviewers of my submitted papers for their help in making the topic clearer and more understandable. Last but certainly not least, I would like to thank the people that made my private life agreeable in order for me to be able to fully concentrate on my work – foremost my partner Julia Semmel, but also my parents Leo and Rolf, and my sister Heike.

Waldbronn, November 2009

CONTENTS

| | |
|--|------------|
| Memorandum on Notation ... | 1 |
| Introduction ... | 3 |
| 1 Polarization in Fiber-Optic Communication ... | 11 |
| 1.1 Classical Description of the Polarization of Light ... | 11 |
| 1.2 The Stokes Parameters & the Poincaré Sphere ... | 14 |
| 1.3 Jones & Mueller Matrices ... | 22 |
| 2 Propagation of Guided Waves ... | 27 |
| 2.1 The Fundamental Evolution Equation ... | 27 |
| 2.2 Linear Effects ... | 32 |
| 2.2.1 Group-Velocity Dispersion ... | 32 |
| 2.2.2 Birefringence ... | 36 |
| 2.2.3 Polarization-Mode Dispersion ... | 49 |
| 2.3 Nonlinear Effects ... | 63 |
| 2.3.1 Four-Wave Mixing ... | 65 |
| 2.3.2 Self-Phase Modulation ... | 67 |
| 2.3.3 Cross-Phase & Cross-Polarization Modulation ... | 69 |
| 3 A Statistical Model of XPoIM ... | 73 |
| 3.1 Random Rotations in Stokes Space ... | 74 |
| 3.2 Estimation of the Parameter \mathcal{U} ... | 85 |
| 3.2.1 Autocovariance of the Stokes Sum ... | 88 |
| 3.2.2 Walk-Off ... | 91 |
| 3.2.3 Mean SOP Decorrelation ... | 96 |
| 3.2.4 Co-Polarized & Mutually Orthogonal Launch ... | 99 |
| 3.2.5 Amplifier Noise ... | 104 |
| 3.3 Application of the Model ... | 106 |
| 3.3.1 Dependence on Fiber Parameters ... | 106 |
| 3.3.2 Dependence on Channel Load ... | 115 |
| 4 The Polarization Ensemble ... | 119 |
| 4.1 The Variance of the DOP ... | 119 |
| 4.2 The Distribution of the DOP ... | 127 |

| | | |
|----------|---|------------|
| 4.3 | Distribution of the SOPs within Single Ensemble Members ... | 129 |
| 5 | Detrimental Effects of XPolM ... | 133 |
| 5.1 | Optical Differential Detection ... | 134 |
| 5.2 | Polarization-Division Multiplex ... | 146 |
| 5.2.1 | ROSNR Penalty Estimate ... | 151 |
| 5.2.2 | Numerical Results ... | 154 |
| | Outlook ... | 165 |
| | Summary ... | 167 |
| | Zusammenfassung ... | 169 |
| | Acronyms ... | 171 |
| | List of Notation ... | 173 |
| | Bibliography ... | 175 |
| | Author's Bibliography ... | 187 |

MEMORANDUM ON NOTATION

THERE are two principal types of vectors that are used throughout this thesis. The first is the complex-valued two-dimensional vector in Jones space, denoted by $|A\rangle$. For it exists a dual, $\langle A|$, which is simply the conjugate transpose of $|A\rangle$. The inner product of two vectors is written as $\langle A|B\rangle$. A lower-case vector $|a\rangle$ is normalized such that $\langle a|a\rangle = 1$, i.e. it has unit norm.

The second vector quantity is a real-valued three-dimensional vector in Stokes space, denoted by a bold uppercase letter, as in \mathbf{S} . Its components are denoted by S_1 , S_2 , and S_3 . It also has a normalized form which is distinguished by a hat accent, as in $\hat{\mathbf{S}}$, and for which $S_1^2 + S_2^2 + S_3^2 = 1$, i.e. it has unit length.

The Fourier transform of either type of vector, and for that matter also for scalars, is denoted by a tilde, as in $|\tilde{U}\rangle$ or $\tilde{\mathbf{S}}$. Where space permitted, the dependence of the Fourier-transformed quantities on frequency is explicitly written, e.g. $\tilde{\mathbf{S}}(z, \omega)$ as opposed to $\mathbf{S}(z, t)$.

A unique vector which is used to transform between Jones and Stokes space is the three-dimensional vector whose components are the three complex Pauli matrices. This *Pauli vector* is denoted by $\vec{\sigma}$.

Matrices are denoted by a bold uppercase letter with an overbar, as in $\bar{\mathbf{M}}$, the identity matrix is $\bar{\mathbf{I}}$.

Probabilistic quantities are denoted by calligraphic capital letters, as in $\mathcal{P}[\cdot]$ for the probability of an event, $\mathcal{E}[\cdot]$ for the time average, or \mathcal{C} for an autocovariance function. The ensemble average is $\langle \cdot \rangle$ to distinguish it clearly from the time average.

Finally, partial derivatives are written as ∂_x which is short for $\partial/\partial x$. Higher-order derivatives are written as powers, e.g. ∂_x^2 for the second derivative with respect to x .

A more detailed overview of the notation used is given on page 173.

CROSS-POLARIZATION MODULATION

INTRODUCTION

TO PLACE this thesis into its proper context, and to honor the people and companies that were instrumental in making fiber-optic communication a reality, this introductory chapter will start with a brief history of fiber optics, as told in considerably greater detail in *City of Light: The Story of Fiber Optics* [Hecht 1999]. For the bibliographical references, the reader is referred to the original text. Once long-haul systems were in place, the analysis and interplay of the various linear and nonlinear effects that affected the transmission of data became a research discipline of its own. The second part of this introduction will briefly outline these effects. This is followed by the history and state-of-the-art of the research on the particular effect that is the topic of the present work: cross-polarization modulation. How this thesis aims to advance the current state-of-the-art is outlined at the end of the chapter.

A SHORT HISTORY OF NEARLY EVERYTHING

The history of fiber optics arguably begins over 150 years ago, when Daniel Colladon first demonstrated total internal reflection in 1841 by guiding light inside a water jet. Around the same time, Jaques Babinet, performing similar experiments, noted that light could be guided this way inside bent glass rods. However, the quality of the available glass at that time being rather poor, the achievable distances were very small, and until the late 1800s total internal reflection was used exclusively as an entertainment attraction. Then, surgeons discovered the usefulness of providing light in hard-to-reach places, by illuminating the insides of the mouth and nose without burning themselves or their patients with arc lamps.

By that time, glass makers had already learned how to draw glass fibers from rods, and such fibers were already finer than silk and could be woven into fabrics. The first scientist to create and examine glass fibers was Charles Vernon Boys in 1887, in his quest for a means to measure very small forces on objects. In the 1930s, glass wool from glass fibers was in commercial production in the *Owens-Illinois Glass Company* and *The Corning Glass Works*.

In the following years, much research went into transmission of images by means of fiber bundles. One product of this research was to apply a transparent, lower-index cladding around the fiber core. Interestingly, this cladding would later help to overcome the major obstacle in the way of using glass fibers for communication at the time: the extremely high attenuation that did not permit light to be transmitted over more than a few meters.

A second difficulty in using light for communication purposes that shifted the bulk of research at that time towards microwave atmospheric transmission and hollow waveguides was the lack of a coherent light source. That changed in 1960, when Theodore Maiman fired the first light pulse from a ruby

cylinder at *Hughes Research Laboratories*. Later the same year, Ali Javan at *Bell Labs* demonstrated the first continuous wave laser, a Helium-Neon gas laser, still the most common gas laser today. The first solid-state semiconductor lasers were demonstrated in 1962, developed nearly simultaneously by groups from *General Electric*, IBM, and MIT, but could operate only for very short times at the temperature of liquid helium. Their progress halted. Only after the invention of the double-heterostructure laser by Mort Panish and Izuo Hayashi at *Bell Labs*, and independently by Zhores Alferov in Leningrad, the first room-temperature semiconductor lasers appeared in 1970. Thereafter, a flurry of research activity led to the first commercially available continuous wave semiconductor lasers operating at room temperature in mid-1975. In 1977, J. Jim Hsieh at MIT developed the first laser emitting at a wavelength of $1.3\ \mu\text{m}$, where optical fiber was extraordinarily transparent, from the quaternary material InGaAsP.

In the autumn of 1965, Charles Kuen Kao calculated the required maximum loss in glass to make optical telecommunication viable to be less than 20 dB/km – at that time, the best available glass had 20 dB loss per 20 meters, largely due to impurities, but Kao was convinced that sufficiently transparent glass would be possible. Years later, in 1969, Kao and Mervin W. Jones measured the intrinsic loss of very pure fused silica to be less than 5 dB/km, finding the ultimately transparent material. In 1970, a team at *The Corning Glass Works* demonstrated the first optical fiber with a loss of less than 16 dB/km, obtained by doping fused silica with Titanium and a post-drawing heat treatment. Titanium was eventually replaced by Germanium, and by 1978 fibers with losses as low as 0.2 dB/km in the near-infrared were being demonstrated by NTT. Kao was awarded the 2009 Nobel Prize in physics for his work on optical fibers.

With the major obstacles to optical communication overcome, many research as well as commercial fiber-optic systems were deployed in the following years. In 1984, *British Telecom* deployed the first submarine fiber-optic cable to the Isle of Wight; in 1986 the first fiber link across the English Channel connected Great Britain and France. At the end of 1988, TAT-8, the first transatlantic fiber-optic cable consisting of single-mode fiber and operating at $1.3\ \mu\text{m}$, began its service.

However, the true breakthrough for long-distance fiber transmission came with the development of the erbium-doped fiber amplifier independently by David Payne and Emmanuel Desurvire in 1987. Already in 1989, gain in excess of 46 dB was reported by NTT. After having been successfully used in terrestrial systems, the first transatlantic cable using optical amplifiers, TAT-12, was deployed in 1996. Optical amplifiers also paved the way for wavelength-division multiplex (WDM) transmission of data. The first half of the 1990s saw a number of research systems making use of WDM come to life, culminating in the first commercial such systems available from *Pirelli*, *Lucent Technologies*, and *Ciena* in 1996. Since then and before, a never-ending stream of record transmission experiments has been reported in the scientific literature, steadily increasing the available capacity and transmission distance on a single optical fiber, with no end yet in sight in the near future.

OPTICAL TRANSMISSION SYSTEMS RESEARCH

With the advent of long-haul systems that send data over thousands of kilometers before being detected and the perpetual increase in data rates, a number of peculiar side effects of the transmission of light through optical fiber were discovered and had to be surmounted in order to provide the steady increase in capacity that was witnessed over the past years. The research and compensation of these effects eventually became a scientific discipline of its own: optical transmission systems research.

During the time in which multi-mode fibers were used (because light could be coupled into them much easier) mode dispersion occurred, as each mode was traveling at a different propagation velocity, broadening the data pulses. The initial remedy were gradient-index fibers, in which all modes propagate with (nearly) the same velocity. Eventually single-mode fibers (actually carrying two degenerate polarization modes) quickly replaced multi-mode fibers for long-haul distances because of their superior transmission capacity. However, even single-mode fibers exhibited chromatic dispersion, also called group velocity dispersion (GVD), the wavelength-dependent propagation velocity of light which broadens pulses until they overlap. There were different approaches to deal with dispersion: Linn Mollenauer proposed optical solitons, special pulse shapes for which dispersion and nonlinearities cancel each other; other research went into dispersion-shifted fibers, in which the material and waveguide dispersion canceled each other in the erbium-doped amplifier wavelength window of 1.55 μm . Solitons never made it past research systems, and the dispersion-shifted fiber proved to be detrimental for future wavelength-division multiplex systems. The current state-of-the-art is the use of non-zero dispersion fiber (regular single-mode fiber or non-zero dispersion-shifted fiber) together with optical or electronic dispersion compensation. Optical compensation is performed with special, dispersion-compensating fibers, developed by *Corning* in the mid-1990s, that have an inverse dispersion profile, and which can be inserted into each repeater span or at the ends of the transmission system. Electronic compensation using linear filters is usually, but not necessarily, done at the receiver end.

With group velocity dispersion out of the way, single-channel data rates were steadily increasing, until a different form of dispersion began to distort signals until they could no longer be reliably detected. Polarization-mode dispersion (PMD) is a result of the supposedly degenerate propagation modes in single-mode fiber not being exactly degenerate and having slightly different propagation velocities. This is exacerbated by these eigenmodes being random and time-dependent, with time constants as small as a few milliseconds. Because fibers at the time could exhibit enormous amounts of polarization-mode dispersion, renewed research went into the improvement of fiber manufacturing processes, with the result that modern fibers nowadays have almost negligible polarization mode dispersion. With the appearance of powerful electronic processors, this effect can now also be compensated sufficiently fast for older «legacy» fiber. However, polarization mode dispersion, even when small, has consequences on the polarization characteristics of signals propagating along the fiber which will become important for the research presented herein.

Another problem in long-haul transmission is the noise accumulation from inline optical amplifiers, which can in principle be countered by using higher optical powers to transmit the data. However,

this directly leads to one of the most disagreeable, because not easily compensated, effects of fiber-optic data transmission: the optical Kerr effect. The Kerr effect is a result of the nonlinear response of the material – in this case glass – to electromagnetic fields. Basically, the electrons in the material can be considered coupled oscillators which are excited by photons from the electromagnetic wave and thus generate their own electromagnetic field. In the direction of light propagation, the superposition of the initial field with the field of the oscillating electrons leads to an apparent phase retardation of the initial field. This phase shift is the cause for the (linear) refractive index and the associated lower velocity of light in the material. However, when the electron excitation becomes very intense, the medium response is no longer approximately linear – similar to a macroscopic spring oscillator or a pendulum that is strongly excited. This nonlinear response introduces an intensity-dependent phase retardation into the passing wave, which corresponds to an intensity-dependent refractive index of the material. This causes a multitude of effects like self-phase modulation (SPM), cross-phase modulation (XPM), four-wave mixing (FWM), and cross-polarization modulation (XPOLM), most of which are some form of nonlinear coupling between multiple wavelengths. It is the dispersion-shifted fiber, originally proposed to combat group velocity dispersion, in which this nonlinear coupling accumulates fatally, because the presence of dispersion promotes a constant change of the multi-channel optical field that prevents the necessary phase conditions for the nonlinear effects to accumulate coherently. For this reason, the use of dispersion-shifted fiber has declined since the introduction of wavelength-division multiplex. For an excellent introduction to these nonlinear transmission effects, and more that are not considered here, the reader is referred to *Nonlinear Fiber Optics* [Agrawal 2007], a seminal book on the topic.

The focus of the present work is the particular nonlinear effect called cross-polarization modulation. In some parts of the literature it is also called nonlinear birefringence [Lee 2002, Khosravani 2001, Agrawal 2007], nonlinear polarization rotation or scattering [Lin 2004, Menyuk 2006, Xie 2009], and even generalized as a polarization side-effect of cross-phase modulation [Collings 2000, Bononi 2003, Xie 2003, Borne 2004]. It has been known for quite some time that the nonlinear coupling of two or more optical signals propagating in optical fiber mediated by the Kerr effect affects not only the phase of these signals, but also their polarization state. This effect enables devices such as Kerr shutters, which combine a pump beam and a polarization filter to modulate a probe by rotating its polarization state [Duguay 1969, Stolen 1973]. For a long time, this nonlinear polarization rotation was neglected in optical transmission systems theory because the components of such systems and the modulation formats used were sufficiently independent of the polarization state. However, with increasing relevance of polarization phenomena, especially in the context of compensation of polarization mode dispersion, polarization-division multiplex, and coherent reception, it has come under increased scrutiny, both theoretically and experimentally.

STATE-OF-THE-ART

The publication that is generally regarded as having introduced cross-polarization modulation into the realm of optical transmission systems is [Mollenauer 1995]. Mollenauer et al. found that collisions of solitons at different wavelengths can «significantly alter the polarization state of the colliding solitons.» By using a simplified propagation equation (the Manakov equation, cf. Section 2.1), they were able to describe the effect analytically as the precession of the polarization states around each other. From their work they concluded that this polarization scattering prohibited polarization-division multiplex transmission, at least with solitons.

Further documented research on the topic did not pick up before the next millennium, when Collings and Boivin experimentally showed nonlinear polarization rotation in a two-channel continuous wave system and also demonstrated the nonlinear depolarization of the center channel of a 5-channel wavelength-division multiplex system [Collings 2000]. They also determined penalties for polarization-sensitive receivers in such systems. Another three years later, Xie et al. first derived the magnitude of XPM-induced polarization scattering in terms of the standard deviation of the scattering angle in a two-channel system, both semi-analytically and experimentally [Xie 2003]. They also showed, with a simplified emulator, that this scattering can induce significant penalties on polarization mode dispersion compensators. In the same year, Bononi et al. presented and thoroughly investigated a «carousel model» of polarization scattering, valid for a two-channel pump-probe system, with which they were able to predict the degree of polarization reduction of the probe in the absence of the polarization effects of polarization mode dispersion [Bononi 2003] (cf. also [Vannucci 2003] and [Corbel 2003]). With it they also showed how polarization mode dispersion and nonlinear polarization rotation cooperate to reduce the effectiveness of degree of polarization-based optical polarization mode dispersion compensators.

Picking up where Mollenauer left off, in 2004 van den Borne et al. coined the term *cross-polarization modulation*, which we adopted within the present work, and showed by simulation and experiment that it has a detrimental influence on polarization-division multiplex (POLDM) long-haul transmission [Borne 2004]. Notably, they used 5- and 7-channel NRZ polarization- and wavelength-division multiplex systems to show a large difference in bit-error rate between single- and dual-polarization transmission. In a follow-up publication the same group found that polarization interleaving of neighboring channels can reduce the nonlinear penalties, as long as there is no or little polarization-mode dispersion to modify the relative polarization states of the individual channels, making this approach useful for short-haul systems mostly [Borne 2005]. We will also examine the merits of this approach within the present work. In the same year, another set of many-channel system experiments was performed by Pachnicke et al. (which has been largely overlooked in the scientific literature likely due to their unusual naming choice of «cross-phase modulation-induced birefringence» for the effect) which determined the nonlinear reduction of the degree of polarization in 9-channel systems with different dispersion maps and various channels spacings [Pachnicke 2005, Pachnicke 2006].

Other notable research on the effect has been performed by Möller et al. who were among the first

to show a setup for demonstrating nonlinear polarization scattering [Möller 2001]. Similarly, Philips and Woodward later experimentally verified the nonlinear evolution of the polarization states described by Mollenauer for a number of specific setups involving single-tone-modulated channels [Philips 2005] (cf. also [Woodward 2004] and [Philips 2006]). Khosravani et al. emphasized the statistical nature of cross-polarization modulation by experimentally determining nonlinear penalty histograms for polarization mode dispersion compensation in a two-channel setup [Khosravani 2001], Pan et al. proposed an intra-bit polarization modulation to alleviate the detrimental effects of nonlinear polarization state fluctuations [Pan 2002, Pan 2004], and Lee et al. also looked at penalties in polarization-mode dispersion-compensated two-channel systems [Lee 2002].

The year 2006 saw the publication of a triplet of intriguing publications. Menyuk and Marks gave an excellent review of the derivation of the nonlinear Schrödinger equation and the Manakov-PMD equation for the description of nonlinear pulse propagation in an optical fiber [Menyuk 2006]. Their derivation is very rigorous and gives significantly more mathematical detail than the derivation presented here. They then applied their results to a number of examples as a verification, including the depolarization in a pump-probe system. However, the focus in 2006 was still mainly on how XPOLM would affect polarization-mode dispersion and its compensation, which is reflected in their examples. Karlsson and Sunnerud also discussed very detailed the effect of cross-polarization modulation on systems in which polarization mode dispersion is being compensated [Karlsson 2006]. Similarly to Bononi et al. they developed a model which incorporated the «walk-off» induced by different group velocities of neighboring wavelength-division multiplexed channels and derived – on a statistical basis – expressions for the nonlinear reduction of the degree of polarization in a two-channel system, which would be adaptable to many wavelength channels. As such, that part of their work may be seen as the predecessor to the theory presented herein. The remainder of their paper deals with PMD-compensated systems. They gave, for the first time, a rule-of-thumb estimation for the required nonlinear degree of polarization to be larger than 0.97 in order to avoid significant penalties. Finally, Boroditsky et al. also examined the influence of cross-polarization modulation on the polarization mode dispersion of the fiber [Boroditsky 2006]. They were the first to experimentally show and analytically describe a statistical modulation of the polarization mode dispersion vector by cross-polarization modulation. Their theory of the combined influence of linear birefringence of the fiber and nonlinear birefringence induced by a pump channel makes use of the methods of stochastic calculus to be able to handle the different random processes involved, which is an intriguing method with potential for further research in this area.

A more detailed summary of the research results on cross-polarization modulation up to approximately the year 2006 can also be found in *Optical Fiber Telecommunications V-A* [Brodsy 2008, ch. 17].

Since then the focus has shifted more and more toward polarization-division multiplex systems, especially in the context of 100-Gigabit Ethernet (e.g. [Duthel 2007] and [Fludger 2008]). With the availability of polarization-diversity coherent receivers which offer access to the amplitude, phase, and polarization information of the incident optical field, polarization demultiplex can be performed as part of the electronic signal processing without adding complexity in the optical path of the receiver. This promises a cheap doubling of the spectral efficiency. The critical aspect in these systems is that a random

modulation of the polarization state of the signal in time can induce significant penalties during the demultiplexing process and can also interfere with the polarization mode dispersion compensation in the electrical domain. There have been several publications recently, dealing expressly with simulation [Borne 2007a, Bertolini 2009] or experimental investigation [Bertran-Pardo 2009, Nelson 2009] of the nonlinear polarization effects in such systems. A particular aspect of these investigations has been the selective upgrade of existing 10 Gbps wavelength-division multiplex systems with higher-speed channels [Borne 2007b, Lefrançois 2007, Bertran-Pardo 2008a]).

THE PRESENT WORK

As described in the previous section, most published theory on the topic deals with simple pump-probe two-channel configurations only, which are relatively simple to describe because they are largely deterministic.¹ However, optical communication systems are not at all deterministic, as the system simulations and experiments, especially with more than one pump channel proved. There were only few works that put their theories on a statistical basis, such as [Karlsson 2006] and [Boroditsky 2006], but they were either very rudimentary or too abstract to enable understanding of the interactions involved in cross-polarization modulation as it occurs in wavelength-division multiplex systems. Therefore, for the systematic and non-deterministic aspects of cross-polarization modulation, the present work starts basically from scratch and should be considered a mostly self-contained work. Wherever possible it incorporates the previous work as summarized above, but the overlap proved to be small.

We start with a basic review of the polarization of light and the various means of its description. This is followed by a review of light propagation in optical fibers. The emphasis here clearly lies on the various linear and nonlinear polarization effects, with most of the other effects described earlier in this chapter mentioned only in passing. We start from the most general description of pulse propagation using the previously mentioned nonlinear Schrödinger equation and derive the Manakov equation for a wavelength-division multiplex system in Jones and Stokes space.

Having established the theoretical ground on which to build our model in Chapter 2, we proceed with a review of random polarization state rotations and their resulting distribution, derive parameters to describe this distribution, including the degree of polarization used in previous works, and figure out how to determine these parameters from the properties of the particular communication systems in Chapter 3. We then apply our model to compare the influence of a number of parameters on the magnitude of cross-polarization modulation. Since our approach is a statistical one, we are not restricted by the number of channels under consideration in any stage of the process – however, for the verification by numerical simulations, we limit ourselves to propagation of 11 channels to balance accuracy and computation time.

¹In a pump-probe environment, a high-power «pump» signal causes nonlinear distortions in a low-power «probe» signal which propagates at a different frequency and, due to its low power, does not cause significant distortions in the pump signal. The evolution of the pump can thus be approximated by the single-channel behavior.

As previous authors realized, the effects of cross-polarization modulation have a random component, due to the dependence on the relative polarization states of the various channels of the system, which evolve apparently randomly as a result of polarization mode dispersion (cf. Section 3.2.3).² Our results given in Chapter 3 incorporate this randomness, but their figures of merit are inherently only mean values for the statistical ensemble of fibers having the same statistical properties. Chapter 4 extends the theory of Chapter 3 to determine the degree of randomness present in the figures of merit and to estimate the distribution of their values. This is a crucial requirement for the estimation of the probability of outages which may occur because some systems of the same class may incur significantly different penalties than the mean values suggest.

The actual topic of transmission penalties is examined in Chapter 5. Various scenarios are examined in which the effects of cross-polarization modulation may possibly prove detrimental, such as differential phase-shift keying and polarization-division multiplex systems. We apply the results of the previous chapters, in which the magnitude of cross-polarization modulation was derived merely in a general way, to particular transmission systems and quantify the penalties incurred in these systems.

The reader should already have a basic understanding of optical communication systems as taught in many graduate classes on the topic. This includes a familiarity with the scalar nonlinear Schrödinger equation, as well as the effects of group velocity dispersion and the various «common» nonlinearities such as self-phase modulation, cross-phase modulation and four-wave mixing. While these are helpful, they are, however, not required to understand the material. Knowledge of advanced modulation formats such as (quaternary) phase-shift keying, including the differential variants, will also be helpful, but is not strictly required.

²The evolution of the polarization states is wholly deterministic when the frequency-dependent birefringence properties of the fiber are known, and it only appears random when they are not which is generally the case for communication fibers.

THIS first, introductory chapter deals with the polarization of light: how it is defined and how it can be described. We will introduce the polarimetric parameters in the Jones space description and derive the Stokes parameters. We also show how optical elements affect polarization and how the effect can be described mathematically. This chapter forms a basis and reference for all topics dealing with the polarization state of an optical signal throughout the remainder of this work.

1.1 CLASSICAL DESCRIPTION OF THE POLARIZATION OF LIGHT

To introduce the concept of the polarization of light, we start with the simplest solution to electromagnetic wave propagation that is permitted by Maxwell's laws and the wave equation derived thereof: the monochromatic plane wave in vacuum. We can describe its complex field as

$$\mathbf{E}(\mathbf{r}, t) = \sqrt{\frac{\mu_0}{\epsilon_0}} \mathbf{A} \exp(i\omega t - i\mathbf{k} \cdot \mathbf{r}) \quad (1.1)$$

where bold symbols denote vector quantities in \mathbb{R}^3 (throughout the present work); other quantities are scalar, and $i = \sqrt{-1}$. The measurable, real-valued field is given by the real part of the complex vector \mathbf{E} . We have introduced a normalization constant chosen such that $\mathbf{A}^\dagger \mathbf{A}$ equals the power flow density, or intensity, of the wave, where † denotes the conjugate transpose. Therefore, \mathbf{A} describes no longer the electric field, but a quantity we will refer to as the *optical field*, having the same orientation as the associated electric field but a different magnitude. In (1.1) $\mathbf{k} = k \mathbf{e}_k$ is the wavevector, with wavenumber $k = \omega \sqrt{\mu_0 \epsilon_0}$ and direction of propagation \mathbf{e}_k . The quantities μ_0 and ϵ_0 are the permeability and permittivity of vacuum, respectively. The wavenumber $k/2\pi$ describes the number of oscillations per unit length in the direction of propagation and is related to the wavelength λ as $k = 2\pi/\lambda$, whereas the angular frequency $\omega/2\pi$ describes the number of field oscillations per unit time. The velocity of propagation of the wavefront (constant phase) is $v_{\text{ph}} = \omega/k = c$, with the speed of light in vacuum c .

A consequence of using the complex exponential notation in (1.1) instead of the real-valued fields is that there exists no consensus regarding the sign of the argument. Texts written by physicists often use the negative carrier frequency convention $\exp(-i\omega t)$ (e.g. [Born 1999], [Agrawal 2007], and [Menyuk 2002]) while engineers often prefer the opposite $\exp(i\omega t)$ (e.g. [Damask 2004, Gordon 2000], and the present work). Either method is valid, as long as the derivations are consistent. The particular choice will affect e.g. the signs of the Stokes parameter S_3 (cf. Section 1.2) and the group-velocity dispersion term (cf. Section 2.1).

Inserting (1.1) into Gauss' law for vacuum ($\text{div } \mathbf{E} = 0$) yields $\mathbf{k} \cdot \mathbf{E} = 0$, that is, the optical field oscillates in a plane orthogonal to the direction of propagation and has no component in that direction. Without loss of generality we assume that the wave travels in the \mathbf{e}_z -direction, so that the field then lies in the \mathbf{e}_x - \mathbf{e}_y -plane. We then disregard the z -component of the field and rewrite (1.1) as

$$|A(z, t)\rangle = A|a\rangle \exp(i\omega t - ikz) \quad (1.2)$$

In writing (1.2), we have made use of Dirac's bra-ket notation (cf. [Dirac 1958]) to distinguish the 2-dimensional *ket* vector

$$|a\rangle = \begin{pmatrix} a_x \\ a_y \end{pmatrix} \quad (1.3)$$

from its 3-dimensional brethren. These 2-dimensional complex vectors are also referred to as *Jones vectors*, after R. Clark Jones who developed the methodology of describing the evolution of polarization in optical systems based on such vectors [Jones 1941]. Jones vectors will be denoted by kets throughout the present work. A *bra* corresponding to the ket (1.3) is the row vector

$$\langle a| = (a_x^* \quad a_y^*) \quad (1.4)$$

where $*$ denotes the complex conjugate. The *braket* $\langle a|a\rangle$ denotes the inner product of $\langle a|$ and $|a\rangle$. The vector $|a\rangle$ in (1.2) is normalized such that $\langle a|a\rangle = 1$, which is denoted by a lower-case letter. It is called the (normalized) *Jones vector*. Its components are generally complex quantities, and we may write

$$|a\rangle = \begin{pmatrix} \cos \chi \\ \sin \chi \exp i\phi \end{pmatrix} \quad (1.5)$$

where we have redefined A in (1.2) to include a possible phase of a_x , so that a_x is a real quantity.

The vector $|a\rangle$ can be decomposed into any set of orthonormal basis (Jones) vectors $|e_1\rangle$ and $|e_2\rangle$ with

$$\langle e_1|e_2\rangle = 0 \quad \text{and} \quad \langle e_1|e_1\rangle = \langle e_2|e_2\rangle = 1 \quad (1.6)$$

so that

$$|a\rangle = a_1|e_1\rangle + a_2|e_2\rangle \quad (1.7)$$

in which $|a\rangle$ becomes independent of any fixed spatial coordinates. In writing (1.3), we have implicitly chosen the basis vectors

$$|e_1\rangle = |e_x\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |e_2\rangle = |e_y\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (1.8)$$

However, we are not limited to this choice, as we will see in Section 1.2. The coefficients a_1 and a_2 are obtained by

$$a_1 = \langle e_1|a\rangle \quad \text{and} \quad a_2 = \langle e_2|a\rangle \quad (1.9)$$

and we may rewrite (1.7) as

$$|a\rangle = |e_1\rangle\langle e_1|a\rangle + |e_2\rangle\langle e_2|a\rangle \quad (1.10)$$

The constructs

$$|e_1\rangle\langle e_1| = |e_x\rangle\langle e_x| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad |e_2\rangle\langle e_2| = |e_y\rangle\langle e_y| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad (1.11)$$

are the so-called *projection operators* which yield vectors in the direction of the basis vectors having a length of the projection of $|a\rangle$ onto these basis vectors.

The quantities χ and ϕ in (1.5) are defined in the $|e_x\rangle$ - $|e_y\rangle$ basis set and are called the *polarimetric parameters*. They describe the partition of the wave's intensity into x - and y -components and the phase difference between these components, respectively. Together they fully define the state of polarization (SOP) of the wave. The SOP defines the motion that the tip of the field vector performs in the plane perpendicular to the direction of propagation. To examine the nature of this motion, we can write the real-valued field components, using (1.2) and (1.5), at some fixed z as

$$A_x(t) = |A| \cos \chi \cos(\omega t + \psi_0) \quad (1.12a)$$

$$A_y(t) = |A| \sin \chi \cos(\omega t + \psi_0 + \phi) \quad (1.12b)$$

where the common phase term $\psi_0 = \arg(A) + kz$. To eliminate the dependence on t and ψ_0 , we determine

$$\frac{A_x(t)}{|A| \cos \chi} \sin \phi = \cos(\omega t + \psi_0) \sin \phi \quad (1.13a)$$

$$\frac{A_x(t)}{|A| \cos \chi} \cos \phi - \frac{A_y}{|A| \sin \chi} = \sin(\omega t + \psi_0) \sin \phi \quad (1.13b)$$

Squaring and adding (1.13), we obtain the elliptical equation

$$\frac{A_x^2}{|A|^2 \cos^2 \chi} + \frac{A_y^2}{|A|^2 \sin^2 \chi} - 2 \frac{A_x A_y}{|A|^2 \cos \chi \sin \chi} \cos \phi = \sin^2 \phi \quad (1.14)$$

Therefore, the motion of the field vector in general is an ellipse which is parameterized by three independent quantities: $|A|$, χ , and ϕ . The scaling factor $|A|$ determines the size of the polarization ellipse in the plane, and χ and ϕ together describe its orientation and shape. There are three categories that an SOP may fall in: circular, linear, and general elliptical. *Circular polarization* is the most restrictive regarding the polarimetric parameters as it requires $\chi = \pm\pi/4$ (equal field magnitude in the x - and y -directions) and a phase shift of $\phi = \pm\pi/2$. *Linear polarization* occurs when both components are in phase ($\phi = 0$) or $\chi = m\pi/2$ (only one component is non-zero). In all other cases, the SOP is elliptical. The direction of rotation of the field vector is determined by the signs of $-\pi/2 \leq \chi \leq \pi/2$ and $-\pi \leq \phi \leq \pi$:

the same sign leads to counter-clockwise, or right-hand, rotation while opposite signs lead to clockwise, or left-hand, rotation (as seen by an observer looking in the $-z$ -direction, into the incoming wave). This sense of direction, however, is lost in the derivation of the elliptical equation (1.14), but can be obtained from (1.12). Figure 1.1 shows the evolution of the optical field for three selected polarization states, and Figure 1.2 shows some examples of various polarization ellipses as seen by an observer looking into the incoming wave.

The polarimetric parameters (χ, ϕ) completely describe the (normalized) polarization ellipse. However, the polarization ellipse is impossible to observe directly due to the short duration on the order of 10^{-15} s of an optical cycle in which the field vector traces this ellipse – a method of describing the polarization of light in terms of observable quantities is required to actually measure the SOP of a light wave. Similarly difficult is the task of determining the degree of polarization (DOP) of a partially polarized wave as it generally occurs in nature in which fully polarized monochromatic light is seldom. In 1852 G. G. Stokes proposed a series of seven relatively simple measurements that yielded four polarization-related parameters from which both the polarization ellipse as well as the DOP could be determined [Stokes 1852].

1.2 THE STOKES PARAMETERS & THE POINCARÉ SPHERE

A general optical field at a fixed position z consisting of a polarized and an unpolarized portion can be written as

$$|A(t)\rangle = A_p |a_p\rangle \exp i\omega t + A_u(t) |a_u(t)\rangle \exp i\omega t \quad (1.15)$$

where the fixed kz phase is again included in the phase of A_p and $A_u(t)$. The index p denotes the polarized portion of the field whose polarimetric parameters are constant for the duration of the measurement, while the index u denotes the unpolarized portion whose amplitude and polarimetric parameters, expressed by $|a_u\rangle$ vary randomly with time.

The first measurement is the total intensity of the wave, which is used as a normalization quantity for the remaining six parameters. The total intensity is

$$\begin{aligned} I_{\text{tot}} = S_0 &= \mathcal{E}[\langle A(t) | A(t) \rangle] \\ &= \mathcal{E}[|A_p|^2 \langle a_p | a_p \rangle + |A_u(t)|^2 \langle a_u(t) | a_u(t) \rangle] \\ &= |A_p|^2 + \mathcal{E}[|A_u(t)|^2] \end{aligned} \quad (1.16)$$

where we used the normalization $\langle a_p | a_p \rangle = \langle a_u | a_u \rangle = 1$, and $\mathcal{E}[\cdot]$ denotes the average over the time of the measurement – which may be as short as a few optical cycles. A reasonable next set of measurements would be the intensity in the x - and y -direction separately. Mathematically, we project the field onto the

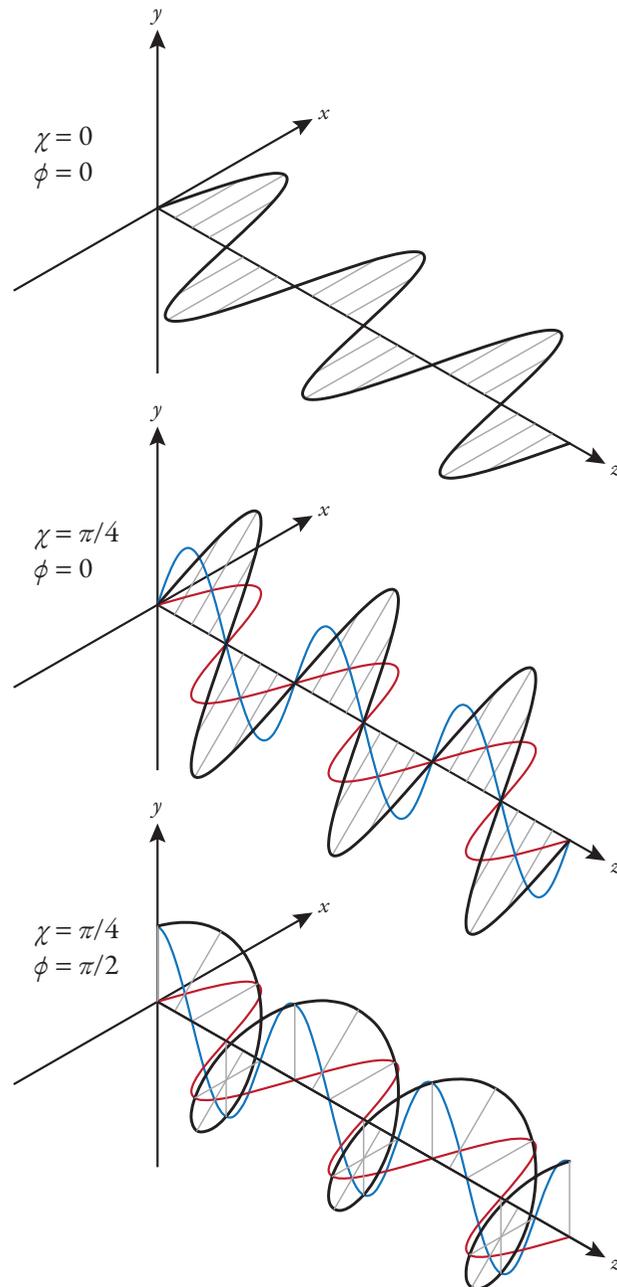


Figure 1.1 shows the optical field (black line) propagating in the z -direction with its x (red line) and y (blue line) components for three selected polarization states: linearly x -polarized, linearly polarized at 45° , and (right) circularly polarized.

CROSS-POLARIZATION MODULATION

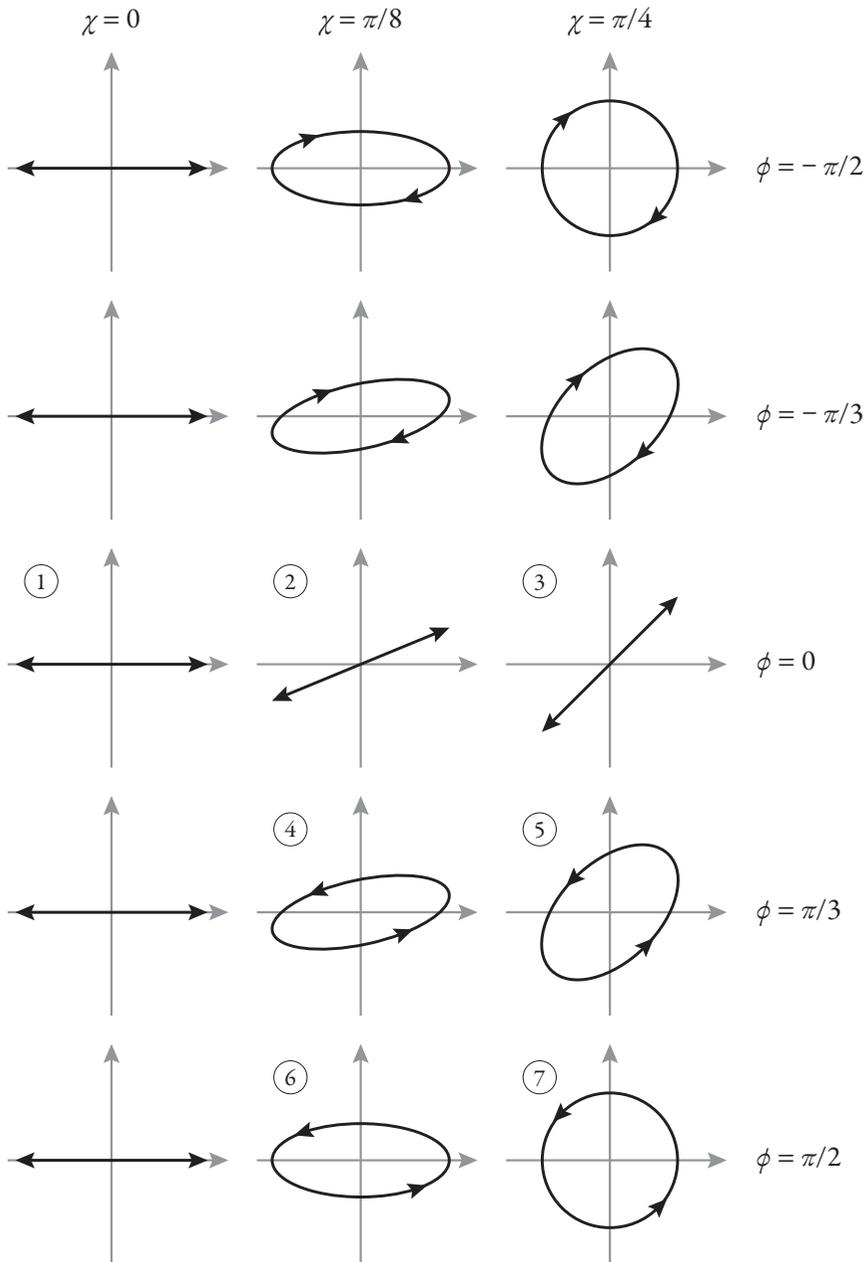


Figure 1.2 shows exemplary a number of SOPs for various combinations of the polarimetric parameters χ and ϕ . Right column top and bottom are circular polarization states, middle row and left column are linear polarization states, others are elliptic.

$|\mathbf{e}_x\rangle$ and $|\mathbf{e}_y\rangle$ basis vectors prior to the measurement, and obtain

$$\begin{aligned} I_x &= \mathcal{E}[\langle \mathbf{A}(t) | \mathbf{e}_x \rangle \langle \mathbf{e}_x | \mathbf{A}(t) \rangle] \\ &= |A_p|^2 \cos^2 \chi_p + \frac{1}{2} \mathcal{E}[|A_u(t)|^2] \end{aligned} \quad (1.17a)$$

$$\begin{aligned} I_y &= \mathcal{E}[\langle \mathbf{A}(t) | \mathbf{e}_y \rangle \langle \mathbf{e}_y | \mathbf{A}(t) \rangle] \\ &= |A_p|^2 \sin^2 \chi_p + \frac{1}{2} \mathcal{E}[|A_u(t)|^2] \end{aligned} \quad (1.17b)$$

where we have used $\mathcal{E}[\cos \chi_u(t)] = \mathcal{E}[\sin \chi_u(t)] = 0$ and $\mathcal{E}[\cos^2 \chi_u(t)] = \mathcal{E}[\sin^2 \chi_u(t)] = 1/2$. By defining the quantity

$$\begin{aligned} S_1 &= I_x - I_y \\ &= |A_p|^2 (\cos^2 \chi_p - \sin^2 \chi_p) \end{aligned} \quad (1.18)$$

$$= |A_p|^2 \cos 2\chi_p \quad (1.19)$$

we eliminate the unpolarized intensity and are left with a measure of the polarized light only. However, when χ_p equals an uneven multiple of $\pi/4$, S_1 will be zero regardless of the fraction of polarized light in the total field $|\mathbf{A}(t)\rangle$, and we cannot distinguish the so-polarized light from the unpolarized light. We must therefore make another set of measurements with the polarizers aligned in precisely those (mutual orthogonal) directions. The orthonormal basis set for this measurement is

$$|\mathbf{e}_{45^\circ}\rangle = \sqrt{\frac{1}{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad |\mathbf{e}_{-45^\circ}\rangle = \sqrt{\frac{1}{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad (1.20)$$

and we measure

$$\begin{aligned} I_{45^\circ} &= \mathcal{E}[\langle \mathbf{A}(t) | \mathbf{e}_{45^\circ} \rangle \langle \mathbf{e}_{45^\circ} | \mathbf{A}(t) \rangle] \\ &= \frac{1}{2} |A_p|^2 (1 + \sin 2\chi_p \cos \phi_p) + \frac{1}{2} \mathcal{E}[|A_u(t)|^2] \end{aligned} \quad (1.21a)$$

$$\begin{aligned} I_{-45^\circ} &= \mathcal{E}[\langle \mathbf{A}(t) | \mathbf{e}_{-45^\circ} \rangle \langle \mathbf{e}_{-45^\circ} | \mathbf{A}(t) \rangle] \\ &= \frac{1}{2} |A_p|^2 (1 - \sin 2\chi_p \cos \phi_p) + \frac{1}{2} \mathcal{E}[|A_u(t)|^2] \end{aligned} \quad (1.21b)$$

where we used $2 \cos \chi \sin \chi = \sin 2\chi$. Again, we subtract both measurements to eliminate the unpolarized intensity:

$$\begin{aligned} S_2 &= I_{45^\circ} - I_{-45^\circ} \\ &= |A_p|^2 \sin 2\chi_p \cos \phi_p \end{aligned} \quad (1.22)$$

With S_1 and S_2 we can fully quantify the linear polarized intensity portion of the optical field. However, when χ_p is an uneven multiple of $\pi/4$ and $\phi_p = \pm\pi/2$, corresponding to circular polarization, both S_1 and S_2 will be zero. With the preceding measurements we thus cannot distinguish circularly polarized light from unpolarized light. Therefore, we need a final set of measurements to quantify the circularly polarized intensity. This can be accomplished by passing the light through a quarter waveplate which converts circular polarization to linear polarization. The field transformation matrix of a quarter waveplate (we add a constant phase of $\pi/4$ to both components to simplify the expression) is

$$\bar{\mathbf{Q}} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \quad (1.23)$$

We can then again apply the polarizers from the previous measurement to isolate the desired components. The measured intensities are

$$\begin{aligned} I_R &= \mathcal{E}[\langle A(t) | \bar{\mathbf{Q}}^\dagger | e_{-45^\circ} \rangle \langle e_{-45^\circ} | \bar{\mathbf{Q}} | A(t) \rangle] \\ &= \mathcal{E}[\langle A(t) | e_R \rangle \langle e_R | A(t) \rangle] \\ &= \frac{1}{2} |A_p|^2 (1 + \sin 2\chi_p \sin \phi_p) + \frac{1}{2} \mathcal{E}[|A_u|^2] \end{aligned} \quad (1.24a)$$

$$\begin{aligned} I_L &= \mathcal{E}[\langle A(t) | \bar{\mathbf{Q}}^\dagger | e_{45^\circ} \rangle \langle e_{45^\circ} | \bar{\mathbf{Q}} | A(t) \rangle] \\ &= \mathcal{E}[\langle A(t) | e_L \rangle \langle e_L | A(t) \rangle] \\ &= \frac{1}{2} |A_p|^2 (1 - \sin 2\chi_p \sin \phi_p) + \frac{1}{2} \mathcal{E}[|A_u|^2] \end{aligned} \quad (1.24b)$$

These equations are to be interpreted as $\bar{\mathbf{Q}}$ being applied to $|A\rangle$, resulting in $\bar{\mathbf{Q}}|A\rangle$, which is then projected onto the $\pm 45^\circ$ axes as before. The corresponding orthonormal basis set derived from (1.24) consists of the left- and right-handed circular polarization states

$$|e_R\rangle = \bar{\mathbf{Q}}^\dagger |e_{-45^\circ}\rangle = \sqrt{\frac{1}{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \quad \text{and} \quad |e_L\rangle = \bar{\mathbf{Q}}^\dagger |e_{45^\circ}\rangle = \sqrt{\frac{1}{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \quad (1.25)$$

Calculating

$$\begin{aligned} S_3 &= I_R - I_L \\ &= |A_p|^2 \sin 2\chi_p \sin \phi_p \end{aligned} \quad (1.26)$$

eliminates the unpolarized portion of the field as before.

As a side note, when using the negative carrier frequency convention, the sign of the imaginary component of $\bar{\mathbf{Q}}$ changes, so that effectively the definitions of the base vectors $|e_R\rangle$ and $|e_L\rangle$ are exchanged. This will invert the sign of S_3 . Therefore, one must either redefine S_3 or associate right-handed elliptical

polarization states with negative S_3 values. This topic is further elaborated in e.g. [Menyuk 2002].

The parameters S_1, S_2, S_3 – together with the total intensity S_0 – have become known as the *Stokes parameters*. We notice that

$$\sqrt{S_1^2 + S_2^2 + S_3^2} = |A_p|^2, \quad (1.27)$$

the total intensity of the polarized light. The fraction of the intensity of the polarized light to the total intensity is called the *degree of polarization* of the wave:

$$D = \frac{|A_p|^2}{S_0} = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0} = \sqrt{s_1^2 + s_2^2 + s_3^2} \quad (1.28)$$

in which the normalized Stokes parameters are defined as

$$s_i = \frac{S_i}{S_0} \quad (1.29)$$

Consequently, for fully polarized light we have $S_0^2 = S_1^2 + S_2^2 + S_3^2$.

Regarding now only the fully polarized part of the wave, each (s_1, s_2, s_3) triplet is uniquely defined by the polarimetric parameters χ_p and ϕ_p and is thus a different way to describe an SOP. Because the DOP of the fully polarized wave is unity, the vector

$$\hat{\mathbf{S}} = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} \quad (1.30)$$

is constrained to the surface of the unit sphere in a 3-dimensional coordinate space called *Stokes space*. The unit sphere of SOPs in Stokes space is called the *Poincaré sphere*, after Henri Poincaré who introduced it in 1892 [Poincaré 1892]. Poincaré did not arrive at his sphere by way of the Stokes parameters, but by projecting the complex plane of the possible values of a_y/a_x onto a sphere using spherical trigonometry, similarly to the stereographic projection known from astronomy. Its usefulness for polarization optics was not recognized internationally before the 1960s, but has been a valuable tool ever since. A detailed derivation following Poincaré can be found in [Goldstein 2003].

Poincaré defined his sphere such that the poles would correspond to the right- (north) and left-circular (south) polarization states while the linear polarization states are on the equator. The northern hemisphere corresponds to right-handed rotation of the field vector around the polarization ellipse and the southern hemisphere to left-handed rotation. The Poincaré sphere is shown in Fig. 1.3, and some of the exemplary polarization states of Fig. 1.2 are marked on the sphere.

We notice from the definition of the Stokes parameters that the SOP representing x -polarized light is antiparallel to the SOP representing y -polarized light:

$$\hat{\mathbf{S}}_x = -\hat{\mathbf{S}}_y \quad (1.31)$$

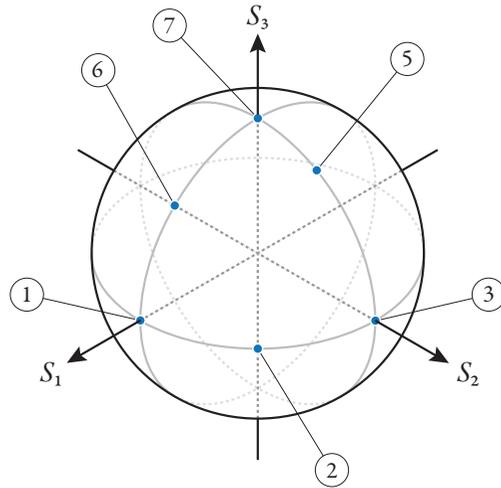


Figure 1.3 The Poincaré sphere. All points on the surface of the sphere correspond to unique polarization states. Linear SOPs are confined to the equator, while right- and left-handed circular polarization states are found at the north and south poles, respectively. Numbered SOPs correspond to the examples of Fig. 1.2.

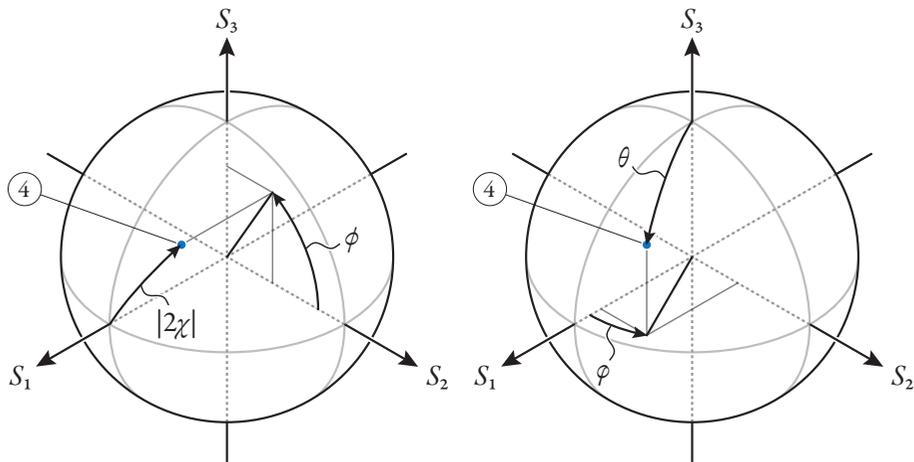


Figure 1.4 shows two variants of identifying an SOP, other than using its coordinates in Stokes space. The first (left) uses the polarimetric parameters χ and ϕ . Note that for negative χ , the angle ϕ is measured from the $-S_2$ axis. The other method (right) uses the spherical polar coordinates θ (colatitude) and ϕ (azimuth).

and also $\hat{\mathbf{S}}_{45^\circ} = -\hat{\mathbf{S}}_{-45^\circ}$ and $\hat{\mathbf{S}}_R = -\hat{\mathbf{S}}_L$. Generally, any pair of orthogonal Jones vectors $|a_+\rangle$ and $|a_-\rangle$, for which [Damask 2004]

$$\chi_- = \chi_+ + \frac{\pi}{2} \quad \text{and} \quad \phi_- = \phi_+ \quad (1.32)$$

results in antiparallel SOPs $\hat{\mathbf{S}}_+ = -\hat{\mathbf{S}}_-$ as a result of the projection.

The vector

$$\mathbf{S} = \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix} \quad (1.33)$$

points into the same direction as the SOP $\hat{\mathbf{S}}$, but has length S_0 , the total intensity. It is called a *Stokes vector*. The Stokes vector is sometimes defined as a 4-vector with the initial vector element S_0 . However, since S_0 is redundant when not dealing with unpolarized light, we use (1.33) as Stokes vector definition. While each optical field vector $A_p|a_p\rangle$ has a corresponding Stokes vector \mathbf{S}_a , there is no such one-to-one mapping from Stokes space into Jones space because the absolute phase information is absent in Stokes space. However, the Jones vector $A_p|a_p\rangle$ can be reconstructed except for an unknown phase ψ_p [Damask 2004, eq. (1.4.19)]:

$$A_p|a_p\rangle = \sqrt{S_0} \exp i\psi_p \begin{pmatrix} \sqrt{\frac{1}{2}(1+s_1)} \\ \sqrt{\frac{1}{2}(1-s_1)} \exp(i \tan^{-1} s_3/s_2) \end{pmatrix} \quad (1.34)$$

We can rewrite the equations (1.17), (1.21), and (1.24) for the normalized Stokes parameters directly in terms of the Jones vector $|a\rangle = |a_p\rangle$ as

$$\begin{aligned} s_1 &= \langle a | \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} | a \rangle - \langle a | \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} | a \rangle \\ &= \langle a | \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} | a \rangle = a_x^* a_x - a_y^* a_y \end{aligned} \quad (1.35a)$$

and analogously

$$s_2 = \langle a | \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} | a \rangle = a_y^* a_x + a_x^* a_y \quad (1.35b)$$

$$s_3 = \langle a | \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} | a \rangle = i(a_y^* a_x - a_x^* a_y) \quad (1.35c)$$

where taking the expectation is no longer necessary, since the fully polarized wave is assumed to have constant polarimetric parameters for the duration of the measurement.

The matrices in (1.35a-c) are known in quantum physics as the *Pauli matrices* σ_i . However, there their numbering is permuted from the associated Stokes parameters. We will follow the convention introduced by Gordon and Kogelnik to «renumber» the Pauli matrices to correspond to the Stokes

parameters [Gordon 2000]:

$$\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad (1.36)$$

One can then define the *Pauli vector* of Pauli matrices

$$\vec{\sigma} = \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} \quad (1.37)$$

which allows for the very concise definition of the SOP $\hat{\mathbf{S}}$ and Stokes vector \mathbf{S} using the equivalent Jones space quantities:

$$\hat{\mathbf{S}} = \langle \mathbf{a} | \vec{\sigma} | \mathbf{a} \rangle \quad \text{and} \quad \mathbf{S} = \langle \mathbf{A} | \vec{\sigma} | \mathbf{A} \rangle \quad (1.38)$$

which is equivalent to (1.35). There are various other possible sets of angles which could describe the polarization ellipse either in Jones or in Stokes space, all with their own merits and drawbacks, and we do not venture to list them all. However, an often used and sensible set that will prove to be very useful has not yet been introduced. When working with points on a sphere, such as the SOPs, it is often useful to give their position by using the spherical polar coordinates φ (azimuth) and θ (zenith, colatitude, or polar angle). In spherical coordinates, the Stokes parameters become

$$s_1 = \cos \theta \cos \varphi \quad (1.39a)$$

$$s_2 = \sin \theta \cos \varphi \quad (1.39b)$$

$$s_3 = \sin \varphi \quad (1.39c)$$

The spherical polar coordinates are shown in Fig. 1.4.

1.3 JONES & MUELLER MATRICES

Any operation that changes the field $|\mathbf{A}\rangle$ with SOP $|\mathbf{a}\rangle$ of the wave into a different field $|\mathbf{B}\rangle$ with SOP $|\mathbf{b}\rangle$ can be represented by a matrix operation

$$|\mathbf{B}\rangle = \bar{\mathbf{J}}|\mathbf{A}\rangle \quad (1.40)$$

where $\bar{\mathbf{J}}$ is a 2×2 complex matrix, called the (general) *Jones matrix*. When we exclude any polarization-dependent gain or loss, we may decompose $\bar{\mathbf{J}}$ into

$$\bar{\mathbf{J}} = T\bar{\mathbf{U}} \quad (1.41)$$

in which T incorporates any gain, loss and phase shift common to both polarization components and $\bar{\mathbf{U}}$ is a complex unitary matrix, i.e. [Riley 2006]

$$\bar{\mathbf{U}}^\dagger \bar{\mathbf{U}} = \bar{\mathbf{I}} \quad (1.42)$$

thus its inverse $\bar{\mathbf{U}}^{-1} = \bar{\mathbf{U}}^\dagger$. Unitary matrices preserve the norm of the (complex) vector they operate upon. When writing the plane wave as $|A\rangle = A \exp(i\omega t - ikz)|a\rangle$ (and analog for $|B\rangle$), we can separate the polarization-independent from the polarization-dependent component of $\bar{\mathbf{J}}$ in (1.40):

$$B = TA \quad (1.43a)$$

$$|b\rangle = \bar{\mathbf{U}}|a\rangle \quad (1.43b)$$

Henceforth, when talking about Jones matrices, we refer to the unitary Jones matrix $\bar{\mathbf{U}}$ unless noted otherwise.

We have used an SOP transformation such as that in (1.43b) without further comment before, in (1.25), where we transformed the basis vectors $|e_{-45^\circ}\rangle$ and $|e_{45^\circ}\rangle$ into $|e_R\rangle$ and $|e_L\rangle$ by means of the unitary matrix $\bar{\mathbf{Q}}^\dagger$, representing the quarter waveplate. Table 1.1 lists several optical elements and their corresponding Jones matrices.

Because the result of the operation $\bar{\mathbf{U}}$ on the SOP $|a\rangle$ is another SOP, $|b\rangle$, both of which can be represented by points on the Poincaré sphere, there must be an operation $\bar{\mathbf{R}}$ in Stokes space isomorphic to $\bar{\mathbf{U}}$:

$$\hat{\mathbf{S}}_b = \bar{\mathbf{R}} \cdot \hat{\mathbf{S}}_a \quad (1.44)$$

The elements of the 3×3 matrix $\bar{\mathbf{R}}$ must be real, as $\hat{\mathbf{S}}_a$ and $\hat{\mathbf{S}}_b$ are real quantities. Since $\bar{\mathbf{R}}$ cannot affect the length of the vector it operates upon, it must represent a rotation. Thus $\bar{\mathbf{R}}$ is an orthogonal matrix with $\bar{\mathbf{R}}^T = \bar{\mathbf{R}}^{-1}$ – the real-valued counterpart to the complex unitary matrix $\bar{\mathbf{U}}$, where T denotes the transpose. Table 1.1 also shows the Stokes rotation matrices corresponding to the exemplary Jones matrices.

The ability to represent the effect of certain optical elements as a rotation in an appropriately defined space was the original motivation for Poincaré to construct his sphere. As matrix algebra was not yet common in optics research at the end of the 19th century, the manual calculations of Jones vectors it replaced was very tedious. By projecting the complex Jones vectors into a real vector space, he greatly simplified the calculation of the effect of such elements. Until today, many polarization effects are easier to visualize, to understand, and to quantify with rotations of SOPs on the Poincaré sphere instead of unitary matrix operations on complex Jones vectors.

This can be most easily demonstrated with a waveplate oriented with its fast axis along the y -polarization. The effect of such a linear retarder in Jones space is to introduce a phase shift ψ between the x - and y -polarized components of the wave (cf. Table 1.1), i.e. it increases the polarimetric phase shift ϕ by ψ . As can be seen in Fig. 1.4 from the definition of ϕ , this corresponds in Stokes space to a rotation around the S_1 -axis, which is the slow axis of the waveplate. Any SOP in Stokes space that coin-

Table 1.1 Some optical elements and their corresponding Jones and Stokes matrices.

| | $\bar{\mathbf{U}}$ | $\bar{\mathbf{R}}$ |
|--|--|--|
| free space | $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ | $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ |
| quarter waveplate (fast axis $ e_y\rangle$) | $\begin{pmatrix} \exp -i\frac{\pi}{4} & 0 \\ 0 & \exp i\frac{\pi}{4} \end{pmatrix}$ | $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$ |
| waveplate (fast axis $ e_y\rangle$, retardation angle ψ) | $\begin{pmatrix} \exp -i\frac{\psi}{2} & 0 \\ 0 & \exp i\frac{\psi}{2} \end{pmatrix}$ | $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \psi & -\sin \psi \\ 0 & \sin \psi & \cos \psi \end{pmatrix}$ |
| waveplate (fast axis $ e_{45^\circ}\rangle$, retardation angle ψ) | $\begin{pmatrix} \cos \frac{\psi}{2} & -i \sin \frac{\psi}{2} \\ -i \sin \frac{\psi}{2} & \cos \frac{\psi}{2} \end{pmatrix}$ | $\begin{pmatrix} \cos \psi & 0 & \sin \psi \\ 0 & 1 & 0 \\ -\sin \psi & 0 & \cos \psi \end{pmatrix}$ |
| waveplate (fast axis $ e_R\rangle$, retardation angle ψ) | $\begin{pmatrix} \cos \frac{\psi}{2} & -i \sin \frac{\psi}{2} \\ i \sin \frac{\psi}{2} & \cos \frac{\psi}{2} \end{pmatrix}$ | $\begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$ |

cides with the fast axis (or the antiparallel slow axis) is not affected by such a rotation – it is therefore an eigenvector of $\bar{\mathbf{R}}$. The corresponding Jones vector must thus also be an eigenvector of the isomorphic Jones matrix $\bar{\mathbf{U}}$. And indeed, if the wave is fully polarized along either the x - or the y -axis, the relative phase ϕ (or its change) becomes meaningless.

The effect of the various retarders in Table 1.1 can be visualized in Stokes space as the rotation of an SOP through the angle ψ around a vector pointing in the direction of the slow axis of the retarder.

The general (real) 4×4 matrix transforming a 4-element Stokes vector representing a partially polarized wave into another 4-element Stokes vector is called the *Mueller matrix* $\bar{\mathbf{M}}$, after H. Mueller who developed his calculus of Stokes vectors in the early 1940s. For a unitary transformation matrix $\bar{\mathbf{U}}$, the corresponding Mueller matrix will look like

$$\bar{\mathbf{M}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \bar{\mathbf{R}} & \\ 0 & & & \end{pmatrix} \quad (1.45)$$

in which $\bar{\mathbf{R}}$ is a submatrix of $\bar{\mathbf{M}}$. The (1,1) element of $\bar{\mathbf{M}}$ shows that the total intensity of the wave, S_0 , is not affected. Since we will be working with fully polarized (at least for the duration of several optical cycles) light, S_0 is defined in terms of the other three Stokes parameters as per (1.28), and we may use the 3-dimensional Stokes vectors and rotation matrix $\bar{\mathbf{R}}^3$.

A general expression which relates the Jones matrix $\bar{\mathbf{U}}$ and the Stokes matrix $\bar{\mathbf{R}}$ can be derived by using (1.38) and (1.43b):

$$\hat{\mathbf{S}}_b = \bar{\mathbf{R}} \cdot \hat{\mathbf{S}}_a = \bar{\mathbf{R}} \langle a | \vec{\sigma} | a \rangle = \langle a | \bar{\mathbf{R}} \vec{\sigma} | a \rangle \quad \text{and} \quad \hat{\mathbf{S}}_b = \langle b | \vec{\sigma} | b \rangle = \langle a | \bar{\mathbf{U}}^\dagger \vec{\sigma} \bar{\mathbf{U}} | a \rangle \quad (1.46)$$

where $\bar{\mathbf{R}} \vec{\sigma}$ is to be interpreted as the product of a 3×3 matrix with the 3-dimensional Pauli vector, resulting in a different 3-dimensional vector (whose elements are again matrices like those of the Pauli vector). Equating the above expressions yields

$$\bar{\mathbf{R}} \vec{\sigma} = \bar{\mathbf{U}}^\dagger \vec{\sigma} \bar{\mathbf{U}} \quad (1.47)$$

It should be noted that the algebraic sign of $\bar{\mathbf{U}}$ is undetermined, even when $\bar{\mathbf{R}}$ is known, and the two Jones matrices differing only by their sign correspond to the same Mueller matrix. For a very thorough treatise on the relations between the various forms of Jones and Stokes matrices and their interrelationship, the reader is referred to [Gordon 2000].

A great advantage of using matrix algebra to calculate the effect of various optical elements on the SOP, either in Jones or Stokes space, is the ability to concatenate transformations: the effect of successive

³For a thorough introduction to Mueller algebra, cf. [Aiello 2006].

transformations can be described by the product of the corresponding matrices. With the initial SOP $|a\rangle$ (\hat{S}_a) and the SOP $|b\rangle$ (\hat{S}_b) after an initial transformation $\bar{U}_{[1]}$ ($\bar{R}_{[1]}$), we can describe the SOP $|c\rangle$ (\hat{S}_c) at the output after a second transformation $\bar{U}_{[2]}$ ($\bar{R}_{[2]}$) as

$$|c\rangle = \bar{U}_{[2]}|b\rangle = \bar{U}_{[2]}\bar{U}_{[1]}|a\rangle = \bar{U}|a\rangle \quad (1.48a)$$

and

$$\hat{S}_c = \bar{R}_{[2]} \cdot \hat{S}_b = \bar{R}_{[2]}\bar{R}_{[1]} \cdot \hat{S}_a = \bar{R} \cdot \hat{S}_a \quad (1.48b)$$

with

$$\bar{U} = \bar{U}_{[2]}\bar{U}_{[1]} \quad \text{and} \quad \bar{R} = \bar{R}_{[2]}\bar{R}_{[1]} \quad (1.48c)$$

In Stokes space, the matrix \bar{R} then describes a series of rotations of the polarization state. And just as matrix multiplication in general is not commutative, we must regard the order in which these rotations are performed.

As we will see in the next chapter, a short piece of single-mode optical fiber acts like a linear waveplate concerning the evolution of the polarization of a signal passing through it, and we will be able to use the above relations to describe the polarization evolution in optical fibers by regarding them as a concatenation of many such short pieces of fiber.

THE SECOND chapter discusses the propagation of modulated signals in optical fibers as described by the nonlinear Schrödinger equation (NLSE) and the Manakov equation derived thereof. Special emphasis is put upon the polarization effects such as birefringence, polarization mode dispersion (PMD), and of course cross-polarization modulation (XPOLM). We will also briefly review group velocity dispersion (GVD) as well as the most important (Kerr) nonlinear effects of self-phase modulation (SPM), cross-phase modulation (XPM), and four-wave mixing (FWM) and their description in the Manakov reference frame. The expressions derived in this chapter will form the basis for the analysis in the subsequent chapters.

2.1 THE FUNDAMENTAL EVOLUTION EQUATION

The spatial and temporal evolution of an electric field in any medium is governed at its most fundamental level by Maxwell's equations. From these derives the electromagnetic wave equation, a second-order partial differential equation (PDE). When applied to the geometry of a cylindrical waveguide such as optical fiber, one obtains for the evolution in the longitudinal direction a first-order PDE which is called nonlinear Schrödinger equation (NLSE) due to its similarity to the equation of the same name in quantum mechanics. We will neither derive nor outline the derivation of the NLSE here.⁴ However, we shall emphasize an important observation made during its derivation: optical single-mode fibers are weakly guiding – i.e. the index difference between core and cladding is small. As a consequence the longitudinal field components are negligible and the propagating wave is approximately transversal. We can thus again use the bra-ket notation to describe the transversal components of the field vectors.

Electromagnetic waves can propagate independently, that is without coupling, in each of two (ideally) degenerate, orthogonal eigenmodes $|R_+(x, y)\rangle$ and $|R_-(x, y)\rangle$ of the fiber, where we assume propagation in the z -direction, so that x and y are the transverse Cartesian coordinates. The eigenmodes $|R_+\rangle$ and $|R_-\rangle$ describe the transversal distribution of the optical field and its orientation. We assume the transversal dependence of both eigenmodes to be equal, so that

$$|R_+(x, y)\rangle = F(x, y)|e_+\rangle \quad \text{and} \quad |R_-(x, y)\rangle = F(x, y)|e_-\rangle \quad (2.1)$$

with the orthonormal basis vectors $|e_+\rangle$ and $|e_-\rangle$, and where $F(x, y)$ describes this transversal depen-

⁴cf. [Agrawal 2007] or [Damask 2004] for very detailed, and [Menyuk 1999] or [Menyuk 2006] for mathematically very rigorous derivations

dence, normalized so that

$$\iint_{-\infty}^{\infty} F^*(x, y) F(x, y) dx dy = 1 \quad (2.2)$$

We can now rewrite the transversal field components of the propagating wave (1.1), analogously to (1.2), for the weakly guiding fiber as

$$|E(x, y, z, t)\rangle = F(x, y) \left[A_1(z, t) |e_+\rangle + A_2(z, t) |e_-\rangle \right] \quad (2.3)$$

Under the implicit assumption that $F(x, y)$ does not change during propagation, i.e. that it describes an eigenmode of the fiber, we can describe the evolution of the optical field solely by its dependence on the longitudinal, or propagation, direction. We can then represent the optical field in the fiber by an equivalent superposition of plane waves described by

$$|A(z, t)\rangle = A_1(z, t) |e_+\rangle + A_2(z, t) |e_-\rangle \quad (2.4)$$

The basis vectors may be Cartesian ($|e_+\rangle = |e_x\rangle$ and $|e_-\rangle = |e_y\rangle$) or any other orthonormal pair in Jones space. Menyuk has shown that that this *plane wave approximation* will lead to correct results for longitudinal length scales that are larger than the transverse extent of the field [Menyuk 1989, Menyuk 2006], which holds for all fiber communication systems. In order to obtain correct results, it is however necessary to properly scale the material- and waveguide-dependent coefficients when going from $|E\rangle$ to $|A\rangle$, to account for the spatial confinement of the actual propagating wave, during the calculation of which the accurate field distribution F in the transverse dimension must be known.

Before writing down the propagation equation for $|A\rangle$, we remark that the width of the modulated spectrum in optical fiber is merely a small fraction of the (carrier) frequency of the optical field. Their ratio is on the order of about 1/100. The fast oscillations of the optical carrier are not relevant for a description of the evolution of the modulated signal and are customarily removed by defining a slowly-varying envelope $|B\rangle$ modulated onto an optical carrier which oscillates at ω_0 , usually defined as the center of the modulated spectrum,

$$|A(z, t)\rangle = |B(z, t)\rangle \exp(i\omega_0 t - i\beta_0 z) \quad (2.5)$$

in which $\beta_0 = \beta(\omega_0)$ is the phase constant in the medium (corresponding to the wavenumber k of (1.2) in a vacuum). This is called the *quasi-monochromatic approximation* and is sketched in Fig. 2.1. In the signal spectrum, the transformation (2.5) transforms the spectrum of $|A\rangle$, a comparatively narrow band around ω_0 , into the baseband, i.e. the spectrum of $|B\rangle$ is located around the zero frequency. The evolution of $|B\rangle$ is governed by the NLSE.

There exist various versions of the NLSE⁵ which differ mainly by the number of propagation effects

⁵When talking about the NLSE, we will always refer to the vector-valued equation. There exists a simpler, scalar version which only describes the evolution of $A(z, t)$ and disregards the polarization effects acting on $|a(z, t)\rangle$ and is thus unsuited

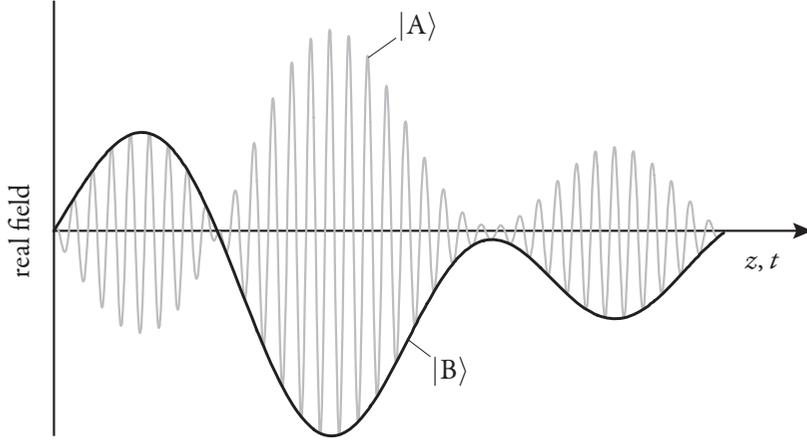


Figure 2.1 illustrates the definition of the slowly-varying envelope $|B\rangle$. The ratio of the highest frequencies occurring in the spectrum of $|B\rangle$ to the carrier frequency ω_0 is on the order of $1/100$ in optical communication systems.

they describe (and often the notation they use). We will use the following equation for the optical field envelope $|B\rangle$, which is essentially [Menyuk 2006, eq. (34)] in bra-ket notation, as working basis:

$$\begin{aligned} \partial_z |B\rangle + \frac{\alpha}{2} |B\rangle + i \frac{\Delta\beta_0}{2} \bar{\mathbf{H}} |B\rangle + \left(\beta_1 \bar{\mathbf{I}} + \frac{\Delta\beta_1}{2} \bar{\mathbf{H}} \right) \partial_t |B\rangle - i \frac{\beta_2}{2} \partial_t^2 |B\rangle \\ - \frac{\beta_3}{6} \partial_t^3 |B\rangle - i\gamma \left[\langle B|B\rangle |B\rangle - \frac{1}{3} \left(\langle B|\sigma_3|B\rangle \right) \sigma_3 |B\rangle \right] = 0 \end{aligned} \quad (2.6)$$

in which ∂_z is an abbreviation for $\partial/\partial z$ and similarly $\partial_t^n = \partial^n/\partial t^n$. Throughout the remainder of the present chapter, we will discuss the various terms appearing in (2.6).

The second term in (2.6) describes the linear, polarization-independent attenuation with power attenuation coefficient α . The terms involving the various derivatives of $\beta(\omega)$ taken at ω_0 , $\beta_n = \partial_\omega^n \beta$, are responsible for the group velocity $v_g = \beta_1^{-1}$ of the wave (or more accurately, the mean group velocity of both eigenmodes), GVD, and third-order dispersion (TOD). The TOD β_3 only becomes relevant when the GVD β_2 is near zero or being compensated. We assume $\beta_3 = 0$ to simplify our equations. It would be straightforward to include β_3 if required. The GVD is examined in Section 2.2.1. The terms containing the coefficients $\Delta\beta_0 = \Delta\beta(\omega_0)$ and $\Delta\beta_1 = \partial_\omega \Delta\beta$ and the matrix $\bar{\mathbf{H}}$ describe birefringence, its frequency derivative (commonly referred to as PMD), and the local fiber eigenmode structure, respectively, and are treated in Sections 2.2.2 and 2.2.3. By not including any higher-order terms in the birefringence $\Delta\beta$, we implicitly assume that GVD and TOD have no polarization dependence. This is in agreement with current literature [Menyuk 2006, Agrawal 2007].

The final term in square brackets expresses the Kerr nonlinearity in vector form, in which γ is the

for our purposes.

fiber nonlinearity coefficient⁶. Other effects such as Raman or Brillouin scattering are not included in the scope of the present work.

It should be remarked that all the material- and waveguide-dependent coefficients in (2.6) describe various properties of the transmission fiber which may change along the propagation path. They are thus implicitly assumed functions of the transmission distance z . Also, as mentioned before, when using the negative frequency sign convention, the GVD term in (2.6) changes sign.

We now introduce the conventional retarded reference frame with $\tau = t - \beta_1 z$ which moves with the mean group velocity v_g at the carrier frequency and obtain the following form of the NLSE for $|B(z, \tau)\rangle$:

$$\begin{aligned} \partial_z |B\rangle + \frac{\alpha}{2} |B\rangle + i \frac{\Delta\beta_0}{2} \bar{\mathbf{H}} |B\rangle + \frac{\Delta\beta_1}{2} \bar{\mathbf{H}} \partial_\tau |B\rangle - i \frac{\beta_2}{2} \partial_\tau^2 |B\rangle \\ - i\gamma \left[\langle B|B\rangle |B\rangle - \frac{1}{3} \left(\langle B|\sigma_3|B\rangle \right) \sigma_3 |B\rangle \right] = 0 \end{aligned} \quad (2.7)$$

from which the β_1 -term has been eliminated. For clarity, we rename the variable τ to t . It is important to note, however, that t now has a different meaning than the t in (2.6), because we changed its frame of reference. Making the transformation

$$|U(z, t)\rangle = \exp\left(\frac{\alpha}{2} z\right) |B(z, t)\rangle \quad (2.8)$$

one obtains the following equation for $|U\rangle$:

$$\begin{aligned} \partial_z |U\rangle + i \frac{\Delta\beta_0}{2} \bar{\mathbf{H}} |U\rangle + \frac{\Delta\beta_1}{2} \bar{\mathbf{H}} \partial_t |U\rangle - i \frac{\beta_2}{2} \partial_t^2 |U\rangle \\ - i \exp(-\alpha z) \gamma \left[\langle U|U\rangle |U\rangle - \frac{1}{3} \left(\langle U|\sigma_3|U\rangle \right) \sigma_3 |U\rangle \right] = 0 \end{aligned} \quad (2.9)$$

By means of (2.8), the attenuation-related term has been eliminated from the NLSE in exchange for a now z -dependent scaling factor of the nonlinear term.

All of the effects described by the various terms in (2.9) (except birefringence) can be separated into *intrachannel* (or in-band) and *interchannel* (or out-of-band) components. The intrachannel effects are observable even when only a single channel is transmitted, while the interchannel effects, as their name implies, require at least two co-propagating wavelength-multiplexed channels. In the case of interchannel nonlinearities, as we will see, the corresponding term in the NLSE decomposes into a large number of terms describing the nonlinear interactions such as cross-polarization modulation (XPOLM) between the various channels of the wavelength-division multiplex (WDM) system. To be able to separate the intrachannel from the interchannel effects, we can no longer describe the full optical spectrum comprising all channels by a single field vector $|U\rangle$. Instead, we consider a number N of co-propagating channels as

⁶This nonlinearity term is derived from the nonlinear polarizability and can take various equivalent forms. This particular form can be found e.g. in [Menyuk 1999, eq. (46)], [McKinstry 2007, eq. (34)], or [Agrawal 2007, eq. (6.6.9)]

baseband signals modulated onto their own respective carriers, and we superimpose their optical fields,

$$|U\rangle = \sum_{\nu=1}^N |U_{\nu}\rangle \exp(i\Delta\omega_{\nu}t) \quad (2.10)$$

in which $|U_{\nu}\rangle = |U_{\nu}(z, t)\rangle$ is a single-channel data signal and

$$\Delta\omega_{\nu} = \omega_{\nu} - \omega_0 \quad (2.11)$$

is the relative frequency offset of the optical carrier frequency ω_{ν} of channel ν from the carrier ω_0 . The spectrum of $|U\rangle$ is still a baseband spectrum as a result of (2.5), but it has been split into smaller bands with slowly-varying envelopes $|U_{\nu}\rangle$ at the carrier frequencies $\Delta\omega_{\nu}$. Usually, ω_0 is chosen so that it coincides with one of the carriers which we index with ρ , so that $\Delta\omega_{\rho} = 0$. We will call channel ρ the *probe channel*.

POLARIZATION-DIVISION MULTIPLEX Before we endeavor to look more closely at the various terms and their relevance in the scope of the present work, we shall introduce the concept of polarization-division multiplex (POLDM) which was first proposed in the context of long-haul soliton transmission [Evangelides 1992]. Basically it consists of modulating the components of an orthogonal expansion of the optical field as in (2.4) with independent data signals in order to double the information that can be transferred within a frequency band and amount of time (cf. also [Winzer 2008]). An immediate drawback is that the optical signal-to-noise ratio (OSNR) per subchannel is decreased by 3 dB, because only one half of the power within the wavelength channel is available to each POLDM subchannel. The SOPs of both subchannels must be orthogonal, but do not have to be the basis vectors of the coordinate system used or any other predefined basis such as the local fiber eigenmodes $|e_{+}\rangle$ and $|e_{-}\rangle$. We shall label them $|e_1\rangle$ and $|e_2\rangle$ for subchannels 1 and 2, respectively, and similarly to (1.7) we can write the composite field as

$$|U\rangle = U_1|e_1\rangle + U_2|e_2\rangle \quad (2.12)$$

At the POLDM receiver, polarizers aligned to $|e_1\rangle$ and $|e_2\rangle$ extract the original signals,

$$|e_1\rangle\langle e_1|U\rangle = U_1|e_1\rangle\langle e_1|e_1\rangle + U_2|e_1\rangle\langle e_1|e_2\rangle = U_1|e_1\rangle \quad (2.13a)$$

and

$$|e_2\rangle\langle e_2|U\rangle = U_1|e_2\rangle\langle e_2|e_1\rangle + U_2|e_2\rangle\langle e_2|e_2\rangle = U_2|e_2\rangle \quad (2.13b)$$

where we used the projection operators (1.11) of Section 1.1 to describe the polarizers mathematically, and also $\langle e_1|e_1\rangle = 1$ and $\langle e_2|e_1\rangle = 0$ for the orthonormal basis.

This, however, only works when the eigenmodes of a linear fiber are perfectly degenerate, i.e. propagate with the same velocity. In the NLSE above there exist three terms that introduce some coupling between the fiber modes which affects proper reception: the birefringence and PMD via the coupling ma-

trix $\bar{\mathbf{H}}$, and the Kerr nonlinearity via σ_3 . How exactly these effect influence the transmission of POLDM signals will be shown in the respective sections of this chapter.

2.2 LINEAR EFFECTS

In order to study the linear effects, we set the nonlinearity coefficient $\gamma = 0$ in (2.9) and work with the simplified (linear) Schrödinger equation

$$\partial_z |U\rangle + i \frac{\Delta\beta_0}{2} \bar{\mathbf{H}} |U\rangle + \frac{\Delta\beta_1}{2} \bar{\mathbf{H}} \partial_t |U\rangle - i \frac{\beta_2}{2} \partial_t^2 |U\rangle = 0 \quad (2.14)$$

which only includes the effects of birefringence, PMD, and GVD. We start with group velocity dispersion which is independent of polarization.

2.2.1 Group-Velocity Dispersion

When there are no effects besides GVD, the evolution equation of the optical field $|U\rangle$ becomes

$$\partial_z |U\rangle - i \frac{\beta_2}{2} \partial_t^2 |U\rangle = 0 \quad (2.15)$$

in which β_2 is called the GVD coefficient. It is usually given in units of ps^2/km , but is more illustrative when shown as $\text{ps}/\text{GHz}/\text{km}$ or $\text{fs}/\text{GHz}/\text{km}$. As such units imply, GVD time-shifts the various spectral components relative to each other while the signal propagates along the fiber. The evolution equation (2.15) can be solved very efficiently in the Fourier domain, in which the time differential becomes a frequency-dependent factor, so that we have $\hat{\mathbf{S}}$

$$|\tilde{U}(z, \Delta\omega)\rangle = |\tilde{U}(0, \Delta\omega)\rangle \exp\left(-i \frac{\beta_2}{2} \Delta\omega^2 z\right) \quad (2.16)$$

where $\Delta\omega = \omega - \omega_0$ as before and $|\tilde{U}\rangle$ is the Fourier transform of $|U\rangle$. To get the time-domain optical field at z , one then only needs to take the inverse Fourier transform of $|\tilde{U}\rangle$ at z .

GVD manifests both as intrachannel and interchannel effect. Within a single channel, it causes broadening and distortion of optical pulses because the spectral components that constitute such a pulse propagate with different velocities along the fiber. The evolution with increasing accumulated dispersion (distance) of a typical pulse as used in this work is shown in Fig. 2.2. Because the spectrum of the pulse is relatively narrow, it maintains a pulse-like shape for a considerable distance before significantly broadening. When the optical field of such a pulse leaks out of its allocated time slot in the sequence of pulses, it interferes with its neighboring pulses, forming interference patterns such as the peak develop-

ing in the zero slot of a 1-0-1 sequence of non-return-to-zero (NRZ) pulses, as shown in Fig. 2.3.

The accumulated GVD in the systems discussed herein will usually be comparatively small wherever the optical power is high enough to cause significant nonlinear distortion. Therefore, we will regularly neglect the intrachannel aspect of GVD. A more important result of GVD, at least in the context of the present work, is the channel-dependent group velocity in a WDM system.

After introducing the retarded reference frame coordinates ζ and τ in the NLSE (2.6), the group delay at the reference frequency ω_0 was eliminated. However, all other frequencies experience a residual group delay due to GVD. To quantify this residual delay, we expand the phase constant $\beta(\Delta\omega)$ into a Taylor series in terms of $\Delta\omega = \omega - \omega_0$ about $\Delta\omega = 0$ (corresponding to ω_0),

$$\beta(\Delta\omega) = \beta(0) + \partial_{\Delta\omega}\beta\Big|_0 \Delta\omega + \frac{1}{2} \partial_{\Delta\omega}^2\beta\Big|_0 (\Delta\omega)^2 \quad (2.17)$$

$$= \beta_0 + \beta_1 \Delta\omega + \frac{\beta_2}{2} (\Delta\omega)^2 \quad (2.18)$$

stopping at the term of second order. The transformation (2.5) into the baseband and the introduction of the retarded frame of reference have already eliminated β_0 and β_1 at $\Delta\omega = 0$. The residual group delays per unit length at the carrier frequencies $\Delta\omega_\nu$ of the WDM channels other than ρ then become

$$\partial_{\Delta\omega}\beta\Big|_{\Delta\omega_\nu} = \partial_{\Delta\omega} \left[\frac{\beta_2}{2} (\Delta\omega)^2 \right]_{\Delta\omega_\nu} = \beta_2 \Delta\omega_\nu \quad (2.19)$$

If we had not neglected the third-order dispersion term in the NLSE (2.7), β_3 would appear in (2.19) as coefficient for the term of second order in $\Delta\omega$. With $\partial_{\Delta\omega}\beta$ we can now determine the *walk-off* Δt_ν between channels ρ and ν simply by integrating the residual group delay per unit length over the propagated distance (thus accounting for possible variations of β with z),

$$\Delta t_\nu(z_1, z_2) = \int_{z_1}^{z_2} \partial_{\Delta\omega}\beta(\zeta)\Big|_{\Delta\omega_\nu} d\zeta = \Delta\omega_\nu \int_{z_1}^{z_2} \beta_2(\zeta) d\zeta \quad (2.20)$$

The walk-off between the probe and interfering channels has a large influence on the magnitude of the interchannel nonlinear distortions, as will be shown in Section 3.2.2. Therefore we introduce a new fiber- and channel-dependent length scale, the walk-off length of channel ν , $L_{\text{WO}\nu}$, which quantifies the length of fiber that must be propagated in order for channel ν to walk off by one symbol from channel ρ ,

$$L_{\text{WO}\nu} = \frac{T_S}{|\beta_2| \Delta\omega_\nu} \quad (2.21)$$

where T_S is the symbol duration (e.g. 100 ps for 10 Gbaud signals) and β_2 is assumed constant within $L_{\text{WO}\nu}$. The walk-off length vs. the channel separation is plotted in Fig. 2.4 for three common values of β_2 . In a standard single-mode fiber ($\beta_2 = -128$ fs/GHz/km), the immediate neighbors of ρ in a

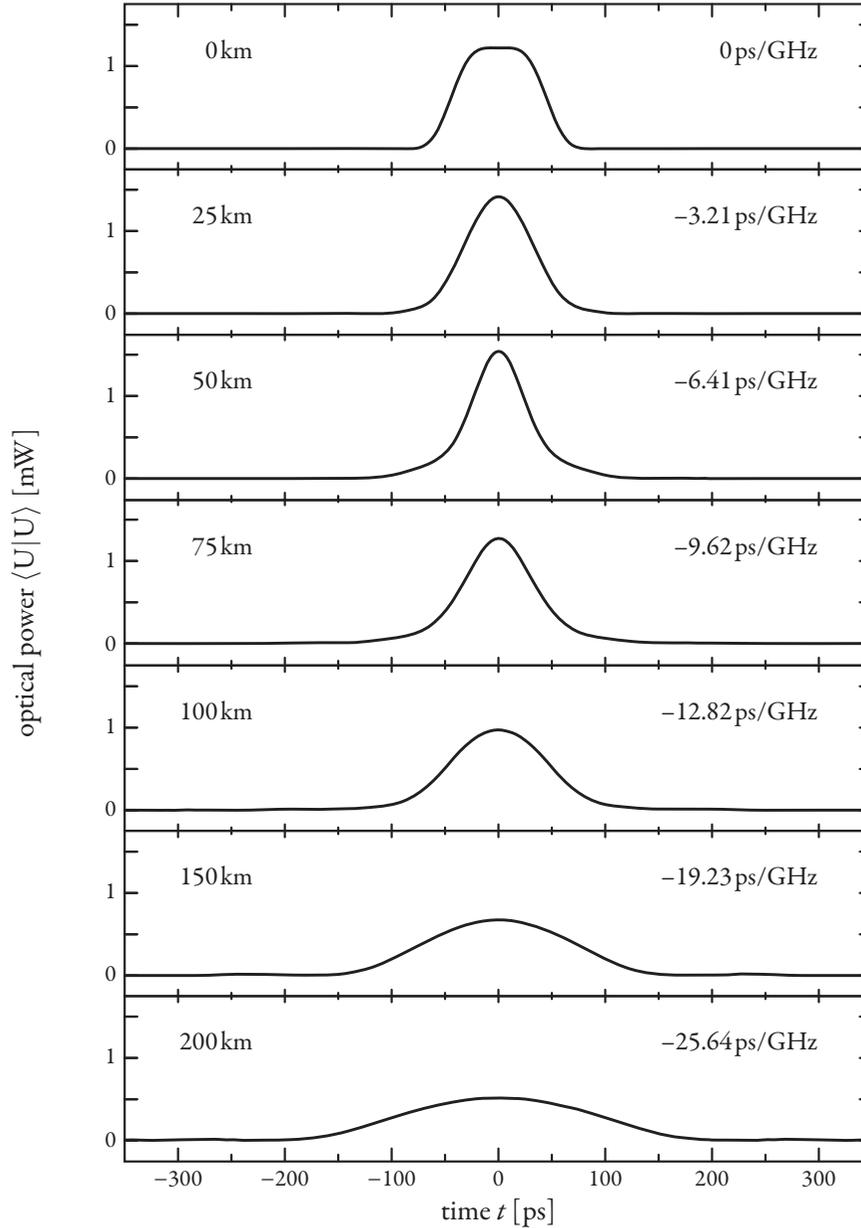


Figure 2.2 shows the evolution of a single pulse as consequence of GVD. The pulse shown here is typical for the data modulation parameters of the interfering channels used in the present work. Pulse width is 100 ps (corresponding to a bit rate of 10 Gbps), modulation current 10-to-90% rise time is 25 ps, and the pulse is optically filtered with a 2nd-order Gaussian filter of 25 GHz bandwidth.

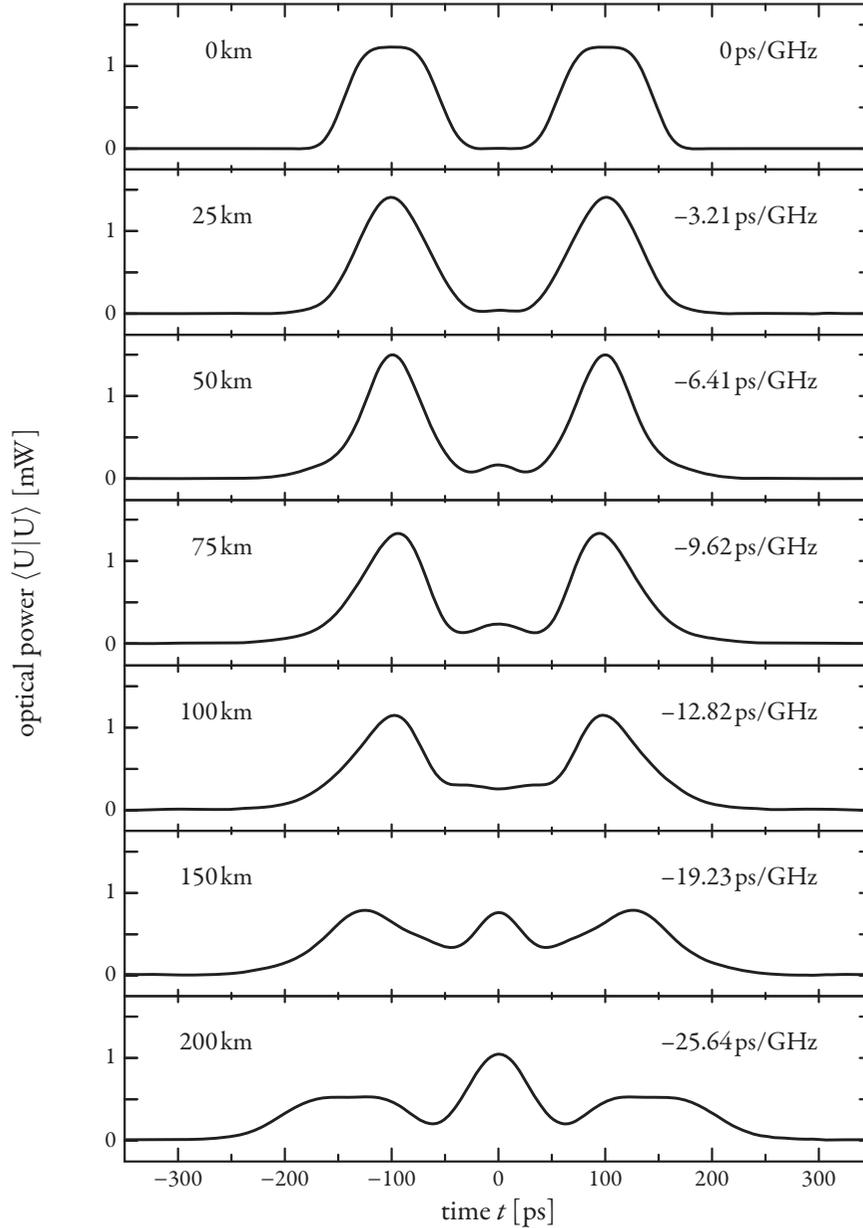


Figure 2.3 shows the evolution of a 1-0-1 pulse sequence as consequence of GVD. Pulse characteristics are the same as in the previous figure. At longer transmission distances, significant parts of the pulses overlap and their individual optical fields interfere with each other, leading to interference patterns in the form of a pronounced peak at the location of the interpulse-gap. This is known as inter-symbol interference (ISI).

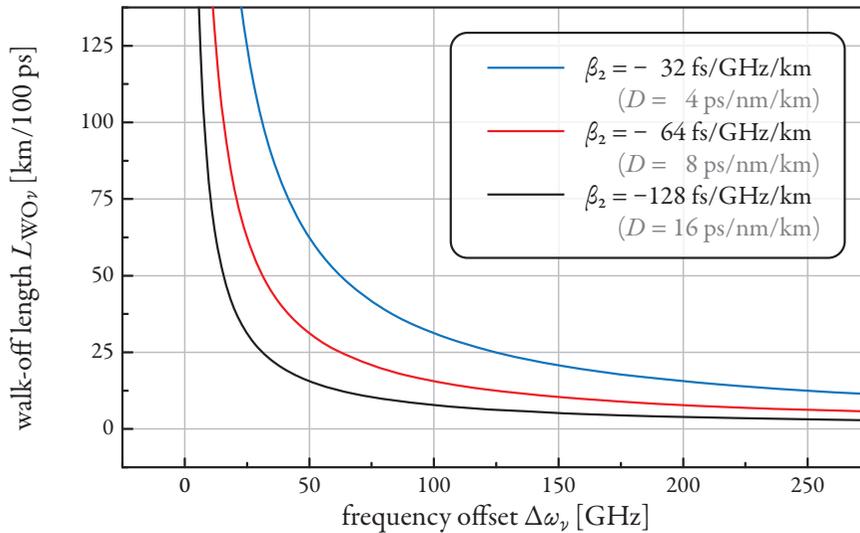


Figure 2.4 shows how the walk-off length for 100 ps symbols depends on the frequency separation of channels ρ and ν for three common fiber GVD parameters.

50 GHz grid have a walk-off length of about 16 km. Such walk-offs are easily calculated by using β_2 with units as described above. However, for historical reasons GVD is in practice almost exclusively quantified in terms of the wavelength instead of frequency. The corresponding parameter is

$$D_{\text{GVD}} = \partial_\lambda \beta_1 = -\frac{2\pi c}{\lambda^2} \beta_2$$

For the SSMF above this corresponds to $D_{\text{GVD}} = 16$ ps/nm/km. For consistency, we will mostly use β_2 to describe GVD throughout this work.

Next, we will (thoroughly) examine the linear polarization effects.

2.2.2 Birefringence

Fiber birefringence at its most basic «refers to slight variations in the indices of refraction along different axes of the fiber» [Keiser 2003]. It is caused by an anisotropy of the medium, which can be inherent due to the crystal structure of a material such as calcite (CaCO_3) or by external causes such as asymmetry of the waveguide, mechanical stress, or applied electrical fields, which can induce birefringence even in inherently isotropic materials such as the glass used for optical fiber.

To isolate birefringence in the linear Schrödinger equation, we again set all non-related coefficients

to zero and obtain the following propagation equation:

$$\partial_z |U(z, t)\rangle + i \frac{\Delta\beta_0}{2} \bar{\mathbf{H}} |U(z, t)\rangle = 0 \quad (2.22)$$

There are two parameters in (2.22) that quantify the birefringence of the fiber. Both depend on the nature and orientation of the cause of the anisotropy, and thus they may change with distance z – albeit slowly compared to the wavelength or the length of the optical pulses – so that $\Delta\beta_0 = \Delta\beta_0(z)$ and $\bar{\mathbf{H}} = \bar{\mathbf{H}}(z)$ is implied even when this dependence is not explicitly given. In the present section we address the effects of frequency-independent birefringence, with $\Delta\beta_0 = \Delta\beta(\Delta\omega = 0)$. The frequency dependence of fiber birefringence $\Delta\beta$ is quantified by $\Delta\beta_1$ – it is the physical reason for PMD, which is described in the subsequent section.

To find the physical meaning of these two parameters, we can (at least symbolically) integrate (2.22) and write

$$|U(z, t)\rangle = \bar{\mathbf{U}}(z) |U(0, t)\rangle \quad (2.23)$$

in which $\bar{\mathbf{U}}$ is a unitary transmission matrix [Gordon 2000]. To gain some insight into birefringence and discover the physical meaning of the parameters, we will initially constrain our description to a single short section of fiber, for which both parameters may be assumed constant, so that

$$\Delta\beta_0(z) = \Delta\beta_0^{[1]} \quad \text{and} \quad \bar{\mathbf{H}}(z) = \bar{\mathbf{H}}_{[1]} \quad (2.24)$$

From here on, all quantities that are valid only when (2.24) holds are designated by a bracketed sub- or superscript number which denotes the index of the fiber section with constant birefringence. Now we can properly integrate (2.22) over the length $L_{[1]}$ of the fiber section and have

$$\bar{\mathbf{U}}(L_{[1]}) = \bar{\mathbf{U}}_{[1]} = \exp\left(-i \frac{\Delta\beta_0^{[1]}}{2} L_{[1]} \bar{\mathbf{H}}_{[1]}\right) \quad (2.25)$$

Because $\bar{\mathbf{U}}_{[1]}$ is unitary, $\bar{\mathbf{H}}_{[1]}$ must be hermitian, $\bar{\mathbf{H}}_{[1]}^\dagger = \bar{\mathbf{H}}_{[1]}$. It is traceless⁷ and can be expanded in terms of the Pauli matrices [Arfken 2005, Sec. 3.4])

$$\bar{\mathbf{H}} = \hat{\mathbf{H}} \cdot \vec{\sigma} = b_1 \sigma_1 + b_2 \sigma_2 + b_3 \sigma_3 = \begin{pmatrix} b_1 & b_2 - ib_3 \\ b_2 + ib_3 & -b_1 \end{pmatrix} \quad (2.26)$$

where the coefficients b_n are real and $b_1^2 + b_2^2 + b_3^2 = 1$ so that $\hat{\mathbf{H}} = (b_1, b_2, b_3)^\top$ is a unit-length Stokes vector.

The yet unexplained matrix $\bar{\mathbf{H}}$ describes the eigenmode structure of the fiber segment. Its associated

⁷The trace of $\bar{\mathbf{H}}$ relates to the common propagation constant β_0 which has been removed in the transformation to the baseband (2.5).

(normalized) eigenvectors are

$$|e_+\rangle = \frac{1}{\sqrt{2+2b_1}} \begin{pmatrix} b_1+1 \\ b_2+ib_3 \end{pmatrix} \quad (2.27a)$$

$$|e_-\rangle = \frac{1}{\sqrt{2-2b_1}} \begin{pmatrix} b_1-1 \\ b_2+ib_3 \end{pmatrix} \quad (2.27b)$$

corresponding to the eigenvalues $+1$ and -1 , respectively. They also form a complete orthonormal basis set in Jones space. Transformation of $|e_+\rangle$ and $|e_-\rangle$ into Stokes space via (1.38) reveals that

$$\hat{\mathbf{H}} = \langle e_+ | \vec{\sigma} | e_+ \rangle \quad \text{and} \quad -\hat{\mathbf{H}} = \langle e_- | \vec{\sigma} | e_- \rangle \quad (2.28)$$

i.e. the eigenvector $|e_+\rangle$ is the Jones space equivalent to the Stokes vector $\hat{\mathbf{H}}$ of the Pauli matrix expansion (2.26). Expression (2.27) for $|e_+\rangle$ is a different way to write (1.34) of Section 1.1, giving an expression for the Jones vector corresponding to a given Stokes vector.

To find $\bar{\mathbf{U}}_{[1]}$, we need to find the matrix exponential $\exp \bar{\mathbf{H}}$, defined for square matrices as the power series expansion

$$\exp \bar{\mathbf{H}} = \sum_n \frac{1}{n!} \bar{\mathbf{H}}^n \quad (2.29)$$

Because it is hermitian, we can diagonalize $\bar{\mathbf{H}}$ with a unitary matrix $\bar{\mathbf{V}}$,

$$\bar{\mathbf{H}} = \bar{\mathbf{V}} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \bar{\mathbf{V}}^\dagger \quad (2.30)$$

where the columns of $\bar{\mathbf{V}}$ are the eigenvectors (2.27) of $\bar{\mathbf{H}}$, and λ_1 and λ_2 are its eigenvalues. We can then write the matrix exponential as [Meyer 2000, Sec. 7.3]

$$\exp \bar{\mathbf{H}} = \bar{\mathbf{V}} \begin{pmatrix} \exp \lambda_1 & 0 \\ 0 & \exp \lambda_2 \end{pmatrix} \bar{\mathbf{V}}^\dagger \quad (2.31)$$

so that with $\lambda_1 = +1$ and $\lambda_2 = -1$

$$\bar{\mathbf{U}}_{[1]} = \bar{\mathbf{V}}_{[1]} \begin{pmatrix} \exp(-i\frac{\Delta\beta_0^{[1]}}{2}L_{[1]}) & 0 \\ 0 & \exp(i\frac{\Delta\beta_0^{[1]}}{2}L_{[1]}) \end{pmatrix} \bar{\mathbf{V}}_{[1]}^\dagger \quad (2.32)$$

The physical meaning of $\bar{\mathbf{U}}$ is the following: The similarity transform $\mathbf{V}^\dagger = \mathbf{V}^{-1}$ transforms the Jones field $|U\rangle$ from the x - y -basis of (2.23) into the basis set defined by the eigenvectors of $\bar{\mathbf{H}}$, (2.27). There both components receive equal but opposite phase shifts. The resulting field is then transformed back into the x - y -basis by means of $\bar{\mathbf{V}}$.

We can avoid the similarity transform by directly expanding $|U\rangle$ at the segment input into a superposition of the eigenmodes of $\bar{\mathbf{H}}_{[1]}$ according to (2.4),

$$|U(0, t)\rangle = U_+(0, t)|e_+^{[1]}\rangle + U_-(0, t)|e_-^{[1]}\rangle \quad (2.33)$$

so that at the segment output

$$|U(L_{[1]}, t)\rangle = \exp\left(-i\frac{\Delta\beta_0^{[1]}}{2}L_{[1]}\right)U_+(0, t)|e_+^{[1]}\rangle + \exp\left(i\frac{\Delta\beta_0^{[1]}}{2}L_{[1]}\right)U_-(0, t)|e_-^{[1]}\rangle \quad (2.34)$$

That is, U_+ and U_- propagate with propagation constants which are equal in magnitude but opposite in sign. They are parameterized by

$$\Delta\beta_0 = \beta_0(|e_+\rangle) - \beta_0(|e_-\rangle) \quad (2.35)$$

being the second birefringence parameter quantifying the strength of birefringence in (2.22). The accumulated phase difference between both eigenmodes is

$$\Delta\psi_{[1]} = \Delta\beta_0^{[1]}L_{[1]} \quad (2.36)$$

The transmission matrix (2.32) is periodic with a period of $\Delta\psi_{[1]} = 2\pi$. The fiber length L_b corresponding to this phase difference is called the *beat length*, such that

$$\Delta\beta_0^{[1]}L_b^{[1]} = 2\pi$$

or more generally,

$$L_b = \frac{2\pi}{\Delta\beta_0} \quad (2.37)$$

That is, within one beat length of (constant birefringence) fiber, the phase difference $\Delta\psi$ that has accumulated between the local eigenstates equals one cycle of the optical carrier (β_0 and $\Delta\beta_0$ refer to the propagation constants at the carrier frequency). Typical beat lengths for contemporary optical fibers are in the range of some 10–30 meters for single-mode transmission fiber and 1–2 mm for polarization-maintaining fiber (PMF) [Galtarossa 2000, Irvine-Halliday 2000]. Comparison with the theory of the previous chapter reveals that the short fiber segment has the same Jones matrix as a waveplate with slow axis $|e_+^{[1]}\rangle$ and retardation angle $\Delta\beta_0^{[1]}L_{[1]}$. The effect of $\bar{\mathbf{U}}$ is thus merely a periodic transformation of the SOP of $|U\rangle$ with the period being the beat length.

STOKES SPACE In general, processes which merely change the polarization of a propagating wave can be better visualized in Stokes space where we deal with real quantities in a 3-dimensional vector space instead of the complex 2-dimensional Jones space. We already know that the unitary matrix $\bar{\mathbf{U}}$ corresponds to a rotational motion in Stokes space. To quantify this motion, we transform (2.22) into Stokes

space by following [Gordon 2000, Sec. 6]. We obtain the evolution of the Stokes vector $\mathbf{S} = \langle \mathbf{U} | \vec{\sigma} | \mathbf{U} \rangle$ with the z -derivative

$$\partial_z \mathbf{S} = \left(\partial_z \langle \mathbf{U} | \right) \vec{\sigma} | \mathbf{U} \rangle + \langle \mathbf{U} | \vec{\sigma} \left(\partial_z | \mathbf{U} \rangle \right) \quad (2.38)$$

Inserting (2.22) and (2.26),

$$\partial_z \mathbf{S} = \langle \mathbf{U} | \left(i \frac{\Delta\beta_0}{2} \hat{\mathbf{H}} \cdot \vec{\sigma} \right) \vec{\sigma} | \mathbf{U} \rangle + \langle \mathbf{U} | \vec{\sigma} \left(-i \frac{\Delta\beta_0}{2} \hat{\mathbf{H}} \cdot \vec{\sigma} \right) | \mathbf{U} \rangle \quad (2.39)$$

Using the spin vector rules [Gordon 2000, eqs. (A.3), (A.4), and (A.13)],

$$\vec{\sigma} (\hat{\mathbf{H}} \cdot \vec{\sigma}) = \hat{\mathbf{H}} \vec{\mathbf{I}} + i \hat{\mathbf{H}} \times \vec{\sigma} \quad (2.40a)$$

$$(\hat{\mathbf{H}} \cdot \vec{\sigma}) \vec{\sigma} = \hat{\mathbf{H}} \vec{\mathbf{I}} - i \hat{\mathbf{H}} \times \vec{\sigma} \quad (2.40b)$$

$$\langle \mathbf{U} | \hat{\mathbf{H}} \times \vec{\sigma} | \mathbf{U} \rangle = \hat{\mathbf{H}} \times \mathbf{S} \quad (2.40c)$$

we have

$$\begin{aligned} \partial_z \mathbf{S} &= \Delta\beta_0 \langle \mathbf{U} | \hat{\mathbf{H}} \times \vec{\sigma} | \mathbf{U} \rangle \\ &= \mathbf{H} \times \mathbf{S} \end{aligned} \quad (2.41)$$

with the *birefringence vector*

$$\mathbf{H} = \Delta\beta_0 \hat{\mathbf{H}} \quad (2.42)$$

which points into the direction of the slow eigenmode and has a length corresponding to the phase change per unit length of fiber.⁸ The expression $\vec{\mathbf{I}}$ in (2.40) is to be interpreted as

$$\vec{\mathbf{I}} = \begin{pmatrix} \bar{\mathbf{I}} \\ \bar{\mathbf{I}} \\ \bar{\mathbf{I}} \end{pmatrix} \quad (2.43)$$

It has the same shape as the Pauli vector. Gordon and Kogelnik write the spin vector rules using the identity matrix $\bar{\mathbf{I}}$, which may lead to some confusion.

Expression (2.41) describes the Stokes space rotation of \mathbf{S} around the local birefringence vector \mathbf{H} . The rotation rate is $|\mathbf{H}| = \Delta\beta_0$ per unit length of fiber. Thus, the optical field cycles through a series of SOPs as the signal propagates along the fiber, as in Jones space. The period of this rotation is again the beat length (2.2.2). This rotational motion can be integrated to yield

$$\mathbf{S}(L_{[1]}, t) = \bar{\mathbf{R}}_{[1]} \mathbf{S}(0, t) \quad (2.44)$$

⁸In much of the literature on the topic, this vector is labeled $\vec{\beta}$. However, we have chosen \mathbf{H} in order to avoid any confusion between its elements b_i and the propagation constant derivatives β_i , and to emphasize that the matrix associated via (2.26) is Hermitian.

where the 3×3 matrix $\bar{\mathbf{R}}_{[1]}$ is the Stokes space counterpart of the Jones matrix $\bar{\mathbf{U}}_{[1]}$ in (2.25). It, too, has a very concise definition [Gordon 2000, eq. (4.9)],

$$\bar{\mathbf{R}}_{[1]} = \exp[L_{[1]} \mathbf{H} \times] = \exp[\Delta\psi_{[1]} \hat{\mathbf{H}} \times] \quad (2.45)$$

with a rotation angle $\Delta\psi_{[1]}$, initially defined in (2.36) as the phase difference between the eigenmodes. The argument of the exponential is called a cross-product operator. It is defined as [Gordon 2000, eq. (4.8)]

$$\hat{\mathbf{H}} \times = \begin{pmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{pmatrix} \quad (2.46)$$

where the b_n are again the components of the normalized Stokes vector $\hat{\mathbf{H}}$. The exponential is interpreted as a power series expansion, as before.

The Stokes space rotation (2.41) is generally more intuitive than the Jones expression (2.34), which describes an abstract phase change between field components defined by eigenvectors of a transmission matrix. If the incident optical field is polarized along one of the birefringence eigenmodes, \mathbf{S} and the rotation axis \mathbf{H} are coincident and no change in polarization occurs, equivalent to the Jones space behavior.

The evolution of an exemplary SOP ($\chi = \pi/4$ and $\phi = 0$, corresponding to $|e_{45^\circ}\rangle$) within one beat length is illustrated in Fig. 2.5.

MULTIPLE SEGMENTS A real fiber (except for PMFs) unfortunately does not behave like a single birefringent segment, as we have assumed until now. As the causes of birefringence change over the fiber length⁹, so does the birefringence vector \mathbf{H} . We can model a real fiber as a concatenation of multiple segments in which the birefringence properties are locally constant and then let these segments become infinitesimally short while increasing their number accordingly. To illustrate the consequences, we will start with the concatenation of just two segments of length $L_{[1]}$ and $L_{[2]}$, and with birefringence vectors $\mathbf{H}_{[1]}$ and $\mathbf{H}_{[2]}$.

If the eigenmode matrices of both segments do not commute because their eigenmodes are different, we have to integrate (2.22) segment-wise and obtain the concatenation matrix (1.48) in Jones space

$$\bar{\mathbf{U}}(L_{[1+2]}) = \bar{\mathbf{U}}_{[1+2]} = \bar{\mathbf{U}}_{[2]} \bar{\mathbf{U}}_{[1]} \quad (2.47)$$

That is, the field is initially transformed by the transmission matrix $\bar{\mathbf{U}}_{[1]}$ of the first segment, and the output field of that segment is consequently transformed by $\bar{\mathbf{U}}_{[2]}$.

The generalization to any number of segments is straightforward and results in an arbitrarily long

⁹These parameters also vary over time, albeit very slowly (compared to the symbol length). For simplicity we omit their explicit dependence on t and instead include that in the concept of the stochastic fiber ensemble, introduced in the next section.

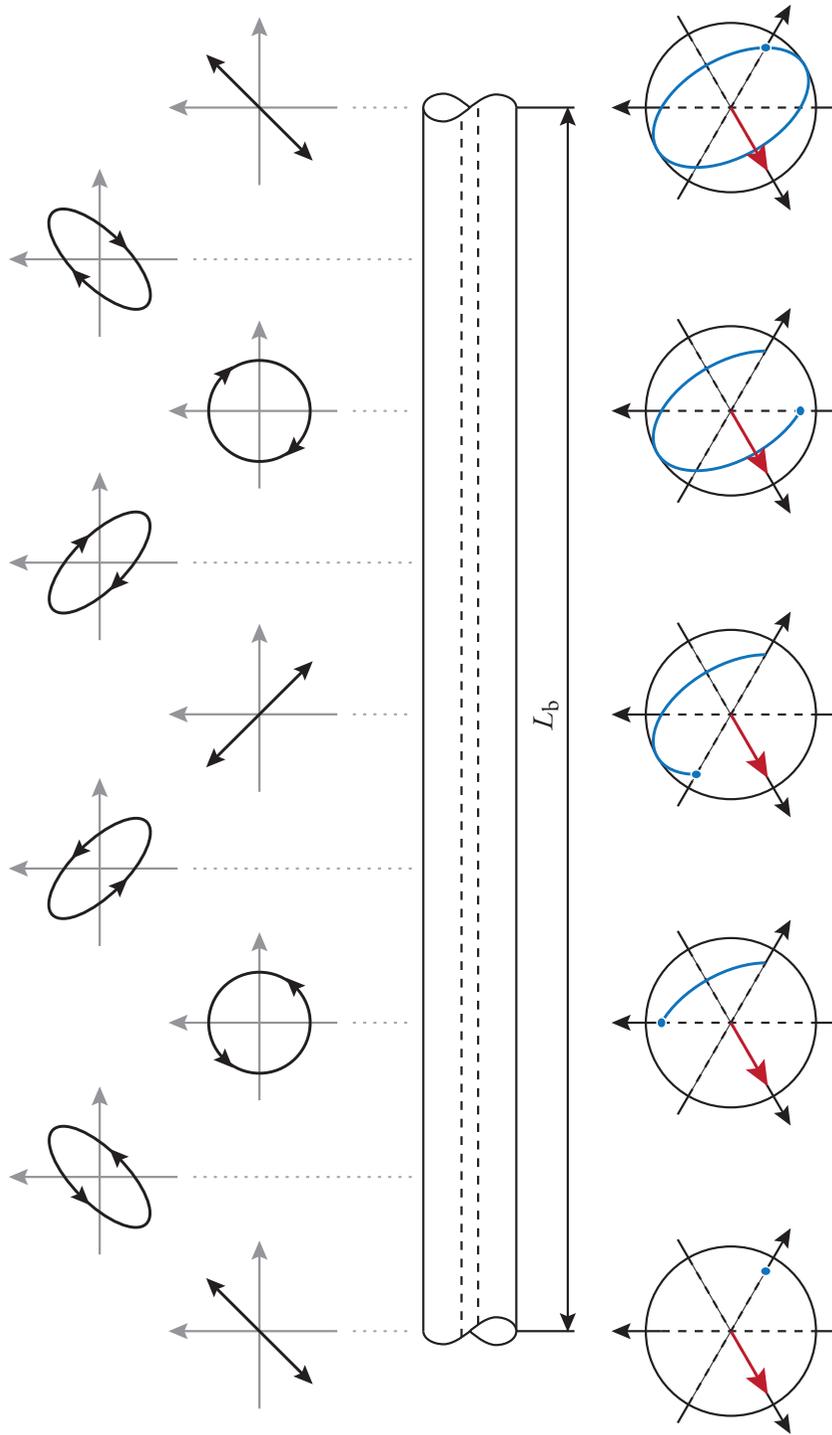


Figure 2.5 illustrates the effect of fiber birefringence on the SOP of a propagating wave within one beat length L_b in Jones (top) and Stokes (bottom) space. The red arrow is the birefringence vector $\hat{\mathbf{H}}$.

expression for $\bar{\mathbf{U}}$. Therefore, we introduce a formalism introduced in [Feynman 1951] and write

$$\bar{\mathbf{U}}(z) = \exp \left[-i \int_0^z \frac{\Delta\beta_0(\zeta)}{2} \bar{\mathbf{H}}(\zeta) d\zeta \right] \quad (2.48)$$

for infinitesimally short segments. Because the matrices $\bar{\mathbf{H}}$ generally do not commute and thus cannot be integrated in a regular manner, the above integral expression has to be understood in a symbolic fashion, as the generalization of (2.47). A formal expression for $\bar{\mathbf{U}}$ can be obtained by expanding the integral into a Magnus series [Magnus 1954, Korneev 2002].

The two-segment behavior is similar in Stokes space, where the SOP first rotates around the birefringence vector $\mathbf{H}_{[1]}$ of the first segment and thereafter around $\mathbf{H}_{[2]}$,

$$\bar{\mathbf{R}}(L_{[1+2]}) = \bar{\mathbf{R}}_{[1+2]} = \bar{\mathbf{R}}_{[2]} \bar{\mathbf{R}}_{[1]} \quad (2.49)$$

as in (1.48). In analogy to the Jones algebra, we could generalize (2.49) into an integral expression similar to [Gordon 2000, Sec. 7], or write the rotation matrix $\bar{\mathbf{R}}$ analog to (2.48),

$$\bar{\mathbf{R}}(z) = \exp \left[\int_0^z \Delta\beta_0(\zeta) \bar{\mathbf{H}}(\zeta) \times d\zeta \right] \quad (2.50)$$

where $\bar{\mathbf{H}}(\zeta) \times$ is the position-dependent cross-product operator as defined in (2.46)

COORDINATE TRANSFORMATION When looking at (2.23), we find that we can define a unitary field transformation

$$|S(z, t)\rangle = \bar{\mathbf{U}}^\dagger(z) |U(z, t)\rangle = |U(0, t)\rangle \quad (2.51a)$$

so that

$$\partial_z |S\rangle = 0 \quad (2.51b)$$

The effect of $\bar{\mathbf{U}}^\dagger$ is to transform the SOP of the field $|U\rangle$ back to the initial $|U(0, t)\rangle$ at every position z by means of a coordinate transformation. The so obtained field $|S\rangle$ no longer changes its SOP in the new coordinates – we have mathematically stopped the birefringence-induced SOP change by proper choice of coordinates.

It is important to note that the field $|S\rangle$ is no longer defined in absolute Jones (x, y) coordinates, sometimes called laboratory coordinates, but in a z -dependent coordinate system which changes depending on the birefringence properties of the fiber. Its coordinates are called local (fiber) coordinates.

It is also straightforward to visualize the coordinate transformation (2.51a) from $|U\rangle$ to $|S\rangle$ in Stokes space: The 3-dimensional coordinate system also rotates around the local birefringence vector \mathbf{H} , however it does so in the opposite direction, as indicated by $\bar{\mathbf{U}}^\dagger = \bar{\mathbf{U}}^{-1}$. The compound motion of the so rotating coordinate system and the Stokes vector \mathbf{S} within this coordinate system result in zero

net motion, in accordance with (2.51b). In this process, however, the coordinate axes in Stokes space, corresponding to certain SOPs in Jones space, lose their fixed meaning. Depending on the orientation of the birefringence vector and the propagated distance z , a fixed point in this newly defined space may correspond to an arbitrary SOP in laboratory coordinates. As we will show presently, not knowing the laboratory SOP poses no problems when using polarization insensitive (direct-detection) receivers, but impairs POLDM systems.

Applying (2.51a) to the NLSE (2.9), we obtain the propagation equation in local coordinates,

$$\partial_z |S\rangle + \frac{\Delta\beta_1}{2} \bar{\Omega} \partial_t |S\rangle - i \frac{\beta_2}{2} \partial_t^2 |S\rangle - i \exp(-\alpha z) \gamma \left[\langle S|S\rangle |S\rangle - \frac{1}{3} \left(\langle S|\bar{Y}|S\rangle \right) \bar{Y}|S\rangle \right] = 0 \quad (2.52a)$$

with

$$\bar{\Omega} = \bar{U}^\dagger \bar{H} \bar{U} \quad (2.52b)$$

and

$$\bar{Y} = \bar{U}^\dagger \sigma_3 \bar{U} \quad (2.52c)$$

where we have used $\partial_t \bar{U} = 0$ – the birefringence properties are assumed constant in time.

The matrices $\bar{\Omega}$ and \bar{Y} describe the eigenmode structures of the linear and nonlinear mode coupling in local coordinates, respectively – the similarity transform \bar{U} in (2.52) quasi-randomizes them in such a way that their eigenvectors are distributed uniformly on the Poincaré sphere, regardless of the underlying distribution of the original eigenvectors of \bar{H} .

POLARIZATION-DIVISION MULTIPLEX In the case of POLDM transmission we do need to know the output SOP $|e_1(L)\rangle$ in laboratory coordinates, related to the subchannel input SOP $|e_1(0)\rangle$ via

$$|e_1(L)\rangle = \bar{U}(L) |e_1(0)\rangle \quad (2.53)$$

where \bar{U} is the fiber transmission Jones matrix (2.48), in order to demultiplex the POLDM subchannels. In a practical system the matrix \bar{U} cannot be predetermined, because the birefringence vector \mathbf{H} is not known for all z along the link (and may also change over time). Therefore some sort of search algorithm must be employed at the receiver to find and track the subchannel output SOPs. This has long been the largest obstacle in the implementation of POLDM systems, because reliable opto- or electro-mechanical implementations were not yet commercially available.¹⁰ For purely electronic demultiplexing, the dual-polarization complex optical field must be known in the electrical domain, which is not possible with direct detection.

¹⁰POLDM has been realized in direct-detection systems, at least in the laboratory, by aligning the polarization filters to suppress an initially inserted RF tone of the unwanted subchannel [Chraplyvy1996] or to minimize broadband RF interference noise [Milivojevic 2005].

If the polarization beam splitter (PBS) at the receiver – be it a beam splitter with preceding adjustable polarization control or a post-detection signal processing splitter based on adaptive linear filters operating on the complex received field – is misaligned to the output subchannel SOPs, one immediately incurs signal fading and crosstalk between the subchannels. If we call the eigenmodes of the misaligned PBS $|e'_1(L)\rangle$ and $|e'_2(L)\rangle$, we can determine the fading and crosstalk with the help of (2.13) by replacing the ideal polarizer aligned to $|e_1(L)\rangle$ with the real one aligned to $|e'_1(L)\rangle$ to obtain

$$|e'_1\rangle\langle e'_1|U\rangle = F U_1|e'_1\rangle + X U_2|e'_1\rangle \quad (2.54a)$$

in which

$$F = \langle e'_1|e_1\rangle \quad \text{and} \quad X = \langle e'_1|e_2\rangle \quad (2.54b)$$

are a complex fading coefficient with $0 \leq |F| \leq 1$ and a complex crosstalk coefficient with $0 \leq |X| \leq 1$. We have dropped the position dependence of all variables for brevity and assume $z=L$ for the receiver throughout the remainder of this section. The coefficients for the other POLDM subchannel U_2 must, of course, be the same as those for U_1 , given above.

Since $|e_1\rangle$ and $|e_2\rangle$ constitute an orthonormal basis, we use Bessel's inequality to write [Riley 2006, eq. (8.21)]

$$|F| = \cos \frac{\Theta}{2} \quad \text{and} \quad |X| = \sin \frac{\Theta}{2} \quad (2.55)$$

so that a misalignment angle $\Theta/2=0$ when the polarizers are perfectly aligned.

From (2.55) we can see that crosstalk will generally cause worse distortions than fading, especially for small misalignment angles $\Theta/2$ when $\cos(\Theta/2) \approx 1$ and $\sin(\Theta/2) \approx \Theta/2$. The misalignment angle $\Theta/2$ is illustrated in Fig. 2.6 for the special case of both $|e_1\rangle$ and $|e'_1\rangle$ being linear polarization states (so that F and X become real quantities).¹¹

The misalignment angle – and thus the magnitude of the fading and crosstalk coefficients – can also be derived from a Stokes space representation of the PBS misalignment, where \hat{S}_1 and \hat{S}'_1 are the Stokes SOPs corresponding to the original $|e_1\rangle$ and the misaligned $|e'_1\rangle$, respectively. We have

$$|\langle e'_1|e_1\rangle|^2 = \cos^2 \frac{\Theta}{2} = \frac{1}{2}(1 + \hat{S}'_1 \cdot \hat{S}_1) \quad (2.56)$$

such that

$$\hat{S}'_1 \cdot \hat{S}_1 = \cos \Theta \quad (2.57)$$

that is, the misalignment angle Θ in Stokes space is twice that in Jones space. To arrive at (2.56), we used

¹¹The occurrence of PBS misalignment in POLDM transmission experiments has resulted in the postulation of the *20/20 rule-of-thumb* for tolerable crosstalk. The rule states that for $|X|^2 = -20$ dB (intensity) crosstalk – corresponding to a misalignment angle of $\Theta/2 \approx 6^\circ$ – the worst-case reduction of the upper rail of the eye pattern is about 20 percent, or 1 dB – the maximum tolerable amount [Nelson 2000].

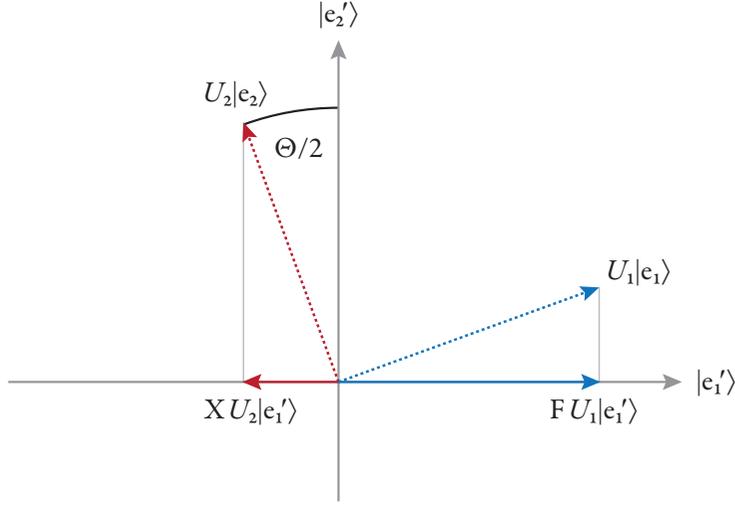


Figure 2.6 illustrates (2.54) in Jones space, assuming that $|e_1\rangle$, $|e_2\rangle$, $|e'_1\rangle$, and $|e'_2\rangle$ are all linear SOPs, so that F and X become real quantities – otherwise a phase shift may be associated with either or both.

$$|\langle e'_1 | e_1 \rangle|^2 = \langle e_1 | e'_1 \rangle \langle e'_1 | e_1 \rangle \quad (2.58a)$$

$$|\langle e'_1 | e_1 \rangle| = \frac{1}{2} (\bar{\mathbf{I}} + \hat{\mathbf{S}}'_1 \cdot \vec{\sigma}) \quad (2.58b)$$

$$\langle e_1 | e_1 \rangle = 1 \quad (2.58c)$$

$$\langle e_1 | \hat{\mathbf{S}}'_1 \cdot \vec{\sigma} | e_1 \rangle = \hat{\mathbf{S}}'_1 \cdot \hat{\mathbf{S}}_1 \quad (2.58d)$$

where (2.58b) and (2.58d) correspond to [Gordon 2000, eq. (3.9) and (A.13)], respectively.

The phase information in the complex F and X is lost when using the misalignment angle as in (2.55) and especially when considering SOPs in Stokes space, as in (2.57). The only way to determine their absolute phases is by using the Jones description (2.54) using the full optical field, not just the SOPs. We can, however, determine how the relative phase between F and X changes as $\hat{\mathbf{S}}_1$, corresponding to $|e_1\rangle$, evolves. The Stokes SOP $\hat{\mathbf{S}}_1$ can change in two ways on the Poincaré sphere: such that the misalignment angle Θ varies, or such that it remains constant – the latter case corresponds to a rotation around the Stokes vector of the misaligned receiver, $\hat{\mathbf{S}}'_1$. Since the first way will also result in a change of the magnitudes of F and X , we are more interested in the second. The corresponding motion of $\hat{\mathbf{S}}_1$ is described by

$$d_\varphi \hat{\mathbf{S}}_1 = \hat{\mathbf{S}}'_1 \times \hat{\mathbf{S}}_1 \quad (2.59)$$

with $d_\varphi = d/d\varphi$ and where φ describes the location of $\hat{\mathbf{S}}_1$ on the circle around $\hat{\mathbf{S}}'_1$ as illustrated in Fig. 2.7. It is only an auxiliary variable and has no absolute reference on this circle. We find that the following

differential equation describes the equivalent change of the normalized Jones vector corresponding to $\hat{\mathbf{S}}_1$:

$$d_\varphi |e_1\rangle = -i \bar{\mathbf{P}} |e_1\rangle \quad (2.60a)$$

with the Hermitian matrix

$$\bar{\mathbf{P}} = |e'_2\rangle\langle e'_2| = \frac{1}{2}(\bar{\mathbf{I}} + \hat{\mathbf{S}}'_2 \cdot \vec{\sigma}) \quad (2.60b)$$

with the projection operator $|e\rangle\langle e|$ introduced in (1.11). The equivalence can be verified by following the steps to get from (2.22) to (2.41). The choice of matrix $\bar{\mathbf{P}}$ is not fixed – any matrix whose eigenvectors correspond to $\hat{\mathbf{S}}'_1$ and $\hat{\mathbf{S}}'_2$ will work. As we will see, our particular choice will result in the phase of F being constant. We can then derive the change of F and X with φ ,

$$d_\varphi F = d_\varphi \langle e'_1 | e_1 \rangle = \left(d_\varphi \langle e'_1 | \right) | e_1 \rangle + \langle e'_1 | \left(d_\varphi | e_1 \rangle \right) \quad (2.61a)$$

$$d_\varphi X = d_\varphi \langle e'_2 | e_1 \rangle = \left(d_\varphi \langle e'_2 | \right) | e_1 \rangle + \langle e'_2 | \left(d_\varphi | e_1 \rangle \right) \quad (2.61b)$$

where we used $X = \langle e'_1 | e_2 \rangle = \langle e'_2 | e_1 \rangle$. Inserting (2.60) and recognizing that $|e'_1\rangle$ and $|e'_2\rangle$ are constant, we have

$$d_\varphi F = -i \langle e'_1 | e'_2 \rangle \langle e'_2 | e_1 \rangle = 0 \quad (2.62a)$$

$$d_\varphi X = -i \langle e'_2 | e'_2 \rangle \langle e'_2 | e_1 \rangle = -i X \quad (2.62b)$$

with the orthonormality conditions $\langle e'_1 | e'_2 \rangle = 0$ and $\langle e'_2 | e'_2 \rangle = 1$. Thus,

$$F = F_0 \quad (2.63a)$$

$$X = X_0 \exp(-i\varphi) \quad (2.63b)$$

i.e. the relative phase between F and X evolves as $\hat{\mathbf{S}}_1$ rotates around $\hat{\mathbf{S}}'_1$, and the circle of polarization states on the Poincaré sphere for which $\hat{\mathbf{S}}'_1 \cdot \hat{\mathbf{S}}_1$ is constant corresponds to all possible phases, even though a one-to-one mapping of an SOP constellation to the relative phase is not possible because the absolute (reference) phases are unknown. We will keep with the notation of the constant phase of F resulting from our choice of $\bar{\mathbf{P}}$ (similar to requiring the x -component of the Jones SOP in (1.5) to be real). It corresponds to the reasonable assumption that in the absence of the second POLDM subchannel, and thus any crosstalk into the first, the phase of the detected signal does not depend on changes of its SOP. We can then arbitrarily set the phase of F to zero, so that F becomes a real quantity and $|F| = F$. We will make use of these results in Chapter 5.

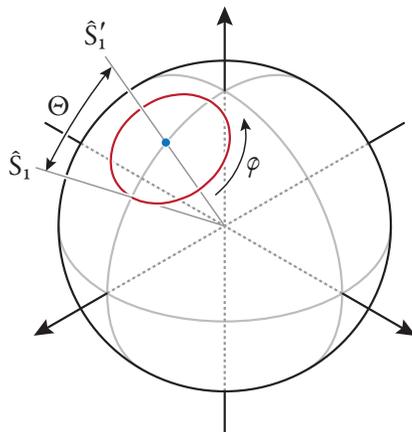


Figure 2.7 illustrates the misalignment angles in Stokes space: the circle around \hat{S}'_1 marks all SOPs \hat{S}_1 for which $\hat{S}'_1 \cdot \hat{S}_1 = \cos \Theta$. An increase of φ corresponds to a right-hand rotation of the position of \hat{S}_1 along the circle.

From (2.54) we see that the crosstalk field not only depends on the complex coefficient X , but also on the instantaneous field U_2 in the interfering POLDM subchannel. Thus, if one could reduce either X or U_2 (or preferably, both), the crosstalk impairment can be significantly improved. The trivial but spectrally inefficient solution would be to have $U_2 = 0$ and transmit a single-polarization signal. A more elegant solution can be obtained when the data-carrying symbols are RZ-shaped, with duty cycles $\leq 50\%$. In this case we can shift one POLDM subchannel with respect to the other by half the symbol duration, $T_S/2$. This way, the (interfering) field $U_2(m \cdot T_S)$ in one subchannel – and hence the crosstalk – will be weak whenever the other subchannel is sampled at $t = (m + 1/2) \cdot T_S$ and vice versa. Since the crosstalk is generated during POLDM demultiplex, this works whenever the pulse shapes approximate return-to-zero (RZ) at the receiver – i.e. when the total accumulated GVD of the transmission link is close to zero. When GVD is compensated electronically, both subchannels must be jointly compensated before the POLDM demultiplexer (DEMUX) in order for interleaving to be beneficial. The interleaving principle is illustrated in Fig. 2.8.

With the renaissance of coherent reception in combination with polarization-diversity (which is required in consequence of the signal SOP uncertainty caused by birefringence) as a prerequisite for extensive electronic post-processing, the vector optical field is available at the receiver, and electronic POLDM demultiplexing essentially «comes for free» (cf. [Winzer 2008] and references in the introductory chapter). This possibility of eliminating the (linear) PBS misalignment electronically (by continuously adapting electronic filters so that X is minimized) in combination with ever increasing demand for higher data rates has sparked the renewed interest in POLDM transmission.

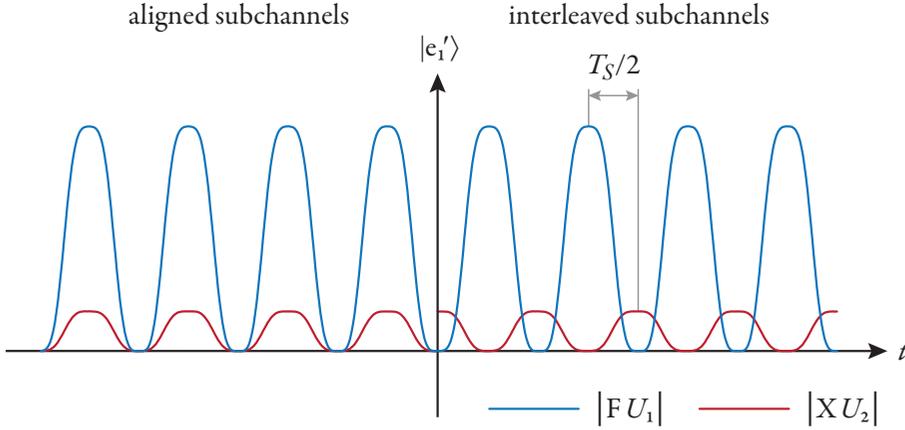


Figure 2.8 illustrates the principle behind polarization interleaving: The magnitude of the crosstalk field $X U_2$ (red) is (ideally) zero whenever the detected subchannel field $F U_1$ (blue) is sampled if both subchannels are temporally shifted by half a symbol, $T_S/2$. See also Fig. 2.6.

2.2.3 Polarization-Mode Dispersion

The topic of polarization mode dispersion is so broad and so much research has gone into it in the past years, that it easily fills entire books and book segments (e.g. [Poole 1997, Kogelnik 2002, Damask 2004, Galtarossa 2005, Brodsky 2008]), among hundreds of research papers – one of the most influential of these being [Gordon 2000]. We shall thus not venture to give an exhaustive description of it, but we will concentrate our discussion upon the basic concepts and those aspects that will be of importance within subsequent chapters of this work.

Note that there are actually two separate effects that are being included in the concept of PMD: on a physical level there exists a frequency dependence of the fiber birefringence, resulting in dispersion of the group velocities of the local fiber eigenmodes, and in a physical and statistical sense there are certain consequences of the longitudinal variation of these parameters. The former is generally referred to as *first-order* PMD, while the latter results in *higher-order* PMD.

We begin our analysis as before by dropping all non-PMD terms from the propagation equation (2.52) and are left with the first-order partial differential equation

$$\partial_z |S\rangle + \frac{\Delta\beta_1}{2} \bar{\Omega} \partial_t |S\rangle = 0 \quad (2.64)$$

The hermitian matrix $\bar{\Omega}$ was introduced in (2.52) and describes the eigenmode structure of the effect, and all of the related theory of the previous section applies to it. It is usually assumed to be frequency-independent [Menyuk 2006]. The other parameter in (2.64) is the magnitude parameter $\Delta\beta_1$, defined as

$$\Delta\beta_1 = \partial_\omega \Delta\beta \Big|_{\Delta\omega=0} \quad (2.65)$$

It describes a linear frequency dependence of the mode dispersion $\Delta\beta(\Delta\omega)$, whose value $\Delta\beta_0$ at $\Delta\omega = 0$ quantifies the birefringence strength. As for GVD in Section 2.2.1, we can obtain the solution of (2.64) by transformation into the frequency domain,

$$\partial_z |\tilde{\mathbf{S}}(z, \Delta\omega)\rangle - i \frac{\Delta\beta_1}{2} \Delta\omega \bar{\mathbf{\Omega}} |\tilde{\mathbf{S}}(z, \Delta\omega)\rangle = 0 \quad (2.66)$$

in which $|\tilde{\mathbf{S}}\rangle$ is the Fourier transform of $|\mathbf{S}\rangle$. Integrating (2.66) yields the expression

$$|\tilde{\mathbf{S}}(z, \Delta\omega)\rangle = \bar{\mathbf{T}}(z, \Delta\omega) |\tilde{\mathbf{S}}(0, \Delta\omega)\rangle \quad (2.67)$$

with a frequency-dependent unitary Jones matrix $\bar{\mathbf{T}}$. Both $\Delta\beta_1$ and $\bar{\mathbf{\Omega}}$ in (2.66) are generally z -dependent quantities. To simplify the investigation of $\bar{\mathbf{T}}$, we start once more with a sufficiently short segment of fiber for which both parameters may be assumed constant,

$$\Delta\beta_1(z) = \Delta\beta_1^{[1]} \quad \text{and} \quad \bar{\mathbf{\Omega}}(z) = \bar{\mathbf{\Omega}}_{[1]} \quad (2.68)$$

For the single section of length $L_{[1]}$ we then have

$$\bar{\mathbf{T}}(L_{[1]}, \Delta\omega) = \bar{\mathbf{T}}_{[1]}(\Delta\omega) = \exp\left(-i \frac{\Delta\beta_1^{[1]}}{2} L_{[1]} \Delta\omega \bar{\mathbf{\Omega}}_{[1]}\right) \quad (2.69)$$

For each frequency component of the signal $|\mathbf{S}\rangle$, this solution is very similar to the single-segment solution for birefringence (2.22). However, each frequency component experiences an individual magnitude of (residual) birefringence, $\Delta\beta_1^{[1]} z \Delta\omega/2$. We can rewrite (2.67) by expanding $|\tilde{\mathbf{S}}\rangle$ into the frequency-independent eigenmodes of $\bar{\mathbf{\Omega}}_{[1]}$, analog to (2.34),

$$|\tilde{\mathbf{S}}(L_{[1]}, \Delta\omega)\rangle = \exp\left[-i \frac{\tau_{[1]}}{2} \Delta\omega\right] \tilde{\mathbf{S}}_+(0, \Delta\omega) |e_+^{[1]}\rangle + \exp\left[i \frac{\tau_{[1]}}{2} \Delta\omega\right] \tilde{\mathbf{S}}_-(0, \Delta\omega) |e_-^{[1]}\rangle \quad (2.70)$$

in which

$$\tau_{[1]} = \Delta\beta_1^{[1]} L_{[1]} \quad (2.71)$$

has units of time, and the PMD eigenmodes for the single segment are the same as those for birefringence. The individual magnitude of residual birefringence for each frequency component of a signal launched into the fiber section also leads to frequency-dependent output SOPs – the signal depolarizes. The amount of linear depolarization can be used as a feedback signal for PMD compensators [Kikuchi 2001]. We can transform (2.70) back into the time-domain and obtain

$$|\mathbf{S}(L_{[1]}, t)\rangle = S_+ \left(0, t - \frac{\tau_{[1]}}{2}\right) |e_+^{[1]}\rangle + S_- \left(0, t + \frac{\tau_{[1]}}{2}\right) |e_-^{[1]}\rangle \quad (2.72)$$

We see that in the time-domain both eigenmodes propagate with different group velocities, acquiring a

time shift $\tau_{[1]}$ between both parts. The quantity τ is called the *differential group delay* (DGD) and the group velocity difference is parameterized by

$$\Delta\beta_1 = \beta_1(|e_+\rangle) - \beta_1(|e_-\rangle) \quad (2.73)$$

where $v_g = \beta_1^{-1}$. The parameter $\Delta\beta_1$ was initially defined in (2.65). The mode splitting (2.72) can lead to severe pulse distortions in square-law detectors when the power-splitting ratio¹²

$$\gamma = \frac{S_+^* S_+}{S_-^* S_-} \quad (2.74)$$

is close to unity and the accumulated DGD between the modes is large. This is shown schematically for $\gamma = 1$ and $\tau_{[1]} = T_S/2$ in Fig. 2.9 (top). The figure also illustrates the time-varying SOP at the output of the fiber segment which occurs as the time-domain equivalent to the frequency-domain depolarization, mentioned above.

STOKES SPACE Since the birefringence (2.23) could be well visualized in Stokes space, we can also illustrate the frequency-domain effect of PMD on the Poincaré sphere. The SOPs of the various frequency components of a signal evolve by rotating around the slow birefringence eigenmode, given by the Stokes vector $\hat{\Omega}_{[1]}$, but the rate of rotation now is dependent on frequency. The evolution equation for the Stokes vector of a specific frequency component $\Delta\omega$ of the signal, $\tilde{\mathbf{S}}(z, \Delta\omega)$, is obtained by transforming (2.66) into Stokes space, analog to (2.41),

$$\partial_z \tilde{\mathbf{S}}(z, \Delta\omega) = \Delta\beta_1 \Delta\omega \hat{\Omega}_{[1]} \times \tilde{\mathbf{S}}(z, \Delta\omega) \quad (2.75)$$

showing a linear dependence of the rate of rotation on $\Delta\omega$. We can integrate (2.75) over the short segment, analog to (2.44), and have

$$\tilde{\mathbf{S}}(L_{[1]}, \Delta\omega) = \bar{\mathbf{R}}_{[1]}(\Delta\omega) \tilde{\mathbf{S}}(0, \Delta\omega) \quad (2.76)$$

in which the rotation matrix $\bar{\mathbf{R}}$ is defined as (cf. (2.45) for the corresponding definition in the context of birefringence)

$$\bar{\mathbf{R}}_{[1]}(\Delta\omega) = \exp[\Delta\beta_1^{[1]} L_{[1]} \Delta\omega \hat{\Omega}_{[1]} \times] = \exp[\Delta\omega \tau_{[1]} \hat{\Omega}_{[1]} \times] \quad (2.77)$$

with the DGD $\tau_{[1]}$ defined in (2.71) and $\hat{\Omega}_{[1]} \times$ being the cross-product operator as in (2.46).

Because different frequency components will have different SOPs at any $z \neq 0$, we should like to

¹²The symbol γ is normally also used for the fiber nonlinearity coefficient. Both have been peacefully coexisting for a long time, since most research on PMD has mostly been carried out in the linear regime. We will not refer to the power-splitting ratio outside the present chapter and since there is thus no opportunity to confuse the two, we will not introduce a different notation.

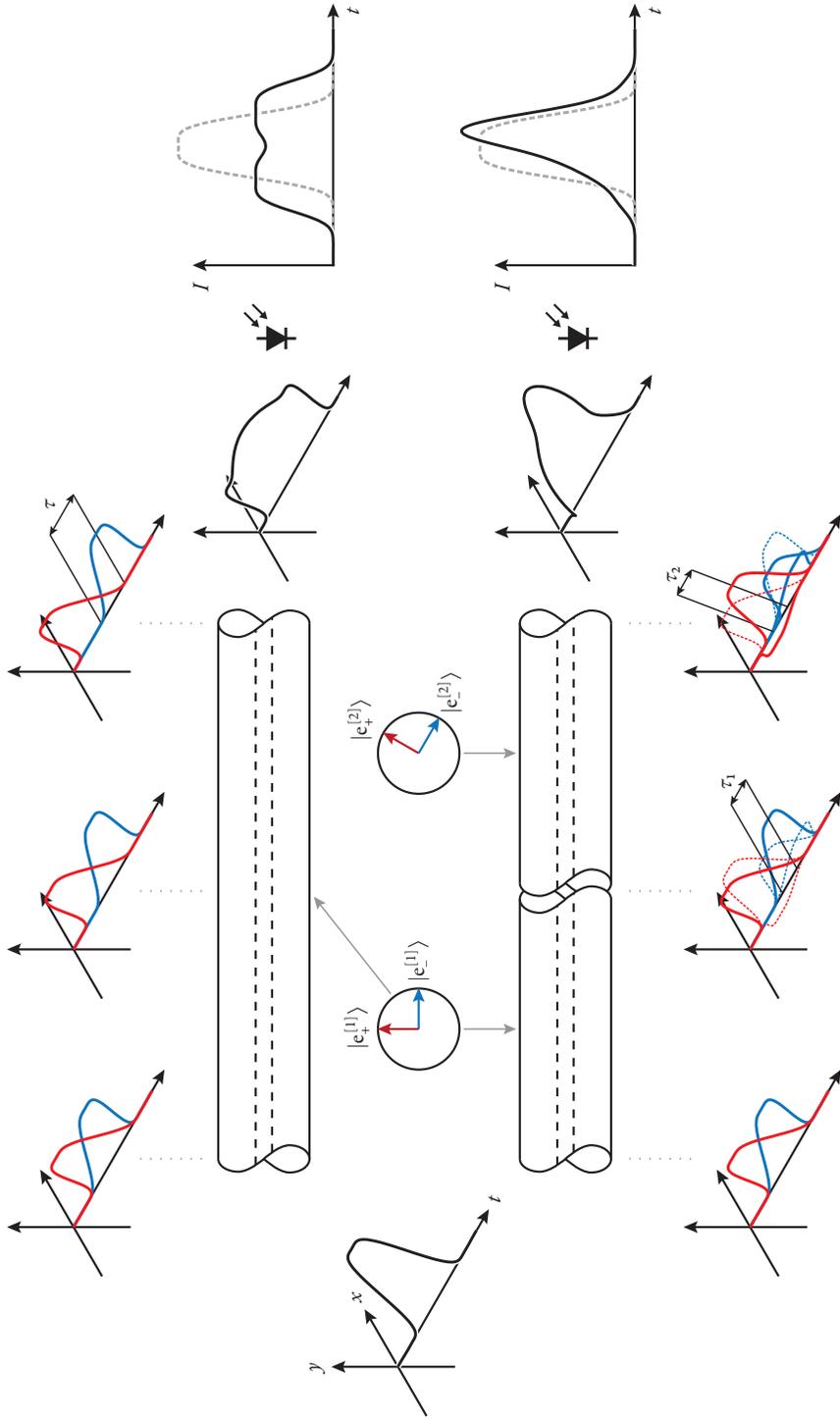


Figure 2.9 illustrates the effect of polarization mode dispersion in a single segment (top) and two concatenated segments (bottom). Within each segment the parameters $\Delta\beta_1$ and $\bar{\mathbf{H}}$ are constant. An input pulse is split into the local eigenmodes which propagate with different group velocities, leading to a time shift at the output. Upon square-law detection, the pulse becomes distorted.

know the exact frequency-dependence of those SOPs, as it is a measure for the rate of depolarization that a signal incurs. In order to determine $\partial_{\Delta\omega}\tilde{\mathbf{S}}$, we could either start in Jones space or in Stokes space. As it is easier to transform the results from Jones to Stokes space than vice versa, we shall follow the derivation in [Gordon 2000, Sec. 5] and start from Jones space. The frequency derivative of the output Jones vector is

$$\partial_{\Delta\omega}|\tilde{\mathbf{S}}\rangle = (\partial_{\Delta\omega}\bar{\mathbf{T}}_{[1]})\bar{\mathbf{T}}_{[1]}^\dagger|\tilde{\mathbf{S}}\rangle \quad (2.78)$$

where $\bar{\mathbf{T}}_{[1]}$ was defined in (2.69). In (2.78), $\bar{\mathbf{T}}_{[1]}^\dagger|\tilde{\mathbf{S}}\rangle$ describes the input Jones vector, which is independent of $\Delta\omega$, via inversion of the channel Jones matrix. The frequency dependence of the output Jones vector is then given by the frequency dependence of the channel Jones matrix. For the short segment of constant PMD we have

$$(\partial_{\Delta\omega}\bar{\mathbf{T}}_{[1]})\bar{\mathbf{T}}_{[1]}^\dagger = -i\frac{\Delta\beta_1^{[1]}}{2}L_{[1]}\bar{\mathbf{\Omega}}_{[1]} \quad (2.79)$$

where we used the unitary property $\bar{\mathbf{T}}^\dagger\bar{\mathbf{T}} = \bar{\mathbf{I}}$ and the derivative [Meyer 2000, Sec. 7.4]

$$\partial_{\Delta\omega}\exp(\Delta\omega\bar{\mathbf{\Omega}}) = \bar{\mathbf{\Omega}}\exp(\Delta\omega\bar{\mathbf{\Omega}}) \quad (2.80)$$

for any diagonalizable matrix $\bar{\mathbf{\Omega}}$. We can now transform (2.78) into Stokes space analog to (2.75), using $\bar{\mathbf{\Omega}} = \hat{\mathbf{\Omega}} \cdot \vec{\sigma}$,

$$\partial_{\Delta\omega}\tilde{\mathbf{S}} = \mathbf{\Omega}_{[1]} \times \tilde{\mathbf{S}} \quad (2.81a)$$

with the Stokes vector

$$\mathbf{\Omega}_{[1]} = \Delta\beta_1^{[1]}L_{[1]}\hat{\mathbf{\Omega}}_{[1]} = \tau_{[1]}\hat{\mathbf{\Omega}}_{[1]} \quad (2.81b)$$

For the simple case of the single segment, we would have obtained the same result by directly differentiating (2.76).

With increasing frequency, the SOP rotates around a Stokes vector $\mathbf{\Omega}_{[1]}$ whose orientation is in the direction of the slower polarization eigenmode. The rate of rotation – related to the length of the vector – is parameterized by the accumulated DGD in the segment. This is illustrated in Fig. 2.10. The vector $\mathbf{\Omega}_{[1]}$ is called the *PMD vector* of the segment. It has a special meaning in the theory of PMD, as we will show later in this section. Any output SOP that is parallel (or anti-parallel) to the PMD vector – and thus the polarization eigenmode $\hat{\mathbf{\Omega}}_{[1]}$ (2.28) of the segment – does not change to first order in frequency according to (2.81a) and thus does not depolarize at all. This corresponds to launching into either $|e_+\rangle$ or $|e_-\rangle$ in the time domain description (2.72) so that a propagating pulse is not split, but only experiences a time delay of $\pm\tau_{[1]}/2$, depending on which eigenmode it was launched into.

MULTIPLE SEGMENTS As we did with birefringence, we will now extend the analysis from the single fiber segment to two concatenated segments, and then generalize to an arbitrary number of segments. And just as with birefringence, the concatenation of two fiber segments can be written in terms of the

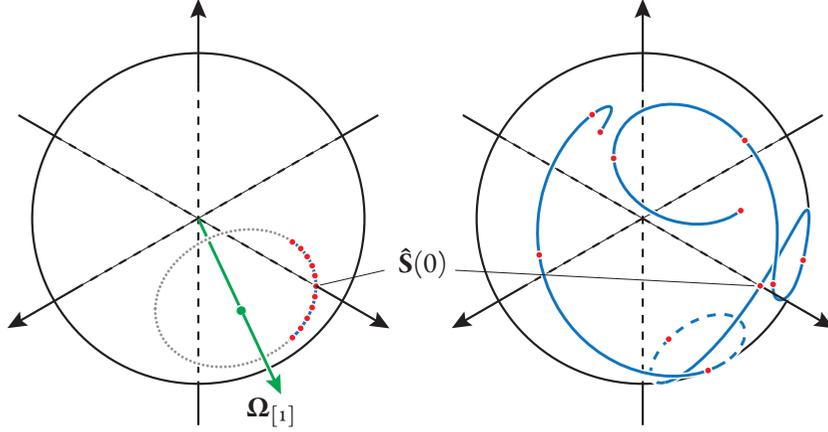


Figure 2.10 illustrates the frequency dependence of the output SOP of a single fiber segment with constant PMD parameters (left) and the concatenation of 100 segments with the same DGD as the single section, but random eigenmodes (right). In both cases, a frequency range of $2\pi \cdot 1/4\tau_{[1]}$ centered on $\Delta\omega = 0$ is plotted, with markers at every $2\pi \cdot 1/20\tau_{[1]}$. For the single segment, the PMD vector as rotation axis is also shown.

product of their unitary Jones transmission matrices,

$$\bar{\mathbf{T}}(L_{[1+2]}, \Delta\omega) = \bar{\mathbf{T}}_{[1+2]}(\Delta\omega) = \bar{\mathbf{T}}_{[2]}(\Delta\omega) \bar{\mathbf{T}}_{[1]}(\Delta\omega) \quad (2.82)$$

which can be solved segment-wise. In the time domain, the counterpart of (2.72) is

$$\begin{aligned} |S(L_{[1+2]}, t)\rangle &= C(+, +) \times S_+ \left(0, t - \frac{\tau_{[1]}}{2} - \frac{\tau_{[2]}}{2}\right) |e_+^{[2]}\rangle \\ &+ C(+, -) \times S_- \left(0, t + \frac{\tau_{[1]}}{2} - \frac{\tau_{[2]}}{2}\right) |e_+^{[2]}\rangle \\ &+ C(-, +) \times S_+ \left(0, t - \frac{\tau_{[1]}}{2} + \frac{\tau_{[2]}}{2}\right) |e_-^{[2]}\rangle \\ &+ C(-, -) \times S_- \left(0, t + \frac{\tau_{[1]}}{2} + \frac{\tau_{[2]}}{2}\right) |e_-^{[2]}\rangle \end{aligned} \quad (2.83)$$

in which

$$|S(0, t)\rangle = S_+(0, t) |e_+^{[1]}\rangle + S_-(0, t) |e_-^{[1]}\rangle \quad (2.84)$$

defines S_+ and S_- as the expansion of the input field into the eigenmodes of the first segment. The output (2.83) depends on the various projection coefficients (describing the relative orientation of the eigenmodes)

$$C(m, n) = \langle e_m^{[2]} | e_n^{[1]} \rangle \quad \text{with } m, n \in \{+, -\} \quad (2.85)$$

and the segment DGDs $\tau_{[1]}$ and $\tau_{[2]}$. The projection coefficients (2.85) result from the output field of segment 1 being projected onto the fiber eigenmodes of segment 2 by means of the projection operators

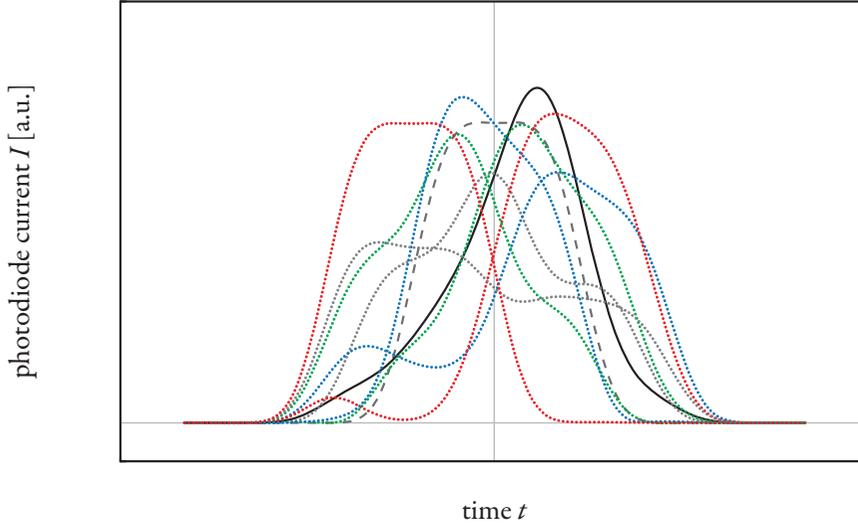


Figure 2.11 shows several exemplary output pulses of a two-segment fiber, given a fixed input pulse (dashed line). The eigenmodes \mathbf{H} of each segment were randomized, leading to random coupling coefficients $C(m, n)$. The segment DGDs were also chosen randomly, keeping the sum of the DGDs of both segments $\tau_{[1]} + \tau_{[2]}$ constant. The solid black line shows the output pulse of the fiber used in Fig. 2.9 (bottom).

$|e_m^{[2]}\rangle\langle e_m^{[2]}|$. Fig. 2.9 (bottom) illustrates the above process and Fig. 2.11 shows some examples for the output field $|S\rangle$ when the orientation of the (linear) birefringence eigenmodes of the segments and their DGD varies, the total DGD $\tau_{[1]} + \tau_{[2]}$ of both segments remaining constant. It is straightforward to extrapolate the equations for any number of concatenated segments. The number of terms will double with each additional segment, leading to a great variety in the possible output fields. For this reason, PMD for realistic fiber links consisting of a large number of such segments with unknown length and eigenmode orientation can generally only be described in a statistical sense.

The simple form of (2.85) demonstrates an advantage of using fiber coordinates over using laboratory coordinates: The SOP of the eigenmodes is the same at any location within any one segment of the fiber when there is no birefringence. In laboratory coordinates the birefringence of one segment would transform the eigenmodes of the other segment (which are assumed to be different) with local transmission matrix \bar{U} (cf. the previous section). Therefore, depending on at which location the projection coefficients are calculated – and this may be done at any one point along the fiber – the eigenmodes of one or more segment must be properly transformed. This can become tedious when there are many segments. However, the coefficients $C(m, n)$ are unaffected by such a transformation.

To obtain the behavior in Stokes space, we integrate the Stokes evolution equation (2.75) segment-wise. We can write analog to (2.49)

$$\bar{\mathbf{R}}(L_{[1+2]}, \Delta\omega) = \bar{\mathbf{R}}_{[1+2]}(\Delta\omega) = \bar{\mathbf{R}}_{[2]}(\Delta\omega) \bar{\mathbf{R}}_{[1]}(\Delta\omega) \quad (2.86)$$

except that now the rotation matrices $\bar{\mathbf{R}}$ depend on frequency $\Delta\omega$. An input SOP rotates first around $\hat{\mathbf{\Omega}}_{[1]}$ describing the first segment and the output vector of that segment then rotates around $\hat{\mathbf{\Omega}}_{[2]}$ describing the second segment.

As before, we can also determine the frequency-derivative of the Stokes vector using (2.78), (2.80), (2.82) and the matrix product rule

$$\partial_{\Delta\omega}(\bar{\mathbf{X}}\bar{\mathbf{Y}}) = \partial_{\Delta\omega}\bar{\mathbf{X}} \cdot \bar{\mathbf{Y}} + \bar{\mathbf{X}} \cdot \partial_{\Delta\omega}\bar{\mathbf{Y}} \quad (2.87)$$

and obtain for the two-segment concatenation

$$(\partial_{\Delta\omega}\bar{\mathbf{T}}_{[1+2]})\bar{\mathbf{T}}_{[1+2]}^\dagger = -i\left(\frac{\Delta\beta_1^{[2]}}{2}L_{[2]}\bar{\mathbf{\Omega}}_{[2]} + \frac{\Delta\beta_1^{[1]}}{2}L_{[1]}\bar{\mathbf{T}}_{[2]}\bar{\mathbf{\Omega}}_{[1]}\bar{\mathbf{T}}_{[2]}^\dagger\right) \quad (2.88)$$

The matrix $\bar{\mathbf{T}}_{[2]}$ in the second term of the right-hand side acts as a similarity transformation related to the residual birefringence of the second segment. Even though we have not written it specifically in (2.88), the matrix $\bar{\mathbf{T}}_{[2]}$ depends on frequency as defined in (2.67). If we limit the discussion to frequencies near the carrier, $\bar{\mathbf{T}}_{[2]}(L_{[2]}, \Delta\omega \approx 0) \approx \bar{\mathbf{I}}$, so that $\bar{\mathbf{T}}_{[2]}$ disappears from (2.88). We can now transform the frequency-derivative into Stokes space as before and have

$$\partial_{\Delta\omega}\tilde{\mathbf{S}}(L_{[1+2]}, \Delta\omega) = \mathbf{\Omega}_{[1+2]} \times \tilde{\mathbf{S}}(L_{[1+2]}, \Delta\omega) \quad (2.89a)$$

with the Stokes vector

$$\mathbf{\Omega}_{[1+2]} = \mathbf{\Omega}_{[1]} + \mathbf{\Omega}_{[2]} = \tau_{[1]}\hat{\mathbf{\Omega}}_{[1]} + \tau_{[2]}\hat{\mathbf{\Omega}}_{[2]} \quad (2.89b)$$

We could also obtain (2.89b) directly in Stokes space by using (2.86) with (2.77) and determining the frequency derivative $\partial_{\Delta\omega}\bar{\mathbf{R}}_{[1+2]}$ at $\Delta\omega = 0$. To obtain the final expression, we make use of the product rule (2.87), $\partial_{\Delta\omega}\tilde{\mathbf{S}}(0, \Delta\omega) = 0$ (constant input SOP), and $\tilde{\mathbf{S}}(L_{[1+2]}, 0) = \tilde{\mathbf{S}}(0, 0)$ (birefringence has been eliminated at $\Delta\omega = 0$), to arrive at (2.89).

A consequence of (2.89) is that after traversing two segments with independent PMD vectors we can still find two anti-parallel polarization states for which $\partial_{\Delta\omega}\tilde{\mathbf{S}} = 0$ – i.e. there is no change in output SOP for small changes in frequency and a narrow-band signal does not depolarize. These states are not the eigenmodes of any of the two segments, and due to $\bar{\mathbf{T}}_{[2]}$ in (2.88) they also become frequency-dependent. In (2.89) they are defined at the output $L_{[1+2]}$ of the fiber, but they can be transformed to the input via (2.75). Unlike the eigenmodes of the single-segment fiber, these SOPs are generally different at the input and output. They are called the *principal states of polarization* (PSP) and their frequency-dependence (or rather that of $\mathbf{\Omega}_{[1+2]}$) is referred to as higher-order PMD. The PSPs form an orthogonal basis set in Jones space which serves well to describe narrow-band PMD phenomena [Poole 1986]. The vector $\mathbf{\Omega}_{[1+2]}$ in (2.89) is the PMD vector of the concatenation of both sections. It is the vector sum of

the two segment PMD vectors in Stokes space. Its length is the composite DGD of both sections,

$$\tau_{[1+2]} = |\mathbf{\Omega}_{[1+2]}| = \sqrt{|\mathbf{\Omega}_{[1]}|^2 + |\mathbf{\Omega}_{[2]}|^2 + 2\mathbf{\Omega}_{[1]} \cdot \mathbf{\Omega}_{[2]}} \quad (2.90)$$

and its direction defines the slow PSP. If an input signal is aligned with an input PSP of the fiber, the signal exits the fiber undistorted (to first order in frequency) and fully polarized with the corresponding output PSP and with a time delay of $\pm\tau_{[1+2]}/2$. The concatenation of both sections thus behaves – to first order in frequency – like a single fiber segment, regarding PMD.

The delay $\tau_{[1+2]}$ depends on the relative alignment in \mathbb{R}^3 of $\mathbf{\Omega}_{[1]}$ and $\mathbf{\Omega}_{[2]}$, and can become zero if they are equal but opposite. This is the principle behind PMD compensation by nulling the PMD vector, where a high-PMD fiber section is aligned such that the total DGD of the concatenation of the transmission fiber and the compensation fiber is zero [Kogelnik 2002].¹³

It is now straightforward to generalize (2.89b) to any number of segments, and in the infinitesimal limit we have

$$\mathbf{\Omega}(z) = \int_0^z \partial_\zeta \mathbf{\Omega}(\zeta) d\zeta = \int_0^z \Delta\beta_1(\zeta) \hat{\mathbf{\Omega}}(\zeta) d\zeta \quad (2.91)$$

or the differential form

$$\partial_z \mathbf{\Omega} = \Delta\beta_1 \hat{\mathbf{\Omega}}(z) \quad (2.92)$$

In much of the literature on PMD, there appears a second term in the right-hand side of (2.92) which describes a rotation of $\mathbf{\Omega}$ around the local birefringence vector \mathbf{H} , as in the Stokes description of birefringence (2.41). Because we have eliminated the birefringence at $\Delta\omega = 0$ by using the local fiber coordinates, this term does not appear here.

Because $\hat{\mathbf{\Omega}}$ will be uniformly distributed on the Poincaré sphere as a result of the similarity transform $\bar{\mathbf{U}}$ – this property of $\bar{\mathbf{U}}$ was a prerequisite to perform the nonlinearity averaging which led to the Manakov equation – (2.92) describes a random walk in \mathbb{R}^3 , to which all the theory of such a stochastic process applies, e.g.

- the DGD $\tau = |\mathbf{\Omega}|$ increases with the square root of the number of segments on average (for equal segment DGD $\tau_{[n]} = \tau_{\text{seg}}$)
- the variance $\text{var}[\mathbf{\Omega}] = \langle \tau^2 \rangle = \sum_n (\tau_{[n]})^2 = N_{\text{seg}} \tau_{\text{seg}}^2$, where N_{seg} is the number of segments
- in the limit of infinitesimal sections, $\mathbf{\Omega}(z)$ can be described by a Wiener process, or Brownian motion, in three dimensions [Papoulis 1991, Ch. 11]
- the distribution of $\mathbf{\Omega}(z)$ is 3-dimensional Gaussian

¹³Another method to compensate PMD is to align $\mathbf{\Omega}_{[2]}$ such that one of the PSPs described by the sum $\mathbf{\Omega}_{[1+2]}$ equals the SOP of the data signal. This allows for more flexibility of the compensating segment, because it relaxes the requirement on the DGD $\tau_{[2]}$ [Kogelnik 2002].

- the distribution of $\tau(z) = |\mathbf{\Omega}(z)|$ is Maxwellian with

$$p(\tau) = \frac{32}{\pi^2 \langle \tau \rangle^3} \tau^2 \exp\left(\frac{-(2\tau)^2}{\pi \langle \tau \rangle^2}\right) \quad (2.93)$$

where $\langle \tau \rangle$ is the ensemble mean value (see below)

- in analogy to Brownian motion or diffusion, we can define a diffusion constant D_{PMD} , so that

$$\langle \tau(z) \rangle^2 = \frac{8}{3\pi} \langle \tau^2(z) \rangle = D_{\text{PMD}}^2 z \quad (2.94)$$

POLARIZATION ENSEMBLE The appearance of *distributions*, averages and variances in the above list suggests that we have to take a look at more than one possible realization of the sequence of segment PMD vectors $\mathbf{\Omega}_{[z]}$. In fact, when describing optical transmission over a fiber that is affected by PMD, we generally have to look at the *stochastic ensemble* of all fibers with equal properties in a statistical sense – in fiber optics, the constant D_{PMD} is generally used to classify the ensemble, because it is defined in terms of the ensemble: the notation $\langle \cdot \rangle$ in (2.94) refers to the average over the ensemble of all possible manifestations of the random walk of the PMD vector. The term ensemble seems to imply that these statistics refer to different physical realizations, different strands, of fiber. But as we mentioned before, both the eigenmode structure and the difference in group delay per unit length may also change with time, albeit very slowly compared to the signal. The fastest changes of fiber PMD measured in the laboratory are on the order of milliseconds – still corresponding to 10^4 bits in a 10 Gbps system [Krummrich 2004]. In buried fiber, they are on the order of minutes or hours [Krummrich 2005, Brodsky 2005, Antonelli 2006]. Thus, as the PMD properties of a single strand of fiber change over time, ensemble statistics begin to apply. We will thus associate these slow changes in the fiber with the statistical ensemble. It is important not to confuse these with the fast changes on the order of the symbol length that are, e.g., averaged over to determine the Stokes parameters in Chapter 1.

Because we know the distribution of τ in the ensemble to be Maxwellian, we can determine the various probabilities that τ will exceed a certain value, e.g. $\mathcal{P}[\tau > 3\langle \tau \rangle] \approx 4 \times 10^{-5}$ simply by integrating the probability density function (PDF). If a system cannot operate stably when an allowable value of τ is exceeded – resulting in an *outage event* – we can define a maximum allowable average $\langle \tau \rangle$ – and related parameter D_{PMD} – so that the probability of an outage is arbitrarily small. Usually, this probability is chosen to be 4×10^{-5} , corresponding to approximately 22 minutes of outage per year, on average.

There exist a number of publications that examine the statistics of the various orders of PMD in great detail using different models. Apart from the books mentioned at the beginning of the section, there are a number of seminal papers on the topic, of which only a small selection can be given here [Poole 1988, Foschini 1991, Menyuk 1994, Wai 1996, Karlsson 2000, Winzer 2004].

We shall, however, examine one particular stochastic aspect of PMD which concerns the relative evolution of the SOPs of the carriers in a WDM system later in Section 3.2.3, after we have introduced the necessary framework to describe random motion on the Poincaré sphere in Chapter 3.

MODELING PMD We have briefly mentioned before that a realistic fiber can be modeled, regarding birefringence and PMD, as the concatenation of segments within which the eigenmode structure and the various strength coefficients may be assumed constant. The eigenmode vectors $\hat{\Omega}$ determine the directions and the group delay $\tau = |\Omega|$ the lengths of the segment PMD vectors that constitute the random walk whose endpoint is the fiber PMD vector, according to (2.89). We also mentioned that in the limit of infinitesimally short sections, the random walk approaches a Brownian motion, or Wiener process, in three dimensions [Papoulis 1991, Ch. 11]. However, there are subtle differences that should not be ignored. First of all, the Maxwellian distribution applies only in the limit of an infinite number of segments. For fewer segments, the distribution differs, but noticeably only when the number of segments becomes small (less than approximately 50). To illustrate the influence of the segment granularity on the statistics of τ , Fig. 2.12 shows the distributions of τ for a varying number of segments, all having an equal segment DGD $\tau_{[n]}$ but random segment eigenmodes $\hat{\Omega}_{[n]}$, normalized to the root-mean-square τ_{RMS} .¹⁴ The difference lies mainly in the distribution tails. As long as the number of segments is finite, there is also a maximum attainable total DGD – the sum of all segment DGDs. It occurs only when the PMD vectors of all segments are aligned, as e.g. in a polarization-maintaining fiber. This maximum event influences the shape of the distributions, as the probability of all DGDs larger than this maximum must be zero.

When we apply the results of Fig. 2.12 to numerical modeling of PMD, we see that, provided we divide the fiber link into more than about 50 segments, we approach the Maxwellian limit statistics sufficiently well, with very small statistical benefits for larger numbers of segments.¹⁵ Thus, compared to the relatively small step size required for numerical solving of the NLSE using the split-step Fourier method [Agrawal 2007, Sec. 2.4], modeling PMD allows for division of the fiber into relatively few segments inside which the PMD-relevant quantities are assumed constant, wherefore this model has been dubbed *coarse-step model*. Within each segment, the split-step method is again used to calculate non-linear propagation.

More recently, time-resolved measurements of PMD in the field have shown that the statistics of single fiber links deviate from the Maxwellian ensemble statistics. This has led to a refinement of the coarse-step model called the *hinge model* of fiber PMD: for most fiber links, especially those with mainly buried fiber, the aggregate PMD vector over long distances is constant in time and only prone to change where the fiber comes out of the ground and might be exposed to environmental influences, as in switching and routing stations – the so-called hinges. We thus again have only a small number of independent segments whose length and orientation themselves are the result of random walks described by the segments between the hinges. The hinge model and its particular statistics are discussed in [Kogelnik 2005, Brodsky 2008]. To model PMD in the present work, we restricted ourselves to the

¹⁴These distributions were derived from basic theory on random walks (sometimes called *random flights* in three dimensions), as e.g. presented in [Hughes 1995, Sec. 2.5], assuming steps of equal length.

¹⁵Field measurements have also shown that the average correlation length in standard single-mode fiber (SSMF) is on the order of some hundreds of meters, resulting in the equivalent of some hundred statistically independent segments per fiber span [Menyuk 1999].

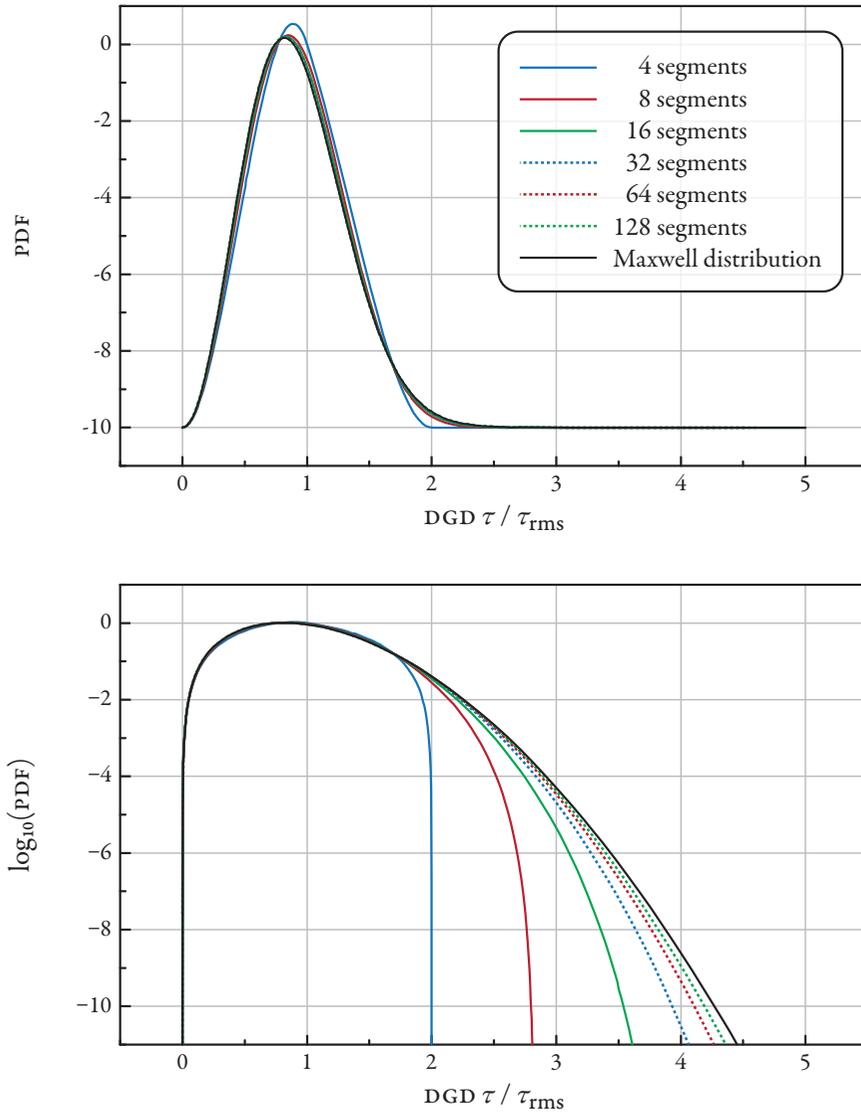


Figure 2.12 shows the probability density functions of the total DGD τ , normalized to its root-mean-square value for a multiple-segment model with varying number of segments, and compares it to the Maxwell limit distribution for infinite segments. Top graph shows the PDFs in a linear scale, bottom graph is logarithmic.

simpler coarse-step model without the time-dependence implied by the hinge model. We will therefore refrain from a thorough discussion of the latter.

Assuming all sections to have equal DGD, $\tau_{[n]} = \tau_{\text{seg}}$, this DGD can be determined from the root-mean-square value at the end of the transmission line by

$$\tau_{\text{seg}} = \sqrt{\frac{\langle \tau^2(L) \rangle}{N_{\text{seg}}}} = \frac{\tau_{\text{RMS}}(L)}{\sqrt{N_{\text{seg}}}} \quad (2.95)$$

where N_{seg} is again the number of segments the fiber is divided into. Further assuming segments of equal length L_{seg} , we have with the definition of the DGD (2.71)

$$\Delta\beta_1^{[n]} = \frac{\tau_{\text{seg}}}{L_{\text{seg}}} = \frac{\tau_{\text{RMS}}(L)}{L} \sqrt{N_{\text{seg}}} = D_{\text{PMD}} \sqrt{\frac{N_{\text{seg}}}{L}} \quad (2.96)$$

that is, if we want the root-mean-square DGD – and thus the general shape of its distribution – and the total fiber length to be constant, the local fiber PMD $\Delta\beta_1$ must increase with the square root of the number of segments we use in the coarse-step model [Marcuse 1997]. It is the same in all segments, if these have equal length and DGD.

However, it has been shown that it is important to avoid equal DGD for all segments [Kogelnik 2002, Sec. 5.5]. The reason for this can be best explained with the Stokes expression of the frequency-dependence of the output SOP (2.81): The output SOP of a segment rotates on the Poincaré sphere with a change in frequency. The output SOP of some frequency components will be identical, corresponding to rotations of one or more full circles. We call the frequency difference between these components the free spectral range (FSR), in analogy with linear filters, and with (2.81) we have

$$\Delta\omega_{\text{FSR}} = \frac{2\pi}{\tau_{\text{seg}}} \quad (2.97)$$

Now if all segments have the same $\tau_{[n]} = \tau_{\text{seg}}$, the output SOPs of any frequencies spaced apart by the FSR will be identical at the output of each segment, and thus also at the output of the fiber. The resulting periodic output SOPs are unphysical. The authors in [Kogelnik 2002] have shown that a Gaussian variation with a distribution width of 20 percent of the segment DGDs can eliminate this unwanted periodicity in PMD simulation and emulation. In the limit of a large number of segments, this does not affect the statistics of $\tau(L)$. When the number of segments is small, however, the distribution of $\tau(L)$ can be obtained with methods described in [Hughes 1995, Sec. 2.5.4].

POLARIZATION-DIVISION MULTIPLEX In the presence of PMD, as we have seen, the output SOP varies with frequency. Therefore, it will be impossible to perfectly align the polarization beam splitter to separate both POLDM subchannels over the whole frequency band. The fading and crosstalk coefficients

F and X in (2.54) become frequency-dependent,

$$F(\Delta\omega) = \langle \mathbf{e}'_1 | \tilde{\mathbf{e}}_1(\Delta\omega) \rangle \quad \text{and} \quad X(\Delta\omega) = \langle \mathbf{e}'_1 | \tilde{\mathbf{e}}_2(\Delta\omega) \rangle \quad (2.98)$$

Once again, this can be better visualized in Stokes space, using the Stokes relation (2.57) for the magnitudes of F and X. When the signal spectrum is narrow and/or the segment DGDs are very small – so that the frequency dependence of the fiber PMD vector can be neglected (the *first-order approximation*) – we can write for the now frequency-dependent misalignment angle (cf. Section 2.2.2 on birefringence)

$$\begin{aligned} \cos \Theta(\Delta\omega) &= \hat{\mathbf{S}}'_1 \cdot \hat{\mathbf{S}}_1(L, \Delta\omega) \\ &= \hat{\mathbf{S}}'_1 \cdot \bar{\mathbf{R}}(L, \Delta\omega) \hat{\mathbf{S}}(0, \Delta\omega) \end{aligned}$$

where the Mueller matrix

$$\bar{\mathbf{R}}(L, \Delta\omega) = \exp[\Delta\omega \mathbf{\Omega}(L) \times]$$

is the generalization of the single-segment rotation matrix (2.77) for the fiber PMD vector $\mathbf{\Omega}(L)$ defined in (2.91). The exponent $\mathbf{\Omega}(L) \times$ is once more the cross-product operator as defined in (2.46).

The deterioration thus not only depends on the spectral width of the channels and the DGD of the link, but also on the relative alignment of the SOPs and the PSPs – launching into the PSPs minimizes both fading and crosstalk, as would be expected.

Assuming the PBS is aligned to the SOP of the channel carrier frequency, we expect the crosstalk (as the more deleterious impairment) to be largest whenever there are transitions between neighboring symbols. For NRZ-modulated subchannels that means that the symbols in both subchannels should be temporally aligned, in order to keep the crosstalk from the neighboring subchannel minimal at the sampling instances [Nelson 2000]. However, as the pulse distortions from PMD become larger, crosstalk will impair detection anywhere in the symbol interval. For RZ-modulated data we can argue along similar lines as in the section on birefringence. Because the crosstalk scales with the instantaneous power in the neighboring subchannel, interleaving both subchannels reduces crosstalk not by minimizing the misalignment Θ , but the crosstalk-source (cf. also [Borne 2005]). However, as PMD increases, symbols spread into the gaps between them where they can cause significant deterioration. The comparatively wider spectrum of RZ versus NRZ further aggravates the depolarization of the subchannels. Above certain levels of PMD, compensation will be required in order to be able to successfully operate a POLDM system. Fortunately, the electronic processing capacities required in POLDM receivers are able to simultaneously compensate PMD.

At this point we have all the basic knowledge about PMD that will be required within the next chapters, and more, so that we may conclude these algebra-dominated remarks about PMD and with it the linear transmission effects, and turn towards the class of fiber nonlinearities summarized under

the label *Kerr nonlinearities*.

2.3 NONLINEAR EFFECTS

We devote the remainder of this chapter to the final term in the propagation equation (2.6), respective (2.52) in local fiber coordinates (cf. discussion on p. 2.2.2) in which the birefringence has been removed mathematically, which describes the fiber nonlinearities resulting from the optical Kerr effect.

The (electro-optical) Kerr effect is a change in the refractive index of a material as a consequence of an applied electrical field. Unlike the Pockels effect, the refractive index change is proportional to the square magnitude of the field. It was discovered as an electrically induced birefringence in 1875 by Rev. John Kerr [Kerr 1875]. For the optical Kerr effect, the light itself is the source of the electric field, and the refractive index change is proportional to the power density or irradiance of the optical field. It was observed in a glass waveguide in 1973 [Stolen 1973], although Kerr also used glass for his experiments.

To examine the consequences of the Kerr effect in optical fiber communication, we neglect the PMD-related term in (2.52) and obtain¹⁶

$$\partial_z |S\rangle - i \frac{\beta_2}{2} \partial_t^2 |S\rangle - i \exp(-\alpha z) \gamma \left[\langle S|S\rangle |S\rangle - \frac{1}{3} (\langle S|\bar{Y}|S\rangle) \bar{Y}|S\rangle \right] = 0 \quad (2.99)$$

where the z -dependent matrix \bar{Y} was defined in (2.52c). This traceless unitary matrix is a rapidly and randomly changing similarity transform of σ_3 . In fiber-optic communication systems, nonlinear distortions accumulate only over large distances, on the order of hundreds of kilometers. Within this distance, all possible eigenstates of \bar{Y} become equally likely as a result of the birefringence-induced linear mixing. In Stokes space, the corresponding eigenvectors cover the Poincaré sphere uniformly. Therefore, it is reasonable to average the term containing \bar{Y} over all possible \bar{Y} . To achieve this, we expand the matrix \bar{Y} into the Pauli matrices, analog to (2.26), and average the term over all corresponding eigenvectors \hat{Y} on the surface of the Poincaré sphere. We thus obtain [Wai 1996, Marcuse 1997]

$$\left\langle (\langle S|\bar{Y}|S\rangle) \bar{Y}|S\rangle \right\rangle = \frac{1}{3} \langle S|S\rangle |S\rangle \quad (2.100)$$

in which $\langle \cdot \rangle$ denotes this averaging process. Inserting (2.100) in (2.99) yields

$$\partial_z |S\rangle - i \frac{\beta_2}{2} \partial_t^2 |S\rangle - i \exp(-\alpha z) \tilde{\gamma} \langle S|S\rangle |S\rangle = 0 \quad (2.101a)$$

¹⁶We include the GVD term because at the optical powers found in fiber communication systems, the Kerr effect accumulates noticeably only over large distances and during propagation GVD can significantly change the signals in a way that affects how nonlinearities evolve.

with

$$\tilde{\gamma} = \frac{8}{9}\gamma \quad (2.101b)$$

This equation is referred to as the *Manakov equation* [Manakov 1974]. When the PMD term is included, it is known as the *Manakov PMD equation* [Wai 1996]. Unlike in (2.7), the nonlinear term no longer depends on the SOP of the field. Also, the fiber nonlinearity acts only with a relative strength of 8/9 due to the birefringence-induced polarization mixing. This reduced nonlinearity has been confirmed experimentally [Chernikov 1996, Andersen 1998].

Because the refractive index is modified as a result of the total optical field propagating through the fiber, the optical Kerr effect causes (nonlinear) coupling between the different wavelength channels in a dense wavelength-division multiplex (DWDM) system. To examine these effects, we expand the optical field $|S\rangle$ into a superposition of co-propagating wavelength channels analog to (2.10) and obtain

$$\langle S|S\rangle|S\rangle = \sum_{\eta} \sum_{\mu} \sum_{\nu} \langle S_{\eta}|S_{\mu}\rangle|S_{\nu}\rangle \exp[i\Delta\omega_{\rho}t] \quad (2.102)$$

in which

$$\Delta\omega_{\rho} = \Delta\omega_{\mu} - \Delta\omega_{\eta} + \Delta\omega_{\nu} \quad (2.103)$$

We are interested in the nonlinear distortions of a single channel only. For this probe channel ρ we set $\Delta\omega_{\rho} = 0$, as we did in Section 2.1 when we first defined the probe channel on page 2.1. We now sort the terms of the triple sum appearing at $\Delta\omega_{\rho} = 0$ grouped by degeneracy,

$$\begin{aligned} \langle S|S\rangle|S\rangle &= \overbrace{\langle S_{\rho}|S_{\rho}\rangle|S_{\rho}\rangle}^{\text{SPM}} + \overbrace{\sum_{\nu \neq \rho} \langle S_{\nu}|S_{\nu}\rangle|S_{\rho}\rangle}^{\text{XPM}} \\ &\quad + \underbrace{\sum_{\nu \neq \rho} |S_{\nu}\rangle \langle S_{\nu}|S_{\rho}\rangle}_{\text{XPM/XPOLM}} + \underbrace{\sum_{\substack{\mu, \nu \neq \rho \\ \rho = \mu + \nu - \eta}} \sum_{\eta} \langle S_{\eta}|S_{\mu}\rangle|S_{\nu}\rangle}_{\text{FWM}} \end{aligned} \quad (2.104)$$

where $\eta = \mu + \nu$ is a short form of (2.103), assuming all channels are arranged in a frequency grid with fixed separation between channels, e.g. the ITU grids of 50 GHz or 100 GHz. SPM, XPM, and XPOLM are degenerate terms, involving channel ρ , of the nonlinear mixing process between up to four wavelengths that is referred to as four-wave mixing, and have been separated from the sum expression.

Next, we will briefly look at these individual nonlinearities appearing as the result of the optical Kerr effect, starting with the most general term, FWM.

2.3.1 Four-Wave Mixing

As the phase of the mixing product resulting from FWM depends on the relative phases of the contributing wavelengths, we have to consider the linear GVD in the mixing channels. To simplify the analysis for illustration purposes, we represent the modulated DWDM channels by continuous wave (CW) signals at the respective carrier frequencies. In this case, we can apply (2.16) directly to the time domain fields $|S\rangle$ and have

$$\partial_z |S_\rho\rangle - i \exp(-\alpha z) \tilde{\gamma} \sum_{\substack{\mu, \nu \neq \rho \\ \rho = \mu + \nu - \eta}} \langle S_\eta | S_\mu \rangle |S_\nu\rangle \times \exp \left[-i \frac{\beta_2 z}{2} (\Delta\omega_\mu^2 + \Delta\omega_\nu^2 - \Delta\omega_\eta^2) \right] = 0 \quad (2.105)$$

where the single sum symbol is used to represent the triple sum. The argument of the second exponential is referred to as the phase matching condition. The mixing products due to FWM can only build up significantly if this phase argument remains (quasi-)constant over long enough distances. This is true only near the zero-dispersion wavelength of optical fibers, and is the reason why dispersion-shifted fibers are detrimental to DWDM transmission and their use has mostly been discouraged for long-haul transmission.

The relative build-up of the FWM mixing product over a single span of fiber is illustrated in Fig. 2.13 for various values of β_2 and the special (three-wave mixing) case $\eta = \rho$, $\mu = \rho + 1$, $\nu = \rho - 1$ in a 50 GHz channel grid. Even for this degenerate case which improves phase matching, the FWM products at the end of the span are very small for fibers with a GVD coefficient as low as $\beta_2 = -32$ fs/GHz/km ($D_{\text{GVD}} = 4$ ps/nm/km). Furthermore, in field-deployed transmission systems, in which the dispersion map is never truly resonant, the phase of the FWM products from different spans will be uncorrelated so that they add randomly.¹⁷ As can also be deduced from (2.105), an increase in the frequency separation between the mixing channels will also decrease the efficiency of the FWM process similar to an increase of β_2 .

Unlike SPM and XPM, discussed next, FWM and its product depend on the SOPs of the involved wavelength channels η , μ , and ν . From (2.105) we see that the FWM efficiency scales with the product $\langle S_\eta | S_\mu \rangle$ of the interferer SOPs, i.e. if channels η and μ are orthogonally polarized, this particular FWM product vanishes – which is not necessarily true for every permutation of η , μ , and ν . The perturbation field then has the same SOP as $|S_\nu\rangle$. Generally, FWM can thus also affect the SOP of the probe channel ρ . However, its magnitude is small compared to the degenerate term that results in XPOLM, when data is transmitted away from the zero-dispersion frequencies. Then, to be able to perform a theoretical analysis of XPOLM, we therefore neglect the influence of the FWM products on the SOP of the probe.

There is one reason why we cannot ignore FWM completely, however. Bosco et al. have shown that using a too large and constant step-size in the split-step Fourier method used to numerically calculate

¹⁷Contrary to computer simulations, truly resonant dispersion maps are very difficult to achieve in practice, and a small residual dispersion per span usually suffices to decorrelate the FWM products.

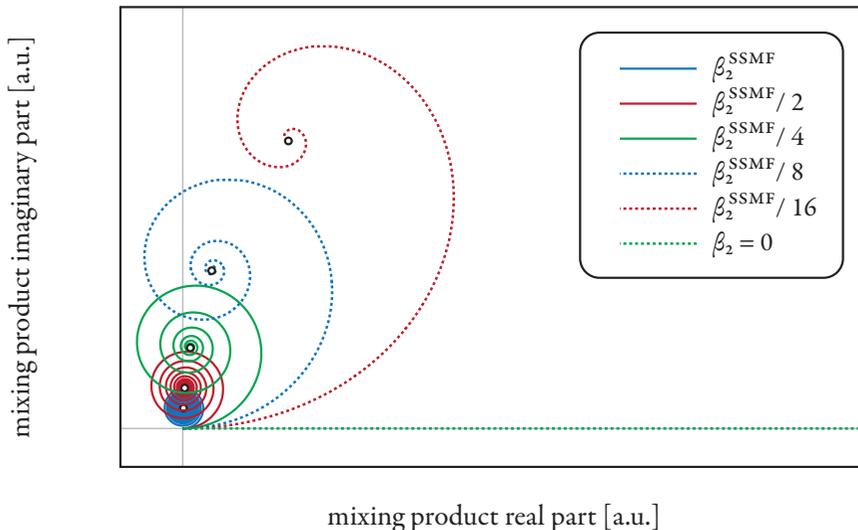


Figure 2.13 illustrates the evolution of the mixing product due to three 10 dBm CW signals at frequencies relative to the channel under test of -50 GHz, 0 , and $+50$ GHz, all having equal SOPs throughout. Fiber parameters used were $\beta_2^{\text{SSMF}} = -128$ fs/GHz/km ($D_{\text{GVD}} = 16$ ps/nm/km), $\alpha = 0.2$ dB/km, $\gamma = 1.31$ (W km) $^{-1}$, $L = 100$ km. Symbols mark the accumulated distortion after propagation of 100 km. All spirals have the same total arc length.

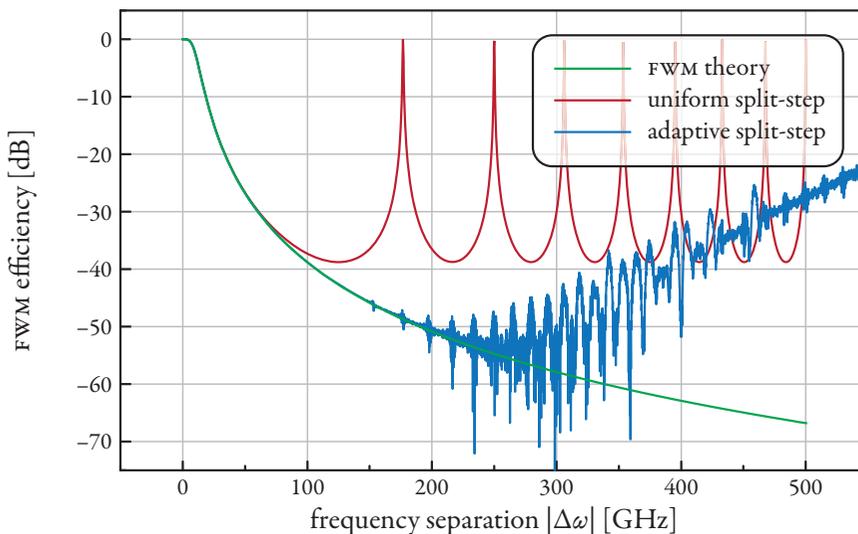


Figure 2.14 shows the FWM efficiency vs. frequency separation according to theory, for split-step simulations using uniform step sizes of 250 m, and for an adaptive step size algorithm (see text). Total propagated fiber length was 100 km; fiber parameters used were $\beta_2 = -128$ fs/GHz/km ($D_{\text{GVD}} = 16$ ps/nm/km), $\alpha = 0.2$ dB/km.

fiber propagation can lead to unphysical, «fictitious» four-wave mixing products at higher frequencies [Bosco 2000]. Figure 2.14 shows the theoretical FWM efficiency, the unphysical FWM efficiency obtained when using equal-sized steps of 250 m length, and the FWM efficiency of the split-step implementation used for the simulations herein. In this implementation the step size starts at 25 m and increases inversely with accumulated attenuation, with a maximum step size fixed at 500 m.¹⁸ This results in approximately 1300 split-steps for a 100 km long fiber span. The FWM efficiency itself is a relative measure of the magnitude of the mixing products after propagation of a certain distance [Shibata 1987].

From Fig. 2.14 we can see that the implementation of the numerical simulations yields correct results up to frequency offsets of 200 GHz and does not peak above -20 dB in the entire examined frequency range. We will thus limit our numerical investigations to systems with 10 interfering channels co-propagating with the probe in a 50 GHz grid, which do not incur significant spurious FWM tones that could alter our results. For a larger system bandwidth with more interfering channels, the number of steps would have to be significantly increased, leading to infeasible simulation durations.

2.3.2 Self-Phase Modulation

As we could see in (2.104), the SPM term is a doubly degenerate form of the general FWM process in which only a single channel participates. Thus the phase-matching condition is always fulfilled and SPM generates nonlinear distortions with high efficiency. The Manakov equation for a single channel can be written as

$$\partial_z |S_\rho\rangle - i \left(\frac{\beta_2}{2} \partial_t^2 + \exp(-\alpha z) \tilde{\gamma} \langle S_\rho | S_\rho \rangle \right) |S_\rho\rangle = 0 \quad (2.106)$$

If the second term could be eliminated, the field $|S_\rho\rangle$ would pass undistorted through the system (in absence of PMD). This is the premise behind the research of *optical solitons*, which was a topic of intense interest in the 1990s (see e.g. [Mollenauer 1997] and [Agrawal 2007, Ch. 5] and the references therein). If the z -dependence of the term was eliminated by setting $\alpha = 0$, a solution exists for $\beta_2 < 0$. The result are pulses that are shaped like the hyperbolic secant function. They are called the fundamental solitons and are shown in Fig. 2.15.

A difficulty with transmitting solitons is the z -dependence of the nonlinear term in (2.106) due to attenuation. In order to keep the z -dependence of the second term in (2.106) minimal, one either needs short amplifier spans to avoid large values of αz (where z denotes the distance from the previous amplifier) or low pulse energies $\langle S_\rho | S_\rho \rangle$. The latter case corresponds to either small values of β_2 or long soliton pulses as a result of the soliton properties [Mollenauer 1997]. To achieve high data rates, having a low

¹⁸The expression that determines the step size is

$$\Delta(z) = \min \left[\Delta_0 \exp\left(\frac{\alpha}{1.5} z\right), \Delta_{\max} \right]$$

with $\Delta_0 = 25$ m, $\Delta_{\max} = 500$ m and where α is the fiber attenuation coefficient. All Δ within a PMD coarse step are equal, and Δ is rounded so that the coarse step length is an integer multiple.

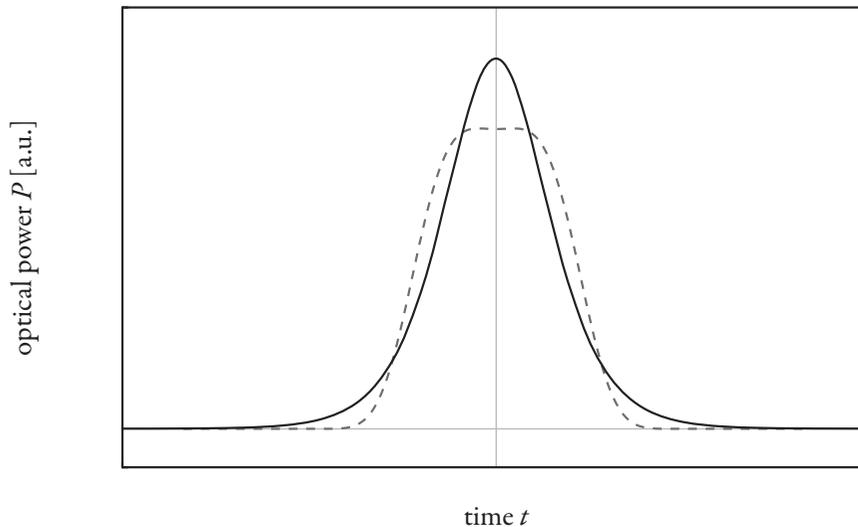


Figure 2.15 compares the pulse shapes of a fundamental soliton and a «regular» NRZ pulse; both pulses have an equal e^{-2} pulse width and the same pulse energy.

magnitude of β_2 is the most practical approach. However, for the required $\beta_2 > -25$ fs/GHz/km the accumulation of FWM products increases significantly (cf. Fig. 2.13) and WDM transmission is strongly impeded [Mamyshev 1996]¹⁹. If dispersion mapping is used in a fiber-optic system, β_2 in (2.106) will become z -dependent. In this case no «true» solitons can be transmitted. A workaround exists in so-called dispersion-managed solitons for which the path-averaged dispersion is compensated by the path-averaged nonlinearity [Smith 1997, Mollenauer 1998]. Such pulses will change their shape during propagation of a fiber span, but will (theoretically) be restored to their original shape at the end of each span. This technique is, however, limited to bit rates of about 10 Gbps in current SSMFs and also suffers from performance loss compared to transmission with fundamental solitons or other suitable modulation formats [Winzer 2006]. For these reasons, among others, high-speed soliton transmission has never been successfully commercialized, and to our knowledge there are currently no commercial high-speed soliton transmission systems in use.

For all other pulse shapes, (2.106) describes the accumulation of distortions during propagation of the fiber, which may also – as for solitons – at least partially cancel each other [Bellotti 1999]. In the remainder of this work, in order to focus on the interchannel nonlinear effects – especially XPOLM – we either use a CW probe channel ρ or reduce the optical power of ρ to below the nonlinearity threshold so that in both cases SPM effects are essentially removed.

For the sake of completeness, and keeping with our approach of separating the various effects into interchannel and intrachannel parts, we should mention that FWM, XPM, and XPOLM have intrachannel

¹⁹This corresponds to $D_{\text{GVD}} \approx 3$ ps/nm/km and is valid for 10 Gbps transmission; at higher bit rates, the limits on β_2 become even more stringent.

equivalents which are usually grouped under the label self-phase modulation. These effects operate like their interchannel variants, however, all participating frequency components belong to the same WDM channel. They can lead to ghost pulses in amplitude-modulated systems (intrachannel FWM) or timing jitter (intrachannel XPM) – intrachannel XPOLM has not yet been a topic of research. Further reading and references on the intrachannel variants can be found in [Essiambre 2002].

2.3.3 Cross-Phase & Cross-Polarization Modulation

We now focus on those FWM mixing products which may also build up quickly because they are degenerate in such a way that the phase-matching condition is always fulfilled. Both have been known for a long time, but XPOLM has only somewhat recently been observed, in the context of colliding solitons [Mollenauer 1995] – it had been theoretically predicted as early as 1974 by Manakov [Manakov 1974]. To start our analysis, we first separate the combined XPM/XPOLM term in (2.104) using the equality [Gordon 2000, eq. 3.9],

$$|S_\nu\rangle\langle S_\nu| = \frac{1}{2} \left(\langle S_\nu | S_\nu \rangle \bar{\mathbf{I}} + \mathbf{S}_\nu \cdot \vec{\sigma} \right) \quad (2.107)$$

where $\bar{\mathbf{I}}$ is the identity matrix, \mathbf{S}_ν is the Stokes vector corresponding to $|S_\nu\rangle$, and $\mathbf{S}_\nu \cdot \vec{\sigma}$ is the Pauli matrix expansion as in (2.26). Then, using only the XPM and XPOLM terms of the triple product (2.104) in the Manakov equation (2.101), the evolution equation describing cross-channel modulation becomes

$$\partial_z |S_\rho\rangle - i \frac{\beta_2}{2} \partial_t^2 |S_\rho\rangle - i \exp(-\alpha z) \tilde{\gamma} \sum_{\nu \neq \rho} \left(\frac{3}{2} \langle S_\nu | S_\nu \rangle + \frac{1}{2} \mathbf{S}_\nu \cdot \vec{\sigma} \right) |S_\rho\rangle = 0 \quad (2.108)$$

The first term in parentheses is polarization-independent and corresponds to the mean XPM, averaged over all possible relative SOPs between probe and interferer. The second term describes the SOP-dependent contribution to XPM which also results in XPOLM, as we will show shortly. For interferers which are co-polarized with the probe, this second term reduces to a phase modulation common to both components of $|S_\rho\rangle$ which increases the total XPM contribution to a factor of 2 compared to SPM. Similarly, for interferers orthogonal to the probe, the XPM contribution is reduced so that it equals the magnitude of SPM. Comparison of (2.108) and (2.106) reveals that, in the Manakov regime, phase modulation due to XPM is on average more efficient than SPM by a factor of 3/2. This is in contrast to the scalar version of the NLSE, in which XPM is more effective than SPM by a factor of 2, thus yielding misleading results.

CROSS-PHASE MODULATION As (2.108) shows XPM, like SPM, creates phase distortions in the probe channel. These phase distortions depend on the cumulate power in the interfering neighbor channels, and they are always positive. In phase-modulated systems, these distortions directly lead to transmission penalties, while in amplitude-modulated systems, the GVD term converts phase changes of the electric

field into amplitude distortions which may be even worse [Weber 2010]. Unlike with solitons, the XPM-induced phase shifts are not deterministic – in the sense that they depend on the uncorrelated signals in other channels – and cannot be used to counteract GVD. However, XPM is only detrimental when the phase distortion differs from one symbol to the next, i.e. the phase distortion from a CW neighbor does not lead to nonlinear penalties – this property is sometimes used in transmission experiments to keep optical amplifiers within operating parameters while varying the number of data-carrying channels in a DWDM system (see e.g. [Renaudier 2009]). Therefore, the statistical properties of the cumulate interferer power, especially its correlation from one probe symbol to the next, will determine to a large degree the actual nonlinear penalties accumulated between transmitter and receiver. Being the result of a stochastic process (in the frame of the probe channel, the modulation of the interfering channels is unknown and thus can be seen as a source of nonlinear phase noise), these penalties are also a random quantity. Compared to the large amount of literature on XPM in general (cf. the references in [Agrawal 2007]), this stochastic aspect has received little attention, even though the dependence on the dynamics of the cumulate interferer power has been noted (e.g. [Essiambre 2006]).

Usually, this dependence is condensed into an interrelation of the XPM penalty and the walk-off length, which we introduced in Section 2.2.1. This relation is quite intuitive: A (short) pulse in an interfering channel which causes a momentous phase change in a certain probe channel symbol, will cause a similar phase change in the temporally neighboring probe channel symbol after both channels have propagated the walk-off length L_{wo} – assuming the short pulse has not significantly changed shape and/or lost energy to attenuation. Hence, in a differentially encoded phase-modulated signal there is little penalty as a result of the interferer pulse. The faster the interfering pulse «walks off» – and thus the more probe channel symbols it affects and the smaller the induced phase change per probe symbol – the smaller will be the associated penalty. Therefore, high-GVD fibers will not only have less penalty due to FWM (see above), but also less penalty due to XPM [Bayvel 2002]. Also, the walk-off length of interfering channels decreases rapidly with an increase in their spectral distance from the probe channel, i.e. the penalty associated with XPM is highest for the channels immediately neighboring the probe [Cartaxo 1999].

CROSS-POLARIZATION MODULATION We have seen in (2.107) and (2.108) that the main topic of the present work, XPOLM, is closely related to XPM – they originate (at least partially) in the same term. In the following chapters we shall show that system design measures often associated with a reduction in XPM generally also reduce the impact of XPOLM. But first we will transform (2.108) into a form that lets us visualize the effect of XPOLM much more clearly.

To do this, we assume a CW probe channel – this lets us ignore the GVD-related term in (2.108).²⁰ We have seen an equation very similar to the XPOLM term in (2.108) before, in (2.22) – with (2.26) – on birefringence. Following the procedure in Section 2.2.2, we can transform (2.108) into Stokes space

²⁰This is only accurate as long as phase distortions due to XPM have not yet significantly broadened the initially single-peak spectrum – however, comparison of the theory obtained under such assumptions and numerical simulations confirm us in this assumption.

by expanding the z -derivative of $\mathbf{S}_\rho = \langle \mathbf{S}_\rho | \vec{\sigma} | \mathbf{S}_\rho \rangle$ and applying the spin vector rules (2.40). Because the pure phase modulation due to XPM is lost in translation to Stokes space, we are left with

$$\partial_z \mathbf{S}_\rho = \exp(-\alpha z) \tilde{\gamma} \mathbf{S}_\Sigma \times \mathbf{S}_\rho \quad (2.109a)$$

with

$$\mathbf{S}_\Sigma(z, t) = \sum_{\nu \neq \rho} \mathbf{S}_\nu(z, t) \quad (2.109b)$$

where we also used the property that the vector cross-product is distributive over addition, $\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C}$. Thus, the interpretation of XPOLM in Stokes space is this: As a result of XPOLM, the state of polarization of the probe channel is modified by a rotation around the instantaneous sum of the Stokes vectors of the interfering channels. Unlike birefringence, it does not depend solely on properties of the fiber, but properties of co-propagating WDM channels, mediated by the fiber nonlinearity. Due to its mathematical similarity with the evolution equation of birefringence, XPOLM is sometimes referred to as *nonlinear birefringence*. Because the interferers evolve as a result of the various effects introduced in the present chapter, and because – as with XPM – it is not the polarization modulation itself, but the difference from one probe channel symbol to the next which will prove detrimental for optical communication, it is again the statistical properties of the interferers which determine their impact via XPOLM. But unlike for XPM, not only the statistics of their cumulate power, but the statistics of the cumulate Stokes vector become relevant – a statistical model of XPOLM, as that proposed in the next chapter, must account for the envelope and polarization evolution of the interfering channels.

CROSS-POLARIZATION MODULATION

IN THE following chapter, we develop a statistical model of XPOLM which can subsequently be used to determine the impact of XPOLM in arbitrary optical communication channels. It is based on a general model of random motion of particles on a (Poincaré) sphere and quantified by statistically analyzing the rotation axis as the generating process for this motion. This analysis comprises a statistical description of the symbol shapes and SOPs in the interfering channels and their evolution. We obtain the probability distribution of the SOPs, which can be parameterized by various quantities, one of which, the degree of polarization (DOP), is directly measurable. We proceed to compare the predictions of our model to numerical simulations of communication systems for a variety of system parameters such as transmission fiber properties, initial SOPs, and the dispersion map, and derive methods to minimize the magnitude of XPOLM.

A NOTE ON NOTATION We have encountered two different kinds of averages until now: in Chapter 1 the time average operator $\mathcal{E}[\cdot]$ which denoted a very short averaging window – possibly not more than a few optical cycles – to determine the Stokes parameters and the degree of polarization of a plane monochromatic wave; and in Section 2.2.3 the polarization ensemble $\langle \cdot \rangle$ was defined as the statistical ensemble of all possible realizations of the evolution of the fiber PMD vector governed by a single diffusion constant D_{PMD} . We will slightly adjust both definitions in order to deal with the nonlinear phenomenon of XPOLM.

The state of polarization of a frequency component other than ω_0 (equal to the carrier frequency of the probe) at any location z is determined by the SOP at the fiber input and the local PMD properties of the fiber, according to the evolution equation (2.75). Because different SOPs of the interfering channels (represented by the SOP of their carrier frequency component) will result in different nonlinear SOP rotations of the probe according to (2.109), we must extend the polarization ensemble known from PMD and incorporating the local PMD properties of the fiber by the initial SOPs of the interfering channels. By the same reasoning, we also include the waveforms (bit patterns, symbol timing, and carrier phase) in the «new» polarization ensemble. We will keep using $\langle \cdot \rangle$ to denote the average over the infinite number of members of this ensemble. The set of possible probe channel SOPs taking into account all ensemble members is referred to as the *SOP population*.

In optical communication systems we are generally dealing with polarized waves which have an instantaneous (i.e. a few optical cycles) DOP of unity according to the definition (1.28) in Chapter 1, resulting from the very short time scale for the averaging process (femtoseconds). An SOP thus is always of unity length and ends on the Poincaré sphere.²¹ However, the SOP of the signal is subject to changes as a result of XPOLM, albeit on a much larger timescale than that used to determine

²¹SOPs of partially polarized light end inside the Poincaré sphere, as the SOP only includes the polarized portion of light.

the Stokes parameters – typically pico- to nanoseconds (a discussion of the rate of this change is performed in Section 5.1). We thus introduce a new time averaging window to capture these variations. The length of this window corresponds to the duration of any member of the polarization ensemble, i.e. as long as the linear polarization evolution of the interfering channels everywhere is unchanged. We will keep using $\mathcal{E}[\cdot]$ to denote this time average, to distinguish it from the polarization ensemble. This definition of the DOP is relatively common when dealing with the phenomena of PMD or XPOLM [Bononi 2003, Vannucci 2003, Kikuchi 2001].

3.1 RANDOM ROTATIONS IN STOKES SPACE

When discussing the polarization properties of a signal, it is usually sufficient to consider the SOPs as defined in (1.30) instead of the Stokes vectors. In the present context, a «single» SOP is defined as $\hat{\mathbf{S}}(z, t)$ for some specific time t in the frame of reference moving with the group velocity of the field at ω_0 . The evolution of such an SOP of the probe channel with propagation along z can be derived from the evolution equation of the probe Stokes vector (2.109) by dividing both sides by $|\mathbf{S}_\rho|$,

$$\partial_z \hat{\mathbf{S}}_\rho = \exp(-\alpha z) \tilde{\gamma} \mathbf{S}_\Sigma \times \hat{\mathbf{S}}_\rho \quad (3.1)$$

This equation describes a sequence of (infinitesimal) rotations around the sum of Stokes vectors \mathbf{S}_Σ as the signal propagates. We have shown in the previous chapter that the length of each of the Stokes vectors contributing to \mathbf{S}_Σ is a random quantity due to data modulation in combination with GVD-induced walk-off (cf. Section 2.2.1). The direction of these Stokes vectors is also random as a result of the PMD-induced polarization evolution (cf. Section 2.2.3). We further assume that all individual channels of the WDM system are launched into the transmission fiber with a random initial polarization. While this may not be entirely true for field-deployed systems, it is a working approximation, as the polarization evolution (2.75) due to PMD – which we will quantify in detail in Section 3.2.3 – randomizes the signal polarization during propagation. Thus, in long-haul transmission systems the actual initial SOP is of minor importance unless PMD is very small. The benefit of the assumption of uniformly distributed initial SOPs is the ability to give our results in closed-form expressions. We will also take a look at the special cases of co-polarized launch and mutual orthogonally polarized launch (also called *polarization interleaving* between spectrally neighboring channels) in Section 3.2.4.

With these assumptions about the Stokes vectors \mathbf{S}_ν , the rotation axis in (3.1) must also be a random vector quantity, and hence the motion $\partial_z \hat{\mathbf{S}}_\rho$ of the probe SOP will be random. Due to this random nature, we will call it a *Brownian motion*, restricted to the surface of the Poincaré sphere. We have also already noted in the preceding section that it is not the XPOLM-induced polarization modulation per se but the time-varying nature of this modulation which is detrimental for optical communication. However, at each location z the Stokes sum \mathbf{S}_Σ will have a time-constant average value around which it varies over time. According to the evolution equation (3.1), this average value will induce a polarization

rotation that is common to the probe SOP at all times t – the Brownian motion above is actually a Brownian motion with drift, where the drift term originates in the average of \mathbf{S}_Σ . Any receiver that is able to adapt to the slow polarization changes due to the time-varying nature of birefringence will also be able to adapt to this additional mean nonlinear rotation, because they change on the same timescale, and the mean is taken over the duration of each polarization ensemble member. We can thus remove the drift term of the vector sum by defining a new SOP variable, $\hat{\mathbf{S}}_p$, which only experiences the variations of \mathbf{S}_Σ relative to its mean,

$$\hat{\mathbf{S}}_p = \bar{\Xi}^{-1} \hat{\mathbf{S}}_\rho \quad (3.2a)$$

with the rotation transformation in Stokes space

$$\bar{\Xi}(z) = \exp \left[\int_0^z \mathcal{E}[\mathbf{S}_\Sigma(\zeta, t)] \times d\zeta \right] \quad (3.2b)$$

where the cross-product operator $\mathcal{E}[\mathbf{S}_\Sigma(\zeta, t)] \times$, which we had introduced in (2.46), is constructed here from the time-averaged Stokes vector sum at each position ζ . The transformation $\bar{\Xi}(z)$ offsets the common nonlinear polarization rotation up to position z . The integral in the exponent in the definition (3.2b) is again to be understood symbolically, as in (2.48), due to the cross-product operators $\mathcal{E}[\mathbf{S}_\Sigma] \times$ at different ζ being non-commutative. Inserting (3.2) into (3.1) yields the evolution equation of the probe SOP without the nonlinear polarization drift,

$$\partial_z \hat{\mathbf{S}}_p(z, t) = \exp(-az) \tilde{\gamma} \left(\mathbf{S}_\Sigma(z, t) - \mathcal{E}[\mathbf{S}_\Sigma(z, t)] \right) \times \hat{\mathbf{S}}_p(z, t) \quad (3.3)$$

In writing (3.3), we have implicitly set $\bar{\Xi}^{-1} \mathbf{S}_\Sigma \approx \mathbf{S}_\Sigma$, i.e. we neglect the effect of the coordinate transformation $\bar{\Xi}$ on the interfering channels, keeping with our premise that we only consider their linear evolution. Eq. (3.3) describes the evolution that causes the detrimental effects of XPOLM. Because the drift term has been removed, the mean change of direction of the SOPs is zero at any position z ,

$$\mathcal{E}[\partial_z \hat{\mathbf{S}}_p(z, t)] = 0 \quad (3.4)$$

The difference between the evolution of $\hat{\mathbf{S}}_\rho$ and that of $\hat{\mathbf{S}}_p$ is illustrated in Fig. 3.1.

To obtain any meaningful, closed-form statements about the distribution of the $\hat{\mathbf{S}}_p$ around its mean on the Poincaré sphere (from which to draw conclusions about the impact of XPOLM) we must assume the infinitesimal increments $\partial_z \hat{\mathbf{S}}_p$ of the Brownian motion to be isotropic for each (z, t) pair. However, in any single member of the polarization ensemble, at a fixed location z , the SOPs of the interfering channels will be fixed in time and only the length of their Stokes vectors changes as a result of data modulation. The increments $\partial_z \hat{\mathbf{S}}_p$ will have zero mean per definition, but will not necessarily be isotropic, depending on the orientation of the individual Stokes vectors. If we extend our observations to the po-

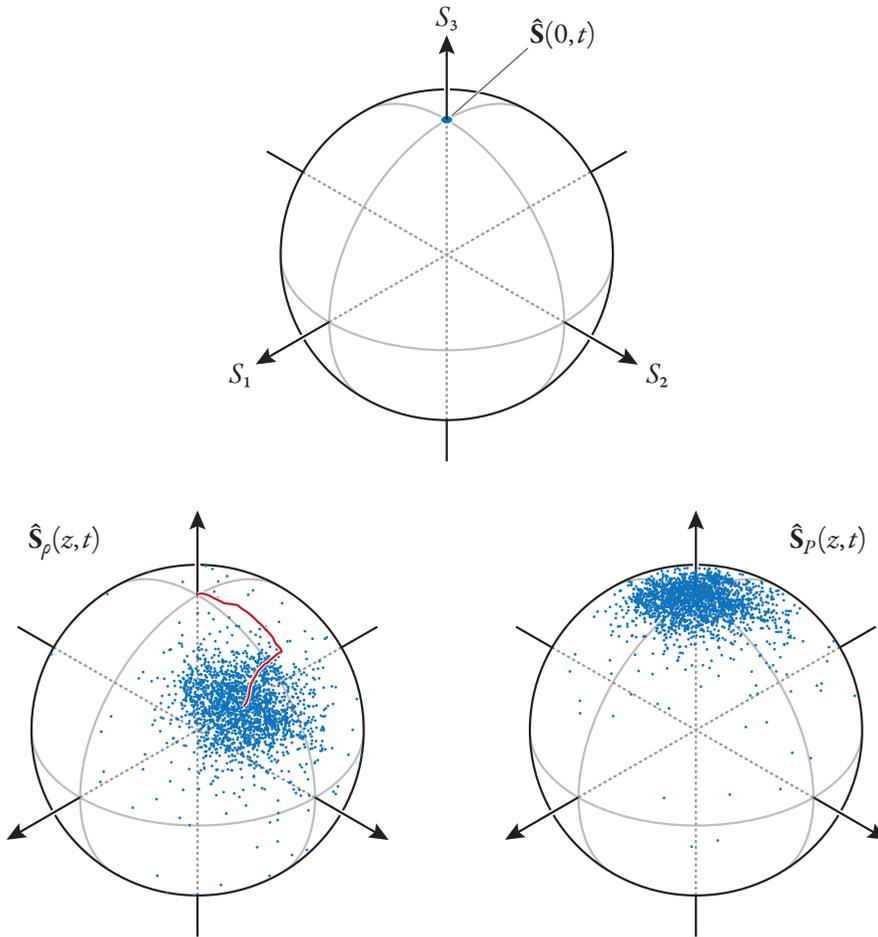


Figure 3.1 illustrates the operation of the transformation (3.2): The top image shows the SOP $\hat{S}_\rho(0, t) = \hat{S}_p(0, t)$ of the probe channel at the transmitter. The bottom left graphic shows the distribution of $\hat{S}_\rho(z, t)$ at $z = 1000$ km (after propagation of 10 spans) as a result of XPOLM. The red line plots the evolution of the mean SOP as a consequence of $\mathcal{E}[\mathbf{S}_\Sigma]$ being greater than zero. The bottom right graphic shows the distribution of $\hat{S}_p(z, t)$ – the motion of the mean SOP has been removed by the coordinate transformation (3.2).

(Figure shows actual simulation data.)

larization ensemble, the $\hat{\mathbf{S}}_y(z, t)$ become distributed uniformly on the unit sphere, and the distribution of $\mathbf{S}_\Sigma - \mathcal{E}[\mathbf{S}_\Sigma]$ will also be uniform. Then all directions of the motion $\partial_z \hat{\mathbf{S}}_P$ in (3.3) must be equally likely and the nonlinear polarization rotation is isotropic as required.

Since we will also investigate the special cases of co-polarized and polarization-interleaved launch SOPs of all channels, we shall also briefly look at the isotropy properties of the increments to $\hat{\mathbf{S}}_P$. Here, the $\hat{\mathbf{S}}_y(z, t)$ will not be uniformly distributed on the unit sphere everywhere, but the population distribution will be symmetric around the initial SOP $\hat{\mathbf{S}}_y(0, t)$, because the polarization diffusion due to PMD itself is isotropic (the PMD vector is assumed to be uniformly distributed on the Poincaré sphere and thus the above considerations about isotropy also apply). Hence, this symmetry condition is actually sufficient for $\partial_z \hat{\mathbf{S}}_P$ to be isotropic – uniformity is not required. In all three cases (uniformly distributed, co-polarized, and mutually orthogonal launch SOPs) – which are the most relevant ones from a practical standpoint – we may use the assumption of isotropic increments in the orientation of $\hat{\mathbf{S}}_P$ and the model of a Brownian diffusion restricted to a spherical surface.

THE BROWNIAN DISTRIBUTION The statistics of a series of random isotropic rotations on the unit sphere were derived under various conditions by Roberts and Ursell in the 1960s [Roberts 1960]. They considered a process consisting of a large number of such rotations to arrive at its resulting probability density function in the limit of an infinite number of such steps and in terms of spherical coordinates (φ, θ) , given as

$$p_B(\varphi, \theta, \mathcal{V}) = \frac{1}{4\pi} \sum_{\mu=0}^{\infty} C(\mu, \mathcal{V}) LP_\mu(\cos \theta) \quad (3.5a)$$

in which

$$C(\mu, \mathcal{V}) = (2\mu + 1) \exp\left(\frac{-\mu(\mu+1)}{4} \mathcal{V}\right)$$

where $LP_\mu(\cdot)$ denote the Legendre polynomials of integer order μ and the azimuth φ and colatitude θ are defined as in Fig. 1.4 when the initial probe SOP is $\hat{\mathbf{S}}_P = (0, 0, 1)^T$ as in Fig. 3.1. This is the population distribution of the SOPs at the end of a long transmission line resulting from our model of random rotations.²² The scalar parameter \mathcal{V} is a variance parameter and fully quantifies the distribution. Roberts and Ursell describe it as «the variance of the corresponding plane motion», i.e. the variance of a plane random walk with isotropic increments of length $|\partial_z \hat{\mathbf{S}}_P|$. The parameter \mathcal{V} plays a central role in our model of XPOLM: Since the distribution of the probe SOPs is given by (3.5), we need only determine \mathcal{V} to fully parameterize it. As we will show in this chapter, it is possible to well approximate \mathcal{V} , only knowing the various fiber parameters and interferer signal shapes.

²² «Long» refers to a length sufficient for the SOPs to approach this limit distribution, similar to how the DGD distribution approaches the Maxwellian with an increasing number of independent segments (cf. Fig. 2.12). This length varies with the length scale on which the Stokes vectors of the interfering channels are randomized. A lower limit is approximately a few spans (cf. Section 3.2.3).

The distribution is normalized such that

$$\int_0^{2\pi} \int_0^\pi p_B(\varphi, \theta, \mathcal{U}) \sin \theta d\theta d\varphi = 1 \quad (3.6)$$

Since $p_B(\varphi, \theta, \mathcal{U})$ does not depend on the azimuth φ , it must be symmetric about the $\theta = 0$ axis, as we would expect as a result of isotropic increments. For $\mathcal{U} = 0$, we have $p_B(\varphi, \theta, 0) = \delta(\theta)$, i.e. the distribution is concentrated in the ‘‘pole’’, and for $\mathcal{U} \rightarrow \infty$, the distribution approaches the uniform distribution on the sphere. The *Brownian distribution* (also known as Roberts-Ursell distribution) is shown in Fig. 3.2 for various values of the parameter \mathcal{U} .

The variance parameter \mathcal{U} is an abstract quantity that cannot be measured directly at the receiver, but can only be obtained by fitting the SOP distribution. There exists another, measurable quantity which we have used before which also describes how dispersed the polarization states of a wave are on the Poincaré sphere: the degree of polarization. Both parameters are directly related.

The DOP D , time-averaged over a single member of the polarization ensemble, is

$$D = \frac{\sqrt{\mathcal{E}[S_1(z, t)]^2 + \mathcal{E}[S_2(z, t)]^2 + \mathcal{E}[S_3(z, t)]^2}}{\mathcal{E}[S_0(z, t)]} = \sqrt{\frac{\mathcal{E}[\mathbf{S}(z, t)] \cdot \mathcal{E}[\mathbf{S}(z, t)]}{\mathcal{E}[S_0(z, t)]^2}} \quad (3.7)$$

For a near-cw channel, we have $\mathcal{E}[S_0] \approx S_0$ and with the definition (1.29) of the SOP $\hat{\mathbf{S}}$, we have

$$D^2 \approx \mathcal{E}[\hat{\mathbf{S}}(z, t)] \cdot \mathcal{E}[\hat{\mathbf{S}}(z, t)] \quad (3.8)$$

If the direction of $\mathcal{E}[\hat{\mathbf{S}}(z, t)]$ is the same as that of the initial SOP $\hat{\mathbf{S}}(0, t)$ and if the signal is fully polarized at $z=0$ so that $\mathcal{E}[\hat{\mathbf{S}}(0, t)] = \hat{\mathbf{S}}(0, t)$, as is the case for $\hat{\mathbf{S}}_p$ (cf. Fig. 3.1), we can write

$$\hat{\mathbf{S}}_p(0, t) = \frac{\mathcal{E}[\hat{\mathbf{S}}_p(z, t)]}{\sqrt{\mathcal{E}[\hat{\mathbf{S}}_p(z, t)] \cdot \mathcal{E}[\hat{\mathbf{S}}_p(z, t)]}} \quad (3.9)$$

where, according to (3.8), the denominator of the fraction is approximately D , so that

$$\mathcal{E}[\hat{\mathbf{S}}_p(z, t)] \approx D \cdot \hat{\mathbf{S}}_p(0, t) \quad (3.10)$$

Inserting (3.10) for one of the factors in (3.8) we have

$$D \approx \mathcal{E}[\hat{\mathbf{S}}_p(z, t)] \cdot \hat{\mathbf{S}}_p(0, t) = \mathcal{E}[\hat{\mathbf{S}}_p(z, t) \cdot \hat{\mathbf{S}}_p(0, t)] \quad (3.11)$$

Assuming a coordinate system (φ, θ) aligned such that the initial SOP $\hat{\mathbf{S}}_p(0, t)$ of the probe channel

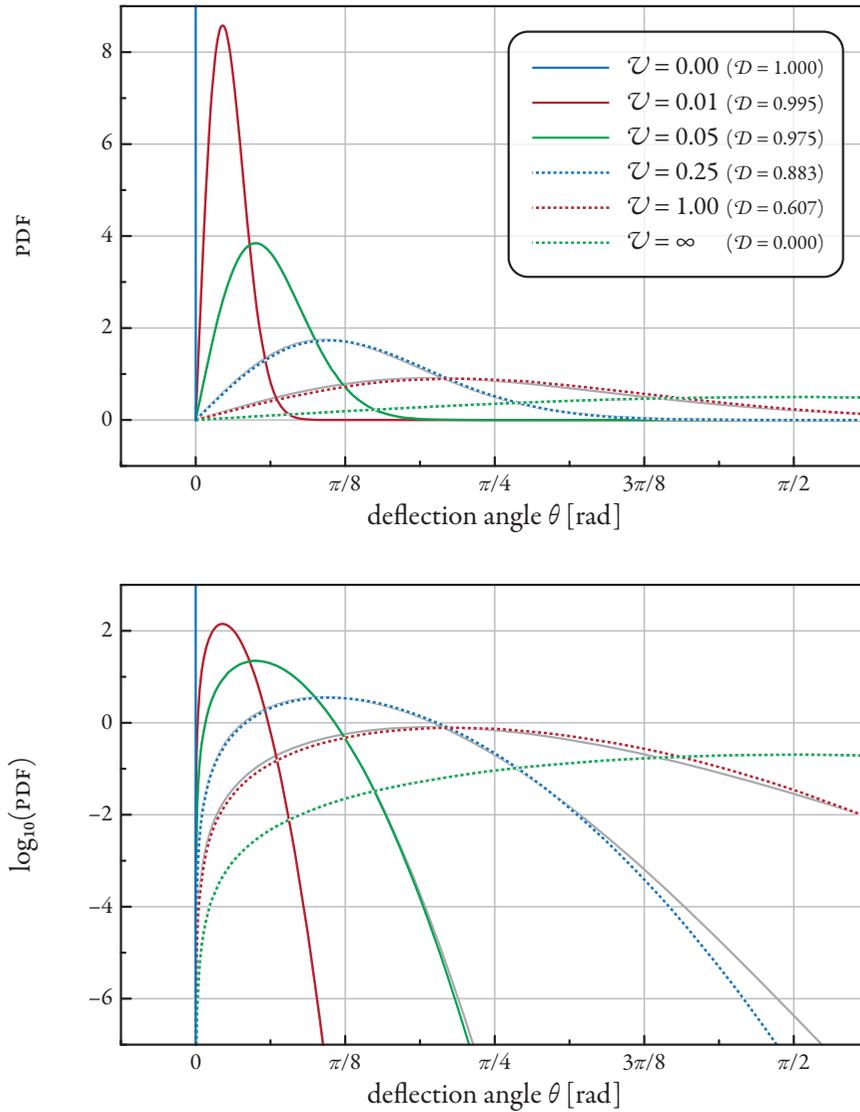


Figure 3.2 shows examples of the Brownian distribution (3.5) for different parameters \mathcal{U} in linear (top) and logarithmic (bottom) scale. The relation between \mathcal{U} and \mathcal{D} is given in (3.14). The light grey lines in the graphs correspond to the equivalent Fisher distributions (3.16). When they cannot be seen, the Brownian and Fisher distributions are congruent for the particular value of \mathcal{U} .

coincides with the north pole (for which $\theta=0$), we can write

$$D(z) \approx \mathcal{E}[\cos \Theta(z, t)] \quad (3.12)$$

The DOP is given by the expectation of $\cos \Theta$.²³ When we again consider the whole polarization ensemble, for which the infinitesimal increments are isotropic, we know that the population of Θ is distributed according to the Brownian distribution (3.5). We can thus take the ensemble average of (3.12) and have

$$\langle \mathcal{E}[\cos \Theta(z, t)] \rangle = \frac{1}{4\pi} \int_0^{2\pi} \int_0^\pi \cos \theta p_B(\varphi, \theta, \mathcal{V}) \sin \theta d\theta d\varphi \quad (3.13a)$$

$$= \exp\left(-\frac{\mathcal{V}}{2}\right) \quad (3.13b)$$

with $\mathcal{V} = \mathcal{V}(z)$. To arrive at this simple form, we have used the orthogonality property of the Legendre polynomials, which eliminates all but one term from the infinite sum in (3.5). Defining the population DOP \mathcal{D} ,

$$\mathcal{D}(z) = \langle D(z) \rangle \quad (3.14)$$

we can write

$$\mathcal{D}(z) \approx \exp\left(-\frac{\mathcal{V}(z)}{2}\right) \quad (3.15)$$

The above equation gives the relation between the abstract variance parameter \mathcal{V} and the DOP. We shall mostly use the latter, because it is a familiar concept in polarization optics. However, care must be taken not to confuse the nonlinear depolarization resulting from XPOLM with the actual (instantaneous) degree of polarization of an optical wave – both differ in the length of the averaging window. Since the variance of a process is a statistical cumulant, it will be helpful when combining the contributions of various sources and thus during the quantitative analysis of XPOLM.

THE FISHER DISTRIBUTION The Brownian distribution above is the equivalent on the unit sphere of the Gaussian distribution in the plane resulting from a planar diffusion process. However, consisting of an infinite sum, it is very unwieldy to derive any dependent quantities from. In their paper [Roberts 1960], Roberts and Ursell also noted the similarity of their distribution to another one derived a few years earlier by Fisher [Fisher 1953]. Fisher postulated a distribution of «elementary errors

²³Another note on notation: Since we have to clearly differentiate between the time- and position-dependent random variable $\Theta(z, t)$ and the spherical coordinate θ as parameter of the distribution of Θ , we use an uppercase letter for the random variable and a lowercase letter for the parameter, as commonly done in texts on probability theory.

over the surface of the unit sphere, which is the field of possible observations», given by

$$p_F(\varphi, \theta, \kappa) = \frac{\kappa}{4\pi \sinh \kappa} \exp(\kappa \cos \theta) \quad (3.16)$$

where κ is a reciprocal measure of the variance, and thus a signal-to-noise ratio parameter. The normalization coefficient ensures that the condition (3.6) is also fulfilled for p_F . Both distributions are related via the DOP, which can be determined for the Fisher distribution analogous to (3.13a) and is obtained as

$$\coth \kappa(z) - \frac{1}{\kappa(z)} = \mathcal{D}(z) = \exp\left(-\frac{\mathcal{V}(z)}{2}\right) \quad (3.17)$$

As we can see in Fig. 3.2, the difference between both distributions may become relevant for comparatively large values of \mathcal{V} (and correspondingly low values of the DOP). However, when dealing with nonlinear distortion we are more interested in the region of small disturbances, as will become clear throughout the remainder of this work, and especially Chapter 5. Therefore we will neglect the small differences between the Brownian and the Fisher distributions and work with whichever is more suitable for the task at hand, which often turns out to be the Fisher distribution.

COMPARISON WITH SIMULATIONS The above probability distributions result from the proposed model of XPOLM as long sequence of random rotations. To verify this model, its predictions must be compared to the actual behavior of the SOPs within a transmission fiber. In the present work this relates to numerical simulations of optical communication systems. In order to verify our results we have created a numerical model of a multi-span DWDM optical communication link in MATLAB. A typical implementation is shown schematically in Fig. 3.3. The DWDM source consists of a probe channel ρ spectrally surrounded by ten interfering channels.²⁴ All channels are then propagated over a number of amplifier spans with in-line optical dispersion compensation – this compensation may be complete on a per-span basis, resulting in a *resonant dispersion map* or may include some residual dispersion per span (RDPS), causing a monotonous increase in the accumulated dispersion at the beginning of each span.

Dispersion maps illustrate the management of GVD by compensating the anomalous GVD accumulated by transmission over most low-loss SMFs within the transmission link [Bergano 2002, Sec. 4.4]. Such a compensation is usually performed at the end of each amplifier span. When dispersion management is used, the local GVD may be large, but the magnitude of accumulated GVD

$$B_2(z) = \int_0^z \beta_2(\zeta) d\zeta \quad (3.18)$$

will remain quite small, as will be the span-averaged GVD. An exemplary dispersion map is plotted in

²⁴A larger number of channels in a 50 GHz spectral grid will lead to spurious FWM, as discussed in Section 2.3.1, when not significantly decreasing the step size, leading to an equally significant increase in simulation duration.

Fig. 3.3. In each span some of the accumulated dispersion of that span is left uncompensated. It has been shown that RDPS can significantly reduce the effects of XPM in similar systems [Chandrasekhar 2007]. Technically, the RDPS may also be negative, resulting in accumulation of positive (span-averaged) dispersion, but this configuration has been shown to be unfavorable [Xiao 2007, cf. Fig. 7e]. If the residual dispersion per span is zero, the dispersion map becomes periodic, or resonant. In field-deployed systems there will always be at least some small RDPS as it is difficult and expensive (and not necessarily beneficial) to match the dispersion compensation to the accumulated GVD for all channels in each span. Thus, truly resonant systems are possible only in computer simulations and thus to be regarded as an idealized limit of low RDPS. Apart from the local dispersion and RDPS there exist two more variables that characterize dispersion maps: the initial GVD added at the transmitter before entering the transmission fiber and residual accumulated GVD remaining uncompensated at the receiver. Both may be beneficial for the fidelity of the signal at the receiver by reducing intrachannel distortions [Essiambre 2002, Sec. 2.4].

Within the scope of the present work we only wish to distinguish the impact of the various interchannel Kerr effects without emphasizing the mostly unrelated class of intrachannel nonlinearities, which may possibly even be compensated electronically in future generations of transmission systems. For this reason, the optical power in the probe channel is reduced by 3 dB compared to the interfering channels. A careful optimization of initial GVD added at the transmitter or residual GVD remaining at the receiver is thus not necessary and not performed. The simple dispersion map is designed such that the accumulated GVD is zero upon the optical field entering and exiting the transmission fiber. Furthermore, 10 Gbps NRZ modulation is chosen for the interfering channels to emphasize the impact of nonlinear interchannel distortion – the comparatively slow amplitude modulation is regarded as a significant source of such distortion [Bayvel 2002, Sec. 4], as we will show in Section 3.2.2.

The nonlinear propagation calculation is done with the well-established split-step Fourier method, already mentioned in Section 2.3.1 on FWM [Agrawal 2007, Sec. 2.4], using the *SSPROP* library developed at the University of Maryland.²⁵ At the receiver, the probe channel is isolated and analyzed. Various relevant parameters of the simulation setup are shown in Table 3.1. An optical filter cascade at the transmitter emulates the concatenation of an interleaver and a multiplexer, to obtain realistic pulses in the fiber. At the receiver, the probe is isolated using a single *DEMUX* filter.

Fig. 3.4 shows exemplary distributions of the SOPs at the output of the system when using a CW input probe channel. Each distribution shown is an ensemble of 500 iterations with random initial SOPs and PMD properties. In the practically most relevant case of a dispersion map that has at least some RDPS, the match is very good, supporting our proposed model of a sequence of random rotations for the effect of XPOLM.

The deviation of the simulated data, especially in the tail of the distribution, can become noticeable

²⁵*SSPROP – Split-Step Fourier Propagation* is free software from the Photonics Research Lab, University of Maryland, available at

<http://www.photonics.umd.edu/software/ssprop/index.html>

We used a customized version 3.0.1 which implemented the birefringence-independent 8/9 coefficient for the fiber nonlinearity according to the Manakov equation (2.101).

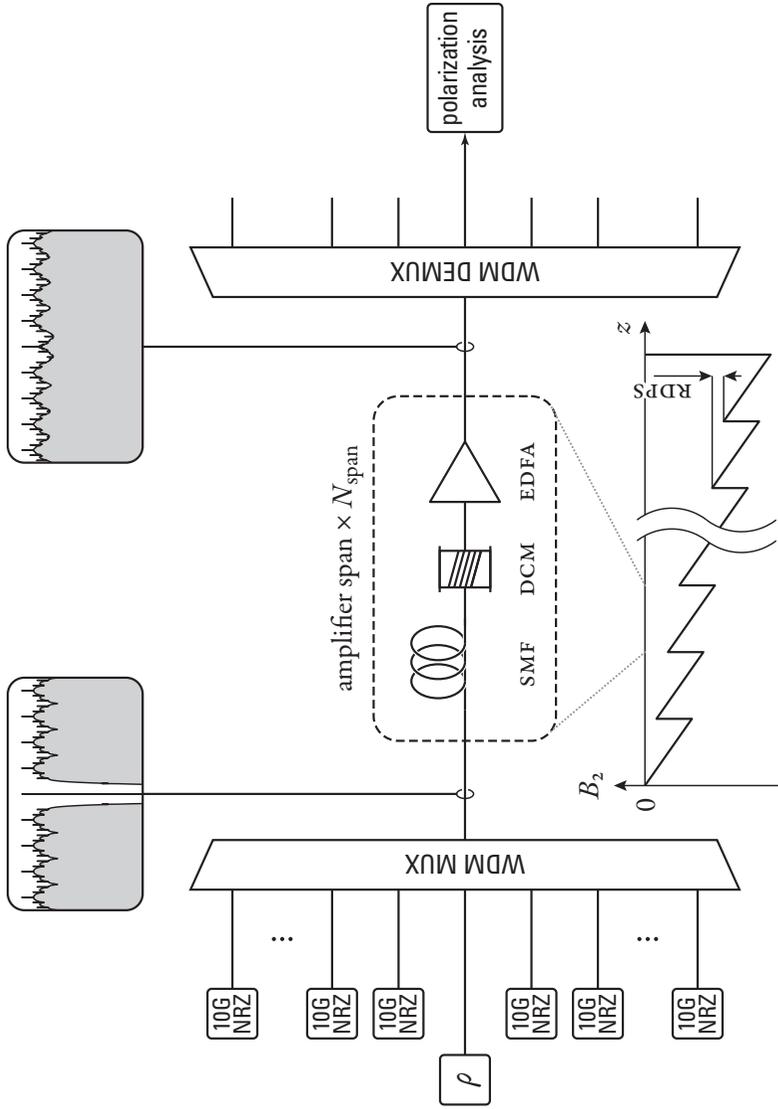


Figure 3.3 illustrates the general setup of the DWDM communication link (see description in the text). Dispersion compensation (DCM) is treated linearly; amplification (EDFA) fully compensates span loss for all channels. The inset shows the evolution of the accumulated GVD B_2 over the length of the transmission line. DWDM spectra are shown at the beginning and end of the transmission, with significant nonlinear broadening visible for all channels.

Table 3.1 Typical system parameters used in split-step simulations and theoretical predictions (where applicable) unless noted otherwise.

| | |
|--------------------------------------|----------------------------------|
| bit rate (interferer) | 10 Gbps |
| modulation format (interferer) | NRZ |
| modulation format (probe) | CW |
| input power (interferer) P_ν | 2 mW |
| input power (probe) P_ρ | $P_\nu/2$ |
| number of interferers | 10 |
| interleaver bandwidth | 38 GHz |
| interleaver shape | 4th-order Gaussian |
| MUX filter bandwidth | 42 GHz |
| MUX filter shape | 4th-order Gaussian |
| probe DEMUX filter bandwidth | 25 GHz |
| probe DEMUX filter shape | 5th-order Gaussian |
| PRBS | $2^{16} - 1$ (subsequences used) |
| simulation time window | 1024 bits |
| laser linewidth | 0 (no phase noise) |
| channel spacing | 50 GHz |
| span length L_{span} | 100 km |
| system length L | 1,000 km |
| SSMF GVD parameter β_2 | -128 fs/GHz/km (16 ps/nm/km) |
| SSMF TOD parameter β_3 | 0 |
| SSMF attenuation α | 0.2 dB/km |
| SSMF nonlinearity parameter γ | 1.31 (W km)^{-1} |
| EDFA noise figure | 1 (no noise added) |
| RDPS (if used) | -2048 fs/GHz (256 ps/nm) |
| mean PMD coarse step length | 1000 m |
| STD PMD coarse step length | 100 m |
| ensemble size | 100 members |

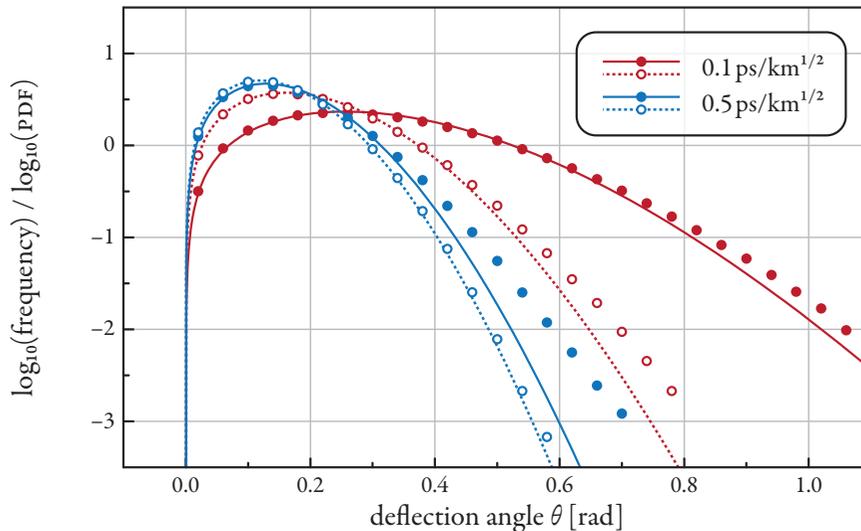


Figure 3.4 shows the SOP distribution for four exemplary systems; symbols represent simulation values (using 500 randomly generated system iterations) and lines are (linear scale) curve fits to the Fisher distribution (3.16) (open symbols / dashed lines: RDPS of -2048 fs/GHz; filled symbols / solid lines: no RDPS). Optical power in the interfering channels was 2 mW. Also see discussion in the text.

in certain system configurations. In general, the match is better when RDPS is used or the number of interferers is increased. Both lead to a greater randomness of the interfering channels' Stokes vectors and hence of the nonlinear polarization rotations of the probe. We have seen a similar difference, due to lack of randomness, in the statistics of the PMD vector as shown in Fig. 2.12, which deviate from the analytic Maxwellian distribution if the number of independent sections becomes too small. Contrary to our claim in Section 2.3.1, FWM does also play a (minor) role in the evolution of the probe SOPs, which we have neglected in our theoretical approach.

A more thorough discussion of the nature of the randomness of the interferers' Stokes vectors leads us directly to an expression predicting the nonlinear depolarization due to XPOLM and thus not only a description of the general SOP distribution shape, but also an analytic prediction of its parameter. This discussion follows in the subsequent section.

3.2 ESTIMATION OF THE PARAMETER \mathcal{U}

A model for XPOLM is only partially useful (so far we can determine the distribution of the SOPs when the DOP is known) if we cannot use it to predict the behavior of the system, and derive methods from it how to improve this behavior. We therefore devote this section to estimating the impact of XPOLM by developing the model from being qualitative to being quantitative.

A most useful property of the Brownian distribution (3.5) is that its parameter, \mathcal{U} , is defined in terms of a mathematically meaningful, though not measurable, quantity: the variance of an equivalent plane random walk. However, to formally link the plane random walk and the XPOLM diffusion process, we must adapt the theory of Roberts and Ursell. A crucial difference between their work and our problem is their assumption of independent, isotropic steps – the XPOLM-induced polarization rotations at two locations z_1 and z_2 need not be and generally will not be independent. According to (3.3), the increments $\partial_z \hat{\mathbf{S}}_P$ are a function of the Stokes vector sum \mathbf{S}_Σ , which changes as a result of GVD and PMD. These changes are not sudden and uncorrelated, and therefore we can expect significant correlation between the $\partial_z \hat{\mathbf{S}}_P$ at z_1 and z_2 .

If

$$\mathbf{W} = \sum_{k=1}^K \mathbf{W}_k \quad (3.19)$$

describes the the position at the end of a planar random walk with real-valued two-dimensional steps \mathbf{W}_k which corresponds to the spherical walk according to Roberts and Ursell, its variance is

$$\mathcal{U} = \langle |\mathbf{W}|^2 \rangle = \langle \mathbf{W} \cdot \mathbf{W} \rangle \quad (3.20a)$$

$$= \left\langle \sum_{k=1}^K \mathbf{W}_k \cdot \sum_{k=1}^K \mathbf{W}_k \right\rangle \quad (3.20b)$$

$$= \sum_{k=1}^K \sum_{l=1}^K \langle \mathbf{W}_k \cdot \mathbf{W}_l \rangle \quad (3.20c)$$

where we have averaged over the ensemble of all possible such walks. In Roberts' and Ursell's walk, all steps are independent, so the cross-expectations vanish and they are left with

$$\mathcal{U} = \sum_{k=1}^K \langle \mathbf{W}_k \cdot \mathbf{W}_k \rangle = \sum_{k=1}^K \langle |\mathbf{W}_k|^2 \rangle \quad (3.21)$$

while we must use the full form (3.20c). We can then make the transition to infinitesimal steps,

$$\mathcal{U}(L) = \int_0^L \int_0^L \langle d\mathbf{W}(z_1) \cdot d\mathbf{W}(z_2) \rangle \quad (3.22)$$

in which $d\mathbf{W}$ is a differential increment of the Brownian motion (on the plane). It is not trivial to accurately relate the correlation between different increments $\partial_z \hat{\mathbf{S}}_P$ at z_1 and z_2 on the surface of the unit sphere to the correlation of the planar increments $d\mathbf{W}$, especially if the SOP $\hat{\mathbf{S}}_P$ changes significantly. We will therefore make an approximation that works well as long as $\hat{\mathbf{S}}_P$ remains near its origin,

$\hat{\mathbf{S}}_P(0, t)$: We assume – only for the determination of the parameter \mathcal{V} – that $\hat{\mathbf{S}}_P(z, t) = \hat{\mathbf{S}}_P(0, t)$, the SOP does not change. Since we are primarily interested in slight to moderate nonlinear distortions, this assumption should be sufficiently justified. We then arbitrarily define the coordinate system so that $\hat{\mathbf{S}}_P(0, t) = (0, 0, 1)^T$, as shown in Fig. 3.1 (top). Inserting this probe SOP in (3.3)

$$\partial_z \hat{\mathbf{S}}_P(z, t) = \Psi_2(z, t) \mathbf{e}_1 - \Psi_1(z, t) \mathbf{e}_2 \quad (3.23)$$

with

$$\Psi(z, t) = \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_2 \end{pmatrix} = \exp(-\alpha z) \tilde{\gamma} \left(\mathbf{S}_\Sigma(z, t) - \mathcal{E}[\mathbf{S}_\Sigma(z, t)] \right) \quad (3.24)$$

where \mathbf{e}_1 , \mathbf{e}_2 , and \mathbf{e}_3 are the base vectors of the coordinate system defined above. The increment $\partial_z \hat{\mathbf{S}}_P$ now only depends on those two components of Ψ which are orthogonal to $\hat{\mathbf{S}}_P$. Furthermore, all $\partial_z \hat{\mathbf{S}}_P$ are now confined to equatorial 1-2 plane and we have reduced the problem to a planar one. Like Roberts and Ursell, we now equate $\partial_z \hat{\mathbf{S}}_P$ and $d\mathbf{W}$ and insert (3.23) into (3.22),

$$\begin{aligned} \mathcal{V}(L) = & \int_0^L \int_0^L \mathbf{e}_1 \cdot \mathbf{e}_1 \left\langle \mathcal{E}[\Psi_2(z_1, t) \Psi_2(z_2, t)] \right\rangle \\ & - 2 \mathbf{e}_1 \cdot \mathbf{e}_2 \left\langle \mathcal{E}[\Psi_2(z_1, t) \Psi_1(z_2, t)] \right\rangle \\ & + \mathbf{e}_2 \cdot \mathbf{e}_2 \left\langle \mathcal{E}[\Psi_1(z_1, t) \Psi_1(z_2, t)] \right\rangle dz_1 dz_2 \end{aligned} \quad (3.25)$$

The ensemble average over all possible walks in (3.22) implies the time average above, because $\partial_z \hat{\mathbf{S}}_P(z, t)$ yields a different walk over z at each time index t . The ensemble average in (3.25) is again the polarization ensemble defined at the beginning of the chapter. Using the orthogonality of the basis vectors \mathbf{e}_n , we can rewrite (3.25),

$$\mathcal{V}(L) = \int_0^L \int_0^L \mathcal{C}_{\Sigma 1}(z_1, z_2) + \mathcal{C}_{\Sigma 2}(z_1, z_2) dz_1 dz_2 \quad (3.26)$$

where $\mathcal{C}_{\Sigma n}$ is the autocorrelation function (ACF) of n -th component of Ψ and thus the autocovariance function (ACOVF) of the n -th component of \mathbf{S}_Σ , with $n \in \{1, 2, 3\}$. It is defined as

$$\mathcal{C}_{\Sigma n}(z_1, z_2) = \left\langle \mathcal{E}[\Psi_n(z_1, t) \Psi_n(z_2, t)] \right\rangle \quad (3.27a)$$

which expands to

$$\begin{aligned} \mathcal{C}_{\Sigma n}(z_1, z_2) = & \exp[-\alpha(z_1 + z_2)] \tilde{\gamma}^2 \\ & \times \left\langle \mathcal{E} \left[\left(\mathcal{S}_{\Sigma n}(z_1, t) - \mathcal{E}[\mathcal{S}_{\Sigma n}(z_1, t)] \right) \cdot \left(\mathcal{S}_{\Sigma n}(z_2, t) - \mathcal{E}[\mathcal{S}_{\Sigma n}(z_2, t)] \right) \right] \right\rangle \end{aligned} \quad (3.27b)$$

Thus, \mathcal{V} is completely determined by the stochastic properties of the Stokes vector sum \mathbf{S}_Σ (or at least two of its Cartesian components). Also, the approximation of a constant $\hat{\mathbf{S}}_p$ we made above in order to be able to write (3.26) in closed form is only relevant as long as the ACOVFs $\mathcal{C}_{\Sigma 1}$ and $\mathcal{C}_{\Sigma 2}$ remain large, which, as we will see, is valid only over short distances, so that our error remains small. It is difficult to accurately estimate the magnitude of this error, but the comparisons to numerical simulations presented throughout this section confirm our approach.

In the case that the distribution of \mathbf{S}_Σ is isotropic (cf. the discussion on p. 75), the ACOVF cannot depend on the particular choice of components of \mathbf{S}_Σ , and we must have

$$\mathcal{C}_{\Sigma 1} = \mathcal{C}_{\Sigma 2} = \mathcal{C}_{\Sigma 3} = \frac{1}{3}\mathcal{C}_\Sigma \quad (3.28)$$

in which $\mathcal{C}_\Sigma = \mathcal{C}_{\Sigma 1} + \mathcal{C}_{\Sigma 2} + \mathcal{C}_{\Sigma 3}$ is the vector autocovariance of \mathbf{S}_Σ , defined as

$$\begin{aligned} \mathcal{C}_\Sigma(z_1, z_2) = \exp[-\alpha(z_1 + z_2)] \tilde{\gamma}^2 \\ \times \left\langle \mathcal{E}[\mathbf{S}_\Sigma(z_1, t) \cdot \mathbf{S}_\Sigma(z_2, t)] - \mathcal{E}[\mathbf{S}_\Sigma(z_1, t)] \cdot \mathcal{E}[\mathbf{S}_\Sigma(z_2, t)] \right\rangle \end{aligned} \quad (3.29)$$

When can rewrite (3.26) using (3.28) as

$$\mathcal{V}(L) = \frac{2}{3} \int_0^L \int_0^L \mathcal{C}_\Sigma(z_1, z_2) dz_1 dz_2 \quad (3.30)$$

Equation (3.30) will serve as defining expression throughout most of the remainder of the present chapter.

3.2.1 Autocovariance of the Stokes Sum

We now know how to determine \mathcal{V} in principle, but we do not yet have any expressions for the particular stochastic properties of the Stokes vector sum \mathbf{S}_Σ . Because our theory is limited to linear evolution of the signals in the interfering channels, only two effects are relevant: GVD (cf. Section 2.2.1) and PMD (cf. Section 2.2.3). We further limit ourselves to investigate only the *interchannel* effects of walk-off and the polarization evolution at the carrier frequency $\Delta\omega_\nu$ of channel ν , neglecting any pulse distortion and depolarization of the interfering channels.

Then we can write \mathbf{S}_Σ as a sum of vectors in the form

$$\mathbf{S}_\Sigma(z, t) = \sum_{\nu \neq \rho} P_\nu(t + \Delta t_\nu) \hat{\mathbf{S}}_\nu(z) \quad (3.31a)$$

where

$$P_\nu(t) = \langle S_\nu(z, t) | S_\nu(z, t) \rangle = |S_\nu(z, t)| \quad (3.31b)$$

$$\hat{S}_\nu(z) = \frac{S_\nu(z, t)}{P_\nu(z, t)} \quad (3.31c)$$

are the instantaneous optical power and the SOP of channel ν , and Δt_ν was defined in (2.20). Because we have neglected the evolution of distortions in the pulse shape due to GVD or PMD, P_ν is only a function of time-dependent data modulation and walk-off, as expressed by the quantity Δt_ν , describing the walk-off from the probe channel due to GVD (using $z_1 = 0$). Similarly, the SOP \hat{S}_ν depends only on the position z , since we have neglected any intra-pulse depolarization due to PMD and assume fully polarized input signals (the special case of polarization-modulated interfering channels is discussed separately). Hence, S_ν is a product of two processes that have different physical causes and must therefore be statistically independent.

When we insert the sum (3.31) into (3.29), we obtain an expression with $(N - 1)^2$ terms. However, all interchannel terms vanish because of the independent data modulation and polarization evolution in each WDM channel. We are then left with

$$\begin{aligned} \mathcal{C}_\Sigma(z_1, z_2) &= \sum_{\nu \neq \rho} \mathcal{C}_\nu(z_1, z_2) \\ &= \Gamma(z_1, z_2) \sum_{\nu \neq \rho} \mathcal{C}_\nu^{\text{WO}}(z_1, z_2) \cdot \mathcal{C}_\nu^{\text{SOP}}(z_1, z_2) \end{aligned} \quad (3.32a)$$

with

$$\mathcal{C}_\nu^{\text{WO}}(z_1, z_2) = \mathcal{E}[P_\nu(t) P_\nu(t + \Delta t_\nu)] - \mathcal{E}[P_\nu(t)]^2 \quad (3.32b)$$

$$\mathcal{C}_\nu^{\text{SOP}}(z_1, z_2) = \langle \hat{S}_\nu(z_1) \cdot \hat{S}_\nu(z_2) \rangle \quad (3.32c)$$

$$\Gamma(z_1, z_2) = \tilde{\gamma}^2 \exp \left[\int_0^{z_1} (g(\zeta) - \alpha(\zeta)) d\zeta + \int_0^{z_2} (g(\zeta) - \alpha(\zeta)) d\zeta \right] \quad (3.32d)$$

We have separated the ACOVFs into a sum over the interfering channels of functions of the walk-off, $\mathcal{C}_\nu^{\text{WO}}$, and of the PMD-induced SOP decorrelation, $\mathcal{C}_\nu^{\text{SOP}}$, and will presently take a closer look at each of them. The coefficient Γ incorporates the accumulated loss and gain at z_1 and z_2 and is the extension of the simple exponential expression for the nonlinearity coefficient that we have used before. In (3.32d), α is the local attenuation and g is local (amplifier) gain. Gain tilt can be included by making Γ channel-dependent.

Before we examine the various functions in (3.32) closer, however, we would like to verify some of the approximations we made in its derivation, especially the linear evolution of the interferers and the reduction of GVD and PMD to their interchannel effects by evaluating them only at the respective carrier

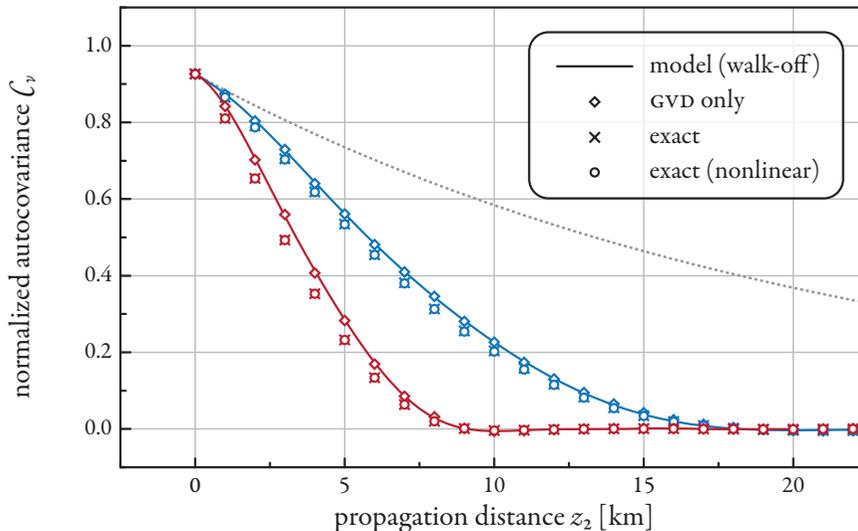


Figure 3.5 shows the single-channel autocovariance functions (normalized to mean-squared optical power) within a single amplifier span for various degrees of approximation for a system as described in Table 3.1: circles mark exact values for nonlinear propagation with 2 mW channel power and $D_{\text{PMD}} = 0.5 \text{ ps/km}^{1/2}$, crosses mark exact values for linear propagation, diamonds correspond to linear propagation without PMD, and solid lines correspond to the walk-off approximation $\mathcal{C}_\nu^{\text{WO}}$ and were obtained by time-shifting a PRBS. Blue color indicates values for interferers at $\Delta\omega_\nu = 2\pi 50 \text{ GHz}$, red color corresponds to $\Delta\omega_\nu = 2\pi 100 \text{ GHz}$. The dashed line shows the contribution of the attenuation term $\Gamma(0, z_2)$.

frequencies.

Fig. 3.5 shows the simulated ACOVFs $\mathcal{C}_\nu(0, z_2)$ of interferers ν within a single span. The ACOVFs are normalized to the mean-squared optical power $\mathcal{E}[P_\nu]^2$ of the interferers. Their initial value at $z_2 = 0$ (where $\Delta t_\nu = 0$) is not unity: $\mathcal{C}_\nu(0, 0) = \mathcal{C}_\nu^{\text{WO}}(0, 0)$ is a measure of the variance of the optical waveform, and for waveforms such as those shown in Fig. 2.2 the normalized ACOVF is $\mathcal{C}_\nu(0, 0)/\mathcal{E}[P_\nu]^2 \approx 0.92$.

As we can see in the figure, the difference between accounting for walk-off only and taking into account pulse distortions due GVD is negligible. A slightly different ACOVF is obtained when we additionally account for PMD – this deviation is described by the SOP decorrelation $\mathcal{C}_\nu^{\text{SOP}}$ (also see subsequent figure). The influence of the nonlinear propagation can be ignored within a single span at the power levels of interest.

Next, Fig. 3.6 shows values of the ACOVF at the beginning of successive amplifier spans, when the accumulated GVD in each span is fully compensated at the end of the span, resulting in a resonant dispersion map, i.e. $\mathcal{C}_\nu^{\text{WO}}(0, z_2) = \mathcal{C}_\nu^{\text{WO}}(0, 0)$ at each data point. In this case, the next most important contribution to the ACOVF comes from the SOP decorrelation term which statistically describes the evolution of the interfering channel SOPs due to PMD. This can be seen by a more or less rapid decrease in $\mathcal{C}_\nu(0, z_2)$, depending on the value of the fiber PMD coefficient. In the simulations without Kerr nonlinearities, there is no difference if only the mean SOP of the interfering channels is used or if

the depolarization and pulse distortion in these channels due to PMD is also taken into account, thus confirming our approach of using a constant envelope in the interfering channels when modeling signal propagation. When fiber nonlinearity is also considered, the ACOVF is reduced compared to the strictly linear case, especially when fiber PMD is very low. This is a result of interferer depolarization and the mean nonlinear SOP motion due to XPOLM, which is very pronounced in these particular systems. However, since in these cases the probe channel will also be significantly depolarized, the combination of these particular dispersion maps comprising low-PMD fiber and high interferer power levels is of little relevance for optical communication, where the nonlinear distortion generally remains small. We thus do not consider this difference a limitation of the model as long as we are merely interested in relatively small perturbations.

3.2.2 Walk-Off

The walk-off ACOVF is the dominating term in the single-span system, when neither PMD nor Kerr nonlinearities significantly affect the Stokes vectors of the interferers. Generally, $\mathcal{C}_\nu^{\text{WO}}$ can be very simply determined by using realistic pulse shapes $P_\nu(t)$ in (3.32b) and calculating $\mathcal{C}_\nu^{\text{WO}}$ for any Δt_ν , as has been done to obtain the «model» curve of Fig. 3.5. Such curves for various popular modulation formats are shown in Fig. 3.7.

This method also works for arbitrary input pulse shapes as those e.g. resulting from GVD pre-distortion / precompensation or OFDM modulation. When the ACOVF $\mathcal{C}_\nu^{\text{WO}}$ depends not only on the time shift Δt_ν accumulated between z_1 and z_2 , but on both coordinates themselves – as is the case for differential phase-shift keying (DPSK) signals which can acquire significant amplitude fluctuations as a result of GVD-mediated PM-AM conversion – a more accurate solution can be obtained by building a $\mathcal{C}_\nu^{\text{WO}}(z_1, z_2)$ matrix for discrete z_1 and z_2 values by performing a single linear simulation run over the dispersion map of interest, which would then be used in the numeric solution of the integral (3.30).

In certain idealized cases, there is also a fully analytic solution for the walk-off ACOVF. When assuming perfectly rectangular NRZ pulses, $\mathcal{C}_\nu^{\text{WO}}$ becomes a linear decreasing function of Δt_ν . When Δt_ν is equal to or larger than the symbol duration T_S , $\mathcal{C}_\nu^{\text{WO}}$ must be zero, because when the modulation is truly random, bit sequences that are shifted by at least one bit are statistically uncorrelated. This lack of correlation must also be reflected in the ACOVF. Using a single rectangular pulse in (3.32b) and imposing the boundary condition, we obtain

$$(NRZ) \quad \mathcal{C}_\nu^{\text{WO}}(z_1, z_2) = \mathcal{E}[P_\nu]^2 \begin{cases} 1 - \frac{|z_2 - z_1|}{L_{\text{WO}\nu}} & |z_2 - z_1| < L_{\text{WO}\nu} \\ 0 & |z_2 - z_1| \geq L_{\text{WO}\nu} \end{cases} \quad (3.33)$$

in which $L_{\text{WO}\nu}$ is the characteristic walk-off length of channel ν , which was defined in (2.21) and which equals the length Δz at which $\Delta t_\nu = T_S$. By using L_{WO} , we can formulate the $\mathcal{C}_\nu^{\text{WO}}$ independently of frequency separation and symbol rate.

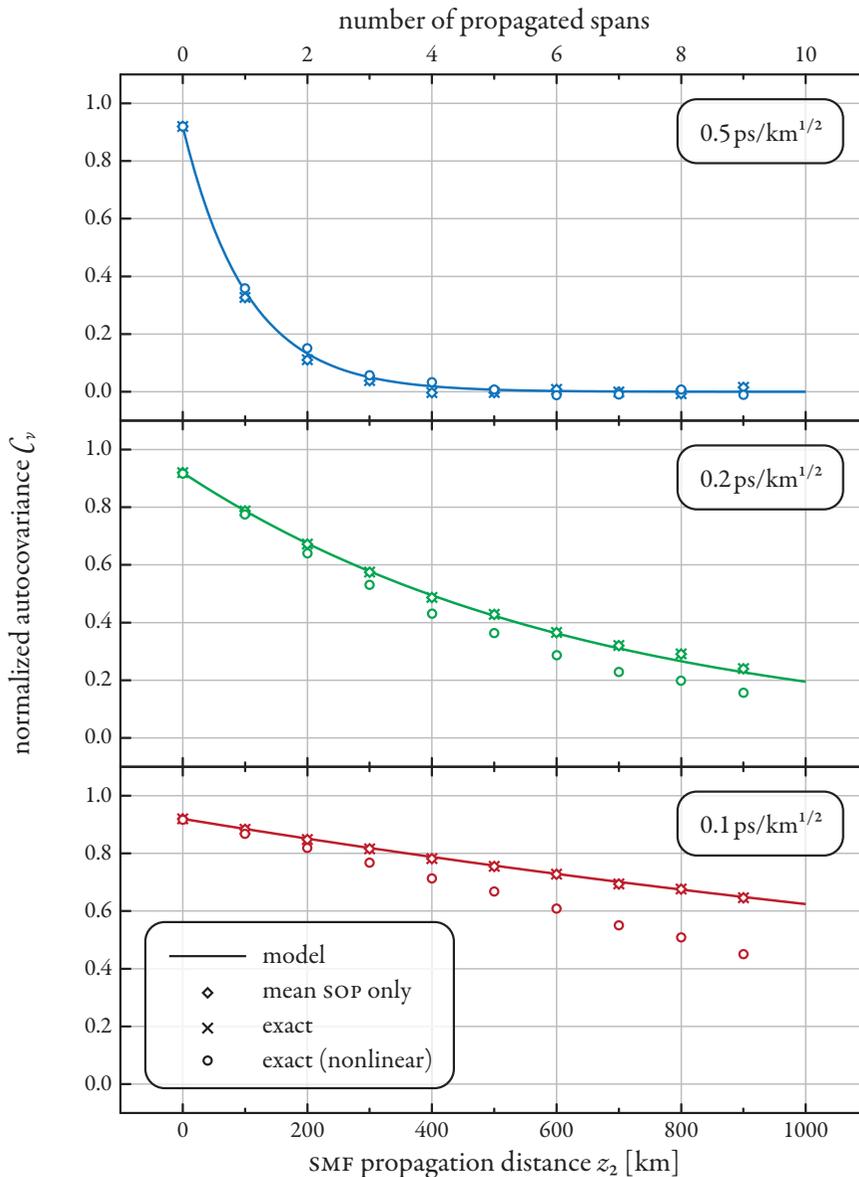


Figure 3.6 shows the single-channel ($\Delta\omega_y = 2\pi 50 \text{ GHz}$) autocovariance functions (normalized to mean-squared optical power) at the beginning of subsequent amplifier spans for various degrees of approximation: circles mark exact values for nonlinear propagation with 2 mW channel power, crosses mark exact values for linear propagation, diamonds correspond to linear propagation when calculating C_y^{SOP} with the mean channel SOPs only, and solid lines correspond to the analytic solution derived on page 96. Ensemble size was 1000 for linear data and 250 for nonlinear data.

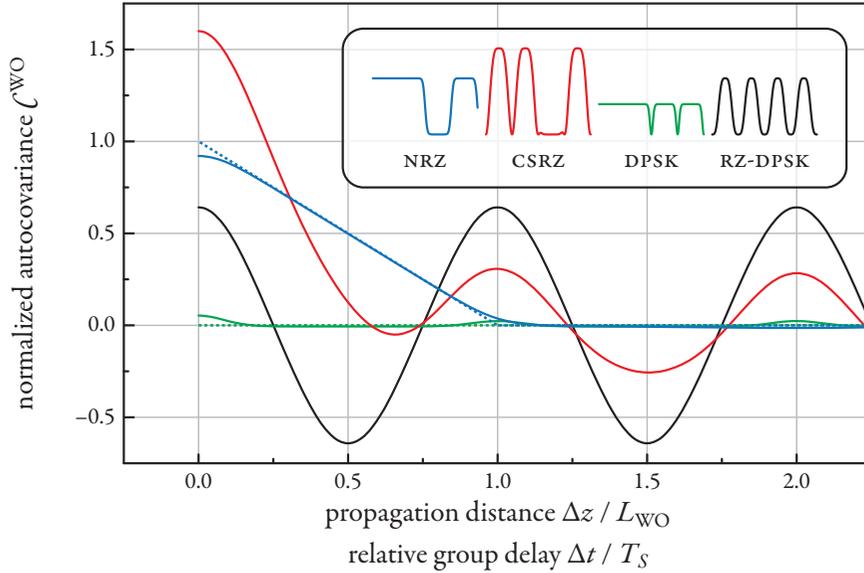


Figure 3.7 shows the walk-off autocovariance functions (normalized to mean-squared optical power) within the first two walk-off lengths for various popular modulation formats, obtained numerically from the pulse shapes $P(t)$. The dotted lines show the approximate functions (3.33) and (3.34) for idealized pulse shapes.

Similarly, for the idealized DPSK signal, which has no amplitude modulation component, we can immediately determine

$$\text{(DPSK)} \quad C_v^{\text{WO}}(z_1, z_2) = 0 \quad (3.34)$$

While realistic DPSK channels may have some residual amplitude modulation, we can conclude that DPSK-modulated interferers cause significantly less XPOLM than their NRZ-modulated counterparts. This, however, is only valid as long as the accumulated GVD remains small and no significant PM-AM occurs. In contrast, the ACOVF C_v^{WO} of RZ-DPSK-modulated signals shows periodic oscillations; however, the integral over the ACOVF equals approximately zero after each walk-off length. Thus, additional RZ modulation is not expected to significantly increase XPOLM, even though there is considerable amplitude variation present in the signal – since that modulation is the same for each symbol, there is no contribution to the variance associated with XPOLM. The carrier-suppressed return-to-zero (CSRZ) signal is a combination of amplitude modulation and RZ pulse shape. Hence, its walk-off ACOVF is a similar combination of the ACOVFs of NRZ and RZ-DPSK modulation: a periodic oscillation is superposed onto a decreasing function; after a walk-off length only the oscillation remains. The integral over this function will always be larger than zero, and thus CSRZ – like all amplitude-modulated formats – is expected to also cause significant amounts of XPOLM.

Both analytic approximations (3.33) and (3.34) are also shown in Fig. 3.7 as dotted lines; they

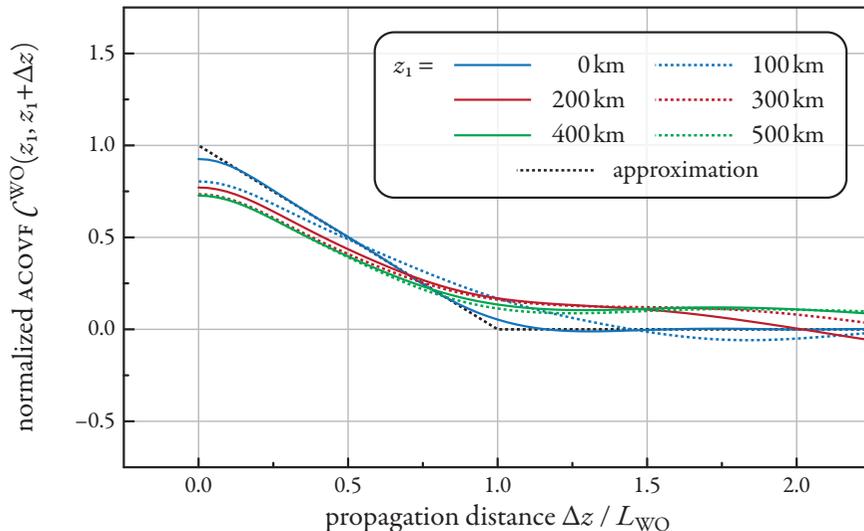


Figure 3.8 shows the accurate walk-off autocovariance functions (normalized to mean-squared optical power) within the first two walk-off lengths for NRZ pulses and various values of accumulated initial GVD, obtained numerically by propagating the signal over a linear SSMF with $\beta_2 = -256$ fs/GHz/km. The approximate function (3.33) is also indicated in the figure.

closely match their counterparts taking into account realistic pulse shapes. Because of the simple analytic expression for the walk-off ACOVF of NRZ-modulated interferers, we use (3.33) exclusively whenever numerically solving the integral equation (3.30) for the XPOLM variance parameter \mathcal{V} .

When the dispersion map is characterized by large amounts of RDPS distributed throughout the transmission link, the pulses in the interfering channels can acquire significant distortion (cf. Figs. 2.2 and 2.3). The NRZ approximation (3.33) may thus no longer hold and we may have to use the ACOVF matrix method, described above, to calculate the nonlinear depolarization accurately. To determine the necessary method, we calculated the appropriate ACOVF matrix \mathcal{C}_v^{WO} for specific (z_1, z_2) pairs in an SSMF without inline dispersion compensation. It is illustrated in Fig. 3.8. Even though the shape of \mathcal{C}_v^{WO} depends on z_1 , the differences remain small. Due to pulse broadening the initial value $\mathcal{C}_v^{WO}(z_1, z_1)$ of the ACOVF decreases, but it shows residual correlation over longer distances. The integral over the function is approximately equal, independent of z_1 . Since the largest value of RDPS used in the present work (-2048 fs/GHz) corresponds to a maximum accumulated GVD value equivalent to ≈ 260 km of uncompensated SSMF in our 10-span link, we are convinced that the simple relation (3.33) is sufficient to describe the ACOVF related to walk-off.

POLARIZATION-MULTIPLEXED INTERFERING CHANNELS When the interfering DWDM channels consist of two POLDM subchannels or, more generally, are polarization-modulated in some way, the separation of the Stokes vector \mathbf{S}_v into the product of its length P_v and SOP $\hat{\mathbf{S}}_v$, as in (3.31a) is not suf-

ficient for an accurate description of their behavior. Particularly, the effect of walk-off on such a signal cannot be described by the time function of its power alone. Instead, the full Stokes vector must be used to determine $\mathcal{C}_\nu^{\text{WO}}$.

Because the physical causes for the walk-off (GVD) and the SOP decorrelation (PMD) remain uncorrelated, the separation of the ACOVF into the product of $\mathcal{C}_\nu^{\text{WO}}$ and $\mathcal{C}_\nu^{\text{SOP}}$ is still valid. However, the walk-off ACOVF now becomes

$$\text{(POLDM)} \quad \mathcal{C}_\nu^{\text{WO}}(z_1, z_2) = \mathcal{E}[\mathbf{S}_\nu(t) \cdot \mathbf{S}_\nu(t + \Delta t_\nu)] - \mathcal{E}[\mathbf{S}_\nu(t)]^2 \quad (3.35)$$

where Δt_ν is again the walk-off accumulated between z_1 and z_2 . We can use (3.35) to numerically determine the ACOVF analogous to Fig. 3.7. This has been done in Fig. 3.9 for various versions of RZ-shaped quaternary phase-shift keying (QPSK) symbols. Polarization-multiplexing QPSK signals is an attractive option for implementation of 100 Gbit Ethernet – the resulting symbol rate of ≈ 25 Gbaud can be transmitted within the 50 GHz ITU grid upon which most deployed long-haul systems are based, so that no significant changes in the hardware within the link are necessary, but only an upgrade of the transmitters and receivers. In this context, Xie has shown that when the interfering channels are also polarization-multiplexed, interleaving both subchannels by half a symbol length can improve transmission fidelity [Xie 2009]. The reason for this improvement is easily understood from our model: because of the RZ pulse shape, each half symbol slot alternately carries data from one POLDM subchannel, so that the SOP oscillates between opposite positions on the Poincaré sphere. The resulting nonlinear polarization rotations approximately cancel each other. In Fig. 3.9 this shows as an oscillating walk-off ACOVF whose integral is almost zero. In fact, the ACOVF is almost the same as that of single-polarization PSK, the difference being that in the latter there is an average movement as described in (3.2) which is absent in the interleaved POLDM signal. However, as we have argued, this average movement is not relevant for the detrimental effects of XPOLM. Xie also notes that this only works in dispersion-managed systems, i.e. within the scope of the walk-off approximation. Without inline GVD compensation, pulses disperse quickly and extend far beyond their time slots – the RZ shape, on which the interleaving improvement is based, is lost.

Figure 3.9 also shows the walk-off ACOVF $\mathcal{C}_\nu^{\text{WO}}$ for aligned POLDM subchannels. Because the interferer SOP in this configuration becomes again data-dependent – the time-dependent phase difference between the subchannels (equivalent to the polarimetric parameter ϕ) determines the SOP $\hat{\mathbf{S}}_\nu(t) - \mathcal{C}_\nu^{\text{WO}}$ is more similar to the ACOVF for NRZ than that of RZ-PSK – it lacks oscillations and becomes approximately zero at $\Delta t = T_S/2$ (since the RZ pulses are shorter). Its integral – and thus the relative magnitude of XPOLM caused by it – is comparable to that of NRZ. This point should be stressed as it might be unsuspected: POLDM signals, when launched with aligned subchannels or after GVD-related pulse distortions have negated the benefit of interleaving, generate similar amounts of XPOLM as NRZ channels – generally considered the worst-case for interchannel nonlinear distortion – at the same symbol rate.

Independent of the modulation format used within the interfering channels, we will show in Chapter 5 that subchannel interleaving within a POLDM probe can lead to a significant reduction of XPOLM-

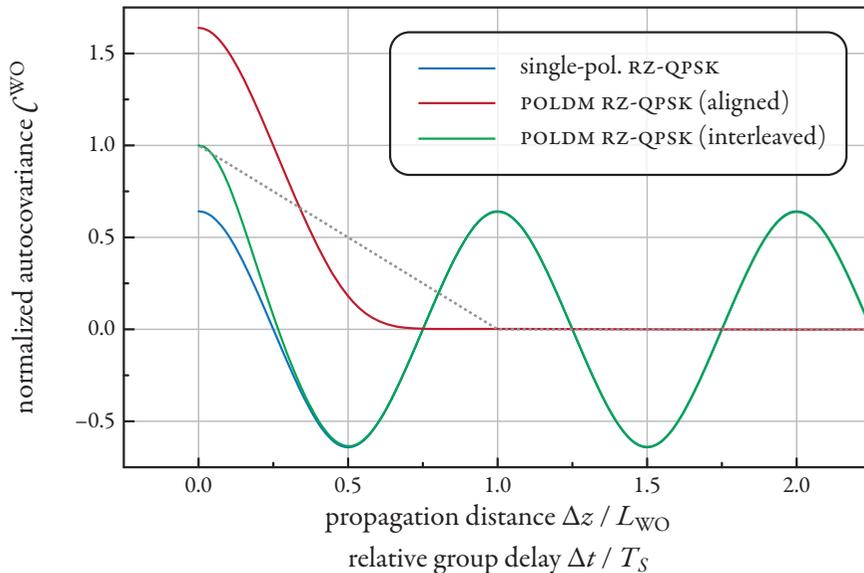


Figure 3.9 compares the walk-off autocovariance functions (normalized to mean-squared optical power) within the first two walk-off lengths for single-polarization and POLDM RZ-QPSK, obtained numerically from the Stokes vectors $\mathbf{S}(t)$, analogously to Fig. 3.7. POLDM subchannel interleaving and alignment are discussed on p. 47. The dashed line shows the approximate NRZ function (3.33) for reference.

related crosstalk. This benefit also does not depend on the particular dispersion map used in the transmission link.

3.2.3 Mean SOP Decorrelation

Whenever C_v^{WO} is greater than zero for some pair of z_1 and z_2 – i.e. while the corresponding signals are still correlated – the evolution of the interferer SOP, described by C_v^{SOP} in (3.32), becomes relevant when determining \mathcal{V} or the DOP. We have mentioned before that the principal reason for changes in the SOP is fiber PMD, and we have explained its mechanisms in Section 2.2.3. If we describe the evolution of the SOP of an optical channel ν by the SOP of its carrier $\hat{\mathbf{S}}_\nu$, at $\Delta\omega_\nu$, we have

$$\hat{\mathbf{S}}_\nu(z) = \frac{\tilde{\mathbf{S}}(z, \Delta\omega_\nu)}{|\tilde{\mathbf{S}}(z, \Delta\omega_\nu)|} \quad (3.36)$$

where $\tilde{\mathbf{S}}$ is again the Fourier transform of \mathbf{S} . To predict the evolution of $\hat{\mathbf{S}}_\nu$ in a statistical sense, we again model the fiber as a concatenation of a large number of independent segments of locally constant

birefringence, and rewrite (2.75) for the SOP evolution in each segment as

$$\partial_z \hat{\mathbf{S}}_v(z) = \Delta\beta_1 \Delta\omega_v \hat{\mathbf{\Omega}} \times \hat{\mathbf{S}}_v(z) = \Delta\omega_v \partial_z \mathbf{\Omega} \times \hat{\mathbf{S}}_v(z) \quad (3.37)$$

where $\Delta\beta_1$ and $\hat{\mathbf{\Omega}}$ are assumed constant within the segment, and we have used (2.92) to write the second equality. The evolution of $\hat{\mathbf{S}}_v$ is thus a rotation around the local PMD vector increment, whose direction is assumed to be distributed uniformly random on the Poincaré sphere.

We can thus describe the evolution of $\hat{\mathbf{S}}_v$ as a sequence of random rotations. Then we can apply the theory of the preceding section also to this evolution. A difference is that there is no dependence of PMD on time, so that any time average is replaced by ensemble averages. Hence, if the SOP $\hat{\mathbf{S}}_v$ at some location z_1 is known, we know its ensemble distribution at z_2 to be the Brownian with a variance parameter $\mathcal{V}_v^{\text{SOP}}$ (to separate it from \mathcal{V} which describes XPOLM). Because the rotation axis is uniformly distributed, we can apply the DOP definition (3.12) and expression (3.15) relating the DOP and \mathcal{V} to write

$$\langle \hat{\mathbf{S}}_v(z_1) \cdot \hat{\mathbf{S}}_v(z_2) \rangle = \exp\left(-\frac{\mathcal{V}_v^{\text{SOP}}(z_1, z_2)}{2}\right) \quad (3.38)$$

The left-hand side of the above equation is the definition of $\mathcal{C}_v^{\text{SOP}}(z_1, z_2)$ which we are seeking. We thus only need to determine $\mathcal{V}_v^{\text{SOP}}$ according to the integral (3.30) where we replace the ACOVF \mathcal{C}_Σ of the sum of interferer Stokes vectors with an expression \mathcal{C}^Ω for the ACOVF of the PMD vector increment $\partial_z \mathbf{\Omega}$, such that

$$\mathcal{V}_v^{\text{SOP}}(z_1, z_2) = \frac{2}{3} \int_{z_1}^{z_2} \int_{z_1}^{z_2} \Delta\omega_v^2 \mathcal{C}^\Omega(\zeta_1, \zeta_2) d\zeta_1 d\zeta_2 \quad (3.39)$$

in which both integrals run from z_1 to z_2 and where \mathcal{C}^Ω is defined as

$$\mathcal{C}^\Omega(\zeta_1, \zeta_2) = \langle \partial_z \mathbf{\Omega}(\zeta_1) \cdot \partial_z \mathbf{\Omega}(\zeta_2) \rangle \quad (3.40)$$

The Green-Kubo relation connects the autocorrelation function of the increments $\partial_z \mathbf{\Omega}$ to the mean square of the length of $\mathbf{\Omega}$ for a Brownian motion / Wiener process [Green 1954, Kubo 1957] (also cf. [Evans 2008, Sec. 4.1]),

$$\int_{z_1}^{z_2} \int_{z_1}^{z_2} \mathcal{C}^\Omega(\zeta_1, \zeta_2) d\zeta_1 d\zeta_2 = \langle \Delta\tau(z_1, z_2)^2 \rangle \quad (3.41a)$$

where

$$\Delta\tau(z_1, z_2) = |\mathbf{\Omega}(z_2) - \mathbf{\Omega}(z_1)| \quad (3.41b)$$

is the length of the PMD vector of the fiber segment between z_1 and z_2 . We further have from (2.94)

$$\begin{aligned} \langle \Delta\tau(z_1, z_2)^2 \rangle &= \frac{3\pi}{8} \langle \Delta\tau(z_1, z_2) \rangle^2 \\ &= \frac{3\pi}{8} D_{\text{PMD}}^2 |z_2 - z_1| \end{aligned} \quad (3.42)$$

in which D_{PMD} is again the fiber PMD parameter. Defining a channel-dependent variable

$$L_{\text{SOP}\nu} \equiv \frac{8}{\pi\Delta\omega_\nu^2 D_{\text{PMD}}^2} \quad (3.43)$$

the SOP variance parameter $\mathcal{V}_\nu^{\text{SOP}}$ of channel ν can be concisely expressed using (3.40) – (3.43) as

$$\mathcal{V}_\nu^{\text{SOP}}(z_1, z_2) = 2 \frac{|z_2 - z_1|}{L_{\text{SOP}\nu}} \quad (3.44)$$

and we can obtain the ACOVF $\mathcal{C}_\nu^{\text{SOP}}$ (3.32c) for the SOP decorrelation with the help of (3.38) as

$$\mathcal{C}_\nu^{\text{SOP}}(z_1, z_2) = \exp\left(-\frac{|z_2 - z_1|}{L_{\text{SOP}\nu}}\right) \quad (3.45)$$

The variable $L_{\text{SOP}\nu}$ describes the decay length of the exponential function quantifying the stochastic SOP correlation – we shall therefore call it the SOP decorrelation length of channel ν . The larger the frequency separation $\Delta\omega_\nu$ between the probe and an interferer ν , the faster the SOP of that interferer decorrelates in a coordinate system that is defined by the birefringence at the probe carrier frequency (cf. Section 2.2.2).

The above equation was used to draw the analytical curves in Fig. 3.6, matching the linear simulation data very well, thus confirming that the intra-pulse depolarization due to PMD is not relevant. Thus, using (3.45) in (3.32) should give a sufficiently good approximation for the polarization part of the Stokes vector ACOVF, which can then be used to estimate the XPOLM-related \mathcal{U} parameter and the DOP reduction for arbitrary fiber systems.

COMPARISON WITH SIMULATIONS To thoroughly test the theory developed in the present section, we compared its predictions to a number of numerical simulations for a variety of parameter combinations. Fig. 3.10 shows the evolution of the nonlinear depolarization for different fiber PMD and GVD values in systems with a resonant dispersion map. As we would expect from the scaling of $\mathcal{C}_\nu^{\text{SOP}}$ with the SOP decorrelation length $L_{\text{SOP}\nu}$, seen in Fig. 3.6, systems with high PMD fiber are less affected by XPOLM due to the fast decorrelation of the interferer SOPs, while the DOP of systems with very low PMD fiber can be significantly more deteriorated. Similarly, we have shown that $\mathcal{C}_\nu^{\text{WO}}$ scales with the walk-off length, and thus systems with large L_{WO} (small β_2) should be more affected, since the integrands \mathcal{C}_ν will be non-zero over larger distances. This behavior is also confirmed in Fig. 3.10, with a very good qualitative

and quantitative match between numerical simulations and the model prediction.

The differences between model prediction and simulation results are primarily due to our assumption of a linear interferer evolution, which causes the SOP decorrelation as described by C_{ν}^{SOP} to be overestimated when fiber PMD is very low (cf. Fig. 3.6). For the low-GVD fiber (bottom graph), additional depolarization as a result of FWM becomes non-negligible. We have seen in Fig. 2.13 that FWM can be much more efficient in low-GVD fiber. The FWM products in the triple sum (2.105) will have the same SOP as channel ν which will generally be different from the SOP of the probe ρ , so that they can be considered as unpolarized noise, which leads to an additional, nonlinear but not XPOLM-related DOP reduction in the simulation results. However, when fiber PMD is very low, the error due to the assumption of linear interferer evolution dominates the influence of FWM; at an intermediate PMD value, both errors nearly cancel each other.

The expressions (3.32b) and (3.32d) predict a quadratic dependence of \mathcal{U} on the optical power in the interfering channel and the fiber nonlinearity parameter $\tilde{\gamma}$. Fig. 3.11 compares the nonlinear DOP reduction vs. the optical power of the interfering channels obtained in numerical simulations to these predictions – the match is very good as long as the approximations hold sufficiently well (i.e. the nonlinear distortions within the interferers can be neglected). This power scaling law has already been implicitly used in Figs. 3.5 and 3.6, in which the curves were normalized by $\mathcal{E}[P_{\nu}]^2$ to make them independent of power. This also means that we can – within the validity range of our model – simply scale the mean nonlinear depolarization \mathcal{D}_0 , determined at some interferer power $\tilde{\gamma}_0 \mathcal{E}[P_0]$ for any fiber parameter combination, to any other nonlinearity parameter or power value $\tilde{\gamma} \mathcal{E}[P]$ simply by using

$$\mathcal{D} = \exp\left(-\frac{\mathcal{U}_0}{2} \cdot \xi\right) = (\mathcal{D}_0)^{\xi} \quad \text{with} \quad \xi = \left(\frac{\tilde{\gamma} \mathcal{E}[P]}{\tilde{\gamma}_0 \mathcal{E}[P_0]}\right)^2 \quad (3.46)$$

where we assumed that $\mathcal{E}[P_{\nu}] = \mathcal{E}[P]$ for all ν , i.e. all interfering channels carry equal optical power.

3.2.4 Co-Polarized & Mutually Orthogonal Launch

We have mentioned before (cf. Section 3.1) that the validity of our model not only extends to launching all channels with uniformly random polarization states, but also to configurations in which channels are launched co-polarized or when neighboring channels initially have orthogonal polarizations, so-called polarization-interleaving. It has been shown that in particular the latter case can improve transmission by suppressing FWM and XPOLM, if fiber PMD is small enough that the orthogonality is maintained sufficiently long [Hansryd 2000, Borne 2005].

From the standpoint of XPOLM there is no difference between co-polarized and mutually orthogonal SOPs since the evolution equation (2.109) is given in Stokes space in which both are anti-parallel, i.e. the rotation axis will be the same and only the direction of rotation will be inverted. Since we care

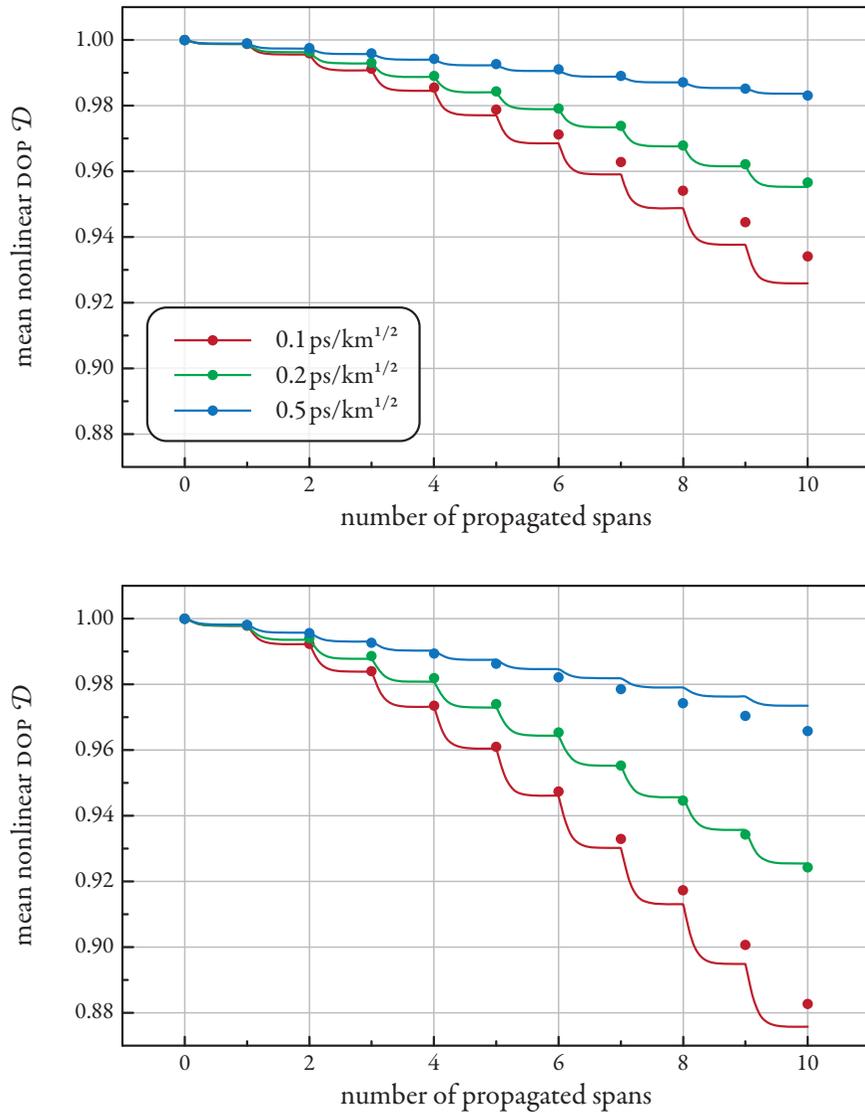


Figure 3.10 shows the evolution of the DOP with propagation over 10 spans of SMF with -128 fs/GHz/km (top) and -64 fs/GHz/km (bottom) with varying amounts of fiber PMD. Symbols denote simulation ensemble mean values and lines denote results obtained using the XPOLM model. The dispersion map was resonant in both cases.

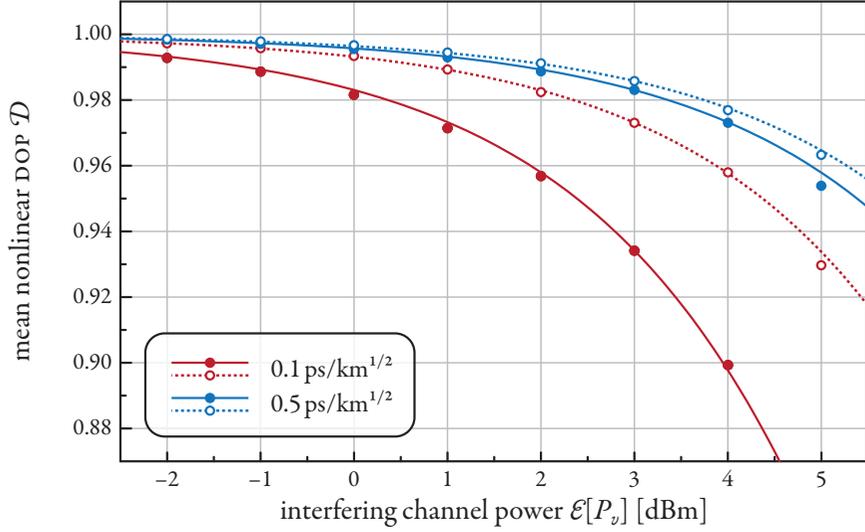


Figure 3.11 shows the dependence of the nonlinear DOP reduction on optical power in the interfering channels after propagation over 10 spans of SSMF with -128 fs/GHz/km for several configurations with a resonant dispersion map (solid lines) and RDPS of -2048 fs/GHz (dashed lines). Symbols denote simulation ensemble mean values, the curves were generated from (3.46) using the DOPs for $P_0 = 3$ dBm.

about these polarization rotations only in a statistical sense, this direction does not matter. We thus use the terms co-polarized and mutually orthogonal interchangeably.

We reiterate that the coefficient $2/3$ in (3.30) was only derived for uniformly distributed interferer SOPs. When these SOPs are distributed differently, as in the current case of interest, this coefficient becomes dependent on that distribution, which in turn will depend on the particular location within the fiber – only when PMD has randomized all channel SOPs sufficiently for them to be regarded isotropic, this coefficient will approach its asymptotic value of $2/3$. The required propagation length can be estimated from (3.43) by demanding about 2-3 SOP decorrelation lengths for each interferer, as we will show. When either z_1 or z_2 are smaller, we must adapt the expression used to determine the variance parameter \mathcal{V} .

The starting point in this case is the integral form (3.26) which shows that we are interested only in those components of the ACOVF of the interferer Stokes sum which lie in the plane orthogonal to the probe. With our definition of Γ in (3.32d) and $\mathcal{C}_v^{\text{WO}}$ in (3.32) we can rewrite (3.26) as

$$\mathcal{V}(L) = \iint_0^L \Gamma(\zeta_1, \zeta_2) \sum_{v \neq p} \mathcal{C}_v^{\text{WO}}(\zeta_1, \zeta_2) \langle \mathbf{s}_{v\perp}(\zeta_1) \cdot \mathbf{s}_{v\perp}(\zeta_2) \rangle d\zeta_1 d\zeta_2 \quad (3.47a)$$

with the vector

$$\mathbf{s}_{v\perp}(\zeta) = \hat{\mathbf{S}}_v(\zeta) - (\hat{\mathbf{S}}_v(\zeta) \cdot \hat{\mathbf{S}}_p) \hat{\mathbf{S}}_p \quad (3.47b)$$

which is the projection of $\hat{\mathbf{S}}_v$ on the plane orthogonal to $\hat{\mathbf{S}}_P$. As before we approximate $\hat{\mathbf{S}}_P$ to be constant, so that its orthogonal plane remains fixed.

For an analysis of the ensemble average in (3.47a), we shall assume that $\zeta_2 > \zeta_1$ – otherwise the indices in all subsequent expressions must be exchanged. We perform this ensemble average over the SOP populations at ζ_1 and ζ_2 in a two-step approach: we first average over $\hat{\mathbf{S}}_v(\zeta_2)$, keeping $\hat{\mathbf{S}}_v(\zeta_1)$ constant. As we will see, this allows us to write $\hat{\mathbf{S}}_v(\zeta_2)$ in terms of $\hat{\mathbf{S}}_v(\zeta_1)$. We subsequently average over all $\hat{\mathbf{S}}_v(\zeta_1)$. We denote the particular average by an index to the ensemble average operator, as e.g. in $\langle \cdot \rangle_2$. The population distribution of $\hat{\mathbf{S}}_v$ at $z = \zeta_1$, initially identical to $\hat{\mathbf{S}}_P$ at $z = 0$, will have evolved – keeping with our previous approximations – as a result of PMD only, and thus we know it to be the Brownian distribution, centered on $\hat{\mathbf{S}}_P$, with a variance parameter $\mathcal{U}_v^{\text{SOP}}(0, \zeta_1)$, as defined in (3.44). Similarly, the population distribution at ζ_2 will be Brownian, centered on $\hat{\mathbf{S}}_v(\zeta_1)$.

We then have, by using (3.47b),

$$\langle \mathbf{s}_{v\perp}(\zeta_1) \cdot \mathbf{s}_{v\perp}(\zeta_2) \rangle_2 = \mathbf{s}_{v\perp}(\zeta_1) \cdot \left[\langle \hat{\mathbf{S}}_v(\zeta_2) \rangle_2 - \left(\langle \hat{\mathbf{S}}_v(\zeta_2) \rangle_2 \cdot \hat{\mathbf{S}}_P \right) \hat{\mathbf{S}}_P \right] \quad (3.48)$$

In (3.32c) we defined an ACOVF related to the PMD-induced SOP decorrelation, with which we can write

$$\langle \hat{\mathbf{S}}_v(\zeta_2) \rangle = \mathcal{C}_v^{\text{SOP}}(\zeta_1, \zeta_2) \hat{\mathbf{S}}_v(\zeta_1) \quad (3.49)$$

for constant $\hat{\mathbf{S}}_v(\zeta_1)$. It was quantified in (3.45). With (3.49) we can rewrite (3.48) as

$$\langle \mathbf{s}_{v\perp}(\zeta_1) \cdot \mathbf{s}_{v\perp}(\zeta_2) \rangle_2 = \mathbf{s}_{v\perp}^2(\zeta_1) \mathcal{C}_v^{\text{SOP}}(\zeta_1, \zeta_2) \quad (3.50)$$

Performing the average over the polarization states at ζ_1 , we have

$$\langle \mathbf{s}_{v\perp}(\zeta_1) \cdot \mathbf{s}_{v\perp}(\zeta_2) \rangle = \mathcal{K}_v(\zeta_1) \mathcal{C}_v^{\text{SOP}}(\zeta_1, \zeta_2) \quad (3.51a)$$

with a scalar divergence coefficient

$$\mathcal{K}_v(\zeta_1) = \langle \mathbf{s}_{v\perp}^2(\zeta_1) \rangle \quad (3.51b)$$

which only depends on the statistical properties of the SOPs at ζ_1 , the smaller of the involved location variables, and $\langle \cdot \rangle = \langle \langle \cdot \rangle_2 \rangle_1$. When approximating the Brownian SOP distribution with the Fisher distribution (3.16), there exists a closed-form solution for \mathcal{K}_v ,

$$\mathcal{K}_v(\zeta_1) = \frac{2 \coth(\varkappa)}{\varkappa} - \frac{2}{\varkappa^2} \quad (3.52)$$

where \varkappa is related to $\mathcal{U}_v^{\text{SOP}}(0, \zeta_1)$ by (3.17).

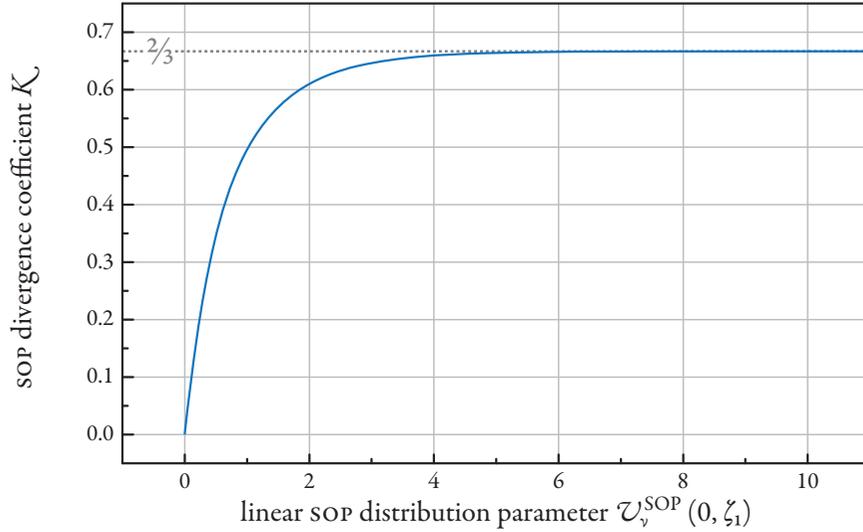


Figure 3.12 plots the dependence of the divergence coefficient \mathcal{K}_v on $\mathcal{U}_v^{\text{SOP}}$. The value of $\mathcal{U}_v^{\text{SOP}}$, depending on fiber location and SOP decorrelation length, can be obtained from (3.44).

We can now rewrite (3.47) in concise form as

$$\mathcal{V}(L) = \int_0^L \int \sum_{v \neq \rho} \mathcal{K}_v(\zeta_1) \mathcal{C}_v(\zeta_1, \zeta_2) d\zeta_1 d\zeta_2 \quad (3.53)$$

with \mathcal{C}_v defined in (3.32). When comparing this to the expression for uniformly distributed SOPs, (3.30), we see that the channel- and position dependent coefficient \mathcal{K}_v has replaced the constant $2/3$, but nothing else has changed.

The dependence of \mathcal{K}_v on $\mathcal{U}_v^{\text{SOP}}$ is shown in Fig. 3.12. We see that after ≈ 2 SOP decorrelation lengths (corresponding to $\mathcal{U}_v^{\text{SOP}} \approx 4$), \mathcal{K}_v is well converged to its asymptotic value of $2/3$. The evolution of the mean DOP as predicted by (3.53) is shown in Fig. 3.13, together with simulation data for co-polarized and polarization-interleaved (neighboring WDM channels are mutually orthogonal) launch SOPs.

As we can see from the figure, the model captures the system behavior very well. Both lower-PMD fibers perform very similar after about 7 spans. The lowest-PMD fiber initially performs best, but will deteriorate after most interferers have surpassed their SOP decorrelation lengths, and the divergence coefficient approaches $2/3$ as in Fig. 3.10. If propagated over more spans it would eventually be significantly outperformed by the higher-PMD fibers as in Fig. 3.10. The model even predicts the crossing of the curves for 0.1 and $0.2 \text{ ps/km}^{1/2}$ at the correct distance. When neighboring channels were launched with mutually orthogonal SOPs, a noticeable difference to co-polarized launch can only be observed for the fiber with lowest PMD. In such a fiber the SOPs of spectrally near channels remain almost constant over

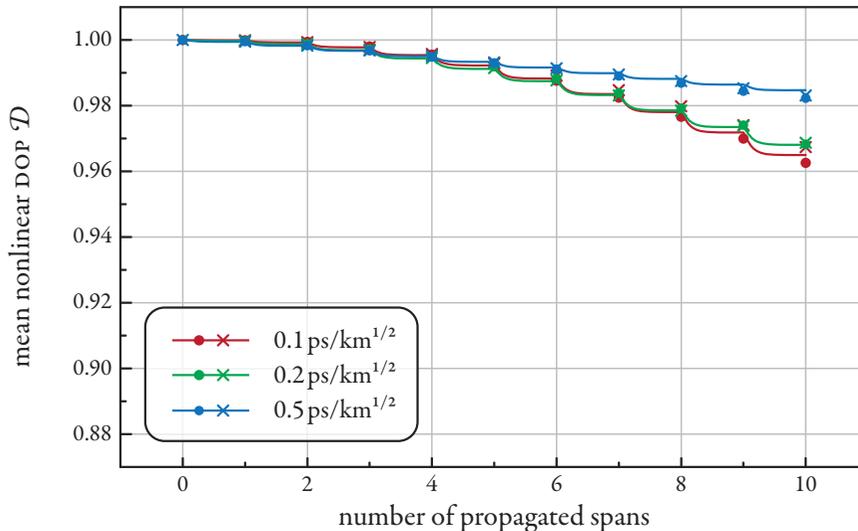


Figure 3.13 shows the evolution of the DOP with propagation over 10 spans of SSMF with $\beta_2 = -128$ fs/GHz/km with varying amounts of fiber PMD. Symbols denote simulation ensemble mean values and lines denote results obtained using (3.53). The dispersion map was in both cases resonant (no RDPS); all channels were initially co-polarized (circles) or polarization-interleaved (crosses).

many spans. Therefore, the FWM products involving initially orthogonal channels are well suppressed throughout. This does not occur when PMD is not negligible and the relative SOPs of neighboring channels are randomized on short length scales, so that for high-PMD fiber there is little difference between co-polarized launch and polarization interleaving.

Overall we feel that the model developed and quantified in the previous sections works very well for a variety of systems and can greatly help in understanding the dependence of XPOLM on the parameters of these systems, and in estimating the impact and various means apart from polarization interleaving of reducing XPOLM for a variety of fiber parameters. This will be done in Section 3.3.

3.2.5 Amplifier Noise

Before we investigate broadly how XPOLM is affected by the various system properties, we will interpose a brief discussion of how the accumulation of optical amplifier noise in the interfering channels affects the dynamics of XPOLM. Until now we have completely neglected the effect of the additive noise imposed by in-line optical amplifiers onto the interferers and implicitly assumed its impact to be negligible. When including such noise in the discussion, we can describe it mathematically in Jones space by

$$|S'(z, t)\rangle = |S(z, t)\rangle + |N(z, t)\rangle \quad (3.54)$$

where $|\mathbf{N}\rangle = (N_1, N_2)^T$ describes complex Gaussian noise in both polarizations with instantaneous optical power

$$P_N(\mathbf{z}, t) = \langle \mathbf{N}(\mathbf{z}, t) | \mathbf{N}(\mathbf{z}, t) \rangle \quad (3.55)$$

distributed equally over both (independent) polarization components. The noise $|\mathbf{N}\rangle$ is unpolarized. In Stokes space, we can model the noise as an additive Stokes vector $\Delta\mathbf{S}_N$ such that

$$\mathbf{S}'(\mathbf{z}, t) = \mathbf{S}(\mathbf{z}, t) + \Delta\mathbf{S}_N(\mathbf{z}, t) \quad (3.56)$$

However, due to the nonlinear relationship between Jones and Stokes vectors, $\Delta\mathbf{S}_N$ is neither Gaussian nor isotropic and has a non-zero mean – it must, however, be symmetric about \mathbf{S} . Its distribution not only depends on $|\mathbf{N}\rangle$, but also on $|\mathbf{S}\rangle$ via the definition of the (instantaneous) Stokes vector components (1.35). Since these combine the Jones vectors in an essential nonlinear way, the exact distribution of $\Delta\mathbf{S}_N$ will likely have to be determined using Monte Carlo methods – however, we do not require the knowledge of the noise distribution in Stokes space to proceed.

To examine the influence of noise on our results, we will use the definition (3.29) of \mathcal{C}_Σ . When including the noise term $\Delta\mathbf{S}_N$, we have

$$\begin{aligned} \mathcal{E}[(\mathbf{S}_\Sigma(z_1, t) + \Delta\mathbf{S}_N(z_1, t)) \cdot (\mathbf{S}_\Sigma(z_2, t) + \Delta\mathbf{S}_N(z_2, t))] \\ = \mathcal{E}[\mathbf{S}_\Sigma(z_1, t) \cdot \mathbf{S}_\Sigma(z_2, t)] + \mathcal{E}[\mathbf{S}_\Sigma(z_1, t) \cdot \Delta\mathbf{S}_N(z_2, t)] \\ + \mathcal{E}[\mathbf{S}_\Sigma(z_2, t) \cdot \Delta\mathbf{S}_N(z_1, t)] + \mathcal{E}[\Delta\mathbf{S}_N(z_1, t) \cdot \Delta\mathbf{S}_N(z_2, t)] \end{aligned} \quad (3.57a)$$

and

$$\mathcal{E}[\mathbf{S}_\Sigma(z_1, t) + \Delta\mathbf{S}_N(z_1, t)] = \mathcal{E}[\mathbf{S}_\Sigma(z_1, t)] + \mathcal{E}[\Delta\mathbf{S}_N(z_1, t)] \quad (3.57b)$$

Because the signal itself and the noise are independent and thus uncorrelated, we have

$$\mathcal{E}[\mathbf{S}_\Sigma(z_1, t) \cdot \Delta\mathbf{S}_N(z_2, t)] = \mathcal{E}[\mathbf{S}_\Sigma(z_1, t)] \cdot \mathcal{E}[\Delta\mathbf{S}_N(z_2, t)] \quad (3.58)$$

Writing \mathcal{C}'_Σ analogous to (3.29) for the primed Stokes vector $\mathbf{S}'_\Sigma = \mathbf{S}_\Sigma + \Delta\mathbf{S}_N$ and inserting above expressions yields

$$\mathcal{C}'_\Sigma(z_1, z_2) = \mathcal{C}_\Sigma(z_1, z_2) + \exp[-\alpha(z_1 + z_2)] \tilde{\gamma}^2 \cdot \mathcal{C}_{\Delta\mathbf{S}}(z_1, z_2) \quad (3.59a)$$

with

$$\mathcal{C}_{\Delta\mathbf{S}}(z_1, z_2) = \mathcal{E}[\Delta\mathbf{S}_N(z_1, t) \cdot \Delta\mathbf{S}_N(z_2, t)] - \mathcal{E}[\Delta\mathbf{S}_N(z_1, t)] \mathcal{E}[\Delta\mathbf{S}_N(z_2, t)] \quad (3.59b)$$

being the noise ACOVF. When the fiber GVD is larger than zero, any difference in z is translated into a

time shift via (2.20), and because we assume the noise spectrum to be white,²⁶

$$\mathcal{C}_{\Delta S}(z_1, z_2) \neq 0 \quad \text{only if} \quad \Delta t = 0 \quad (3.60)$$

with Δt from (2.20), which includes inline GVD compensation. Compared to S_{Σ} , the contribution of ΔS_N to the \mathcal{U} integral (3.30) will be negligible due to the very short correlation time of the noise. Even if the noise is neither exactly white nor its spectral width infinite, we assume its correlation time to be sufficiently small to not influence the results significantly. This assumption is comparable to ignoring the presence of XPM-induced nonlinear phase noise when dealing with XPM phenomena, which has been shown to occur in 10 Gbps probe channels, but is often neglected [Kim 2003, Ho 2004]. The properties and dynamics of this *nonlinear polarization noise* might become relevant if the interfering channels have very low inherent Stokes vector variance, as e.g. DPSK signals do, so that the relative impact of the amplifier noise is increased. However, it is beyond the scope of the present work to examine them.

3.3 APPLICATION OF THE MODEL

In the present section, we will use our model to determine various dependencies of XPOLM on the transmission system which cannot be done with such great variety of parameter combinations using numerical simulations, due to the great amount of calculation time associated with the latter. We will start by quantifying the XPOLM impact versus different properties of the transmission fiber, and then show how residual dispersion per span and/or co-polarized launch SOPs can be used to significantly lower this impact for certain ranges of fiber properties. We will also determine how XPOLM scales with the number of interfering channels that co-propagate with the probe. Using the model, it is not particularly time-consuming to increase this number to beyond 20, while for numerical simulations this would require large simulation bandwidths which are generally associated with prohibitively long simulation durations. With the present model of XPOLM it is also possible to compare the relative impact of interfering channels, depending on their frequency separation from the probe channel, in terms of their contribution to the variance parameter \mathcal{U} .

3.3.1 Dependence on Fiber Parameters

An immediate application of the model derived in the first sections of the current chapter is the comparison of the impact of XPOLM, as measured by the nonlinear depolarization, as it depends on the fiber parameters describing the linear effects that affect the behavior of the interfering channels along

²⁶This assumption is not entirely accurate when optical filters are used to reduce out-of-band optical amplifier noise. However, when there are many co-propagating WDM channels, the noise autocorrelation function (as Fourier transform of the power spectral density) is sufficiently narrow to be considered a Dirac delta function for our purposes.

the transmission path. This is shown in Fig. 3.14 (top), where the nonlinear DOP reduction is plotted versus the GVD and the PMD parameters of the transmission fiber, after 10 spans in a resonant dispersion map. The fiber nonlinearity parameter γ is held constant for illustrative purposes (cf. Table 3.1).

Another similarly useful quantity is the maximum allowable input power into the interfering channels at which the DOP does not drop below some pre-determined value. We call this quantity the *non-linear threshold* (NLT). The NLTs corresponding to the top graph of Fig. 3.14 are shown in its bottom graph for a nonlinear DOP of $\mathcal{D} = 0.97$ in the probe channel. The choice of 0.97 is rather arbitrary at the moment, since we have not yet determined the impact of a certain DOP reduction on the transmission quality. We use it here because it has been estimated by Karlsson and Sunnerud that a DOP reduction to 0.97 can significantly impede PMD compensation [Karlsson 2006]. Because \mathcal{U} is proportional to $(\tilde{\gamma} \mathcal{E}[P])^2$, the threshold powers given in Fig. 3.14 (bottom), when converted to mW, can be easily converted to any other target DOP value \mathcal{D} by making use of relation (3.46) to obtain

$$\text{NLT}_{\mathcal{D}} = \sqrt{\frac{\ln \mathcal{D}}{\ln 0.97}} \text{NLT}_{0.97} \quad (3.61)$$

where $\ln(\cdot)$ is the natural logarithm and nonlinear threshold $(\text{NLT})_{0.97}$ is the nonlinear threshold for a mean DOP of 0.97 from Fig. 3.14. Thus we can easily rescale the NLTs in Fig. 3.14 for any required DOP value, e.g. the NLT values for $\mathcal{D} = 0.90$ will be nearly 3 dB higher than those plotted in the figure, while those for $\mathcal{D} = 0.99$ will be 2.4 dB lower. The results may be less accurate when scaled to very low DOP values when the nonlinear perturbation can no longer be considered small, but – as will become clear in Chapter 5 – these highly-nonlinear cases are generally not of practical interest.

From Fig. 3.14 we learn that low-PMD NZDSFs are affected significantly more than are older, high-PMD SSMFs. They have NLTs which are lower by more than 5 dB. Additionally, fibers that have a low GVD parameter generate significantly more FWM products, so that the NLTs will further decrease with the magnitude of β_2 when taking both XPOLM and FWM into account.

Low PMD fiber results in a comparatively high magnitude of XPOLM in the resonant dispersion map, in which the accumulated GVD is fully compensated within each span and the ACOVF $\mathcal{C}_v^{\text{WO}}$ is restored to its maximum value at the span boundaries. This is to be expected when looking, e.g., at Fig. 3.6 where high PMD is seen to decorrelate the nonlinear polarization rotations from one span to the next while they remain almost fully correlated when PMD is low. The behavior is significantly different in systems that use RDPS to reduce the impact of interchannel nonlinearities. Fig. 3.15 shows the $\text{NLT}_{0.97}$ in systems equal to the one of Fig. 3.14, but with RDPS values of -1024 and -2048 fs/GHz, respectively. These correspond to an uncompensated walk-off (for the channels neighboring the probe) of one-half and one full symbol length per span. With increasing RDPS the pulse shapes in the interfering channels will start to vary from one span to the next and the assumption of a constant envelope eventually becomes invalid. In this case we must calculate the ACOVF matrix as described in Section 3.2.2. However, at the RDPS values used herein, the ACOVFs related to the walk-off remain very similar to those used in Fig. 3.7 so that we keep using the approximation (3.33).

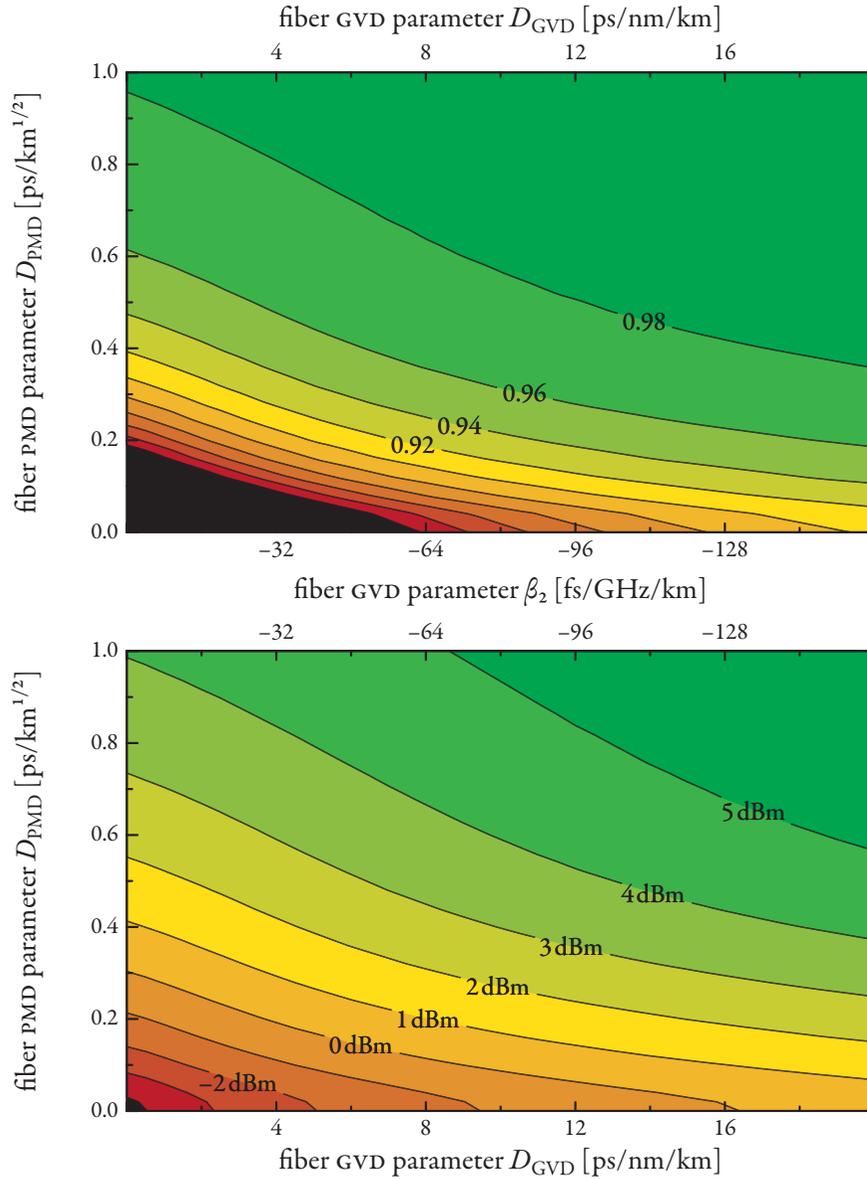


Figure 3.14 plots the mean nonlinear DOP reduction \mathcal{D} in the probe channel (top) and the nonlinear threshold, defined as the interfering channel power required to obtain $\mathcal{D} = 0.97$, versus the fiber parameters for GVD and PMD. The transmission system consisted of 10 spans of SMF with full in-line dispersion compensation at the end of each amplifier span (resonant dispersion map without RDPS). The dark areas in the bottom right corners of the graphs denote DOP values below 0.80 and NLTs below -4 dBm.

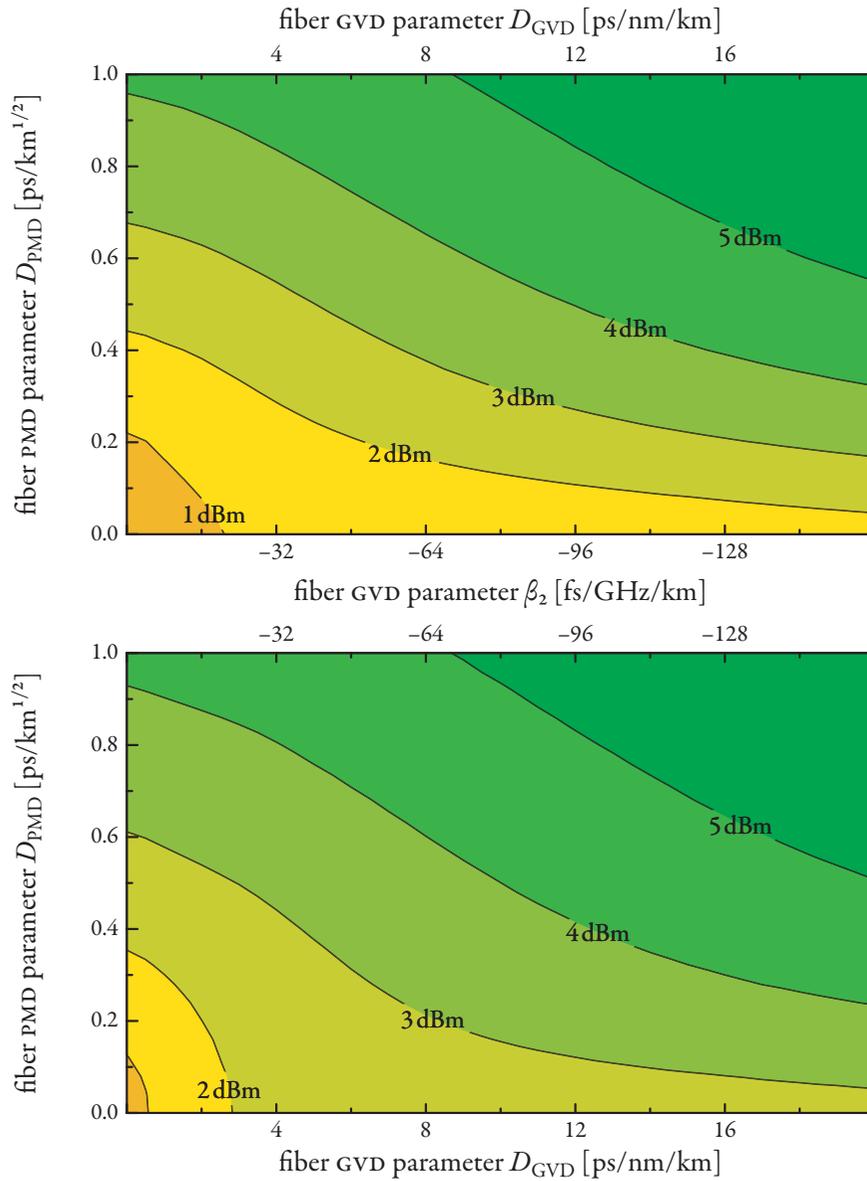


Figure 3.15 plots the nonlinear threshold, defined as the interfering channel power required to obtain $\mathcal{D}=0.97$ in the probe channel, versus the fiber parameters for GVD and PMD. The transmission system consisted of 10 spans of SMF, with residual dispersion per span of -1024 fs/GHz (top) and -2048 fs/GHz (bottom), corresponding to walk-off Δt of one-half and a full symbol in the neighboring interferer. In those cases where the accumulated GVD per span was smaller than the RDPS value, anomalous dispersion is added instead of dispersion compensation.

We see that having non-negligible RDPS values in an optical communication system with NRZ interferers can indeed significantly suppress XPOLM by decorrelating the nonlinear polarization distortions from one span to the next. Statistically, it is similar to high PMD, and thus is less effective in high-PMD systems than it is in low-PMD systems. As Fig. 3.16 shows, the evolution of the DOP remains only weakly dependent on the fiber PMD parameter when RDPS is large enough (one symbol length of uncompensated walk-off per span in this case). However, while the decorrelation between spans due to high PMD is only an ensemble average behavior – meaning that there can be certain members in the ensemble for which the spans do not decorrelate – the decorrelation due to RDPS is almost deterministic (assuming random bit sequences) for each member of the ensemble and thus more reliable. This is also reflected in the histograms of the DOP reduction at the output of the system, shown in Fig. 3.17. While the mean DOP reduction of a system with high PMD is almost ignorant of the presence of RDPS, the variance of the distribution with RDPS is much lower, due to a reduction of the dependence of C_v on C_v^{SOP} , as a result of zero-valued C_v^{WO} for more (or most) z_1 - z_2 pairs. We will discuss this variance of the nonlinear depolarization in more detail in Section 4.1 and also take a look at the differences of the SOP distribution within single members of the polarization ensemble between resonant dispersion maps and those employing RDPS in Section 4.3.

For sake of completeness, Fig. 3.18 shows the nonlinear threshold $\text{NLT}_{0.97}$ in the case of co-polarized (or polarization-interleaved) launch for a system with a resonant dispersion map and one employing RDPS. High-PMD fibers will be little affected by the launch conditions, as we have mentioned previously, but the improvement for low-PMD fiber can be significant. Especially when combined with RDPS, co-polarized launch will result in an almost homogenous distribution of the NLTs over all fiber types.

Given an interferer modulation format, the walk-off ACOVF is parameterized mainly by the walk-off length L_{WO} , as long as accumulated GVD does not significantly modify the pulse shapes, as e.g. eventually happens with PSK-modulated data signals as a result of PM-AM conversion. From (2.21) we can see that L_{WO} is inversely proportional to $|\beta_2| \Delta\omega_v$. Thus, doubling the frequency separation between channels is equivalent to doubling the fiber GVD parameter. Similarly, the SOP decorrelation function (3.45) for a particular interferer is fully parameterized by L_{SOP} , which is inversely proportional to $\Delta\omega_v^2 D_{\text{PMD}}^2$. Here also, doubling the frequency separation between channels is equivalent to doubling the fiber PMD parameter. This makes Figs. 3.14 and 3.15 easily scalable to different values of the frequency separation between channels in the WDM system.

Using our XPOLM model, we can also quickly identify sensible values of RDPS to be used, as we would e.g. expect an increase in walk-off beyond one symbol per span in each of the interfering channels to be of little extra benefit. Fig. 3.19 plots the mean nonlinear DOP reduction versus RDPS for various fiber PMD values. The amount of RDPS is given in terms of the equivalent length L_{RDPS} of uncompensated SMF length per span. Formally, we have

$$L_{\text{RDPS}} = \frac{\text{RDPS}}{\beta_2} \quad (3.62)$$

when RDPS is specified in units of fs/GHz, corresponding to accumulated dispersion. The exact shape of

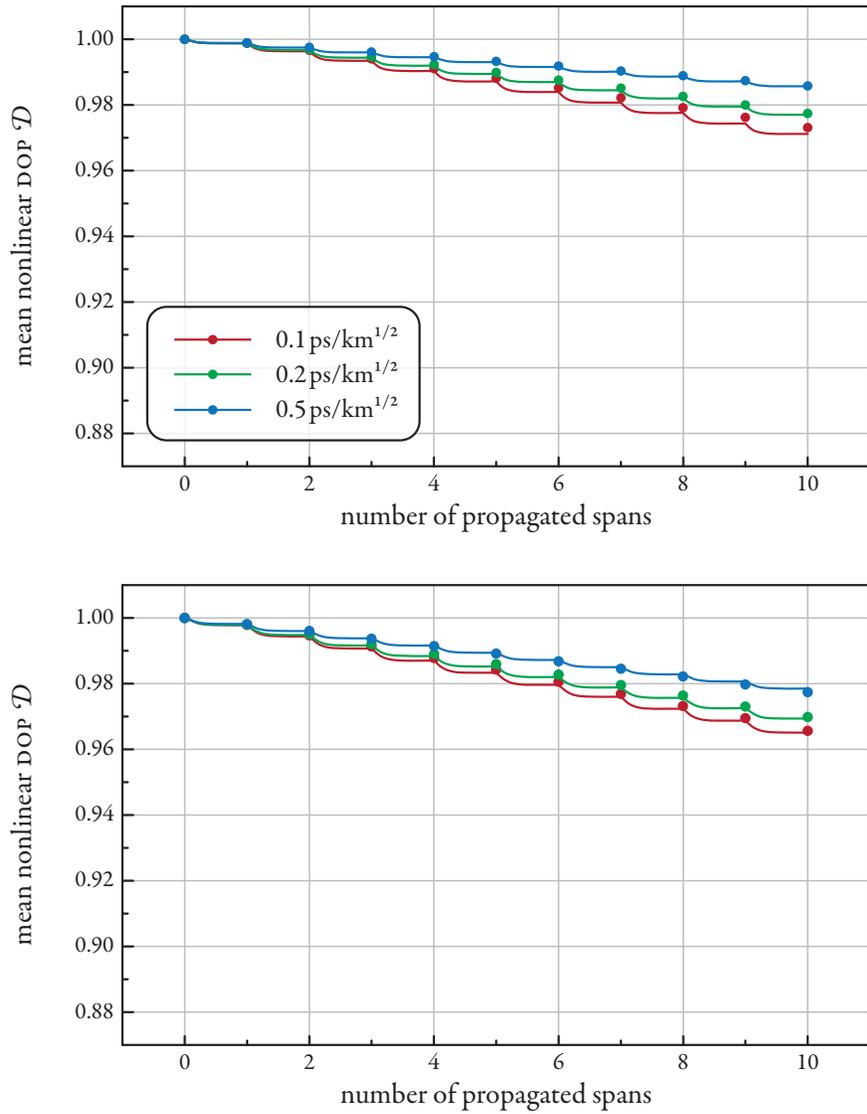


Figure 3.16 shows the evolution of the DOP with propagation over 10 spans of SMF with $\beta_2 = -128$ fs/GHz/km (top) and -64 fs/GHz/km (bottom) with varying amounts of fiber PMD, and residual dispersion per span of -2048 fs/GHz, corresponding to walk-off Δt of one symbol in the interferer neighboring the probe. All other parameters are taken from Table 3.1. Symbols denote simulation ensemble mean values and lines denote model results.

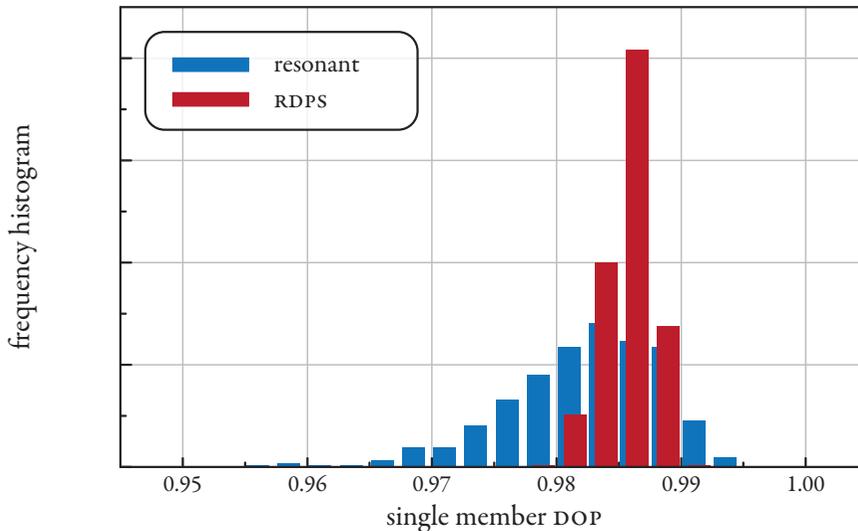


Figure 3.17 compares the DOP histograms obtained in 500 numerical simulation runs each with a resonant dispersion map (mean DOP $\mathcal{D} = 0.982$) and RDPS of -2048 fs/GHz ($\mathcal{D} = 0.986$). Fiber GVD was $\beta_2 = -128$ fs/GHz/km and PMD was $D_{\text{PMD}} = 0.5$ ps/km^{1/2} in both cases, corresponding to Figs. 3.10 and 3.16.

the curves depends on the walk-off ΔCOVF and thus on the modulation format used – for the curves in Fig. 3.19, we used the linear approximation (3.33) for ideal NRZ pulses. We can see that the effectiveness of RDPS slows down at an uncompensated walk-off corresponding to about one full NRZ symbol length, as expected. Further increases in RDPS have a small additional positive effect, because they not only decorrelate $\mathcal{C}_v^{\text{WO}}$ at the beginning of consecutive spans (where the optical power is highest), but also between locations in the fiber which are spaced apart by non-integer multiples of the amplifier span length. Ideally, one would have no in-line dispersion compensation at all.²⁷ In this case, there would be no two locations within the transmission system that have the same value of accumulated dispersion $B_2(z) = \int_0^z \beta_2(\zeta) d\zeta$ (and thus a relative walk-off $\Delta t_v(z_1, z_2) = 0$ and high value of $\mathcal{C}_v^{\text{WO}}$). As long as there is at least some RDPS, there exist location pairs with equal B_2 , but there are significantly less such pairs than in a fully resonant map. This is shown schematically in Fig. 3.20: without in-line dispersion compensation no such pair exists; with small amounts of RDPS, the number of pairs is greatly reduced compared to the resonant dispersion map (1 pair vs. 10 pairs, respectively).

²⁷We should note that such a configuration is not optimal from the viewpoint of intrachannel nonlinearities, especially intrachannel FWM and XPM in high-speed systems, due to the large amount of overlapping symbols [Essiambre 2002]. Also, large amounts of accumulated GVD can be detrimental for amplitude-modulated systems, because of significant FM-AM conversion efficiency [Bayvel 2002]. Both effects are not captured correctly when the probe channel is CW. Thus, in realistic systems, the amount of permissible RDPS will be subject to restrictions other than those examined here.

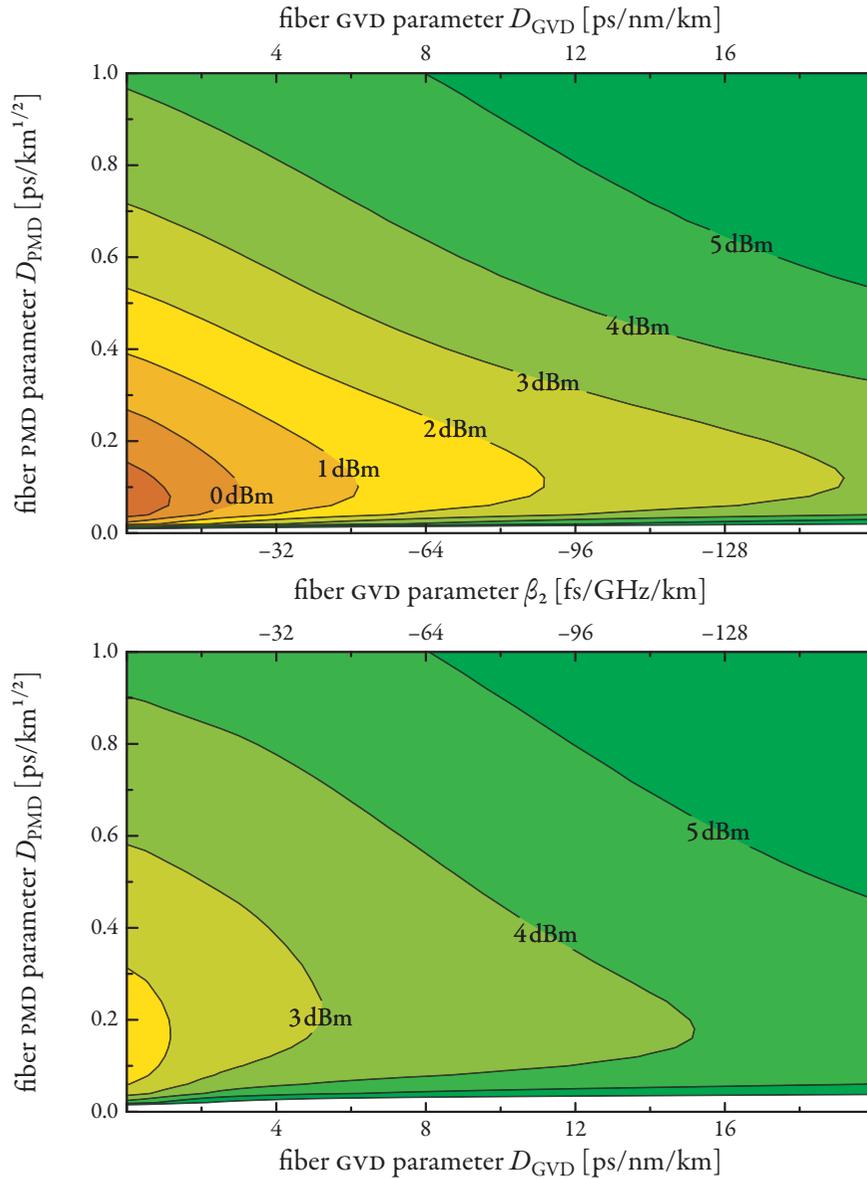


Figure 3.18 plots the nonlinear threshold, defined as the interfering channel power required to obtain $\mathcal{D}=0.97$ in the probe channel, versus the fiber parameters for GVD and PMD. The transmission system consisted of 10 spans of SMF with a resonant dispersion map (top) and RDPS of -2048 fs/GHz (bottom). All channels are assumed co-polarized at launch.

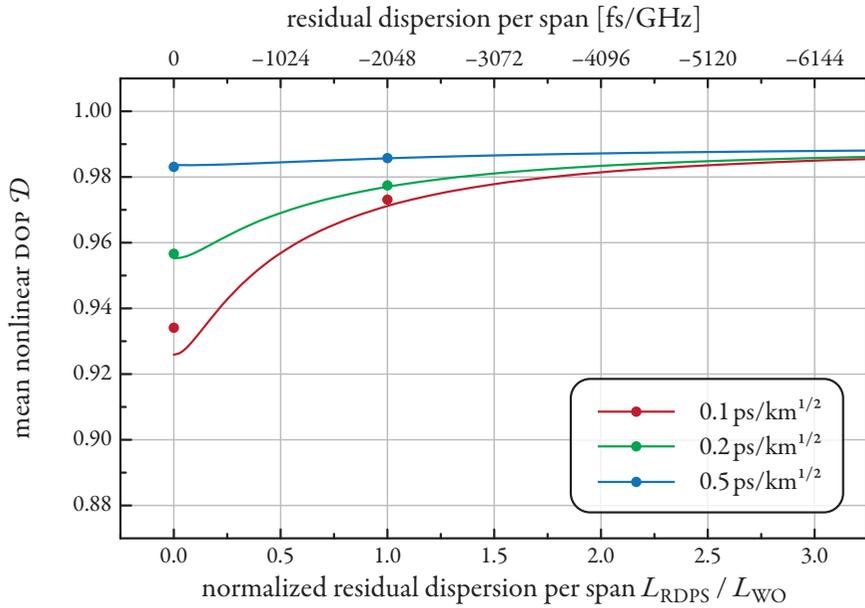


Figure 3.19 plots the mean nonlinear DOP reduction versus the amount of RDPS after propagation over 10 spans of SSMF with $\beta_2 = -128$ fs/GHz/km for various fiber PMD parameters, obtained with the XPOLM model. RDPS is given in terms of uncompensated SSMF length L_{RDPS} and is normalized to the walk-off length L_{WO} of the interfering channels neighboring the probe.

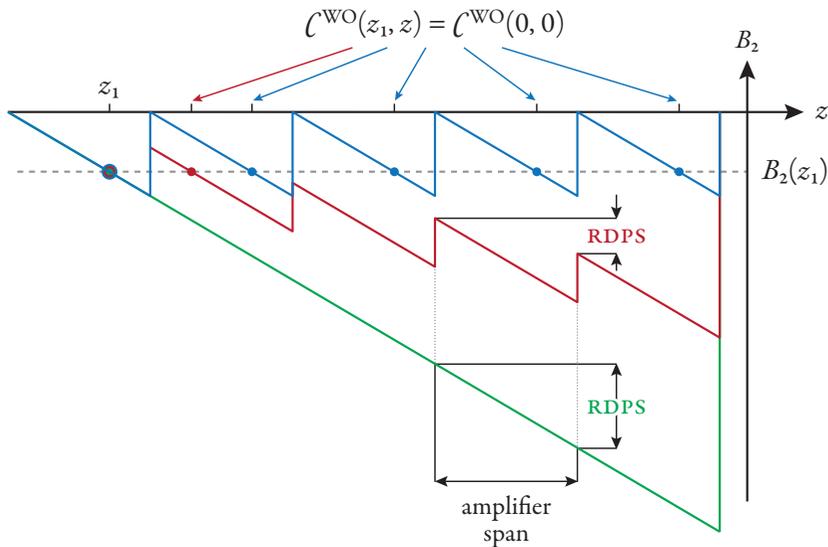


Figure 3.20 illustrates how RDPS helps to reduce interchannel nonlinearities by avoiding argument pairs (z_1, z_2) for which C_v^{WO} is large (because the walk-off Δt_v between them becomes zero).

3.3.2 Dependence on Channel Load

Another application in which our simple model can replace extensive simulation and/or measurements is the determination of the relative impact of the various interfering channels. We have mentioned before on several occasions that the channels immediately neighboring the probe are the most relevant regarding its nonlinear distortions. We shall elaborate on that presently.

The impact of each channel is determined by its contribution to \mathcal{U} via the variance integral (3.30) and the ACOVF (3.32a). The fiber location pairs (z_1, z_2) for which the integrand corresponding to each channel is greater than zero are determined in turn by its walk-off length $L_{\text{WO},\nu}$, and its SOP decorrelation length $L_{\text{SOP},\nu}$, which depend linearly and squared on the frequency separation $\Delta\omega_\nu$ between the probe ρ and the interferer ν , respectively. We can thus easily – for any combination of the system parameters β_2 , D_{PMD} , the number of spans, and the RDPS – plot a per-channel contribution to \mathcal{U} of a channel spaced at any particular $\Delta\omega_\nu$. This has been done in Fig. 3.21 for SSMF and different PMD parameters and RDPS values. We also plotted the relative contribution when all channels are launched co-polarized or polarization-interleaved.

We can see that for the isotropic launch the channels nearest to the probe have a significantly larger contribution to \mathcal{U} than all other interferers. Without RDPS, the difference between a channel at $\Delta\omega_\nu = 2\pi 50$ GHz and one at 100 GHz is often a factor of 3 or 4. When fiber PMD is low, it might be as low as 2. Also, with RDPS it is often less than 2.²⁸ Channels further away contribute comparatively little. This is also evident from Fig. 3.22 which plots the DOP reduction versus the number of interferers (leaving the optical power of each interferer constant). An increase in the channel load beyond 11 (including the probe) does not significantly affect the mean nonlinear depolarization – this fact prompted us to use 10 interferers in our numerical simulations in an attempt to balance simulation time and accuracy.

These analytical results are confirmed by measurements performed by Bertran-Pardo et al., which determined an improvement in transmission depending on the number of interferers removed from the channel map, starting with the ones closest to the probe. In [Bertran-Pardo 2008b, Fig. 4] we see that the removal of the innermost neighbors has the largest impact on the system, which thereafter only slowly improves. It is difficult and expensive to perform a necessary amount of measurements to be able to make statements about ensemble behavior, and Bertran-Pardo et al. only measured a single ensemble member, but the predicted behavior is already evident.

In Fig. 3.21 (bottom) we can also see why co-polarized or polarization-interleaved launch can significantly reduce the impact of XPOLM when fiber PMD is low. Such launch conditions can significantly reduce the impact of the spectrally closest channels – those that contribute most with isotropic launch – since the polarization rotations due to XPOLM are greatly suppressed as long as probe and interferer have (anti-)parallel Stokes vectors. The SOPs of spectrally distant channels diffuse over short distances mostly regardless of the magnitude of fiber PMD, as the SOP decorrelation length scales with the square of $\Delta\omega_\nu$,

²⁸That is a primary factor for which SOP distributions in Fig. 3.4 match the Brownian distribution well – when the contribution of multiple channels is better balanced, the overall randomness is greater than when only one or two channels dominate this statistics.

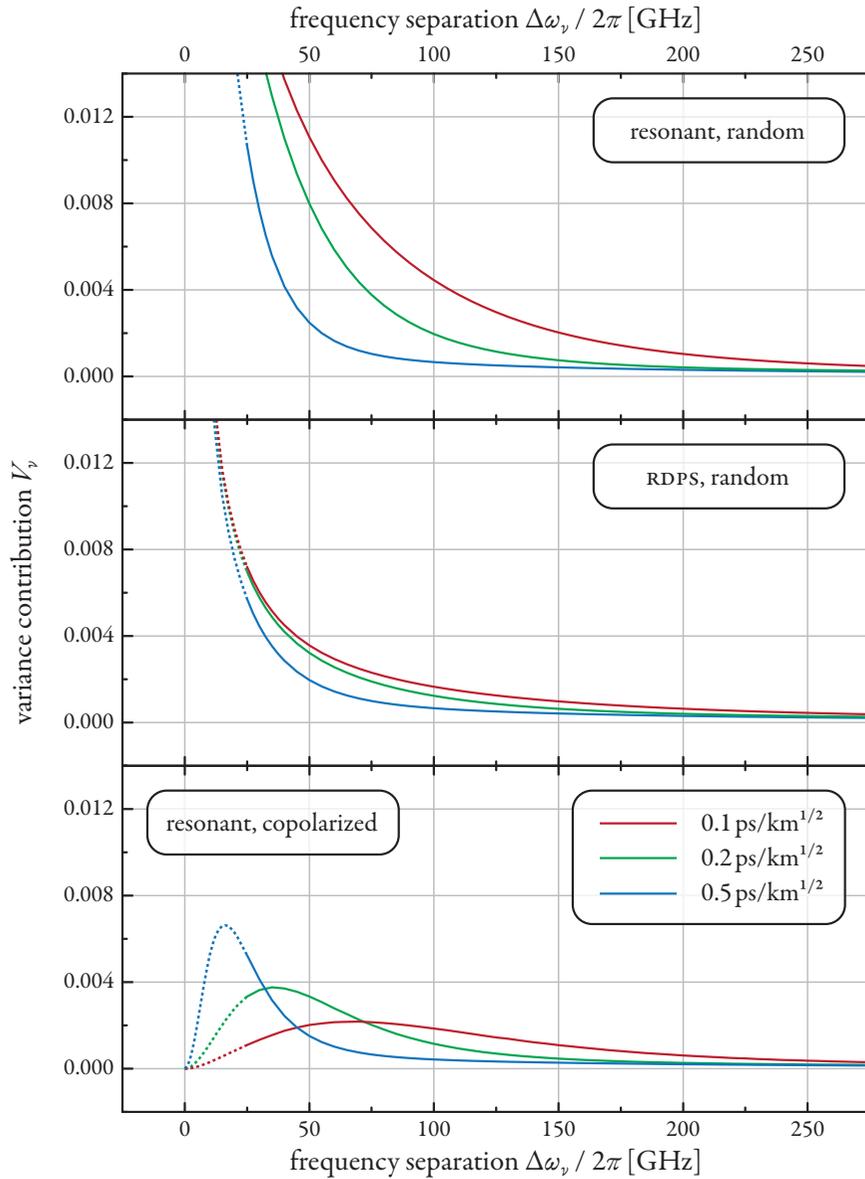


Figure 3.21 compares the dependence of the contribution of a single WDM channel ν and its ACOVF \mathcal{C}_ν to the \mathcal{U} integral (3.30) on the frequency separation $\Delta\omega_v$ of that channels from the probe channel ρ after propagation of 10 spans for three system types: resonant dispersion map with randomly polarized launch (top), residual dispersion per span of -2048 fs/GHz with randomly polarized launch (middle), and resonant map with copolarized launch of all channels (bottom). Channel power was 0 dBm.

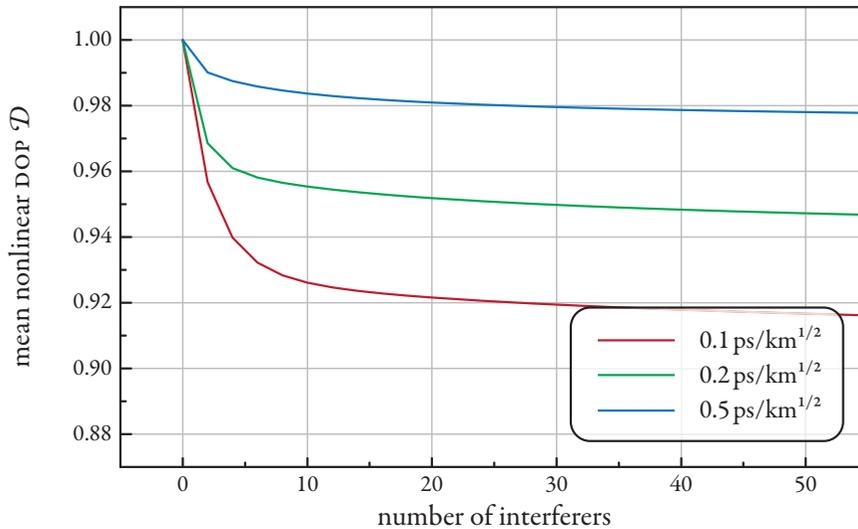


Figure 3.22 shows the mean nonlinear DOP after propagation of 10 spans versus the number of interfering channels, obtained using the present model, for various values of the fiber PMD parameter in a resonant dispersion map with randomly polarized launch (worst case interference).

and for those there is little difference between isotropic and predetermined launch. However, in any case, the contribution of each channel to \mathcal{U} is never larger for co-polarized or polarization-interleaved launch than it is for isotropic launch. This was also visible in Fig. 3.13, where clearly the improvement was larger for low fiber PMD values.

Once again for the sake of completeness, Fig. 3.23 shows the comparison between theory and numerical simulation for the co-polarized and polarization-interleaved launch into a system with significant RDPS. The total nonlinear perturbation is so small that there is no visible difference between either condition, regardless of fiber PMD. Otherwise, the simulation results and the model prediction match very well.

In the present chapter we have shown that the ensemble distribution of the polarization states of a probe channel which is subjected to XPOLM is equal (or close) to the Brownian distribution on the unit sphere, as it results from random polarization rotations during propagation along the fiber. Mathematically, the Brownian distribution is well approximated by the Fisher distribution, which is preferable to be used for calculations due to its simpler probability density function. Both distributions are described by a single parameter which is related to the nonlinear degree of polarization reduction of the probe.

The nature of the interfering channels causing XPOLM is not relevant for the Brownian distribution shape, it only affects the resulting DOP. We have chosen 10 Gbps NRZ channels to demonstrate the validity of the random rotations model, but other data rates and formats may be used as well. Because XPOLM accumulates much slower for e.g. DPSK channels (cf. Fig. 3.7), the expected nonlinear threshold

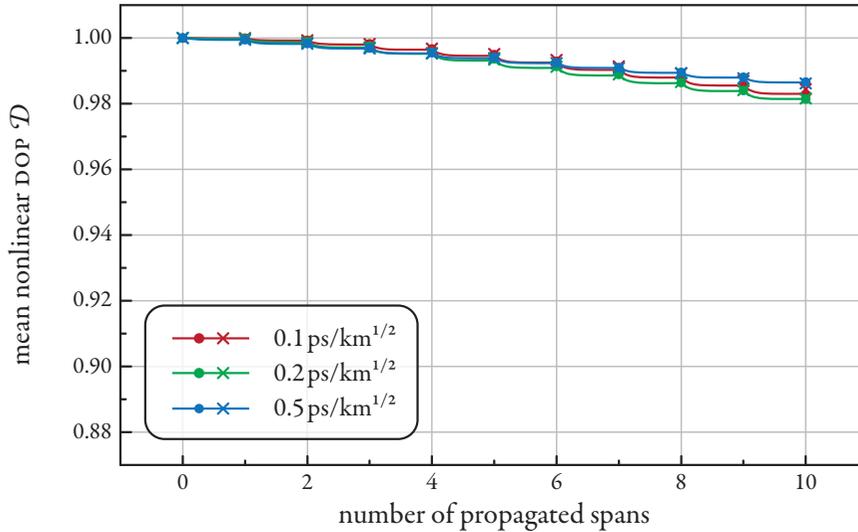


Figure 3.23 compares the evolution of the DOP between theory and simulations for propagation over 10 spans of SSMF with $\beta_2 = -128$ fs/GHz/km with varying amounts of fiber PMD, and residual dispersion per span of -2048 fs/GHz. The system is thus equal to that of Fig. 3.19, except that all channels were launched co-polarized (circles) or mutually orthogonal, so-called polarization interleaving (crosses).

to obtain a particular DOP reduction in the probe will be significantly higher. However, the distribution of the SOPs at that DOP will be the same, regardless of the nature of its cause. This way we can decouple the discussion of the actual transmission penalties due to XPOLM being caused by the SOP diffusion somewhat from its origins, as we may simply presume Brownian statistics of the SOPs in any further penalty analysis. This is being done in Chapter 5.

Before we can attend to the transmission penalty analysis, we must yet spend some time on a closer examination of the polarization ensemble. Until now, we have (with some accuracy) attempted to predict the mean nonlinear depolarization of the ensemble as a whole, which is a useful measure of the susceptibility of a system (including the dispersion map, fiber types and channel map) to XPOLM in general and to devise and compare methods to reduce this susceptibility. But as we have already hinted at in the histograms in Fig. 3.17, the DOP of single members of this ensemble – corresponding to polarization-related changes in time or within a class of systems – can deviate significantly from the mean value. Ensemble members / systems which exhibit large deviations (towards lower DOP values) will certainly be affected differently than the «mean» system and may incur much higher transmission penalties. When this additional penalty exceeds a certain limit – e.g. when forward error correction (FEC) is no longer able to correct an increasing number of bit errors – such systems may be regarded as experiencing an *outage*, similar to PMD outages. The probability of such outage events is a very important quantity for the operators of transmission systems, and thus the ensemble statistics deserve their own chapter in this text, which follows next.

AS WE HAVE seen in the previous chapter, the distribution of the DOP values within the ensemble can have a significant variance, and this variance also depends on the nature of the transmission system (cf. Fig. 3.17 – RDPS reduces this variance considerably). In the present chapter, we will not yet determine any outage probabilities in terms of transmission penalty, but we shall describe the distribution of the DOP values and determine its variance. We will also show that the distribution of SOPs within a single ensemble member is not necessarily equal to the Brownian distribution of the SOPs in the ensemble as a whole. We will introduce the concept of *stochastic degrees of freedom* to describe both the DOP and SOP distribution from a less mathematical and more descriptive point of view, which is still useful to compare the influence of the various system parameters on the statistics of XPOLM and to aid in the estimation of penalties.

4.1 THE VARIANCE OF THE DOP

Unfortunately, we do not have a closed-form expression for the DOP of a single member of the polarization ensemble, but only for its mean \mathcal{D} , given in (3.15) as a function of the variance parameter \mathcal{U} which is itself a mean over the ensemble, via the integral (3.30) over the ACOVF (3.32). Similar to the relation between D and \mathcal{D} we can introduce the parameter V for a single member of the polarization ensemble, so that

$$\mathcal{U}(L) = \langle V(L) \rangle \quad (4.1)$$

V can be derived from the general expression (3.47) as

$$V(L) = \int_0^L \int_0^L \Gamma(\zeta_1, \zeta_2) \sum_{\nu \neq \rho} \mathcal{C}_\nu^{\text{WO}}(\zeta_1, \zeta_2) [\mathbf{S}_{\nu\perp}(\zeta_1) \cdot \mathbf{S}_{\nu\perp}(\zeta_2)] d\zeta_1 d\zeta_2 \quad (4.2)$$

and will be different for every combination of $\mathbf{S}_{\nu\perp}(\zeta_1) \cdot \mathbf{S}_{\nu\perp}(\zeta_2)$. It no longer includes the ensemble averaging. Here, $\mathbf{S}_{\nu\perp}$ is again the projection of the SOP of channel ν onto the plane orthogonal to the probe, which was defined in (3.47b). We will determine the variance of V which can then be related to the variance of the DOP D . However, strictly we cannot apply expression (3.15) relating \mathcal{D} and \mathcal{U} to a single member of the ensemble, because it is only valid for the Brownian population distribution. Even if their exponential relation would hold, in order to properly relate the variances of V and D , we would also need the distribution law of V . To be able to make at least qualitatively correct statements about the DOP distribution and because we are most interested in DOP values near unity, we assume the following

linear relation between V and D , corresponding to the power series expansion of (3.15), truncated after the linear term:

$$D(z) \approx 1 - \frac{V(z)}{2} \quad (4.3)$$

The variance $\text{var}[D]$ of the DOP can then be determined approximately from

$$\begin{aligned} \text{var}[D] &= \langle D^2 \rangle - \langle D \rangle^2 \\ &= \frac{1}{4} (\langle V^2 \rangle - \langle V \rangle^2) = \frac{1}{4} \text{var}[V] \end{aligned} \quad (4.4)$$

We thus seek the ensemble variance $\text{var}[V]$ of the integral (4.2) whose mean is \mathcal{U} . However, just as this mean variance \mathcal{U} of the integral XPOLM process became a double integral over the autocovariance function in Section 3.2, so the variance of this double integral becomes a quadruple integral over an appropriate autocovariance function. This new ACOVF is analogously

$$\begin{aligned} \mathcal{C}\mathcal{C}_v^{\text{SOP}}(\zeta_1, \zeta_2, \xi_1, \xi_2) &= \left\langle \left[\mathbf{s}_{v\perp}(\zeta_1) \cdot \mathbf{s}_{v\perp}(\zeta_2) \right] \left[\mathbf{s}_{v\perp}(\xi_1) \cdot \mathbf{s}_{v\perp}(\xi_1) \right] \right\rangle \\ &\quad - \left\langle \mathbf{s}_{v\perp}(\zeta_1) \cdot \mathbf{s}_{v\perp}(\zeta_2) \right\rangle \left\langle \mathbf{s}_{v\perp}(\xi_1) \cdot \mathbf{s}_{v\perp}(\xi_2) \right\rangle \end{aligned} \quad (4.5)$$

and

$$\begin{aligned} \text{var}[V(L)] &= \int \int \int \int_0^L \Gamma(\zeta_1, \zeta_2) \Gamma(\xi_1, \xi_2) \sum_{\nu \neq \rho} \mathcal{C}_\nu^{\text{WO}}(\zeta_1, \zeta_2) \mathcal{C}_\nu^{\text{WO}}(\xi_1, \xi_2) \\ &\quad \times \mathcal{C}\mathcal{C}_\nu^{\text{SOP}}(\zeta_1, \zeta_2, \xi_1, \xi_2) d\xi_1 d\xi_2 d\zeta_1 d\zeta_2 \end{aligned} \quad (4.6)$$

where all four integrals have to be evaluated from 0 to L .

Thus, for every pair (ζ_1, ζ_2) , or interval, that contributes to the integral (3.47), we need to determine the ACOVF with every other possible pair (ξ_1, ξ_2) , and integrate over all such contributions. The functions Γ and $\mathcal{C}_\nu^{\text{WO}}$ in (4.6) were defined in (3.32), respectively. We are left with finding a solution for $\mathcal{C}\mathcal{C}_\nu^{\text{SOP}}$ to be able to calculate the integral.

The factors in the second term of the definition of $\mathcal{C}\mathcal{C}_\nu^{\text{SOP}}$ have already been calculated in Section 3.2.4 and are given in (3.51). We will assume randomly polarized launch SOPs for simplicity, so that we have

$$\left\langle \mathbf{s}_{v\perp}(\zeta_1) \cdot \mathbf{s}_{v\perp}(\zeta_2) \right\rangle = \frac{2}{3} \mathcal{C}_\nu^{\text{SOP}}(\zeta_1, \zeta_2) \quad (4.7)$$

with $\mathcal{C}_\nu^{\text{SOP}}$ defined in (3.45), and analogous for the ξ .

For a derivation of the first term in the definition of $\mathcal{C}\mathcal{C}_\nu^{\text{SOP}}$ we assume two cases of the most general positions ζ_1, ζ_2, ξ_1 , and ξ_2 as shown in Fig. 4.1: either the two intervals are (partially) overlapping or they are disjunct. Any different order of the four positions is included in the solution by simply exchanging

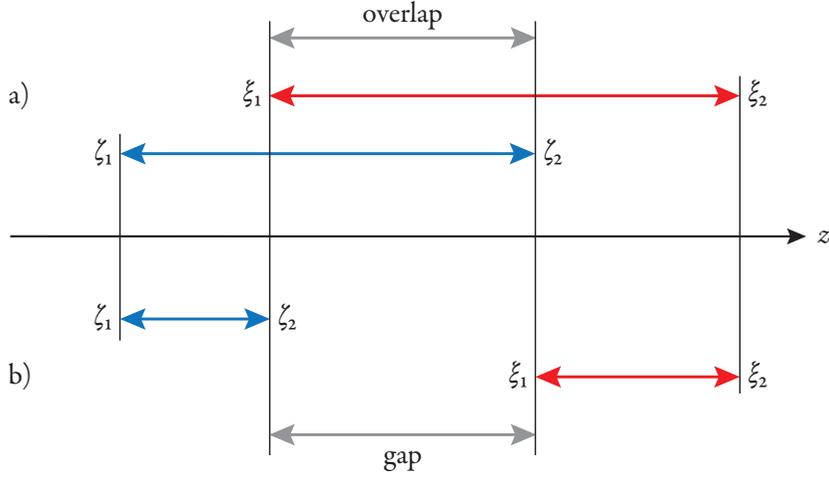


Figure 4.1 illustrates the relation between the position parameters ζ_1 , ζ_2 , ξ_1 , and ξ_2 as used in the derivations within the present section for two different cases: a) the intervals are overlapping, and b) the intervals are non-overlapping.

the relevant variables.²⁹ Any special (degenerate) cases are also included, as we will show. Our method is similar to the two-step approach employed in Section 3.2.4 for co-polarized launch: instead of the ensemble average over the SOPs at all four locations, we initially average only over all possible $\hat{\mathbf{S}}_\nu(\zeta_1)$, given a certain $\hat{\mathbf{S}}_\nu(\xi_1)$, and all possible $\hat{\mathbf{S}}_\nu(\xi_2)$, given a certain $\hat{\mathbf{S}}_\nu(\zeta_2)$.³⁰ Thus we have

$$\langle \hat{\mathbf{S}}_\nu(\zeta_1) \rangle = \mathcal{C}_\nu^{\text{SOP}}(\zeta_1, \xi_1) \hat{\mathbf{S}}_\nu(\xi_1) \quad (4.8)$$

and analogously for the (ζ_2, ξ_2) pair. We now replace $\mathbf{S}_{\nu\perp}(\zeta_1)$ and $\mathbf{S}_{\nu\perp}(\xi_2)$ in (4.5) using the expansion (3.47b), and make use of the relation (4.8) for $\hat{\mathbf{S}}_\nu(\zeta_1)$ and $\hat{\mathbf{S}}_\nu(\xi_2)$ to obtain

$$\begin{aligned} \mathcal{C}_\nu^{\text{SOP}} = & \left(\langle [\mathbf{S}_{\nu\perp}(\zeta_2) \cdot \mathbf{S}_{\nu\perp}(\xi_1)]^2 \rangle - \left[\frac{2}{3} \mathcal{C}_\nu^{\text{SOP}}(\zeta_2, \xi_1) \right]^2 \right) \\ & \times \mathcal{C}_\nu^{\text{SOP}}(\zeta_1, \xi_1) \mathcal{C}_\nu^{\text{SOP}}(\zeta_2, \xi_2) \quad \text{if } \zeta_1 \leq \xi_1 \leq \zeta_2 \leq \xi_2 \quad (4.9) \end{aligned}$$

²⁹We can freely exchange either ζ_1 with ζ_2 , ξ_1 with ξ_2 , or both ζ variables with both ξ variables without altering the result due to the symmetry in the equations.

³⁰Without elaboration or proof, we state that the linear SOP evolution due to PMD is a continuous Markov process with isotropic transition probabilities. As such it is reversible and the reverse process will have the identical probabilistic structure as the original process [Primak 2004, Sec. 5.5], and the probability distribution of $\hat{\mathbf{S}}_\nu(\zeta_1)$ given $\hat{\mathbf{S}}_\nu(\xi_1)$ is the same as that in the reversed case.

where we made use of the properties

$$\mathcal{C}_\nu^{\text{SOP}}(\zeta_1, \zeta_2) = \mathcal{C}_\nu^{\text{SOP}}(\zeta_1, \xi_1) \mathcal{C}_\nu^{\text{SOP}}(\xi_1, \zeta_2) \quad (4.10a)$$

$$\mathcal{C}_\nu^{\text{SOP}}(\xi_1, \xi_2) = \mathcal{C}_\nu^{\text{SOP}}(\xi_1, \zeta_2) \mathcal{C}_\nu^{\text{SOP}}(\zeta_2, \xi_2) \quad (4.10b)$$

which derive from the definition of $\mathcal{C}_\nu^{\text{SOP}}$ in (3.45). The ensemble average can be calculated numerically using the property that $\hat{\mathbf{S}}_\nu(\xi_1)$ must be distributed according to a Brownian distribution centered on $\hat{\mathbf{S}}_\nu(\zeta_2)$ – as a result of PMD – with a variance parameter $\mathcal{V}_\nu^{\text{SOP}}(\zeta_2, \xi_1)$, given in (3.44). If we substitute the Fisher distribution for the Brownian, (4.9) has an analytical solution:

$$\mathcal{C}_\nu^{\text{SOP}} = \left(\frac{22 + 24\kappa^2 - 2\kappa \coth \kappa - 20\kappa^2 \coth^2 \kappa}{45\kappa^2} \right) \mathcal{C}_\nu^{\text{SOP}}(\zeta_1, \xi_1) \mathcal{C}_\nu^{\text{SOP}}(\zeta_2, \xi_2) \quad \text{if } \zeta_1 \leq \xi_1 \leq \zeta_2 \leq \xi_2 \quad (4.11)$$

in which κ is related to $\mathcal{V}_\nu^{\text{SOP}}(\xi_1 - \zeta_2)$ via (3.17).

The various terms in (4.9) all have a physical and/or statistical meaning and are not only abstract mathematical constructs. We shall thus briefly explain them, in the hope of furthering an understanding of these quite complex expressions. We start with the degenerate case $\zeta_1 = \xi_1$ and $\zeta_2 = \xi_2$, in which both intervals are fully overlapping, and $\mathcal{C}_\nu^{\text{SOP}}(\zeta_1, \xi_1) = \mathcal{C}_\nu^{\text{SOP}}(\zeta_2, \xi_2) = 1$.

The term in parentheses in (4.9) is then related to the variance of the contribution of the location pair (ζ_2, ξ_1) to the V -integral (4.2). As the length of this section increases, so does its variance, as both $\hat{\mathbf{S}}_\nu(\zeta_2)$ and $\hat{\mathbf{S}}_\nu(\xi_1)$ become uncorrelated as a result of PMD-related SOP diffusion within the interval. This term only depends on the accumulated PMD in the interval, as expressed by $\mathcal{V}_\nu^{\text{SOP}}(\zeta_2, \xi_1)$, and can be computed numerically by first integrating over all possible $\hat{\mathbf{S}}_\nu(\xi_2)$, given by the Brownian distribution around $\hat{\mathbf{S}}_\nu(\zeta_2)$, and subsequently over all possible $\hat{\mathbf{S}}_\nu(\zeta_2)$, distributed uniformly on the Poincaré sphere.

Increasing the degeneracy further so that $\zeta_2 = \xi_1$ and the interval has length zero, the ACOVF $\mathcal{C}_\nu^{\text{SOP}}$ is determined to be equal to $4/45$ – it is greater than zero because we are not seeking the variance of the squared SOP $\hat{\mathbf{S}}_\nu^2(\zeta_1)$, but the variance of its projection onto the plane orthogonal to the probe, which will depend on the angle between $\hat{\mathbf{S}}_\nu(\zeta_1)$ and the probe – both of which being distributed uniformly – resulting in a finite variance contribution.

When there are overlapping and non-overlapping sections (non-degenerate case) of the intervals as shown in Fig. 4.1a, the covariance must decrease, because the non-overlapping sections introduce an uncertainty into the expectation (4.5). Even when the overlap is large, if the non-overlapping section between ζ_2 and ξ_2 is large enough that PMD has greatly decorrelated the SOPs between these locations, then the product $\hat{\mathbf{S}}_\nu(\xi_1) \cdot \hat{\mathbf{S}}_\nu(\xi_2)$ cannot have significant correlation with $\hat{\mathbf{S}}_\nu(\zeta_1) \cdot \hat{\mathbf{S}}_\nu(\zeta_2)$, as quantified by $\mathcal{C}_\nu^{\text{SOP}}$. This additional uncertainty is expressed by the ACOVFs $\mathcal{C}_\nu^{\text{SOP}}(\zeta_1, \xi_1)$ and $\mathcal{C}_\nu^{\text{SOP}}(\zeta_2, \xi_2)$ of the SOPs in the non-overlapping sections in (4.9).

If $\zeta_1 \leq \zeta_2 < \xi_1 \leq \xi_2$, i.e. there is no overlap section common to both intervals, as illustrated in

Fig. 4.1b, the relations between the various $\hat{\mathbf{S}}_\nu$ are somewhat altered. In this case (4.8) becomes

$$\langle \hat{\mathbf{S}}_\nu(\zeta_1) \rangle = \mathcal{C}_\nu^{\text{SOP}}(\zeta_1, \zeta_2) \hat{\mathbf{S}}_\nu(\zeta_2) \quad (4.12)$$

where the distribution of $\hat{\mathbf{S}}_\nu(\zeta_1)$ no longer depends on $\hat{\mathbf{S}}_\nu(\xi_1)$, but on (the closer) $\hat{\mathbf{S}}_\nu(\zeta_2)$. An analogous expression exists for the other interval. Then, (4.9) becomes

$$\mathcal{C}_\nu^{\text{SOP}} = \left(\left\langle [\mathbf{s}_{\nu\perp}(\zeta_2) \cdot \mathbf{s}_{\nu\perp}(\xi_1)]^2 \right\rangle - \frac{4}{9} \right) \mathcal{C}_\nu^{\text{SOP}}(\zeta_1, \zeta_2) \mathcal{C}_\nu^{\text{SOP}}(\xi_1, \xi_2) \quad (4.13)$$

if $\zeta_1 \leq \zeta_2 < \xi_1 \leq \xi_2$

which again has an analytical solution when we use the Fisher distribution instead of the Brownian:

$$\mathcal{C}_\nu^{\text{SOP}} = \left(\frac{12 + 4\kappa^2 - 12\kappa \coth \kappa}{45\kappa^2} \right) \mathcal{C}_\nu^{\text{SOP}}(\zeta_1, \zeta_2) \mathcal{C}_\nu^{\text{SOP}}(\xi_1, \xi_2) \quad (4.14)$$

if $\zeta_1 \leq \zeta_2 < \xi_1 \leq \xi_2$

where κ is related via (3.17) to $\mathcal{U}_\nu^{\text{SOP}}(\xi_1 - \zeta_2)$. As long as PMD has not fully decorrelated the SOPs in the length separating the intervals, there still exists some residual correlation, which is captured in (4.13). Fig. 4.2 plots the function $\mathcal{C}_\nu^{\text{SOP}}$ versus the distance $\xi_1 - \zeta_2$, normalized to the accumulated PMD in terms of $\mathcal{U}_\nu^{\text{SOP}}(\zeta_2, \xi_1)$, for the degenerate case of $\zeta_1 = \xi_1$ and $\zeta_2 = \xi_2$. Negative values represent non-overlapping intervals as described by (4.13), and the degeneracy then corresponds to $\zeta_1 = \zeta_2$ and $\xi_1 = \xi_2$.

But regardless of how much we can simplify the involved functions, determining the fourfold integral (4.6) will still require prohibitive amounts of time. We thus need a further simplification to make the variance $\text{var}[D]$ calculable.

In order to do this, we initially regard the case in which ζ_1, ζ_2, ξ_1 , and ξ_2 are locations within the same span. We can see in (4.6) that the ACOVF $\mathcal{C}_\nu^{\text{SOP}}$ scales with the walk-off-related ACOVFs $\mathcal{C}_\nu^{\text{WO}}$ of both intervals and the attenuation at the four involved locations as quantified by the functions Γ . If either interval becomes larger than the walk-off length, the variance contribution will be zero. If the intervals are disjoint and spaced apart significantly, at least one of both attenuation-related functions will be nearly zero. Thus, most contributions to $\text{var}[V]$ come from small intervals at the beginning of the span. We now propose that PMD can be neglected for all such intervals and reasonable PMD values, so that the SOP $\hat{\mathbf{S}}_\nu$ at all locations within a span is the same, resulting in $\mathcal{C}_\nu^{\text{SOP}}(\zeta_1, \zeta_2, \xi_1, \xi_2) = \mathcal{C}_\nu^{\text{SOP}}(\zeta_1, \zeta_1, \xi_1, \xi_1) = 4/45$ when $\zeta_1, \zeta_2, \xi_1, \xi_2$ are all within the same span.

When extended to multi-span systems, this means that for the purpose of determining $\mathcal{C}_\nu^{\text{SOP}}$ we can write generally

$$\mathcal{C}_\nu^{\text{SOP}}(\zeta_1, \zeta_2, \xi_1, \xi_2) = \mathcal{C}_\nu^{\text{SOP}}(Z_1 \cdot L_{\text{span}}, Z_2 \cdot L_{\text{span}}, \Xi_1 \cdot L_{\text{span}}, \Xi_2 \cdot L_{\text{span}}) \quad (4.15a)$$

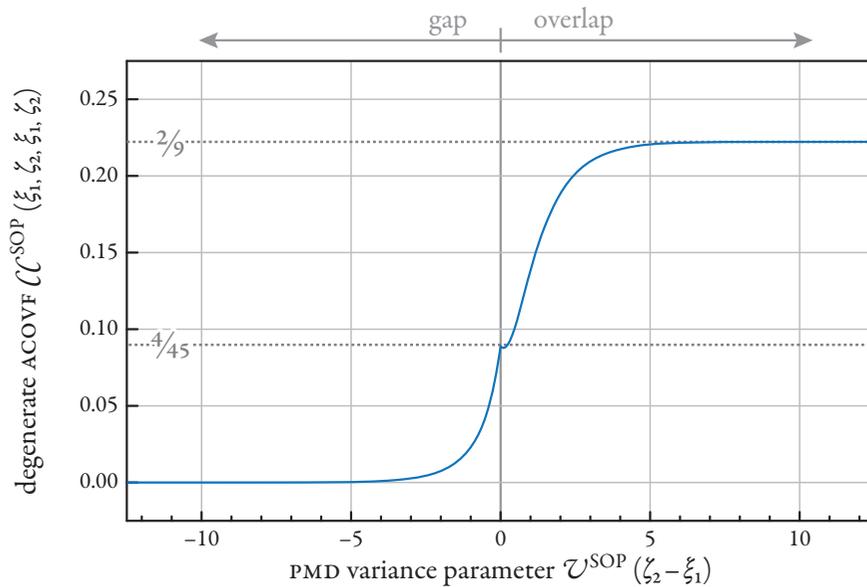


Figure 4.2 plots the autocovariance function $\mathcal{C}_v^{\text{SOP}}$ versus the interval length, given in terms of the accumulated PMD. Positive values of $\mathcal{T}_v^{\text{SOP}}$ correspond to the degenerate case in which $\zeta_1 = \xi_1$ and $\zeta_2 = \xi_2$ (full overlap of the intervals). Negative values correspond to $\zeta_1 = \zeta_2 < \xi_1 = \xi_2$ (no overlap). Also cf. Fig. 4.1.

with

$$Z_1 = \left\lfloor \frac{\zeta_1}{L_{\text{span}}} \right\rfloor \quad (4.15b)$$

where $\lfloor \cdot \rfloor$ is the floor operation, rounding down to the nearest integer. The expressions for Z_2 , Ξ_1 , and Ξ_2 are analogous. These variables essentially are span indices, starting from 0 in the first span, and ending at $N_{\text{span}} - 1$. We shall use the short form $\mathcal{C}_v^{\text{SOP}}(Z_1, Z_2, \Xi_1, \Xi_2)$ for the right-hand side of (4.15a) from here on. By using (4.15), the integral expression (4.6) reduces to a sum over the span indices,

$$\text{var}[V(L)] \approx \sum_{Z_1, Z_2, \Xi_1, \Xi_2=0}^{N_{\text{span}}-1} \sum_{\nu \neq \rho} \mathcal{C}_\nu^{\text{WO}}(Z_1, Z_2, \Xi_1, \Xi_2) \times \mathcal{C}_\nu^{\text{SOP}}(Z_1, Z_2, \Xi_1, \Xi_2) \quad (4.16a)$$

where

$$\begin{aligned}
 \mathcal{C}\mathcal{C}_\nu^{\text{WO}}(Z_1, Z_2, \Xi_1, \Xi_2) &= \int_0^{L_{\text{span}}} \int_0^{L_{\text{span}}} \int_0^{L_{\text{span}}} \int_0^{L_{\text{span}}} \Gamma(\zeta_1, \zeta_2) \Gamma(\xi_1, \xi_2) \\
 &\quad \times \mathcal{C}_\nu^{\text{WO}}(\zeta_1 + Z_1 L_{\text{RDPS}}, \zeta_2 + Z_2 L_{\text{RDPS}}) \\
 &\quad \times \mathcal{C}_\nu^{\text{WO}}(\xi_1 + \Xi_1 L_{\text{RDPS}}, \xi_2 + \Xi_2 L_{\text{RDPS}}) d\xi_1 d\xi_2 d\zeta_1 d\zeta_2
 \end{aligned} \tag{4.16b}$$

is a quadruple integral that runs only over the length of a single span.

Fig. 4.3 compares the results of our approximation to data from the numerical simulations. The match is remarkably good when considering the great number of approximations made throughout the present work to arrive at (4.16).³¹

We can see that, as hinted at in Fig. 3.17, the variance is significantly reduced in the presence of RDPS. However, even though the mean nonlinear depolarization may be similar (cf. Fig. 3.19), the variance scales differently with RDPS. This is elaborated in Fig. 4.4 which shows the reduction of the ratio \mathcal{R} of the standard deviation (STD) of V to its mean \mathcal{U} with an increase in RDPS – we use V and the standard deviation, because both the mean and the STD of V scale with the square of the mean interferer power (cf. Section 3.2.3), so that this power dependence is eliminated from \mathcal{R} .

In Fig. 4.4, we recognize that the high-PMD fiber initially has a higher relative STD \mathcal{R} . The reason is that the low-PMD fiber has almost constant interferer SOPs throughout all spans – the variance of its V parameter is mainly determined by the launch SOPs (cf. Fig. 4.2 for small values of $\mathcal{U}_\nu^{\text{SOP}}$), and this dependence cannot be removed by any amount of RDPS. Therefore, \mathcal{R} decreases only slightly with RDPS.

In the high-PMD fiber system, due to the faster linear motion of the interferer SOPs on the Poincaré sphere, the average correlation between the contribution to V of the various spans is lower, but its variance is much higher (cf. Fig. 4.2 for large values of $\mathcal{U}_\nu^{\text{SOP}}$), thus resulting in a large variance of V itself. Since increasing RDPS reduces this variance by deterministically decorrelating the single-span contributions, \mathcal{R} quickly decreases with RDPS in the high-PMD system.

Also, from the match between our predictions (which took only polarization effects into account) and the simulations we estimate that the influence of the actual bit patterns and the timing / phases in the interfering channels (which were not part of the model) is of minor consequence for the variance of the DOP distribution – this is in contrast to XPM, in which the variations are traceable to exactly those influences while polarization effects play a subordinate role [Weber 2010]. This has not been verified by simulation, however.

³¹In the light of the results in Fig. 4.3, one may reason that the double integral (3.30) to calculate \mathcal{U} could also be replaced by a double sum in the same fashion as in the present section. We leave such an investigation to the interested reader.

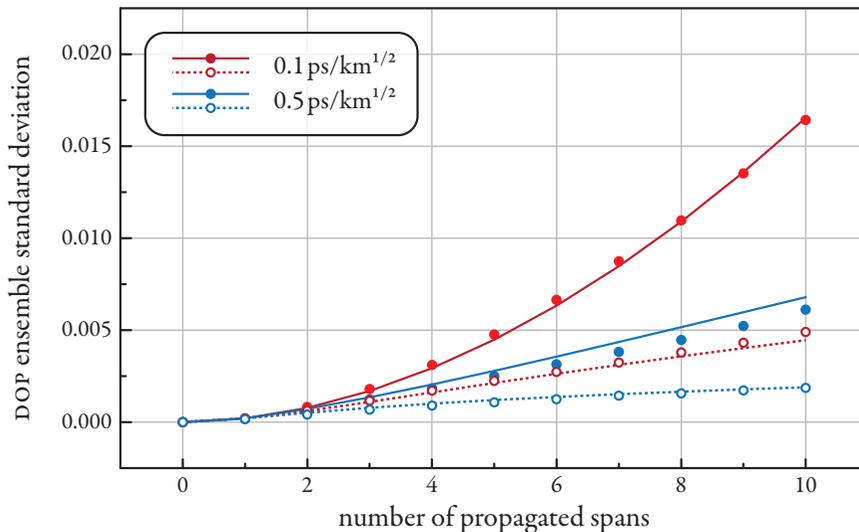


Figure 4.3 plots the evolution of the standard deviation with propagated spans, obtained from (4.16) (lines) and from numerical simulation ensemble averages (symbols). Solid lines and filled circles represent resonant dispersion maps, and dashed lines and open circles represent systems with RDPS of -2048 fs/GHz. All other parameters are those of Table 3.1.

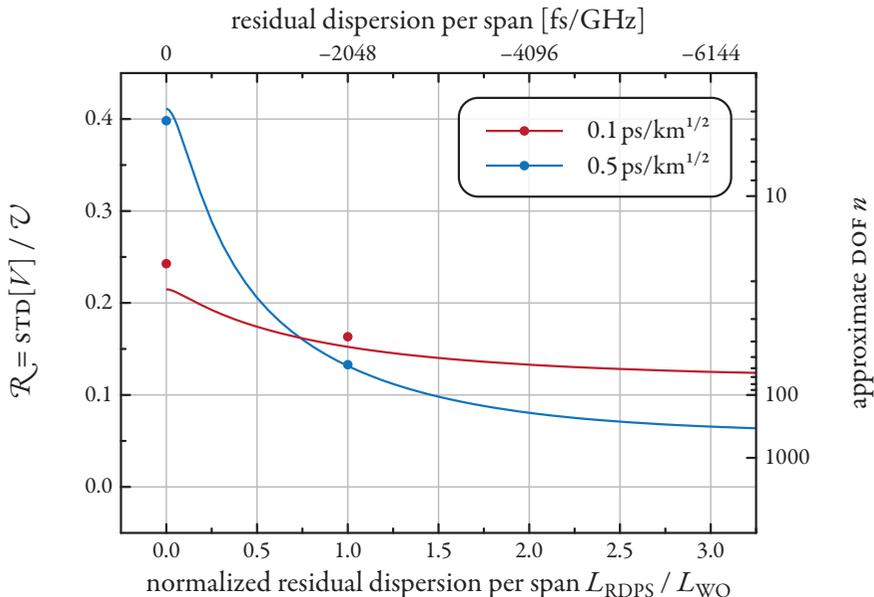


Figure 4.4 plots $\mathcal{R}_\sigma = \sqrt{\text{var}[V]} / \mathcal{U}$, versus the amount of RDPS for 10 spans of SSMF with $\beta_2 = -128$ fs/GHz/km and two different PMD values. The curves are based on the XPOLM model, in particular (3.30) and (4.16). Symbols combine the data from Figs. 3.10, 3.16, and 4.3. Axis quantity n on the right-hand side is discussed in Section 4.2.

4.2 THE DISTRIBUTION OF THE DOP

We have seen in Fig. 3.17 that the DOP distribution does not necessarily equal one of the more common statistical distributions in use. Thus, it will not be sufficient to know the variance of the DOP to estimate the probability of the DOP being below some lower limit of safe operation, whatever this particular limit may be (we shall discuss the concept of an outage and an associated outage probability within the next chapter). We thus need some way of determining, estimating or approximating the distribution of the parameter that describes the impact of XPOLM on our system, namely the nonlinear depolarization.

We do know the distribution of the SOP population – i.e. the SOPs of all members of the polarization ensemble – which is the Brownian parameterized by the mean DOP \mathcal{D} or the variance parameter \mathcal{U} , interchangeably. If we now consider the DOP of each ensemble member to be the sample mean of a number n of random samples Θ_i from this Brownian distribution,

$$D(z) = \frac{1}{n} \sum_{i=1}^n \cos \Theta_i(z) \quad (4.17)$$

corresponding to the expectation in (3.12), then $D(z)$ will have a variance related to the population variance $\text{var}[\cos \Theta]$ by

$$\text{var}[D(z)] = \frac{\text{var}[\cos \Theta(z)]}{n} \quad (4.18)$$

From the Brownian PDF we can determine the population variance as

$$\begin{aligned} \text{var}[\cos \Theta] &= \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} \cos^2 \theta p_B(\varphi, \theta, \mathcal{U}) \sin \theta d\theta d\varphi \\ &\quad - \left\langle \mathcal{E}[\cos \Theta(z, t)] \right\rangle^2 \end{aligned} \quad (4.19a)$$

$$= \frac{1}{3} + \frac{2}{3} \exp\left(-\frac{3\mathcal{U}}{2}\right) - \exp(-\mathcal{U}) \quad (4.19b)$$

We have analytically determined the parameter \mathcal{U} used above in Section 3.2 and $\text{var}[D]$ in the previous section, so that we can determine n via (4.18).

If we now approximate (4.19b) by its power series expansion up to the second-order term (since the variance is a squared quantity), we obtain

$$n \approx \frac{1}{\mathcal{R}^2} \quad (4.20)$$

where \mathcal{R} is the normalization used in Fig. 4.4, so that n is quasi-independent of \mathcal{U} and the optical power of the interferers (as long as \mathcal{U} remains reasonably small). Without the approximation for $\text{var}[\cos \Theta]$,

there is actually a residual dependence on \mathcal{U} – however, we have in several places approximated the exponential relationship between \mathcal{U} and the DOP to not higher than the linear term of its series expansion, so that this small residual dependence is not captured in the analytical model anyway. The difference in n between our approximation and the exact value is around 1 for DOPs near unity.

We will call n the stochastic *degrees of freedom* (DOF) of XPOLM for a particular system. In general, the higher the DOF of a system, the better will be the match to the quantities derived throughout the present work, because such a large DOF implies a low amount of correlation between the interferer Stokes vectors along the transmission path and ensures that the motion of the probe SOPs is sufficiently random. Until now we have hinted at this only by the somewhat diffuse concept of randomness – the DOF is a quantitative measure for this randomness.

As we have discussed in the previous section, we can see in Fig. 4.4 that systems with RDPS generally have a higher number of DOF than those without, because RDPS stochastically decorrelates the contribution of each interferer from one span to the next. This not only leads to a much narrower DOP distribution around the mean \mathcal{D} , but also to a better match of the SOP distribution to the Brownian and as a result the predicted DOP after some number of spans (cf. figures in Section 3.3, especially the histograms in Fig. 3.17).

Our treatment in (4.17) of the ensemble member DOP as sample mean from the SOP population with n degrees of freedom also points the way to an estimate – even though only a semi-analytical one – of the DOP distribution: by repeatedly taking n random samples from the population with parameter \mathcal{U} and averaging according to (4.17), we can quickly and numerically build the distribution of D with any required accuracy – this may even be sped up significantly using variance reduction methods such as importance sampling, et cetera.³² A fully analytical solution may possibly be obtained using the characteristic function method or something similar, but we shall defer such an analysis for now. We can state at the present that this distribution will approach the Gaussian distribution with an increasing number of DOFs, owing to the central limit theorem [Soong 2004, Sec. 9.1.1].

Fig. 4.5 compares the histograms of 500 simulations runs for four different combinations of fiber PMD and RDPS (the remainder of the parameters being those of Table 3.1) with our model, using the mean and variance of the simulation DOPs to determine the number of degrees of freedom and subsequently the PDF of the DOP distribution by means of a large number of Monte Carlo trials. The so obtained curves match the simulation data very well in almost all cases. The noticeable difference in the case of the resonant dispersion map using low-PMD fiber – the DOP is closer to the Gaussian than the predicted distribution – leads us to speculate that there occurs non-negligible additional nonlinear

³²To generate samples of an arbitrary distribution, we need to find the inverse cumulative density function (CDF) F^{-1} of that distribution, so that $x = F^{-1}(u)$ has the required distribution when u is uniformly distributed between 0 and 1. In the case of the Fisher distribution (3.16) with parameter \varkappa , we have

$$F^{-1}(u) = \arccos\left(\frac{1}{\varkappa} \ln\left(\sinh \varkappa \left[\frac{\exp \varkappa}{\sinh \varkappa} - 2u\right]\right)\right)$$

For the Brownian distribution, there exists no analogous closed-form expression, so that we use the Fisher distribution to generate our results.

depolarization due to FWM, which is not part of the present model.

We now have a method for not only determining the average nonlinear depolarization, but for semi-analytically estimating the probability with which a given transmission system will have a DOP below a certain bound. However, until now this boundary value is a somewhat arbitrary concept because we do not yet know if and how the nonlinear DOP reduction affects an optical transmission system. In the final chapter, we will look at the effects of XPOLM in terms of the quality of transmitted signals.

4.3 DISTRIBUTION OF THE SOPS WITHIN SINGLE ENSEMBLE MEMBERS

Before we investigate the effect of XPOLM on realistic data transmission systems, we should like to interpose some reflections on the distribution of SOPS within a single member of the polarization ensemble. In the previous section we have seen that the whole of the SOPS of an ensemble member can be represented – at least in determining their average – by a number n samples randomly chosen from the (global) ensemble distribution, where n is the number of stochastic degrees of freedom. Small n lead to large variations of the DOP, relative to its mean.

The Brownian distribution is described in (3.5) as a symmetric distribution on the Poincaré sphere. Randomly sampling from this distribution needs not result in a symmetric distribution, nor do the corresponding angles Θ of a finite-sized sample necessarily have to be Brownian distributed with mean \mathcal{D} . The larger the sample size n , however, the better the sample will approach the ensemble statistics. We want to see if this is true also for the SOP distributions observable in our simulations.

Figs. 4.6 and 4.7 show some examples of the SOP distribution for a system in which $n = 8$ and one with $n = 58$, respectively, which have a nearly equal mean nonlinear depolarization \mathcal{D} . Both figures plot the SOPS of the first 12 iterations performed for both ensembles, centered on the mean SOP. Clearly, the resonant system with fewer DOFs shows much greater variety, and asymmetry, as its wider DOP distribution suggests (cf. Fig. 4.5), while the distributions within the RDPS ensemble are quite homogeneous in shape and their resulting DOP.

Obviously, there are more distinct SOPS present than the number of DOFs suggests. However, statistically it does make no difference, as not all of the occurring SOPS are uncorrelated or independent. However, as a more detailed analysis of the sample distribution is beyond our means, we will be contented to point out this behavior. A better knowledge of the member statistics may also be helpful in a discussion of transmission penalties, as endeavored in the following chapter, to evaluate outage probabilities: Since any single real-world instance of a transmission system with certain polarization ensemble properties may be in some state – corresponding to an ensemble member – for an arbitrary amount of time, it would be useful to identify the probability with which particularly poor constellations (relating to transmission errors) occur.

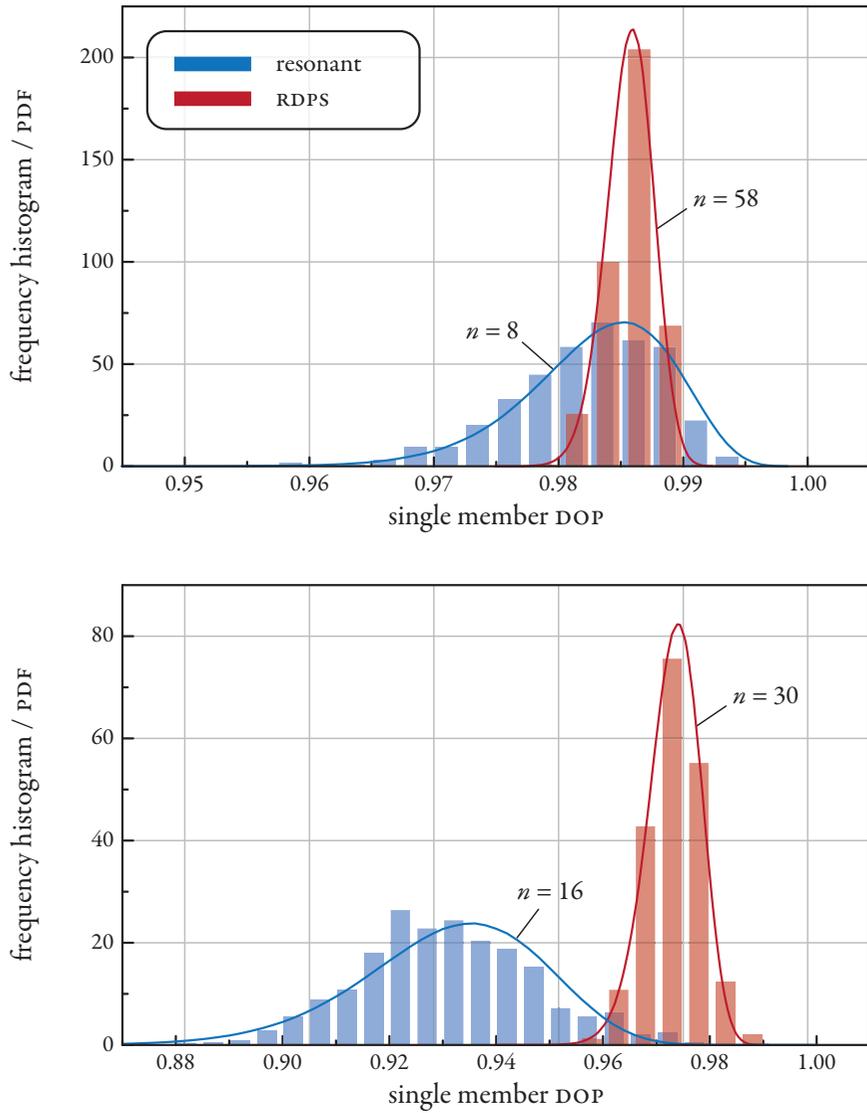


Figure 4.5 compares the DOP histograms obtained in 500 numerical simulation runs each with a resonant dispersion map and RDPS of -2048 fs/GHz for systems containing SSMF fiber with $D_{\text{PMD}} = 0.5 \text{ ps/km}^{1/2}$ (top) and $D_{\text{PMD}} = 0.1 \text{ ps/km}^{1/2}$ (bottom) to the semi-analytically obtained PDFs (the DOP mean and variance were extracted from simulation data to parameterize the PDFs). Note that the axes of both figure parts are scaled differently.

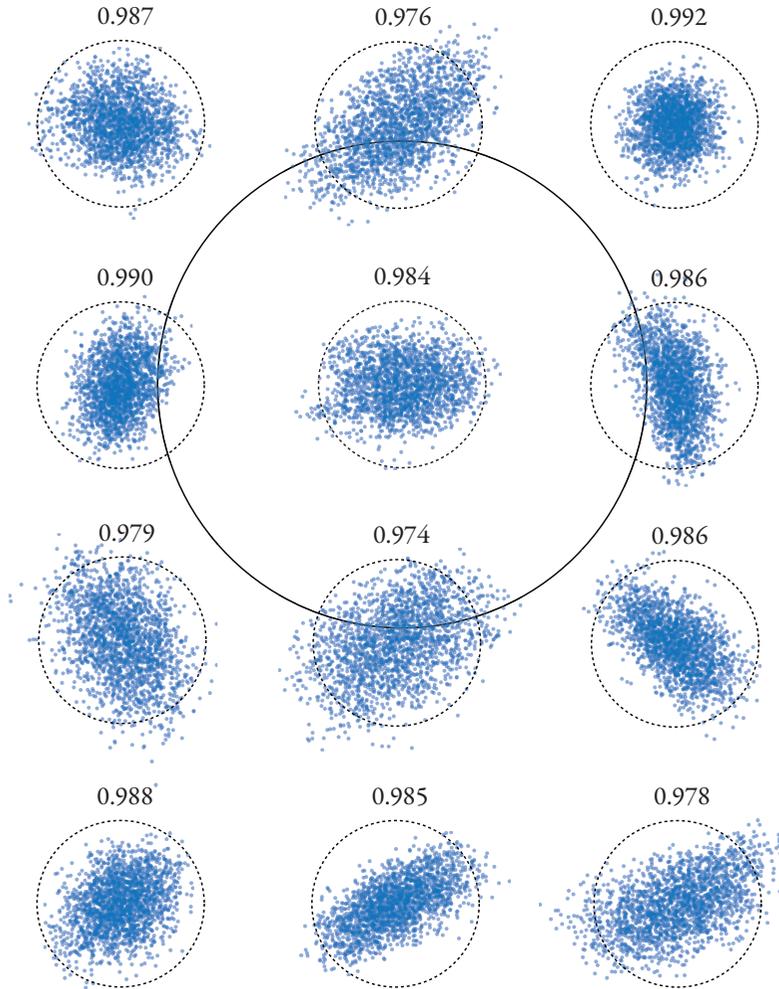


Figure 4.6 illustrates some exemplary SOP distributions within single members of the polarization ensemble in a system with high-PMD SSMF ($D_{\text{PMD}} = 0.5 \text{ ps/km}^{1/2}$, $\beta_2 = -128 \text{ fs/GHz/km}$, other parameters as in Table 3.1) with resonant dispersion map, in which DOF $n \approx 8$ and $\mathcal{D} = 0.982$: dashed circles show 20° deviation from mean SOP, the full Poincaré sphere is shown for system no. 5; numbers indicate the DOF for each ensemble member.

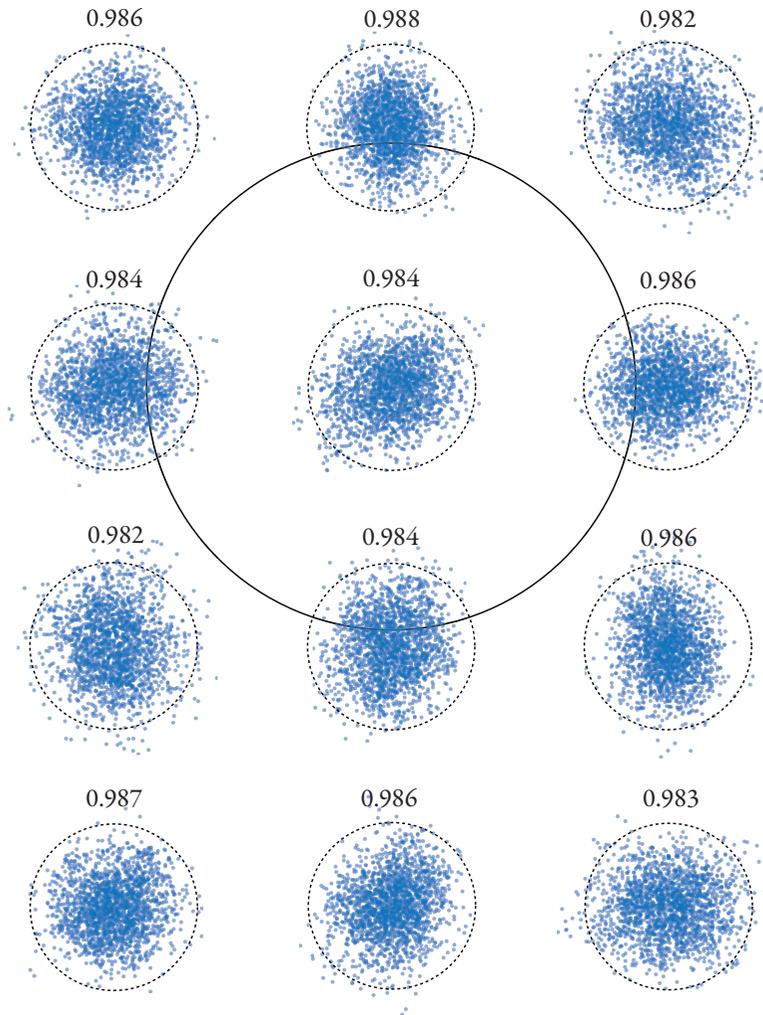


Figure 4.7 illustrates some exemplary SOP distributions within single members of the polarization ensemble in a system as in Fig. 4: SOP-distributions-n8, but with RDPS of -2048 fs/GHz, in which $\text{DOF } n \approx 58$ and $\mathcal{D} = 0.986$: dashed circles show 20° deviation from mean SOP, the full Poincaré sphere is shown for system no. 5; numbers indicate the DOP for each ensemble member.

WITHIN THE preceding chapters we have looked at XPOLM in optical communication systems exclusively in terms of the statistics of the nonlinear depolarization in a CW probe channel. This has provided us with extensive insights into this particular nonlinear process and how it is affected by the dispersion map of a system, the PMD values of the fibers used, the channel map, and the nature of the co-propagating channels. However, until now we can only quantify XPOLM in a meaningful way so as to compare its impact in various systems. We do not yet know how it affects optical communication over long distances, what amounts of XPOLM are permissible, and what amounts will likely lead to an increase in transmission errors.

The subject matter presented in this chapter is concerned with the effect of XPOLM-induced depolarization on various optical receivers. As such it depends on the statistics of the SOPs of the channel under consideration, the probe channel. However, as we have seen in the previous chapters, the nature of the SOP distribution is largely independent of the exact properties of the underlying transmission system. Only the nonlinear depolarization \mathcal{D} – and if the ensemble statistics become relevant, also the degrees of freedom n – depend on these properties, and these two quantities are sufficient to describe the consequences of XPOLM on the SOPs of the probe channel.³³ Hence, the theory of Chapters 3 and 4 is somewhat separated and independent (not in a statistical sense) from the present discussion.

XPOLM EFFECTS ON THE PMD VECTOR There are actually two mechanisms by which the nonlinear birefringence described in Section 2.3.3 and underlying XPOLM can affect an optical signal: In the preceding chapters we have described and quantified the effect of cross-polarization modulation and the resulting depolarization of an optical signal in great detail; however, the PMD vector corresponding to a WDM channel also depends on the local birefringence.

Our analysis of PMD assumed that linear birefringence had been compensated by an appropriate modification of the local coordinate system. In such a case, our results (2.91) and (2.92) for the PMD vector were particularly simple expressions. However, with the addition of a nonlinear birefringence term to the evolution equation (2.64), the PMD vector at each optical frequency will experience an additional nonlinear rotation in each subsequent segment, according to the instantaneous value of nonlinear birefringence in that segment, very similar to what happens to the probe SOPs. This results in a rapidly varying PMD vector at the fiber output which cannot be followed by PMD compensators. Compensation of only the mean PMD vector leaves the effects of residual PMD imparted onto the signal, which was analytically estimated in [Boroditsky 2006]. If the residual PMD becomes too large, the benefit of PMD compensation may be partially or completely lost [Karlsson 2006].

³³Actually, we shall see in Section 5.1 that there exists a third relevant quantity describing the rate of change of the probe SOP on very short timescales.

In conjunction with our analysis in the previous chapters, we will restrict ourselves herein to a discussion of the direct effects of XPOLM, i.e. the depolarization of the data signal itself. In order to observe any effect, an optical receiver (or possibly the modulation format itself) must in some way be polarization-sensitive – the effect on an NRZ probe using direct detection will not be observable. There are three major constellations in which XPOLM may become relevant: polarization-modulated signals or polarization-division multiplex (POLDM) as a special case thereof, optical differential detection, and any sort of polarization filtering at the receiver (this also includes coherent detection with a single-polarization local oscillator, as the orthogonal polarization is not detected). The latter is unlikely to appear in practice due to the unpredictability of the output SOP (cf. Section 2.2.2 on birefringence). Coherent systems usually employ polarization diversity to avoid the fading associated with SOP misalignment (cf. e.g. [Glance 1987] and [Hui 2008, Sec. 2.7.3]). We shall thus concentrate on the former two.

5.1 OPTICAL DIFFERENTIAL DETECTION

In differential detection, a delayed symbol of the data channel is used as reference signal (similar to a local oscillator) for another symbol of the same wavelength channel. This demodulation is performed with a so-called delay-line interferometer (DLI), illustrated in Fig. 5.1. The incoming signal is split into equal parts, one part is delayed, usually by the length of a symbol T_S , and then both parts interfere in the output directional coupler (DC). Mathematically,

$$\begin{pmatrix} |A_{\text{out},1}(t)\rangle \\ |A_{\text{out},2}(t)\rangle \end{pmatrix} = \bar{\mathbf{U}}_{\text{DC}} \bar{\mathbf{U}}_{\text{DL}} \bar{\mathbf{U}}_{\text{DC}} \begin{pmatrix} 0 \\ |A_{\text{in}}(t)\rangle \end{pmatrix} \quad (5.1a)$$

with

$$\bar{\mathbf{U}}_{\text{DC}} = \begin{pmatrix} \sqrt{1/2} & i\sqrt{1/2} \\ i\sqrt{1/2} & \sqrt{1/2} \end{pmatrix} \quad (5.1b)$$

$$\bar{\mathbf{U}}_{\text{DL}} = \begin{pmatrix} \exp(-i\omega_0 T_S) & 0 \\ 0 & 1 \end{pmatrix} \quad (5.1c)$$

being the unitary and polarization-independent transmission matrices of the directional coupler and the delay line, respectively. The Jones-space field $|A_{\text{in}}\rangle$ is the input into the first DC and the $|A_{\text{out},m}\rangle$ are the outputs of the DLI, ω_0 is the optical carrier frequency. The transmission matrices work on the two inputs and outputs of the optical elements rather than on the polarization components of the optical field. We then have

$$|A_{\text{out},1}(t)\rangle = \frac{i}{2} \left(|A_{\text{in}}(t)\rangle + |A_{\text{in}}(t - T_S)\rangle \right) \quad (5.2a)$$

$$|A_{\text{out},2}(t)\rangle = \frac{1}{2} \left(|A_{\text{in}}(t)\rangle - |A_{\text{in}}(t - T_S)\rangle \right) \quad (5.2b)$$

Both output signals are subsequently detected using a photodiode and then subtracted from one another in a configuration that is known as *balanced detection* [Winzer 2006]. The current into the decision circuit then is

$$I(t) = \frac{R}{2} \left(\langle A_{\text{in}}(t) | A_{\text{in}}(t - T_S) \rangle + \langle A_{\text{in}}(t - T_S) | A_{\text{in}}(t) \rangle \right) \quad (5.3)$$

where R is the responsivity of the photodiode, which we will assume to be unity for simplicity. We can now expand the single-polarization field as in (2.4) on page 28 and write

$$|A_{\text{in}}(t)\rangle = A_{\text{in}}(t) |e_1(t)\rangle \quad (5.4)$$

We now allow for a change in signal SOP within the time T_S , so that

$$|A_{\text{in}}(t - T_S)\rangle = A_{\text{in}}(t - T_S) |e_1(t - T_S)\rangle \quad (5.5)$$

The decision current then is

$$I(t) = F(t) \cdot |A_{\text{in}}(t)| \cdot |A_{\text{in}}(t - T_S)| \cdot \cos \Delta\varphi \quad (5.6)$$

with the real-valued fading factor F from (2.54) on page 45 and the field phase difference $\Delta\varphi$. We have shown in Section 2.2.2 that

$$F(t) = \cos \frac{\Delta\Theta(t)}{2} = \sqrt{\frac{1}{2} \left[1 + \hat{\mathbf{S}}_P(L, t) \cdot \hat{\mathbf{S}}_P(L, t - T_S) \right]} \quad (5.7)$$

when $\Delta\Theta$ is the misalignment angle in Stokes space. However, unlike Θ of Chapter 3, $\Delta\Theta$ here denotes the relative angle between the probe SOPs $\hat{\mathbf{S}}_P(L, t)$ and $\hat{\mathbf{S}}_P(L, t - T_S)$, and not the deviation from some mean SOP.

Expression (5.7) gives us the relation to determine the mean value and PDF of F if we know the PDF of $\Delta\Theta$. We now argue that the latter distribution is also Brownian, with a corresponding variance parameter $\mathcal{U}^{\Delta\Theta}$ (which will be a function of T_S): In (3.2) on page 75 we introduced a transformation $\bar{\mathbf{E}}$ which ensured that the mean probe SOP remained constant. Similarly, we could define an analogous transformation that fixes any probe SOP $\hat{\mathbf{S}}_P(z, \tau)$ in place for all z , we would obtain $\hat{\mathbf{S}}_P(L, \tau - T_S)$ as the result of a large number of random rotations around the Stokes vector difference $\mathbf{S}_\Sigma(z, \tau - T_S) - \mathbf{S}_\Sigma(z, \tau)$, which would according to Section 3.1 have a Brownian PDF. As is shown in Fig. 5.2, the distribution of $\Delta\Theta$ obtained in numerical simulations is indeed approximately Brownian. Again using the Fisher distribution to approximate the Brownian, we have with (5.7),

$$\langle \mathcal{E}[F] \rangle \approx \frac{1}{\sqrt{2}} + \frac{1}{2\sqrt{2}} \mathcal{C}^{\Delta\Theta}(T_S) \quad (5.8a)$$

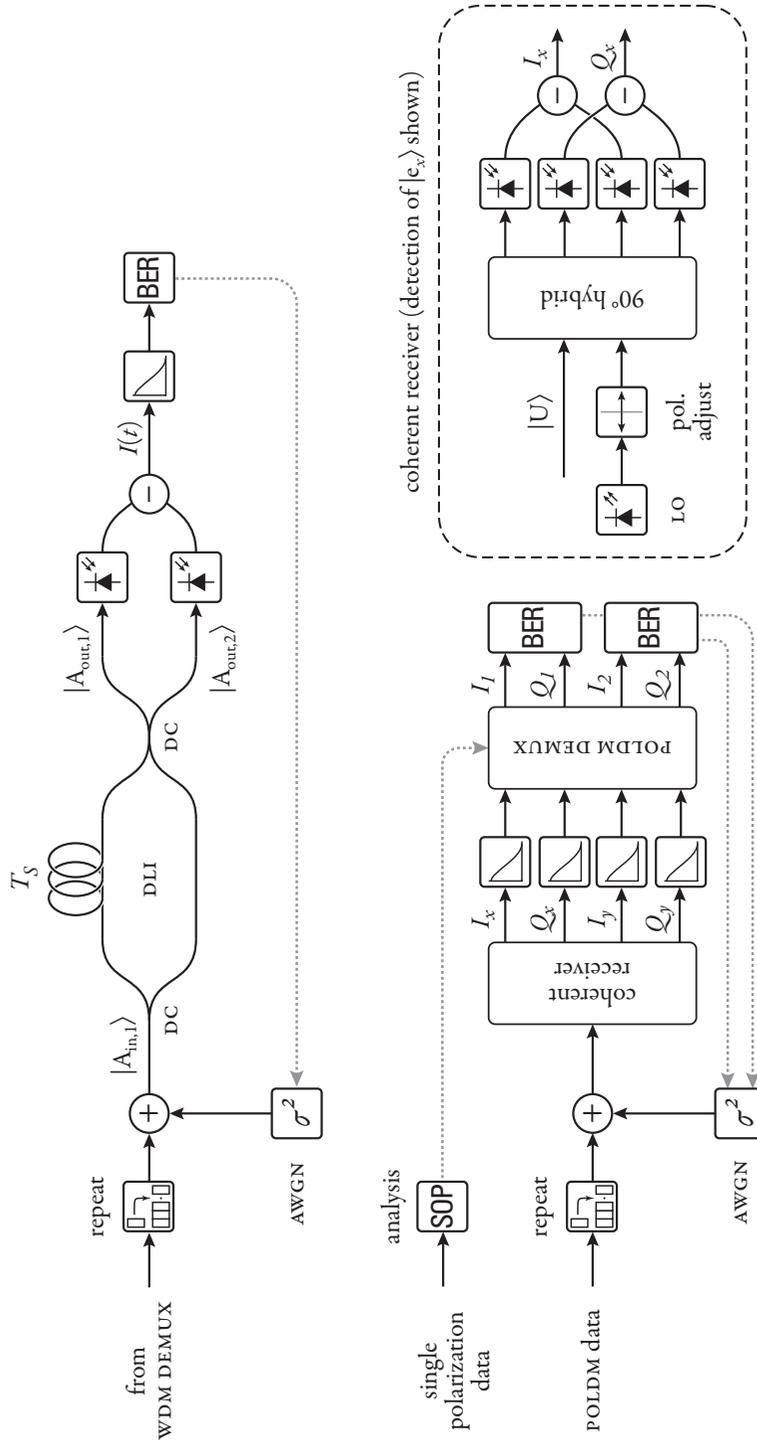


Figure 5.1 schematically illustrates the receivers for DPSK reception (top) and coherent polarization-diversity detection (bottom). Coherent receiver for $|e_y\rangle$ differs only in the alignment of the *polarization adjust* and may use same LO source. Coherent detection is used and explained in Section 5.2. Both receivers are described in more detail in the text.

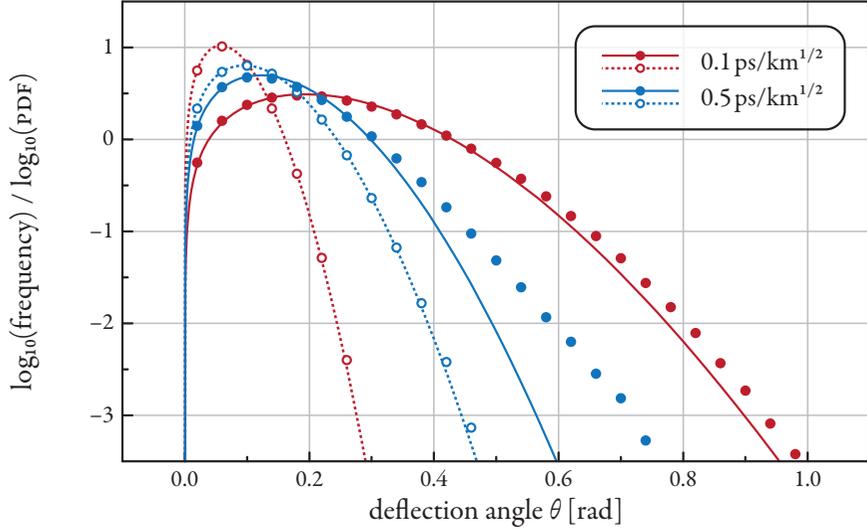


Figure 5.2 shows the $\Delta\Theta$ distribution for four exemplary systems; symbols represent simulation values (using 500 randomly generated system iterations) and lines are (linear scale) curve fits to the Fisher distribution (3.16) (open symbols / dashed lines: RDPS of -2048 fs/GHz; filled symbols / solid lines: no RDPS).

where

$$\begin{aligned}
 \zeta^{\Delta\Theta}(T_S) &= \left\langle \mathcal{E}[\hat{\mathbf{S}}_p(L, t) \cdot \hat{\mathbf{S}}_p(L, t - T_S)] \right\rangle \\
 &= \left\langle \mathcal{E}[\cos \Delta\Theta] \right\rangle \\
 &= \exp\left(-\frac{\mathcal{V}^{\Delta\Theta}(T_S)}{2}\right)
 \end{aligned} \tag{5.8b}$$

is the ensemble-averaged autocorrelation function of all probe SOP pairs which are spaced apart T_S in time. The final equality in (5.8b) serves as the definition of the variance parameter $\mathcal{V}^{\Delta\Theta}$, whose relation to $\Delta\Theta$ is the same as that of \mathcal{V} to Θ in (3.13). We have used

$$\cos \frac{x}{2} = \sqrt{\frac{1 + \cos x}{2}} = \sqrt{1 - \frac{1 - \cos x}{2}} \approx 1 - \frac{1 - \cos x}{4} = \frac{3}{4} - \frac{\cos x}{4}$$

for small arguments x of the cosine in (5.8a) to give the expectation in closed form. The distribution of F can be derived via (5.7) and the Brownian PDF (3.5). It is shown in Fig. 5.3 for various values of $\mathcal{V}^{\Delta\Theta}$. We can see that small fading coefficients (corresponding to strong fading) appear with significant probability only for moderate values of $\mathcal{V}^{\Delta\Theta}$, with approximately $\zeta^{\Delta\Theta}(T_S) < 0.9$.

Thus, to estimate the influence of XPOLM on DPSK reception we need to determine the average correlation between the SOPs at successive sampling instants, i.e. spaced apart by time T_S . Because the fading appears to be rather small, and we have seen in Section 2.2.2 that compared to polarization

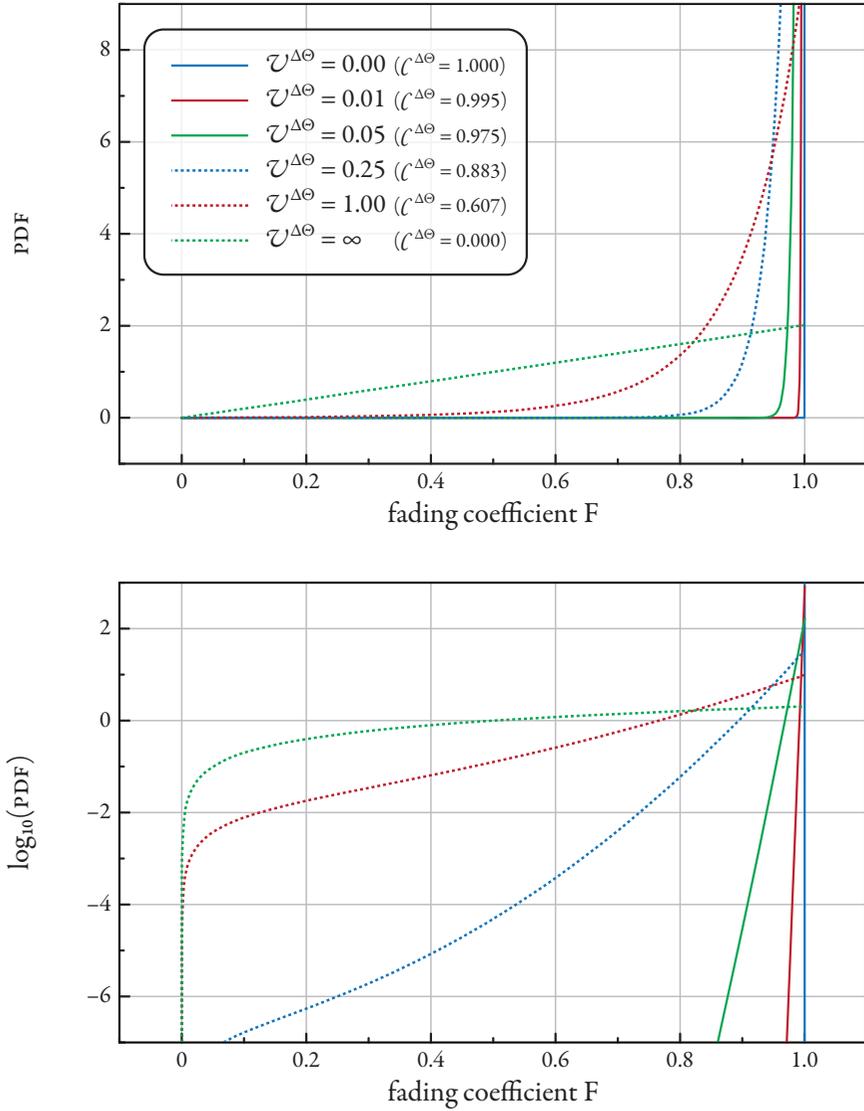


Figure 5.3 shows the population distribution of the fading coefficient F for different values of the variance parameter $\mathcal{V}^{\Delta\Theta}$ respective the ACF of the Stokes vectors in linear (top) and logarithmic (bottom) scale.

misalignment crosstalk it is almost negligible, we will estimate the fading significance with exemplary numerical simulations to assess if we need to dive into the required mathematics to determine $\mathcal{V}^{\Delta\Theta}$ analytically from the Stokes vector statistics.

Because we have already done extensive numerical simulations (to verify our model in Chapters 3 and 4) we can make use of data obtained there for our estimate. In order to do this, we assume our CW probe channel to be a DPSK-modulated channel consisting only of “0” data bits, which corresponds to a constant-phase and -amplitude optical signal at the modulator output. Since we are not interested in penalties related to PMD, GVD, or SPM, this approximation should affect the result very little. Fortunately, the decision threshold at the DPSK receiver is independent of the optical signal, always being at $I_{\text{th}} = 0$, and the DPSK eye diagram is approximately symmetric with respect to the threshold at the sampling instants. Therefore it should not matter significantly in terms of the bit error ratio (BER) if the decision current is unipolar or bipolar, and we can obtain an error estimate given only the upper rail of the «eye». Fig. 5.4 compares the eye diagrams in both cases. At equal average optical power, the required optical signal-to-noise ratio (ROSNR) value of the pseudo-random DPSK signal is lower than that of the all-zero sequence, because the optical filters at the transmitter cause the eye to «open» in the center of the symbol slot. However, the distortions after propagation of 10 amplifier spans are comparable, resulting in similar penalties.

REQUIRED OSNR We quantified the results in Fig. 5.4 using the required optical signal-to-noise ratio to achieve some pre-defined BER. The ROSNR allows to separate the linear and nonlinear signal distortions acquired during propagation from the stochastic effects of amplifier noise. When amplifier noise is included in the numerical propagation of the optical signals, the particular noise samples – randomly chosen at the input – can significantly affect the signal fidelity at the output and may distort or veil other effects. It is also difficult for system operators to estimate their operating margin without great computational cost, because a large number of samples is needed for accurate results. The ROSNR is a measure purely of the signal distortions that are not noise-related, and the difference between the required OSNR value and the achievable OSNR, which is a function of the input power, span loss, and amplifier noise figures.³⁴ gives operators a measure of operating margin. To determine the ROSNR value in computer simulations, a signal is propagating through an optical system which is identical to the system of interest but whose optical amplifiers add no noise to the signal, i.e. the amplifiers have a noise figure of unity. Instead, all optical noise is added at the output of the transmission fiber – the signal is being *noise-loaded*. The amount of noise added can be varied arbitrarily, and one usually aims to find the noise level in terms of the OSNR such that after demodulation, detection, and filtering a certain target BER is achieved. Be-

³⁴The achievable OSNR in dB in a multi-span optical transmission system is given by [Zyskind 1997]

$$\text{OSNR} \approx 58 + P_{\text{out}} - L - NF - 10 \log_{10} N$$

where P_{out} is the output power per channel in dBm, L is the span loss, NF is the amplifier noise figure (both in dB), and N is the number of amplifier spans. The 58 dB constant corresponds to the shot noise limit normalized to the 0.1 nm reference bandwidth of the OSNR [Krummrich 2002].

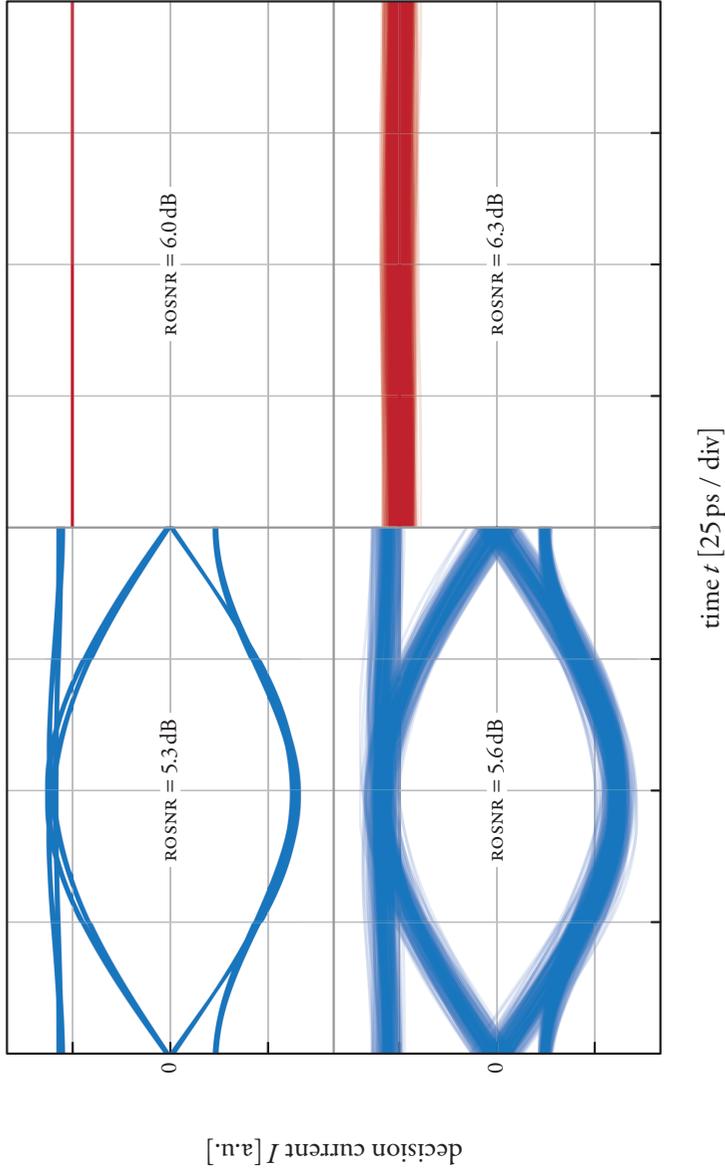


Figure 5.4 illustrates the different eye diagrams for pseudo-random DPSK (left) and all-zero DPSK (right), the latter of which is used to estimate inter-channel nonlinear penalties in a DPSK channel. Probe channel input power was -10 dBm in both cases, to avoid SPM effects. Shown are back-to-back eye diagrams (above) and signals after propagation of 10 spans in the system of Table 3.1. Interferer input power was 3 dBm.

cause contemporary error-correction mechanism can compensate bit error ratios of greater than 10^{-3} , we use this (arbitrary) value as our target BER [Djordjevic 2008].

Often, the signal distortion is quantified as an ROSNR *penalty*. This penalty is computed relative to the back-to-back ROSNR value, i.e. the signal having undergone the various multiplexer (MUX) and DEMUX filtering processes, demodulation, etc., without having actually been transmitted over an optical fiber. In this way, we can further separate the actual distortion accumulated within the transmission fiber from the modulation format or transmitter implementation. By plotting the ROSNR (penalty) versus the input power – of either the probe, the interferers, or both – we can additionally distinguish the linear penalties due to GVD and PMD, which will be visible at all power levels, from the nonlinear penalties, which will rise sharply after a certain threshold power is exceeded. This nonlinear threshold is usually defined as the input power (of the interfering channels in the present case) at which the ROSNR penalty reaches 1 dB – in contrast to Section 3.3, in which we designated the power at which $\mathcal{D} = 0.97$ as the NLT. Both are conceptually very similar.

Models exist to semi-analytically determine the ROSNR of a given optical signal, taking into account optical and electrical filtering, based on the noise PDFs at the decision instant (cf. e.g. [Forestieri 2000]). These models work reasonably well for NRZ-modulated signals, with adaptations for DPSK available. However, at the high BERs that we are considering, simple Monte-Carlo error counting becomes feasible. Its computational cost is approximately equal, but it is much more straightforward to implement and thus less prone to errors. In order for the ROSNR interpolation to be sufficiently accurate, we average the BER over 128 noise iterations per output signal / OSNR value combination, each consisting of 2^{10} symbols. This results in an accuracy for the so-obtained ROSNR to within approximately 0.1 dB.³⁵ Fig. 5.1 also schematically illustrates the process used to determine the DPSK ROSNR value. The power of the added noise (additive white Gaussian noise) is adapted such that the BER equals 10^{-3} . The ROSNR then is the ratio of average signal power to total noise power (in both polarizations, normalized to 0.1 nm bandwidth).

There are, of course, some disadvantages to avoiding amplifier noise within the transmission fiber. Foremost, it is not possible in this way to evaluate any signal-noise interaction within the fiber. One effect that is often cited in this context is nonlinear phase noise, also called Gordon-Mollenauer effect after its discoverers [Gordon 1990, Ho 2003]. Especially relevant for phase-modulated signals, it describes the accumulation of a nonlinear phase shift over multiple amplifier spans which depends on the instantaneous amplitude of the amplifier noise imposed on the signal. However, it would seem that this effect is so small as to be difficult to observe in the field, at least in high-speed DPSK systems [Borne 2007c]. We have shown in Section 3.2.5 that the effect of optical noise on the statistics of XPOLM is negligible. We are thus confident that our results in terms of the ROSNR (penalty) will give a satisfactory quantification of the interchannel distortions experienced by our probe.

³⁵The standard deviation of the ROSNR values is about 0.05–0.08 dB, as obtained by determining the ROSNR value multiple times for the same data set.

NUMERICAL RESULTS Fig. 5.5 shows the ACF $\mathcal{C}^{\Delta\Theta}(T_S)$, which quantifies the average correlation of the SOPs separated by a time interval T_S , for various exemplary system setups with and without RDPS. We notice that in all cases the ACF more or less rapidly approaches an asymptotic value, which corresponds approximately to the square of the DOP of the signal. The rate of this approach depends on the similarity of the infinitesimal nonlinear polarization rotations $\partial_z \hat{\mathbf{S}}_P(t)$ that adjacent symbols indexed by t and $t - T_S$ experience: mostly, it is dominated by the rate of the walk-off of the neighboring channels. The more probe symbols – where probe «symbols» refer to time slots of length T_S since the CW signal does not have an intrinsic separation into symbols – each interferer pulse passes during propagation, the longer the probe SOPs will be correlated. Hence, the initial rate of decline of the ACFs becomes steeper when fiber GVD becomes smaller. Without RDPS, the magnitude of fiber PMD plays only a minor role in determining the time shift over which the ACF reaches its asymptotic value. However, when RDPS is used, the number of passed probe symbols is greatly increased, as the maximum accumulated dispersion becomes much larger. The distortion originating in a certain interferer pulse affects one (or more) probe symbol in one span, the neighboring symbol in the following span, and so forth. This can lead to long correlation times of the probe SOPs. However, in high-PMD fibers, the nonlinear polarization rotations in different spans (originating in the same interferer symbol) are again statistically uncorrelated because the SOP rotation axis (the Stokes vector sum) is randomized from span to span. Thus, the slower decline of $\mathcal{C}^{\Delta\Theta}$ in the presence of RDPS is mostly negated.

The symbols in Fig. 5.5 highlight the values at $T_S = 100$ ps, the relevant quantity for differential reception of 10 Gbaud optical signals. We have seen in Fig. 5.3 that small fading coefficients do not occur with significant probability for approximately $\mathcal{C}^{\Delta\Theta}(T_S) > 0.9$. Even the smallest observed value of $\mathcal{C}^{\Delta\Theta}(T_S)$ greatly exceeds this threshold, though the penalties due to XPM in a resonant system with interferers at $\mathcal{E}[P_v] = 2$ mW as used in Fig. 5.5 are sufficient to make a reception with a BER smaller than 10^{-3} all but impossible. Similarly to the DOP in Chapter 4, the ACF $\mathcal{C}^{\Delta\Theta}$ has a finite-width distribution about its mean value – two such exemplary distributions are shown in Fig. 5.6 – as a result of the individual SOP evolution in each polarization ensemble member. Again, the variance of this distribution is much larger when the dispersion map is resonant. Overall, there is a strong similarity to the corresponding DOP histograms for the same pair of systems as shown in Fig. 3.17 (both are derived from the same set of simulation data).

However, even when the average ACF is comparatively low, the individual member ACFs are still unlikely to cross the (somewhat arbitrarily chosen) threshold of 0.9, and then only for systems with a resonant dispersion map in which the NRZ interferer power is decreased below 2 mW due to XPM. Therefore, from the SOP ACF we do not expect the fading due to XPOLM to play a significant role in DPSK transmission.

To verify this, we adjust the interferer power such that the average ROSNR penalty (mainly due to XPM and XPOLM) equals approximately 1 dB, corresponding to a typical system operating margin allowance for interchannel nonlinearities. We call the particular interferer power at this point the nonlinear threshold for DPSK transmission. It is shown in Fig. 5.7 where the mean ROSNR penalty for the polarization ensemble (obtained by averaging the ROSNR of 100 iterations) is plotted versus the DWDM

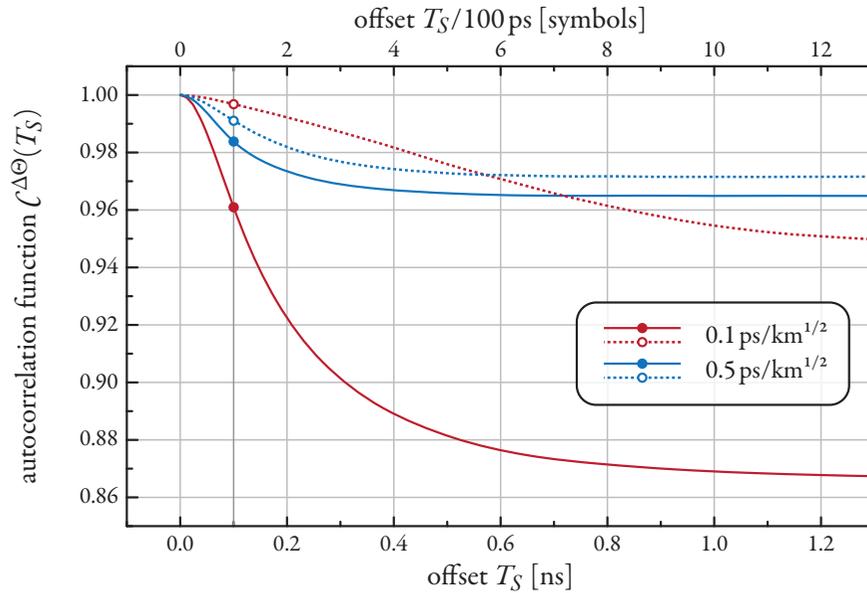


Figure 5.5 shows the numerically obtained ACF $C^{\Delta\Theta}(T_S)$ of the SOPs of the probe channel, defined in (5.8b) for different example systems with parameters from Table 3.1. Solid lines denote resonant dispersion maps and dashed lines denote systems with RDPS of -2048 fs/GHz . Symbols highlight the values at $T_S = 100 \text{ ps}$, which are most relevant for the 10 Gbps DPSK probe under consideration. Interfering channels were launched with 3 dBm optical power per channel.

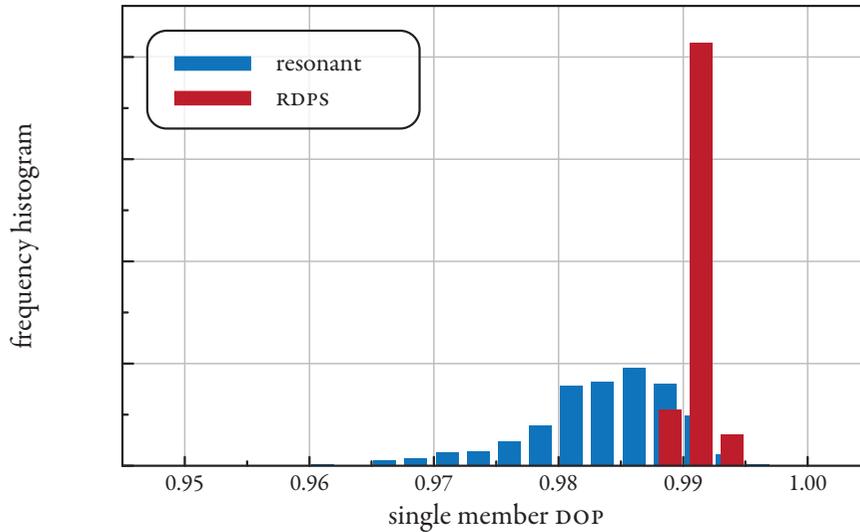


Figure 5.6 shows ensemble histograms of the ACF $C^{\Delta\Theta}(T_S = 100 \text{ ps})$ for the systems of Fig. 5.5 with $D_{\text{PMD}} = 0.5 \text{ ps/km}^{1/2}$.

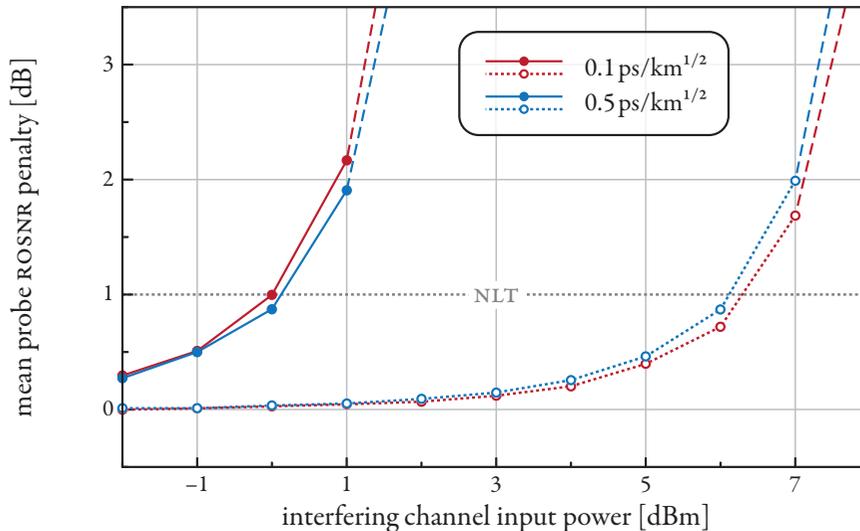


Figure 5.7 plots the mean probe channel ROSNR penalty versus the interferer power for the DPSK system of Fig. 5.1, as obtained by numerical simulations. Dashed lines and open symbols describe systems with RDPS, solid lines and filled symbols describe systems with a resonant dispersion map. The back-to-back reference ROSNR for the penalty is 6 dB (cf. Fig. 5.4). Ensemble size was 100.

interferer input power. We subsequently plot the individual ROSNR penalties at the (approximate) NLT versus the ACF of the SOPs for each member of the simulated ensemble. This allows us to visually determine any dependence of the ROSNR on the instantaneous value of $\zeta^{\Delta\Theta}$, which we used to compress the relevant SOP statistics into a single number, within and between the various polarization ensembles.

As can be seen in Fig. 5.8, and as predicted, there exists no such dependence, at least within each polarization ensemble. This can also be mathematically verified using e.g. Pearson's correlation coefficient [Pearson 1920, Winter 2009a]. However, comparison of Figs. 5.7 and 5.5 shows a small dependence of the mean ROSNR penalty on the mean ACF of the SOPs: Both, for the resonant systems and those with RDPS, the PMD value that results in a lower ACF $\zeta^{\Delta\Theta}$ also results in a slightly higher ROSNR penalty (at the same interferer power). However, this dependence is small because the ROSNR penalty in these DPSK systems is dominated by the effects of XPM. It is difficult to imagine a set of system parameters for which XPOLM may become an effect dominating the ROSNR for DPSK transmission.

Furthermore, we notice that the variance within the ROSNR penalty distribution is larger for resonant systems and, within those, larger for the system with the lower PMD fiber. Since the statistics of XPM do not depend on PMD, one may be tempted to attribute this behavior to XPOLM (cf. Fig. 5.6). However, since these fluctuations exhibit no dependence on the instantaneous value of $\zeta^{\Delta\Theta}$, we propose that these fluctuations be due to four-wave mixing. Since the magnitude of the mixing product depends on the relative SOPs of the participating channels, as described in (2.105), the fast SOP evolution due to

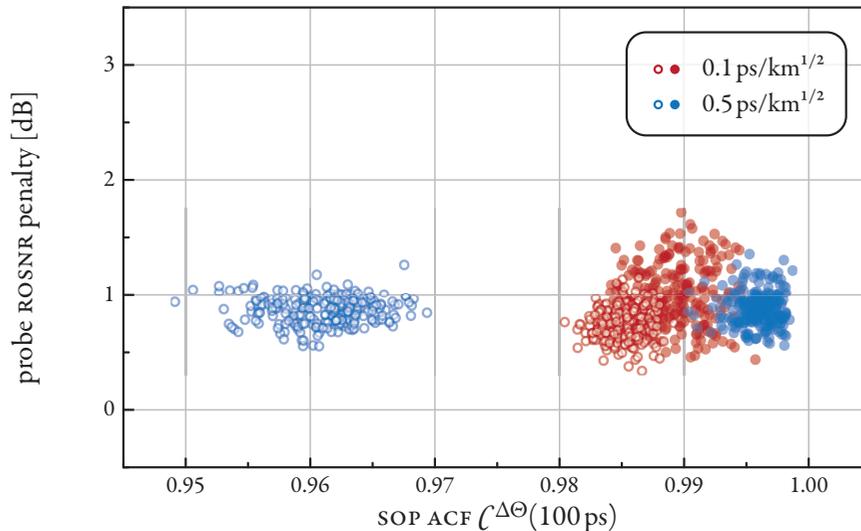


Figure 5.8 plots the individual ensemble member probe channel ROSNR penalty versus the ACF of the SOPs for the DPSK system of Fig. 5.1, as obtained by numerical simulations. Open symbols describe systems with RDPS (interferer input power: 6 dBm), filled symbols describe systems with a resonant dispersion map (interferer input power: 0 dBm). The back-to-back reference ROSNR for the penalty is 6 dB. Ensemble size was 250.

high values of PMD will lead to certain averaging and thus reduced variance, absent in low-PMD fibers.³⁶

The probe channel modulation assumed within the present section was characterized by a symbol rate of 10 Gbaud. The current state-of-the-art in transmitter and receiver technology permits phase-encoded transmission at up to 40 Gbaud using commercially available equipment. System operators seeking to upgrade their legacy 10 Gbps ASK with individual high-rate DPSK channels will be interested if the above results also hold at higher symbol rates. Referring to Fig. 5.5, we see that as the offset T_S between symbols becomes shorter, the ACF of the SOPs increases – the distortion due to XPOLM-related depolarization decreases. We can verify this easily, since the CW signal is suitable to be interpreted at any data rate. If it is decoded at a symbol rate of 40 Gbaud, the back-to-back ROSNR reference increases to ≈ 11.9 dB. The nonlinear thresholds (at 1 dB ROSNR penalty) for low- and high-PMD fiber in a dispersion map with RDPS are 8.4 and 8.3 dBm, respectively. This is approximately 2 dB above the value for 10 Gbaud, showing that XPM distortions are also decreased because the differential nonlinear phase shift between neighboring symbols is also reduced in consequence of the faster walk-off. Their difference between the NLTs has decreased slightly as a result of the rate change. Even though this is only anecdotal evidence, it confirms the predictions of the underlying theory.

Concluding, we determine that the time-dependent SOPs in the XPOLM-affected probe lead to fading in DPSK receivers. However, this fading is small enough not to be significant, compared with the

³⁶Consequently, the penalty variance obtained with scalar simulation of the NLSE, corresponding to zero PMD, gives only an upper margin.

distortion experienced as a consequence of XPM, and can therefore be neglected in margin allotments. Furthermore, as long as no other polarization effects (e.g. PMD) are of importance, it will be sufficient for simulations of such systems to be performed in a scalar instead of vectorial manner.³⁷

We obtain quite different results concerning the significance of XPOLM for polarization-division multiplex transmission, which is one of the prime candidates in the quest to increase the bit rate per DWDM channel and to raise the efficiency with which the usable spectrum offered by optical amplifiers is used. All but required to implement the current evolution of optical Ethernet (at 100 Gbps per 50 GHz wavelength channel), it is the topic of the following section.

5.2 POLARIZATION-DIVISION MULTIPLEX

We have seen in the previous section that the nonlinear polarization rotations from one symbol to the next are generally not large enough to generate XPOLM-related penalties in differentially encoded data signals. In polarization-division multiplex (POLDM) systems on the other hand, the relevant SOP is not that of the previous symbol, but the mean SOP of the incoming signal. In terms of the description of POLDM in Section 2.2.2, the receiver SOP $\hat{\mathbf{S}}'_1 = \mathcal{E}[\hat{\mathbf{S}}_1]$. Thus, the misalignment angle Θ in (2.57) is described by the Brownian distribution derived and quantified in Chapter 3. As such the angular excursions of the symbol SOPs may be significantly larger than in the differential detection case, causing significant fading and crosstalk, as derived in Section 2.2.2. Also, leveraging the results of the previous section, large SOP excursions may persist over the duration of multiple symbols, leading to long error bursts. Indeed, Borne et al. have shown as early as in 2004 that POLDM transmission systems suffer from XPOLM-related excess transmission penalties [Borne 2004], and others have only recently published experimental results on a variety of transmission systems (e.g. [Bertran-Pardo 2008a, Bertran-Pardo 2008b, Nelson 2009]). We shall devote the present section to an investigation of possible penalties in polarization-division multiplex systems induced by XPOLM.

We have shown in Section 2.2.2 that symbol-wise deviations from the mean SOP do result not only in fading, but also crosstalk between the POLDM subchannels. In single-polarization transmission, this crosstalk is irrelevant, because the receiver is not polarization-sensitive or the orthogonal subchannel is discarded. The impact of the crosstalk generally exceeds that of fading, which was shown to be negligible in the previous section. In the present analysis we will therefore neglect the fading and concentrate on the crosstalk – even though whenever the crosstalk X in (2.55) will be large, the fading F will also be more pronounced – in both POLDM subchannels – further exacerbating the crosstalk impact on the already impaired subchannel.

Given the angular SOP deviation Θ on the Poincaré sphere, we can use (2.54) and (2.55) from

³⁷However, doing so will exaggerate the nonlinear phase change due to XPM, since usually the reduced nonlinearity coefficient from (2.101) is not taken into account, and also because the phase shift due to XPM is larger than that due to SPM only by a factor of 3/2 in the Manakov approximation (2.108) instead of 2 which appears in the scalar NLSE.

page 45 to write the crosstalk field as

$$U_{21}(t) = X(t) U_2(t) = |U_2(t)| \cos \frac{\Theta(t)}{2} \exp(i\psi_2 + i\psi_X) \quad (5.9)$$

where U_{21} is the crosstalk from subchannel 2 into subchannel 1 generated during the POLDM demultiplex. The phase $\psi_2 + \psi_X$ of this field remains indefinite: it depends on the phase ψ_2 of the optical field in the interfering subchannel and the particular alignment of the Stokes vectors. As was shown in Section 2.2.2, a symmetric distribution of the Stokes vectors around some fixed mean will result in a uniform distribution of the crosstalk phase ψ_X . Simulation results confirm this assumption, at least in the sense of the polarization ensemble, in which the SOP population is symmetric in Stokes space, as shown exemplarily in Fig. 5.9 (top).

The PDF of the crosstalk magnitude $|X| = \sin \Theta/2$ population can in principle be derived with the distribution function (3.5) of the angle Θ using standard methods of probability theory. If we again replace this PDF with the simpler Fisher distribution (3.16), we can obtain a closed form solution:

$$p_F(|X|, \kappa) = 4\kappa N(\kappa) |X| \exp(-2\kappa |X|^2) \quad (5.10a)$$

with the normalization parameter

$$N(\kappa) = \frac{\exp \kappa}{2 \sinh \kappa} \quad (5.10b)$$

The above distribution is the Rayleigh distribution truncated at $|X| = 1$ – i.e. the crosstalk cannot be greater than its source. The normalization parameter ensures

$$\int_0^1 p_F(|X|, \kappa) d|X| = 1 \quad (5.11)$$

For $\kappa > 4$ we have $N(\kappa) \approx 1$ and (5.10) is properly Rayleigh-distributed. Figure 5.10 shows the distribution for different values of the DOP, after it has been converted to the respective κ using (3.17). Crosstalk coefficients greater than 0.2 occur with significant probability even at relatively high DOP values.

With our assumption of a uniformly distributed phase ψ_X of the crosstalk coefficient, we can deduce from the approximate Rayleigh distribution of the coefficient magnitude that the distribution of both components (in-phase and quadrature) of the complex coefficient X themselves must be Gaussian, each with zero mean and variance $(4\kappa)^{-1}$. To determine the properties of the crosstalk field U_{21} from (5.9), we also need the distribution of $|U_2|$ which is statistically independent of $|X|$, since their variations will have (mostly) separate causes.³⁸ At the sampling instants, U_2 can be approximated by a complex

³⁸Part of the amplitude variations of U_2 will be due to XPM via PM-AM conversion, which will be related to the XPOLM

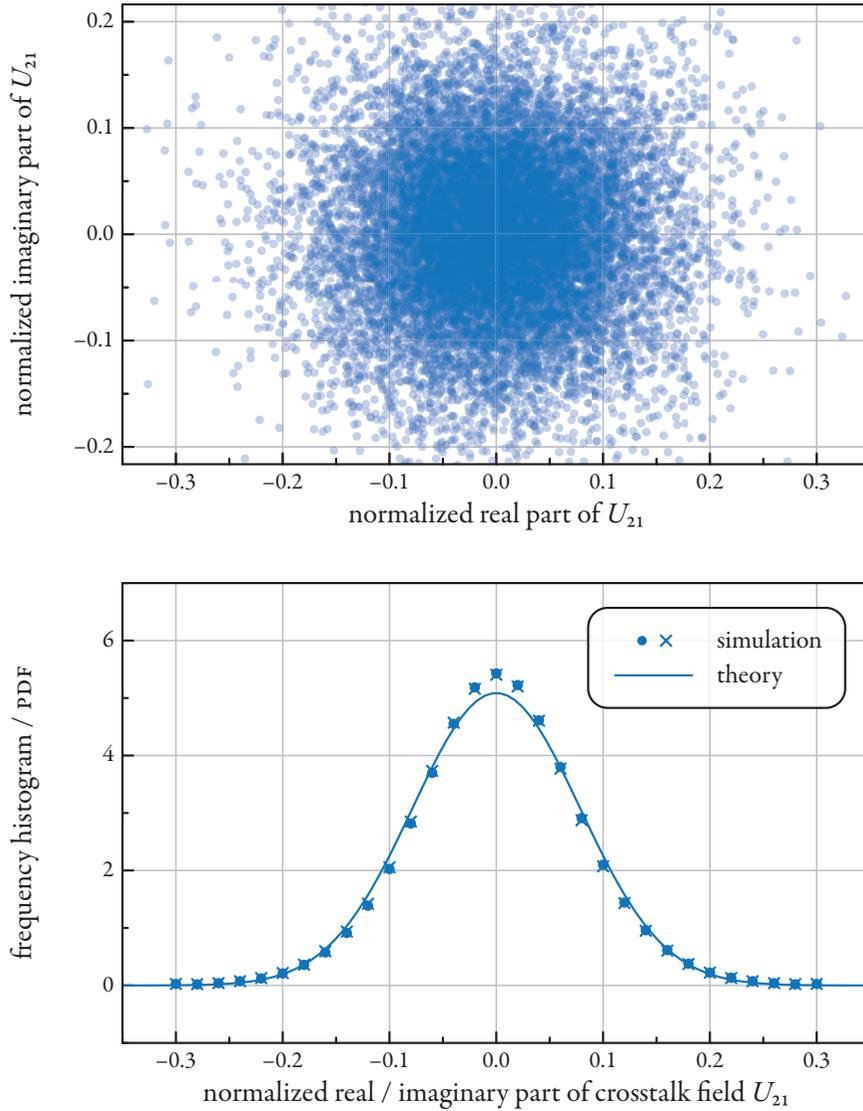


Figure 5.9 shows a sample distribution of 2^{14} random data points of the crosstalk field in the complex plane (top) and the ensemble distribution for the real and imaginary components of the field (bottom) for 500 iterations of the transmission link of Fig. 4.7 (SSMF with high PMD and RDPS). Curve in the bottom figure shows the prediction according to (5.12) for the ensemble mean DOP of $\mathcal{D} \approx 0.975$. All values are normalized to the RMS probe field $\mathcal{E}[P_2]^{1/2}$. The instantaneous crosstalk phase in the top figure is determined relative to the instantaneous phase of the signal in the data subchannel in accordance with the definitions in Section 2.2.2 (page 46ff.)

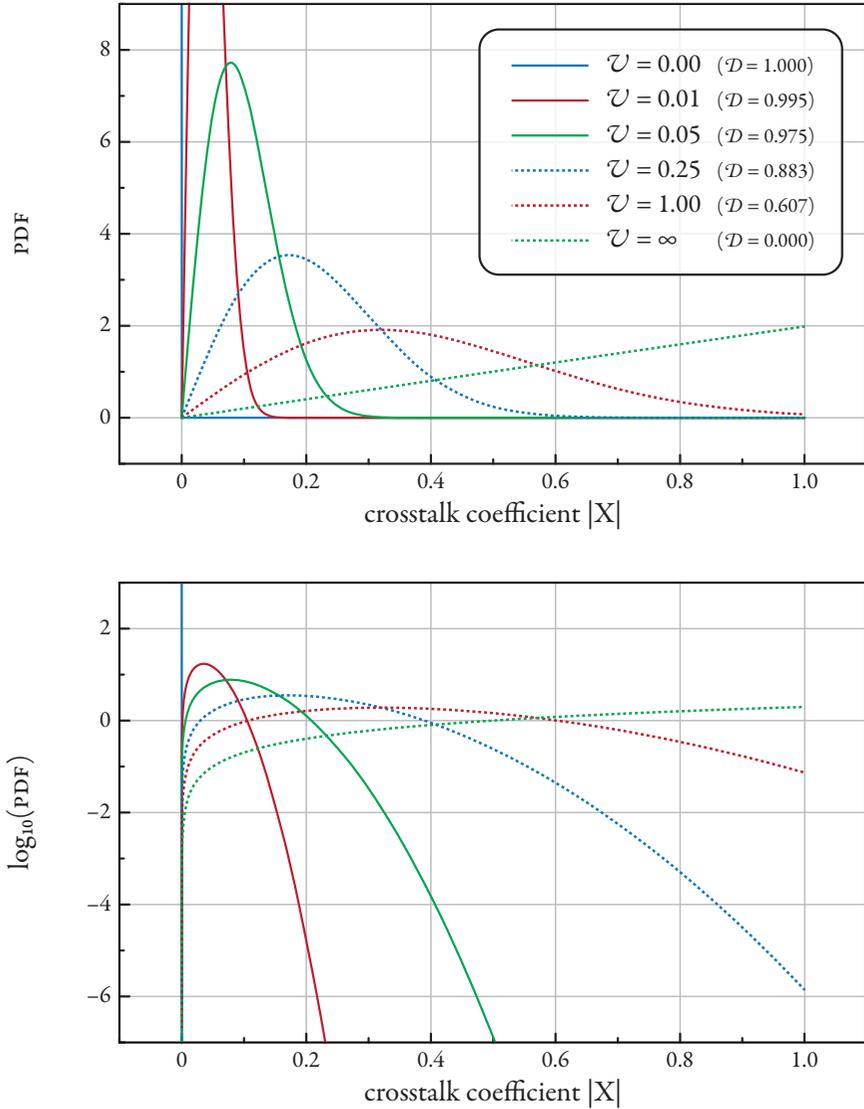


Figure 5.10 shows the population distribution of the crosstalk coefficient $|X|$ for different values of the variance parameter \mathcal{U} respective the DOP of the probe channel in linear (top) and logarithmic (bottom) scale.

Gaussian with mean $\mathcal{E}[U_2]$. In the case of higher-order modulation such as QPSK, the constellation diagram will be a superposition of several such complex Gaussian distributions, related to the various symbols in the modulation alphabet.

The crosstalk field, given as the product of the interfering subchannel field and the complex crosstalk coefficient will then have a uniformly distributed phase, independent of the data modulation, owing to the phase statistics of X . The crosstalk field magnitude will be the magnitude of the product of a Rayleigh-distributed variable ($|X|$) and a Rice-distributed variable ($|U_2|$).³⁹ The exact PDF can be found in [Simon 2002]. However, we will approximate the subchannel field at the sampling instant to be (nearly) constant, with only a small variance, so that $\text{var}[U_2] \ll \mathcal{E}[U_2]^* \mathcal{E}[U_2]$. Then the crosstalk field $U_{21} = X U_2$ will also be approximately complex Gaussian, independent of the actual underlying distribution of U_2 , with a compound variance of [Aroian 1947]

$$\begin{aligned} \text{var}[U_{21}] &= \mathcal{E}[U_2]^* \mathcal{E}[U_2] \text{var}[X] + \text{var}[U_2] \text{var}[X] \\ &= \frac{1}{2\kappa} \left(\mathcal{E}[U_2]^* \mathcal{E}[U_2] + \text{var}[U_2] \right) = \frac{1}{2\kappa} \mathcal{E}[P_2] \end{aligned} \quad (5.12)$$

where we used $\mathcal{E}[X] = 0$ because of distribution symmetry in the complex plane resulting from the uniformly distributed phase ψ_X , and $\text{var}[U] = \mathcal{E}[U^* U] - \mathcal{E}[U]^* \mathcal{E}[U]$. Otherwise, the distribution of U_2 will affect the distribution of the crosstalk U_{21} and its variance. Figure 5.9 shows the distribution of the real and imaginary crosstalk components, obtained in numerical simulations for an exemplary transmission system, together with curves for the Gaussian distribution with the variance from (5.12). These were obtained by determining the crosstalk from a CW signal, for which above condition holds, into an «empty» POLDM subchannel ($U_1(z=0) = 0$), after polarization demultiplexing the signals obtained in the course of the verification of our model in Chapter 3. The match is very good.

The crosstalk generated during POLDM demultiplexing as a result of XPOLM-induced SOP fluctuation is thus additive and approximately Gaussian, with an average power corresponding to the variance (5.12) which is only a function of the interfering subchannel statistics at the sampling instants and the nonlinear depolarization, as described by κ or equivalently the DOP \mathcal{D} . Unlike the additive white Gaussian noise (AWGN) used for noise-loading the signal in order to determine its ROSNR (cf. Section 5.1), the demultiplex noise is not necessarily white. The spectrum of X itself is actually narrow (comparable to the spectral broadening due to XPM). While we do not derive it directly, one can see from the ACFs of the SOP in Fig. 5.5 that the SOP (and hence X) changes rather slowly – even more so in systems with RDPS – and thus is not expected to have very high frequency components. However, the spectrum of the crosstalk field U_{21} in (5.9) is obtained from the convolution of the spectrum of the interfering subchannel U_2 and the spectrum of X . The crosstalk spectrum thus depends to a large degree on the spectral properties of the interfering subchannel. Since we expect that POLDM will primarily be used

distortion, but the degree of the correlation is sufficiently small not to affect the results.

³⁹The Rice distribution describes the magnitude of a complex Gaussian with mean (corresponding to the Rayleigh distribution which assumes zero mean). The square of a Rice-distributed variable has a non-central χ^2 distribution with two degrees of freedom [Proakis 2001].

with high-speed, higher-order modulation, this spectrum would be nearly white.

Typical spectra of U_{21} and U_2 are shown in Fig. 5.11 as obtained from the same set of simulations as the SOP distributions in Figs. 4.6 and 4.7, in which the signal subchannel is launched CW, having a single-peak spectrum. The relatively smooth curves were obtained by averaging the power spectral density (PSD) over all 500 iterations. Fig. 5.12 additionally shows the crosstalk spectrum when the single-polarization signal is QPSK-modulated. The spectral shape of the crosstalk U_{21} closely imitates that of the signal subchannel U_2 , from which we deduce by inverse application of the convolution theorem that the spectral width of X itself is indeed very narrow. Comparing the spectra in Fig. 5.11, we can also see from the much more broadened spectrum how the resonant system is significantly more affected by XPM and FWM (the optical power of the DWDM channels being equal in both plots).

CROSSTALK MITIGATION BY SYMBOL INTERLEAVING As we have already shown in Section 2.2.2, time-interleaving RZ-modulated symbols in both POLDM subchannels can reduce the crosstalk due to a constant SOP misalignment as results e.g. from uncompensated birefringence by minimizing U_2 in (5.9) and (2.54), taken at the instants at which the subchannel of interest is sampled at its maximum. The same principle can be applied when the misalignment is due to XPOLM and of a time-varying nature, where interleaving theoretically minimizes $\mathcal{E}[P_2]$ in (5.12), and thus the crosstalk variance.

On the other hand, it has been shown that such interleaving can lead to a reduced PMD tolerance, especially for NRZ-shaped pulses [Nelson 2000], but also for RZ-QPSK [Borne 2007a]. However, in the presence of PMD compensation, which can be implemented electronically within the POLDM receiver, this argument may be insubstantial. We avoid the PMD-related discussion altogether by using fibers with zero PMD in our numerical simulations (see Section 5.2.2), since, according to our theory of the previous chapters, the presence of PMD will only significantly affect the magnitude, but not the statistics, of XPOLM. Then, using the power scaling law derived in Section 3.3, we can easily convert the nonlinear DOP reduction in the absence of PMD into a corresponding value for arbitrary fiber PMD, including an associated change in the nonlinear threshold of the DWDM interferers.

Throughout the remainder of the present chapter we will compare results for interleaving with those where the symbol slots are aligned in both subchannels to demonstrate its merits.

5.2.1 ROSNR Penalty Estimate

If we approximate the crosstalk generated at the POLDM demultiplexer as a result of XPOLM as additive white Gaussian noise with the power given in (5.12), we can determine an estimate of the ROSNR penalty in a straightforward manner: the allowable optical amplifier noise is reduced by the amount of generated XPOLM noise in order to achieve the same total noise level and thus OSNR and BER. When

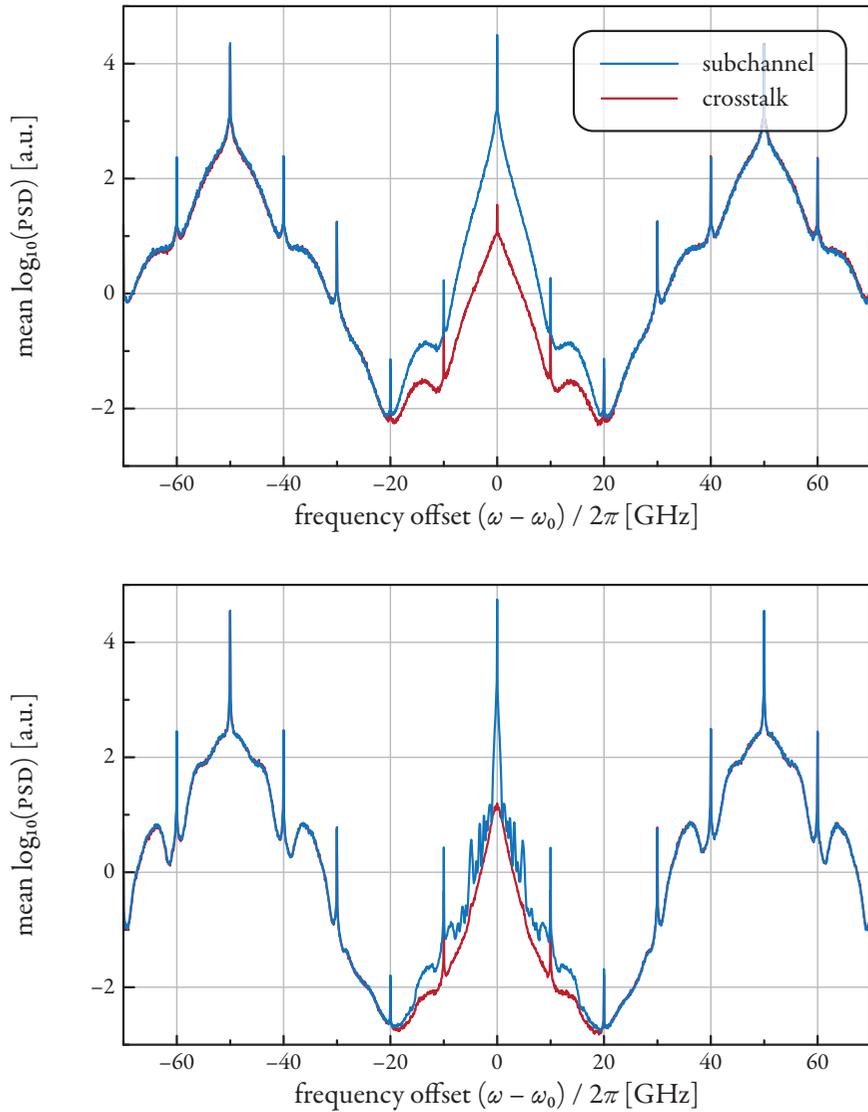


Figure 5.11 shows the power spectral density of the single-polarization CW signal and the crosstalk into the orthogonal SOP after POLDM demultiplexing using the mean signal SOP for a system with resonant dispersion map (top) and one with RDPS of -2048 fs/GHz (bottom). Data was averaged over an ensemble of 500 each. Also see discussion in the text.

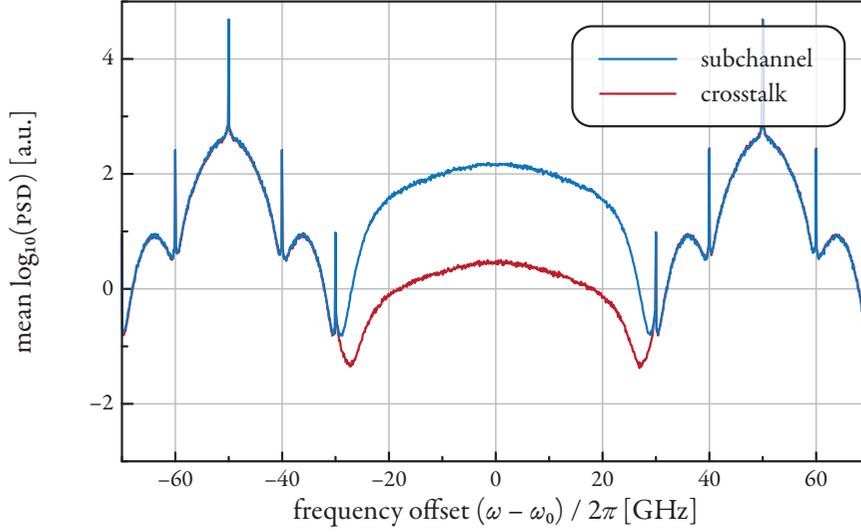


Figure 5.12 shows the power spectral density of the single-polarization RZ-QPSK signal and the crosstalk into the orthogonal SOP after POLDM demultiplexing using the mean signal SOP for a system RDPS of -2048 fs/GHz. Interferer optical power was 4 dBm, probe power was 1 dBm; averaged over an ensemble of 500.

using linear units for the ROSNR,

$$ROSNR_{XPOLM} = \frac{1}{ROSNR_0^{-1} - C_{BW} \frac{\text{var}[U_{21}]}{\mathcal{E}[P_1]}} \quad (5.13a)$$

where P_1 denotes the power in the subchannel of interest at the sampling instants and $ROSNR_0$ is the value required in the absence of XPOLM-related DEMUX noise. If both subchannels carry the same optical power and their symbols are aligned, $\mathcal{E}[P_1] = \mathcal{E}[P_2]$ and with (5.12) the second term in the denominator reduces to $C_{BW}/2\kappa$, where κ is the concentration parameter of the Fisher SOP distribution. When interleaving is used with RZ symbols, $\mathcal{E}[P_1] > \mathcal{E}[P_2]$, and thus the excess ROSNR penalty is reduced. The factor C_{BW} accounts for the ROSNR being normalized to a bandwidth of 12.5 GHz (0.1 nm),

$$C_{BW} = \frac{12.5 \text{ GHz}}{\Delta f_{DWDM}} \quad (5.13b)$$

where Δf_{DWDM} is the spectral width of a DWDM channel. Should the XPOLM demultiplex noise, due to its spectral properties, be less affected by narrow-band optical filtering, as Fig. 5.12 suggests (matched filters for the signal will also be matched for the XPOLM crosstalk), then the impact of XPOLM noise relative to amplifier noise is increased. Such filtering effects can also be incorporated by adapting C_{BW} appropriately. We determined the correction factor numerically. For 25 Gbps QPSK subchannel mod-

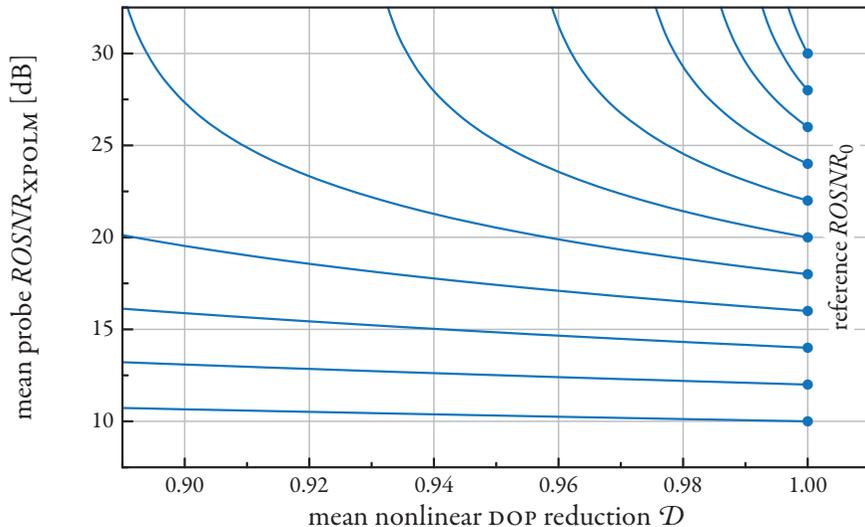


Figure 5.13 illustrates (5.13) for various base values of the ROSNR. Base values are those in the absence of XPOLM and are shown as symbols at $\mathcal{D} = 1$. The higher the base value, the larger the penalty due to XPOLM.

ulation and the receiver filter cascade as described in Table 5.1, the coefficient C_{BW} and thus the impact of crosstalk compared to white noise, is increased by about 12 percent.

As we can see from (5.13), the ROSNR penalty does not only depend on the nonlinear DOP reduction via the Fisher distribution parameter \varkappa , but also on the base ROSNR value. It will become more pronounced as the channel symbol rates increase and as modulation formats become more complex, both associated with an increase of ROSNR_0 . Figure 5.13 shows the dependence of the ROSNR value on the nonlinear DOP for various base ROSNR values. Modulation formats with low OSNR requirements will thus be less affected by XPOLM crosstalk (and fading) than those with higher ROSNR_0 values. We would thus expect the currently much-researched higher-order modulation formats, which make nearly full use of the OSNR offered by the transmission line, to be affected more seriously.

5.2.2 Numerical Results

To verify our estimates, we again performed numerical simulations in which we examined an ensemble of equivalent systems (from the standpoint of polarization averages / distributions) for each parameter combination. Since our estimates predict a possibly significant impact of XPOLM on POLDM transmission, we replace the CW probe – which has helped us considerably in the verification of our model and in gaining a better understanding by isolating the interchannel Kerr nonlinearities – by a POLDM probe channel with RZ-shaped, differential-encoded, QPSK-modulated pulses at 25 Gbaud, while keep-

ing the 10 Gbps NRZ interfering channels as a source of significant XPM and XPOLM.⁴⁰ The former being a prime candidate for 100 Gbit Ethernet transmission (not accounting for a possible line rate increase due to overhead for FEC or line coding), we determine penalties in the more realistic scenario of a progressive system upgrade from 10 to 100 Gbps DWDM channels. The generation of RZ-differential quaternary phase-shift keying (DQPSK) pulses is straightforward and described e.g. in [Seimetz 2009]. Our receiver configuration is shown in Fig. 5.1 (bottom).

At its core is a coherent receiver in a polarization-diversity configuration, i.e. the incoming signal is split and input to a pair of coherent receivers, the LO in one of which is polarized in $|e_x\rangle$ and the other in $|e_y\rangle$ – usually, these use a common laser source polarized at $|e_{45^\circ}\rangle$ followed by a PBS aligned to $|e_x\rangle$ and $|e_y\rangle$. The output photocurrents are derived e.g. in [Seimetz 2009] and can be written very compactly in bra-ket notation,

$$I_x = \frac{1}{2} \left(\langle U | U_{LOx} \rangle + \langle U_{LOx} | U \rangle \right) = \text{re} \left[\langle U | U_{LOx} \rangle \right] \quad (5.14a)$$

$$Q_x = \frac{i}{2} \left(\langle U | U_{LOx} \rangle - \langle U_{LOx} | U \rangle \right) = \text{im} \left[\langle U | U_{LOx} \rangle \right] \quad (5.14b)$$

$$I_y = \frac{1}{2} \left(\langle U | U_{LOy} \rangle + \langle U_{LOy} | U \rangle \right) = \text{re} \left[\langle U | U_{LOy} \rangle \right] \quad (5.14c)$$

$$Q_y = \frac{i}{2} \left(\langle U | U_{LOy} \rangle - \langle U_{LOy} | U \rangle \right) = \text{im} \left[\langle U | U_{LOy} \rangle \right] \quad (5.14d)$$

where $|U_{LOx}\rangle = U_{LO}|e_x\rangle$ is the local oscillator output for the x polarization component, and correspondingly $|U_{LOy}\rangle$, $\text{re}[\cdot]$ is the real part, $\text{im}[\cdot]$ denotes the imaginary part, and we neglected the photodiode responsivity. All quantities in (5.14) are time-dependent, even if this dependence was not explicitly stated. If we assume an ideal homodyne receiver, we can thus reconstruct the optical field $|U\rangle$, transformed into the baseband, as

$$U_{LO}|A\rangle = U_{LO} \left(A_x |e_x\rangle + A_y |e_y\rangle \right) = U_{LO} \begin{pmatrix} I_x + iQ_x \\ I_y + iQ_y \end{pmatrix} \quad (5.15)$$

where $|U\rangle = |A\rangle \exp(i\omega_\rho t)$ with carrier frequency ω_ρ of the POLDM probe. Without receiver phase noise, U_{LO} is a constant complex scalar, which we shall arbitrarily set to unity to simplify our equations. Disregarding the electrical filtering of the photocurrents to reduce noise, the POLDM DEMUX then calculates the projection of $|A\rangle$ on the POLDM subchannel SOPs according to (2.13), e.g. for subchannel 1,

$$|e'_1\rangle \langle e'_1 | A \rangle = A'_1 |e'_1\rangle = \left(\langle e'_1 | e_x \rangle A_x + \langle e'_1 | e_y \rangle A_y \right) |e'_1\rangle \quad (5.16)$$

where the brackets yield complex projection coefficients – these are normalized so that the sum of their

⁴⁰The differential encoding allows us to forego the implementation of phase recovery algorithms which become necessary in coherent detection as there exists no reference phase. However, this increases the ROSNR by about 3 dB since the optical noise fields of the two pulses contributing to each received symbol add.

squares is unity. However, the DEMUX does not a priori know the correct SOP $|e'_1\rangle$ due to the various linear and nonlinear polarization effects discussed in Chapters 2 and 3 and must determine it from the incoming signal. Due to the time-varying nature of the signal SOP, this will involve some form of averaging, so that we may write for the POLDM DEMUX alignment

$$|e'_1\rangle = \mathcal{E}[|e_1(t)\rangle] \quad (5.17)$$

where $|e_1\rangle$ is the actual, time-dependent SOP of the incoming signal. In the presence of XPOLM there will always exist misalignment between $|e_1\rangle$ and $|e'_1\rangle$ due to the rapid SOP motion, which leads to fading and crosstalk, as derived in Section 2.2.2.

In our simulations, the subchannel SOPs $|e'_1\rangle$ and $|e'_2\rangle$ are obtained from parallel simulations of a single-polarization signal through the exact same polarization ensemble member, i.e. a system with the same initial SOPs, bit sequences, etc. in all channels except the probe, in which the second subchannel is empty, $U_2(z=0) = 0$. At the receiver, we simply calculate (5.17) over all transmitted symbols in the probe and obtain the receiver SOP $|e'_1\rangle$ which minimizes Θ and thus the crosstalk X . In a real-world implementation, where we do not have the luxury of such duplicate transmission, the best subchannel SOPs must be found by an appropriate search algorithm from the POLDM signal itself. Often, the constant-modulus algorithm is used for this purpose [Godard 2003, Leven 2008]. The presence of XPOLM makes it harder to find this (quickly changing) SOP and the averaging windows of these algorithms may significantly affect the impact of XPOLM.

The additional single-channel transmission also provides a reference ROSNR value – resulting from XPM/FWM alone without XPOLM crosstalk – as well as the instantaneous value of the nonlinear depolarization D by which XPOLM was quantified in the previous chapters - this DOP value cannot be determined from the compound POLDM signal.

In order to concentrate on the relevant nonlinear effects in the present context, we use lasers without phase noise for both the transmitter and the receiver LO, eliminating additional noise sources. Furthermore, we use DQPSK modulation instead of QPSK. In this way we avoid the need for implementation of complicated carrier recovery schemes, such as the Viterbi-Viterbi method [Viterbi 1983], well knowing that the so-obtained results are not directly transferable to laboratory or field-deployed systems that use QPSK. Also, to avoid the (in our case uninteresting) penalty associated with PMD – which we had outlined on page 61 – the simulation SSMFs did not cause PMD at all. The actual impact of PMD depends to a large degree on the abilities of the employed PMD compensation, which we do not model in the scope of the present work. As we used significant RDPS of $-2048 \text{ ps/km}^{1/2}$, the impact of PMD on the magnitude \mathcal{D} of XPOLM remains slight (cf. Figs. 3.16 and 3.15). However, to compensate for absent fiber PMD, we randomly inserted two reconfigurable optical add-drop multiplexer (ROADM)s into each interferer channel which effectively enforce SOP randomization, not unlike what we would see as the result of PMD (the statistics of both being significantly different, of course).

The ROSNR can then be obtained, as before, by noise-loading the received signal in repeated Monte Carlo trials. We use 128 noise instances for each ensemble member of 2^{11} symbols length to obtain the

Table 5.1 System parameters changed in split-step simulations of the present section unless noted otherwise.

| | |
|---|----------------------------------|
| modulation format (probe) | POLDM RZ-QPSK |
| probe MUX filter bandwidth | 45 GHz |
| probe MUX filter shape | 5th-order Gaussian |
| probe DEMUX filter bandwidth | 45 GHz |
| probe DEMUX filter shape | 5th-order Gaussian |
| ROADM add/drop filter bandwidth | 40 GHz |
| ROADM add/drop filter shape | rectangular |
| PRBS (in each I and Q tributary) | $2^{16} - 1$ (subsequences used) |
| simulation time window | 2048 bits |
| fiber PMD parameter | 0 ps/km ^{1/2} |

ROSNR for a BER of 10^{-3} with sufficient accuracy. We average the ROSNR values for the in-phase and quadrature signal components, obtaining a pair of ROSNR values – one for each POLDM subchannel – per simulation iteration.

The system parameters differing from those in Table 3.1 as a result of the changes in the probe channel are summarized in Table 5.1.

The top parts of Figs. 5.14 and 5.15 show the ROSNR penalty versus the NRZ interferer input power for a system using RDPS and one with a resonant dispersion map, respectively, determined relative to the back-to-back ROSNR value of 18.7 (POLDM) and 15.7 (single-polarization).⁴¹ The lowest-penalty curve is the XPM reference penalty for single-polarization transmission. As we have shown in Section 5.1, there is no statistically significant correlation between this penalty and XPOLM-related fading. The other curves in the figure represent the penalties for POLDM transmission using temporally aligned and interleaved subchannels. In the case of subchannel alignment the mean ROSNR penalty is significantly larger than that of interleaved subchannels. This means that the optical power within the co-propagating NRZ channels must be reduced by up to 2 dB to enable POLDM transmission.

In order to determine the reason for the unexpectedly poor performance of subchannel interleaving, especially in the system employing RDPS – ideally, interleaving would negate any penalty associated with XPOLM crosstalk – we must look at the statistics of P_2 , the power in the interfering subchannel, at the sampling instants. As we expect the signal shapes and their distortions in both POLDM subchannels to be very similar to the equivalent single subchannel signal, eye diagrams of the single polarization signal at the fiber output are plotted in Fig. 5.16. We can discern from the resonant system that there exist significant pulse distortions and pattern effects from optical filtering, leading to non-zero power

⁴¹The inherent 3 dB ROSNR penalty for POLDM transmission is due to the optical power available for each subchannel being only half the total power used to determine the signal-to-noise ratio. Also, theoretical and measured back-to-back ROSNR values for systems similar to the one implemented herein are as low as 15.5 dB, hinting that the externally provided filter data in Table 5.1 may not be optimal [Borne 2009].

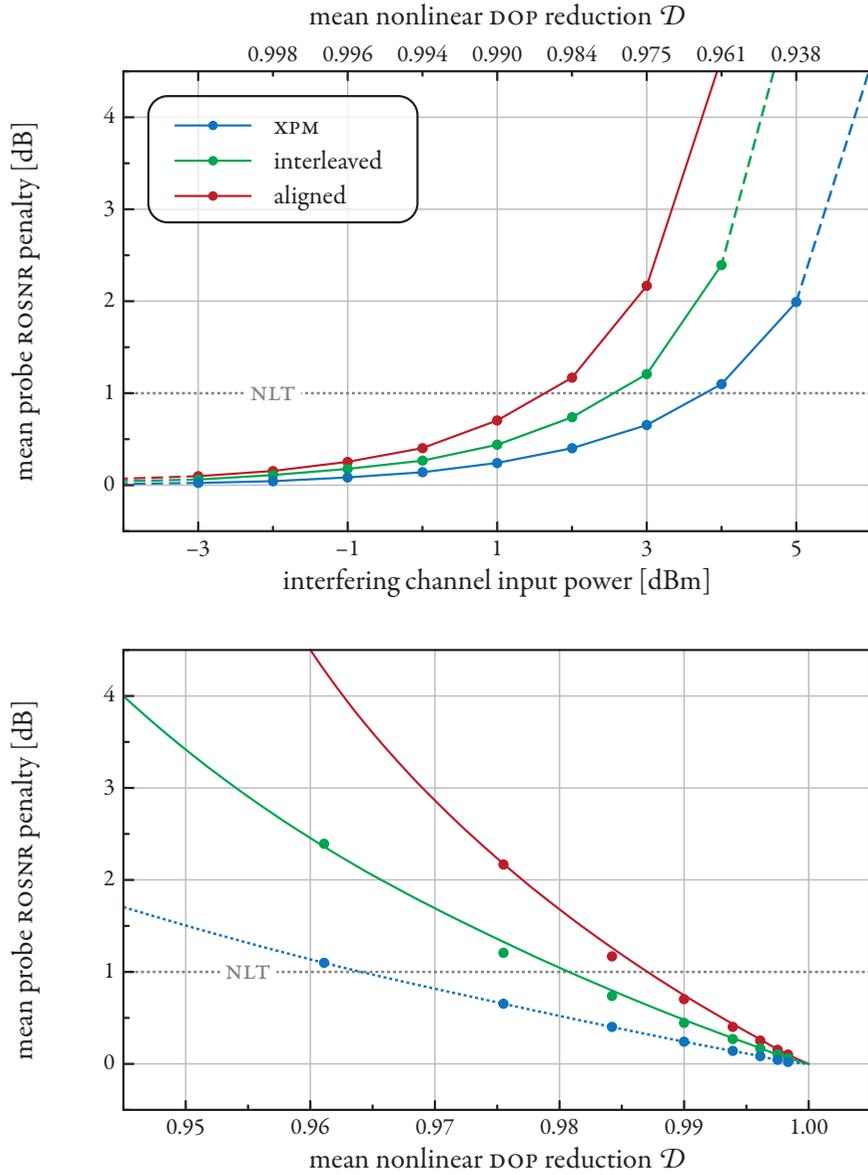


Figure 5.14 plots the mean probe channel ROSNR penalty versus the interferer power (top) and versus the nonlinear depolarization \mathcal{D} (bottom) for the single-polarization, aligned-subchannel POLDM, and interleaved-subchannel POLDM systems, as obtained by numerical simulations. The dispersion map had an RDPS value of -2048 fs/GHz. Short dashes denote interpolation. The back-to-back reference ROSNR for the penalty is 15.7 dB (single-polarization) or 18.7 dB (POLDM). Ensemble size was 100.

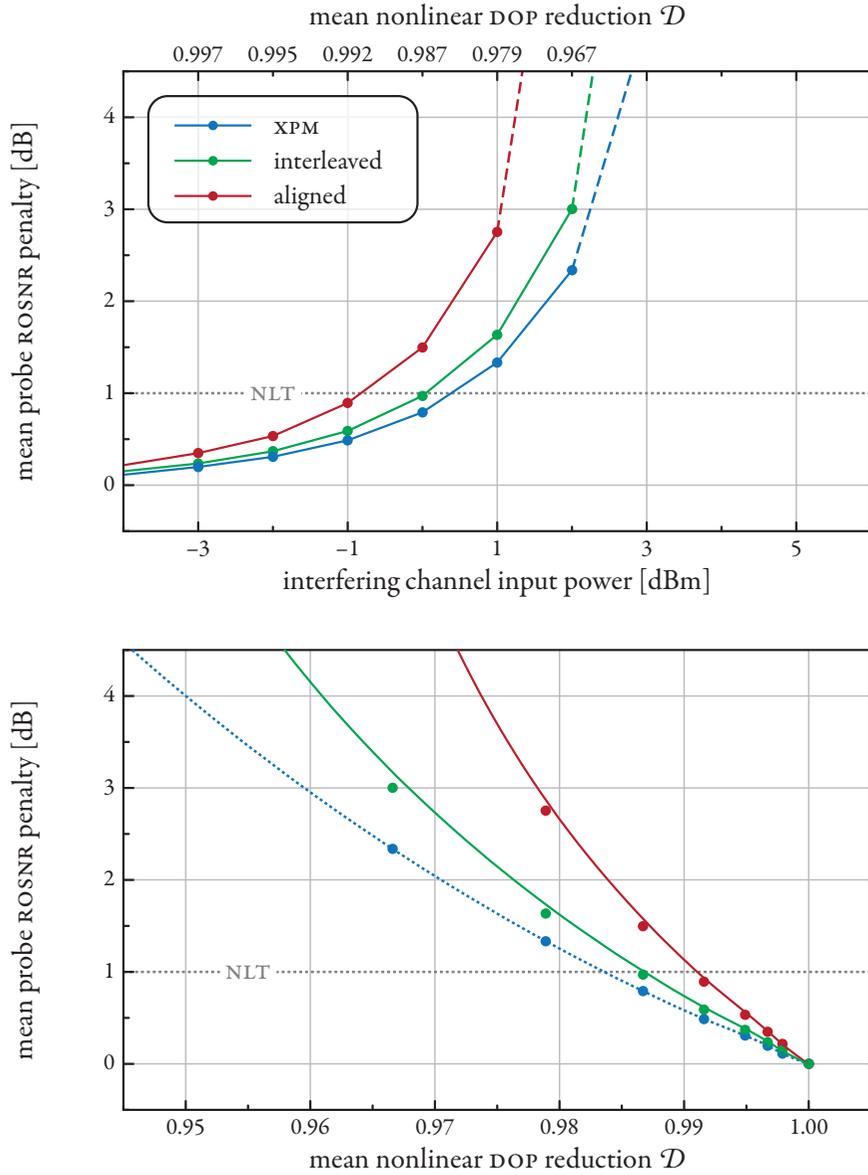


Figure 5.15 plots the mean probe channel ROSNR penalty versus the interferer power (top) and versus the nonlinear depolarization \mathcal{D} (bottom) for the single-polarization, aligned-subchannel POLDM, and interleaved-subchannel POLDM systems, as obtained by numerical simulations. The dispersion map was resonant (no RDPS). Short dashes denote interpolation. The back-to-back reference ROSNR for the penalty is 15.7 dB (single-polarization) or 18.7 dB (POLDM). Ensemble size was 100.

between adjacent symbols. Furthermore, for the system employing RDPS, there is significant additional distortion from intrachannel nonlinearities, leading to large timing jitter and further increasing the optical power present in the inter-symbol spaces [Essiambre 2002]. It occurs because in the RDPS system the optical power in the probe is much higher than in the resonant system (it is coupled to the NRZ channel input power to avoid emphasizing the impact of FWM), and the accumulated dispersion B_2 can become very large so that many symbols overlap due to pulse broadening (cf. Section 2.2.1). The result is a non-Gaussian, multi-modal power distribution with a high variance of the field power in the inter-symbol spaces, so that the the distribution of $X U_2$ (or rather $X A_2$) is no longer Gaussian and (5.12) does not apply.

We can, however, determine the variance $\text{var}[U_{21}]$ needed in (5.13) directly from the simulations by demultiplexing the single-polarization signal and analyzing the crosstalk appearing in the «empty» POLDM subchannel, similar to what we did to obtain the spectra in Figs. 5.11 and 5.12. We obtain exemplary values of

$$\begin{aligned} \text{var}[U_{21}] &\approx 1.30 \frac{1}{2\chi} \mathcal{E}[P_2] && \text{for the RDPS map at } \mathcal{E}[P_\rho] = 1 \text{ dBm} \\ \text{var}[U_{21}] &\approx 1.05 \frac{1}{2\chi} \mathcal{E}[P_2] && \text{for the resonant map at } \mathcal{E}[P_\rho] = -3 \text{ dBm} \end{aligned}$$

The non-Gaussian correction factors depend on the amount of distortion within the POLDM probe and thus on its launch power and/or filter history. However, at similar ROSNR penalties the factor, and thus the crosstalk, for the resonant system in which the interchannel distortions dominate will generally be smaller. Optimization of the dispersion map regarding pre-distortion may also affect the accumulated distortions within the probe and thus the correction factor.

So, even though the ratio $\mathcal{E}[P_2]/\mathcal{E}[P_1]$, which appears when inserting (5.12) into (5.13), is in our simulations almost equal independent of the dispersion map (≈ 0.3), the large correction factor for the RDPS system reduces the efficiency of interleaving, as observed when comparing Figs. 5.14 and 5.15.

The bottom parts of these figures plot the ROSNR penalty versus the nonlinear probe depolarization. Even though the XPM curves (interpolated from simulation data) seem to indicate a penalty dependence on the DOP due to the depolarization of the signal, we know that the actual penalty is due to XPM and the depolarization is just an irrelevant «side effect» as discussed in Sec. 5.1 – therefore the line is shown dashed. However, when a second POLDM subchannel is added, additional penalties due to polarization crosstalk appear, which do depend on the DOP. The solid curves in the figures, describing this additional penalty, were semi-analytically obtained with (5.13), using the corresponding XPM penalties from the single-polarization cases and the correction factors determined above. The match with the simulation results is very good. We notice that the DOP requirements in order not to exceed the NLT are between 0.98 and 0.99, depending on subchannel alignment and dispersion map – the higher required DOP values for the resonant system are partially due to the faster increase of XPM distortions compared to nonlinear depolarization in such a dispersion map. We can then use these DOP threshold values in (3.61) and Fig. 3.15 to determine the corresponding thresholds for the interferer powers using our proposed

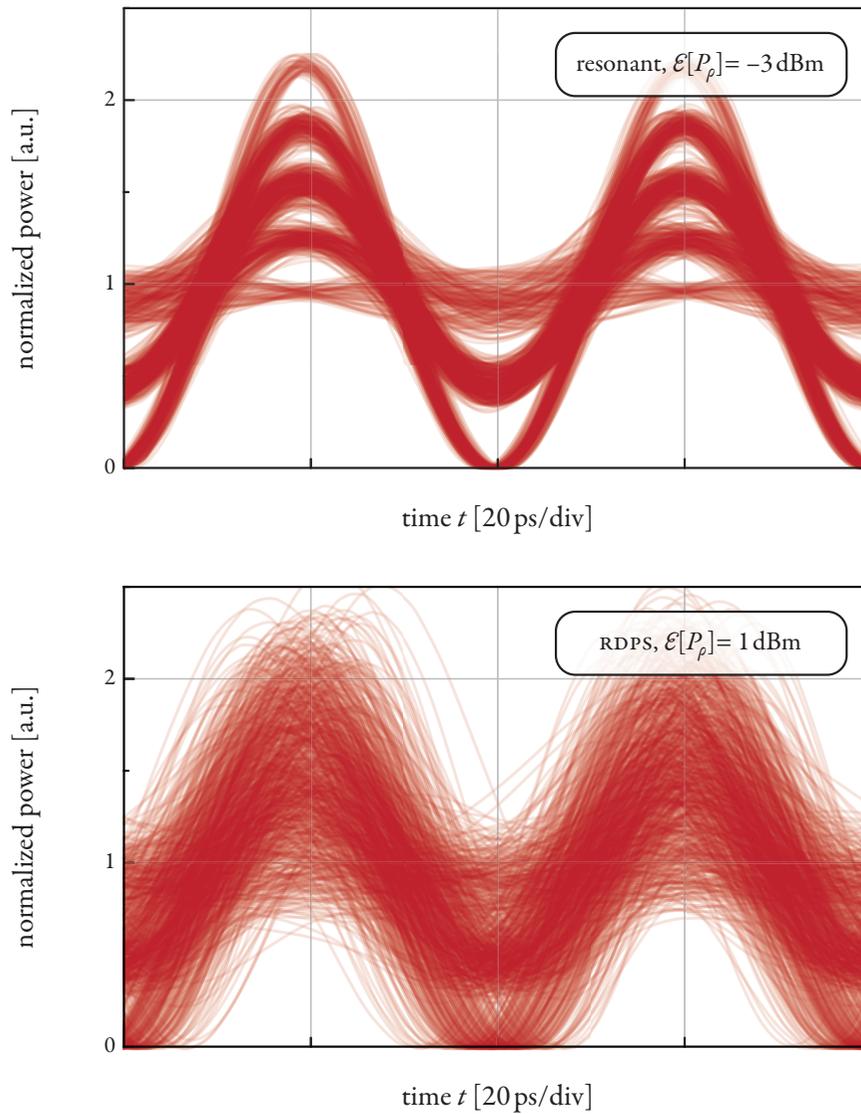


Figure 5.16 shows the eye diagrams of the single-polarization signal at the output of the transmission line for a system with resonant dispersion map (top) and one using RDPS (bottom). Sampling instants are either in the center of the symbols (subchannel alignment) or in the center of inter-symbol space (subchannel interleaving) when such a channel is the source of POLDM crosstalk.

model for any combination of fiber GVD and PMD. Because all NLT figures in Section 3.3 were referenced to a DOP of 0.97, the threshold powers shown there must be decreased by 0.9 dB and 2.4 dB for DOPs of 0.98 and 0.99, respectively.

ERROR BURSTS We have shown that the correlation of the probe SOP can be correlated over longer periods of time in Fig. 5.5, especially when RDPS is present in the dispersion map. Thus, a probe symbol whose SOP is significantly deflected from the mean SOP (assumed to be an eigenstate of the POLDM DEMUX) is likely to have neighboring symbols which are equally disturbed. This probability increases with the amount of RDPS present. It will also differ from one ensemble member to the next, as characterized by the distribution in Fig. 5.6 – however, the width of this distribution decreases with increasing RDPS. It is beyond the scope of the present work to perform a more detailed analysis of the probability of error bursts and the burst properties, apart from these general remarks. Because of the widespread proliferation of FEC in optical transmission systems, we expect that this topic will come under increased scrutiny in the near future.

ENSEMBLE STATISTICS AND OUTAGE STATISTICS We have already briefly touched on the subject of differences within the polarization ensemble in the previous paragraph. These variations apply not only to the ACF $\zeta^{\Delta\Theta}$ of the SOPs, but also to the DOP, as we have discussed in Chapter 4, and the SOP distribution itself, as shown in Figs. 4.6 and 4.7. We thus also expect variations of the ROSNR within the polarization ensemble, which we will briefly discuss next.

To illustrate the ensemble behavior and XPOLM dependence, we plot the individual ROSNR penalties versus the individual nonlinear DOP values at three different interferer power levels in Fig. 5.17. The scatter plots for the single-polarization signal are always independent of the nonlinear depolarization D , confirming the result of Section 5.1. On the other hand, the scatter plots for the POLDM signals show a clear dependence of the penalty on the DOP within the ensemble, emphasizing the relationship between the transmission penalty and XPOLM. For illustration, we have also plotted the semi-analytical estimates, similar to the bottom parts of Figs. 5.14 and 5.15, which correspond well to the mean values of the simulation scatter plots.

As the DOP approaches unity, the theoretically predicted and simulation ROSNR values approach the XPM/FWM limit. However, the dependence of the ROSNR on the DOP has significant statistical variation in the numerical simulations, partly because it depends on the particular SOP distribution within the probe channel and also because the ROSNR is a very nonlinear quantity which is prone to strong fluctuations as the penalty increases. This variation scales with the DOP reduction – and thus the XPOLM magnitude. At equal mean ROSNR penalties, the penalty distributions are very similar (cf. XPM at 4 dBm, interleaved at 3 dBm, and aligned at 2 dBm). However, the penalties due to XPOLM remain dependent on the individual DOP D , while those due to XPM do not.

As with PMD, where the penalty also varies from one ensemble member to the next, we find it sensible to define an outage in terms of exceeding a certain ROSNR threshold (another outage condition could be defined in terms of the burst properties of errors, as discussed previously). Arbitrarily setting

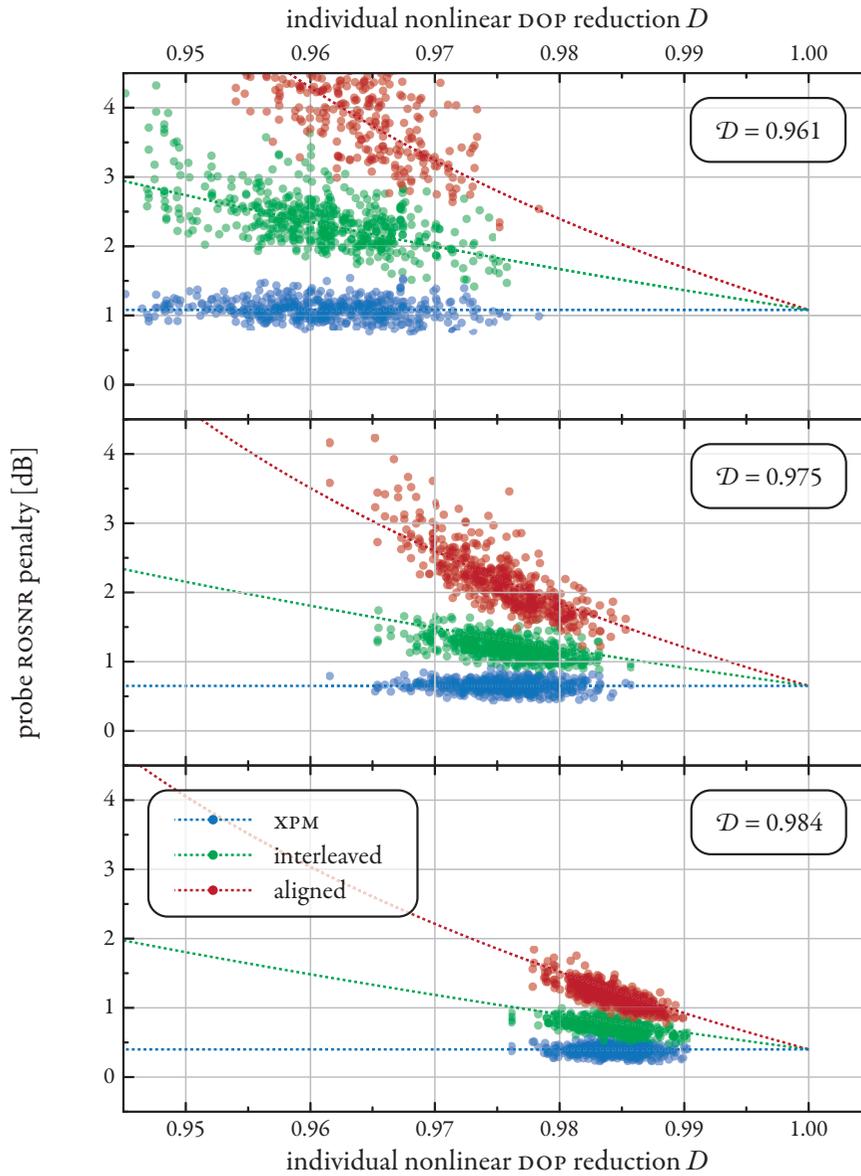


Figure 5.17 plots the individual member ROSNR penalties versus the nonlinear DOP reduction for the single-polarization signal (blue), POLDM with symbol-aligned subchannels (red), and POLDM with interleaved subchannels (green) at interferer powers of 4 dBm (top), 3 dBm (middle), and 2 dBm (bottom). The curves show theoretical predictions according to Section 5.2.1 for comparison.

the outage threshold to an ROSNR 2 dB (3 dB) above the back-to-back value, we see that when provisioning the transmission link to allow for ≈ 1 dB XPM/FWM penalty mean, the outage probability for the ensembles in Fig. 5.17 (top) will be 0.87 (0.08) and 1.00 (0.96) for interleaved and aligned POLDM subchannels, respectively, which are not acceptable for a system operator. However, when provisioning the link so that the average compound (XPM, FWM, and XPOLM) nonlinear interchannel penalty is ≈ 1 dB, no outages occur within the simulation ensemble of 500 values. However, the width of the distribution is not infinitely narrow, and with finite probability outages may still occur. Unfortunately, due to the many factors that enter into the determination of the ROSNR (penalty), it is at present not possible to give an analytic expression for these distributions and the associated outage probabilities.

As long as the penalties are grouped sufficiently tight around the semi-analytic predictions, as in the bottom part of Fig. 5.17, we may use the theory of Section 4.2 to estimate the probabilities of those DOP values that will likely cause outages (e.g. $D < 0.973$ for aligned subchannels at 2 dBm interferer power), though there will still be some uncertainty about the result.

A different way to a semi-analytic solution is related to the method employed in Section 4.2: Using the stochastic degrees of freedom of a particular system, one might generate a set of SOPs for a particular member. From this (finite) set, using (5.9) and assumed statistics of the interfering subchannel A_2 , one might generate another (infinite) set of crosstalk samples, which are generally non-Gaussian within a single ensemble member. These samples could then be used in an analytic prediction as in [Forestieri 2000] to determine the ROSNR value for that particular member. The propagation of a single probe channel without DWDM interferers might suffice to generate the necessary statistics of A_1 and A_2 . This kind of crosstalk generation would be much faster than simulating a very large number of systems to generate accurate frequency histograms. While leaving the actual building of the histograms in this way to the interested reader, we can say from Fig. 4.4 that systems with a resonant dispersion map will show larger fluctuations due to their low number of DOFs than systems with RDPS.

Concluding, we have shown that the nonlinear depolarization induced by XPOLM can lead to significant penalties for POLDM transmission and must be properly accounted for. In the exemplary system investigated in the preceding section, the excess ROSNR penalty due to XPOLM crosstalk generated when demultiplexing the POLDM subchannels can be several dB above the XPM-related penalty. The optical power of the interfering NRZ channels must be reduced by up to 2 dB in order to keep the ROSNR penalty attributed to interchannel nonlinearities at the same level as for single-polarization transmission. This is a significant reduction. Interleaving both POLDM subchannels results in the related penalties to be approximately halved – this could be further improved by avoiding leakage of optical power from the symbols into the neighboring gaps, thus increasing the extinction ratio of the modulation at the receiver.

OUTLOOK

Yet even the longest and most thorough work on a topic as unexplored and vast as cross-polarization modulation cannot answer all questions. The present work, though long, is no exception. We have often mentioned in the text when a further exploration into a particular direction, even if desirable, was not possible due to various, mostly time-related constraints and will summarize the most interesting of these to conclude our work. Foremost, we would liked to have performed a critical evaluation of the effect of XPOLM, and its induced nonlinear birefringence, on PMD. Both Karlsson and Boroditsky have shown that PMD compensation will be impaired by a nonlinear depolarization of the PMD vector. Since PMD compensation has moved into the electrical domain and become feasible and required for transmission at ever higher data rates, a possible nonlinear deterioration of this linear effect would be of great interest.

PMD also affects the effectiveness of time-interleaving both subchannels in POLDM systems, as it was shown elsewhere that uncompensated PMD causes linear crosstalk between the subchannels, and interleaving increases the magnitude of this crosstalk. However, the transmission fibers in the simulations in Section 5.2 on POLDM systems had no PMD because time constraints made it impossible to implement a realistic PMD compensator. The true benefit of time-interleaving can only be determined when PMD is properly accounted for, especially in light of the nonlinear PMD discussion above. Another aspect of great influence in real-world systems is the algorithm that extracts the subchannel SOPs at the receiver, such as the constant-modulus algorithm. How application of this algorithm alters the results given in Section 5.2, and how these depend on the various parameters of the algorithm, such as the averaging window, will likely be the topic of continued research.

Furthermore, we have restricted ourselves throughout to interfering channels that used the NRZ modulation format at a modest bit rate of 10 Gbps to verify our model. As we have shown, such modulation causes significant cross-channel distortions, both XPOLM and XPM. While we have predicted significantly less distortion from modulation formats such as (RZ-)DPSK or higher data rates, we were not able to quantify their impact, either theoretically or by means of simulations within the time allotted to the research project whose results are summarized herein. A particular interesting aspect is that the benefit of residual dispersion per span (RDPS) may be negated in the case of DPSK modulation, as the associated increased PM-AM conversion may cause significantly larger penalties than in the absence of RDPS. This is especially relevant, as inline dispersion compensation is being made obsolete by receiver electronics.

Also, we have been able to show that large SOP excursions can persist over the duration of several symbols in the probe channel, so that the associated crosstalk in POLDM systems may lead to longer error bursts. A closer examination of the dynamics of XPOLM may shed some light on the expected effect on forward error correction.

Finally, while we have made some progress regarding the ensemble statistics by introducing the de-

degrees of freedom as a property for a particular transmission system, there has been much left unanswered. It is not clear if these DOFs could be used to predict the ensemble distribution of transmission penalties in e.g. POLDM systems by a similar method as the one used to semi-numerically determine the DOP distribution within the ensemble; or if any other relevant statistics besides the DOP can be derived by means of the DOF.

While this is only a selection of questions that occurred during the present work, it demonstrates that there is much yet to explore on the topic of cross-polarization modulation. Since POLDM is already moving from research to implementation, and data and baud rates are expected to keep increasing, XPOLM may well be a topic of interest for years to come.

Going beyond XPOLM, the methods and expressions derived herein may possibly be adapted to an analysis of XPM, whose nature and causes are very similar. Since it is not dependent on polarization (at least not in the Manakov description), the resulting expressions and relations may prove even simpler than those for XPOLM, so that they might even have been a sensible starting point from which to extend to cross-polarization modulation.

SUMMARY

The present work discusses a nonlinear effect which occurs during propagation of data signals over optical fiber, especially in WDM systems in which many such signals are transmitted simultaneously using different wavelengths of light. This particular effect, cross-polarization modulation (XPOLM), causes signal-dependent changes in the birefringence of the fiber, comparable to how cross-phase modulation modulates the refractive index. This results, among other consequences, in a rapid modulation of the state of polarization (SOP) of the propagating signals, as a result of which the data signal depolarizes.

Starting with the derivation of a mathematical framework on which all subsequent analysis is based, we created a formalism and model for the changes in the SOPs due to cross-polarization modulation. We derived the statistical behavior of a probe channel under the influence of XPOLM from a large number of interfering channels under almost arbitrary conditions. We have made extensive use of the methods of probability theory due to the exceedingly large number of states such a system can be in. Our results were therefore given in terms of probabilities for the polarization properties of the probe signal after traversing an optical link — most importantly the distribution of the polarization states of the probe, in the sense of the population distribution which includes the whole ensemble of polarization parameters (launch polarization, birefringence and PMD properties of the fiber) and is parameterized by the nonlinear degree of polarization (DOP) reduction.

We believe that the expressions for the magnitude of XPOLM degradation, as described by the nonlinear depolarization, are sufficiently straightforward to allow considerable insight into the nonlinear process and its dependence on various system parameters. The integrals in the expressions can be solved numerically sufficiently fast to allow evaluation of a large number of different systems within a reasonable amount of time. We have also demonstrated how to determine the nonlinear threshold power corresponding to a given required minimum DOP value, and how various means of mitigation, such as co-polarized or polarization-interleaved launch as well as employing residual dispersion per span, can significantly reduce XPOLM degradations in low-PMD fibers.

To probe beyond the all-encompassing population distribution, we also derived the distribution of nonlinear impairments (in terms of the nonlinear depolarization) within the polarization ensemble using a combination of analytic and semi-analytic means. Such a distribution is helpful in determining the percentage of systems that are affected significantly more than the statistical average or some pre-defined limit, or alternatively finding the required average so that the percentage of systems so affected is negligible.

Finally, we have examined the consequences of nonlinear depolarization of a data signal in two cases which required completely different approaches. These are intended as relevant examples for further work on the topic, showing how the knowledge of the distribution of SOPs and the nonlinear depolarization can be used to estimate the impact on transmission fidelity. The first case was DPSK transmission,

in which the optical fields of adjacent data symbols must interfere, which is only possible if their SOPs are sufficiently parallel. We found that the nonlinear SOP changes are not rapid enough to cause significant penalties. The second case was polarization-division multiplex transmission, in which both polarization subchannels must be separated at the receiver. We found that the nonlinear depolarization can cause significant crosstalk between both subchannels. The penalties incurred as a result may be significantly higher than those related to XPM, even for phase-sensitive QPSK transmission.

ZUSAMMENFASSUNG

Die vorliegende Arbeit diskutiert einen nichtlinearen Effekt, der bei der Übertragung von Datensignalen über optische Fasern auftritt; insbesondere in WDM Systemen, in denen viele solcher Signale gleichzeitig auf verschiedenen Wellenlängen übertragen werden. Dieser spezielle Effekt, die Kreuz-Polarisations-Modulation, verursacht signalabhängige Änderungen in der Doppelbrechung der Faser, vergleichbar mit der Modulation des Brechungsindex durch die Kreuz-Phasen-Modulation. Dies führt, unter anderem, zu einer schnellen Modulation des Polarisationszustandes des sich ausbreitenden Signals und der daraus resultierenden Depolarisierung.

Wir entwickelten, anfangend mit der Herleitung einer mathematischen Grundstruktur auf der die weitere Analyse basiert, einen Formalismus und ein statistisches Modell für die Änderungen der Polarisationszustände durch die Kreuz-Polarisations-Modulation. Wir leiteten das statistische Verhalten eines Testkanals unter fast beliebigen Bedingungen her, der den Effekten von XPOLM einer großen Anzahl von Nachbarkanälen ausgesetzt ist. Aufgrund der überwältigenden Anzahl von Zuständen, in denen sich solch ein System befinden kann, machten wir umfassenden Gebrauch von den Methoden der Wahrscheinlichkeitslehre. Unsere Ergebnisse sind deshalb in Form von Wahrscheinlichkeiten der Polarisations-eigenschaften des Testkanals nach dem Durchlaufen der optischen Übertragungstrecke gegeben – die wichtigsten hierbei sind die Verteilung der Polarisationszustände des Testkanals im Sinne der Gesamtheitsverteilung, die das vollständige Ensemble der Polarisationsparameter (Ausgangspolarisation, Doppelbrechungs- und PMD-Eigenschaften der Faser) umfasst und durch die nichtlineare Reduktion des Polarisationsgrades quantifiziert wird.

Wir glauben, dass die Ausdrücke für die Stärke der XPOLM-Störung, die durch die nichtlineare Depolarisierung beschrieben wird, ausreichend einfach sind, um einen umfassenden Einblick in diesen nichtlinearen Prozess und seine Abhängigkeit von verschiedensten Systemparametern zu gewähren. Die Integrale in den Ausdrücken können numerisch schnell genug gelöst werden, um eine große Anzahl von Systemen in begrenzter Zeit zu untersuchen. Wir haben auch demonstriert, wie man die nichtlineare Schwelle bestimmen kann, die zu einem gegebenen Wert der nichtlinearen Depolarisation gehört, und wie verschiedene Möglichkeiten der Störungsminderung, wie z.B. parallel oder paarweise orthogonal polarisierte Ausgangspolarisationen oder Restdispersion in den Verstärkerabschnitten, die Beeinträchtigungen durch XPOLM in Fasern mit niedriger PMD reduzieren können.

Um einen Schritt weiter zu gehen als die alles-umfassende Gesamtheitsverteilung, haben wir auch die Verteilung der nichtlinearen Störung (gegeben durch die nichtlineare Depolarisierung) innerhalb des Polarisations-Ensembles mittels analytischer und semi-analytischer Methoden untersucht. Die Kenntnis solch einer Verteilung ist hilfreich bei der Bestimmung des Anteils von gleichartigen Übertragungssystemen, die deutlich stärker als das statistische Mittel oder eine vorgegebene Grenze beeinträchtigt sind, oder alternativ für die Bestimmung des Mittelwertes, der zu einem vernachlässigbaren

Anteil extrem stark beeinträchtigt Systeme führt.

Zum Abschluss haben wir die Auswirkungen der nichtlinearen Depolarisierung eines Datensignals in zwei Fällen untersucht, die eine komplett unterschiedliche Herangehensweise notwendig machten. Diese sind als Beispiele für weitere Arbeiten auf dem Gebiet gedacht, die zeigen sollen, wie die Kenntnis der Verteilung der Polarisationszustände und der nichtlinearen Depolarisation benutzt werden kann, um auf die Beeinträchtigung der Übertragungsqualität zu schließen. Der erste Fall behandelt die DPSK-Übertragung, bei der die optischen Felder benachbarter Symbole miteinander interferieren müssen. Solch eine Interferenz ist nur möglich, wenn sich die Polarisationszustände ausreichend ähneln. Wir fanden heraus, dass die Änderungen der Polarisationszustände mit der Zeit nicht schnell genug sind, um die Signalqualität deutlich zu beeinträchtigen. Der zweite Fall behandelte die Übertragung mit Polarisationsmultiplex, bei der beide orthogonalen Subkanäle am Empfänger sauber getrennt werden müssen. Wir fanden heraus, dass die nichtlineare Depolarisation signifikantes Übersprechen der beiden Subkanäle verursachen kann. Die daraus resultierenden Beeinträchtigungen können deutlich schwerer ausfallen als die durch die Kreuz-Phasen-Modulation, sogar bei Übertragung mit dem phasenempfindlichen QPSK-Modulationsformat.

ACRONYMS

| | |
|-------|--|
| ACF | autocorrelation function |
| ACOVF | autocovariance function |
| ADC | analog-digital converter |
| ASK | amplitude-shift keying |
| AWGN | additive white Gaussian noise |
| BER | bit error ratio |
| CDF | cumulative density function |
| CSRZ | carrier-suppressed return-to-zero |
| CW | continuous wave |
| DC | directional coupler |
| DCM | dispersion compensation module |
| DEMUX | demultiplexer |
| DGD | differential group delay |
| DLI | delay-line interferometer |
| DOF | degrees of freedom |
| DOP | degree of polarization |
| DPSK | differential phase-shift keying |
| DQPSK | differential quaternary phase-shift keying |
| DSF | dispersion-shifted fiber |
| DWDM | dense wavelength-division multiplex |
| EDFA | Erbium-doped fiber amplifier |
| FEC | forward error correction |
| FSR | free spectral range |
| FWM | four-wave mixing |
| GVD | group velocity dispersion |
| ISI | inter-symbol interference |
| ITU | International Telecommunication Union |
| LO | local oscillator |
| MUX | multiplexer |
| NLSE | nonlinear Schrödinger equation |
| NLT | nonlinear threshold |
| NRZ | non-return-to-zero |
| NZDSF | non-zero dispersion-shifted fiber |
| OFDM | orthogonal frequency-division multiplex |

| | |
|-------|---|
| OSNR | optical signal-to-noise ratio |
| PBS | polarization beam splitter |
| PDE | partial differential equation |
| PDF | probability density function |
| PDL | polarization-dependent loss (also gain) |
| PMD | polarization mode dispersion |
| PMF | polarization-maintaining fiber |
| POLDM | polarization-division multiplex |
| PRBS | pseudo-random bit sequence |
| PSD | power spectral density |
| PSK | phase-shift keying |
| PSP | principal states of polarization |
| QPSK | quaternary phase-shift keying |
| RDPS | residual dispersion per span |
| RMS | root mean square |
| ROADM | reconfigurable optical add-drop multiplexer |
| ROSNR | required optical signal-to-noise ratio |
| RZ | return-to-zero |
| SOP | state of polarization |
| SPM | self-phase modulation |
| SMF | single-mode fiber |
| SSMF | standard single-mode fiber |
| STD | standard deviation |
| TOD | third-order dispersion |
| WDM | wavelength-division multiplex |
| XPM | cross-phase modulation |
| XPOLM | cross-polarization modulation |

LIST OF NOTATION

| | |
|--------------------------------|---|
| $(\cdot)^*$ | complex conjugate of a number |
| $(\cdot)^T$ | transpose of a matrix or vector |
| $(\cdot)^\dagger$ | complex-conjugate transpose of a matrix or vector |
| $ A\rangle$ | 2-dimensional Jones (<i>ket</i>) vector (p. 12) |
| $\langle A $ | conjugate transpose <i>bra</i> vector (p. 12) |
| $ \tilde{U}\rangle$ | Fourier transform of $ U\rangle$ (p. 32) |
| $ a\rangle$ | normalized Jones vector (p. 12) |
| $ e\rangle$ | normalized Jones vector which is part of a basis set (p. 12) |
| $\langle a b\rangle$ | inner product of $\langle a $ and $ b\rangle$ (p. 12) |
| $ a\rangle\langle a $ | projection operator (p. 13) |
| $\mathcal{E}[\cdot]$ | time average / expectation (pp. 14, 74) |
| $\mathcal{P}[\cdot]$ | probability of an event (p. 58) |
| $\langle \cdot \rangle$ | ensemble average / expectation (pp. 58, 73) |
| ∂_x^n | partial derivative $\partial^n / \partial x^n$ (p. 29) |
| B_2 | accumulated group velocity dispersion (p. 83) |
| \mathcal{C} | autocovariance function (ACOVF) (p. 88) |
| $\mathcal{C}_\nu^{\text{WO}}$ | walk-off related ACOVF (p. 89) |
| $\mathcal{C}_\nu^{\text{SOP}}$ | SOP decorrelation-related ACOVF (p. 89) |
| D | degree of polarization (DOP) (pp. 19, 78) |
| \mathcal{D} | population DOP (p. 80) |
| D_{GVD} | fiber GVD parameter (p. 36) |
| D_{PMD} | fiber PMD parameter (p. 58) |
| $\hat{\mathbf{H}}$ | birefringence vector (p. 40) |
| $\hat{\mathbf{H}}\times$ | cross-product operator of $\hat{\mathbf{H}}$ (p. 41) |
| $\bar{\mathbf{H}}$ | Jones matrix describing the birefringence eigenstates (p. 37) |
| $\bar{\mathbf{I}}$ | identity matrix (p. 69) |
| L_b | beat length (p. 39) |
| $L_{\text{SOP}\nu}$ | SOP decorrelation length of channel ν (p. 98) |
| $L_{\text{WO}\nu}$ | 1-symbol walk-off length of channel ν (p. 33) |
| $\bar{\mathbf{R}}$ | 3×3 rotation matrix (p. 23) |
| S_i | Stokes parameter i with $i \in \{0, 1, 2, 3\}$ (p. 14ff.) |
| s_i | normalized Stokes parameter i with $i \in \{1, 2, 3\}$ (p. 19) |
| \mathbf{S} | Stokes vector (p. 21) |
| $\mathbf{S}_{\nu\perp}$ | projection of $\hat{\mathbf{S}}_\nu$ onto the plane orthogonal to $\hat{\mathbf{S}}_p$ (p. 101) |

| | |
|--------------------------------|---|
| $\hat{\mathbf{S}}$ | state of polarization (SOP) in Stokes space (p. 19) |
| $\hat{\mathbf{S}}_P$ | state of polarization (SOP) of the probe channel with mean nonlinear motion removed (p. 75) |
| $\bar{\mathbf{T}}$ | frequency-dependent unitary transmission matrix (p. 50) |
| T_S | symbol duration = 1 / baud rate (p. 33) |
| $\bar{\mathbf{U}}$ | 2×2 unitary Jones matrix (p. 23) |
| \mathcal{V} | population variance parameter associated with XPOLM (p. 87) |
| $\mathcal{V}_\nu^{\text{SOP}}$ | population variance parameter associated with the SOP decorrelation (p. 97) |
| α | fiber attenuation coefficient (p. 29) |
| β_n | propagation coefficient $\partial_\omega^n \beta$ (p. 29) |
| Γ | accumulated loss and gain (p. 89) |
| Δt_ν | walk-off between channels ρ and ν (p. 33) |
| $\Delta \beta_0$ | differential phase per unit length parameter (p. 39) |
| $\Delta \beta_1$ | differential group delay per unit length parameter (p. 49) |
| $\Delta \Theta$ | differential angle in Stokes space (p. 135) |
| $\Delta \omega_\nu$ | frequency offset of channel ν from the probe channel ρ (p. 31) |
| θ | spherical zenith/colatitude coordinate (p. 22) |
| Θ | angular deviation of an SOP from the mean (p. 80) |
| $\bar{\mathbf{E}}$ | transformation matrix to remove the mean nonlinear SOP motion (p. 75) |
| σ_i | Pauli matrix i with $i \in \{1, 2, 3\}$ (p. 22) |
| $\vec{\sigma}$ | Pauli vector (p. 22) |
| τ | differential group delay (p. 50) |
| φ | spherical azimuth coordinate (p. 22) |
| ϕ | polarimetric parameter in Jones notation (p. 12) |
| χ | polarimetric parameter in Jones notation (p. 12) |
| $\mathbf{\Omega}$ | PMD vector (p. 53) |
| $\bar{\mathbf{\Omega}}$ | PMD eigenmode matrix in local fiber coordinates (p. 44) |

BIBLIOGRAPHY

- [Agrawal 2007] G. P. Agrawal, *Nonlinear Fiber Optics*, 4th ed., ser. Optics and Photonics. Academic Press, 2007.
- [Aiello 2006] A. Aiello and J. P. Woerdman, “Linear algebra for Mueller calculus,” Mar 2006. [Online] Available: <http://arxiv.org/abs/math-ph/0412061>
- [Andersen 1998] J. K. Andersen, J. W. Lou, G. A. Nowak, T. Xia, M. N. Islam, R. M. Fortenberry, S. A. Newton, “Path average measurements of optical fiber nonlinearity using solitons,” *Journal of Lightwave Technology*, vol. 16, no. 12, pp. 2328–2335, Dec 1998.
- [Antonelli 2006] C. Antonelli, A. Mecozzi, M. Brodsky, M. Boroditsky, “A simple analytical model for PMD temporal evolution,” in *Optical Fiber Communication Conference and Exposition (OFC) and National Fiber Optic Engineers Conference (NFOEC)*, 2006, paper OWJ4.
- [Arfken 2005] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, 6th ed. Elsevier, 2005.
- [Aroian 1947] L. A. Aroian, “The probability function of the product of two normally distributed variables,” *The Annals of Mathematical Statistics*, vol. 18, no. 2, pp. 265–271, 1947.
- [Barlow 1981] A. J. Barlow, J. J. Ramskov-Hansen, D. N. Payne, “Birefringence and polarization-mode dispersion in spun single-mode fibers,” *Applied Optics*, vol. 20, no. 17, pp. 2962–2968, Sep 1981.
- [Bayvel 2002] P. Bayvel and R. Killey, “Nonlinear optical effects in WDM transmission,” in *Optical Fiber Telecommunications IV B: Systems and Impairments*, I. P. Kaminov and T. Li, Eds. Academic Press, 2002.
- [Bellotti 1999] G. Bellotti, A. Bertaina, S. Bigo, “Dependence of self-phase modulation impairments on residual dispersion in 10-Gb/s-based terrestrial transmissions using standard fiber,” *IEEE Photonics Technology Letters*, vol. 11, no. 7, pp. 824–826, Jul 1999.
- [Bergano 2002] N. S. Bergano, “Undersea communication systems,” in *Optical Fiber Telecommunications IV B: Systems and Impairments*, I. P. Kaminov and T. Li, Eds. Academic Press, 2002.
- [Bertolini 2009] M. Bertolini, P. Serena, N. Rossi, A. Bononi, “Monte Carlo estimation of PDM-QPSK/OOK and DQPSK/OOK hybrid systems tolerance against nonlinear effects,” *IEEE Photonics Technology Letters*, vol. 21, no. 1, pp. 15–17, Jan 2009.

BIBLIOGRAPHY

- [Bertran-Pardo 2008a] O. Bertran-Pardo, J. Renaudier, H. Mardoyan, P. Tran, G. Charlet, S. Bigo, “Investigation of design options for overlaying 40Gb/s coherent PDM-QPSK channels over a 10Gb/s system infrastructure,” in *Optical Fiber Communication Conference and Exposition (OFC) and National Fiber Optic Engineers Conference (NFOEC)*, 2008, paper OTuM5.
- [Bertran-Pardo 2008b] O. Bertran-Pardo, J. Renaudier, G. Charlet, H. Mardoyan, P. Tran, S. Bigo, “Nonlinearity limitations when mixing 40-Gb/s coherent PDM-QPSK channels with preexisting 10-Gb/s NRZ channels,” *IEEE Photonics Technology Letters*, vol. 20, no. 15, pp. 1314–1316, Aug 2008.
- [Bertran-Pardo 2009] O. Bertran-Pardo, J. Renaudier, G. Charlet, P. Tran, H. Mardoyan, M. Salsi, S. Bigo, “Experimental assessment of interactions between nonlinear impairments and polarization-mode dispersion in 100-Gb/s coherent systems versus receiver complexity,” *IEEE Photonics Technology Letters*, vol. 21, no. 1, pp. 51–53, 2009.
- [Bononi 2003] A. Bononi, A. Vannucci, A. Orlandini, E. Corbel, S. Lanne, S. Bigo, “Degree of polarization degradation due to cross-phase modulation and its impact on polarization-mode dispersion compensators,” *Journal of Lightwave Technology*, vol. 21, no. 9, pp. 1903–1913, Sep 2003.
- [Born 1999] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th ed. Cambridge University Press, 1999.
- [Borne 2004] D. van den Borne, N. E. Hecker-Denschlag, G.-D. Khoe, H. de Waardt, “Cross phase modulation induced depolarization penalties in 2×10 Gbit/s polarization-multiplexed transmission,” in *30th European Conference on Optical Communication (ECOC)*, 2004, paper Mo.4.5.5.
- [Borne 2005] D. van den Borne, S. L. Jansen, S. Calabro, N. E. Hecker-Denschlag, G.-D. Khoe, H. de Waardt, “Reduction of nonlinear penalties through polarization interleaving in 2×10 Gb/s polarization-multiplexed transmission,” *IEEE Photonics Technology Letters*, vol. 17, no. 6, pp. 1337–1339, Jun 2005.
- [Borne 2007a] D. van den Borne, S. L. Jansen, E. Gottwald, P. M. Krummrich, G.-D. Khoe, H. de Waardt, “1.6-b/s/Hz spectrally efficient transmission over 1700 km of SSMF using 40×85.6 -Gb/s POLMUX-RZ-DQPSK,” *Journal of Lightwave Technology*, vol. 25, no. 1, pp. 222–232, Jan 2007.
- [Borne 2007b] D. van den Borne, C. R. S. Fludger, T. Duthel, T. Wuth, E. D. Schmidt, C. Schullien, E. Gottwald, G.-D. Khoe, H. de Waardt, “Carrier phase estimation for coherent equalization of 43-Gb/s POLMUX-NRZ-DQPSK transmission with 10.7-Gb/s NRZ neighbours,” in *33rd European Conference on Optical Communication (ECOC)*, 2007, paper We7.2.3.
- [Borne 2007c] D. van den Borne, private communication at *The 20th Annual Meeting of the IEEE Lasers and Electro-Optics Society*, Oct 2007.

BIBLIOGRAPHY

- [Borne 2009] D. van den Borne, V. A. J. M. Sleiffer, M. S. Alfiad, S. L. Jansen, T. Wuth, "POLMUX-QPSK modulation and coherent detection: the challenge of long-haul 100G transmission," in *35th European Conference on Optical Communication (ECOC)*, 2009, paper Tu.3.4.1.
- [Bosco 2000] G. Bosco, A. Carena, V. Curri, R. Gaudino, P. Poggiolini, 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.
- [Boroditsky 2006] M. Boroditsky, M. Bourd, M. Tur, "Effect of nonlinearities on PMD," *Journal of Lightwave Technology*, vol. 24, no. 11, pp. 4100–4107, Nov 2006.
- [Brodsky 2005] M. Brodsky, M. Boroditsky, P. Magill, N. J. Frigo, M. Tur, "A 'hinge' model for the temporal dynamics of polarization mode dispersion," in *17th Annual Meeting of the IEEE Lasers and Electro-Optics Society (LEOS)*, 2004, paper MJ5.
- [Brodsky 2008] M. Brodsky, N. J. Frigo, M. Tur, "Polarization mode dispersion," in *Optical Fiber Telecommunications V A: Components and Subsystems*, I. P. Kaminow, T. Li, and A. E. Willner, Eds. Academic Press, 2008.
- [Cartaxo 1999] A. V. T. Cartaxo, "Cross-phase modulation in intensity modulation-direct detection WDM systems with multiple optical amplifiers and dispersion compensators," *Journal of Lightwave Technology*, vol. 17, no. 2, pp. 178–190, Feb 1999.
- [Chandrasekhar 2007] S. Chandrasekhar and X. Liu, "Impact of channel plan and dispersion map on hybrid DWDM transmission of 42.7-Gb/s DQPSK and 10.7-Gb/s OOK on 50-GHz grid," *IEEE Photonics Technology Letters*, vol. 19, no. 22, pp. 1801–1803, Nov 2007.
- [Chernikov 1996] S. V. Chernikov and J. R. Taylor, "Measurement of normalization factor of n_2 for random polarization in optical fibers," *Optics Letters*, vol. 21, no. 19, pp. 1559–1561, Oct 1996.
- [Chraplyvy 1996] A. R. Chraplyvy, A. H. Gnauck, R. W. Tkach, J. L. Zyskind, J. W. Sulhoff, A. J. Lucero, Y. Sun, R. M. Jopson, F. Forghieri, R. M. Derosier, C. Wolf, A. R. McConnick, "1-Tb/s Transmission Experiment," *IEEE Photonics Technology Letters*, vol. 8, no. 9, pp. 1264–1266, Sep 1996.
- [Collings 2000] B. C. Collings and L. Boivin, "Nonlinear polarization evolution induced by cross-phase modulation and its impact on transmission systems," *IEEE Photonics Technology Letters*, vol. 12, no. 11, pp. 1582–1584, Nov 2000.
- [Corbel 2003] E. Corbel, J. P. Thiéry, S. Lanne, S. Bigo, A. Vannucci, A. Bononi, "Experimental statistical assessment of XPM impact on optical PMD compensator efficiency," in *Optical Fiber Communications Conference and Exhibit (OFC)*, 2003, paper ThJ2.

BIBLIOGRAPHY

- [Damask 2004] J. N. Damask, *Polarization optics in telecommunications*, Springer series in optical sciences. Springer Science and Business Media, 2004.
- [Dirac 1958] P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4th ed., ser. The International Series of Monographs on Physics. Oxford University Press, 1958.
- [Djordjevic 2008] I. Djordjevic, L. Minkov, H. Batshon, “Mitigation of linear and nonlinear impairments in high-speed optical networks by using LDPC-coded turbo equalization,” *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 6, pp. 73–83, Aug 2008.
- [Duguay 1969] M. A. Duguay and J. W. Hansen, “An ultrafast light gate,” *Applied Physics Letters*, vol. 15, no. 6, pp. 192–194, Sep 1969.
- [Duthel 2007] T. Duthel, C. R. S. Fludger, D. van den Borne, C. Schulien, E.-D. Schmidt, T. Wuth, E. de Man, G. D. Khoe, H. de Waardt, “Impairment tolerance of 111Gbit/s POLMUX-RZ-DQPSK using a reduced complexity coherent receiver with a T-spaced equaliser,” in *33rd European Conference on Optical Communication (ECOC)*, 2007, paper Mo1.3.2.
- [Essiambre 2002] R.-J. Essiambre, G. Raybon, B. Mikkelsen, “Pseudo-linear transmission of high-speed TDM signals: 40 and 160 Gb/s,” in *Optical Fiber Telecommunications IV B: Systems and Impairments*, I. P. Kaminov and T. Li, Eds. Academic Press, 2002.
- [Essiambre 2006] R.-J. Essiambre, P. J. Winzer, X. Q. Wang, W. Lee, C. A. White, E. C. Burrows, “Electronic predistortion and fiber nonlinearity,” *IEEE Photonics Technology Letters*, vol. 18, no. 17, pp. 1804–1806, Sep 2006.
- [Evangelides 1992] S. G. Evangelides, L. F. Mollenauer, J. P. Gordon, N. S. Bergano, “Polarization multiplexing with solitons,” *Journal of Lightwave Technology*, vol. 10, no. 1, pp. 28–35, Jan 1992.
- [Evans 2008] D. J. Evans and G. Morriss, *Statistical Mechanics of Nonequilibrium Liquids*, 2nd. ed. Cambridge University Press, 2008.
- [Feynman 1951] R. P. Feynman, “An operator calculus having applications in quantum electrodynamics,” *Physical Review*, vol. 84, no. 1, pp. 108–128, Oct 1951.
- [Fisher 1953] R. Fisher, “Dispersion on a sphere,” *Proceedings of the Royal Society of London, A*, vol. 217, no. 1130, pp. 295–305, May 1953.
- [Fludger 2008] C. R. S. Fludger, T. Duthel, D. van den Borne, “Coherent equalization and POLMUX-RZ-DQPSK for robust 100-GE transmission,” *Journal of Lightwave Technology*, vol. 26, no. 1, pp. 64–72, Jan 2008.

BIBLIOGRAPHY

- [Forestieri 2000] E. Forestieri, “Evaluating the error probability in lightwave systems with chromatic dispersion, arbitrary pulse shape and pre- and postdetection filtering,” *Journal of Lightwave Technology*, vol. 18, no. 11, pp. 1493–1503, Nov 2000.
- [Foschini 1991] G.J. Foschini and C.D. Poole, “Statistical theory of polarization dispersion in single mode fibers,” *Journal of Lightwave Technology*, vol. 9, no. 11, pp. 1439–1456, Jan 2008.
- [Galtarossa 2000] A. Galtarossa, L. Palmieri, A. Pizzinat, M. Schiano, T. Tambosso, “Measurement of local beat length and differential group delay in installed single-mode fibers,” *Journal of Lightwave Technology*, vol. 18, no. 10, pp. 1389–1394, Oct 2000.
- [Galtarossa 2005] A. Galtarossa and C.R. Menyuk, Eds., *Polarization Mode Dispersion*, ser. Optical and Fiber Communications Reports. Springer, 2005.
- [Glance 1987] B. Glance, “Polarization independent coherent optical receiver,” *Journal of Lightwave Technology*, vol. 5, no. 2, pp. 274–276, Feb 1987.
- [Godard 2003] D. Godard, “Self-recovering equalization and carrier tracking in two-dimensional data communication systems,” *IEEE Transactions on Communications*, vol. 28, no. 11, pp. 1867–1875, Jan 2003.
- [Goldstein 2003] D. Goldstein, *Polarized Light*, 2nd ed. Marcel Dekker, 2003.
- [Gordon 1990] J.P. Gordon and L.F. Mollenauer, “Phase noise in photonic communications systems using linear amplifiers,” *Optics Letters*, vol. 15, no. 23, pp. 1351–1353, Dec 1990.
- [Gordon 2000] 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.
- [Green 1954] M.S. Green, “Markoff random processes and the statistical mechanics of time-dependent phenomena. II. Irreversible processes in fluids,” *Journal of Chemical Physics*, vol. 22, no. 3, pp. 398–413, Mar 1954.
- [Hansryd 2000] J. Hansryd, H. Sunnerud, P.A. Andrekson, M. Karlsson, “Impact of PMD on four-wave-mixing-induced crosstalk in WDM systems,” *IEEE Photonics Technology Letters*, vol. 12, no. 9, pp. 1261–1263, Sep 2000.
- [Hecht 1999] J. Hecht, *City of Light: The Story of Fiber Optics (Sloan Technology Series)*. Oxford University Press, 2004.
- [Ho 2003] K.-P. Ho, “Performance degradation of phase-modulated systems due to nonlinear phase noise,” *IEEE Photonics Technology Letters*, vol. 15, no. 9, pp. 1213–1215, Sep 2003.

BIBLIOGRAPHY

- [Ho 2004] K.-P. Ho, "Error probability of DPSK signals with cross-phase modulation induced nonlinear phase noise," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 10, no. 2, pp. 421–427, Jun 2004.
- [Hughes 1995] B. D. Hughes, *Random Walks and Random Environments, Vol. 1: Random Walks*. Clarendon Press, 1995.
- [Hui 2008] R. Hui and M. O'Sullivan, *Fiber Optic Measurement Techniques*. Academic Press, 2008.
- [Irvine-Halliday 2000] D. Irvine-Halliday, M. Raziullah Khan, P.-G. Zhang, "Beat-length measurement of high-birefringence polarization-maintaining optical fiber using the DC Faraday magneto-optic effect," *Optical Engineering*, vol. 39, no. 5, pp. 1310–1315, May 2000.
- [Jeruchim 2000] M. C. Jeruchim, P. Balaban, K. S. Shanmugan, *Simulation of Communication Systems: Modeling, Methodology, and Techniques*, 2nd ed., ser. Information Technology: Transmission, Processing, and Storage. Kluwer Academic Publishers, 2000.
- [Jeunhomme 1989] L. B. Jeunhomme, *Single-mode Fiber Optics: Principles and Applications*, 2nd ed., ser. Optical Engineering. Marcel Dekker, 1990.
- [Jones 1941] R. C. Jones, "A new calculus for the treatment of optical systems," *Journal of the Optical Society of America*, vol. 31, no. 7, pp. 488–493, Jul 1941.
- [Karlsson 2000] M. Karlsson, J. Brentel, 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.
- [Karlsson 2006] M. Karlsson and H. Sunnerud, "Effects of nonlinearities on PMD-induced system impairments," *Journal of Lightwave Technology*, vol. 24, no. 11, pp. 4127–4137, Nov 2006.
- [Keiser 2003] G. Keiser, "Optical Fiber Communications," in *Wiley Encyclopedia of Telecommunications*, J. G. Proakis, Ed. John Wiley & Sons, Inc., 2003.
- [Kerr 1875] J. Kerr, "A new relation between electricity and light: dielectrified media birefringent," *Philosophical Magazine and Journal of Science*, vol. 50, no. 332, pp. 337–348, Nov 1875.
- [Khosravani 2001] R. Khosravani, Y. Xie, L. S. Yan, Y. W. Song, A. E. Willner, C. R. Menyuk, "Limitations to first-order PMD compensation in WDM systems due to XPM-induced PSP changes," in *Optical Fiber Communication Conference and Exhibit (OFC)*, 2001, paper WAA5.
- [Kikuchi 2001] N. Kikuchi, "Analysis of signal degree of polarization degradation used as control signal for optical polarization mode dispersion," *Journal of Lightwave Technology*, vol. 19, no. 4, pp. 481–486, Apr 2001.

BIBLIOGRAPHY

- [Kim 2003] H. Kim, “Cross-phase-modulation-induced nonlinear phase noise in WDM direct-detection DPSK systems,” *Journal of Lightwave Technology*, vol. 21, no. 8, pp. 1770–1774, Aug 2003.
- [Kogelnik 2002] H. Kogelnik, R. M. Jopson, L. E. Nelson, “Polarization-mode dispersion,” in *Optical Fiber Telecommunications IV B: Systems and Impairments*, I. P. Kaminov and T. Li, Eds. Academic Press, 2002.
- [Kogelnik 2005] H. Kogelnik, P. J. Winzer, L. E. Nelson, R. M. Jopson, M. Boroditsky, M. Brodsky, “First-order PMD outage for the hinge model,” *IEEE Photonics Technology Letters*, vol. 17, no. 6, pp. 1208–1210, Jun 2005.
- [Korneev 2002] N. Korneev, “Application of Magnus series for polarization evolution in fibers,” *Revista mexicana de física*, vol. 48, no. 3, pp. 250–254, Jun 2002.
- [Krummrich 2002] P. Krummrich, E. Gottwald, K. Kotten, H. Geiger, C. Glingener, C. Scheerer, G. Fischer, “Übertragungsstrecken mit Wellenlängenmultiplexbetrieb,” in *Optische Kommunikationstechnik: Handbuch für Wissenschaft und Industrie*, E. Voges and K. Petermann, Eds. Springer, 2002.
- [Krummrich 2004] P. M. Krummrich and K. Kotten, “Extremely fast (microsecond timescale) polarization changes in high speed long haul WDM transmission systems,” in *Optical Fiber Communications Conference and Exhibit (OFC)*, 2004, paper FI3.
- [Krummrich 2005] P. M. Krummrich, E. D. Schmidt, W. Weiershausen, A. Mattheus, “Field trial results on statistics of fast polarization changes in long haul WDM transmission systems,” in *Optical Fiber Communication Conference and Exposition (OFC) and National Fiber Optic Engineers Conference (NFOEC)*, 2005, paper OThT6.
- [Kubo 1957] R. Kubo, “Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems,” *Journal of the Physical Society of Japan*, vol. 12, no. 6, pp. 570–586, Jun 1957.
- [Lee 2002] J. H. Lee, K. J. Park, C. H. Kim, Y. C. Chung, “Effects of nonlinear crosstalk in optical PMD compensation,” *IEEE Photonics Technology Letters*, vol. 14, no. 8, pp. 1082–1084, Aug 2002.
- [Lefrançois 2007] M. Lefrançois, F. Houndonougbo, T. Fauconnier, G. Charlet, S. Bigo, “Cross comparison of the nonlinear impairments caused by 10Gbit/s neighboring channels on a 40Gbit/s channel modulated with various formats, and over various fiber types,” in *Optical Fiber Communication Conference and Exposition (OFC) and National Fiber Optic Engineers Conference (NFOEC)*, 2007, paper JThA44.

BIBLIOGRAPHY

- [Leven 2008] A. Leven, N. Kaneda, Y.-K. Chen, “A real-time CMA-based 10 Gb/s polarization demultiplexing coherent receiver implemented in an FPGA,” in *Optical Fiber Communication Conference and Exposition (OFC) and National Fiber Optic Engineers Conference (NFOEC)*, 2008, paper OTuO2.
- [Lin 2004] Q. Lin and G. P. Agrawal, “Vector theory of cross-phase modulation: Role of nonlinear polarization rotation,” *IEEE Journal of Quantum Electronics*, vol. 40, no. 7, pp. 958–964, Jul 2004.
- [Magnus 1954] W. Magnus, “On the exponential solution of differential equations for a linear operator,” *Communications on Pure and Applied Mathematics*, vol. 7, no. 4, pp. 649–673, Nov 1954.
- [Mamyshev 1996] P. V. Mamyshev and L. F. Mollenauer, “Pseudo-phase-matched four-wave mixing in soliton wavelength-division multiplexing transmission,” *Optics Letters*, vol. 21, no. 6, pp. 396–398, Mar 1996.
- [Manakov 1974] S. V. Manakov, “On the theory of two-dimensional stationary self-focusing of electromagnetic waves,” *Soviet Physics – JETP*, vol. 38, no. 2, pp. 248–253, Feb 1974.
- [Marcuse 1997] D. Marcuse, C. R. Menyuk, 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, Sep 1997.
- [McKinstrie 2007] C. J. McKinstrie, H. Kogelnik, G. G. Luther, L. Schenato, “Stokes-space derivations of generalized Schrödinger equations for wave propagation in various fibers,” *Optics Express*, vol. 15, no. 17, pp. 10 964–10 983, Aug 2007.
- [Menyuk 1989] C. R. Menyuk, “Pulse Propagation in an Elliptically Birefringent Kerr Medium,” *IEEE Journal of Quantum Electronics*, vol. 25, no. 12, pp. 2674–2682, Dec 1989.
- [Menyuk 1994] C. R. Menyuk and P. K. A. Wai, “Polarization evolution and dispersion in fibers with spatially varying birefringence,” *Journal of the Optical Society of America B*, vol. 11, no. 7, pp. 1288–1296, Jul 1994.
- [Menyuk 1999] C. R. Menyuk, “Application of multiple-length-scale methods to the study of optical fiber transmission,” *Journal of Engineering Mathematics*, vol. 36, no. 1-2, pp. 113–136, Aug 1999.
- [Menyuk 2002] C. R. Menyuk, B. S. Marks, I. T. Lima, J. Zweck, Y. Sun, G. M. Carter, D. Wang, “Polarization effects in long-haul undersea systems,” in *Undersea Fiber Communication Systems*, J. Chesnoy, Ed., ser. Optics and Photonics. Academic Press, 2002.
- [Menyuk 2006] 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.

- [Meyer 2000] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, 2000.
- [Milivojevic 2005] B. Milivojevic, A. F. Abas, A. Hidayat, S. Bhandare, D. Sandel, R. Noé, M. Guy, M. Lapointe, “1.6-b/s/Hz 160-Gb/s 230-km RZ-DQPSK polarization multiplex transmission with tunable dispersion compensation,” *IEEE Photonics Technology Letters*, vol. 17, no. 2, pp. 495–497, Feb 2005.
- [Mollenauer 1995] L. F. Mollenauer, J. P. Gordon, F. Heismann, “Polarization scattering by soliton-soliton collisions,” *Optics Letters*, vol. 20, no. 20, pp. 2060–2062, Oct 1995.
- [Mollenauer 1997] L. F. Mollenauer, J. P. Gordon, P. V. Mamyshev, “Solitons in high bit-rate long-distance transmission,” in *Optical Fiber Telecommunications III A*, I. P. Kaminov and T. L. Koch, Eds. Academic Press, 1997.
- [Mollenauer 1998] L. F. Mollenauer, P. V. Mamyshev, “Massive wavelength-division multiplexing with solitons,” *IEEE Journal of Quantum Electronics*, vol. 34, no. 11, pp. 2089–2102, Nov 1998.
- [Möller 2001] L. Möller, L. Boivin, S. Chandrasekhar, L. L. Buhl, “Setup for demonstration of cross channel-induced nonlinear PMD in WDM system,” *Electronics Letters*, vol. 37, no. 5, pp. 306–308, Mar 2001.
- [Nelson 2000] L. E. Nelson and H. Kogelnik, “Coherent crosstalk impairments in polarization multiplexed transmission due to polarization mode dispersion,” *Optics Express*, vol. 7, no. 10, pp. 350–361, Nov 2000.
- [Nelson 2009] L. E. Nelson, S. L. Woodward, S. Foo, X. Zhou, M. D. Feuer, D. Hanson, D. McGhan, H. Sun, M. Moyer, M. O’Sullivan, P. D. Magill, “Performance of a 46-Gbps dual-polarization QPSK transceiver with real-time coherent equalization over high PMD fiber,” *Journal of Lightwave Technology*, vol. 27, no. 3, pp. 158–167, Feb 2009.
- [Pachnicke 2005] S. Pachnicke, J. Reichert, E. Voges, “Impact of POLXPM on Optical Transmission Systems with Arbitrary Dispersion Compensation,” in *Opto-Electronics and Communications Conference (OECC)*, 2005.
- [Pachnicke 2006] S. Pachnicke, E. Voges, E. De Man, E. Gottwald, J. Reichert, “Experimental investigation of XPM-induced birefringence in mixed-fiber transparent optical networks,” in *Optical Fiber Communication Conference and Exposition (OFC) and National Fiber Optic Engineers Conference (NFOEC)*, 2006, paper JThB9.
- [Pan 2002] Z. Pan, Q. Yu, A. E. Willner, “XPM-induced polarization-state fluctuations in WDM systems and their mitigation,” in *Optical Fiber Communication Conference and Exhibit (OFC)*, 2002, paper ThA7.

BIBLIOGRAPHY

- [Pan 2004] Z. Pan, Q. Yu, Y. Arieli, A. E. Willner, “The effects of XPM-induced fast polarization-state fluctuations on PMD compensated WDM systems,” *IEEE Photonics Technology Letters*, vol. 16, no. 8, pp. 1963–1965, Aug 2004.
- [Papoulis 1991] A. Papoulis, Probability, Random Variables, and Stochastic Processes, 3rd ed., ser. Communications and signal processing McGraw-Hill, 1991.
- [Pearson 1920] K. Pearson, “Notes on the history of correlation,” *Biometrika*, vol. 13, no. 1, pp. 25–45, Oct 1920.
- [Philips 2005] M. R. Phillips and S. L. Woodward, “Cross-polarization modulation: Theory and measurement of a two-channel WDM system,” *IEEE Photonics Technology Letters*, vol. 17, no. 10, pp. 2086–2088, Oct 2005.
- [Philips 2006] M. R. Phillips, S. L. Woodward, R. L. Smith, “Cross-Polarization Modulation: Theory and Measurement in Subcarrier-Modulated WDM Systems,” *Journal of Lightwave Technology*, vol. 24, no. 11, pp. 4089–4099, Nov 2006.
- [Poincaré 1892] H. Poincaré, *Théorie Mathématique de la Lumière*, vol. 2 Gauthiers-Villars, 1892.
- [Poole 1986] C. D. Poole and R. E. Wagner, “Phenomenological approach to polarisation dispersion in long single-mode fibres,” *Electronics Letters*, vol. 22, no. 19, pp. 1029–1030, Sep 1986.
- [Poole 1988] C. D. Poole, “Statistical treatment of polarization dispersion in single-mode fiber,” *Optics Letters*, vol. 13, no. 8, pp. 687–687, 1988.
- [Poole 1997] C. D. Poole and J. Nagel, “Polarization effects in lightwave systems,” in *Optical Fiber Telecommunications III A*, I. P. Kaminov and T. L. Koch, Eds. Academic Press, 1997.
- [Primak 2004] S. Primak, V. Kontorovich, V. Lyandres, *Stochastic Methods and their Applications to Communications* John Wiley & Sons, Inc., 2004.
- [Proakis 2001] J. G. Proakis, *Digital Communications*, 4th ed. McGraw-Hill, 2001.
- [Renaudier 2009] J. Renaudier, O. Bertran-Pardo, G. Charlet, M. Salsi, M. Bertolini, P. Tran, H. Mardoyan, S. Bigo, “On the required number of WDM channels when assessing performance of 100 Gb/s coherent PDM-QPSK overlaying legacy systems,” in *35th European Conference on Optical Communication (ECOC)*, 2009, paper Tu3.4.5.
- [Riley 2006] K. F. Riley, M. P. Hobson, S. J. Bence, *Mathematical Methods for Physics and Engineering*, 3rd ed. Cambridge University Press, 2006.
- [Roberts 1960] P. H. Roberts and H. D. Ursell, “Random walk on a sphere and on a Riemannian manifold,” *Philosophical Transactions for the Royal Society of London. Series A*, vol. 252, no. 1012, pp. 317–356, Mar 1960.

BIBLIOGRAPHY

- [Seimetz 2009] M. Seimetz, *High-Order Modulation for Optical Fiber Transmission*, ser. Springer series in optical sciences Springer, 2009.
- [Shibata 1987] N. Shibata, R. Braun, R. Waarts, “Phase-mismatch dependence of efficiency of wave generation through four-wave mixing in a singlemode optical fiber,” *IEEE Journal of Quantum Electronics*, vol. 23, no. 7, pp. 1205–1210, Jul 1987.
- [Simon 2002] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables*, ser. Springer international series in engineering and computer science Springer, 2002.
- [Smith 1997] N. J. Smith, N. J. Doran, W. Forysiak, F. M. Knox, “Soliton transmission using periodic dispersion compensation,” *Journal of Lightwave Technology*, vol. 15, no. 10, pp. 1808–1822, Oct 1997.
- [Soong 2004] T. T. Soong, *Fundamentals of Probability and Statistics for Engineers* John Wiley & Sons, 2004.
- [Stokes 1852] 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, Feb 1852.
- [Stolen 1973] R. H. Stolen and A. Ashkin, “Optical Kerr effect in glass waveguide,” *Applied Physics Letters*, vol. 22, no. 6, pp. 294–296, Mar 1973.
- [Vannucci 2003] A. Vannucci, A. Bononi, A. Orlandini, “A simple formula for the degree of polarization degraded by XPM and its experimental validation,” in *Optical Fiber Communications Conference and Exhibit (OFC)*, 2003, paper ThJ1.
- [Viterbi 1983] A. J. Viterbi and A. M. 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.
- [Wai 1996] 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.
- [Weber 2010] C. Weber, “Electronic Precompensation of Dispersion and Nonlinearities in Fibre-Optic Transmission Systems,” Ph.D. thesis, Technische Universität Berlin, Berlin, 2010. Available: <http://opus.kobv.de/tuberlin/>
- [Wiersema 2008] U. F. Wiersema, *Brownian Motion Calculus*, ser. Wiley finance series. John Wiley & Sons, 2008.

BIBLIOGRAPHY

- [Winter 2009a] M. Winter, K. Petermann, D. Setti, “Investigation of nonlinear DPSK fading due to cross-polarization modulation,” *Optical Fiber Communication Conference and Exposition (OFC) and National Fiber Optic Engineers Conference (NFOEC)*, 2009, paper OTuD3.
- [Winter 2009b] M. Winter, C.-A. Bunge, D. Setti, K. Petermann, “A statistical treatment of cross-polarization modulation in DWDM systems,” *Journal of Lightwave Technology*, vol. 27, no. 17, pp. 3739–3751, Sep 2009.
- [Winzer 2004] P. J. Winzer, H. Kogelnik, K. Ramanan, “Precise outage specifications for first-order PMD,” *IEEE Photonics Technology Letters*, vol. 16, no. 2, pp. 449–451, 2004.
- [Winzer 2006] P. J. Winzer and R. J. Essiambre, “Advanced optical modulation formats,” *Proceedings of the IEEE*, vol. 94, no. 5, pp. 952–985, May 2006.
- [Winzer 2008] P. J. Winzer and R. J. Essiambre, “Advanced optical modulation formats,” in *Optical Fiber Telecommunications V B: Systems and Networks*, I. P. Kaminov, T. Li, A. E. Willner, Eds. Academic Press, 2008.
- [Woodward 2004] S. L. Woodward, C. Roby, M. R. Phillips, P. D. Magill, M. Boroditsky, “RF Measurement and theory of cross-polarization modulation,” in *Optical Fiber Communications Conference and Exhibit (OFC)*, 2004, paper ThF3.
- [Xiao 2007] X. Xiao, S. Gao, Y. Tian, C. Yang, “Optimization of the net residual dispersion for self-phase modulation-impaired systems by perturbation theory,” *Journal of Lightwave Technology*, vol. 25, no. 3, pp. 929–937, Mar 2007.
- [Xie 2003] C. J. Xie, L. Möller, D. C. Kilper, L. F. Mollenauer, “Effect of cross-phase-modulation-induced polarization scattering on optical polarization mode dispersion compensation in wavelength-division-multiplexed systems,” *Optics Letters*, vol. 28, no. 23, pp. 2303–2305, Dec 2003.
- [Xie 2009] C. Xie, “Interchannel nonlinearities in coherent polarization-division-multiplexed quadrature-phase-shift-keying systems,” *IEEE Photonics Technology Letters*, vol. 21, no. 5, pp. 274–276, Mar 2009.
- [Zyskind 1997] J. L. Zyskind, J. A. Bagel, H. D. Kidorf, “Erbium-Doped Fiber Amplifiers for Optical Communications,” in *Optical Fiber Telecommunications III B*, I. P. Kaminov and T. L. Koch, Eds. Academic Press, 1997.

AUTHOR'S BIBLIOGRAPHY

M. Winter, D. Kroushkov, K. Petermann, "Polarization-multiplexed transmission system outage due to nonlinearity-induced depolarization," in *36th European Conference on Optical Communication (ECOC)*, Turino, Italy, invited paper Th.10.E.3, 19–23 Sep 2010.

M. Winter, D. Kroushkov, K. Petermann, "Cross-polarization modulation in polarization-division multiplexed transmission systems," in *The IEEE Photonics Society Summer Topical Meeting*, Playa del Carmen, Mexico, invited paper MA2.1, 19–21 Jul 2010.

M. Winter, K. Petermann, D. Setti, "Cross-polarization modulation in polarization-division multiplex transmission," *IEEE Photonics Technology Letters*, vol. 22, no. 8, pp. 538–540, Apr 2010.

M. Winter, C.-A. Bunge, D. Setti, K. Petermann, "A statistical treatment of cross-polarization modulation in DWDM systems," *Journal of Lightwave Technology*, vol. 27, no. 17, pp. 3739–3751, Sep 2009.

M. Winter, D. Setti, K. Petermann, "Interchannel nonlinearities in polarization-multiplexed transmission," in *35th European Conference on Optical Communication (ECOC)*, Vienna, Austria, paper Th.10.4.4, 20–24 Sep 2009.

M. Winter, K. Petermann, D. Setti, "Investigation of nonlinear DPSK fading due to cross-polarization modulation," in *Optical Fiber Communication Conference and Exposition (OFC) and National Fiber Optic Engineers Conference (NFOEC)*, San Diego, USA, paper OTuD3, 22–26 Mar 2009.

C. Weber, C.-A. Bunge, M. Winter, K. Petermann, "Fibre nonlinearities in 10 and 40 Gb/s electronically dispersion precompensated WDM transmission," in *Optical Fiber Communication Conference and Exposition (OFC) and National Fiber Optic Engineers Conference (NFOEC)*, San Diego, USA, paper OTuD2, 22–26 Mar 2009.

J. K. Fischer, M. Winter, K. Petermann, "Scaling of nonlinear threshold with fiber type and channel spacing in WDM transmission systems," in *Optical Fiber Communication Conference and Exposition (OFC) and National Fiber Optic Engineers Conference (NFOEC)*, San Diego, USA, paper JThA41, 22–26 Mar 2009.

M. Winter, C.-A. Bunge, K. Petermann, D. Setti, "Analysis of cross-polarization modulation in dispersion-managed DWDM systems," in *34th European Conference on Optical Communication (ECOC)*, Brussels, Belgium, paper We.3.E.3, 21–25 Sep 2008.

- M. Winter, C.-A. Bunge, K. Petermann, D. Setti, "Impairments in polarization-multiplexed DWDM channels due to cross-polarization modulation," in *The 21th Annual Meeting of the IEEE Lasers and Electro-Optics Society*, Newport Beach, USA, paper WH3, 9–13 Nov 2008.
- M. Winter, C.-A. Bunge, K. Petermann, D. Setti, "Analysis of cross-polarization modulation in dispersion-managed DWDM systems," in *The 20th Annual Meeting of the IEEE Lasers and Electro-Optics Society*, Orlando, USA, paper WFF2, 21–25 Oct 2007.
- G. Lehmann, R. H. Derksen, C. Schubert, M. Winter, "100 Gigabit Ethernet transmission - physical layer issues," in *Optical Fiber Communication Conference and Exposition (OFC) and National Fiber Optic Engineers Conference (NFOEC)*, Anaheim, USA, paper OThD5, 25–29 Mar 2007.
- A. Kilian, R. Hauffe, M. Winter, P. Runge, J. Kuhmann, C. G. Greisen, S. Weichel, L. Shiv, M. Heschel, "Silicon micro-machined hermetic packaging technology for optical subassemblies," in *The 19th Annual Meeting of the IEEE Lasers and Electro-Optics Society*, paper MJ4, Oct 2006.
- M. Winter, R. Hauffe, A. Kilian, P. Runge, D. Reznik, C. G. Greisen, S. Weichel, L. Shiv, M. Heschel, "Application and RF performance of integrated resistor structures on non-planar topologies in micro-machined silicon packages for transmitter optical subassemblies," in *Proceedings of the 56th Electronic Components and Technology Conference (ECTC)*, pp. 1560–1566, 30 May–2 Jun 2006.
- M. Winter, R. Hauffe, A. Kilian, "Simplified optical coupling and alignment scheme for cost-effective 10 Gbit/s TOSA modules based on edge emitters hermetically packaged in micro-machined silicon structures," in *Optical Fiber Communication Conference and Exposition (OFC) and National Fiber Optic Engineers Conference (NFOEC)*, Anaheim, USA, paper OThU8, 6–11 Mar 2005.
- R. Hauffe, A. Kilian, M. Winter, L. Shiv, G. Elger, M. Heschel, J. Kuhmann, S. Isaacs, S. Weichel, P. Gaal, H. Korth, A. Hase, "Optimized micro-via technology for high density and high frequency (>40GHz) hermetic through-wafer connections in silicon substrates," in *Proceedings of the 55th Electronic Components and Technology Conference (ECTC)*, pp. 324–330, 31 May–3 Jun 2005.
- A. Hodžić, N. Hecker-Denschlag, M. Winter, S. L. Jansen, K. Saucke, K. Petermann, "10 Gbit/s-based NRZ DWDM systems using polarisation switching in single wavelength channel," *Electronics Letters*, vol. 39, no. 18, pp. 1329–1330, 2003.

PATENTS

- M. Winter, "Optical I-Q-Modulator," European patent application no. 08104643.5 - 2415
- M. Winter, R. Hauffe, A. Kilian "Optical Module Hermetically Packaged in Micro-Machined Structures," United States patent no. US 7,543,999 B2