APPROVAL SHEET

Title of Thesis: **Polarization effects in dense WDM system**

Name of Candidate: Ding Wang
                     Doctor of Philosophy, 2000

Dissertation and Abstract Approved: _____

                                      Professor Curtis R. Menyuk
                                      Computer Science and Electrical Engineering

Date Approved: _____

Curriculum Vitæ

Name: Ding Wang

Permanent Address: Jizhong Wang,
Affiliated Hospital of North Sichuan Medical College,
Nanchong, Sichuan,
P. R. China, 637000.

Degree and date to be conferred: Ph.D., May, 2000.

Date of Birth: June 8, 1972.

Place of Birth: Nanchong, Sichuan, P. R. China.

Secondary education: Nanchong Middle School
Nanchong, Sichuan, P. R. China,
July, 1989.

Collegiate institutions attended: 1989–1993, University of Electronic Science and Technology of China (UESTC), B.S., July, 1993,
1993–1995, Institute of Applied Physics at UESTC,
1995–1997, University of Maryland Baltimore County, M.S., August, 1997,
1993–2000, University of Maryland Baltimore County, Ph.D., May, 2000.

Major: Photonics.

Professional publications:

[1] D. Wang, J. Wang, F. Xu, and Y. Liu, "Matrix analysis of thermal blooming in liquid and its new phenomenon," J. Univ. Electronic Science & Technology of China, vol. 20, pp. 155–157, 1994.

[2] D. Wang, E. A. Golovchenko, A. N. Pilipetskii, and C. R. Menyuk, "The nonlinear optical loop mirror demultiplexer based on standard communication fiber," J. Lightwave Tech., vol. 15, pp. 642–646, 1997.

[3] C. R. Menyuk, D. Wang, and A. N. Pilipettski, "Repolarization of polarization scrambled optical signals due to polarization dependent loss," Photonics Tech. Lett., vol. 9, pp. 1247–1249, 1997.

[4] C. R. Menyuk, D. Wang, and A. N. Pilipettski, "Repolarization of polarization scrambled optical signals due to polarization dependent loss," European Conf. Opt. Comm., pp. 3.275–3.278, UK, 1997.

[5] D. Wang and C. R. Menyuk "A reduced model for the polarization evolution in WDM systems," Optics Lett., vol. 23, pp. 1677–1679, 1998.

[6] D. Wang and C. R. Menyuk, "Polarization evolution due to the Kerr Nonlinearity and chromatic dispersion," J. Lightwave Technol., vol. 17, pp. 2520–2529, 1999.

[7] D. Wang , E. A. Golovchenko, and A. N. Pilipetskii, "100 Gbit all-optical NOLM demultiplexer based on standard communication fiber," Optical Society of America Annual Meeting, Rochester, NY, 1996.

[8] E. A. Golovchenko, D. Wang, A. N. Pilipetskii, and C. R. Menyuk, "Raman effect in the standard communication fiber," Optical Society of America Annual Meeting, Rochester, NY, 1996.

[9] D. Mahgereftec, H-Yun Yu, D. L. Butler, J. Goldhar, D. Wang, E. A. Golovchenko, A. M. Pilipetskii, and C.R. Menyuk, "Effect of polarization mode dispersion on Raman effect in standard communication fiber," Conf. on Lasers and Electro-optics, pp. 447–448, Baltimore, MD, 1997.

[10] D. Wang, E. A. Golovchenko, A. N. Pilipetskii, and C. R. Menyuk, "Filter effect on the channel crosstalk in soliton transmission," Optical Society of America Annual Meeting, Long Beach, CA, 1997.

[11] D. Wang and C. R. Menyuk, "Nonlinear evolution of the polarization states in WDM channels," Opt. Fiber Comm. Conf., pp. 125–127, San Jose, CA, 1998.

[12] D. Wang and C. R. Menyuk, "Evolution of polarization states in WDM systems," Optical Society of America Annual Meeting, Baltimore, MD, 1998.

[13] C. R. Menyuk, V. S. Grigoryan, R-M Mu, D. Wang, and T. Yu "Modeling high-data-rate optical fiber communication systems," short course at Opt. Fiber Comm. Conf., San Diego, CA, 1999.

[14] D. Wang and C. R. Menyuk, "Polarization effects in dense WDM system," Conf. on Lasers and Electro-optics pp. 39–40, Baltimore, MD, 1999.

[15] Y. Sun, D. Wang, P. Sinha, G. Carter, and C. R. Menyuk "Evolution of polarization in the experimental loop of dispersion managed solitons," Optical Society of America Annual Meeting, Santa Clara, CA, 1999.

[16] C. R. Menyuk and D. Wang, "A model for channel fading in undersea system," workshop at European Conf. Opt. Comm., Nice, France, 1999.

[17] D. Wang and C. R. Menyuk, "Outage due to polarization effects in a transoceanic system," Opt. Fiber Comm. Conf., Baltimore, MD, 2000.

[18] C. R. Menyuk, D. Wang, Ronald Holzöhner, I. T. Lima, Jr., E. Ibragimov, and V. S. Grigoryan, "Polarization mode dispersion in optical transmission systems," tutorial at Opt. Fiber Comm. Conf., Baltimore, MD, 2000.

[19] Y. Sun, D. Wang, P. Sinha, G. Carter, and C. R. Menyuk "Polarization evolution in a 107 km dispersion-managed recirculating loop," Conf. on Lasers and Electro-optics, San Francisco, CA, 2000.

Abstract

Title of Dissertation: **Polarization effects in dense WDM system**

Ding Wang, Doctor of Philosophy, 2000

Dissertation directed by: Professor Curtis R. Menyuk
Computer Science and Electrical Engineering

Wavelength division multiplexed (WDM) systems have become increasingly important in long-distance optical communications since the invention of the erbium-doped fiber amplifier. Polarization effects play an important role in these systems since they can lead to channel fading in which an entire channel may drop out. To investigate fading, it is necessary to study a large number of different cases of the randomly varying birefringence in optical fibers to determine the Q-values. It is not feasible to do so using complete simulations; so, my colleague and I have developed a reduced model in which one only follows the Stokes vectors for each channel. I have validated this model to the extent possible using full simulations. I have also validated it using single channel, recirculating loop experiments. Finally I applied this model to determining the distribution of Q-values and the probability of fading over the lifetime of an undersea cable system for realistic parameters.

Polarization effects in dense WDM system

by

Ding Wang

Dissertation submitted to the Faculty of the Graduate School
of the University of Maryland in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2000

Dedication

To my parents and sister

Acknowledgements

I wish to express my sincere gratitude to my dissertation advisor, Professor C. R. Menyuk for his guidance and encouragement throughout the course of my master's thesis work. His insightful vision, enthusiasm, quest for perfection, and unwaivering dedication to scientific research are infectious and inspiring. It has been my good fortune and highest honor to have had him as my academic advisor.

I would like to express my sincere appreciation to the other dissertation committee members, Dr. Gary Carter, Dr. F. S. Choa, Dr. Li Yan, and Dr. Yan-Hua Shih for their time and effort.

Any accomplishment of mine, be it academic or non-academic, will always be dedicated to my wife, my parents, and my sister.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Polarization effects in optical fiber transmission systems are due to birefringence in both the optical fiber itself and the components like the WDM couplers that are used in current amplifier systems. Birefringence in the optical fiber is due to the accidental loss of degeneracy of the two orthogonal polarization modes that exist in a single-mode fiber and is both weak and randomly varying. By contrast, materials like $LiNbO_3$ that are used in optical components often have a strong polarization dependence in which case these components are strongly birefringent along a fixed axis. Polarization effects in both optical components and optical fibers have become important in recent years because the advent of the erbium-doped fiber amplifiers implies that these effects accumulate over hundreds of kilometers in terrestrial systems and thousands of kilometers in undersea systems.

It is generally assumed that the principal polarization effects that lead to transmission impairment are polarization dependent loss, polarization dependent gain, and polarization mode dispersion [1]. Polarization dependent loss is caused by the strong polarization dependence in optical components such as WDM couplers, gratings, fiber

1

couplers, isolators, filters, and optical switches. Polarization dependent gain is caused by polarization hole-burning in the Er-doped fiber amplifiers. Finally, polarization mode dispersion is principally caused by the randomly varying birefringence in the optical fiber, although the differential group delay of the optical components can contribute under some circumstances. The combination of these effects contribute to channel fading [2] which may sometimes lead to the loss of an entire channel in a long-distance, high-data-rate wavelength-division-multiplexed (WDM) system. Polarization scrambling greatly ameliorates the effects of polarization impairments [3], but under some circumstances the signal can repolarize [4], [5], leading once more to polarization effects.

Today, WDM systems can carry 100 10 Gbits/sec channels through a 10,000 km distance [6]. The effect of the randomly varying birefringence is constantly changing in these systems as the internal fiber orientations change on a time scale that can vary between milliseconds and hours. It is not computationally feasible to carry out full simulations for the many channels and many realizations that are required to accurately calculate the penalties. Moreover, as optical networks become increasingly prevalent, these difficulties will be compounded by the greater complexity of these new systems. Finding effective ways to calculate the penalties due to polarization effects in long point-to-point systems and in optical networks has become a very important issue. In this thesis, I will propose a new model, based on following just the Stokes parameters of each channel rather than the full behavior in the time domain. I will carefully validate this model to the extent possible using full simulations and recirculating loop experiments. Finally, I will apply this model to calculate the polarization penalties in trans-oceanic systems with up to 40 channels, simulating up to

10^6 fiber realizations. While the application in this thesis is to point-to-point systems, we expect this model to be of great utility in modeling future optical networks due to its simplicity and efficiency.

Chapter 2

Polarization mode dispersion (PMD)

Polarization model dispersion (PMD) is caused by the birefringence in optical fibers, and the random rotation of their axes along the the transmission line. The birefringence causes a group velocity difference in two orthogonal polarizations, so that different frequency components experience different phase changes. These different phase changes, coupled to the random changes of the fiber axes, eventually lead to pulse spreading and loss of information. To the first order, I can physically explain and calculate this pulse broadening using the concept of principal states [7]. A narrowband input signal of a particular optical fiber can be projected onto two orthogonally-polarized principal states. These states depend on both the central frequency of the signal and the length of propagation as well as the fiber in question. Effectively, each input pulse splits into two pulses, each of which propagates at a slightly different speed so that the original pulse spreads. If at the input the signal can be completely projected into just one principal state, it will not be broadened by PMD to the first order. Current terrestrial fiber networks are still using old fibers which may have large PMD so that PMD emulation and compensation are very im-

portant topics [24]. However the quality of fiber is improving very fast, so that the PMD of current fiber can be as low as 0.02 ps/km$^{1/2}$ or even lower.

Service providers generally install new fiber when they create new trans-oceanic systems, rather than attempting to retrofit their legacy systems in contrast to what is done with terrestrial systems. Pulse spreading due to PMD is not a problem in systems using current fiber up to data rates of 40 Gbits/s per WDM channel. Since my research focuses on trans-oceanic systems at data rates up to 10 Gbits/s per channel, which is the maximum in current systems, I will not be concerned with pulse spreading due to PMD in this dissertation.

This limitation is required in order to only consider a single set of Stokes parameters for the signal in each channel (as well as an additional set for the noise). However, this limitation does not imply that PMD is unimportant. Over the wide bandwidth of a WDM system, PMD will lead to a differential rotation of different frequency components. Hence, if the initial states of different channels are lined up in some way, for example, if every other channel is initially co-polarized and polarized orthogonally to its neighbors [25], then PMD will eventually randomize the orientations. This randomization plays an important role in determining the polarization penalties, as I will show later in this dissertation.

2.1   Theory

In real optical fibers, both the orientation of the axes of birefringence and the magnitude of the birefringence change in a complex but continuous way that is not well known. All that is really known is that all possible orientations must occur with

equal probability because otherwise there would be permanent preferred axes and that the local birefringence must be nearly linear (as opposed to elliptical) because otherwise the strength of the cross-phase modulation would not equal two-thirds of the strength of the self-phase modulation [9]. Wai and Menyuk [9], [25], [26] have shown that with these minimal assumptions, very different models for the statistical behavior of a fiber yield nearly identical results for the evolution of the statistical quantities that characterize the electric field of the light as long as two key statistical parameters that characterize the fiber remain unchanged. These two parameters are the fiber correlation length, the length scale on which the fiber loses memory of the orientation of the axes of birefringence, and the fiber's average beat length. Wai and Menyuk [9], [25], [26] have shown that with these minimal assumptions, very different models for the statistical behavior of a fiber yield nearly identical results for the evolution of the statistical quantities that characterize the electric field. They showed that the polarization state of each frequency component of the light evolves randomly on the Poincaré sphere, ultimately covering it uniformly, but that the equatorial and azimuthal rates of diffusion are different in general, depending on the values of the fiber's correlation length and average beat length. Wai and Menyuk [9], [25], [26] also showed that the fiber's correlation length, the equatorial diffusion length, and the field correlation length that characterizes the usual linear PMD are all the same. The azimuthal diffusion length is irrelevant to the behavior of the usual linear PMD. This length does affect what Wai and Menyuk [9] referred to as "nonlinear PMD," but Marcuse, *et al.* [16] later showed that this effect is irrelevant in modern-day communication system. While my colleagues and I have shown that this effect becomes important in short-pulse devices like nonlinear loop mirrors [27], [8]. We will not

consider it any further in this dissertation.

Fiber correlation lengths vary from less than a meter in modern-day, low-PMD fiber to around 100 meters for high-PMD fiber. Average beat lengths vary in the range of 10–30 m. By contrast, typical nonlinear and dispersive length scales are typically hundreds or even thousands of kilometers [15]. This wide discrepancy of length scales is the reason that nonlinear PMD is irrelevant in modern-day communication systems; this effect only becomes important when the fiber correlation length and the nonlinear scale length become comparable At the same time, it is a severe challenge to the modeler who wishes to study PMD. In principle, to accurately calculate the effects of the randomly varying birefringence, one must resolve the random motion of the light's polarization states on the Poincaré sphere. To do so, one must take steps that are on the order of meters or less. By contrast, to resolve the dispersive and nonlinear evolution it would be sufficient to take steps that are on the order of kilometers or more. Thus, it would be more than a thousand times slower to run a vector computer code that includes polarization effects than it would be to run a scalar computer code that ignores these effects — an unacceptable computational cost! At the same time, the wide discrepancy in length scales suggests that it should be possible to average over the rapid variations on the Poincaré sphere when one is not interested in the details of this evolution, as is typically the case when modeling communications systems. Motivated by this consideration, Wai, *et al.* [9] and Evangelides, *et al.* [28] independently proposed variants of the coarse step method. In the variant that I shall use in this dissertation, originally proposed by Marcuse, *et al.* [16], one keeps the birefringence fixed during one computational step, whose size is determined by the dispersive and nonlinear length scales and is typically on the order of several

kilometers. Each computational step is followed by a random rotation on the Poincaré sphere. Marcuse, *et al.* [16] showed that this approach yields equivalent results for the PMD statistics as long as

$$(\Delta\beta) = (2h_{\text{fiber}}/\zeta)^{1/2}(\Delta\beta)_{\text{actual}}, \tag{2.1}$$

where, h_{fiber} is the fiber correlation length, ζ is the step size, $(\Delta\beta)_{\text{actual}}$ is the actual average fiber propagation constant difference due to the birefringence. I note that $\Delta\beta_{\text{actual}} = \omega\Delta n_{\text{eff}}/c$, where Δn_{eff} is the birefringence difference. The quantity $\Delta\beta$ defined by Eq. 2.1 is the propagation constant difference that I use in my simulations. The reduction in the magnitude of $\Delta\beta$ relative to its actual value compensates for the orientation of the birefringence being held constant for a length that equals the step size rather than a length on the order of h_{fiber}. The coarse step method can be viewed as a specific application of multiple length scale techniques [28].

Rather than directly measuring the PMD, most experiments measure the differential group delay (DGD) as a function of distance. The PMD is defined as $\langle\Delta\tau\rangle/Z^{1/2}$, where $\Delta\tau$ is the DGD, Z is the propagation distance, and the brackets indicate an ensemble of measurements. Theory indicates that the measured DGDs are Maxwellian-distributed if the ensemble is sufficiently large [31], with a mean that is given by

$$\left(\frac{64h_{\text{fiber}}Z}{3\pi}\right)^{1/2}\Delta\beta'_{\text{actual}}, \tag{2.2}$$

where $\Delta\beta' = d\Delta\beta/d\omega \simeq \Delta n_{\text{eff}}/c$. Substituting Eq. 2.1 into Eq. 2.2, I find that Eq. 2.2 becomes

$$\left(\frac{32\zeta Z}{3\pi}\right)^{1/2}\Delta\beta', \tag{2.3}$$

when using the coarse step method. I note that $Z = N\zeta$, where N is the number of

steps in the propagation. From the measured PMD or average DGD, one can infer an appropriate value for $\Delta\beta'$ to be used in the simulations.

Given a value of $\Delta\beta'$ that is inferred from Eq. 2.3, one may proceed to calculate the evolution of the Jones vector $\mathbf{A}(\omega)$ as a function of distance z at each frequency. The Jones vector $\mathbf{A}$ is a complex 2-vector in which each component corresponds to the complex field for one of the polarizations [29]. It is related to the Stokes vector $\mathbf{S} = (S_1, S_2, S_3)$ by the relationships

$$S_1 = \mathbf{A}^\dagger \sigma_3 \mathbf{A}, \qquad S_2 = \mathbf{A}^\dagger \sigma_1 \mathbf{A}, \qquad S_3 = -\mathbf{A}^\dagger \sigma_2 \mathbf{A}, \qquad (2.4)$$

where

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \qquad (2.5)$$

Dividing the Stokes vector by its magnitude, we obtain the corresponding point on the Poincaré sphere $\mathbf{s} = (s_1, s_2, s_3)$. In the coarse step method, one finds explicitly that [16]

$$\mathbf{A}(\omega, z = Z) = \mathsf{R}_N \mathsf{R}_\omega \mathsf{R}_{N-1} \mathsf{R}_\omega \cdots \mathsf{R}_2 \mathsf{R}_\omega \mathsf{R}_1 \mathsf{R}_\omega \mathbf{A}(\omega, z = 0). \qquad (2.6)$$

The matrices R_j, $j = 1 \ldots N$, are defined as

$$\mathsf{R}_j = \begin{pmatrix} \cos(\theta_j/2) \exp[i(\phi_j + \psi_j)/2] & \sin(\theta_j/2) \exp[i(\phi_j - \psi_j)/2] \\ -\sin(\theta_j/2) \exp[-i(\phi_j - \psi_j)/2] & \cos(\theta_j/2) \exp[-i(\phi_j + \psi_j)/2] \end{pmatrix}, \qquad (2.7)$$

where the ϕ_j and ψ_j are random variables chosen independent at each j from uniform distributions in the range $[0, 2\pi]$ and θ_j is a random variable chosen independently at each j such that $\cos(\theta_j/2)$ is uniformly distributed in the range $[-1, 1]$. This choice corresponds to a random rotation with a uniform probability distribution on the Poincaré sphere. The matrix R_ω is defined as

$$\mathsf{R}_\omega = \begin{pmatrix} \exp(\Delta\beta'\omega\zeta/2) & 0 \\ 0 & \cos(\Delta\beta'\omega\zeta/2) \end{pmatrix}, \qquad (2.8)$$

	$\zeta = 1$ km	$\zeta = 3$ km	$\zeta = 5$ km	$\zeta = 9$ km	$\zeta = 15$ km
mean level crossing (ps)	24.90	24.99	25.02	25.05	25.18
extrema point (ps)	25.49	25.55	25.61	25.59	25.76

Table 2.1: Measured DGD using the coarse step method

Using the method of measuring PMD in the fiber proposed by Poole [30], I can show numerically that the coarse step method is a very accurate method for modeling PMD in optical fibers. In order to illustrate numerically the validity of the coarse step method for modeling PMD in optical fibers, I consider an example in which I set the expected DGD $\langle \Delta\tau \rangle = 25$ ps over a length of 9000 km, corresponding to a PMD of 0.25 ps/km$^{1/2}$. I also set $h_{\text{fiber}} = 50$ m for each of 1000 fiber realizations at several different step sizes. I then calculated the DGD using both the mean level crossing method and the extrema point method described by Favin and Poole [30]. I show the average results in Table 2.1. The agreement is good even when ζ is as large as 15 km.

2.2 Stokes model of PMD

Since I am interested in the evolution of the polarization of an entire communication channel, I will be focusing on the evolution of the averaged Stokes parameters for each separate channel m in a WDM system. I first define $\mathbf{U}(z,t)$ as the Jones vector in the time domain. It is the inverse Fourier transform of $\mathbf{A}(z,\omega)$ defined in Sec. 2.1. I next write $\mathbf{U}$ as a sum of contributions over n channels, obtaining

$$\mathbf{U} = \sum_{m=1}^{n} \mathbf{U}^{(m)} \exp[ik^{(m)}z - i\omega^{(m)}t], \qquad (2.9)$$

where $k^{(m)}$ and $\omega^{(m)}$ are the central wavenumber and frequency of the m-th channel with respect to the central wavenumber and frequency of $\mathbf{U}$, and $\mathbf{U}^{(m)}$ is the corresponding wave envelope. The definition of the Stokes parameters for each channel is

$$
\begin{aligned}
S_0^{(m)} &= \frac{1}{T} \int_{t_1}^{t_2} \left[\left| u_x^{(m)}(t) \right|^2 + \left| u_y^{(m)} \right|^2 \right] dt, \\
S_1^{(m)} &= \frac{1}{T} \int_{t_1}^{t_2} \left[\left| u_x^{(m)}(t) \right|^2 - \left| u_y^{(m)} \right|^2 \right] dt, \\
S_2^{(m)} &= \frac{2}{T} \int_{t_1}^{t_2} \mathrm{Re} \left[u_x^{(m)}(t) u_y^{(m)*}(t) \right] dt, \\
S_3^{(m)} &= \frac{2}{T} \int_{t_1}^{t_2} \mathrm{Im} \left[u_x^{(m)}(t) u_y^{(m)*}(t) \right] dt, \quad\quad (2.10)
\end{aligned}
$$

where $T = t_2 - t_1$, while $u_x^{(m)}(t)$ and $u_y^{(m)}(t)$ are the wave envelopes in two orthogonal polarizations. I am assuming that T is very large compared to a single bit period and that the channel becomes statistically stationary when T is large so that this definition is meaningful. In general, the relationship between the Stokes parameters defined here and the Stokes vector defined in Eq. 2.4 is somewhat complex. However, we will treat the final three Stokes parameters $S_1^{(m)}$– $S_3^{(m)}$ just like a Stokes vector at single frequency $\omega = \omega^{(m)}$, the central frequency of channel m. Since I will only be following one set of Stokes parameters for the signal in each channel in the WDM system, I must convert from the Jones representation to the Mueller representation so that I can deal with partially polarized channels. PMD in the fiber will cause the polarization states of different channels to evolve differently. However the polarization states of each single channel have a uniform evolution—in this model the evolution of the polarization state of the central frequency of the channel. Hence, I ignore intra-channel PMD. This approximation is reasonable as long as the accumulated DGD is not large compared to the pulse width. Since there is no PDL in the optical

fiber, in contrast to the amplifiers where I will take into account the PDL separately, and since the polarization-independent loss must be exactly compensated by the gain over the length of the transmission line, so that I may ignore the spatially varying gain and loss, I find $S_0^{(m)}(z + \zeta) = S_0^{(m)}(z)$. Using Eqs. (2.4)–(2.8), I find that the Stokes vector portion of the Stokes parameters $\mathbf{S}^{(m)} = (S_1^{(m)}, S_2^{(m)}, S_3^{(m)})^t$ transforms according to the relationship

$$\mathbf{S}^{(m)}(z + \zeta) = \mathsf{M}_R^{(m)}\mathsf{M}_j(z)\mathbf{S}^{(m)}(z), \tag{2.11}$$

where

$$\mathsf{M}_R^{(m)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\Delta\beta'\omega^{(m)}\zeta) & -\sin(\Delta\beta'\omega^{(m)}\zeta) \\ 0 & \sin(\Delta\beta'\omega^{(m)}\zeta) & \cos(\Delta\beta'\omega^{(m)}\zeta) \end{pmatrix}, \tag{2.12a}$$

$$\mathsf{M}_j =$$
$$\begin{pmatrix} \cos\theta_j & \sin\theta_j\cos\psi_j & -\sin\theta_j\sin\psi_j \\ -\sin\theta_j\cos\phi_j & \cos\theta_j\cos\phi_j\cos\psi_j - \sin\phi_j\sin\psi_j & -\cos\theta_j\cos\phi_j\sin\psi_j - \sin\phi_j\cos\psi_j \\ -\sin\theta_j\sin\phi_j & \cos\theta_j\sin\phi_j\cos\psi_j + \cos\phi_j\sin\psi_j & -\cos\theta_j\sin\phi_j\sin\psi_j + \cos\phi_j\cos\psi_j \end{pmatrix}.$$
$$\tag{2.12b}$$

I stress that θ_j, ϕ_j, ψ_j are the same for all channels.

Chapter 3

Polarization dependent loss (PDL)

PDL is due to the polarization dependence of the transmission in some devices. It is caused by physical effects such as bulk dichroism (polarization-dependent absorption), fiber bending, angled optical interfaces, and oblique reflection [10]. It can be measured as follows: One starts with a polarized light source whose degree of polarization is 1. Fixing the power of the input light to the device under test from this source, one varies the polarization state of the input light over the entire Poincaré sphere, and measures the output power. From the maximum $P_{\max}$ and minimum powers $P_{\min}$, one obtains the PDL as

$$\text{PDL} = 10 \log_{10} \frac{P_{\max}}{P_{\min}}. \tag{3.1}$$

In Table 3.1, I show the PDL of some typical devices that are widely used in optical fiber communications.

In long distance transmission systems, the PDL will accumulate along the transmission line; so, the requirement for the PDL in each individual device is very strict; it must be below 0.1 dB. PDL can cause fluctuations of the signal-to-noise ratio. If the signal is polarization scrambled, then the fluctuations will be reduced, but they can-

Component	PDL (dB)	Details
Optical connector	0.02	straight type
	0.06	angled type
3 dB coupler	0.1–0.2	Single window
	0.15–0.3	1300/1500 type
10 dB coupler	0.02	through path
	0.1	−10 dB path
Isolator	0.05–0.3	
Circulator	0.1–0.2	
DWDM multiplexer	0.05–0.1	

Table 3.1: PDL of some typical devices

not be completely eliminated. My colleagues and I have studied this problem, and we found that a completely depolarized signal may repolarize when there is PDL in the transmission line, which can seed further degradations due to polarization dependent gain [11].

3.1 Repolarization of a single channel system

Here, I will consider the simple case of a single channel in which polarization dependent loss elements located at amplifiers alternate with lengths of optical fiber, as shown in Fig. 3.1. These results have been published in ref. [11]. I will derive an analytical formula for the probability distribution function of the degree of polarization and compare the formula to Monte Carlo simulations. I will show that if

each amplifier has a polarization dependent loss of $x_{\text{PDL}} = 0.1$ dB, which is a typical value for current systems, then a significant amount of repolarization will occur. I will then discuss the scaling of the repolarization with x_{PDL} and the implications for communications systems.
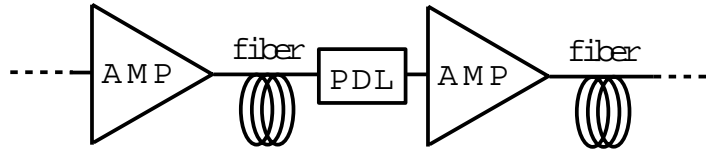
Figure 3.1: Model of PDL in the transmission line.

The effect of a polarization dependent loss element is to cause excess loss in one of two orthogonal polarizations. Using the Jones vector notation defined earlier, where I take the second component to be in the direction of maximum loss, I may write

$$
\begin{pmatrix} u_x(t) \\ u_y(t) \end{pmatrix}_{\text{after}} = \begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} u_x(t) \\ u_y(t) \end{pmatrix}_{\text{before}}
\tag{3.2}
$$

where α is related to x_{PDL} through the relationship, x_{PDL} (in dB) $= -20\log_{10}\alpha$. From Eq. (2.10) and Eq. (3.2) I find

$$
\begin{aligned}
S_{0,\text{after}} &= \frac{1+\alpha^2}{2}S_{0,\text{before}} + \frac{1-\alpha^2}{2}S_{1,\text{before}}, \\
S_{1,\text{after}} &= \frac{1-\alpha^2}{2}S_{0,\text{before}} + \frac{1+\alpha^2}{2}S_{1,\text{before}}, \\
(S_2 + iS_3)_{\text{after}} &= \alpha(S_2 + iS_3)_{\text{before}},
\end{aligned}
\tag{3.3}
$$

where I recall that the Stokes parameters are averaged over time. The effect of the fiber between the polarization dependent loss elements is to randomly rotate the Stokes vector $S_{\text{pol}} = (S_1, S_2, S_3)$ on the Poincaré sphere. I thus obtain the following

iterative equations,

$$
\begin{aligned}
S_{0,\text{after}} &= \frac{1+\alpha^2}{2} S_{0,\text{before}} + \frac{1-\alpha^2}{2} S_{\text{pol,before}} \cos\theta, \\
S_{1,\text{after}} &= \frac{1-\alpha^2}{2} S_{0,\text{before}} + \frac{1+\alpha^2}{2} S_{\text{pol,before}} \cos\theta, \\
(S_2 + iS_3)_{\text{after}} &= \alpha S_{\text{pol,before}} \sin\theta \exp(i\phi).
\end{aligned}
\tag{3.4}
$$

where

$$
(S_{\text{pol,before}})^2 = (S_{1,\text{before}})^2 + (S_{2,\text{before}})^2 + (S_{3,\text{before}})^2.
\tag{3.5}
$$

The $\cos\theta$ are independent and identically distributed (i.i.d.) random variables, uniformly distributed in the range $[-1, 1]$ so that $\langle \cos^2\theta \rangle = 1/3$. Similarly, the ϕ are i.i.d. random variables that are also independent of the $\cos\theta$ and are uniformly distributed in the range $[0, 2\pi]$. Equations (3.4) and (3.5) together constitute a random process, and our goal is to solve them simultaneously in order to determine $f(d_{\text{pol}})$, the probability distribution function for $d_{\text{pol}} = S_{\text{pol}}/S_0$.

Equations (3.4) and (3.5) do not take into account polarization independent loss and gain. In reality, the amplifiers in optical fiber telecommunications systems are set to operate in saturation so that the total power S_0 is maintained as close to constant as possible [12]. Since d_{pol} is not affected by the polarization independent loss and gain, we do not include it, but the reader should be aware that S_{pol} and S_0 are not separately meaningful. Only the ratio $S_{\text{pol}}/S_0 = d_{\text{pol}}$ is meaningful. Since d_{pol} is not affected by the total power level, I will simply set $S_{0,\text{input}} = 1$. I also set $S_{\text{pol,input}} = 0$, which is equivalent to assuming $d_{\text{pol,input}} = 0$, so that the channel is initially polarization scrambled.

To make further theoretical progress, I next eliminate $\cos\theta$ and ϕ from Eq. (3.4)

and (3.5), obtaining

$$(S_{0,n+1})^2 - (S_{\mathrm{pol},n+1})^2 = \alpha^2[(S_{0,n})^2 - (S_{\mathrm{pol},n})^2] = \alpha^{2(n+1)}, \tag{3.6}$$

where n refers to the n-th PDL element. This relation suggests replacing $S_{0,n}$ and $S_{\mathrm{pol},n}$ with the new variables $x^n = S_{0,n}/\alpha^n$, which now obey the equations

$$
\begin{aligned}
x_{n+1} &= \frac{1+\alpha^2}{2\alpha}x_n + \frac{1-\alpha^2}{2\alpha}y_n\cos\theta, \\
x_n^2 - y_n^2 &= 1,
\end{aligned}
\tag{3.7}
$$

where $d_{\mathrm{pol},n} = y_n/x_n$. Using Eq. (3.7) yields

$$x_{n+1} = \frac{1+\alpha^2}{2\alpha}x_n + \frac{1-\alpha^2}{2\alpha}(x_n^2-1)^{1/2}\cos\theta. \tag{3.8}$$

Since α is close to one, the change in each step is small, and it is reasonable to replace Eq. (3.8) with the stochastic differential equation

$$\dot{x} = \beta x + (x^2-1)^{1/2}\xi, \tag{3.9}$$

where $\dot{x} = dx/dn$ is the first-order stochastic differential operator, ξ is a white noise process with variance $\sigma_\xi^2 = (1-\alpha^2)^2/12\alpha^2$, and $\beta = (1-\alpha)^2/2\alpha$. Since Eq. (3.8) is a forward difference equation, Eq. (3.9) must be interpreted in the sense of Ito, which implies that Eq. (3.9) represents a diffusion process, and the evolution of the distribution of x, $f(x)$, is governed by the Fokker-Planck equation [35]

$$\frac{\partial f_x}{\partial n} + \beta\frac{\partial}{\partial x}x f_x - \frac{\sigma_\xi^2}{2}\frac{\partial^2}{\partial x^2}(x^2-1)f_x = 0. \tag{3.10}$$

Changing variables from x to γ, where $x = \cosh\gamma$, one finds that $f_\gamma = f_x[x(\gamma)]dx/d\gamma = f_x[x(\gamma)]/\sinh\gamma$ is governed by the Fokker-Planck equation

$$\frac{\partial f_\gamma}{\partial n} + (\beta - \frac{1}{2}\sigma_\xi^2)\coth(\gamma)f_\gamma - \frac{1}{2}\sigma_\xi^2\frac{\partial^2 f_\gamma}{\partial\gamma^2} = 0, \tag{3.11}$$

where $d_{\text{pol}} = \tanh \gamma$. In the limit of interest to us in which $\gamma \leq 1$, one may replace $\coth \gamma \simeq \gamma^{-1}$, in which case Eq. (3.11) has the solution

$$
\begin{aligned}
f_\gamma(\gamma) &= \frac{2}{(2\sigma_\xi^2 n)^{1/2} \Gamma(\beta/\sigma_\xi^2)} \left(\frac{\gamma^2}{2\sigma_\xi^2 n}\right)^{(\beta/\sigma_\xi^2)-(1/2)} \exp\left(-\frac{\gamma^2}{2\sigma_\xi^2 n}\right) \\
&\simeq \frac{4}{\pi^{1/2}} \frac{\gamma^2}{(2\sigma_\xi^2 n)^{3/2}} \exp\left(-\frac{\gamma^2}{2\sigma_\xi^2 n}\right)
\end{aligned}
\tag{3.12}
$$

with $\beta/\sigma_\xi^2 = 6\alpha/(1+\alpha)^2 \simeq 3/2$. From Eq. (3.12), one may obtain $f(d_{\text{pol}})$ using the relationship

$$
f(d_{\text{pol}}) = f_\gamma(\gamma) d\gamma/d(d_{\text{pol}}).
\tag{3.13}
$$

Figure 3.2 compares the analytical expression for $f(d_{\text{pol}})$ from Eqs. (3.12) and (3.13) to a Monte Carlo solution of Eqs. (3.4) and (3.5), setting $x_{\text{PDL}} = 0.1$. I used 10^6 representations of the $\cos \theta$. Two points are apparent. The first is that the theory is indistinguishable from the simulation. The second is that when $n = 300$, corresponding to trans-oceanic distances, a significant amount of repolarization occurs.
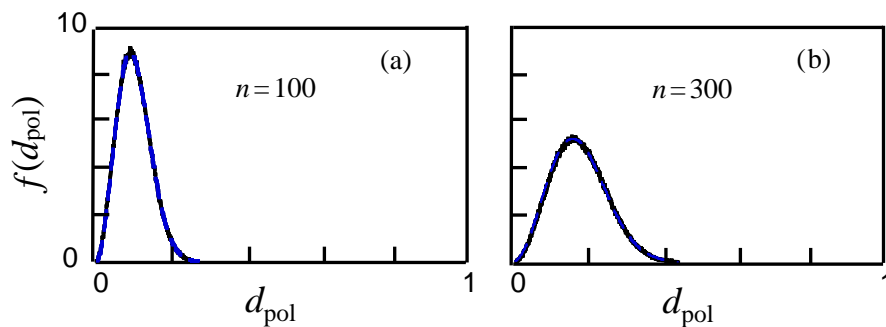
Figure 3.2: Comparison of the distribution function $f(d_{\text{pol}})$ obtained by Monte Carlo simulation of the original difference equations to the theoretically calculated function. The parameters are $x_{\text{PDL}}=0.1$ dB. (a) $n = 100$. (b) $n = 300$. The two approaches yield indistinguishable results.

A thorough discussion of the impact that the repolarization has on long-distance,

high-data rate systems must include the effects of polarization dependent gain and polarization mode dispersion. I do that later in this dissertation. Nonetheless, it is possible to immediately draw from this simple calculation a number of results that are useful for system design. First, polarization dependent loss can lead to a significant repolarization of a polarization scrambled signal. This repolarization can then seed further signal degradation due to polarization dependent gain and polarization mode dispersion; so, there is clear motivation to reduce x_{PDL} as much as possible. From Eq. (3.12), using $d_{\mathrm{pol}} \simeq \gamma$, I find $d_{\mathrm{pol,max}} \simeq 0.09 n^{1/2} x_{\mathrm{PDL}}$ and $\langle d_{\mathrm{pol}} \rangle = 0.11 n^{1/2} x_{\mathrm{PDL}}$, where $d_{\mathrm{pol,max}}$ is the mean of $f(d_{\mathrm{pol}})$. For the example presented in Fig. 3.2, in which $n = 300$ and $x_{\mathrm{PDL}} = 0.1$, I find that $d_{\mathrm{pol,max}} = 0.16$ and $\langle d_{\mathrm{pol}} \rangle = 0.19$.

How low must one make x_{PDL} before it can be neglected entirely? I may estimate this value very roughly as follows: After polarization scrambling, the residual d_{pol} is about 15%; so, it is reasonable to assume that if the additional repolarization is under 15%, then it can be neglected. If I assume that a fiber cable has a lifetime of 20 years and, somewhat conservatively, that it changes its polarization state every hour so that a different value of d_{pol} is sampled, then the cable passes through 2×10^5 states in its lifetime. Thus, if the probability of obtaining a repolarization of 0.15 is less than 5×10^{-6}, the repolarization can be ignored. Integrating Eq. (3.12), I find that $d_{\mathrm{pol}} \geq 35 n^{1/2} x_{\mathrm{PDL}}$ with probability 5×10^{-6}; so with $n = 300$, our criterion becomes $x_{\mathrm{PDL}} \leq 0.025$ which is about a half of the current best value.

3.2   Stokes parameter model for multichannel systems

To describe multi-channel systems, I proceed in close analogy to the single-channel calculation in Sec. 3.1. However, the fiber changes the polarization states of the different channels by slightly different amounts due to PMD. Thus, we may write the change that occurs in the Stokes parameters at each amplifier due to PDL as

$$
\begin{aligned}
S_{0,\text{after}}^{(m)} &= \frac{1+\alpha^2}{2} S_{0,\text{before}}^{(m)} + \frac{1-\alpha^2}{2} S_{\text{pol,before}}^{(m)} \cos\theta^{(m)}, \\
S_{1,\text{after}}^{(m)} &= \frac{1-\alpha^2}{2} S_{0,\text{before}}^{(m)} + \frac{1+\alpha^2}{2} S_{\text{pol,before}}^{(m)} \cos\theta^{(m)}, \\
(S_2 + iS_3)_{\text{after}}^{(m)} &= \alpha S_{\text{pol,before}}^{(m)} \sin\theta^{(m)} \exp(i\phi^{(m)}).
\end{aligned}
\tag{3.14}
$$

where m indicates the channel number. Between amplifiers, we account for the PMD as described in Eq. (2.11). I note that several iterations of the coarse step method are required between amplifiers. Typically, I used a 1 km spacing. I experimented with spacings up to 3 km, and I found identical results.

Chapter 4

Polarization dependent gain (PDG)

Currently in erbium doped fiber amplifiers (EDFAs), there are two kinds of PDG. The first kind of PDG is induced by laser diode pumping [13]. If the pump light is polarized, then the laser medium will be selectively excited. So in some polarization states, the incoming signal will experience a bigger gain, and in other polarization states the gain will be smaller. The maximum gain difference is about 0.12 dB. The polarization of this type of PDG is always fixed as long as the polarization of the pump light is fixed; so, its effect is similar to PDL. This kind of PDG has been nearly eliminated in practical systems by pumping in two orthogonal polarizations, and I will not consider it further.

The second kind of PDG is due to polarization hole burning induced by the incoming signal. The gain in the polarization orthogonal to the incoming signal is larger than the gain in the polarization of the incoming signal. EDFAs have a response time of several milliseconds. Changes in the incoming polarization states that occur on a shorter time scale, due for example to bit-synchronous polarization scrambling, do not affect the orientation of the PDG. However, changes that occur in the incoming

polarization states on a longer time scale, due for example to the temperature-induced changes in the fiber length, do affect the orientation of the PDG. Since slow changes of fiber lengths are constantly occurring in real systems, the orientation of this type of PDG is constantly changing. The amount of PDG in a single amplifier is only about 0.07 dB for an EDFA with 3 dB of gain compression and becomes larger as the amplifier goes deeper into gain compression. The magnitude of the polarization hole burning is proportional to the degree of polarization, d_{pol}, of the incoming signal. (To determine d_{pol}, one must average the Stokes parameters on a time scale that is long compared to the bit period and short compared to the response time of the EDFA.)

Polarization hole burning can lead to a significant degradation of the signal-to-noise ratio [1]. To mitigate this effect, an effective approach is to polarization scramble the incoming signal [3]. As I showed however in Sec. 3.1, PDL can lead to repolarization of polarization-scrambled signals, which in turn can induce a degradation in the signal-to-noise ratio due to PDG. This effect is already known to be important in single-channel systems [1]. By contrast, I will show later in this dissertation that PDG does not have a large effect on the signal-to-noise ratio in multi-channel systems with more than about 10 channels. Physically, PMD leads to a large amount of randomization of the overall polarization state to which the EDFA responds, reducing d_{pol}. A more significant effect is the random variation of the relative powers in the different channels due to the combination of PDL and PMD.

4.1 Theory and model

Finally, I turn to a model for PDG and noise growth in the amplifiers. I will model the PDG much like the PDL, except that the direction of maximum gain must be chosen self-consistently with the existing signal in a given system. Thus, if I ignore the noise contribution for the moment, returning to it later, and I write

$$S_0^{\text{total}} = \sum_m S_0^{(m)}, \qquad \mathbf{S}^{(\text{total})} = \sum_m \mathbf{S}^{(m)}, \tag{4.1}$$

where the sum indicates an addition over all the channels, I find the total degree of polarization $d_{\text{pol}} = |\mathbf{S}^{(\text{total})}|/S_0^{(\text{total})}$ and the total state of polarization, $\mathbf{s} = \mathbf{S}^{(\text{total})}/|\mathbf{S}^{(\text{total})}|$. I now write

$$\begin{pmatrix} u_x(t)^{(m)} \\ u_y(t)^{(m)} \end{pmatrix}_{\text{after}} = \mathsf{R} \begin{pmatrix} 1 & 0 \\ 0 & g^{1/2} \end{pmatrix} \mathsf{R}^{-1} \begin{pmatrix} u_x(t)^{(m)} \\ u_y(t)^{(m)} \end{pmatrix}_{\text{before}}, \tag{4.2}$$

where g is the polarization dependent gain, normalized to the gain in the polarization state of the input signal. The value of g may be written [1]

$$g = 10^{\frac{PDG * d_{\text{pol}}}{10}}. \tag{4.3}$$

The rotation matrix R is determined by the overall polarization state of the signal and noise since it is this polarization state that determines the orientation of the PDG, while R^{-1} is the inverse of R. It will be useful later to define angles φ, ψ, and ϕ that are related to R via the relationship

$$\mathsf{R} = \begin{pmatrix} \cos(\varphi/2)\exp[-i(\psi+\phi)/2] & -\sin(\varphi/2)\exp[i(\psi-\phi)/2] \\ \sin(\varphi/2)\exp[-i(\psi-\phi)/2] & \cos(\varphi/2)\exp[i(\psi+\phi)/2] \end{pmatrix}. \tag{4.4}$$

I find that the elements of R are related to **s** through the relationship

$$s_1 = |r_{11}|^2 - |r_{12}|^2 = \cos\varphi, \qquad s_2 + is_3 = 2r_{11}r_{12}^* = \sin\varphi \exp(-i\phi). \qquad (4.5)$$

I note that it is not possible to determine ψ from **s**. However, the angle ψ does not affect the evolution of the Stokes parameters of the individual channels in any way and so can be safely ignored.

I once again transform from the Jones representation to the Stokes representation using Eq. (2.10) to do the appropriate time average for each channel. Defining

$$\alpha = \frac{g^{1/2} + 1}{2}, \qquad\qquad \varepsilon = \frac{g^{1/2} - 1}{2}, \qquad (4.6)$$

I then obtain

$$
\begin{aligned}
S_{0,\text{after}}^{(m)} &= S_{0,\text{before}}^{(m)}\frac{g+1}{2} - \mathbf{s} \cdot \mathbf{S}_{\text{pol,before}}^{(m)}, \\[2mm]
S_{1,\text{after}}^{(m)} &= -2\alpha\varepsilon s_1 S_{0,\text{before}}^{(m)} + 2\varepsilon^2 s_1 s_2 S_{2,\text{before}}^{(m)} + 2\varepsilon^2 s_1 s_3 S_{3,\text{before}}^{(m)} \\[1mm]
&\quad +(\alpha^2 + \varepsilon^2)S_{1,\text{before}}^{(m)} - 2\varepsilon^2 s_2^2 S_{1,\text{before}}^{(m)} - 2\varepsilon^2 s_3^2 S_{1,\text{before}}^{(m)}, \\[2mm]
S_{2,\text{after}}^{(m)} &= -2\alpha\varepsilon s_2 S_{0,\text{before}}^{(m)} + 2\varepsilon^2 s_1 s_2 S_{1,\text{before}}^{(m)} + 2\varepsilon^2 s_2 s_3 S_{3,\text{before}}^{(m)} \\[1mm]
&\quad +(\alpha^2 + \varepsilon^2)S_{2,\text{before}}^{(m)} - 2\varepsilon^2 s_1^2 S_{2,\text{before}}^{(m)} - 2\varepsilon^2 s_3^2 S_{2,\text{before}}^{(m)}, \\[2mm]
S_{3,\text{after}}^{(m)} &= -2\alpha\varepsilon s_3 S_{0,\text{before}}^{(m)} + 2\varepsilon^2 s_1 s_3 S_{1,\text{before}}^{(m)} + 2\varepsilon^2 s_2 s_3 S_{2,\text{before}}^{(m)} \\[1mm]
&\quad +(\alpha^2 + \varepsilon^2)S_{3,\text{before}}^{(m)} - 2\varepsilon^2 s_1^2 S_{3,\text{before}}^{(m)} - 2\varepsilon^2 s_2^2 S_{3,\text{before}}^{(m)}, \qquad (4.7)
\end{aligned}
$$

where I recall that $\mathbf{S}_{\text{pol}}^{(m)} = (S_1^{(m)}, S_2^{(m)}, S_3^{(m)})$. After rearrangement, I obtain

$$
\begin{aligned}
\mathbf{S}_{\text{pol,after}}^{(m)} &= -\frac{g-1}{2}S_{0,\text{before}}^{(m)} \cdot \mathbf{s} + \frac{g+1}{2}\mathbf{S}_{\text{pol,before}}^{(m)} + \frac{(g^{1/2}-1)^2}{2}\mathbf{s} \times (\mathbf{s} \times \mathbf{S}_{\text{pol,before}}^{(m)}) \\[2mm]
S_{0,\text{after}}^{(m)} &= \frac{g+1}{2}S_{0,\text{before}}^{(m)} - \frac{g-1}{2}\mathbf{s} \cdot \mathbf{S}_{\text{pol,before}}^{(m)}. \qquad (4.8)
\end{aligned}
$$

4.2   Combining PDG with PMD, PDL, and ASE noise

We account for the ASE noise by following four noise Stokes parameters ($S_{0,\text{noise}}^{(m)}$, $\mathbf{S}_{\text{noise}}^{(m)}$) at each m. We must track these Stokes parameters separately from the signal Stokes parameters because they are random variables while the signal Stokes parameters are deterministic. Since the ASE noise is unpolarized, each amplifier will cause the following change in the Stokes parameters

$$
\begin{aligned}
S_{0,\text{noise,after}}^{(m)} &= S_{0,\text{noise,before}}^{(m)} + 2n_{\text{sp}}(G-1)B^{(m)}h\nu, \\
\mathbf{S}_{\text{noise,after}}^{(m)} &= \mathbf{S}_{\text{noise,before}}^{(m)},
\end{aligned}
\tag{4.9}
$$

where n_{sp} is the spontaneous emission factor, G is the amplifier gain, $h\nu$ is the energy of a single photon, and $B^{(m)}$ is the optical bandwidth of the m-th channel. These Stokes parameters are affected by the PMD, PDL, and PDG in exactly the same way as the signal Stokes parameters and participate in determining the degree of polarization and total Stokes parameters. Additionally, if there is any part of the gain bandwidth of the Er-doped fiber amplifier that is not included in one of the optical channels, then this noise energy will participate in the total energy balance. We may write for this additional portion

$$
S_{0,\text{noise,after}}^{(\text{add})} = S_{0,\text{noise,before}}^{(\text{add})} + 2n_{\text{sp}}(G-1)B^{(\text{add})}h\nu,
\tag{4.10}
$$

and we will assume that this contribution is unpolarized. We now write

$$
\begin{aligned}
S_0^{(\text{total})} &= \sum_m S_0^{(m)} + \sum_m S_{0,\text{noise}}^{(m)} + \sum_m S_{0,\text{noise}}^{(\text{add})}, \\
\mathbf{S}^{(\text{total})} &= \sum_m \mathbf{S}^{(m)} + \sum_m \mathbf{S}_{\text{noise}}^{(m)}.
\end{aligned}
\tag{4.11}
$$

The degree of polarization may now be written $d_{\mathrm{pol}} = |\mathbf{S}^{(\mathrm{total})}|/S_0^{(\mathrm{total})}$. The final step in the procedure is to take into account the effect of gain saturation by assuming that the total power at the output of the amplifier is fixed at a value $\mathcal{S}$. We then renormalize $S_0^{(m)}, \mathbf{S}^{(m)}, S_{0,\mathrm{noise}}^{(m)}, \mathbf{S}_{\mathrm{noise}}^{(m)}, \mathbf{S}_0^{(\mathrm{add})}$ by the factor $\mathcal{S}/S_0^{(\mathrm{total})}$ which takes into account the renormalization of the total power that occurs in real systems due to gain saturation in the amplifiers. I summarize the complete procedure schematically in Fig. 4.1. This procedure is repeated iteratively from amplifier to amplifier.
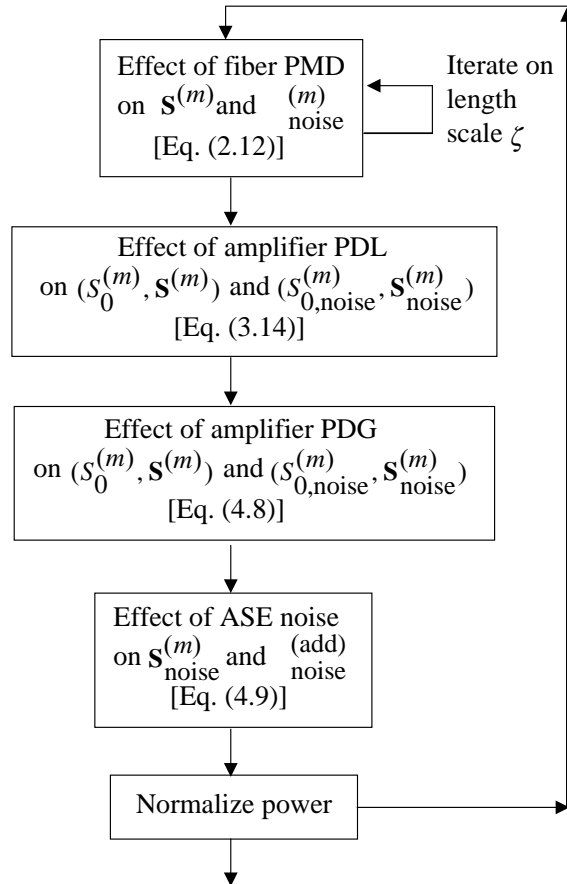
Figure 4.1: Structure of the model.

From the calculated signal and noise Stokes parameters, it is possible to determine $Q^{(m)}$—the so-called Q factor—for each channel m and from that to infer the penalty due to PDL and PDG. To calculate this penalty, I first note that the $Q^{(m)}$ that I calculate from this model is not meaningful by itself because this reduced model does not take into account the degradation due to nonlinearity and dispersion. What is meaningful is the *difference* between the $Q^{(m)}$ values that I calculate when PDL and PDG are present and when they are absent for a fixed value of PMD. The assumption in this approach is that the degradation due to nonlinearity and dispersion is not changed by increasing the PDL and PDG. This assumption is physically reasonable because PDL and PDG are gross effects that are insensitive to the time variations of the signal, while nonlinearity and dispersion affect the individual bits. We will validate this assumption explicitly in Chaps. 5 and 6.

To calculate $Q^{(m)}$ for a particular choice of PMD, PDL, and PDG, I first obtain the signal-to-noise (SNR) ratio of channel m which is equal to

$$\text{SNR}^{(m)} = \frac{P_{\text{peak}}}{P_{\text{ave}}} \frac{S_{0,\text{signal}}^{(m)}}{S_{0,\text{noise}}^{(m)}}. \tag{4.12}$$

where the ratio $P_{\text{peak}}/P_{\text{ave}}$ is the ratio of the peak power in a mark to the average power in the signal channel. For the standard NRZ format, this ratio is 2; for the standard RZ format, this ratio is 4; and, for the CRZ format of Bergano, *et al.* [33], this ratio is approximately 5.3. In principle, these ratios are lowered somewhat by the electrical filtering in the receiver; however, we have found that our results are insensitive to these corrections.I may then use a formula relating the Q-factor to the signal-to-noise ratio assuming that the noise is Gaussian-distributed [15]

$$Q^{(m)} = \frac{\text{SNR}^{(m)}}{\sqrt{2\text{SNR}^{(m)} + 1} + 1} \sqrt{\frac{2B_{\text{opt}}}{B_{\text{elec}}}}, \tag{4.13}$$

where B_{opt} is the optical bandwidth and B_{elec} is the electrical bandwidth. I note that this expression includes contributions from both the spontaneous-spontaneous beat noise and the signal-spontaneous beat noise. Physically, there are $2B_{\text{opt}}/B_{\text{elec}}$ noise modes that the electrical detector at the end of transmission line receives. Therefore, the signal-spontaneous beat noise $S_{0,\text{sig}-\text{spon}}$ is given by $(2S_0 S_{0,\text{noise}})^{1/2}(B_{\text{elec}}/2B_{\text{opt}})^{1/2}$, while the spontaneous-spontaneous beat noise $S_{0,\text{spon}-\text{spon}}$ just equals $S_{0,\text{noise}}$. Noting that the noise power in the marks is given by $S_{0,\text{sig}-\text{spon}} + S_{0,\text{spon}-\text{spon}}$ while the noise power in the spaces is just given by $S_{0,\text{spon}-\text{spon}}$, I obtain Eq. (4.13).

Chapter 5

Numerical Approach and Simulation Model

5.1 Full simulation

In order to validate the Stokes model of polarization effects, I built a full model using the coupled nonlinear Schrödinger equation [14]

$$
\begin{aligned}
-i\frac{\partial u_x}{\partial z} &= ik'\frac{\partial u_x}{\partial t} - \frac{1}{2}k''\frac{\partial^2 u_x}{\partial t^2} + \frac{1}{6}k'''\frac{\partial^3 u_x}{\partial t^3} + \gamma\left(|u_x|^2 + \frac{2}{3}|u_y|^2\right)u_x \\
&\quad + \frac{\gamma}{3}u_y^2 u_x^* \exp[-2i(k_0 - l_0)z] + i\frac{\alpha}{2}u_x, \qquad (5.1a) \\
-i\frac{\partial u_y}{\partial z} &= il'\frac{\partial u_y}{\partial t} - \frac{1}{2}l''\frac{\partial^2 u_y}{\partial t^2} + \frac{1}{6}l'''\frac{\partial^3 u_y}{\partial t^3} + \gamma\left(|u_y|^2 + \frac{2}{3}|u_x|^2\right)u_y \\
&\quad + \frac{\gamma}{3}u_x^2 u_y^* \exp[2i(k_0 - l_0)z] + i\frac{\alpha}{2}u_y, \qquad (5.1b)
\end{aligned}
$$

where $\mathbf{U}(z,t) = [u_x(z,t),\ u_y(z,t)]^t$ is the time-dependent Jones vector, t is the retarded time, k' and l' equal the inverse group velocities in the x and y polarizations respectively, k'' and l'' are the corresponding second order dispersion coefficients, k'''

and l''' are the corresponding third order dispersion coefficients, γ is the nonlinear coefficient, k_0 and l_0 are the propagation constants in the x and y polarizations, and α is the loss coefficient of the fiber.

I take the polarization dependent loss into account using Eq. (3.2) directly. I assume that the PDL in every amplifier is lumped, *i.e.*, it exists in one device such as an isolator inside the amplifier. I add this effect in the frequency domain just before I amplify the attenuated signal and add amplifier noise.

The PDG is determined by the degree of polarization and the state of polarization of the total input signal. To determine these quantities, I first recall that the Stokes parameters $[S_0(\omega),\ S_1(\omega),\ S_2(\omega),\ S_3(\omega)]$ are defined in terms of the Jones vector in the frequency domain $\mathbf{A}(\omega)$ as

$$S_0(\omega) = \mathbf{A}(\omega)^\dagger \mathbf{A}(\omega), \qquad S_1(\omega) = \mathbf{A}(\omega)^\dagger \sigma_3 \mathbf{A}(\omega),$$
$$S_2(\omega) = \mathbf{A}(\omega)^\dagger \sigma_1 \mathbf{A}(\omega), \qquad S_3(\omega) = -\mathbf{A}(\omega)^\dagger \sigma_2 \mathbf{A}(\omega). \tag{5.2}$$

Since these Stokes parameters are defined at a single frequency, they obey the relation $S_0^2(\omega) = S_1^2(\omega) + S_2^2(\omega) + S_3^2(\omega)$. I may now integrate over these frequency-dependent Stokes parameters to obtain $S_0^{(\text{total})}$ and $\mathbf{S}^{(\text{total})}$ by analogy to Eq. (4.1), and from them the total degree of polarization $d_{\text{pol}} = |\mathbf{S}^{(\text{total})}|/S_0^{(\text{total})}$ and the total state of polarization $\mathbf{s} = \mathbf{S}^{(\text{total})}/|\mathbf{S}^{(\text{total})}|$. I must now project the Jones vector

$$\mathbf{A}(\omega) = \begin{pmatrix} \widetilde{u}_x(\omega) \\ \widetilde{u}_y(\omega) \end{pmatrix}, \tag{5.3}$$

where $\widetilde{u}_x(\omega)$ and $\widetilde{u}_y(\omega)$ are the Fourier transforms of $u_x(t)$ and $u_y(t)$, onto the axes parallel and perpendicular to the total polarization state of the incoming signal. I

first write the unit vector in the parallel and perpendicular directions as

$$\widehat{\mathbf{e}}_{\parallel} = \begin{pmatrix} \cos(\varphi/2) \\ \sin(\varphi/2)\exp(i\phi) \end{pmatrix}, \qquad \widehat{\mathbf{e}}_{\perp} = \begin{pmatrix} \sin(\varphi/2) \\ -\cos(\varphi/2)\exp(i\phi) \end{pmatrix}, \qquad (5.4)$$

where φ and ϕ are related to the total polarization state as shown in Eq. (4.4). To project $\mathbf{A}(\omega)$ onto these unit vectors I write

$$\begin{aligned} \mathbf{A}(\omega) &= [\widetilde{u}_x(\omega)\cos(\varphi/2) + \widetilde{u}_y(\omega)\sin(\varphi/2)\exp(-i\phi)]\,\widehat{\mathbf{e}}_{\parallel} \\ &\quad + [\widetilde{u}_x(\omega)\sin(\varphi/2) - \widetilde{u}_y(\omega)\cos(\varphi/2)\exp(-i\phi)]\,\widehat{\mathbf{e}}_{\perp} \end{aligned} \qquad (5.5)$$

Writing $\mathbf{A}(\omega) = \widetilde{u}_{\parallel}(\omega)\widehat{\mathbf{e}}_{\parallel} + \widetilde{u}_{\perp}(\omega)\widehat{\mathbf{e}}_{\perp}$, I find explicitly that

$$\widetilde{u}_{\parallel}(\omega) = \widetilde{u}_x(\omega)\cos(\varphi/2) + \widetilde{u}_y(\omega)\sin(\varphi/2)\exp(-i\phi), \qquad (5.6\text{a})$$

$$\widetilde{u}_{\perp}(\omega) = \widetilde{u}_x(\omega)\sin(\varphi/2) - \widetilde{u}_y(\omega)\cos(\varphi/2)\exp(-i\phi). \qquad (5.6\text{b})$$

I can now calculate the effects of the polarization-independent amplifier gain G, the polarization-dependent gain g, and the noise on $\widetilde{u}_{\parallel}(\omega)$ and $\widetilde{u}_{\perp}(\omega)$. The amplified fields are given by

$$\widetilde{u}_{\parallel}^{a}(\omega) = G^{1/2}\widetilde{u}_{\parallel}(\omega) + r_{\parallel}, \qquad (5.7\text{a})$$

$$\widetilde{u}_{\perp}^{a}(\omega) = g^{1/2}[G^{1/2}\widetilde{u}_{\perp}(\omega) + r_{\perp}], \qquad (5.7\text{b})$$

where $u_{\parallel}^{a}$ and $u_{\perp}^{a}$ indicate the amplified fields, $u_{\parallel}$ and $u_{\perp}$ indicate the initial unamplified fields, and $r_{\parallel}$ and $r_{\perp}$ are the ASE noise contributions in the parallel and perpendicular directions. The average power in these noise contributions equals $n_{\mathrm{sp}}h\nu(G-1)$, where n_{sp} is the spontaneous emission factor, h is Planck's constant, and ν is the signal's central frequency. We add the noise power in our simulations using the Monte

Carlo method [32]. We may now return to the original coordinate system by writing

$$\mathbf{A}^a = \begin{pmatrix} \tilde{u}_x^a \\ \tilde{u}_y^a \end{pmatrix} = \begin{pmatrix} \tilde{u}_\parallel^a \cos(\varphi/2) + \tilde{u}_\perp^a \sin(\varphi/2) \\ \tilde{u}_\parallel^a \sin(\varphi/2) \exp(i\phi) - \tilde{u}_\perp^a \cos(\varphi/2) \exp(i\phi) \end{pmatrix}. \tag{5.8}$$

In order to explain the connection with these results of the signal and noise powers calculated in the reduced model as described in Eqs. (4.9)–(4.11), I begin by considering the first Stokes parameter

$$S_{1,\text{total}}^a = \int_{-\infty}^{\infty} (|\tilde{u}_x^a|^2 - |\tilde{u}_y^a|^2) \frac{d\omega}{2\pi}. \tag{5.9}$$

This quantity equals the first Stokes parameter as defined by integrating over the time domain in Eq. (2.10). Substituting Eq. (5.8) into Eq. (5.9), I obtain

$$S_{1,\text{total}}^a = \int_{-\infty}^{\infty} \left[(|\tilde{u}_\parallel^a|^2 - |\tilde{u}_\perp^a|^2) \cos\varphi + (u_\parallel^a u_\perp^{a*} + u_\parallel^{a*} u_\perp^a) \sin\varphi \right] \frac{d\omega}{2\pi}. \tag{5.10}$$

Next, substituting Eq. (5.7) into Eq. (5.10), I find

$$S_{1,\text{total}}^a = S_1^a + S_{1,\text{noise}}^a, \tag{5.11}$$

where

$$S_1^a = \int_{-\infty}^{\infty} G \left[(|\tilde{u}_\parallel|^2 - g|\tilde{u}_\perp|^2) \cos\varphi + g^{1/2}(\tilde{u}_\parallel \tilde{u}_\perp^* + \tilde{u}_\parallel^* \tilde{u}_\perp) \sin\varphi \right] \frac{d\omega}{2\pi}, \tag{5.12a}$$

$$S_{1,\text{noise}}^a = \int_{-\infty}^{\infty} H(\omega) \left[(|r_\parallel|^2 - g|r_\perp|^2) \cos\varphi + g^{1/2}(r_\parallel r_\perp^* + r_\parallel^* r_\perp) \sin\varphi \right] \frac{d\omega}{2\pi}. \tag{5.12b}$$

Physically, S_1 corresponds to the signal portion of $S_{1,\text{total}}$ and $S_{1,\text{noise}}$ corresponds to the noise portion. The factor $H(\omega)$ equals 1 over the bandwidth of the amplifier and zero elsewhere. In this calculation, I have assumed that the cross terms between the signal and the noise integrate to zero over the bandwidth of the amplifier. This physically reasonable assumption was also implicitly used in our reduced model. Of

course, in a single mark the contribution of the signal-spontaneous beat noise is typically larger than the contribution of the spontaneous-spontaneous beat noise. In our complete model, the signal-spontaneous beat noise is automatically present. In our reduced model, it is included in Eq. (4.13) by using the relation for each channel

$$\frac{S_{0,\text{spon}-\text{spon}}^{\prime(m)}}{S_{0,\text{sig}-\text{spon}}^{\prime(m)}} = \left(\frac{S_{0,\text{noise}}^{\prime(m)}}{2S_0^{(m)}}\right)^{1/2}\left(\frac{2B_{\text{elec}}}{B_{\text{opt}}}\right)^{1/2} \tag{5.13}$$

that I previously discussed in Sec. 4.2. Knowing the evolution of the Stokes parameters for the signal and noise separately, it is thus possible to estimate the evolution of their beat terms.

To evaluate Eq. (5.12a), I return to the original $(\widetilde{u}_x,\ \widetilde{u}_y)^t$ coordinate system, and I find

$$
\begin{aligned}
S_1^a &= \int_{-\infty}^{\infty}\Big\{-2\alpha\varepsilon\cos\varphi(|\widetilde{u}_x|^2+|\widetilde{u}_y|^2)+[(\alpha^2+\varepsilon^2)-2\varepsilon^2\sin^2\varphi](|\widetilde{u}_x|^2-|\widetilde{u}_y|^2)\\
&\quad+2\varepsilon^2\cos\varphi\sin\varphi[\exp(i\phi)\widetilde{u}_x\widetilde{u}_y^*+\exp(-i\phi)\widetilde{u}_x^*\widetilde{u}_y]\Big\}\frac{d\omega}{2\pi},
\end{aligned}
\tag{5.14}
$$

where I recall $\alpha = (g^{1/2}-1)/2$ and $\varepsilon = (g^{1/2}+1)/2$. Dividing S_1^a into its contributions from the different channels by analogy to what we did in the reduced model, I arrive at the equation for $S_{1,\text{after}}^{\prime(m)}$ that I derived earlier in the reduced model, displayed in Eq. (4.7). The equations for the other Stokes parameters that are shown in Eq. (4.7) can be derived from the full model in a similar fashion. Likewise, I can derive similar equations for the noise Stokes parameters. Thus, I have found that the reduced model can be obtained from the full model by ignoring the spontaneous-signal beat contributions which are then restored to the reduced model by using the relationship Eq. (5.13).

5.2 Evolution of the degree of polarization due to PMD

If the fiber PMD is sufficiently large, the polarization of the optical signal in each channel will be scrambled, leading to depolarization of the channel. In the Stokes model, since the signal in each channel is only described by four Stokes parameters, PMD can not lead to depolarization of a single channel. It can only lead to depolarization of a group of channels. It is important to show that the value of PMD in the undersea systems that I am modeling is not so large that single-channel depolarization due to PMD becomes important. To estimate when the single-channel PMD becomes important, I note that if I require the DGD to be no greater than 10% of the beat period, then after 9,000 km of propagation, I find that the PMD coefficient should be no larger than approximately 0.1 ps/km$^{1/2}$ for a channel operating at 10 Gbits/sec. To verify this, I ran full simulations to study the PMD-induced depolarization of a single channel propagating through 10,000 km. I set the PMD to 0.1 ps/km$^{1/2}$, and I shut off the nonlinearity, chromatic dispersion, the fiber loss, and the ASE noise, leaving only the PMD. I found that the depolarization of the channel is very small, as shown in Fig. 5.1. In modern-day trans-oceanic systems, the PMD is under 0.1 ps/km$^{1/2}$; so, I need not be concerned by depolarization due to PMD inside a single channel.

Indeed, it is possible today to manufacture fibers with PMD values of 0.025 ps/km$^{1/2}$, so that even single channels with a bit rate of 40 Gbits/sec can propagate over trans-oceanic distances without significant repolarization. In this case, the Stokes parameter model would still be valid.
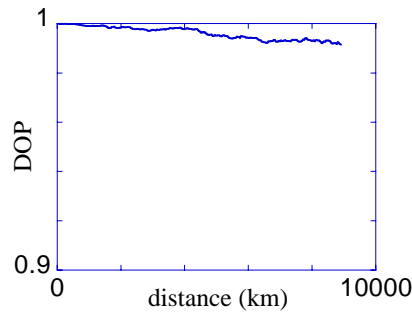
Figure 5.1: Evolution of the degree of polarization of a single channel due to PMD

5.3 Polarization evolution due to nonlinearity and chromatic dispersion

5.3.1 Introduction

As I discussed in Sec. 4.2, I neglect the effects of nonlinearity and dispersion when calculating the penalties due to PDL and PDG using the reduced model. In order for this assumption to be valid, the nonlinearity and dispersion should have no effect on the degree of polarization of a single channel. In this section, I will verify that this condition is obeyed for physically reasonable parameters corresponding to trans-oceanic transmission. This condition, just like the condition that PMD should not by itself polarization-scramble a single channel that I verified in the last section, is not sufficient to assure the validity of the reduced model. I will present a more complete validation in Chap. 6. However, since this condition must be obeyed for the reduced model to be valid, and since its breakdown is not difficult to detect computationally, it is useful for finding the parameter values beyond which the reduced model fails.

In order to clarify this problem, I numerically investigated the evolution of the degree of polarization of individual channels and their Stokes parameters in a WDM

system in which Kerr nonlinearity and chromatic dispersion are taken into account but in which polarization mode dispersion as well as PDL and PDG are neglected. I compared the results to a version of the reduced model in which I also neglected PDL and PDG but in which I kept the effect of the nonlinearity in a simplified form. In this case, the reduced model predicts no change in the degree of polarization of each of the channels so that an initially polarization-scrambled channel does not repolarize and an initially polarized signal will not depolarize. I will show shortly that the reduced model is strictly valid in this case in the limit of strong dispersion management.

5.3.2   Derivation of the Mean field Equation

Our starting point is the Manakov equation

$$i\frac{\partial \mathbf{U}}{\partial z} - \frac{\beta''}{2}\frac{\partial^2 \mathbf{U}}{\partial t^2} + \frac{8}{9}\gamma(\mathbf{U}^\dagger \mathbf{U})\mathbf{U} = 0, \tag{5.15}$$

where $\mathbf{U} = (u_x, u_y)^t$ represents the complex envelope of the two polarizations, β'' is the dispersion coefficient, γ is the nonlinear coefficient, and z and t are distance along the fiber and retarded time. This equation may be derived from Eq. (5.1) by averaging over the rapid variations of the polarization states on the Poincaré sphere, neglecting the terms that contribute to PMD [16]. Earlier experimental [19] and theoretical [9], [15] work has shown that this equation accurately describes nonlinear and dispersive light propagation in standard communication fiber with rapidly and randomly varying birefringence when polarization mode dispersion can be neglected. I will neglect PMD in most of this section in order to determine the impact of the Kerr effect and chromatic dispersion on the polarization states of the channels in a WDM systems in the absence of other effects, although I will briefly discuss its impact at the end of this section. I also note that Eq. (5.15) does not include the spatially varying gain and loss that is present in real systems; however, this effect is easily included and I will show how to do that at the end of this section.

Starting from Eq. (2.9), I write $\mathbf{U}$ as a sum of contributions over n channels, obtaining

$$\mathbf{U} = \sum_{m=1}^{n} \mathbf{U}^{(m)} \exp[ik^{(m)}z - i\omega^{(m)}t], \tag{5.16}$$

Substituting Eq. (5.16) into Eq. (5.15), I find

$$i\frac{\partial \mathbf{U}^{(m)}}{\partial z} \quad - \quad \frac{\beta''}{2}\frac{\partial^2 \mathbf{U}^{(m)}}{\partial t^2} + \frac{8}{9}\gamma(\mathbf{U}^{\dagger(m)}\mathbf{U}^{(m)})\mathbf{U}^{(m)} + \frac{8}{9}\gamma \sum_{q=1,\neq m}^{n} (\mathbf{U}^{\dagger(q)}\mathbf{U}^{(q)})\mathbf{U}^{(m)}$$

$$+\frac{8}{9}\gamma \sum_{q=1,\neq m}^{n} (\mathbf{U}^{\dagger(q)}\mathbf{U}^{(m)})\mathbf{U}^{(q)} = 0, \tag{5.17}$$

where, consistent with our assumption that the dispersion between channels is large, I neglect four wave mixing terms. The definition of the Stokes vector of each channel follows Eq. (2.10). Using Eq. (5.17) to determine the evolution of the Stokes parameters, I find $dS_0^{(m)}/dz = 0$, and I find for $dS_1^{(m)}/dz$ that

$$\frac{dS_1^{(m)}}{dz} = i\frac{8}{9}\frac{\gamma}{T}\int_{t_1}^{t_2} \left[\left(u_x^{(m)}u_y^{(m)*} + u_x^{(m)*}u_y^{(m)}\right) \sum_{q=1,\neq m}^{n} \left(u_x^{(q)}u_y^{(q)*} - u_x^{(q)*}u_y^{(q)}\right) \right.$$
$$\left. - \left(u_x^{(m)}u_y^{(m)*} - u_x^{(m)*}u_y^{(m)}\right) \sum_{q=1,\neq m}^{n} \left(u_x^{(q)}u_y^{(q)*} + u_x^{(q)*}u_y^{(q)}\right) \right] dt. \tag{5.18}$$

In a highly dispersive system, the channels for which $q \neq m$ rapidly pass through the m-th channel in the time domain. Consequently, the evolution of the m-th channel is only affected by the averaged time variation of the $q \neq m$ channels so that I may effectively treat them as continuous waves. This assumption is the heart of the mean field approximation. I thus replace

$$u_x^{(q)}u_y^{(q)*} - u_x^{(q)*}u_y^{(q)} \rightarrow \frac{1}{T}\int_{t_1}^{t_2} (u_x^{(q)}u_y^{(q)*} - u_x^{(q)*}u_y^{(q)})dt, \tag{5.19}$$

from which I conclude

$$\frac{dS_1^{(m)}}{dz} = \frac{8}{9}\gamma \sum_{q=1}^{n} \left(S_2^{(m)}S_3^{(q)} - S_3^{(m)}S_2^{(q)}\right). \tag{5.20}$$

I can find similar expressions for $dS_2^{(m)}/dz$ and $dS_3^{(m)}/dz$, so that I finally obtain

$$\frac{d\mathbf{S}^{(m)}}{dz} = \frac{8}{9}\gamma\mathbf{S}^{(m)} \times \sum_{q=1}^{n} \mathbf{S}^{(q)}. \tag{5.21}$$

The effect of dispersion does not appear in Eq. (5.21); only the nonlinearity appears, and the equations are analogous to the equations that govern nonlinear polarization

rotation of continuous wave beams [20]. However, the local dispersion plays a critical role since it must be large enough so that each channel only affects its neighbors through its Stokes parameters. I note as well that there is no change in $S_0^{(m)}$, *i.e.*, $dS_0^{(m)}/dz = 0$. From Eq. (5.21), it follows that the only effect of combining the Kerr effect with large dispersion is to rotate the polarization states of the different WDM channels. In particular, there is no change in the degree of polarization! This version of the reduced model contains the effect of nonlinear polarization rotation, while the reduced model presented in Chap. 3 does not. We can safely exclude the nonlinear polarization rotation because its effect is very small in trans-oceanic systems. In Chap. 6, we will directly validate the reduced model presented in Chap. 3.

It is interesting to notice that even though Eq. (5.21) is nonlinear, a complete analytical solution may be found:

$$\mathbf{S}^{(m)} = a_1 \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} + a_2 \begin{pmatrix} r \\ s \\ t \end{pmatrix} \exp[i(8/9)\gamma cz] + a_3 \begin{pmatrix} r^* \\ s^* \\ t \end{pmatrix} \exp[-i(8/9)\gamma cz], \tag{5.22}$$

$$r = c_1 c_3 + icc_2, \qquad s = c_2 c_3 - icc_1, \qquad t = c_3^2 - c^2,$$

$$c_1 = \sum_{q=1}^{n} S_{10}^{(q)}, \qquad c_2 = \sum_{q=1}^{n} S_{20}^{(q)}, \qquad c_3 = \sum_{q=1}^{n} S_{30}^{(q)},$$

$$c^2 = c_1^2 + c_2^2 + c_3^2, \qquad a_1 = D_1/\Delta, \qquad a_2 = D_2/\Delta, \qquad a_3 = D_3/\Delta,$$

$$\Delta = \begin{vmatrix} c_1 & r & r^* \\ c_2 & s & t \\ c_3 & t & t \end{vmatrix}, \qquad D_1 = \begin{vmatrix} S_{10}^{(m)} & r & r^* \\ S_{20}^{(m)} & s & s^* \\ S_{30}^{(m)} & t & t \end{vmatrix}, \qquad D_2 = \begin{vmatrix} c_1 & S_{10}^{(m)} & r^* \\ c_2 & S_{20}^{(m)} & s^* \\ c_3 & S_{30}^{(m)} & t \end{vmatrix},$$

$$D_3 = \begin{vmatrix} c_1 & r & S_{10}^{(m)} \\ c_2 & s & S_{20}^{(m)} \\ c_3 & t & S_{30}^{(m)} \end{vmatrix},$$

where $S_{10}^{(q)}$, $S_{20}^{(q)}$, $S_{30}^{(q)}$, $S_{10}^{(m)}$, $S_{20}^{(m)}$, and $S_{30}^{(m)}$ are the initial values of the Stokes vector in

the q-th and m-th channels. This result is intrinsically significant because the number of large-dimensional nonlinear systems for which exact solutions can be found is small.

When fiber loss and lumped gain at the amplifier are introduced into the transmission line, Eq. (5.15) becomes

$$i\frac{\partial \mathbf{U}}{\partial z} - \frac{\beta''}{2}\frac{\partial^2 \mathbf{U}}{\partial t^2} + \frac{8}{9}\gamma(\mathbf{U}^\dagger\mathbf{U})\mathbf{U} = -i\alpha\mathbf{U} + ig\sum_{p=1}^{N_{\mathrm{amp}}}\delta(z - pL_a)\mathbf{U}, \tag{5.23}$$

where α is the loss coefficient, g is the amplifier gain, L_a is the amplifier spacing, and N_{amp} is the total number of amplifiers. Since the loss and gain occur periodically in real systems on a length scale that is short compared to the scale on which the nonlinearity and average chromatic dispersion operate, they will induce periodic oscillations in the amplitude without changing the state of polarization. Since the Stokes vector components also periodically oscillate as a result, $\mathbf{S}^{(q)}$ in Eq. (5.21) will be smaller than its initial value right after an amplifier during most of the propagation, and $\mathbf{S}^{(m)}$ will evolve more slowly than if $\mathbf{S}^{(q)}$ always had this initial value. Defining $\mathbf{U}(z,t) = \bar{\mathbf{U}}(z,t)\exp[\alpha(z - pL_a)]$ in the interval $pL_a < z < (p+1)L_a$, so that I normalize the amplitude throughout the interval between amplifiers to its value at the beginning of the interval, and defining $\bar{\mathbf{S}}^{(m)}(z)$ from $\bar{\mathbf{U}}(z,t)$ analogous to Eq. (5.18), I find that

$$\frac{d\bar{\mathbf{S}}^{(m)}}{dz} = \frac{8}{9}\gamma\bar{\mathbf{S}}^{(m)} \times \sum_{q=1}^{n}\bar{\mathbf{S}}^{(q)}\exp[-2\alpha(z - pL_a)]. \tag{5.24}$$

Defining now, $\bar{z} = \bar{z}_p + \int_{pL_a}^{z}\exp[-2\alpha(z' - pL_a)]dz' = \bar{z}_p + (1/2\alpha)[1 - \exp[-2\alpha(z - pL_a)]]$, where $\bar{z}_p$ is a constant, I conclude

$$\frac{d\bar{\mathbf{S}}^{(m)}(\bar{z})}{d\bar{z}} = \frac{8}{9}\gamma\bar{\mathbf{S}}^{(m)}(\bar{z}) \times \sum_{q=1}^{n}\bar{\mathbf{S}}^{(q)}(\bar{z}), \tag{5.25}$$

which has the same form as Eq. (5.21). If I let $\bar{z}_p = (p/2\alpha)[1 - \exp(-2\alpha L_a)] = pL_{\text{eff}}$ then $\bar{z}(z)$ varies continuously, and Eq. (5.25) can be interpreted as an evolution equation dependent upon this new variable. The variation of $\bar{z}(z)$ is somewhat complicated by the exponential variation, but in most realistic settings, the nonlinear evolution is slow compared to the periodic evolution due to the fiber loss and lumped gain so that it is sufficient to approximate [12]

$$\frac{d\bar{\mathbf{S}}^{(m)}(\bar{z})}{dz} = \frac{L_{\text{eff}}}{L_a}\frac{8}{9}\gamma\bar{\mathbf{S}}^{(m)}(z) \times \sum_{q=1}^{n}\bar{\mathbf{S}}^{(q)}(z). \tag{5.26}$$

I will validate this in the following part.

When I include the effect of polarization mode dispersion, Eq. (5.15) becomes

$$i\frac{\partial\mathbf{U}}{\partial z} - \frac{\beta''}{2}\frac{\partial^2\mathbf{U}}{\partial t^2} + \frac{8}{9}\gamma(\mathbf{U}^\dagger\mathbf{U})\mathbf{U} = -i\Delta\beta'\bar{\sigma}\frac{\partial\mathbf{U}}{\partial t}, \tag{5.27}$$

where $\Delta\beta'$ indicate the inverse group velocity difference between the fast and slow axes while $\bar{\sigma}$ is a rapidly varying unitary matrix that takes into account the changes in orientation of the birefringent axes [16]. I solve this equation using the coarse step method [16].

5.3.3   Basic Numerical Model

In our simulations, I used standard split-step methods to solve Eq. (5.15). In each channel, I used synchronous phase modulation and in some cases synchronous amplitude modulation of non-return to zero (NRZ) signals, much as described by Bergano, *et al.* [3], to both polarization scramble the signals and minimize pulse distortion. The functional form that I used for the initial field at the entry to the fiber is

$$u_x^{(m)}(t) \ = \ A_x(t)\exp[i\phi_x(t)], \tag{5.28a}$$

$$u_y^{(m)}(t) \ = \ A_y(t)\exp[i\phi_y(t)], \tag{5.28b}$$

when the signal has no amplitude modulation and

$$u_x^{(m)}(t) \ = \ A_x(t)\exp[i\phi_x(t)]\cos(\Omega t/2 + \pi/2), \tag{5.29a}$$

$$u_y^{(m)}(t) \ = \ A_y(t)\exp[i\phi_y(t)]\cos(\Omega t/2 + \pi/2), \tag{5.29b}$$

when the signal is amplitude-modulated, where $\phi_x(t) = \delta_x + a_x\cos(\Omega t + \psi_x + \pi/2)$ and $\phi_y(t) = \delta_y + a_y\cos(\Omega t + \psi_y + \pi/2)$. Here, I let $A_x(t) = c_x H(t)$ and $A_y(t) = c_y H(t)$, where c_x and c_y are constant coefficients, while $H(t) = 1$ in the time slots of the marks and $H(t) = 0$ in the time slots of the spaces except when making a transition from a space to a mark or from a mark to a space. When making a transition from a space to a mark, I set $H(t) = \{1 + \tanh[(t - t_0)/t_{\rm rise}]\}/2$, where t_0 is the boundary between the space and the mark and $t_{\rm rise} = 10$ ps for a 5 Gbits/sec signal and $t_{\rm rise} = 5$ ps for a 10 Gbits/sec signal. When making a transition from a mark to a space, I define $H(t) = \{1 + \tanh[-(t - t_0)/t_{\rm rise}]\}/2$. The coefficients a_x and a_y give the strength of the phase modulation. For all the simulations in this paper, I set $a_x = 3.307$ and $a_y = 0.903$ so that the difference $a_x - a_y$ nearly equals to 2.405, the first zero

of the zeroth Bessel function. With this choice, setting $c_x = c_y$, an ideal square pulse would be completely depolarized. In our simulations, since the pulses are never perfectly square, even without amplitude modulation, there is always a small residual degree of polarization like in the experiments of Bergano *et al.* The sum $a_x + a_y$ was chosen consistent with the experiments of Bergano, *et al.* [3]. The phase modulation frequency Ω corresponds to the bit rate, and ψ_x and ψ_y describe the relative phase between the phase modulation and the data bits. By varying ψ_x, ψ_y, c_x, and c_y, I can adjust the degree of polarization to any desired value. I used a 64 bit pattern in each channel, chosen so that the number of marks and spaces is identical and so that strings of marks of varying sizes are in the pattern. I experimented with a number of different strings in several cases and verified that our results are insensitive to this choice. We used 8192 node points for each channel in all cases which experimentation showed was adequate.

I chose the system parameters as follows: the central wavelength $\lambda = 1.58$ μm, the Kerr coefficient $n_2 = 2.6 \times 10^{-20}$ m^2/W, and a total fiber length $L = 10,000$ km. For our basic simulations, I used a dispersion map with a map length of 1,000 km like that of Bergano, *et al.* [3], but I also investigated the effect of shortening the map, and I will report these results in Sec. 5.3.4. In every map period, there is a span of length L_1 that consists of standard fiber with dispersion value $D_1(\lambda)$ and a span of length L_2 that consists of dispersion-shifted fiber with dispersion value $D_2(\lambda)$. I chose the lengths of these two portions so that $L_1 : L_2 = 1 : 8.5$. At the central frequency $\lambda_0 = 1.58$ μm, I choose the corresponding dispersion values so that they are in the ratio $8.5 : -1$, and the net cumulative dispersion is zero. Typical experimental systems that model undersea systems would have a value of $D_1 = -2$ ps/nm-km and

$D_2 = 17$ ps/nm-km, but I often varied these values significantly in our simulations to compare our results to the mean field approach. I note that values in terrestrial systems are often substantially higher. When $\lambda \neq \lambda_0$, third-order dispersion implies that the net dispersion is non-zero. In our simulations, I used $dD/d\lambda = 0.07$ ps/nm²-km. In our basic simulations, I did not use amplitude modulation, and the signaling rate is 5 Gbits/sec, corresponding to a bit period of 200 ps. The channel spacing is 0.5 nm.

In Fig. 5.2, I show a two-channel simulation. I note that because $a_x \neq a_y$, and ψ_x and ψ_y are different for the two channels, the two channels have slightly different values of S_2 and S_3 and hence slightly different degrees of polarization. The degree-of-polarization is non-zero in this case because $A_x \neq A_y$. Figure 5.2.a shows the theoretically predicted result from Eq. (5.23), and Fig. 5.2.b shows the evolution of the Stokes parameters with standard values of the dispersion. There are significant quantitative discrepancies between the mean field theory and the simulation, and these discrepancies only completely disappear when the local dispersion becomes quite large as shown in Fig. 5.2.c. Nonetheless, the Stokes parameters still oscillate around their initial values in Fig. 5.2.b, although with somewhat different frequencies and amplitudes than in Fig. 5.2.a. There are no long-term drifts in the Stokes parameters from the theoretically-predicted values. Thus, I would anticipate that there is little change in the degree of polarization of the channels, and this expectation is borne out as shown in Fig. 5.3 where I show the degree of polarization, $[S_1^2 + S_2^2 + S_3^2]^{1/2}/S_0$, for each channel over 10,000 km. The change is only about 0.02. In particular, if the degree-of-polarization is initially near zero for both channels, which I obtain by setting $A_x = A_y$ in this case, then they undergo little repolarization as shown

Figure 5.2: Evolution of the Stokes vector components as a function of distance in a 5 Gbits/sec system. The dispersion map length is 1000 km, and the channel spacing is 0.5 nm. The solid lines are the Stokes components of channel one; the dashed lines are the Stokes components of channel two. (a) Analytical result. (b) Simulation result, $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. (c) Simulation result, $D_1 = -20$ ps/nm-km, $D_2 = 170$ ps/nm-km. Other simulation parameters are $\lambda = 1550$ nm for channel one, $\lambda = 1550.5$ nm for channel two; $\psi_x = 0.7\pi$ and $\psi_y = 0$ for channel one, $\psi_x = 0$ and $\psi_y = 0.7\pi$ for channel two; the peak power in the x-polarization is 0.24 mW for channel one and is 0.2 mW for channel two; the peak power in the y-polarization is 0.2 mW for channel one and 0.24 mW for channel two.

Figure 5.3: Evolution of the degree of polarization as a function of distance with $D_1 = -2$ ps/nm-km, $D_2 =17$ ps/nm-km. Other parameters are the same as in Fig. 5.2.

in Fig. 5.4. As the number of channel increases, the agreement between the mean field theory and simulation improves somewhat because the presence of multiple channels
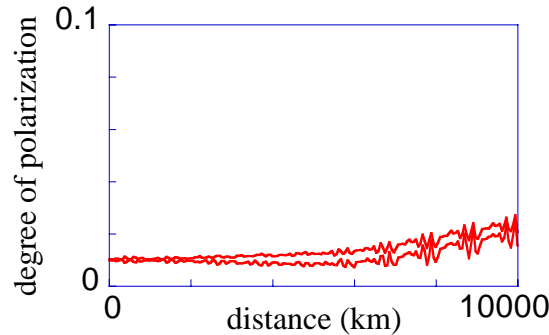
Figure 5.4: parameters are the same as in Fig. 5.1 except that the peak power in both the x- and y-polarization is 0.2 mW for both channels.

leads to better averaging over the different channels. In Fig. 5.5, I show the evolution with three channels. The oscillation periods of the Stokes parameters are reduced with respect to the case with two channels because I kept the power in each channel nearly the same as in the two-channel case. In Fig. 5.6, I show the evolution of the Stokes parameter S_1 with seven channels. I only show one of the Stokes parameters
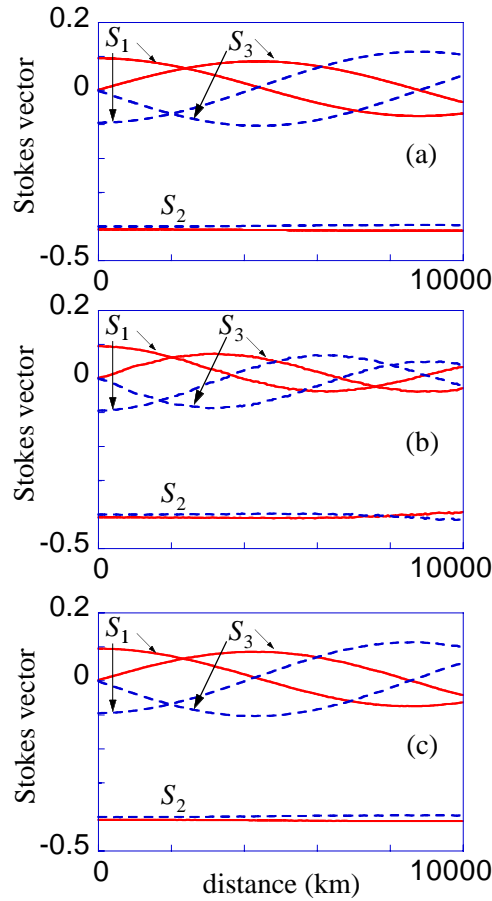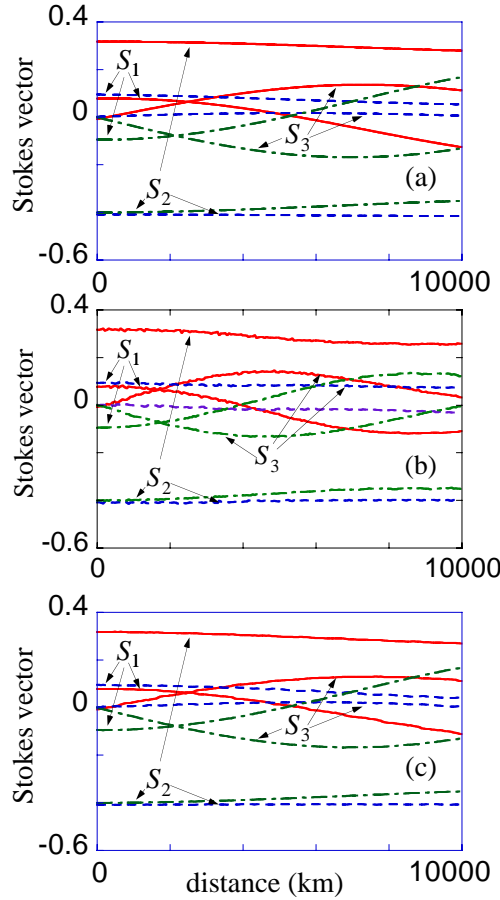
Figure 5.5: Evolution of the Stokes vector components as a function of distance in a 5 Gbits/sec system. The dispersion map is 1000 km and the channel spacing is 0.5 nm. The solid lines are the Stokes components of channel one, the dashed lines are the Stokes components of channel two, and the dash-dotted lines are the Stokes parameters of channel one. (a) Analytical result. (b) Simulation result, $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. (c) Simulation result, $D_1 = -20$ ps/nm-km, $D_2 = 170$ ps/nm-km. Other simulation parameters are $\psi_x = 0.7\pi$, $\psi_y = 0$ for channel one; $\psi_x = 0$, $\psi_y = 0.7\pi$ for channel two; $\psi_x = 0.4\pi$, $\psi_y = 0$ in channel three; the peak power in the x-polarization is 0.24 mW for channel one, 0.2 mW for channel two, and 0.23 mW for channel three; the peak power in the y-polarization is 0.2 mW for channel one, 0.24 mW for channel two, and 0.23 mW for channel three. Other parameters are the same as in Fig. 5.2.

Figure 5.6: Evolution of the Stokes vector component S_1 of seven channels as a function of distance in a 5 Gbits/sec system. The dispersion map length is 1000 km and the channel spacing is 0.5 nm. (a) Analytical result. (b) Simulation result, $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. (c) Simulation result, $D_1 = -10$ ps/nm-km, $D_2 = 85$ ps/nm-km. The channels are centered around $\lambda = 1550$ μm. Other parameters are: $\psi_x = 0$, $\psi_y = 0$ for all channels.
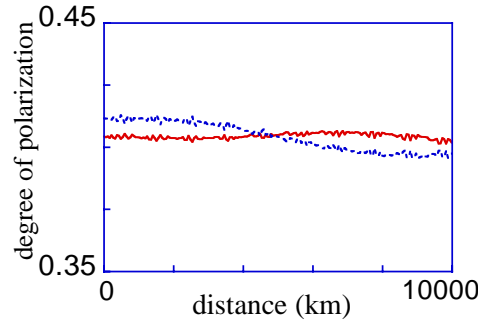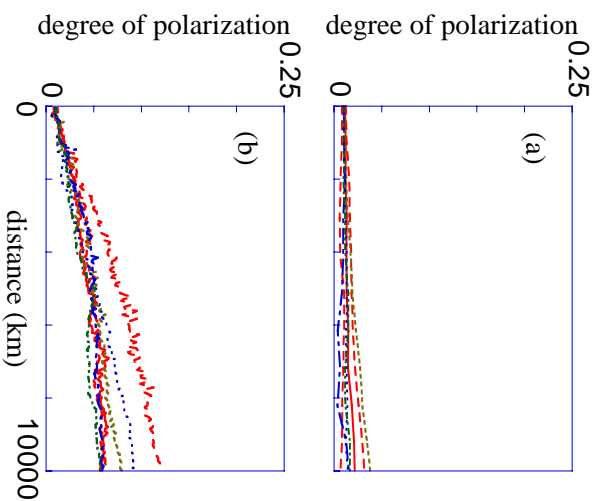
Figure 5.7: Evolution of the degree of polarization as a function of distance with $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. The peak power in the x- and y-polarization is 0.2 mW for all channels. (a) PMD=0 ps/km$^{1/2}$. (b) PMD=0.1 ps/km$^{1/2}$.

because the figure would be too busy if we included all three. Finally, in Fig. 5.7, I show the evolution of the degree of polarization of all seven channels. The worst channel changes by less than 0.03.

In Fig. 5.7, I also show the effect of adding a moderate polarization mode dispersion of 0.1 ps/km$^{1/2}$. I observe a slight increase in the repolarization in all of the channels. This increase occurs because even moderate polarization mode dispersion changes the polarization state slightly from what it would have been across the spectrum of the channel. As a consequence, the Stokes vectors are slightly shifted as a function of frequency from their values in the absence of polarization mode dispersion. I note that this increased repolarization is too small to affect our conclusions when repolarization due purely to nonlinearity and chromatic dispersion becomes too large

to be tolerated, indicating an unacceptable parameter regime.

5.3.4 Effects of Parameter Variation

Increasing the data rate:

In Fig. 5.8, I show the evolution of a two-channel system when the signaling rate is increased to 10 Gbits/sec, corresponding to a bit period of 100 ps. In this case, the channel spacing is 1 nm. I see that the agreement between the mean field theory and simulations at a typical dispersion value is better at this data rate than it was at 5 Gbits/sec. Moreover, at dispersion values of $D_1 = -10$ ps/nm-km, $D_2 = 85$ ps/nm-km, the agreement is already complete. The wider channel spacing required to avoid intersymbol interference and the smaller bit period lead to better averaging. As before, when I increase the number of channels, I find that agreement improves.

Adding gain and loss:

I add gain and loss to our basic model by assuming a loss coefficient $\alpha = 0.2$ db/km and an amplifier spacing of 50 km. As shown in Fig. 5.9, the results are essentially identical to Fig. 5.8, except that the oscillations are stretched out by a factor L_a/L_{eff} as predicted by Eq. (5.26). I have left out the bars above the Stokes vector in Fig. 5.9.

Adding amplitude modulation:

I now add amplitude modulation to our basic model as shown in Eqs. (5.29a) and (5.29b). The results are shown in Fig. 5.10. The behavior is similar to that shown in Fig. 5.2. At realistic values of the dispersion, shown in Fig. 5.10.b the mean field theory differs quantitatively from the actual simulation, but shows no large drifts, and when the dispersion is 20 times larger, as shown in Fig. 5.10.c, the results become
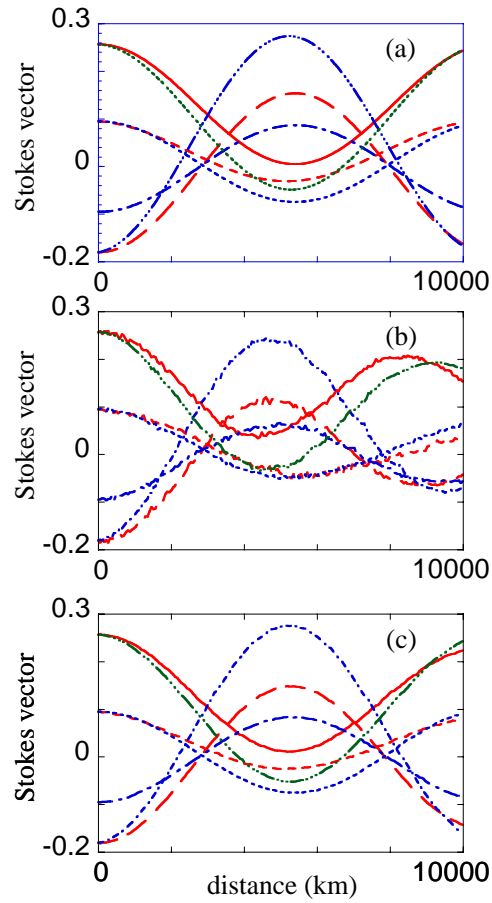
Figure 5.8: Evolution of the Stokes vector components as a function of distance for a 10 Gbits/sec system. The dispersion map length is 1000 km, and the channel spacing is 1 nm. The solid lines are the Stokes components of channel one; the dashed lines are the Stokes components of channel two. (a) Analytical result. (b) Simulation result, $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. (c) Simulation result, $D_1 = -10$ ps/nm-km, $D_2 = 85$ ps/nm-km. Channel one is at $\lambda = 1550$ μm, and channel two is at $\lambda = 1551$ μm. Other parameters are the same as in Fig. 5.2.
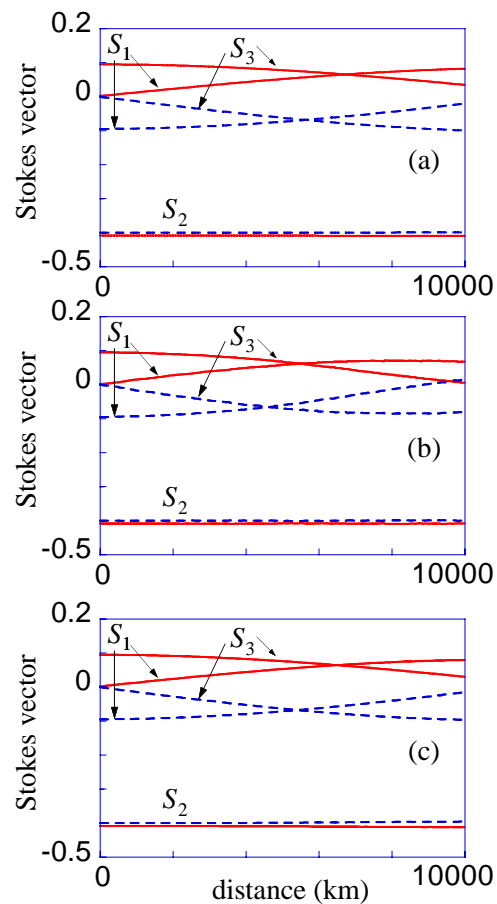
Figure 5.9: Evolution of the Stokes vector components as a function of distance with gain and loss added. The system parameters are the same as in Fig. 5.2.
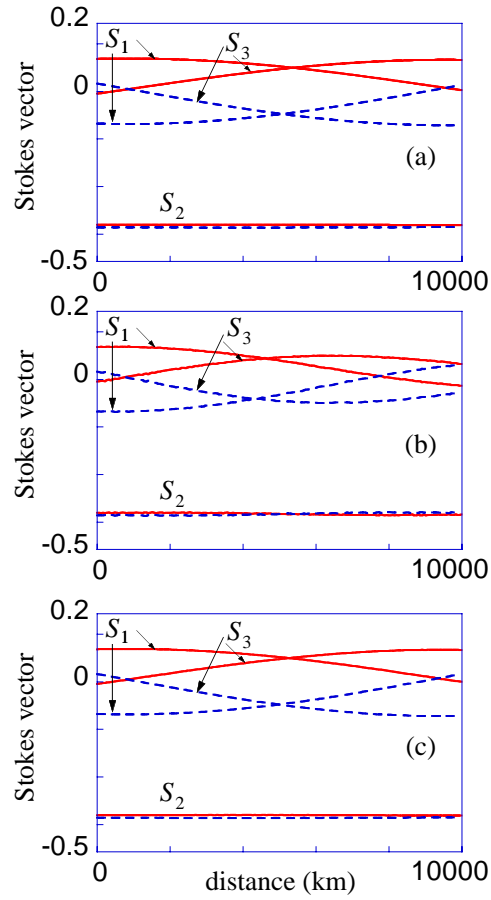
Figure 5.10: Evolution of the Stokes vector components as a function of distance with amplitude modulation added. The system parameters are the same as in Fig. 5.1. The dispersion map is 1000 km, and the channel spacing is 0.5 nm. The amplitude modulation is added. The solid lines are the Stokes components of channel one; the dashed lines are the Stokes components of channel two. (a) Analytical result. (b) Simulation result, $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. (c) Simulation result, $D_1 = -20$ ps/nm-km, $D_2 = 170$ ps/nm-km. The other parameters are the same as Fig. 5.2.

identical. I conclude that the detailed pulse shape is not important in the evolution of the polarization states.

Changing channel spacing:

Increasing the channel spacing while keeping the dispersion map fixed increases the relative velocity of the channels and is nearly equivalent to increasing the magnitude of the local dispersion values by the same factor. In Fig. 5.11, I show the result of setting the channel spacing to 2.5 nm. Agreement between the theory and the simulation is almost complete.

However, when I reduce the channel spacing I find large changes in the polarization evolution due to the nonlinear interaction between channels. Phase modulation expands the spectral width of each channel, leading to channel crosstalk. Simulations indicate that the channel spacing is limited to 0.3 nm before these effects become intolerably large. From a purely theoretical standpoint, these results imply that the mean field approach can only be used for channel spacing of 0.3 nm or more.

Reducing the map length:

In this case, the averaging becomes worse because the channels move back and forth through each other with a smaller period. Consequently, the agreement between the mean field theory and simulation deteriorates as shown in Fig. 5.12. While there is improvement as the number of channels increases, significant differences remain with up to seven channels. In Fig. 5.13, I show the degree of polarization for each channel in a seven channel system. I see that there is a non-negligible repolarization in some cases, amounting to a change in the degree of polarization of nearly 0.1 in the worst case. Thus, the mean field approach fails in this case.
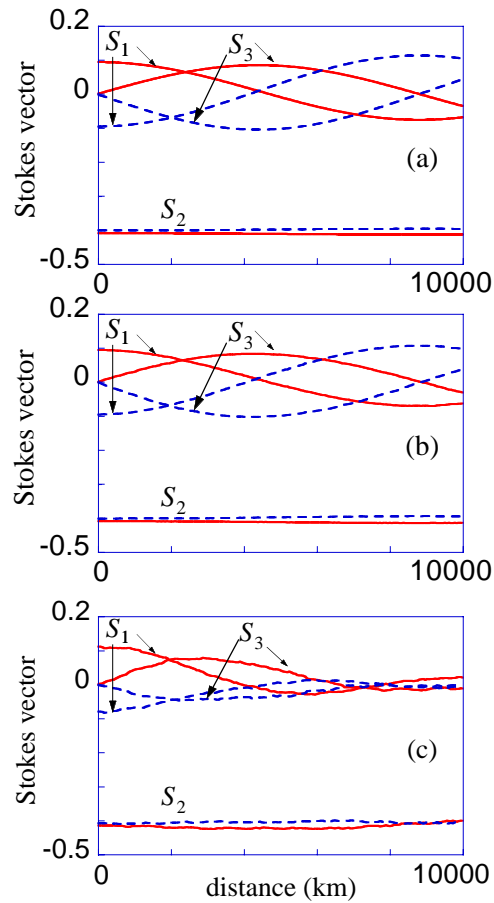
Figure 5.11: Evolution of the Stokes vector components as a function of distance with variation of the channel spacing. (a) Analytical result. (b) Simulation result. Channel one is at $\lambda = 1550$ nm, and channel two is at $\lambda = 1552.5$ nm. (c) The channel spacing is 0.3 nm. Other parameters are the same as in Fig. 5.2.
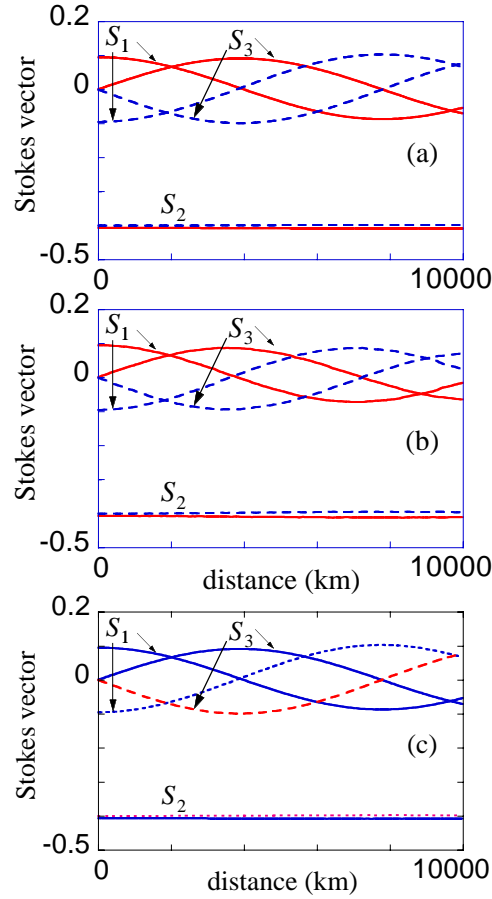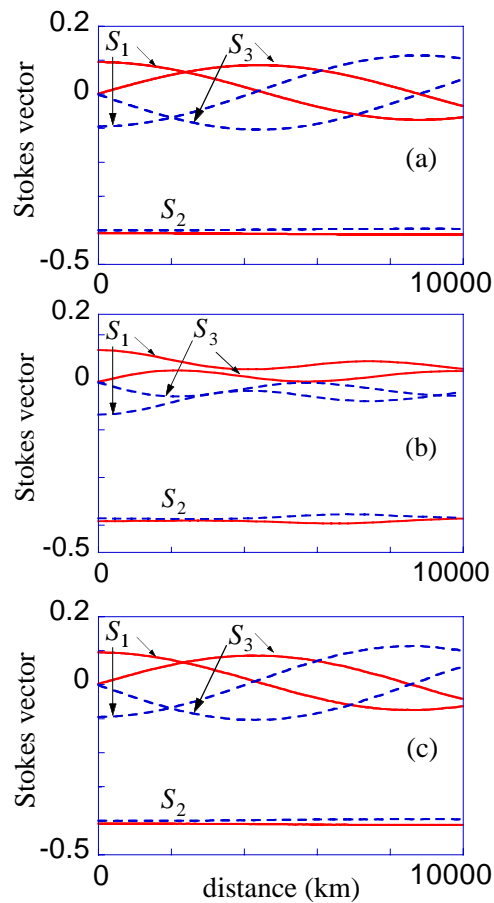
Figure 5.12: Evolution of the Stokes vector components as a function of distance in a 5 Gbits/sec system. The dispersion map length is 200 km. (a) Analytical result. (b) Simulation result, $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. (c) Simulation result, $D_1 = -20$ ps/nm-km, $D_2 = 170$ ps/nm-km. The other parameters are the same as in Fig. 5.2.
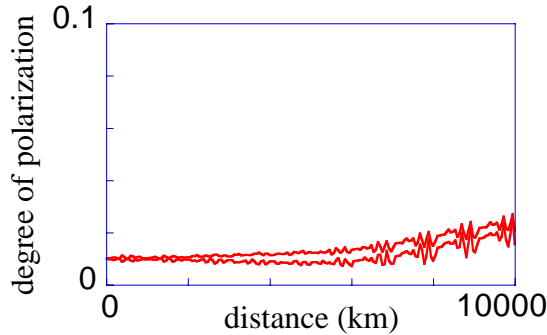
Figure 5.13: Evolution of the degree of polarization of 7 channels as a function of distance with $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. The dispersion map length is 200 km. Other system parameters are the same as in Fig. 5.6.

Adding filtering:

In soliton WDM systems, in-line Fabry-Perot filters are sometimes added to reduce the timing jitter due to the Gorden-Haus effect and soliton collisions [10]. Here, I examine its effect on the polarization state of a polarization-scrambled, NRZ WDM signal. When the filter acts on each channel, the effect is to disturb the phase relationship between the different frequency components which make up the channel's signal so that the signal will repolarize. The transmission function $\hat{f}(\omega)$ of the filter may be written

$$\hat{f}(\omega) = \frac{1 - R}{1 - R \exp[i(\omega - \omega_f)z_d/c\tau_0]}, \qquad (5.30)$$

where $R = 0.03$ is the reflectivity, ω_f is the central frequency of the filter which is centered on one of the channels, and the the free spectral range $z_d/c\tau_0$ is chosen equal to the channel spacing. As shown in Fig. 5.14, there is a large amount of repolarization, virtually ruling out the use of in-line filters in real systems with an NRZ pulse format. From a purely theoretical standpoint, the mean field approach would have to be modified before it could be used with filters.
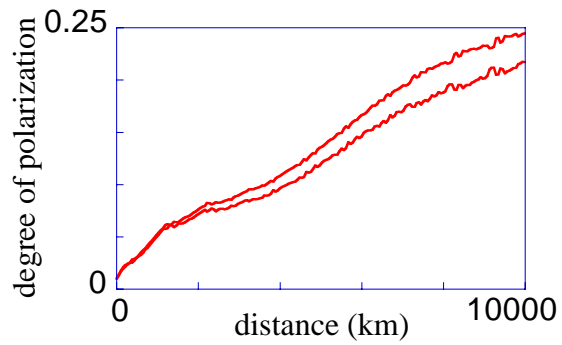
Figure 5.14: Evolution of the degree of polarization of two channels as a function of distance with $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. The Fabry-Perot filter is added. Other parameters are the same as in Fig. 5.2.

Chapter 6

Comparison of the Stokes model and the full model

I will now compare the Stokes model to the full model for both single-channel and eight-channel systems at a data rate of 10 Gbits/sec per channel. In the WDM studies, I used a 1 nm channel spacing. I have also carried out comparisons at 5 Gbits/sec that I will not present here because at the current time there is little commercial interest in systems at 5 Gbits/sec. In the full model, I studied non-return to zero (NRZ), return to zero (RZ), and chirped returned to zero (CRZ) data formats. I used a periodic dispersion map that consisted of one section of a single mode fiber whose dispersion D_1 at $\lambda_0 = 1.55$ μm is 16 ps/nm-km and whose length is 264 km, and another section of dispersion-shifted fiber whose dispersion D_2 at $\lambda_0 = 1.55$ μm is -2 ps/nm-km and whose length is 33 km. In both sections, I used a dispersion slope of 0.07 ps/nm^2-km. I used pre- and post-dispersion compensation, split equally, to compensate for the excess dispersion in channels for which $\lambda \neq \lambda_0$. At the end of the transmission line, I optically filtered the separate channels with a 60 GHz

tenth-order Bessel filter in my WDM simulations, and I then electrically filtered each channel using a 10 GHz tenth-order Bessel filter. I used clock recovery to determine the boundaries of the time slots in each channel. PMD, PDL, PDG, and ASE noise are included in the full model as described in Chap. 5. For each set of parameters I ran 20 cases, each of which corresponds to a different realization of the random variations of the birefringence axes of the fiber and a different realization of the ASE noise. However, I chose the bit string to be the same in each channel in all 20 cases in order to avoid Q-variations from case to case due to pattern dependences in the limited strings of 64 bits per channel that I could keep in my simulations.

For each set of parameters, I set the decision level for the marks and spaces empirically, so I can get the best SNR. I determined the SNR and then calculated the Q factor using Eq. (4.13) which I reproduce here,

$$Q^{(m)} = \frac{\text{SNR}^{(m)}}{\sqrt{2\text{SNR}^{(m)} + 1} + 1}\sqrt{\frac{2B_{\text{opt}}}{B_{\text{elec}}}}.$$
(6.1)

As I noted previously in Sec. 4.2, $Q^{(m)}$ is not meaningful when calculated using the reduced model. Only the difference $\Delta Q^{(m)}$ between the Q factor with and without PDL and PDG is meaningful. Consequently, I find this difference $\Delta Q^{(m)}$ for all 20 cases and from that I find the mean $\langle \Delta Q^{(m)} \rangle$ and the standard deviation $\sigma_Q^{(m)}$ for comparison to the reduced model.

Given the large random variation of the signal-spontaneous beat noise from bit-to-bit which leads to significant variations from case-to-case, 20 cases is not really sufficient to accurately determine $\langle \Delta Q^{(m)} \rangle$ and $\sigma_Q^{(m)}$. My colleagues and I only used 20 cases because for each set of parameters in which we ran an eight-channel simulation, the simulation required between 26 and 27 CPU hours on an SGI Onyx. Additionally,

there can be a large amount of variation due to pattern dependences since we only keep 64 bits in each channel. To illustrate the large variability that can occur due to pattern dependences, we ran 5 randomly chosen bit patterns for one parameter set with a single channel. The results, shown in Fig. 6.1 indicate that the variation due to pattern dependences is quite significant in this case. While the variability is significantly larger in this case than in other cases that we checked, some variability is always present. Thus, a comparison of the reduced model to the full model should be viewed as a demonstration of consistency rather than a complete check of the reduced model.
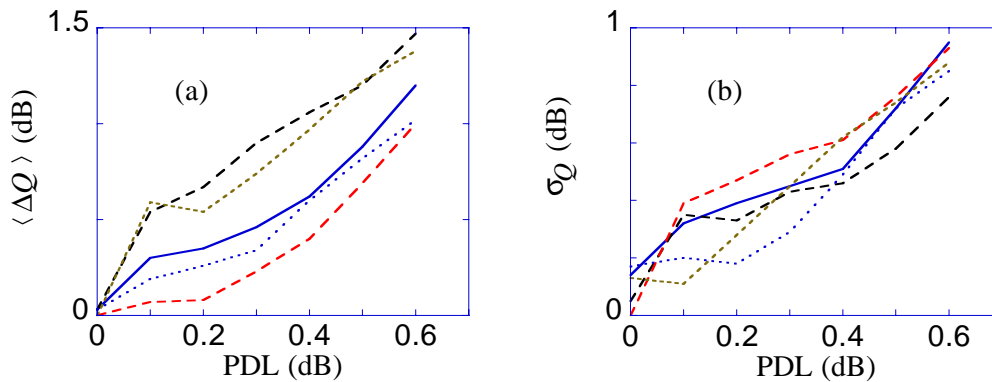
Figure 6.1: Pattern dependence demonstration with five different patterns. PMD = 0.1 ps/km$^{1/2}$, PDG = 0.06 dB; NRZ pulse format, polarization scrambling is added, (a) $\langle \Delta Q \rangle$, (b) σ_Q.

6.1   A single channel with the NRZ format

I first compare the full model to the Stokes model in the simple case when the pulse modulation format is NRZ and there is only a single channel. I show the results as a function of the PDL in Figs. 6.2 and 6.3, setting the PMD = 0.1 ps/km$^{1/2}$

and the PDG = 0.0 dB and 0.06 dB respectively. The agreement between the two models is quite good. The PDL values that I compared are 0.1, 0.2,..., 0.6 dB. We note that when $\sigma_Q = 1$ that the expected deviation of the full model from its mean is approximately $1/\sqrt{19} = 0.23$ since we only have 20 realizations at each value of the PDL. Thus I find that the deviation between the full model and the Stokes model lies within the expected statistical error of the full model. Indeed, given the

Figure 6.2: Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD = 0.1 ps/km$^{1/2}$, PDG = 0.0 dB; (a) $\langle \Delta Q \rangle$, (b) σ_Q. Solid lines indicate the Stokes model, and dashed lines indicate the average of the full model.

small number of realizations for the full model and the small number of bits in each realization, my colleagues and I consider the Stokes model to be at least as reliable as the full model. I note that the difference between the two models appears to be systematic rather than purely random since the full model consistently yielded lower values than the Stokes model for both $\langle \Delta Q \rangle$ and σ_Q as I varied the PDL. This systematic deviation is not surprising because the realizations with different values of PDL are not truly independent. First, we used the same bit pattern for all realizations at all values of PDL in order to avoid artificially enlarging σ_Q because of the effect

of pattern dependences. I have found that when I change the bit pattern, I change the dependence of $\langle \Delta Q \rangle$ and $\langle \sigma_Q \rangle$ on the PDL so that in some cases the full model yields larger values than the Stokes model due to the variability from the pattern dependences that I already showed in Fig. 6.1. Second, I reinitialized the random number generator for each new value of the PDL. Comparing Figs. 6.2 and 6.3, it
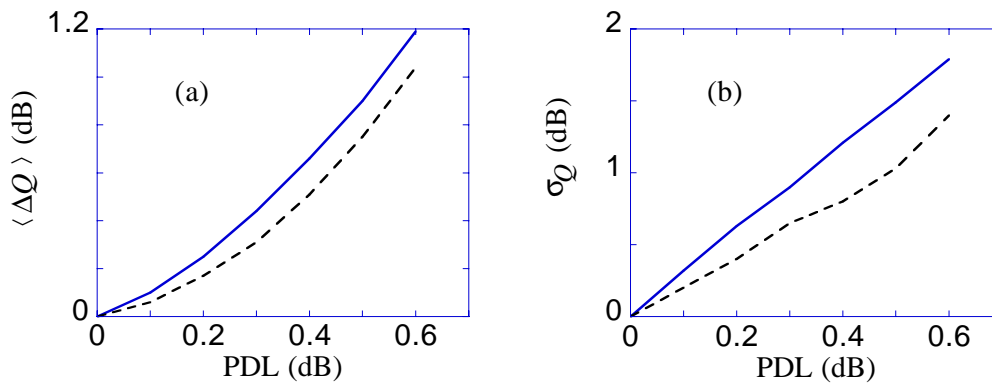

Figure 6.3: Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD = 0.1 ps/km$^{1/2}$, PDG = 0.06 dB; (a) $\langle \Delta Q \rangle$, (b) σ_Q. Solid lines indicate the Stokes model, and dashed lines indicate the average of the full model.

is apparent that PDG adds a substantial penalty almost independent of the PDL. When the PDL is 0.6 dB but the PDG is 0.0 dB, $\langle \Delta Q \rangle$ is only 1.2 dB. By contrast, when the PDG is 0.06 dB, $\langle \Delta Q \rangle$ is consistently almost 4.0 dB regardless of the PDL. However, σ_Q only increases slightly with non-zero PDG.

6.2   A single channel with the NRZ format adding polarization scrambling

Neal Bergano has proposed [3] to use polarization scrambling to reduce the degradation induced by PDG. So in this section we add polarization scrambling to the signal, as described by Eqs. (5.28) and (5.29), so that the signal is completely depolarized. I show results in Fig. 6.4 corresponding to a PDG of 0 dB. In Fig. 6.5, I add a PDG of 0.06 dB. As expected, we find a significant decrease in both $\langle \Delta Q \rangle$ and σ_Q in these two figures.

Figure 6.4: Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD = 0.1 ps/km$^{1/2}$, PDG = 0.0 dB; (a) $\langle \Delta Q \rangle$, (b) σ_Q. Solid lines indicate the Stokes model and dashed lines indicate the average of the full model.

6.3   A single channel with the RZ format

In this section, I discuss the evolution of the RZ pulse format when a single channel is transmitted. The parameters of the RZ pulses are the same as in Sec. 5.3.3. I show
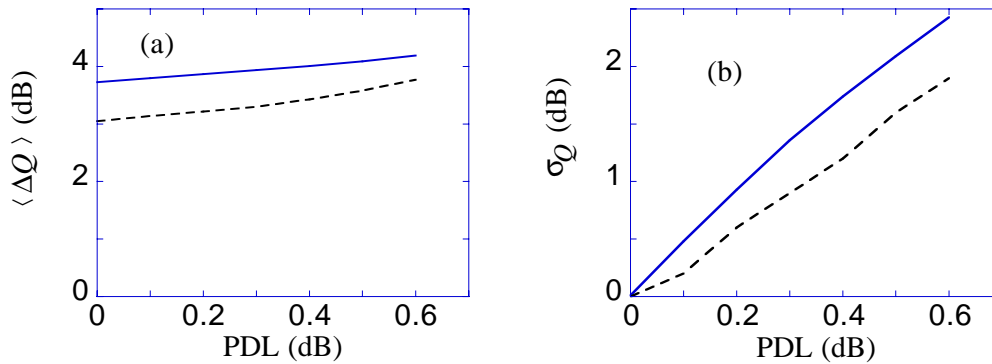
Figure 6.5: Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD = 0.1 ps/km$^{1/2}$, PDG = 0.06 dB; (a) $\langle \Delta Q \rangle$, (b) σ_Q. Solid lines indicate the Stokes model and dashed lines indicate the average of the full model. Shown here is one of the cases that appeared in Fig. 6.1.

the results in Fig. 6.6. Then I added PDG in each amplifier, so that the PDG equals to 0.06 dB. I show the result in Fig. 6.7. The results are similar to corresponding results with the NRZ format. I note that the variability due to pattern dependences is lower in this case than in the case of the NRZ format.
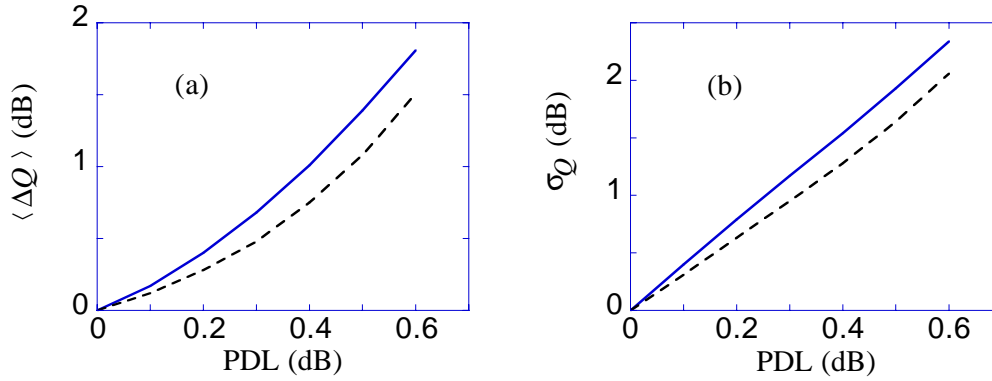
Figure 6.6: Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD = 0.1 ps/km$^{1/2}$, PDG = 0.0 dB; (a) $\langle \Delta Q \rangle$, (b) σ_Q. Solid lines indicate the Stokes model and dashed lines indicate the average of the full model.
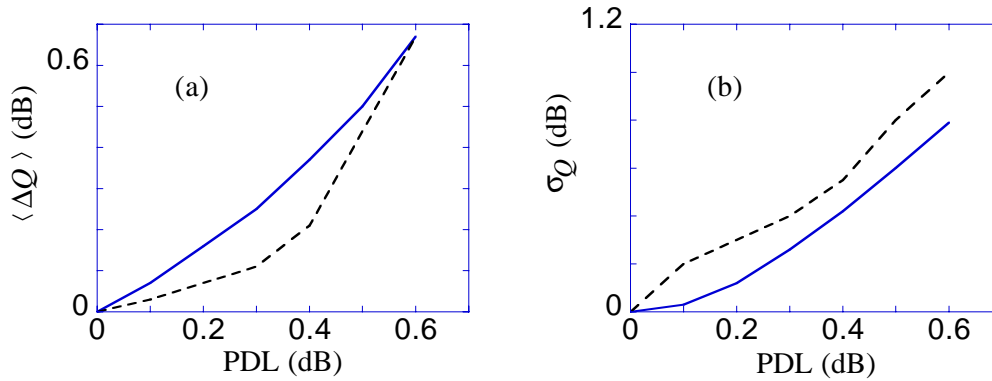

Figure 6.7: Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD = 0.1 ps/km$^{1/2}$, PDG = 0.06 dB; (a) $\langle \Delta Q \rangle$, (b) σ_Q. Solid lines indicate the Stokes model and dashed lines indicate the average of the full model.

6.4 Eight channels with the NRZ format

I now discuss an eight channel system that uses the NRZ format. I show the comparison between the full model and the reduced model in Fig. 6.8 when the PDG is 0 dB. The error bars on the dashed line show the standard deviation of the 8 channels. Comparison to Fig. 6.2 shows that the degradation is slightly higher than with a single channel, but the result is similar. With PDG included, I show the comparison
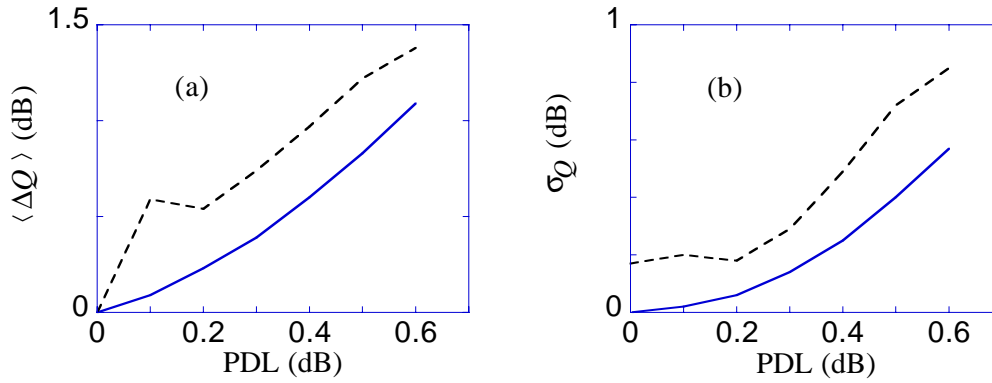

Figure 6.8: Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD = 0.1 ps/km$^{1/2}$, PDG = 0.0 dB; (a) $\langle \Delta Q^{(m)} \rangle$, (b) $\sigma_Q^{(m)}$. Solid lines indicate the Stokes model, dashed lines indicate the average of the full model, and the error bars indicate the standard deviation of values for all eight channels.

in Fig. 6.9. PDG increases the degradation, but its effect is much smaller than in the case of a single channel.

Figure 6.9: Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD = 0.1 ps/km$^{1/2}$, PDG = 0.06 dB; (a) $\langle \Delta Q^{(m)} \rangle$, (b) $\sigma_Q^{(m)}$. Solid lines indicate the Stokes model, dashed lines indicate the average of the full model, and the error bars indicate the standard deviation of values for all eight channels.

6.5 Eight channels with the NRZ format adding polarization scrambing

In this section I add phase scrambling, just like in Sec. 6.2. I show the results in Figs. 6.10 and 6.11. As expected, the degradation is smaller than in the cases without phase scrambling. I note that the variance of the Q factor in Fig. 6.11 is smaller than that in Fig. 6.10 when the PDL is large. The reason is that the PDG tends to depolarize the signal in contrast to the PDL which tends to repolarize the signal. Thus, when PDG is smaller, there is less depolarization and the variance is larger.
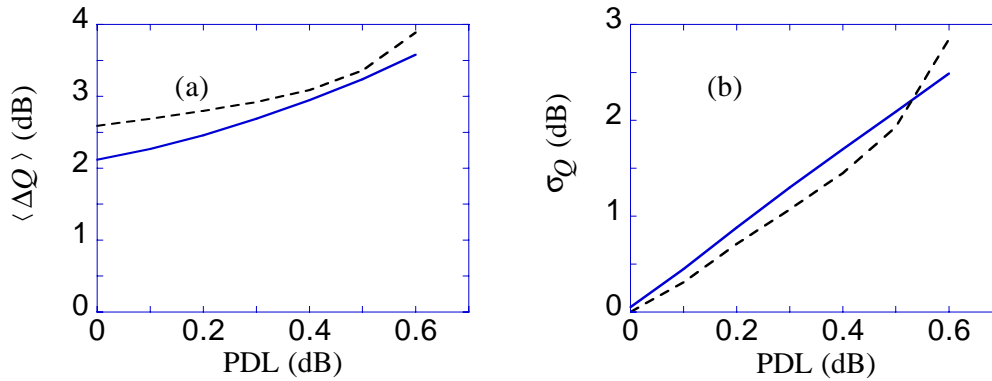
Figure 6.10: Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD = 0.1 ps/km$^{1/2}$, PDG = 0.0 dB; (a) $\langle \Delta Q^{(m)} \rangle$, (b) $\sigma_Q^{(m)}$. Solid lines indicate the Stokes model, dashed lines indicate the average of the full model, and the error bars indicate the standard deviation of values for all eight channels.
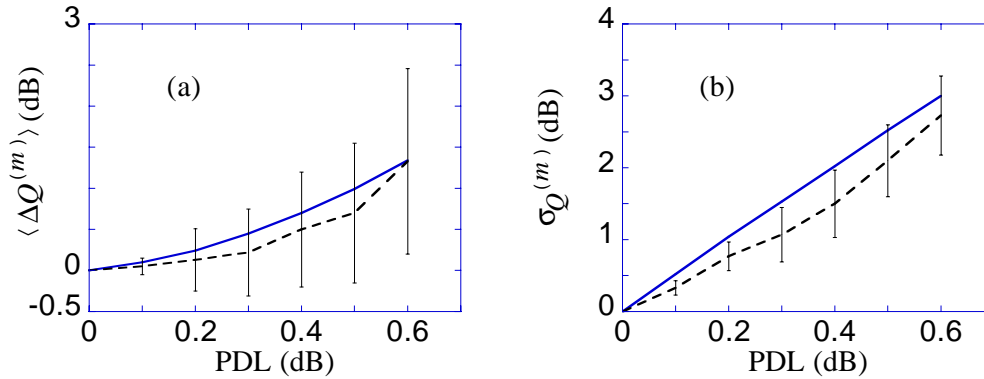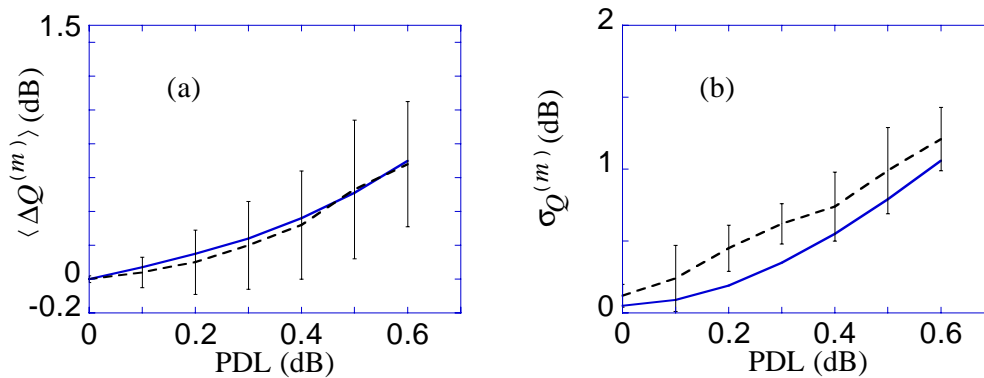

Figure 6.11: Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD = 0.1 ps/km$^{1/2}$, PDG = 0.06 dB; (a) $\langle \Delta Q^{(m)} \rangle$, (b) $\sigma_Q^{(m)}$. Solid lines indicate the Stokes model, dashed lines indicate the average of the full model, and the error bars indicate the standard deviation of values for all eight channels.

6.6   Eight channels with the RZ format

Next, I study the RZ pulse format without PDG. The pulse shape is the same as in Sec. 6.3. I show the result in Fig. 6.12. Again, the degradation is similar to what I found in the single channel case, shown in Fig. 6.6, but it is slightly larger. When I
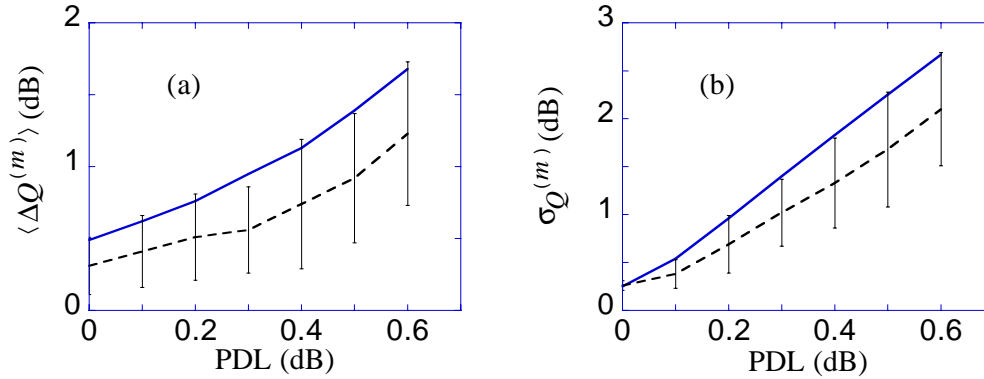

Figure 6.12: Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD $= 0.1$ ps/km$^{1/2}$, PDG $= 0.0$ dB; (a) $\langle \Delta Q^{(m)} \rangle$, (b) $\sigma_Q^{(m)}$. Solid lines indicate the Stokes model, dashed lines indicate the average of the full model, and the error bars indicate the standard deviation of values for all eight channels.

add PDG $= 0.06$ dB, I find as before that there is less additional degradation than in the single channel case, as shown in Fig. 6.13.

6.7   Eight channels with the CRZ format

Finally, I study a chirped RZ (CRZ) format because it is becoming widely used in long-distance communications systems. The pulses in this format have a raised cosine shape, just like the RZ pulses that I have already studied, but, additionally, I add the same phase modulation that Bergano, *et al.* [33] used in their experiments. To do
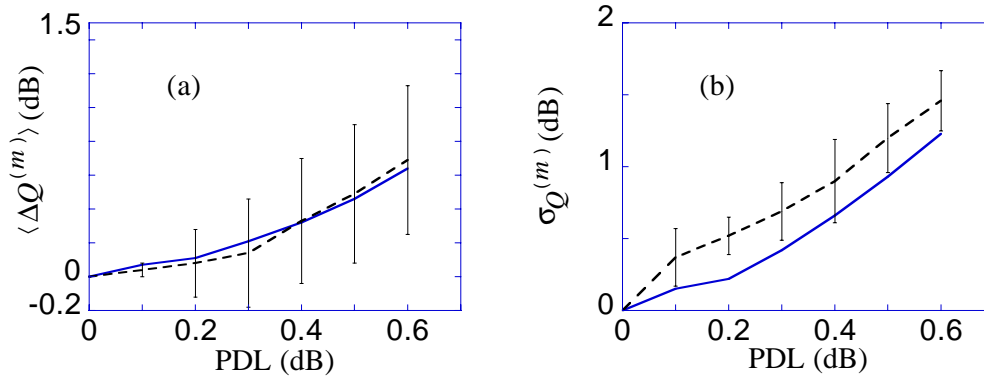
Figure 6.13: Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD = 0.1 ps/km$^{1/2}$, PDG = 0.06 dB; (a) $\langle \Delta Q^{(m)} \rangle$, (b) $\sigma_Q^{(m)}$. Solid lines indicate the Stokes model, dashed lines indicate the average of the full model, and the error bars indicate the standard deviation of values for all eight channels.

that, I multiply $u(z = z_0, t)$ by $\exp(i\phi)$, the synchronous optical phase modulation ϕ is adjusted to give the maximum chirp (i.e. $|d^2\phi/dt^2|$) aligned with the center of the pulse. I show the results in Fig. 6.14.

When the PDL is large, the difference between the full and the reduced models becomes large. This occurs because when the PDL is large, the signal is strongly distorted, and the clock recovery is not as effective as when the signal is less distorted, so that it is hard to measure the SNR accurately.

When there is PDG in the amplifier, PDG = 0.06 dB, I obtain the result shown in Fig. 6.15. The PDG induces additional degradation, particularly when the PDL is small, but its effect is much smaller than in the single channel case.
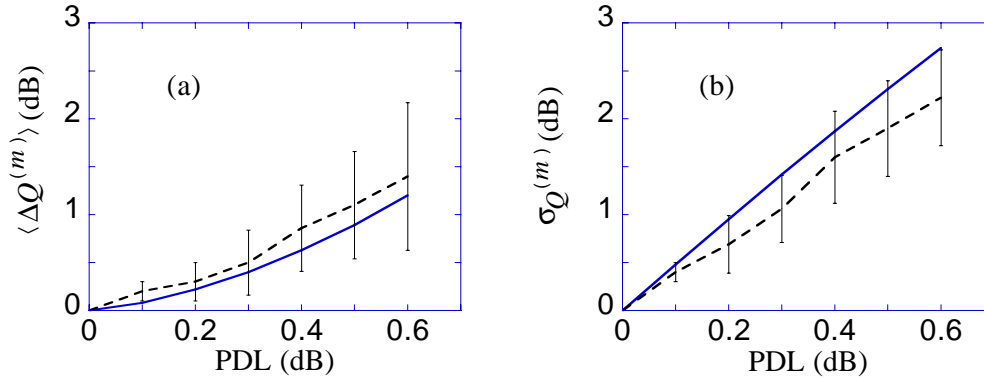
Figure 6.14: Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD = 0.1 ps/km$^{1/2}$, PDG = 0.0 dB; (a) $\langle \Delta Q^{(m)} \rangle$, (b) $\sigma_Q^{(m)}$. Solid lines indicate the Stokes model and dashed lines indicate the average of the full model, and the error bars indicate the standard deviation of values for all eight channels.
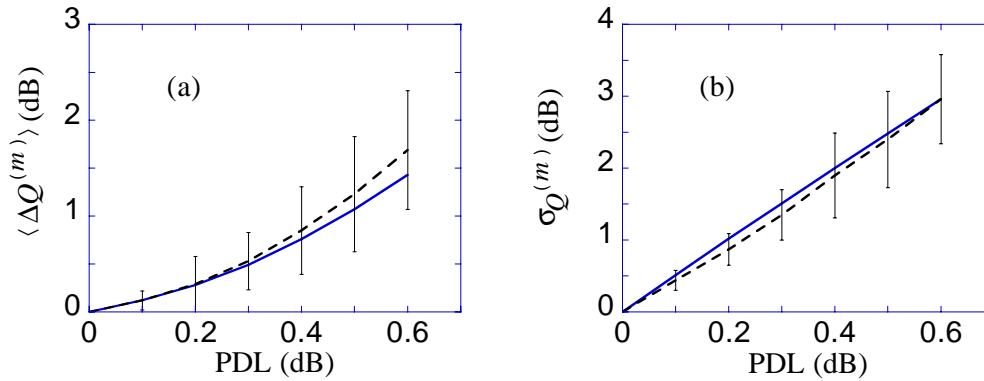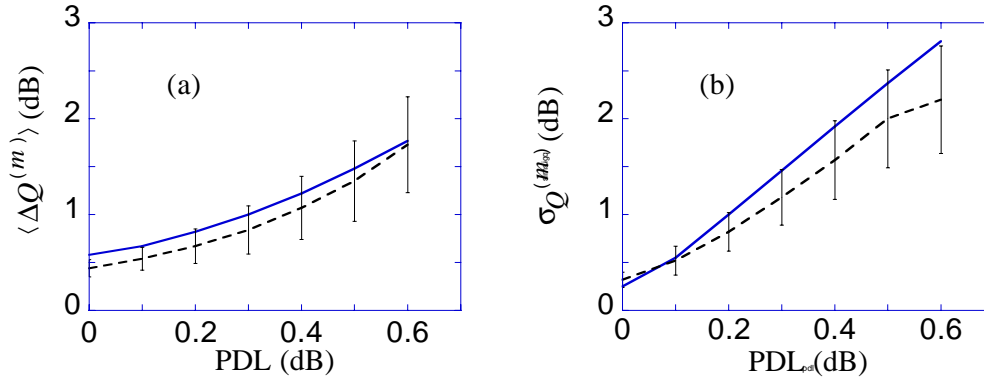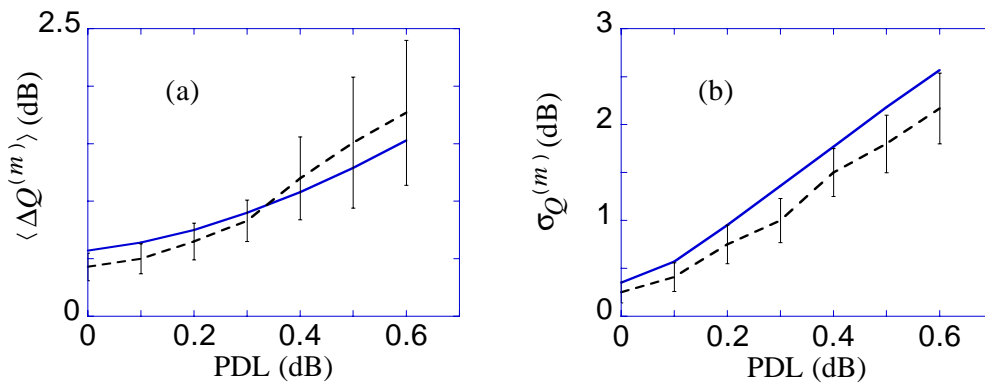

Figure 6.15: Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD = 0.1 ps/km$^{1/2}$, PDG = 0.06 dB; (a) $\langle \Delta Q^{(m)} \rangle$, (b) $\sigma_Q^{(m)}$. Solid lines indicate the Stokes model and dashed lines indicate the average of the full model, and the error bars indicate the standard deviation of values for all eight channels.

Chapter 7

Comparison between the Stokes model and experiments

In this chapter, I will compare the Stokes model to experiments that my colleagues and I carried out in a recirculating loop configuration. The basic configuration is described by Carter, *et al.* [21]. This recirculating loop is a little over 100 km long. Recirculating loops are a relatively simple and efficient approach to studying long-haul transmission systems [21]–[23]. However, it has long been known that it is necessary to use polarization controllers in loops that are shorter than about 500 km in order to obtain reasonable results because the measured bit error rate (BER) depends sensitively on the state of polarization controllers. The polarization behavior of short recirculating loops can be quite different from real transmission systems. Yet, this behavior had never been carefully characterized until my colleagues and I undertook to study it. Thus, the work that I will report here not only serves to validate the Stokes model, but it provides valuable insight into the behavior of the recirculating loops themselves. I will compare the evolution of the DOP in the experimental system

to the Stokes model. I will look at cases in which a 10 Gbits/sec pseudo-random signal is propagating and cases in which there is no initial signal and the light in the recirculating loop grows from noise.

In detail, the experimental system is a dispersion-managed recirculating loop that includes 100 km of dispersion-shifted fiber with a normal dispersion of -1.1 ps/nm-km at 1551 nm and $\simeq 7$ km of standard fiber with an anomalous dispersion of 16.7 ps/nm-km at 1551 nm. The entire loop comprises one period of the dispersion map. The PMD of the fiber is below 0.1 ps/km$^{1/2}$. A single 2.8 nm bandwidth optical filter and five EDFAs are in the loop.

My colleagues and I investigated the polarization evolution inside the loop using a commercial polarization analyzer, an HP 8509B, to measure the DOP. By sampling the Stokes parameters as a function of the propagation time, we could determine the DOP as a function of the propagation time or, equivalently, distance. We first measured the evolution of the DOP and the BER simultaneously up to 20,000 km while varying the polarization controllers. We observed that the polarization evolution
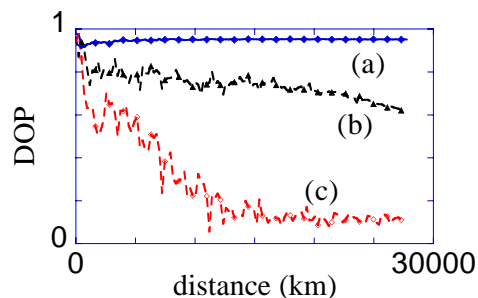
Figure 7.1: Evolution of the DOP corresponding to different BERs. Curve (a) corresponds to BER 10^{-9}, curve (b) corresponds to BER 10^{-6}, and curve (c) corresponds to BER 10^{-2}.

inside the loop was closely correlated to the BER. As shown in Fig. 7.1, the signal

was highly polarized when the BER was less than 1×10^{-10} at 20,000 km. When we set the polarization controllers so that the BER increased, the signal was increasingly depolarized with distance.

If the signal is initially polarized, it is difficult to separate the effects of PDL, PMD, and PDG. In our loop, the PMD is quite small and can be neglected. In order to separate the effects of PDL and PDG, we measured the DOP evolution inside the loop with pure ASE noise, which could be treated as a depolarized signal. First, we set the polarization controllers in the loop to obtain a BER less than 1×10^{-10} at 20,000 km. I show the evolution of the DOP in this case as a starred line in Fig. 7.2.a. Next, we shut off the signal and let the ASE noise propagate inside the loop, as shown in Fig. 7.2.b by the dashed line.
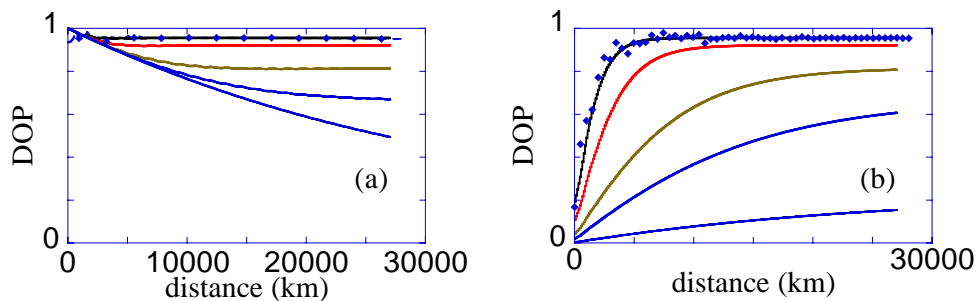
Figure 7.2: Evolution of the DOP; (a) signal plus noise, (b) noise only. The experimental results are shown as starred lines. The theoretical curves correspond, in order of decreasing DOP, to PDLs of 0.45, 0.25, 0.15, 0.05, and 0.01 dB.

In order to quantify the PDL, we opened our loop into a 107 km straight line and then measured the PDL of the entire line. Varying the polarization states of the input signal so that we covered the entire Poincaré sphere, we measured the output power as a function of the output polarization state. The difference between the maximum

power and minimum power equaled the PDL. We found that the PDL equaled 0.35 dB. This PDL is the lumped PDL of the whole loop.

In order to compare the Stokes model to these experiments, I had to modify the model in order to take into account the periodic nature of the recirculating loop. This periodicity is important because the PDL contributions are no longer random but repeat with the same period as the loop. Using the modified model, I found that when the PDL equals 0.45 dB, the results of the model are in exact agreement with the experimental results, shown in Fig. 7.2 as a starred line. The agreement between the model and the experiment is acceptable because we estimated the error in the measurement to equal approximately 0.1 dB, and the open loop did not contain the loop switches and couplers used in the closed loop experiments.

Keeping the same setup of the polarization controller in the Stokes model that yielded the lowest BER, I reduced the PDL from 0.45 dB to 0.25 dB, 0.15 dB, 0.05 dB, and finally 0.01 dB. I then calculated the DOP evolution in each case. I found that as the PDL became smaller, the noise played an increasingly important role, leading to an increased signal depolarization, as shown in Fig. 7.2.a. When the PDL equals 0.01 dB, the signal DOP falls below 0.5. The repolarization of the noise when there is no signal also becomes smaller as the PDL decreases. When the PDL equals 0.01 dB, I found that the DOP after 27,000 km just equals 0.2 as shown in Fig. 7.2.b.

Chapter 8

Application of the Stokes model to trans-oceanic systems

Having validated the Stokes model, I will now use it to calculate the outage probability in trans-oceanic systems. The outage probability is defined as the probability that the penalty $\Delta Q^{(m)}$ exceeds a preset value, typically 2.5 dB or 3.0 dB, referred to here as the allowed degradation level. This calculation will allow system designers to determine the required margin for polarization effects more precisely than they have in the past. The outage probability should not be confused with the probability of failure in a given time. To calculate the probability of failure per unit time, it is necessary to know the rate at which the polarization states of the transmission line change and become uncorrelated. This number is not well-known, but it has been estimated that an undersea system will pass through on the order of 10^5–10^6 independent states in a twenty year lifetime [11]. Since the outage probability is a rapidly decreasing function of $\Delta Q^{(m)}$, I will show that there is little ambiguity in simply demanding that the outage probability be less than 10^{-6}.

In the studies that I present in this chapter, I used 10^5 realizations for each set of parameters. I found that when the PDL is small, the distribution of $\Delta Q^{(m)}$ was always well-described by a Gaussian distribution. When estimating outage probabilities of 10^{-6} or less, we used this Gaussian fit.

8.1 Effect of increasing the number of channels

The number of WDM channels in trans-oceanic systems is rapidly increasing. While the effect of PMD on a single channel is typically small, the PMD does change the polarization states of the different channels with respect to one another. In other words, the PMD changes the angular separation of the channels on the Poincaré sphere. As a consequence of the interaction of the PMD and the PDL, different channels will undergo different amounts of loss when they pass through a device with PDL. Since the gain saturation in the amplifiers is tuned to effectively restore the total signal power in all the channels, some channels gain power at the expense of the others. This effect leads to a random walk in the power of each channel and can cause one or more channels to fade. I will show that this mechanism is the primary cause of fading in systems with more than approximately ten channels, in contrast to single-channel systems in which PDG is the primary cause of fading.

To investigate this issue, I considered a system in which the channel spacing and the optical filter bandwidth equaled 0.6 nm. I set the other system parameters as follows: PMD = 0.1 ps/km$^{1/2}$, PDL = 0.0 dB, and PDG = 0.06 dB. Figure 8.1 shows that as the number of channel increases, the importance of PDG decreases as expected from the argument in the preceding paragraph.
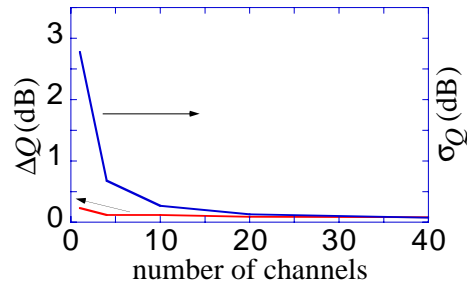
Figure 8.1: The degradation and variance of Q factor as a function of number of channels.

Next, I set the PDG equal to zero, leaving only the effects of PMD and PDL in the model. In this case, I set the channel spacing to 1.0 nm and the optical filter bandwidth to 0.5 nm. I set the other system parameters: PMD $= 0.1$ ps/km$^{1/2}$ and PDL $= 0.1$ dB. Increasing the number of channels, I find that if $\Delta Q_{\text{allowed}}$, the allowed degradation level for any single channel, is set equal to 2.5 dB, then the outage probability dramatically increases from 6.5×10^{-13} in the case of a single channel to 3.0×10^{-4} when there are many channels. With only three channels, the outage probability already exceeds 10^{-5}. If I raise $\Delta Q_{\text{allowed}}$ to 3.0 dB, then the maximum outage probability falls to 2.3×10^{-6}, a decrease of more than two orders of magnitude.
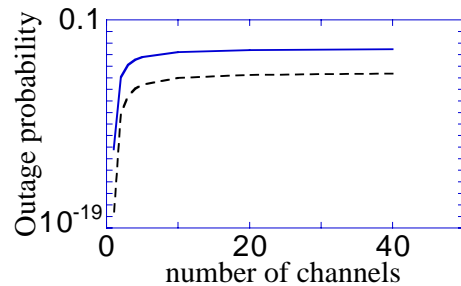
Figure 8.2: Outage probability as a function of number of channels. Solid line is for the 2.5 dB decision level; dashed line is for the 3 dB decision level.

8.2 Effect of Parameters Variations

Varying the channel spacing:

When I reduce the channel spacing from 1 nm to 0.5 nm and keep the other parameters the same, I find that the outage probability becomes smaller as shown in Fig. 8.3. For example, when the number of channels is equal to 40, and I use a 2.5 dB
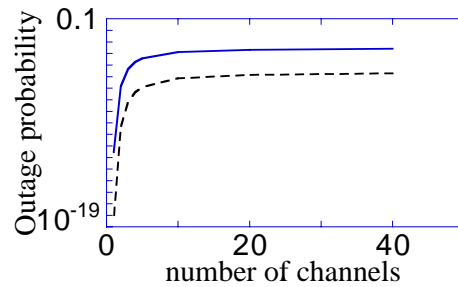
Figure 8.3: Outage probability as a function of the number of channels. Solid line is for the 2.5 dB decision level; dashed line is for the 3 dB decision level.

decision level, the outage probability is 2.6×10^{-4} compared to 3.0×10^{-4} in Fig. 8.2. Hence, a smaller channel spacing reduces the penalty due to PDL.

Increasing the amplifier spacing:

When I increase the amplifier spacing from 33 km to 45 km and then 50 km, I find that the average value of Q decreases due to the additional ASE noise that is added to the total signal. However, the outage probability decreases because the number of PDL elements along the transmission line is reduced, as shown in Fig. 8.4. When the number of channels is 40, the outage probability drops from 3.0×10^{-4} to 1.3×10^{-5}, and 2.8×10^{-6}, respectively. So, when one designs a WDM system and chooses the amplifier spacing, one has to take into account both factors. If the PDL is the same in each amplifier, then a short amplifier spacing will introduce less noise but a higher
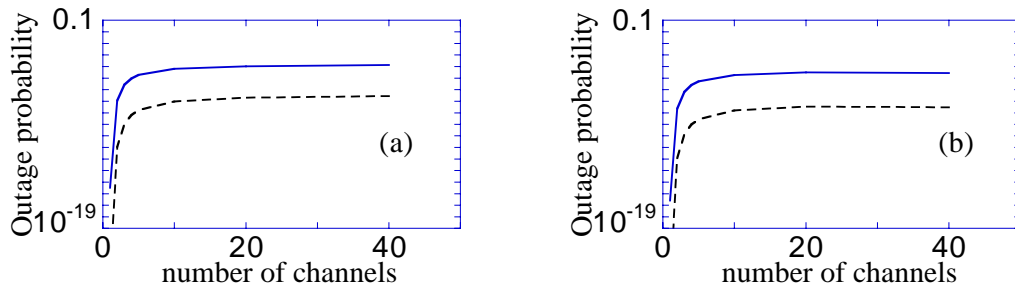
Figure 8.4: Outage probability as a function of the number of channels. Amplifier spacing equals (a) 45 km, (b) 50 km. Solid line is for the 2.5 dB decision level; dashed line is for the 3 dB decision level.

outage probability. By contrast, a long amplifier spacing will introduce more noise but a smaller outage probability.

Increasing the number of PDL elements:

In the previous cases, I assumed that there was only one PDL element in each amplifier, but if there is more than one PDL element, as is possible in EDFAs because there are several discrete elements in front of the amplifying fiber, then I can expect that the PDL is enhanced. For example, if there are two PDL elements in front of the amplifying fiber, I obtain the result shown in Fig. 8.5.a When the PDL in each
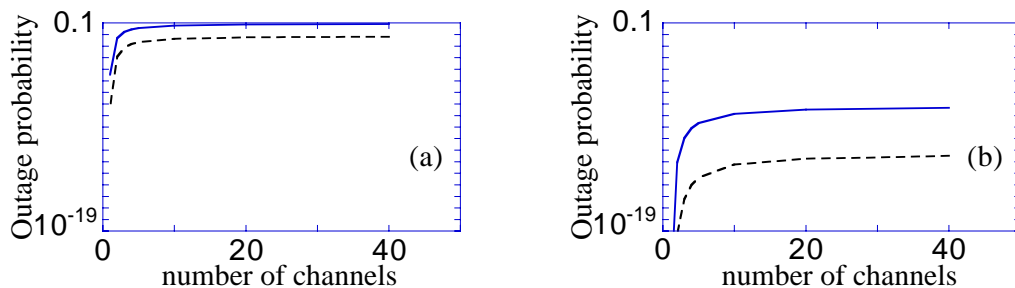
Figure 8.5: Outage probability as a function of the number of channels, with 2 PDL elements, each with (a) PDL = 0.1 dB and (b) PDL = 0.05 dB. Solid line is for the 2.5 dB decision level; dashed line is for the 3 dB decision level.

element is 0.1 dB, I find that the outage probability is nearly 0.1, which is too large to be acceptable in system design. If the PDL in each element is only 0.05 dB, then as shown in Fig. 8.5.b. The outage probability drop rapidly to 4.6×10^{-9} when the number of channels is 40. If one of the PDL elements is in front of the amplifying fiber and the other one is after, the results look the same as when both PDL elements are in front of the amplifying fiber.

When I increase the number of PDL elements in front of the Er-doped fiber to three, each PDL = 0.05 dB, I can see in Fig. 8.6 that the outage probability in a WDM system with 40 channels is just 7.7×10^{-6} which is still far below the value when there is only one PDL element in every amplifier and the PDL is 0.1 dB.
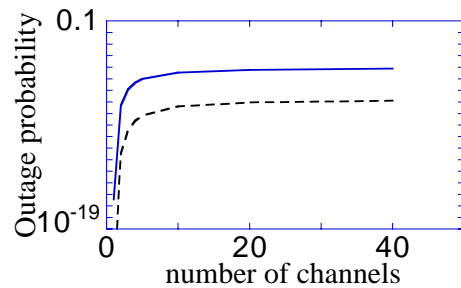
Figure 8.6: Outage probability as a function of the number of channels with 3 PDL elements in each amplifier, PDL = 0.05 dB. Solid line is for the 2.5 dB decision level; dashed line is for the 3 dB decision level.

Consequently, reducing the PDL in each of the elements in an EDFA is an effective way to reduce the outage probability. It might seem surprising at first that an amplifier with 3 PDL elements, each of which has a PDL of 0.05 dB, performs better than an amplifier with 1 PDL element whose PDL is 0.1 dB. The reason is that the PDL of each of the three elements is randomly aligned, reducing its effect.

Adding polarization scrambling:

When I add polarization scrambling to each channel, so that S_1, S_2, and S_3 are all zero, I find that the outage probability is nearly zero even when the number of channels is large, using the same parameter set as in Fig. 8.2.

Neighboring channels orthogonal to each other:

It is increasingly common to reduce the interaction between neighboring channels by alternating the initial polarization state in each channel, so that the polarization states in neighboring channels are orthogonal to each other [33]. Using my model, I find that the outage probability, shown in Fig. 8.7, is nearly the same as in Fig. 8.2 for the same degradation level. Thus, orthogonal injection does not reduce the outage

Figure 8.7: Outage probability as a function of the number of channels. Solid line is for the 2.5 dB decision level, dashed line is for the 3 dB decision level.

probability due to PDL although it raises Q.

Adding PDG:

As I showed in Sec. 8.1, the effect of PDG becomes insignificant when there are more than approximately ten channels in a WDM system. To further investigate this issue, I added a PDG of 0.07 dB to the calculation whose results I show in Fig. 8.2. I show these new results in Fig. 8.8. Instead of a small outage probability when the number of channels is small, as shown in Fig. 8.2, I find the outage probability

becomes large for a small number of channels and then decreases to its final value. The dramatic increase in the outage probability due to PDG when the number of channels is small is due to the faster growth of ASE noise that is induced. The outage probability then decreases as the number of channels becomes larger because the PMD between the channels leads to an averaging of the polarization states so that the DOP for the *total* signal is nearly zero, and the PDG leads to almost no excess ASE growth. When the number of channels equals 40, the outage probability is 2.2×10^{-4} which is actually smaller than the corresponding value of 3.0×10^{-4} when there is no PDG. The reason for this paradoxical decrease in the outage probability is that the PDG tends to compensate for the effects of PDL on channels that experience excess loss.

Figure 8.8: Outage probability as a function of the number of channels, PDG = 0.07 dB. Solid line is for the 2.5 dB decision level; dashed line is for the 3 dB decision level.

Adding PDG when neighboring channels are orthogonal:

In some current WDM systems, designers are using orthogonal polarization injection to reduce the interaction of neighboring channels [33]. I find that if I include PDG in the amplifiers, orthogonal injection does not help too much in reducing the outage probability due to PDG. The outage probability in this case is almost the

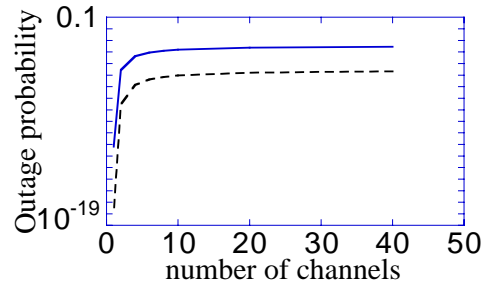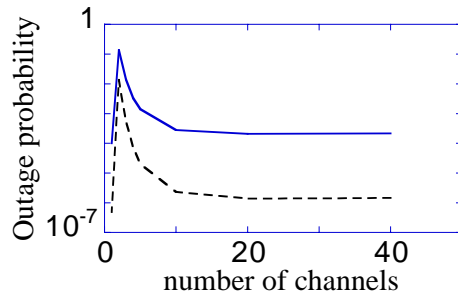same as in Fig. 8.8, as shown in Fig. 8.9.
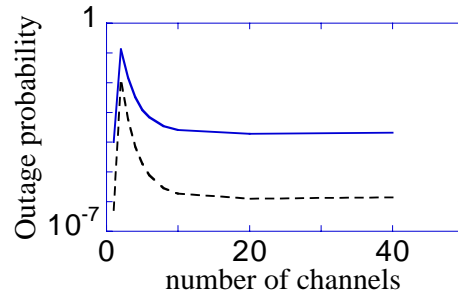
Figure 8.9: Outage probability as a function of the number of channels, PDG = 0.07 dB. Solid line is for the 2.5 dB decision level; dashed line is for the 3 dB decision level.

Increasing the average power in each channel with PDG:

When I keep all the parameters the same as in Fig. (8.8), and I increase the average power in each channel, I find that the outage probability increases when there are ten or fewer channels and decreases when the number of channels is larger, as shown in Fig. 8.10. I found earlier that PDG by itself increases the outage probability with ten or fewer channels while it decreases the outage probability with more than ten channels. Thus, raising the intensity enhances the effect of PDG. To understand this phenomenon, I first recall that the intensity only affects the transmission through gain saturation in the Stokes model. With a larger signal, the ASE noise grows for a longer distance before gain saturation affects its magnitude. Since PDG leads to excess growth of the noise components that are orthogonal to the total signal, the ratio of the ASE noise power orthogonal to the total signal to the ASE noise power parallel to the total signal will be larger when the initial signal is larger. Hence, the effect of PDG is enhanced. Quantitatively, I find that with 40 channels the outage probability decreases from 2.2×10^{-4} when the average power is 0.3 mW, as in Fig. 8.8,

to 1.1×10^{-4} when the average power is 0.5 mW, as in Fig. 8.10.
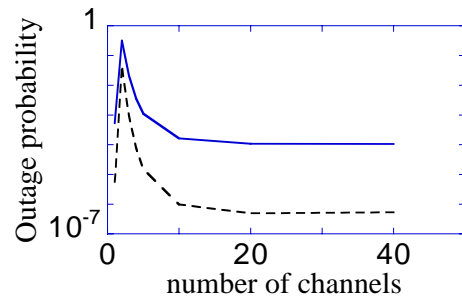
Figure 8.10: Outage probability as a function of the number of channels, PDG = 0.07 dB. The average power is 0.5 mW. Solid line is for the 2.5 dB decision level; dashed line is for the 3 dB decision level.

Chapter 9

Conclusion

Polarization effects — the combination of PMD, PDL, and PDG — can lead to large penalties in trans-oceanic transmission systems. They are likely to also lead to significant penalties in future optical networks in which signals remain in the optical domain for thousands of kilometers. Because the polarization states in optical fibers slowly evolve and because polarization effects can cause an entire channel to fade, it is important to be able to calculate the outage probability in these systems and thus assure that they are designed with adequate margins. It is impossible to use standard simulation techniques to determine the appropriate margin. The full simulation of a WDM system over 9,000 km requires more than 1 hour of CPU time on an SGI Onyx with only eight channels and only one fiber realization. Simulating the approximately 100,000 realizations that are required to obtain a statistical distribution for the degradation of the Q factor due to polarization effects using full simulations is clearly impossible. A different approach is needed.

In this dissertation, I presented a reduced model and validated it to the extent possible. In this approach, one only follows the Stokes parameters for the signal and

for the noise in each channel in a WDM system. In this case, it is possible to study a 40-channel WDM system using 100,000 realizations with 12 hours of CPU time on an SGI Onyx. This study, while lengthy, is certainly feasible. After validating this model, I used it to calculate the outage probability for different amounts of PDL and PDG and with a realistic value of the PMD for different numbers of channels.

I began by presenting mathematical models that describe the polarization effects that appear in optical fiber transmission systems. The PMD is modeled using the coarse step method. In addition to presenting the theory, I also presented simulation results that show that this method accurately reproduces the expected PMD statistics. PMD is an effect that occurs in the optical fibers and causes the polarization states of different wavelength channels to diffuse with respect to one another. An important premise of the reduced model is that the PMD is not so large that it causes the polarization states inside a single channel to diffuse apart. We verified that this assumption is reasonable for modern-day, trans-oceanic communication systems. PDL and PDG are effects that are due to polarization-dependent elements inside ED-FAs. PDL occurs in components, particularly WDM couplers, that are outside the amplifying element, and it can be simply modeled as a loss whose strength depends on the incoming polarization states. By contrast, PDG occurs in the amplifying element and is due to polarization hole burning in the polarization state aligned with the incoming light and is proportional to the DOP. It is slightly more complex model than PDL because the polarization state with maximum gain must be determined self-consistently with the incoming light. It is necessary to keep a separate set of Stokes parameters for both the ASE noise and the signal in each channel because the Stokes parameters for the noise are random variables while the Stokes parameters

for the signal are deterministic. From the Stokes parameters for the signal and the noise, it is possible to calculate an SNR and from that the Q factor for each channel. This Q factor is not meaningful by itself since this model does not take into account the effects of nonlinearity and dispersion. What is meaningful is the degradation ΔQ that occurs when PDL and PDG are added to the system.

After describing the reduced Stokes parameter model, I reviewed the full model which is based on the Manakov-PMD equation and described how to solve it using the coarse step method. I showed analytically and demonstrated explicitly that a PMD value of 0.1 ps/km$^{1/2}$ do not degrade the DOP of individual channels at 10 Gbits/sec per channel. Hence, I concluded that this PMD is too low to lead to mutual diffusion of the polarization states inside a single channel — a basic premise of the Stokes parameter model. This PMD value is on the high end of values that are found in modern-day trans-oceanic systems. Next, I used the Manakov equation to study changes in the DOP due to chromatic dispersion and the Kerr nonlinearity. I showed that for realistic parameters changes in the DOP due to dispersion and nonlinearity are negligible — another basic premise of the Stokes parameter model.

Finally, I used the full model to validate the reduced Stokes parameter model to the extent possible by comparing. It was only possible in practice to run full simulations with 20 realizations for each choice of parameters and up to 8 channels with 64 bits per channel. With this number of realizations and this number of bits per channel, the expected statistical deviations of the calculated values from the true values for the mean degradation $\langle \Delta Q \rangle$ and the standard deviation σ_Q is large. Thus, the results of the full model are no more reliable than the results of the reduced model, and their comparison should be viewed as a consistency check. I compared

the models for one- and eight-channel systems, for different modulation formats, and different values of the PMD, PDL, and PDG. In all the cases that I examined, the results of the full and the reduced models were consistent.

I also validated the reduced Stokes parameter model by comparison to experiments in a recirculating loop. My colleagues and I studied the evolution of the DOP both with a continuous train of solitons and with no initial signal so that the noise is allowed to grow. I obtained excellent agreement between the reduced model and the experiments by choosing PDL = 0.45 dB in the loop. This choice is consistent with the experimentally measured value.

Finally, I used the Stokes model to simulate polarization effects in trans-oceanic WDM systems. I calculated the outage probability with an allowed degradation level of 2.5 dB and 3 dB as a function of the number of WDM channels in the system. I found an important qualitative difference between systems with less than ten channels and systems with more. In the former case, the principal degradation mechanism is excess ASE noise due to PDG. It is possible to mitigate this effect using polarization scrambling, but PDL can lead to repolarization of the transmitted light which then again degrades due to PDG. By contrast, PDG is unimportant in systems with more than ten channels because PMD rapidly randomizes the polarization states of the channels, reducing the total DOP to nearly zero. In this case, the principal degradation mechanism is that the gain-saturated amplifiers only control the total power in the transmission line, but the powers in individual channels can vary. The combination of PMD and PDL leads to different channels undergoing different amounts of loss in the PDL elements. As a result, the power in each channel undergoes a random walk, leading to degradation of some of the channels. I investigated in detail

the effects of changing the channel spacing, increasing the amplifier spacing, varying the PDL, varying the PDG, adding polarization scrambling, setting up the channels so that they are initially orthogonal to their neighbors, and increasing the average power in each channel.

In future work, this model should be carefully compared to data from trans-oceanic systems.

While this model was developed to study trans-oceanic systems, it seems likely that this model will be of use in simulating future optical networks. Since optical fiber communications networks are becoming increasingly complicated, there is an increasing need for effective reduced models. The MONET consortium created a reduced model that is similar to ours, but it does not take into account polarization effects [34]. It only follows the first Stokes parameter, which corresponds to the total power, in each wavelength bin. Other groups are interested in constructing similar models. The theory and the algorithms that I have presented here will be a good starting point for future optical network modeling and design.

Bibliography

[1] E. Lichtman, "Limitations imposed by polarization-dependent gain and loss on all-optical ultralong communication systems," J. Lightwave Technol., vol. 13, pp. 906–913, 1995.

[2] C. D. Poole, R. W. Tkach, A. R. Chraplyvy, and D. A. Fishman, "Fading in lightwave systems due to polarization-mode dispersion," IEEE Photon. Technol. Lett., vol. 3. pp. 68–70, 1991.

[3] N. S. Bergano, "Wavelength division multiplexing in long-haul transmission systems," J. Lightwave Technol., vol. 14, pp. 1299–1308, 1996.

[4] C. R. Menyuk, D. Wang, and A. N. Pilipetskii, "Repolarization of polarization-scrambled optical signals due to polarization dependent loss," IEEE Photon. Technol. Lett., vol. 9, pp. 1247–1249, 1997.

[5] D. Wang and C. R. Menyuk, "Reduced model for the evolution of the polarization states in wavelength-division-multiplexed channels," Opt. Lett., vol. 23, pp. 1677–1679, 1998.

[6] T. Naito, N. Shimojoh, T. Tanaka, H. Nakamoto, M. Doi, T. Ueki, and M. Suyama, "1 Terabit/s WDM transmission over 10,000 km," European Conf. Opt. Comm., paper PD2-1, Nice, France, Sept. 1999.

[7] C. D. Poole and R. E. Wagner,"Phenomenological approach to polarisation dispersion in long single-mode fibers," Electron. Lett., vol. 22, pp. 1029–1030, 1986.

[8] D. Wang, *Nonlinear Optical Loop Mirror Based on Standard Communication Fiber*, M.S. Thesis, UMBC, Aug., 1997.

[9] P. K. A. Wai and C. R. Menyuk, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence," J. Lightwave Technol., vol. 14, pp. 148–157, 1996.

[10] A. E. Amari, N. Gisin, B. Perny, H. Zbinden, and C. W. Zimmer, "Statistical prediction and experimental verification of concatenations of fiber optic components with polarization dependent loss," J. Lightwave Technol., vol. 16, pp. 332–339, 1998.

[11] C. R. Menyuk, D. Wang, and A. N. Pilipettski,"Repolarization of polarization-scrambled optical signals due to polarization dependent loss," Photon. Technol. Lett., vol. 9, pp. 1247–1249, 1997.

[12] J. Schesser, S. M. Abbott, R. L. Easton, and M. S. Stix, "Design requirements for the current generation of undersea cable systems," AT&T Tech. J., vol. 74, pp. 16–32, 1995.

[13] V. J. Mazurczyk and J. L Zyskind, "Polarization hole burning in Erbium-doped fiber amplifiers," Conf. Lasers Electro-Opto., paper CPD26, Baltimore, MD, May 1993.

[14] C. R. Menyuk, "Stability of solitons in birefringent optical fibers. I: equal propagation amplitudes," Opt. Lett., vol. 12, pp. 614–616, 1987.

[15] G. P. Agrawal, *Nonlinear Fiber Optics*. San Diego: Academic Press, 1995.

[16] D. Marcuse, C. R. Menyuk, and P. K. A. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," J. Lightwave Technol., vol. 15, pp. 1735–1746, 1997.

[17] D. Wang and C. R. Menyuk, "A reduced model for the evolution of the polarization states," Opt. Lett., vol. 23, pp. 1677–1679, 1998.

[18] The authors thank D. Chowdhury and J. Chesnoy for providing information about their computational benchmarks.

[19] L. F. Mollenauer, J. P. Gordon, and S. G. Evangelides, "The sliding-frequency guiding filter: An improved form of soliton jitter control," Opt. Lett., vol. 17, pp. 1575–1578, 1992.

[20] P. D. Maker and R. W. Terhune, "Study of optical effects due to an induced polarization third order in the electric field strength," Phys. Rev., vol. 137, pp. A801–A818, 1965.

[21] G. M. Carter, R. M. Mu, V. S. Grigoryan, C. R. Menyuk, T. F. Carruthers, M. L. Dennis, and I. N. Duling III, "Transmission of dispersion-managed solitons at 20 Gbit/s over 20000 km," Electron. Lett., vol. 35, pp. 233–234, 1999.

[22] N. S. Bergano and C. R. Davidson, "Circulating loop transmission experiments for the study of long-haul transmission systems using erbium-doped fiber amplifiers," J. Lightwave Technol., vol. 13, pp. 879–888, 1995.

[23] K. Suzuki, H. Kubota, A. Sahara, and M. Nakazawa, "40 Gbit/s single channel optical soliton transmission over 70,000 km using in-line synchronous modulation and optical filtering," Electron. Lett., vol. 34, pp. 98–99, 1998.

[24] R. Noé, D. Sandel, M. Yoshida-Dierolf, S. Hinz, V. Mirvoda, A. Schöda, C. Glingener, E. Gottwald, C. Scheerer, G. Fischer, T. Weyrauch, and W. Haase, "Polarization mode dispersion compensation at 10, 20, and 40 Gb/s with various optical equalizers," J. Lightwave Technol., vol. 17, pp. 1602–1615, 1999.

[25] C. R. Menyuk and P. K. A. Wai, "Polarization evolution and dispersion in fibers with spatially varying birefringence," J. Opt. Soc. Am. B, vol. 11, pp. 1288–1296, 1994.

[26] P. K. A. Wai and C. R. Menyuk, "Polarization decorrelation in optical fibers with randomly varying birefringence," Opt. Lett., vol. 19, pp. 1517–1519, 1994.

[27] D. Wang, E. A. Golovchenko, A. N. Pilipettsk, M. Arend, and C. R. Menyuk, "The nonlinear optical loop mirror demultiplexer based on standard communication fiber," J. Lightwave Tech., vol. 15, pp. 642–646, 1997.

[28] S. G. Evangelides, Jr., L. F. Mollenauer, J. P. Gordon, and N. S. Bergano, "Polarization multiplexing with solitons," J. Lightwave Technol., vol. 10, pp. 28–35, 1992

[29] J. D. Jackson, *Classical Electrodynamics*. John Wiley & Sons, Inc., 1998.

[30] C. D. Poole and D. L. Favin, "Polarization-mode dispersion measurements based on transmission spectra through a polarizer," J. Lightwave Tech., vol. 12, pp. 917–929, 1994.

[31] C. D. Poole, J. H. Winters, and J. A. Nagel, "Dynamical equation for polarization dispersion," Opt. Lett., vol. 16, pp. 372–374, 1991.

[32] See, *e.g.*, R.-M. Mu, V. S. Grigoryan, C. R. Menyuk, E. A. Golovchenko, and A. N. Pilipetskii, "Timing-jitter reduction in a dispersion-managed soliton system," Opt. Lett., vol. 12, pp. 930–932, 1998.

[33] N. S. Bergano, C. R. Davidson, M. Ma, A. Pilipetskii, S. G. Evangelides, H. D. Kidorf, J. M. Darcie, E. Golovchenko, K. Rottwitt, P. C. Corbett, R. Menges, M. A. Mills, B. Pedersen, D. Peckham, A. A. Abramov, and A. M. Vengsarkar, "320 Gb/s WDM transmission (64×5 Gb/s) over 7,200 km using large mode fiber spans and chirped return-to-zero signals," Opt. Fiber Comm. Conf., paper PD12-1, San Jose, Feb. CA, 1998.

[34] R. E. Wagner, R. C. Alferness, A. A. M. Saleh, and M. S. Goodman, "MONET: Multiwavelength optical networking," J. Lightwave Technol., vol. 14, pp. 1349–1355, 1996.

[35] L Arnold, *Stochastic Differential Equations: Theory and Applications*. New York: Wiely, 1974.