APPROVAL SHEET

Title of Thesis: Investigation of the performance degradation due to polarization effects in optical fiber communications systems

Name of Candidate: Ivan T. Lima, Jr.
Doctor of Philosophy, 2003

Dissertation and Abstract Approved: _____
Professor Curtis R. Menyuk
Computer Science and Electrical Engineering

Date Approved: _____

Curriculum Vitae

Ivan T. Lima Jr.
1411 Centennial Blvd., ECE 101
Fargo, ND 58105-5285
E-mail: Ivan.Lima@ndsu.nodak.edu

December 2, 2003

Birth:
Juazeiro, Bahia, Brazil, in October 16, 1971.

Education:

Ph.D. Electrical Engineering: Photonics
University of Maryland, Baltimore County, December 2003
Dissertation: Investigation of the performance degradation due to polarization effects in optical fiber communications systems
Advisor: Dr. Curtis R. Menyuk

M.Sc. Electrical Engineering: Electronics and Communications
State University of Campinas, Brazil, March 1998
Thesis: Analysis of the frequency selectivity of two-dimensional periodic dielectric gratings
Advisor: Dr. Attilio J. Giarola

B.Sc. Electrical Engineering: Electronics and Communications
Federal University of Bahia, Brazil, December 1995

Employment:

8/2003–Present Assistant Professor, Department of Electrical and Computer Engineering, North Dakota State University, Fargo, ND
Teaching, laboratory and course developement, research, advising, and service.

8/1998–8/2003 Research Assistant, Computer Science and Electrical Engineering
University of Maryland Baltimore County, Baltimore, MD
Research novel techniques for the mitigation of the effects that limit the capacity of single channel and wavelength-division-multiplexed optical fiber communications systems.

12/1994–02/1996	Advisor, State Superintendance of Bahia, Bank of Brazil (Banco do Brasil) Analysis and development of database software. From 1993 to 1996 was under the direct supervision of the Presidency of the Bank. On leave from 03/1996 to 01/1997 to pursue graduate education.
6/1992–11/1994	Assistant, State Superintendance of Bahia, Bank of Brazil Analysis and development of database software.
6/1990–5/1992	Computer Programmer, Data Processing Center of Salvador, Bahia, Bank of Brazil Analysis and development of computer software.
10/1989–5/1990	Administrative Assistant, Central Branch of Salvador, Bahia, Bank of Brazil
10/1986–9/1989	Office Aid Minor, Branch of Araguaina, Tocantins, Bank of Brazil

Consulting:

5/2000	PhotonEx Corporation, Maynard, Massachusetts
12/2000	KDDI Corporation, Japan

Awards:

10/2003 – LEOS Graduate Student Fellowship Award from the Lasers & Electro-Optics Society of the Institute of Electrical and Electronics Engineers. The fellowship award was presented at the LEOS Annual Meeting (LEOS 2003).

9/2003 – Venice Summer School 2002 Award for paper (co-author) that appeared in the proceedings of the 29^{th} European Conference on Optical Communication (ECOC 2003). The award was presented at ECOC 2003.

4/1998 – Graduate Scholarship at Ph.D. degree level from the National Research Council (CNPq) of the Brazilian Ministry of Science and Technology for graduate education at the University of Maryland Baltimore County, Baltimore, MD, USA. Period: 9/1998 to 8/2000.

6/1996 – Graduate Scholarship at M.Sc. degree level from CAPES of the Brazilian Ministry of Education for graduate education at the State University of Campinas, Campinas, Brazil. Period: 3/1996 to 3/1998.

Languages:

— English: Read, write, and speak fluently.
— Portuguese: Read, write, and speak fluently.
— Spanish: Read and speak.

Computer skills:

— Languages: C++, Fortran 77, Matlab, Maple, dbase III, Clipper, Assembly Z-80, BASIC.
— Applications: Latex 2ε, Gnuplot, Acrobat, Ghostview, PowerPoint.
— Operating Systems: Linux, and other UNIX flavors, Windows, DOS, Palm.

Professional Societies:
— Lasers & Electro-Optics Society (LEOS) of the Institute of Electrical and Electronics Engineers (IEEE)
— Optical Society of America (OSA)

Archival journal publications:

1) I. T. Lima, Jr., A. O. Lima, G. Biondini, C. R. Menyuk, and W. L. Kath, "A comparative study of single-section polarization-mode dispersion compensators," *IEEE/OSA Journal of Lightwave Technology*, Special issue-PMD in May 2004.

2) Y. Sun, I. T. Lima, Jr., A. O. Lima, H. Jiao, J. Zweck, L. Yan, C. R. Menyuk, and G. M. Carter, "System performance variations due to partially polarized noise in a receiver," *IEEE Photonics Technology Letters*, Vol. 15, No. 11, pp. 1648–1650, November 2003.

3) J. Zweck, I. T. Lima, Jr., Yu Sun, A. O. Lima, C. R. Menyuk, and G. M. Carter, "Modeling receivers in optical communication systems with polarization effects," *Optics & Photonics News*, Vol. 14, No. 11, pp. 30–35, November 20003.

4) I. T. Lima, Jr., A. O. Lima, J. Zweck, and C. R. Menyuk, "Performance characterization of chirped return-to-zero modulation format using an accurate receiver model," *IEEE Photonics Technology Letters*, Vol. 15, No. 4, pp. 608–610, April 2003.

5) A. O. Lima, I. T. Lima, Jr., C. R. Menyuk, G. Biondini, and W. L. Kath, "Statistical analysis of the performance of PMD compensators using multiple importance sampling," to appear in the December 2003 issue of *IEEE Photonics Technology Letters*.

6) Y. Sun, A. O. Lima, I. T. Lima, Jr., J. Zweck, L. Yan, C. R. Menyuk, and G. M. Carter, "Statistics of the system performance in scrambled recirculating loop with PDL and PDG," *IEEE Photonics Technology Letters*, Vol. 15, No. 8, pp. 1067–1069, August 2003.

7) I. T. Lima, Jr., A. O. Lima, J. Zweck, and C. R. Menyuk, "Efficient computation of outage probabilities due to polarization effects in a WDM system using a reduced Stokes model and importance sampling," *IEEE Photonics Technology Letters*, Vol. 15, No. 1, pp. 45–47, January 2003.

8) A. O. Lima, I. T. Lima, Jr., C. R. Menyuk, and T. Adali, "Comparison of penalties resulting from first-order and all-order polarization mode dispersion in optical fiber transmission systems," *Optics Letters*, Vol. 28, No. 5, pp. 310–311, March 2003.

9) B. S. Marks, I. T. Lima, Jr., and C. R. Menyuk, "Autocorrelation function for PMD emulators with rotators," *Optics Letters*, Vol. 27, No. 13, July 2002.

10) I. T. Lima, Jr., G. Biondini, B. S. Marks, W. L. Kath, and C. R. Menyuk, "Analysis of PMD compensators with fixed DGD using importance sampling," *IEEE Photonics Technology Letters*, Vol. 14, No. 5, pp. 627–629, May 2002.

11) A. O. Lima, I. T. Lima, Jr., T. Adali and C. R. Menyuk, "A Novel Polarization Diversity Receiver for PMD Mitigation," *IEEE Photonics Technology Letters*, Vol. 14, No. 4, pp. 465–467, April 2002.

12) I. T. Lima, Jr., R. Khosravani, P. Ebrahimi, E. Ibragimov, A. E. Willner, and C. R. Menyuk, "Comparison of Polarization-Mode Dispersion Emulators," *IEEE/OSA Journal of Lightwave Technology*, Vol. 19, No. 12, pp. 1872–1881, December 2001.

13) C. R. Menyuk, R. Holzlöhner, and I. T. Lima, Jr., "Advances in Modeling of Optical Fiber Transmission Systems," *IEEE LEOS Newsletter*, pp. 21–23, October 2001.

14) Y. Sun, I. T. Lima, Jr., H. Jiao, J. Wen, H. Xu, H. Ereifej, C. R. Menyuk, and G. Carter, "Study of System Performance in a 107 km Dispersion Managed Recirculating Loop Due to Polarization Effects," *IEEE Photonics Technology Letters*, Vol. 13, No. 9, pp. 966–968, September 2001.

15) R. Khosravani, I. T. Lima, Jr., P. Ebrahimi, E. Ibragimov, A. E. Willner, and C. R. Menyuk, "Time and Frequency Domain Characteristics of Polarization-Mode Dispersion Emulators," *IEEE Photonics Technology Letters*, Vol. 13, No. 2, pp. 127–129, February 2001.

16) I. T. Lima, Jr. and A. J. Giarola, "Frequency Selective Properties of Two-Dimensional Dielectric Gratings: TE and TM Polarizations," *International Journal of Infrared and Millimeter Waves*, Vol. 21, No. 3, pp. 447–459, March 2000.

17) G. V. Grigoryan, I. T. Lima, Jr., T. Yu, V. S. Grigoryan, and C. R. Menyuk, "Using Color to Understand Light Transmission," *Optics & Photonics News*, Vol. 11, No. 8, pp. 44–50, August 2000.

18) I. T. Lima, Jr. and A. J. Giarola,"An Integral Equation Analysis of Two-Dimensional Dielectric Gratings," *Microwave and Optical Technology Letters*, Vol. 20, No. 5, pp. 329–333, March 1999.

Book chapter:

1) C. R. Menyuk, B. S. Marks, I. T. Lima, Jr., J. Zweck, Y. Sun, G. M. Carter, and D. Wang "Polarization effects in long-haul undersea systems", in *Undersea Fibre Communication Systems*, José Chesnoy, ed., Academic Press: San Diego, CA, 2002, pp. 269–305.

Invited papers at conferences:

1) I. T. Lima, Jr., G. Biondini, B. S. Marks, W. L. Kath, and C. R. Menyuk, "Optimization of a PMD Compensator with Constant Differential Group Delay Using Importance Sampling," (invited), *in Proceedings of the Conference on Laser and Electro-Optics (CLEO) 2001*, paper CFE3, Baltimore, Maryland, USA, May 6–11, 2001.

2) C. R. Menyuk, R. Holzlöhner, I. T. Lima, Jr., B. S. Marks, and J. Zweck, "Advances in modeling high data rate optical fiber communication systems," (invited) *SIAM Conference on Computational Science and Engineering (CSE03)*, San Diego, CA, 2003.

3) J. Zweck, I. T. Lima, Jr., R. Holzlöhner, C. R. Menyuk, "New advances in modeling optical fiber communication systems", (invited), *in Proceedings of the Integrated Photonics Research OSA Topical Meeting and Exhibit*, paper IThB1, Vancouver, Canada,

july 17–19, 2002.

4) W. L. Kath, G. Biondini, I. T. Lima, Jr., B. S. Marks and C. R. Menyuk, "Calculations of outage probabilities due to PMD using importance sampling," (invited) *in Proceedings of the LEOS Annual Meeting 2001*, paper WAA2, San Diego, California, USA, November 14–15, 2001.

5) C. R. Menyuk, R. Holzlöhner, and I. T. Lima, Jr., "New approaches for modeling high data rate optical fiber communication systems," (invited) *OSA Annual Meeting 2001*, paper ThFF1, Long Beach, California, USA, October 14–18, 2001.

6) C. R. Menyuk, R. Holzlöhner, and I. T. Lima, Jr., "Advances in Modeling of Optical Fiber Transmission Systems," (invited) *in Proceedings of the IEEE LEOS Summer Topical Meeting 2001*, paper MD1.2, Copper Mountain, Colorado, USA, July 30–August 1, 2001.

Contributed papers at conferences:

1) I. T. Lima, Jr. and A. O. Lima, "Computation of the probability of power penalty and Q-penalty outages due to PMD," *in Proceedings of the LEOS Annual meeting 2003*, paper TuQ1, Tucson, Arizona, USA, October 26–30, 2003.

2) A. O. Lima, I. T. Lima, Jr., J. Zweck, and C. R. Menyuk, "Efficient computation of PMD-induced penalties using Multicanonical Monte Carlo simulations," *in Proceedings of the 29th European Conference on Optical Communication (ECOC) 2003*, Rimini, Italy, September 21–25, 2003.

3) I. T. Lima, Jr., L.-S. Yan, B. S. Marks, C. R. Menyuk, and A. E. Willner, "Experimental verification of the penalty produced by the polarization effects in fiber recirculating loops," *in Proceedings of the Conference on Lasers and Electro Optics (CLEO) 2003*, paper CThD2, Baltimore, Maryland, USA, June 1–6, 2003.

4) I. T. Lima, Jr., A. O. Lima, J. Zweck, and C. R. Menyuk, "An accurate formula for the Q-factor of a fiber transmission system with partially polarized noise," *in Proceedings of the Conference on Lasers and Electro Optics (CLEO) 2003*, paper CThJ2, Baltimore, Maryland, USA, June 1–6, 2003.

5) H. Jiao, I. T. Lima, Jr., A. O. Lima, Y. Sun, J. Zweck, L. Yan, C. R. Menyuk, and G. M. Carter, "Experimental validation of an accurate receiver model for systems with unpolarized noise," *in Proceedings of the Conference on Lasers and Electro Optics (CLEO) 2003*, paper CThJ1, Baltimore, Maryland, USA, June 1–6, 2003.

6) S. E. Minkoff, J. Zweck, A. O. Lima, I. T. Lima, Jr., and C. R. Menyuk, "Numerical Simulation and Analysis of Fiber Optic Compensators," *Society for Industrial and Applied Mathematics (SIAM) Annual Meeting*, Montreal, Canada, June 16-20, 2003.

7) I. T. Lima, Jr., A. O. Lima, J. Zweck, and C. R. Menyuk, "Computation of the Q-factor in optical fiber systems using an accurate receiver model," *in Proceedings of the Optical Fiber Communication Conference and Exposition (OFC) 2003*, paper MF81, Atlanta, Georgia, U.S.A, March 23–28, 2003.

8) A. O. Lima, I. T. Lima, Jr., B. S. Marks, C. R. Menyuk, and W. L. Kath, "Performance analysis of single-section PMD compensators using multiple importance sampling," *in Proceedings of the Optical Fiber Communication Conference and Exposition (OFC) 2003*, paper ThA3, Atlanta, Georgia, USA, March 23–28, 2003.

9) Y. Sun, I. T. Lima, Jr., A. O. Lima, H. Jiao, J. Zweck, L. Yan, C. R. Menyuk, and G. M. Carter, "Effects of partially polarized noise in a receiver," *in Proceedings of the Optical Fiber Communication Conference and Exposition (OFC) 2003*, paper MF82, Atlanta, Georgia, USA, March 23–28, 2003.

10) Y. Sun, A. O. Lima, I. T. Lima, Jr., L. Yan, J. Zweck, C. R. Menyuk, and G. M. Carter, "Accurate Q-factor distributions in optical transmission systems with polarization effects," *in Proceedings of the Optical Fiber Communication Conference and Exposition (OFC) 2003*, paper ThJ4, Atlanta, Georgia, USA, March 23–28, 2003.

11) J. Zweck, S. E. Minkoff, A. O. Lima, I. T. Lima, Jr., and C. R. Menyuk, "A comparative study of feedback controller sensitivity to all orders of PMD for a fixed DGD compensator," *in Proceedings of the Optical Fiber Communication Conference and Exposition (OFC) 2003*, paper ThY2, Atlanta, Georgia, USA, March 23–28, 2003.

12) I. T. Lima, Jr., A. O. Lima, J. Zweck, and C. R. Menyuk, "Computation of the penalty due to the polarization effects in a wavelength-division multiplexed system using a reduced Stokes model with a realistic receiver," *Venice Summer School on Polarization Mode Dispersion (VSS) 2002*, Venice, Italy, June 24–26, 2002.

13) A. O. Lima, I. T. Lima, Jr., T. Adali, and C. R. Menyuk, "Compensation of polarization mode dispersion in optical fiber transmission systems using a polarization diversity receiver," *Venice Summer School on Polarization Mode Dispersion (VSS) 2002*, Venice, Italy, June 24–26, 2002.

14) A. O. Lima, I. T. Lima, Jr., T. Adali, and C. R. Menyuk, "Comparison of power penalties due to first- and all-order PMD distortions," *in Proceedings of the 28th Euro-*

pean Conference on Optical Communication (ECOC) 2002, paper 7.1.2, Copenhagen, Denmark, September 8–12, 2002.

15) I. T. Lima, Jr., A. O. Lima, Y. Sun, J. Zweck, B. S. Marks, G. M. Carter, and C. R. Menyuk, "Computation of the outage probability due to the polarization effects using importance sampling," *in Proceedings of the Optical Fiber Communication Conference and Exhibit (OFC) 2002*, paper TuI7, Anaheim, California, USA, March 17–22, 2002.

16) B. S. Marks, I. T. Lima, Jr., and C. R. Menyuk, "Autocorrelation function for PMD emulators with rotators," *in Proceedings of the Conference on Laser and Electro-Optics (CLEO) 2002*, paper CWH5, Long Beach, California, USA, May 20–23, 2002.

17) A. O. Lima, T. Adali, I. T. Lima, Jr., and C. R. Menyuk, "Polarization diversity and equalization for PMD mitigation," *in Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP) 2002*, Orlando, Florida, USA, May 13–17, 2002.

18) A. O. Lima, T. Adali, I. T. Lima, Jr., and C. R. Menyuk, "Polarization diversity receiver for PMD mitigation," *in Proceedings of the Optical Fiber Communication Conference and Exhibit (OFC) 2002*, paper WI7, Anaheim, California, USA, March 17–22, 2002.

19) Y. Sun, B. S. Marks, I. T. Lima, Jr., K. Allen, G. M. Carter, and C. R. Menyuk, "Polarization state evolution in recirculating loops," *in Proceedings of the Optical Fiber Communication Conference and Exhibit (OFC) 2002*, paper ThI4, Anaheim, California, USA, March 17–22, 2002.

20) G. Biondini, W. L. Kath, I. T. Lima, Jr., and C. R. Menyuk, "A simulation technique for rare polarization mode dispersion events," *OSA Annual Meeting 2001*, paper ThNN1, Long Beach, California, USA, October 14–18, 2001.

21) I. T. Lima, Jr., G. Biondini, B. S. Marks, W. L. Kath, and C. R. Menyuk, "Analysis of Polarization Mode Dispersion Compensators Using Importance Sampling", *in Proceedings of the Optical Fiber Communication Conference and Exhibit (OFC) 2001*, paper MO4, Anaheim, California, USA, March 17–22, 2001.

22) A. O. Lima, I. T. Lima, Jr., T. Adali, and C. R. Menyuk, "PMD Mitigation Using Diversity Detection," *in Proceedings of the IEEE LEOS Summer Topical Meeting 2001*, paper MD3.3, Copper Mountain, Colorado, USA, July 30–1 August 2001.

23) Y. Sun, I. T. Lima, Jr., H. Jiao, J. Wen, H. Xu, H. Ereifej, C. R. Menyuk, and G. Carter, "Variation of System Performance in a 107 km Dispersion Managed Recirculating Loop Due to Polarization Effects," *in Proceedings of the Conference on Lasers and Electro-Optics (CLEO) 2001*, paper CFE4, Baltimore, Maryland, USA, May 6–11, 2001.

24) I. T. Lima, Jr., R. Khosravani, P. Ebrahimi, E. Ibragimov, A. E. Willner, and C. R. Menyuk, "Polarization Mode Dispersion Emulator," *in Proceedings of the Optical Fiber Communication Conference and Exhibit (OFC) 2000*, paper ThB4, Baltimore, Maryland, USA, March 5–10, 2000.

25) C. R. Menyuk, D. Wang, R. Holzlöhner, I. T. Lima, Jr., E. Ibragimov, and V. S. Grigoryan, "Polarization Mode Dispersion in Optical Transmission systems," (tutorial), *in Proceedings of the Optical Fiber Communication Conference and Exhibit (OFC) 2000*, Baltimore, Maryland, USA, March 5–10, 2000.

26) I. T. Lima, Jr. and A. J. Giarola, "Frequency Selective Properties of Arrays of Rectangular Dielectric Waveguides," *in Proceedings of the VIII Brazilian Symposium of Microwave and Optoelectronics*, pp. 330–334, Joinville, SC, Brazil, July 13–15, 1998.

27) I. T. Lima, Jr. and A. J. Giarola, "Electromagnetic Wave Propagation in a Periodic Array of Rectangular Dielectric Waveguides," *in Proceedings of the 1997 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference*, pp. 413–418, Natal-RN, Brazil, August 11–14, 1997.

28) I. T. Lima, Jr. and A. J. Giarola, "Analysis of Two-Dimensional Dielectric Gratings for the Design of Dichroic Structures," *in Proceedings of the 22nd International Conference on Infrared and Millimeter Wave*, p. 342, Wintergreen, Virginia, USA, July 20–25, 1997.

29) I. T. Lima, Jr. and A. J. Giarola, "Electromagnetic Wave Propagation in Two Dimensional Anisotropic Dielectric Gratings," *in Proceedings of the 1997 IEEE-AP-S International Symposium*, pp. 2400–2403, Montreal, Canada, July 13–18, 1997.

Patents and other intellectual properties:

1) R. Khosravani, P. Ebrahimi, I. T. Lima, Jr., E. Ibragimov, A. E. Willner, and C. R. Menyuk, "Polarization Mode Dispersion Emulator," Patend US 6,542,650 B2, issued on April 1, 2003. Licensed to Phaeton Communications, Fremont, CA.

2) I. T. Lima, Jr., and C. R. Menyuk, "Optical Communication Systems Simulator," Software licensed to Science Applications International Corporation, McLean, Virginia.

Abstract

Title of Dissertation: Investigation of the performance degradation due to polarization effects in optical fiber communications systems

Ivan T. Lima, Jr., Doctor of Philosophy, 2003

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

The polarization effects of polarization mode dispersion, polarization-dependent loss, and polarization-dependent gain can strongly impact the performance of optical fiber communications systems. These effects, combined with gain saturation in the amplifiers, couple many channels together in a wavelength-division-multiplexed system. Thus, when investigating the performance degradation theoretically, it is not currently feasible to use full computer simulations. One must use a reduced model, such as the reduced Stokes model that was developed earlier by Wang and Menyuk. In this Ph.D. dissertation, I report two contributions that I made with the collaboration of my colleagues to the field of optical fiber systems modeling. These contributions

can be used in combination with the reduced Stokes model to accurately and efficiently compute the penalty produced by the polarization effects in optical fiber systems. First, my colleagues and I developed an importance sampling technique to accurately determine the probability density functions for the Q-penalty and from that, the outage probability. The results differ significantly from those previously obtained using Gaussian extrapolations. Second, we developed an accurate receiver model. Prior work used an *ad hoc* model, while my work introduces a model based on a calculation of the first- and second-order moments of the current probability density functions with realistic optical and electrical filters. In addition, the receiver model that we developed properly accounts for the effect of partially polarized noise in the performance of the system, while previous work did not. We validated the reduced Stokes model in combination with the receiver model by comparison to full time-domain simulations. To do so, we introduced a new validation procedure that ensured that we were comparing the same fiber realizations in both the full and the reduced models. We also validated the results experimentally at the University of Maryland Baltimore County. The validation work shows a good agreement with the models.

Investigation of the performance degradation due to polarization effects in optical fiber communications systems

by

Ivan T. Lima, Jr.

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
2003

Dedication

To my wife Aurenice, my parents Ivan and Eliza, and my brothers Luciano and Robertinho

Acknowledgements

I wish to express my gratitude to Dr. Curtis Menyuk, my dissertation advisor, for his contributions to my graduate education, for his continuous support throughout my graduate education at UMBC, and for having granted me the opportunity to identify the research projects that I carried out for my Ph.D. dissertation.

I am grateful to Dr. Gary Carter, Dr. Alan Willner, Dr. William Kath, and Dr. Gino Biondini for giving to me the opportunity to collaborate with their research groups in cutting-edge research projects in photonics.

I am grateful to all my colleagues and friends at UMBC: Aurenice Lima, Dr. Edem Ibragimov, Dr. Vladimir Grigoryan, Dr. John Zweck, Dr. Brian Marks, Dr. Li Yan, Lyn Randers, Dr. Tao Yu, Yu Sun, Dr. Ding Wang, Dr. Ruo-mei Mu, Hai Xu, Dr. Ronald Holzlöhner, Oleg Sinkin, Hua Jiao, Jiping Wen, Jonathan Hu, Anshul Kalra, and Aniweta Onuorah, who kindly shared their knowledge with me. I am particularly grateful to Dr. John Zweck for his helpful comments on preliminary versions of my Ph.D. dissertation.

I would like to thank the National Research Council of Brazil (CNPq) for partial financial support during the first two years of my graduate education at UMBC.

I would like to express my immense gratitude to my wife Aurenice, my parents Ivan and Eliza, and my brothers Luciano and Robertinho, for having unconditionally supported me throughout the long and difficult journey of my education. I dedicated this dissertation to them.

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter 1

Introduction

The rapid advance of optical fiber communications systems has revolutionized telecommunications on a global scale and has played a major role in the advent of the information era. The invention of the laser diode in the 1960s [1], [2], low loss optical fiber in the 1970s [3], and the erbium-doped fiber amplifier in the 1980s [4] allowed engineers to develop long-haul optical fiber communications systems. Terrestrial and trans-oceanic optical fiber systems currently transport the great majority of inter-metropolitan and international voice and data traffic in the world. Theoretical capacities in optical fiber communication systems are tens of THz. Nonetheless, four major impairments—nonlinearity, chromatic dispersion, polarization effects, and amplified-spontaneous-emission noise—limit the capacities and transmission distances of optical fiber communications systems [5]. Among these effects, polarization effects are the only ones that can produce randomly varying signal distortions and power penalties that can lead to system outages on a time scale that varies between milliseconds and hours [6]. A penalty in an optical fiber communications system is a performance degradation that increases the probability of error in the decoded signal.

This dissertation is concerned with investigating polarization effects, as well as their interactions with other transmission and receiver effects and gain saturation in the amplifiers.

The three principal polarization effects are polarization mode dispersion (PMD), polarization-dependent loss (PDL), and polarization-dependent gain (PDG) [7], [8]. PMD is primarily due to the randomly varying birefringence in optical fibers [9], although birefringent components can make an important contribution in some systems [10]. The beat length due to the birefringence in fibers and the autocorrelation length of the birefringence variations are on the order of tens of meters. As a consequence, the polarization state at any frequency changes on a scale of tens of meters [11]. This variation is rapid compared to typical dispersive and nonlinear length scales, which vary from tens to thousands of kilometers. By itself, this variation has no deleterious effect on communications systems. However, neighboring frequencies in the signal undergo slightly different variations, and, as a consequence, two frequencies that are initially in the same polarization state will ultimately drift apart in polarization. The rate at which this drift occurs is proportional to the frequency separation. Hence, frequencies that are far apart drift apart in polarization state faster. In particular, PMD will cause the polarization states of different channels in a wavelength-division-multiplexed (WDM) system to drift apart long before the frequencies inside a single channel drift apart. We refer to the drift of the polarization states of different WDM channels as inter-channel PMD. When PMD is large enough to depolarize the frequency components within an individual channel, it can lead to waveform distortions. We refer to the polarization drift of the frequencies inside a

single channel as intra-channel PMD.

There has been considerable interest in intra-channel PMD in the optical communications research community in recent years because intra-channel PMD can be a major source of impairments in systems with a per-channel data rate of 10 Gbit/s and higher [12]. However, even in systems with low to moderate PMD, in which intra-channel PMD can be ignored, inter-channel PMD can still play a very important role in WDM systems. In this dissertation, I will be almost exclusively concerned with systems in which intra-channel PMD can be neglected. However, in my studies of WDM systems, inter-channel PMD must be taken into account.

The second polarization effect that I will consider, PDL, is primarily due to the polarization sensitivity of amplifier components—notably the isolators and WDM couplers [7]. While it is simple to model, its effect can be subtle. In some recirculating loop systems, it can lead to partial or total polarization of the noise that artificially improves the performance of the recirculating loop relative to what would be observed in a straight-line system. This issue will emerge at several points in this Ph.D. dissertation.

PDG, the third and final polarization effect that I will consider, is due to polarization hole burning in the erbium-doped fiber amplifiers (EDFAs) that are commonly used in optical fiber communications systems. The gain in EDFAs is slightly inhomogeneous, and, as a consequence, the gain in the polarization state orthogonal to the signal is slightly higher than in the polarization state of the signal when the EDFA is saturated. In single-channel systems, this effect leads to exponential growth of the noise [13]. This effect can be mitigated by scrambling the polarization state of

the signal to reduce the degree-of-polarization (DOP) to nearly zero [14]. In WDM systems, this effect is typically not important, since inter-channel PMD ensures that the DOP of the system is close to zero. If one interleaves channels in orthogonal polarization states at the transmitter, then the DOP will be close to zero throughout the transmission, and the role of PDG becomes negligible as the number of channels increases. In an example considered by Wang and Menyuk [8] of a 10 Gbit/s channel system with a 50 Gbit/s channel spacing, the authors found that PDG is negligible when the system has more than 10 channels. In this dissertation, I will also take PDG into account, even though I found that PDG never plays a significant role for the problems I considered.

Another effect that one must take into account is gain saturation in the amplifiers. Amplifiers typically operate in a gain saturated regime because, by doing so, they damp out power fluctuations [5]. However, since we assume that the gain in EDFAs is nearly constant with frequency, all channels are equally affected by changes in the gain. As a consequence, all the channels in a WDM system are coupled together. Schematically, if some channels suffer excess loss due to PDL in a component, then the total power out of that EDFA will be somewhat diminished and in the next EDFA the gain will be somewhat higher since this EDFA will be slightly less saturated. All channels except the ones that suffered excess loss will emerge from this EDFA with slightly more power than from the previous EDFA. In effect, some channels have transferred some of their power to the other channels.

Modeling the effect of gain saturation in a WDM system with many channels is a significant computational challenge because all of the channels interact. This situa-

tion is quite different from modeling the Kerr effect, which only links together nearby channels [5]. It is not feasible to model such systems using full time-domain simulations. One typically wants to model the system with many thousands of realizations to obtain the full probability density function (pdf) of the penalties produced by the polarization effects. It is not currently feasible to calculate this number of realizations with full time-domain simulations.

For these reasons, Wang and Menyuk introduced a reduced Stokes model that follows four signal Stokes parameters and four noise Stokes parameters for each channel [8]. This model takes into account the phenomena of inter-channel PMD, PDL, and PDG, as well as gain saturation in the amplifiers. The computation time is a small fraction of what is required for full time-domain simulations, so that the calculations described in the previous paragraph become feasible. This reduced model is based on the assumption that PMD is moderate, so that intra-channel PMD can be neglected. It is also based on the assumption that PMD, PDL, and PDG in combination with gain saturation are slow time effects compared to the bit time, so that they affect all bits equally, and nonlinearity and chromatic dispersion can be ignored. Due to this difference in time scales, it is assumed that the penalty due to the combination of PMD, PDL, and PDG can be calculated separately from the penalty due to chromatic dispersion and noise, and that these penalties can be added. Wang and Menyuk carried out studies in which they found the limits of validity for these assumptions, and they validated the reduced model by comparison to full time-domain simulations in the limit where the reduced model applies. All current commercial systems that are not limited by intra-channel PMD fall within this limit. I note,

however, that the validation carried out by Wang and Menyuk used a receiver model that did not correctly account for the signal-dependent noise—the beating between the signal and the noise in photodetectors—in direct detection receivers. Nonetheless, I showed that the original assumptions made by Wang and Menyuk are true when I employ an accurate receiver model that I developed, which accurately accounts for signal-dependent noise at the receiver.

In this Ph.D. dissertation, I describe two contributions that I made with the collaboration of my colleagues to the field of optical fiber systems modeling. These contributions are used to extend and to significantly improve the accuracy of previous work by Wang and Menyuk, as I show in this dissertation. In the first contribution, I developed an importance sampling technique to accurately calculate the outage probability due to polarization effects. The outage probability is the probability that the penalties due to polarization effects will exceed a specified margin—typically 1 dB or 2 dB. This probability is usually required to be very low—typically 10^{-5} or 10^{-6}. As a consequence, a direct calculation of the outage probabilities using standard Monte Carlo simulations is not currently feasible even with the reduced model. Instead, Wang and Menyuk [8] used Gaussian extrapolation [15]. In this dissertation, I describe an importance sampling technique that I developed to calculate the exact pdf of the penalty and, from that, the outage probability due to polarization effects. Using this technique, I found that Gaussian extrapolation can lead to significant errors in the computation of the outage probability [16]. Importance sampling is a method for biasing Monte Carlo simulations that is described in [17]. However, to successfully apply importance sampling, the biasing technique must be appropriate for each

particular problem.

My second contribution is the introduction of a new accurate receiver model. The receiver model that I developed accurately relates the optical signal-to-noise ratio and the polarization states of the signal and the noise of a channel at the receiver to the Q-factor. The Q-factor is a widely used performance indicator that is highly correlated to the bit-error ratio. Wang and Menyuk related the optical signal-to-noise ratio of a channel to the Q-factor by using an *ad hoc* extension of the receiver model of Marcuse [18] and Humblet and Azizoğlu [19]. Additionally, the number of noise modes used by that model was determined in an *ad hoc* fashion; it was assumed that the extinction ratios were perfect; and the partial polarization of the noise at the receiver was neglected. Wang and Menyuk used the same receiver model in both their reduced model and full model simulations. So, the validity of the receiver model was never tested. I found that the performance of a channel has a strong dependence on the relative polarization states of signal and noise when the noise is partially polarized.

Here, I describe a method based on one by Winzer, *et al.* [20] to calculate the exact first- and second-order moments of the pdf for the marks and the spaces, given realistic optical and electrical filters, that my colleagues and I extended to allow for arbitrarily polarized noise. With this approach, one can accurately calculate the Q-factor and, consequently, the Q penalty. For systems without pattern dependencies, the Q-factor is defined as the ratio of the difference between the mean currents of marks and spaces to the sum of the standard deviations of the marks and spaces due to noise [18]. In Sec 5.2, I describe how one can accurately compute the Q-factor in the presence of pattern dependencies. The enhancement factor, which was intro-

duced by Wang and Menyuk in an *ad hoc* manner, is a parameter that determines how efficiently a given pulse format and receiver configuration convert optical signal-to-noise ratio into electrical signal-to-noise ratio prior to the decision circuit. In this dissertation, I show how the enhancement factor can be exactly taken into account. Moreover, I show how one can take into account arbitrary extinction ratios and accurately account for the polarization states of signal and noise when computing the performance degradation due to polarization effects. This model yields very good agreement with experiments [21]–[23].

In addition, I present a new procedure for validating the reduced model using full model simulations. In [8], Wang and Menyuk essentially compared the optical signal-to-noise ratios at the receiver for both models, since the polarization states of the signal and noise were not accounted for. In the comparison, Wang and Menyuk used different fiber realizations for the reduced and the full models. As a consequence, they could only verify consistency with the two models to within the statistical error of the full model. This statistical error was large because the number of realizations that can be kept in the full model is small. In the new procedure presented here, I used the same fiber realizations for both the reduced model and the full model. The same statistical fluctuations are thus present in both models, and any deviation between the two is due to nonlinear polarization rotation in the full model. In fact, I demonstrated that the deviation is small for realistic systems.

In collaboration with Dr. Carter's research group at the University of Maryland Baltimore County (UMBC), we carried out a series of experiments that validated the receiver model that I developed when combined with the reduced Stokes model. The

receiver model was validated using back-to-back and fiber recirculating loop studies. A key prediction of the new receiver model is that the relationship between the optical signal-to-noise ratio (OSNR) and the Q-penalty is not unique when the noise is partially polarized. Our experimental results verified this prediction. The reduced model was also validated at UMBC for single-channel soliton systems.

The remainder of this Ph.D. dissertation is organized as follows: In Chapter 2, I describe how I model the polarization state of light, and I also describe the polarization effects that can lead to impairment in optical fiber communication systems. In Chapter 3, I describe the equations of the full time-domain model and the reduced Stokes model that I use in this dissertation. In addition, I describe the algorithms that I use to numerically solve these equations. In Chapter 4, I derive a receiver model that provides an explicit relationship between the Q-factor and the optical signal-to-noise ratio in optical fiber communication systems for arbitrary pulse shape, realistic receiver filters, and arbitrarily polarized noise. My colleagues and I validate the receiver model that I developed by comparison to Monte Carlo simulations and back-to-back experiments. I also define the enhancement factor and three other parameters that explicitly quantify the relative performance of different modulation formats in a receiver. In Chapter 5, I describe the technique that I developed that uses Monte Carlo simulations with importance sampling to compute the probability density function of the Q-factor and the outage probability for a channel in a WDM optical fiber communication system due to the combination of PDL, inter-channel PMD, and PDG. I also show how to combine multiple distributions using importance sampling with biased polarization-induced penalty to compute the outage probability due to polarization

effects. In Chapter 6, I present results from fiber recirculating loop experiments that validate the use of the reduced Stokes model with the accurate receiver model that I introduced in Chapter 4.

Chapter 2

Polarization effects and gain saturation

2.1 Polarization state representations

The state of polarization of light is defined by the time evolution of the orientation of the electric-field vector $\boldsymbol{\mathcal{E}}(\mathbf{r}, t)$, a three-vector whose elements depend on both the time t and the position $\mathbf{r}$. For monochromatic light, the three components of $\boldsymbol{\mathcal{E}}(\mathbf{r}, t)$ vary sinusoidally in time with different amplitudes and phases. If the direction of propagation is parallel to the z-axis, the electric-field vector of a monochromatic wave of angular frequency ω propagating in an isotropic media with refractive index n is given by [24]

$$\boldsymbol{\mathcal{E}}(z, t) \;=\; \mathrm{Re}\left\{(E_x\hat{\mathbf{x}} + E_y\hat{\mathbf{y}})\exp\left[i(kz - \omega t)\right]\right\}, \tag{2.1}$$

where E_x and E_y are the complex envelopes of the electric-field in the x and y directions, respectively, $k = \omega n/c$ is the wavenumber, and ω is the angular frequency. The parameter c is the speed of light in vacuum.

In optical fibers the light is mostly confined to the core due to total internal reflection of the light rays in the core-cladding boundary [24]. In single-mode fibers, which

are used in modern day optical-fiber communications systems, the difference between the refractive indexes of the core and the cladding is so small that the guided light rays are nearly paraxial—parallel to the fiber axis. The longitudinal components of the electric and the magnetic fields are small compared to the transverse components. As a consequence, light waves are approximately transverse electromagnetic waves, and the two transverse components of the electric field, which are linearly polarized, are sufficient to completely characterize the polarization state of light. The evolution of the electric field vector in (2.1) generally traces out an ellipse in a plane that is transverse to the direction of propagation. The state of polarization of the light is characterized by either the Jones vector or the Stokes parameters, both of which are determined by the transverse components of the electric field.

The Jones vector is a complex valued two-vector that consists of the complex envelopes of the transverse components of the electric field vector in (2.1). The Jones vector completely characterizes the intensity, the phase, and the polarization state of a monochromatic wave, and is typically represented as [25]

$$\mathbf{E} = \begin{bmatrix} E_x \\ E_y \end{bmatrix}, \tag{2.2}$$

where E_x and E_y are the complex envelopes of two orthogonal transverse components of the electric field whose intensity is equal to $|E_x|^2 + |E_y|^2$. Typically, E_x and E_y represent the envelopes of the electric field in the horizontal and the vertical linear polarization states, respectively, which form a particular basis of orthogonal polarization states. For this reason, in this chapter I define $[1,\, 0]^t$ and $[0,\, 1]^t$ as normalized Jones vectors in the horizontal and the vertical polarization states, respectively, unless

otherwise stated, where ^{t} is the transpose operator. However, any pair of normalized Jones vectors that are orthogonal to the direction of propagation and to each other can determine an arbitrary basis of Jones vectors. In the basis used in (2.2), the normalized Jones vectors $\left[1/\sqrt{2},\, -i/\sqrt{2}\right]^t$ and $\left[1/\sqrt{2},\, i/\sqrt{2}\right]^t$ represent the right and the left circular polarization states, respectively, which are a linear combination of the horizontal and the vertical polarization states. These two polarization states are circular because the time evolution of the electric field vector of the light traces out a circle transverse to the propagation direction. All possible normalized Jones vectors correspond to all possible degrees of ellipticity of the time evolution of the electric field vector. The normalized Jones vector representation of the polarization state of light is not unique. A normalized Jones vector represents the phase of monochromatic light, as well as its polarization. For example, both normalized Jones vectors $[1,\, 0]^t$ and $[i,\, 0]^t$ represent the horizontal polarization state. The phase of the monochromatic waves must be taken into account when using full time-domain models of the light propagation in optical fibers. For this reason, Jones vectors are useful when running full time-domain simulations, as I will show in Sec. 3.1.1. We note that the complete Jones vector has intensity information, as well as phase and polarization state information. The normalized Jones vector does not have intensity information.

The Stokes parameters are the components of a real-valued four-vector that completely characterizes the intensity and the polarization state of either a monochromatic light wave or the superposition of light waves at different frequencies. If at least two frequency components of the light have different normalized Jones vectors, the light is said to be partially polarized. Therefore, the Stokes parameters can also

represent the polarization state of partially polarized light. The Stokes parameters $\boldsymbol{\mathcal{S}} = [S_0,\, S_1,\, S_2,\, S_3]^t$ of monochromatic light can be computed from the elements of the Jones vector as [26]

$$S_0 \,=\, \mathbf{E}^\dagger\, \sigma_0\, \mathbf{E}, \quad S_1 \,=\, \mathbf{E}^\dagger\, \sigma_3\, \mathbf{E}, \quad S_2 \,=\, \mathbf{E}^\dagger\, \sigma_1\, \mathbf{E}, \quad S_3 \,=\, -\mathbf{E}^\dagger\, \sigma_2\, \mathbf{E}, \tag{2.3}$$

where

$$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{2.4}$$

are the Pauli matrices [27]. We also define a Stokes vector that consists of the last three Stokes parameters,

$$\mathbf{S} \,=\, \begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix}. \tag{2.5}$$

The Stokes parameters of an arbitrary number of monochromatic light waves with different frequencies are equal to the sum of the Stokes parameters S_{0,ω_i} and $\mathbf{S}_{\omega_i}$ of each frequency component with angular frequency ω_i,

$$S_0 = \sum_i S_{0,\omega_i}\,; \qquad \mathbf{S} = \sum_i \mathbf{S}_{\omega_i}. \tag{2.6}$$

The polarized component of the light can be represented by the normalized Stokes vector, which is a three-vector $\mathbf{s} = \mathbf{S}/\,|\mathbf{S}|$ that consists of the Stokes vector divided by its magnitude. The normalized Stokes vectors $[1,\, 0,\, 0]^t$ and $[-1,\, 0,\, 0]^t$ represent the horizontal and the vertical polarization states, respectively, while $[0,\, 0,\, 1]^t$ and $[0,\, 0,\, -1]^t$ represent the circular right and the circular left polarization states, respectively. Each possible normalized Stokes vector represents a unique polarization

state and corresponds to a point on the surface of a unit sphere that is referred to as the Poincaré sphere. Points on the equator of the sphere correspond to linear polarizations; the two poles correspond to circular polarizations; intermediate points correspond to various degrees of ellipticity. The length of the Stokes vector, $|\mathbf{S}|$, represents the intensity of the polarized component of the light. The difference between the total intensity of the light and the intensity of the polarized component of the light represents the intensity of the unpolarized component of the light, $S_0 - |\mathbf{S}|$.

The degree-of-polarization (DOP) of the light is the ratio of the intensity of the polarized component of the light to total intensity of the light. The DOP can be expressed in terms of the elements of the Stokes vectors and the Stokes parameters as

$$\text{DOP} = \frac{|\mathbf{S}|}{S_0} = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0}. \tag{2.7}$$

Unpolarized light has a DOP equal to 0, while completely polarized light has DOP equal to 1. The light is partially polarized when the DOP is a number between 0 and 1. The Stokes vector completely characterizes the state of polarization of the light only if the light is completely polarized. If the light is partially polarized or unpolarized, all four Stokes parameters are required to characterize the state of polarization of the light.

2.2 The transmission matrix of optical devices

Both the polarization state and the intensity of light can be modified after the light passes through an optical device. The effect that an optical device produces on both

the polarization state and the intensity of light can be modeled either by a frequency-dependent Jones matrix or by a Müller matrix. In this Ph.D. dissertation, I indicate a Jones matrix with an overbar and the corresponding Müller matrix without an overbar. For example, the Jones matrix $\bar{\mathsf{T}}$ of a device produces the same effect on the input Jones vector that the corresponding Müller matrix T of the device produces on the input Stokes parameters. The output Jones vector $\mathbf{E}_{\text{out}}$ of a given device is related to the input Jones vector $\mathbf{E}_{\text{in}}$ by the equation [24]

$$\mathbf{E}_{\text{out}} \;=\; \bar{\mathsf{T}}\,\mathbf{E}_{\text{in}}, \tag{2.8}$$

where $\bar{\mathsf{T}}$ the 2×2 complex valued Jones matrix of the device.

The Müller matrix of a device is either a real-valued 3×3 or a 4×4 matrix. We use the former only when the device does not affect the intensity of the light in a way that is dependent on its polarization state. The 3×3 Müller matrix relates the output Stokes vector to the input Stokes vector through an equation of the form

$$\mathbf{S}_{\text{out}} \;=\; \mathsf{T}\,\mathbf{S}_{\text{in}}, \tag{2.9}$$

which is equivalent to the Jones matrix $\bar{\mathsf{T}}$ defined in (2.8).

If a device produces a change in both the polarization state and the intensity of the light, this device must be modeled by a 4×4 Müller matrix, which relates the output Stokes parameters to the input Stokes parameters. The output Stokes parameters $\boldsymbol{\mathcal{S}}_{\text{out}}$ of this device are related to the input Stokes parameters $\boldsymbol{\mathcal{S}}_{\text{in}}$ by the equation,

$$\boldsymbol{\mathcal{S}}_{\text{out}} \;=\; \mathsf{T}\,\boldsymbol{\mathcal{S}}_{\text{in}}. \tag{2.10}$$

Once a 3×3 Müller matrix of a device is known, assuming that the device can be mathematically represented by such a matrix, its corresponding 4×4 Müller matrix can be obtained using the relationship,

$$\mathsf{T}_{4\times4} = \begin{bmatrix} \det^{1/3}(\mathsf{T}_{3\times3}) & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \mathsf{T}_{3\times3} & \\ 0 & & & \end{bmatrix}. \tag{2.11}$$

If, however, a 4×4 Müller matrix of a device is known, it is only possible to represent this device using a 3×3 Müller matrix if the first row and the first column of the 4×4 Müller matrix are equal to the first row and the first column of the matrix that is on the right-hand side of (2.11), which is equivalent to the earlier statement that to be represented by the 3×3 Müller matrix, the device must not change the intensity of the light in a way that is dependent on the input polarization state. A 3×3 Müller matrix can produce polarization-independent gain or loss, which is represented by $\det^{1/3}(\mathsf{T}_{3\times3})$ in (2.11).

Once the Jones Matrix $\bar{\mathsf{T}}$ of a device is known, the Müller matrix T of this device can be computed using (2.3). Since the output Jones vector of this device is related to the input Jones vector by $\mathbf{E}_{\text{out}} = \bar{\mathsf{T}}\,\mathbf{E}_{\text{in}}$, as in (2.8), one can substitute (2.8) into (2.3) to express the components of the output Stokes parameters as

$$\begin{aligned} S_{0,\text{out}} &= \left(\bar{\mathsf{T}}\,\mathbf{E}_{\text{in}}\right)^{\dagger} \sigma_0 \left(\bar{\mathsf{T}}\,\mathbf{E}_{\text{in}}\right), \quad S_{1,\text{out}} = \left(\bar{\mathsf{T}}\,\mathbf{E}_{\text{in}}\right)^{\dagger} \sigma_3 \left(\bar{\mathsf{T}}\,\mathbf{E}_{\text{in}}\right), \\ S_{2,\text{out}} &= \left(\bar{\mathsf{T}}\,\mathbf{E}_{\text{in}}\right)^{\dagger} \sigma_1 \left(\bar{\mathsf{T}}\,\mathbf{E}_{\text{in}}\right), \quad S_{3,\text{out}} = -\left(\bar{\mathsf{T}}\,\mathbf{E}_{\text{in}}\right)^{\dagger} \sigma_2 \left(\bar{\mathsf{T}}\,\mathbf{E}_{\text{in}}\right), \end{aligned} \tag{2.12}$$

where the matrices σ_0, σ_1, σ_2, and σ_3 are the Pauli matrices defined in (2.4). Equation (2.12) can be re-expressed so that the effect produced by the matrix $\bar{\mathsf{T}}$ is separated

from the input Jones vector $\mathbf{E}_{\rm in}$, to obtain

$$\begin{aligned} S_{0,\rm out} &= \mathbf{E}_{\rm in}^{\dagger}\left(\bar{\mathsf{T}}^{\dagger}\,\sigma_0\,\bar{\mathsf{T}}\right)\mathbf{E}_{\rm in}, \quad S_{1,\rm out} = \mathbf{E}_{\rm in}^{\dagger}\left(\bar{\mathsf{T}}^{\dagger}\,\sigma_3\,\bar{\mathsf{T}}\right)\mathbf{E}_{\rm in}, \\ S_{2,\rm out} &= \mathbf{E}_{\rm in}^{\dagger}\left(\bar{\mathsf{T}}^{\dagger}\,\sigma_1\,\bar{\mathsf{T}}\right)\mathbf{E}_{\rm in}, \quad S_{3,\rm out} = -\mathbf{E}_{\rm in}^{\dagger}\left(\bar{\mathsf{T}}^{\dagger}\,\sigma_2\,\bar{\mathsf{T}}\right)\mathbf{E}_{\rm in}. \end{aligned} \tag{2.13}$$

The four rows of the Müller matrix T that corresponds to the Jones matrix $\bar{\mathsf{T}}$ are obtained using (2.13) by solving each of the four independent sets of linear equations given by the matrices

$$\begin{aligned} \bar{\mathsf{T}}^{\dagger}\,\sigma_0\,\bar{\mathsf{T}} &= \beta_0^{(0)}\sigma_0 + \beta_1^{(0)}\sigma_3 + \beta_2^{(0)}\sigma_1 - \beta_3^{(0)}\sigma_2, \\ \bar{\mathsf{T}}^{\dagger}\,\sigma_3\,\bar{\mathsf{T}} &= \beta_0^{(1)}\sigma_0 + \beta_1^{(1)}\sigma_3 + \beta_2^{(1)}\sigma_1 - \beta_3^{(1)}\sigma_2, \\ \bar{\mathsf{T}}^{\dagger}\,\sigma_1\,\bar{\mathsf{T}} &= \beta_0^{(2)}\sigma_0 + \beta_1^{(2)}\sigma_3 + \beta_2^{(2)}\sigma_1 - \beta_3^{(2)}\sigma_2, \\ \bar{\mathsf{T}}^{\dagger}\,\sigma_2\,\bar{\mathsf{T}} &= \beta_0^{(3)}\sigma_0 + \beta_1^{(3)}\sigma_3 + \beta_2^{(3)}\sigma_1 - \beta_3^{(3)}\sigma_2. \end{aligned} \tag{2.14}$$

Once the coefficients β in (2.14) are computed, one can express the Müller matrix that relates the output Stokes parameters to the input Stokes parameters as

$$\mathsf{T} = \begin{bmatrix} \beta_0^{(0)} & \beta_1^{(0)} & \beta_2^{(0)} & \beta_3^{(0)} \\ \beta_0^{(1)} & \beta_1^{(1)} & \beta_2^{(1)} & \beta_3^{(1)} \\ \beta_0^{(2)} & \beta_1^{(2)} & \beta_2^{(2)} & \beta_3^{(2)} \\ \beta_0^{(3)} & \beta_1^{(3)} & \beta_2^{(3)} & \beta_3^{(2)} \end{bmatrix}, \tag{2.15}$$

so that $\boldsymbol{\mathcal{S}}_{\rm out} = \mathsf{T}\,\boldsymbol{\mathcal{S}}_{\rm in}$, as in (2.10).

The Jones matrix of a linear polarizer, which is a device that transmits only the horizontal polarization state, is given by

$$\bar{\mathsf{T}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \tag{2.16}$$

This device transforms a given input Jones vector $[E_x,\, E_y]^t$ to the output Jones vector $[E_x,\, 0]^t$. Therefore, a linear polarizer affects both the polarization state and

the intensity of the light that passes through the device. The 4×4 Müller matrix of this device is given by

$$\mathsf{T} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{2.17}$$

which transforms any given input Stokes parameters $[S_0,\, S_1,\, S_2,\, S_3]^t$ into the output Stokes parameters $[S_0/2 + S_1/2,\, S_0/2 + S_1/2,\, 0,\, 0]^t$.

A polarization rotator, on the other hand, is a device that rotates the plane of polarization of a plane wave around the propagation axis without modifying the intensity of the wave. For example, the rotation matrix

$$\bar{\mathsf{T}} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \tag{2.18}$$

transforms a linearly polarized wave with Jones vector $[\cos\theta_1,\, \sin\theta_1]^t$ into a linearly polarized wave with Jones vector $[\cos\theta_2,\, \sin\theta_2]^t$, where $\theta_2 = \theta_1 + \theta$. The 3×3 Müller matrix of this polarization rotator is given by

$$\mathsf{T} = \begin{bmatrix} \cos 2\theta & -\sin 2\theta & 0 \\ \sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{2.19}$$

which is equivalent to the Jones matrix $\bar{\mathsf{T}}$ in (2.18). The Müller matrix in (2.19) produces a rotation on the Poincaré sphere by the angle 2θ about the S_3-axis.

2.3 Polarization mode dispersion

2.3.1 Physical origin of polarization mode dispersion

It has been known for many years that all single-mode optical fibers are actually bimodal due to the presence of birefringence, which breaks the two-fold degeneracy of the HE_{11} mode [28]. The birefringence difference between the two local eigenmodes is very weak in absolute terms—typically, $\Delta n/n \simeq 10^{-7}$. However, the wavelength of light is very small, $\lambda \simeq 1.55$ μm, so that typical beat lengths are on the order of 3 to 30 meters. The beat length is the length in which the phase difference between the polarization eigenmodes changes by 2π at a given wavelength, such as 1.55 μm. This scale, which is typically tens of meters, is very small compared to typical dispersive scale lengths, nonlinear scale lengths, and system lengths, all of which are typically hundreds or even thousands of kilometers. At the same time, the orientation of the axes of birefringence changes randomly over length scales that vary from a fraction of a meter to a hundred meters, depending on the fiber type. Since the magnitude of an effect is inversely proportional to its scale length, the birefringence should be considered large but rapidly varying, relative to the system scale lengths of interest [29]. Birefringence is a linear effect that changes the phase difference between the light components that propagate in the two local eigenmodes of the fiber, thereby changing the polarization state of light. The phase difference between the light components that propagate in the two local eigenmodes of polarization is frequency dependent. Over a restricted wavelength range the deviation of the phase difference between the two polarization components of the light varies linearly with frequency. The length

scale on which a single frequency changes its polarization state equals the length scale on which the orientation of the optical fiber birefringence changes, which is also the same order of magnitude as the beat length [11]; so, the polarization state of the light is also rapidly changing with distance. By contrast, the changes in the polarization states for all frequencies in the communication band of a typical WDM optical fiber system are *nearly* identical, so that the polarization states in two different frequencies drift apart slowly if they start out the same [29]. It is this differential drift that is the physical source of polarization-mode dispersion (PMD).

The length scale on which the differential drift of the evolution of the polarization states inside a communication band occurs varies over a wide range, since it depends on the bandwidth as well as the properties of the optical fiber; however, typical values range from tens of kilometers to tens of thousands of kilometers. When the differential drift is large enough to affect the polarization states inside the bandwidth of a single wavelength channel, the pulses can spread in the time domain. This pulse spreading can lead to an increase in the bit-error ratio, where the bit-error ratio is the probability of error in the decoded bit at the receiver. This effect becomes more important as the data rate of a single channel increases because the bandwidth of a single channel increases with the data rate. In this dissertation, I focus on systems whose PMD is small enough so that it does not produce significant waveform distortions. In these systems, the mean of the accumulated differential-group delay of the transmission line does not exceed 10% of the bit period [30], [31]. These systems include present-day trans-oceanic fiber transmission systems operating at 10 Gbit/s per channel and long-haul terrestrial communications systems that use low-PMD optical fibers that were

deployed after 1990. However, PMD still has an important effect on these systems since it causes the polarization states of different wavelength channels to gradually drift apart, as I will describe in the next section.

2.3.2 Modeling of polarization mode dispersion

PMD is caused by the randomly varying strength and the orientation of the axis of birefringence in optical fibers [9]. Poole and Favin [32] and others [11], [33] showed that a model that consists of a concatenation of a large number of sections of constant birefringence is sufficient to model the PMD statistics if the polarization state is randomly rotated prior to each section. This random rotation produces a random coupling of the light in the two local principal states of polarization of a birefringent section, which are the local fiber modes. Marcuse, *et al.* [27] analyzed the coarse-step method, which uses large birefringent section lengths to model PMD in optical fibers. They showed that the length of the sections could be much larger than both the beat length and the correlation length without affecting the PMD statistics. If we consider only the linear PMD term in the the coupled nonlinear Schrödinger equation [34], the Jones matrix $\bar{\mathsf{F}}(\omega)$ of an optical fiber that consists of N constant birefringent sections may be written as [27]

$$\bar{\mathsf{F}}(\omega) = \prod_{n=1}^{N} \bar{\mathsf{P}}\,\bar{\mathsf{S}}_n, \tag{2.20}$$

where $\bar{\mathsf{S}}_n$ is a unitary Jones matrix that produces the random mode coupling in the n-th section, and $\bar{\mathsf{P}}$ models the propagation through a constant birefringent section. To efficiently solve the Manakov-PMD equation using the coarse-step method [27],

we randomly rotate the polarization states uniformly on the Poincaré sphere at the beginning of each section. Therefore, we define the Jones matrix $\bar{\mathsf{S}}_n$ as

$$\bar{\mathsf{S}}_n = \begin{bmatrix} \cos(\xi_n/2)\exp[i(\psi_n+\phi_n)/2] & i\sin(\xi_n/2)\exp[i(\psi_n-\phi_n)/2] \\ i\sin(\xi_n/2)\exp[-i(\psi_n-\phi_n)/2] & \cos(\xi_n/2)\exp[-i(\psi_n+\phi_n)/2] \end{bmatrix}, \quad (2.21)$$

which models the random mode coupling prior to the n-th birefringence section. The parameters ξ_n, ψ_n, and ϕ_n are random variables that are independent at each n and from each other. In (2.21), the probability density functions (pdf) of the angles ψ_n and ϕ_n are uniformly distributed between 0 and 2π, while the pdf of the quantities $\cos\xi_n$ are uniformly distributed between -1 and 1. The Jones matrix

$$\bar{\mathsf{P}} = \begin{bmatrix} \exp(-i\omega b'\Delta_z) & 0 \\ 0 & \exp(i\omega b'\Delta_z) \end{bmatrix}, \quad (2.22)$$

models the transmission through a birefringent section, where Δ_z is the length of the n-th section, and ω is the angular frequency. The parameter b' is the linear birefringence per unit length, which is given by [27]

$$b' = \frac{D_{\mathrm{PMD}}}{2\,\Delta_z^{1/2}}, \quad (2.23)$$

where D_{PMD} is the PMD coefficient (in ps/km$^{1/2}$).

The 3×3 Müller matrix S_n equivalent to the Jones matrix $\bar{\mathsf{S}}_n$ in (2.21) is comprised of elementary rotations around the x-axis and the y-axis [35], [36] of the Poincaré sphere,

$$\mathsf{S}_n \;=\; \mathsf{R}_x(\psi_n)\,\mathsf{R}_y(\xi_n)\,\mathsf{R}_x(\phi_n)\,. \quad (2.24)$$

Consequently, the matrices $\bar{\mathsf{S}}_n$ in (2.21) correspond to uniformly distributed rotations on the Poincaré sphere. The Müller matrix P equivalent to the Jones matrix

$\bar{\mathsf{P}}$ in (2.22) is simply a rotation around the x-axis of the Poincaré sphere: $\mathsf{P} = \mathsf{R}_x\left(-2\omega b'\Delta_z\right)$. The 4×4 Müller matrix that is equivalent to (2.24) can be defined from the 3×3 Müller matrix as in (2.11).

2.4 Polarization-dependent loss

Some devices produce a polarization-dependent attenuation on the input light, which leads to polarization-dependent loss (PDL). Physical effects that produce PDL include bulk dichroism (polarization-dependent absorption), fiber bending, angled optical interfaces, and oblique reflection [7]. In optical fiber communication systems, there are several sources of PDL: optical connectors, couplers, isolators, circulators, and demultiplexers [7].

The differential attenuation α produced by the PDL of a device is defined as

$$\alpha = \frac{P_{\min}}{P_{\max}}, \tag{2.25}$$

where $P_{\min}$ and $P_{\max}$ are the minimum and maximum optical power that are measured at the output of the device under test, given that polarized light is launched in all possible polarization states at the input of the device. It follows that $0 \leq \alpha \leq 1$.

The effect of PDL is to cause excess loss in one of two orthogonal polarization states. Using the Jones vector notation, one can define the transmission matrix of a PDL element with the highest loss in arbitrary polarization state as

$$\bar{\mathsf{S}}_{\mathrm{PDL}} = \bar{\mathsf{R}}_{\mathrm{PDL}}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & \alpha^{1/2} \end{bmatrix} \bar{\mathsf{R}}_{\mathrm{PDL}}, \tag{2.26}$$

where the matrix $\bar{\mathsf{R}}_{\mathrm{PDL}}$ is a unitary Jones matrix that rotates the polarization state with highest loss to the vertical polarization state $[0,\, 1]^t$. Therefore, a PDL element whose highest loss is in the vertical polarization state has the polarization rotation matrix $\bar{\mathsf{R}}_{\mathrm{PDL}}$ equal to a 2×2 identity matrix.

The effect of PDL on the Stokes parameters of the light $\boldsymbol{\mathcal{S}} = [S_0,\, S_1,\, S_2,\, S_3]^t$ is given by the 4×4 Müller matrix

$$\mathsf{S}_{\mathrm{PDL}} \;=\; \mathsf{R}_{\mathrm{PDL}}^{-1} \begin{bmatrix} (1+\alpha)/2 & (1-\alpha)/2 & 0 & 0 \\ (1-\alpha)/2 & (1+\alpha)/2 & 0 & 0 \\ 0 & 0 & \alpha^{1/2} & 0 \\ 0 & 0 & 0 & \alpha^{1/2} \end{bmatrix} \mathsf{R}_{\mathrm{PDL}}, \tag{2.27}$$

where the matrices $\mathsf{S}_{\mathrm{PDL}}$ and $\mathsf{R}_{\mathrm{PDL}}$ are equivalent respectively to the matrices $\bar{\mathsf{S}}_{\mathrm{PDL}}$ and $\bar{\mathsf{R}}_{\mathrm{PDL}}$ in (2.26).

2.5 Polarization-dependent gain

There are two kinds of polarization-dependent gain (PDG) in erbium-doped fiber amplifiers (EDFAs). The first kind is induced by laser diode pumping, which can produce a polarization sensitive excitation of the erbium atoms if the pump laser is polarized. The gain difference between the polarization state with maximum and with minimum gain due to this kind of PDG is about 0.12 dB; however, this kind of PDG is nearly eliminated in practical systems by pumping in two orthogonal polarization states. The second kind of PDG is due to the polarization-dependent saturation of the EDFA, which is also known as polarization-hole burning. The polarization-dependent saturation produces a larger gain in the light component that is in the polarization

state that is orthogonal to the polarized component of the light. This effect is proportional to the DOP of the incoming light. The intensity and the polarization state of the polarized component of the light are represented by the Stokes vector **S**. The PDG coefficient g_{PDG} is only about 0.07 dB for an EDFA with 3 dB of gain compression, but it becomes larger as the amplifier goes deeper into gain compression [37]. An amplifier with 3 dB of gain compression produces a gain in the signal that is 3 dB smaller than when the gain is unsaturated.

In single-channel systems, the presence of PDG in EDFAs can significantly degrade the optical signal-to-noise ratio (OSNR) in the transmission because the noise that is orthogonal to the polarization state of the signal is enhanced. However, the effect of PDG is negligible in modern-day WDM systems [8], since it is common to launch neighboring channels in orthogonal polarization states to mitigate both the PDG and the nonlinear interaction between the channels, and PMD tends to randomize the polarization state of the channels so that the overall DOP remains at a low level. Therefore, the DOP and, consequently, the PDG effect is small throughout the entire transmission line.

The effect of PDG is similar to the effect of PDL; however, with PDG the polarization state with excess gain is dependent on the polarization state of the incoming light. Using the Jones vector notation, one can define the transmission matrix of a PDG element as

$$\bar{\mathsf{S}}_{\mathrm{PDG}} = \bar{\mathsf{R}}_{\mathrm{PDG}}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & g^{1/2} \end{bmatrix} \bar{\mathsf{R}}_{\mathrm{PDG}}, \tag{2.28}$$

where g is the polarization dependent gain, and $\bar{\mathsf{R}}_{\mathrm{PDG}}$ is a Jones matrix that rotates the polarization state of the incoming light to the horizontal polarization $[1,\, 0]^t$. As

a consequence, the polarization state that is orthogonal to the polarization state of the incoming light is transformed to the vertical polarization $(0, 1)$, which undergoes an excess gain. The polarization dependent gain g is given by [7],

$$\log_{10} g = \frac{1}{10} \left(\text{PDG} \times \text{DOP} \right), \tag{2.29}$$

where DOP here refers to the degree-of-polarization of the incoming light. Note that $g > 1$, since PDG > 0. The Müller matrix that describes the effect of PDG on the Stokes parameters $\boldsymbol{\mathcal{S}} = [S_0, S_1, S_2, S_3]^t$ of the incoming light is similar to the corresponding Müller matrix for the PDL in (2.27),

$$\mathsf{S}_{\text{PDG}} = \mathsf{R}_{\text{PDG}}^{-1} \begin{bmatrix} (1+g)/2 & (1-g)/2 & 0 & 0 \\ (1-g)/2 & (1+g)/2 & 0 & 0 \\ 0 & 0 & g^{1/2} & 0 \\ 0 & 0 & 0 & g^{1/2} \end{bmatrix} \mathsf{R}_{\text{PDG}}. \tag{2.30}$$

The Müller matrices S_{PDG} and R_{PDG} in (2.30) are equivalent to the Jones matrices $\bar{\mathsf{S}}_{\text{PDG}}$ and $\bar{\mathsf{R}}_{\text{PDG}}$ in (2.28), respectively.

2.6 Gain saturation

In long-haul WDM systems, the combined effects of PMD, PDL, PDG, and the gain saturation in the optical amplifiers produce performance fluctuations that can lead to outages. PMD causes the polarization states of different WDM channels to drift apart. Then, the channels undergo different attenuations as they pass through the PDL elements in the optical amplifiers. The optical amplifiers are designed to operate in the saturated regime, which mitigates for the average power fluctuations in the transmission line. Because the gain saturation is a frequency independent effect, the

excess loss that some channels undergo due to the PDL elements are not compensated by the optical amplifiers, so that the optical power of the channels that undergo excess loss due to PDL is effectively transfered to the other channels. As a consequence, different channels have different optical signal-to-noise ratios prior to the receiver. In addition, the PDL elements partially polarize the noise, which significantly affects the performance of each channel.

The gain in erbium-doped fiber amplifiers (EDFAs) depends on a large number of parameters, such as the erbium-ion concentration, amplifier length, core radius, and pump power [38]. A three-level rate-equation is commonly used to model EDFAs, since stimulated emission terminates in the ground state in EDFAs. However, the evolution of the optical power in EDFAs can be approximated by the differential equation [39],

$$\frac{\mathrm{d}P(z)}{\mathrm{d}z} = \frac{g_0 P(z)}{1 + P(z)/P_s}, \tag{2.31}$$

where g_0 is the unsaturated differential gain, P_s is the intrinsic saturation power of the erbium-doped fiber, and $P(z)$ is the signal power at a distance z along the erbium-doped fiber in the amplifier. The saturated gain G of an amplifier with length L is given by

$$G = \frac{P_{\mathrm{out}}}{P_{\mathrm{in}}}, \tag{2.32}$$

where $P_{\mathrm{in}} = P(z = 0)$ and $P_{\mathrm{out}} = P(z = L)$. An amplifier is said to operate at 3 dB gain compression when the gain G is 3 dB smaller than the small signal gain ($P_{\mathrm{out}} \ll P_s$). The gain of a given amplifier increases with the reduction of the input power P_{in} of this amplifier. This increase in the gain recovers on the order of 0.6 dB of the power for every 1 dB reduction of the input power due to excess losses in the

previous stages when the gain compression is about 5 dB [10]. The gain saturation in optical amplifiers effectively produces an automatic gain control that partially compensates for the power fluctuations along the transmission.

Wang and Menyuk [8] used a gain saturation model that is simpler than the one in (2.31). In their model, the output power of an EDFA is assumed to be virtually independent of the input power, except for the added spontaneous-emission noise. Once the target output power $P_{\text{out,target}}$ is defined, in the absence of random power fluctuations, the gain of an amplifier with input power P_{in} can be obtained from the expression [38],

$$P_{\text{out,target}} = G\,P_{\text{in}} + 2n_{\text{sp}}(G-1)\sum_{m} h\nu^{(m)}B^{(m)}, \tag{2.33}$$

where n_{sp} is the spontaneous emission factor, G is the amplifier gain, which compensates for the losses in the fiber, and $h\nu^{(m)}$ and $B^{(m)}$ are the energy of a single photon and the optical bandwidth of the m-th channel, respectively. The parameter h is the Planck's constant [38], and $\nu^{(m)}$ is the central frequency of the m-th channel. Using (2.33), one can compute the saturated gain, which is given by

$$G = \frac{P_{\text{out,target}} + 2n_{\text{sp}}\sum_m h\nu^{(m)}B^{(m)}}{P_{\text{in}} + 2n_{\text{sp}}\sum_m h\nu^{(m)}B^{(m)}}. \tag{2.34}$$

In the saturation model in (2.34), each amplifier recovers all the power fluctuation at the input of the amplifier, which can be caused by excess losses in the previous stages. On the other hand, the saturation model in (2.31) recovers only about 0.6 dB for every 1 dB of power reduction in the input due to excess losses. However, my colleagues and I numerically verified that both models produce similar results when

the number of amplifiers is large and the optical amplifers are set up to compensate for the average losses in the system. For this reason, we used the gain saturation model in (2.34) throughout the work reported in this dissertation unless otherwise stated.

Chapter 3

Optical fiber transmission models

3.1 Full time-domain model

3.1.1 Manakov-PMD equation

Light propagation in optical fibers is described by the coupled nonlinear Schrödinger (CNLS) equation, its extensions, and its reductions [34]. Menyuk and Wai [11] developed the Manakov-PMD equation,

$$i\frac{\partial \mathbf{\Psi}}{\partial z} - \frac{1}{2}\beta''\frac{\partial^2 \mathbf{\Psi}}{\partial t^2} - \frac{i}{6}\beta'''\frac{\partial^3 \mathbf{\Psi}}{\partial t^3} + \frac{8}{9}n_2 k_0 \left|\mathbf{\Psi}\right|^2 \mathbf{\Psi} - i\frac{\gamma\left(z\right)}{2}\mathbf{\Psi} = -i\, b'\, \bar{\sigma}\frac{\partial \mathbf{\Psi}}{\partial t}, \tag{3.1}$$

which is a reduction of the CNLS equation, in which nonlinear PMD is ignored as is appropriate in communications systems [27]. The Jones vector $\mathbf{\Psi}$ is the electric field envelope transformed to a basis that freezes the evolution of the polarization state of the field $\mathbf{\Psi}$ at the center frequency [27]. The parameter $b = (\beta_1 - \beta_2)$ is the difference between the propagation constants of the fast mode β_1 and the slow mode β_2 of the fiber, and the primes represent derivatives with respect to the angular frequency ω, where $\beta'' = \beta_1'' \approx \beta_2''$ is the chromatic dispersion and β''' is its frequency derivative. The quantity b' that I define here is the continuous analog of the quantity b' that I

previously defined after (2.22) for the coarse step method [27]. The parameter $\gamma(z)$ is the attenuation, and n_2 is the nonlinear coefficient of the fiber [34]. The term on the right-hand side of (3.1) corresponds to the usual linear PMD, where $\bar{\sigma}$ is a unitary Jones matrix that is related to the orientation of the randomly varying axes of birefringence [29]. The derivation of the Manakov-PMD equation was based on the observation that while single frequency, continuous wave (CW) light at low intensity undergoes a rapid random evolution on the Poincaré sphere, the trajectories on the Poincaré sphere of different frequencies in the bandwidth of a typical communication signal drift apart slowly. One efficient technique to solve the Manakov-PMD equation is the coarse-step method, where the polarization states are scattered uniformly on the Poincaré sphere at the beginning of each step and the birefringence strength and orientation is kept constant within each step. I described this technique in Sec. 2.3.2. The steps can be orders of magnitude larger than the length scale over which the birefringence orientation changes in optical fibers, as long as it is still short compared to the nonlinear and dispersive scale lengths. Thus, solving for $\mathbf{\Psi}$ using the coarse-step method is much faster than solving for $\mathbf{\Psi}$ using a method that resolves the details of the variations of the signal's state of polarization.

Equation (3.1) is based on the Jones representation of the electric field envelope. This representation is appropriate when carrying out studies in which the time-domain is resolved. Later in this chapter I will introduce a reduced model in which I average over the time domain of each wavelength channel. In this case, a Stokes parameter representation is appropriate, since in general the degree-of-polarization (DOP) within a single wavelength channel does not equal one.

3.1.2 Split-step Fourier method

The Manakov-PMD equation (3.1) is a coupled nonlinear partial differential equation that does not generally have analytical solutions. A numerical approach is necessary to solve the Manakov-PMD equation. A commonly used method to solve this equation and its reductions is the split-step Fourier method [38]. In this Chapter, I show how to apply the split-step Fourier method to solve the Manakov-PMD equation. Since the split-step Fourier method is efficient and flexible, I applied this method to solve the full time-domain transmission model in this dissertation.

The Manakov-PMD equation in (3.1) can be represented as

$$\frac{\partial \mathbf{\Psi}}{\partial z} = \left(\hat{\mathsf{D}} + \hat{\mathsf{N}}\right) \mathbf{\Psi}, \tag{3.2}$$

where $\hat{\mathsf{D}}$ is a linear differential matrix operator that accounts for PMD, dispersion and attenuation and $\hat{\mathsf{N}}$ is a nonlinear matrix operator that accounts for the fiber nonlinearity. The operators $\hat{\mathsf{D}}$ and $\hat{\mathsf{N}}$ are given by

$$\hat{\mathsf{D}} = \left(-\frac{i}{2}\beta'' \frac{\partial^2}{\partial t^2} + \frac{1}{6}\beta''' \frac{\partial^3}{\partial t^3} - \frac{\gamma}{2}\right) \bar{\mathsf{I}} - ib'\bar{\sigma}\frac{\partial}{\partial t} \tag{3.3}$$

and

$$\hat{\mathsf{N}} = i\frac{8}{9} n_2 k_0 \left|\mathbf{\Psi}\right|^2 \bar{\mathsf{I}}, \tag{3.4}$$

where $\bar{\mathsf{I}}$ is the identity matrix. The formal solution of the first-order nonlinear differential equation in (3.2) for the propagation of the field from z to $z + \Delta_z$ is given by

$$\mathbf{\Psi}(z + \Delta_z) = \exp\left[\Delta_z \left(\hat{\mathsf{D}} + \hat{\mathsf{N}}\right)\right] \mathbf{\Psi}(z). \tag{3.5}$$

However, it is not practical to solve (3.5) numerically because $\hat{\mathsf{D}}$ is a linear operator in which frequency-domain field components of the field evolve independently but time domain components are coupled together, while $\hat{\mathsf{N}}$ is a nonlinear operator in which time-domain field components evolve independently but frequency-domain components are coupled together. The linear and the nonlinear operators are non-commuting operators that act together throughout the fiber, as described in (3.5). However, one can obtain an approximate solution for $\mathbf{\Psi}$ after the field propagates small distance Δ_z through the fiber by assuming that the operators $\hat{\mathsf{D}}$ and $\hat{\mathsf{N}}$ are separable. In addition, we assume that the strength and the orientation of the birefringence is constant throughout the distance Δ_z. More precisely, one can propagate (3.2) from z to $z+\Delta_z$ in three steps using the symmetrized split-step Fourier method that mathematically corresponds to

$$\mathbf{\Psi}(z+\Delta_z) \;=\; \exp\left(\frac{\Delta_z}{2}\hat{\mathsf{D}}\right)\exp\left(\Delta_z\hat{\mathsf{N}}\right)\exp\left(\frac{\Delta_z}{2}\hat{\mathsf{D}}\right)\mathbf{\Psi}(z). \tag{3.6}$$

In the first step, the linear operator acts alone ($\hat{\mathsf{N}}=0$) propagating from z to $z+\Delta_z/2$ after the input field $\mathbf{\Psi}(z)$ undergoes a uniformly distributed rotation on the Poincaré sphere as in (2.21) to model the random mode coupling due to PMD using the coarse-step method. After this random rotation, we assume that the local birefringence orientation defined by $\bar{\sigma}$ corresponds to the horizontal and the vertical polarization states. More precisely,

$$\bar{\sigma} \;=\; \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{3.7}$$

Since the operator $\hat{\mathsf{D}}$ is linear, the propagation of (3.2) from z to $z+\Delta_z/2$ can be efficiently carried out in the frequency domain because each frequency component of

the field $\mathbf{\Psi}$ evolves independently. More specifically,

$$\mathbf{\Psi}(z+\Delta_z/2,t) \;=\; \exp\left(\frac{\Delta_z}{2}\,\hat{\mathsf{D}}\right)\mathbf{\Psi}(z,t) = \mathcal{F}^{-1}\left\{\exp\left[\frac{\Delta_z}{2}\,\hat{\mathsf{D}}(\omega)\right]\;\mathcal{F}\left[\mathbf{\Psi}(z,t)\right]\right\}, \quad (3.8)$$

where $\mathcal{F}\left[\cdot\right]$ and $\mathcal{F}^{-1}\left\{\cdot\right\}$ denote the forward and inverse Fourier transform, and

$$\hat{\mathsf{D}}(\omega) \;=\; \left(\frac{i}{2}\beta''\omega^2 + \frac{i}{6}\beta'''\omega^3 - \frac{\gamma}{2}\right)\bar{\mathsf{I}} - ib'\omega\bar{\sigma} \quad (3.9)$$

is the linear differential operator in (3.3) acting on a monochromatic plane wave of frequency ω. In order to be consistent with our choice of negative carrier frequencies in (2.1), we must define the Fourier transform, so that

$$\mathcal{F}\left[g(t)\right] \;\equiv\; \int_{-\infty}^{+\infty} dt\, g(t)\exp(i\omega t), \quad (3.10)$$

$$\mathcal{F}^{-1}\left\{G(\omega)\right\} \;=\; \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi}\, G(\omega)\exp(-i\omega t). \quad (3.11)$$

I note that while this convention is standard in the physics literature, it is opposite of the convention that is used in most of the electrical engineering literature. Both conventions are used in the optical communications literature. A thorough discussion of this issue may be found in Menyuk, *et al.* [40]. An important practical issue in the numerical implementation of (3.8) is that the usual convention for fast Fourier transforms [41] corresponds to the opposite signs in (3.10) and (3.11). Thus, the Fourier transform that is defined in (3.10) corresponds to the inverse fast Fourier transform, and vice versa.

In the second step, the nonlinearity acts alone ($\hat{\mathsf{D}} = 0$) starting with the result of the first step to propagate $\mathbf{\Psi}$ from z to $z + \Delta_z$. Therefore, the propagation of the nonlinear operator of (3.2) from z to $z + \Delta_z$ can be efficiently carried out in

the time domain because each component of the field $\mathbf{\Psi}$ in the time domain evolves independently. Specifically, one finds

$$\mathbf{\Psi}(z + \Delta_z) \;=\; \exp\left(\Delta_z\, \hat{\mathsf{N}}\right) \mathbf{\Psi}(z). \tag{3.12}$$

Finally, in the third step the dispersion acts alone once again ($\hat{\mathsf{N}} = 0$) starting with the result of the second step to propagate $\mathbf{\Psi}$ from $z + \Delta_z/2$ to $z + \Delta_z$ as in (3.8). The symmetry of the numerical evaluation of the linear and nonlinear operators in the symmetrized split-step Fourier algorithm produces a numerical result that has second-order accuracy with respect to the step size Δ_z [38]. Consequently, the use of the split-step Fourier method in combination with the fast Fourier transform [41] can be up to two orders of magnitude faster than simple finite-difference implementations with the same accuracy [38]. While more sophisticated implementations of finite difference methods can produce comparable speeds to the split-step Fourier method [42], these implementations are complex and cannot deal easily with changes in the dispersion relations.

3.2 Reduced Stokes Model

Wang and Menyuk [8] introduced a reduced Stokes model to efficiently compute the penalty produced by the combined effects of inter-channel PMD, PDL, and PDG, as well as gain saturation in the amplifiers. Since we are interested in the evolution of the polarization state of an entire optical channel, the reduced model follows only the evolution of the Stokes parameters of each channel in a WDM system. We first define $\mathbf{E}(z, t)$ to be the sum of the complex envelopes of all the channels in the time

domain,

$$\mathbf{E}(z,t) = \sum_{m=1}^{M} \mathbf{E}^{(m)}(z,t) \exp\left[ik^{(m)}z - i\omega^{(m)}t\right], \tag{3.13}$$

where $k^{(m)}$ and $\omega^{(m)}$ are the central wavenumber and frequency of the m-th channel, measured with respect to the central wavenumber and frequency of $\mathbf{E}(z,t)$. The variable $\mathbf{E}^{(m)}(t)$ is the corresponding wave envelope. The Stokes parameters of the m-th channel are defined by

$$\begin{aligned}
S_0^{(m)}(z) &= \frac{1}{T}\int_{t_1}^{t_2} \left[\left|E_x^{(m)}(z,t)\right|^2 + \left|E_y^{(m)}(z,t)\right|^2\right] dt \\
S_1^{(m)}(z) &= \frac{1}{T}\int_{t_1}^{t_2} \left[\left|E_x^{(m)}(z,t)\right|^2 - \left|E_y^{(m)}(z,t)\right|^2\right] dt \\
S_2^{(m)}(z) &= \frac{2}{T}\int_{t_1}^{t_2} \mathrm{Re}\left[E_x^{(m)}(z,t)\, E_y^{(m)*}(z,t)\right] dt \\
S_3^{(m)}(z) &= \frac{2}{T}\int_{t_1}^{t_2} \mathrm{Im}\left[E_x^{(m)}(z,t)\, E_y^{(m)*}(z,t)\right] dt
\end{aligned} \tag{3.14}$$

where $T = t_2 - t_1$ is the interval of integration for the computation of the Stokes parameters, $E_x^{(m)}$ and $E_y^{(m)}$ are the complex wave envelopes in two orthogonal polarization state, and the $*$ indicates complex conjugate. I assume that the interval T is sufficiently large compared to the bit period so that the Stokes parameters are independent of the bit pattern. The parameter $S_0^{(m)}$ gives the total power of the channel, while the vector magnitude $\left|\mathbf{S}^{(m)}\right|$ gives the polarized power in the channel, where $\mathbf{S}^{(m)} = \left[S_1^{(m)}, S_2^{(m)}, S_3^{(m)}\right]^t$. The unpolarized power is the difference between the total power and the unpolarized power, $S_0^{(m)} - \left|\mathbf{S}^{(m)}\right|$. I associate the Stokes parameters of each channel with the central frequency of the channel $\omega^{(m)}$. In the

reduced model, one must keep track of the evolution of the four Stokes parameters for the signal $\boldsymbol{\mathcal{S}}^{(m)}_{\text{signal}}$ and four Stokes parameters for the noise $\boldsymbol{\mathcal{S}}^{(m)}_{\text{noise}}$ for each individual channel. It is necessary to follow the noise and signal parameters separately so that one can compute both the polarized and unpolarized components of the signal and of the noise at the receiver for each channel.

To model PMD, PDL, and PDG in the system I use the 4×4 Müller matrices that I described in Sec. 2.2. A system realization is simulated using the reduced Stokes model by concatenating the Müller matrices of the elements of the system to relate the output Stokes parameters of all the channels to their corresponding input Stokes parameters for a given random realization of the random mode coupling due to PMD. In addition to PDL and PDG, the EDFAs add unpolarized optical noise to the system that is dependent on the amplifier gain. Therefore, an EDFA will also produce the following change in the Stokes parameters of the noise:

$$\begin{aligned} S^{(m)}_{0,\text{noise,after}} &= G\, S^{(m)}_{0,\text{noise,before}} + 2\, n_{\text{sp}}(G-1)B^{(m)}h\nu \\ \mathbf{S}^{(m)}_{\text{noise,after}} &= G\, \mathbf{S}^{(m)}_{\text{noise,before}}, \end{aligned} \tag{3.15}$$

where n_{sp} is the spontaneous emission factor, G is the amplifier gain, which compensates for the losses in the fiber, $h\nu$ is the energy of a single photon, and $B^{(m)}$ is the optical bandwidth of the m-th channel. The gain G is determined using the gain saturation model that I described in Sec. 2.6.

The reduced Stokes model only gives the optical signal-to-noise ratio (OSNR) and the polarization states of the signal and the noise of each channel in the receiver. However, one must compute the Q-factor in order to compute the Q-penalty due to

the polarization effects. In addition to the OSNR and the polarization states of the signal and the noise, a receiver model must also account for the modulation format and the shapes of the optical and the electrical filters in the receiver to compute the Q-factor and, consequently, the Q-penalty produced by the polarization effects in each channel. In Chapter 4, I describe the receiver model that I developed, which can be used in combination with the reduced Stokes model to accurately compute penalties due to the combination of polarization effects and gain saturation in the amplifiers.

Chapter 4

Accurate receiver model

4.1 Performance measures

A fundamental problem in the design of optical fiber transmission systems is to achieve a desired bit-error ratio (BER), with a given outage probability, after a signal is transmitted through the system. The outage probability is the probability that the penalty due to one of the system impairments, such as the polarization effects, exceeds a specified margin. Typically, optical fiber transmission systems are designed to operate with a BER below 10^{-9}. To work properly, these systems must operate with a BER substantially smaller than 10^{-9} to reduce the probability of exceeding this BER limit. As a consequence, direct BER measurements may require hours of transmission experiments [15], and direct BER computation may require days of standard Monte Carlo simulations [43]. Since direct BER measurements can be difficult to make, a widely used performance measure is the Q-factor [18]. The Q-factor is a function of the means and the standard deviations of the received electric currents in the marks and in the spaces [44], which can be viewed experimentally in the time domain using an oscilloscope; it can be used to give an approximate value

for the BER under the assumption that the electric currents in the marks and in the spaces at the receiver are both Gaussian-distributed. Even though the actual distributions of the marks and of the spaces are generalized χ^2 distributions, rather then being Gaussian distributions, this approach can provide a good estimate of the BER in many cases [15], [19]. The optical signal-to-noise ratio (OSNR) is another commonly used performance indicator that is even easier to measure than the Q-factor [19], [30]. However, the relationship between the OSNR and the Q-factor is not straightforward, since the Q-factor depends on the shape of the optical pulses after transmission, the relationship between the Stokes vectors of the signal and the noise, and on the characteristics of the receiver.

The design and performance evaluation of optical fiber communication systems relies just as much on the accuracy and efficiency of receiver models as it does on accurate and efficient modeling of the transmission line [20]. Accurate receiver modeling is especially important when comparing modulation formats. Marcuse [18] and Humblet and Azizoğlu [19] derived widely used approximate expressions for the Q-factor as a function of the signal-to-noise ratio (SNR) of the electric current and as a function of the OSNR, respectively. In both [18] and [19], the authors assume that the receiver consists of a rectangular optical filter, a square-law photodetector, and an integrate-and-dump electrical filter. They also assume that the optical signals have a perfect extinction ratio and that the optical noise is Gaussian and white prior to the optical filter.

I generalize Marcuse's and Humblet and Azizoğlu's results by deriving a formula for the Q-factor in terms of the OSNR for an optical signal in a single polarization

state and an arbitrary pulse shape immediately prior to the receiver and an arbitrary optical extinction ratio in the spaces, for arbitrary optical and electrical filters in the receiver, and for arbitrarily polarized noise. To do so, I calculate the moments of the electric current in the receiver, using an approach that was introduced earlier by Winzer, *et al.* [20] to calculate the BER. In [20], Winzer *et al.* investigated the performance advantage of impulsive amplitude-shift-keyed formats. They limited their study to systems with unpolarized optical noise.

I extend the method in [20] to account for arbitrarily polarized noise. The polarization state of the noise can be significantly affected by the presence of polarization-dependent components, such as components with polarization-dependent loss (PDL) and polarization-dependent gain (PDG) [7]. I will demonstrate that, for a fixed OSNR, the Q-factor can vary widely depending both on the degree of polarization (DOP) of the noise and on the angle between the Stokes vectors of the signal and the noise. The formula for the Q-factor that I present in this dissertation can be used in combination with reduced models of the transmission line to accurately quantify how the performance of the system depends on the combined effects of polarization-mode dispersion (PMD), PDL, PDG, and the gain saturation of optical amplifiers. I will also show that the use of a receiver model that does not account for the polarization state of the noise, as in [8], leads to an inaccurate computation of the outage probability due to polarization effects.

In order to correctly account for the pulse shape in the formula for the Q-factor, I define an enhancement factor. This factor explicitly quantifies how efficiently the combination of a pulse shape and a receiver translates the OSNR into the SNR

of the electric current in the receiver. I validate the formula for the Q-factor by comparison with both experiments and Monte Carlo simulations. I also use the formula for the Q-factor to explicitly quantify the advantage in the receiver of using a chirped return-to-zero (CRZ) modulation format rather than a return-to-zero (RZ) or a non-return-to-zero (NRZ) format with the same mean optical power and receiver characteristics [45]. Just as in [18] and [19], the computational cost of computing the Q-factor using the formula that I propose is orders of magnitude less than the cost of accurately computing the Q-factor in the time domain using Monte Carlo simulations. However, this model does not take into account nonlinear signal-noise interactions during transmission that can both amplify and color the noise prior to the receiver in long-haul fiber transmission systems [43], [46].

In Sec. 4.2, I derive the formula for the Q-factor. In (4.39), I express the Q-factor in terms of the SNR of the electric current, and in (4.51), I express the Q-factor in terms of the OSNR and the polarization state of the noise relative to that of the signal. In Appendix A, I show the derivation of the integral expressions that I use in this chapter. In Secs. 4.3 and 4.4, I present a validation of the receiver model discussed in Sec. 4.2 by comparison to Monte Carlo simulations of the receiver and back-to-back experiments, respectively. In both sections, the validation is performed with unpolarized optical noise and with partially polarized noise prior to the receiver. Finally, in Sec. 4.5, I use the receiver model to compare the performance of different modulation formats.

4.2 Receiver model

4.2.1 Electric current in the receiver

In this section, I derive an expression that relates the Q-factor to the OSNR, and I introduce the enhancement factor. The enhancement factor quantifies the relative performance of different modulation formats and receivers [8]. I begin by recalling the expressions for the mean and variance of the electrically filtered current in the receiver as in [20], which I generalize to account for arbitrarily polarized optical noise. I emphasize that one cannot simply compute the variance of the electric current due to the noise beating with itself by summing the variance of the current that is produced by the noise component that is co-polarized with the signal with the variance of the current that is produced by the noise component that is orthogonal to the polarization state of the signal. This approach is not correct because the components of the noise that are parallel and perpendicular in Jones space to the signal may be correlated if the noise is partially polarized. For example, if the noise is completely polarized in the $+45^\circ$ linear polarization state, which corresponds to the unit Jones vector $\left(1/\sqrt{2},\, 1/\sqrt{2}\right)$, the noise can also be represented as the sum of two components with the same power in the horizontal and the vertical linear polarization states. Nonetheless, one cannot compute the variance of the electric current due to noise by summing the variances of the current that are produced by the noise in the horizontal and the vertical polarization components because these two components are not uncorrelated in general. However, the coherency matrix of the noise is a Hermitian matrix, which implies that there is always an orthogonal basis for Jones space in

which the two orthogonal components of the noise are uncorrelated. Therefore, one may derive formulae for the mean and the variance of the electric current in the receiver by expressing both the signal and the noise in an orthonormal basis for Jones space in which the two orthogonal components of the noise are uncorrelated. In this basis, the variance of the electric current due to the noise is equal to the sum of the variances of the current that are produced by each of the two orthogonal polarization components of the noise.

At the receiver, I assume that the noise from the optical amplifiers is the dominant source of noise, as is the case in an optical communications system with an optically preamplified receiver [39]. Prior to the optical filter in the receiver, I assume that the signal is polarized and that the noise is a Gaussian white noise process that has been generated by optical amplifiers. I let $\mathbf{e}_s(t)$ and $\mathbf{e}_n(t)$ denote the Jones vectors of the electric field envelopes of the signal and noise respectively prior to the receiver, where t is time. Since most optical transmission systems have polarization-dependent components that can affect the polarization state of the noise, I assume that the optical noise entering the receiver has an arbitrary polarization state.

The receiver model that I use consists of an optical filter with a transfer function $H_o(\omega)$ and corresponding impulse response $h_o(t)$, a square-law photodetector, and an electrical filter with a transfer function $H_e(\omega)$ and corresponding impulse response $h_e(t)$. The transfer function $H_e(\omega)$ can account for linear effects in other electrical components of the receiver, including the photodetector. Then the electrical current at the detection point in the receiver is given by

$$i\,(t) \;=\; R\,|[\mathbf{e}_s(t) + \mathbf{e}_n(t)] * h_o(t)|^2 * h_e(t), \tag{4.1}$$

where R is the responsivity of the photodetector and the convolution of two arbitrary functions $g(t)$ and $h(t)$ is defined by $g(t) * h(t) = \int_{-\infty}^{+\infty} g(\tau)h(t-\tau)\mathrm{d}\tau$.

4.2.2 Noise correlation

I assume that the optical noise prior to the optical filter at the receiver is a Gaussian white noise process that has been generated by optical amplifiers. Therefore, the optical noise is delta-correlated with independent and identically distributed real and imaginary Gaussian probability density functions with zero mean [47]. Hence, the autocorrelation function of the optical noise is given by

$$\langle \mathbf{e}_n(t) \cdot \mathbf{e}_n^*(t') \rangle \;=\; N_{\mathrm{ASE}}\, \delta\,(t - t')\,, \tag{4.2}$$

where $\mathbf{v} \cdot \mathbf{w}^* = v_1\, w_1^* + v_2\, w_2^*$ is the standard Hermitian inner product of two Jones vectors $\mathbf{v} = [v_1, v_2]^t$ and $\mathbf{w} = [w_1, w_2]^t$, which is independent of the choice of orthonormal basis of Jones space. The bracket $\langle \cdot \rangle$ indicates an average over all noise realizations, and N_{ASE} is the total power spectral density of the noise prior to the receiver. The effect of the optical filter on the input light is mathematically equivalent to the convolution of the impulse response of the optical filter $h_o(t)$ with the Jones vector of the input light. Therefore, the optically filtered noise $\mathbf{e}_{n_o}(t)$ can be defined as $\mathbf{e}_{n_o}(t) = \mathbf{e}_n(t) * h_o(t)$. Using (4.2) and the impulse response of the optical filter $h_o(t)$, I find that the autocorrelation function of the optically filtered noise $\mathbf{e}_{n_o}(t)$ is given by

$$\langle \mathbf{e}_{n_o}(t) \cdot \mathbf{e}_{n_o}^*(t') \rangle \;=\; N_{\mathrm{ASE}}\, r_o\,(t' - t)\,, \tag{4.3}$$

where

$$r_o(\tau) = \int_{-\infty}^{+\infty} h_o(\tau')\, h_o^*(\tau + \tau')\mathrm{d}\tau', \tag{4.4}$$

is the autocorrelation function of the optical filter. The quantity

$$r_o(0) = \int_{-\infty}^{+\infty} |h_o(\tau')|^2\, \mathrm{d}\tau' \tag{4.5}$$

is the power-equivalent spectral width of the optical filter [24], which we also denote by B_o.

To account for the effect of the polarization-dependent components along the transmission, I must first derive expressions for the autocorrelation function of the noise in the two orthogonal polarization state components that are uncorrelated. To do so, I compute the temporal coherency matrix [24], which is a 2×2 complex Hermitian matrix that mathematically describes both the time correlation and the polarization state of light. Assuming that the optical noise entering the receiver has an arbitrary polarization state, the temporal coherency matrix $\mathsf{J}_n(\tau)$ of the optically filtered noise is defined by

$$\mathsf{J}_n(\tau) \;=\; \begin{bmatrix} \langle e_{n_{o,x}}(t)\, e^*_{n_{o,x}}(t+\tau)\rangle & \langle e_{n_{o,x}}(t)\, e^*_{n_{o,y}}(t+\tau)\rangle \\ \langle e_{n_{o,y}}(t)\, e^*_{n_{o,x}}(t+\tau)\rangle & \langle e_{n_{o,y}}(t)\, e^*_{n_{o,y}}(t+\tau)\rangle \end{bmatrix}, \tag{4.6}$$

where $e_{n_{o,x}}(t)$ and $e_{n_{o,y}}(t)$ are the components of the Jones vector of the optically filtered electric field of the noise $\mathbf{e}_{n_o}(t)$ in an orthonormal basis $\{\hat{\mathbf{a}}_x, \hat{\mathbf{a}}_y\}$ of Jones space, and τ is a time delay between the field components. I assume that the noise process is stationary. Therefore, $\mathsf{J}_n(\tau)$ does not depend on the time t. Assuming that the differential-group delay between two orthogonal components of the complex envelope of the noise field due to PMD in the transmission line is small compared to

the width of the impulse response of the optical filter $h_o(t)$, the PMD will not cause a significant depolarization of the polarized component of the noise in the bandwidth of the optical filter. As a consequence, all the elements of the temporal coherency matrix $\mathsf{J}_n(\tau)$ have approximately the same time delay dependence $r_o(\tau)$ given by (4.3). In this case, the state of polarization of the filtered noise can be represented by the coherency matrix $\mathsf{J}_n = \mathsf{J}_n(0)$ of the optically filtered noise defined as [25]

$$\mathsf{J}_n \;=\; \left[\begin{array}{cc} \langle e_{n_{o,x}}(t)\, e^*_{n_{o,x}}(t)\rangle & \langle e_{n_{o,x}}(t)\, e^*_{n_{o,y}}(t)\rangle \\ \langle e_{n_{o,y}}(t)\, e^*_{n_{o,x}}(t)\rangle & \langle e_{n_{o,y}}(t)\, e^*_{n_{o,y}}(t)\rangle \end{array} \right], \tag{4.7}$$

and the temporal coherency matrix becomes

$$\mathsf{J}_n(\tau) \;=\; \frac{r_o(\tau)}{r_o(0)}\, \mathsf{J}_n. \tag{4.8}$$

The elements $J_{n_{xx}} = \langle e_{n_{o,x}}(t)\, e^*_{n_{o,x}}(t)\rangle$ and $J_{n_{yy}} = \langle e_{n_{o,y}}(t)\, e^*_{n_{o,y}}(t)\rangle$ in (4.7) are the intensity of the filtered noise in the $\hat{\mathbf{a}}_x$ and $\hat{\mathbf{a}}_y$ polarizations, respectively, while $J_{n_{xy}} = \langle e_{n_{o,x}}(t)\, e^*_{n_{o,y}}(t)\rangle$ is a measure of the correlation between the components of the electric field in the $\hat{\mathbf{a}}_x$ and $\hat{\mathbf{a}}_y$ polarizations [25]. The optical intensity of the filtered noise $I_{\rm tot}$ is equal to the trace of the matrix J_n,

$$I_{\rm tot} = \mathrm{Tr}\, \mathsf{J}_n = J_{n_{xx}} + J_{n_{yy}} = \langle e_{n_{o,x}}(t)\, e^*_{n_{o,x}}(t)\rangle + \langle e_{n_{o,y}}(t)\, e^*_{n_{o,y}}(t)\rangle = N_{\rm ASE}\, B_o, \tag{4.9}$$

and the degree of polarization of the filtered noise is given by

$$\mathrm{DOP}_n = \frac{I_{\rm pol}}{I_{\rm tot}} = \left(1 - \frac{4\,\mathrm{Det}\,\mathsf{J}_n}{(\mathrm{Tr}\,\mathsf{J}_n)^2}\right)^{1/2}, \tag{4.10}$$

where $I_{\rm pol}$ is the optical intensity of the polarized part of the filtered noise [25]. Note that the quantities $I_{\rm pol}$, $I_{\rm tot}$ and DOP_n do not depend on the choice of orthonormal basis of Jones space.

The polarization state of the optically filtered noise can also be characterized by the Stokes parameters $\boldsymbol{\mathcal{S}}_n = [S_{n_0}, S_{n_1}, S_{n_2}, S_{n_3}]^t$ of the noise, where S_{n_0} is the average power of the noise after it is filtered by an optical filter in the receiver. In Chapters 2 and 3, I showed how the Stokes parameters of the light can be computed from the Jones vector. The filtered noise can be decomposed as the sum of a polarized part with Stokes parameters $[|\mathbf{S}_n|, S_{n_1}, S_{n_2}, S_{n_3}]^t$ and an unpolarized part with Stokes parameters $[S_{n_0} - |\mathbf{S}_n|, 0, 0, 0]^t$, where $\mathbf{S}_n$ is the Stokes vector of the filtered noise. The DOP of the noise is the power ratio of the polarized part of the noise to the total noise that I introduced in (4.10), which can be expressed using the Stokes parameters by $\mathrm{DOP}_n = |\mathbf{S}_n| / S_{n_0}$. The Stokes parameters of the optically filtered noise can be expressed in terms of the noise coherency matrix by the formula

$$\begin{aligned} S_{n_0} &= J_{n_{xx}} + J_{n_{yy}}, \\ S_{n_1} &= J_{n_{xx}} - J_{n_{yy}}, \\ S_{n_2} &= J_{n_{xy}} + J_{n_{yx}}, \\ S_{n_3} &= i\left(J_{n_{yx}} - J_{n_{xy}}\right). \end{aligned} \tag{4.11}$$

Since the coherency matrix J_n is Hermitian, there is an orthonormal basis $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2\}$ of Jones space in which J_n is diagonal with $J_{n_{11}} \geq J_{n_{22}}$ [25]. This basis simply consists of the unit length eigenvectors of J_n. In the corresponding frame for Stokes space, the Stokes vector of the filtered noise is $\mathbf{S}_n = [S_{n_0}\,\mathrm{DOP}_n, 0, 0]^t$. Consequently, in Jones space the unit vectors $\hat{\mathbf{a}}_1$ and $\hat{\mathbf{a}}_2$ are respectively parallel and perpendicular to the polarized part of the filtered noise. In the basis $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2\}$, the electric field of the

filtered noise is given by

$$\mathbf{e}_{n_o}(t) = e_{n_{o,1}}(t)\,\hat{\mathbf{a}}_1 + e_{n_{o,2}}(t)\,\hat{\mathbf{a}}_2, \tag{4.12}$$

where the components $e_{n_{o,1}}(t)$ and $e_{n_{o,2}}(t)$ of the filtered noise are uncorrelated, since J_n is diagonal. Moreover, in this basis, (4.10) simplifies to

$$\mathrm{DOP}_n = \left(J_{n_{1\,1}} - J_{n_{2\,2}}\right) / \left(J_{n_{1\,1}} + J_{n_{2\,2}}\right). \tag{4.13}$$

Using (4.6)–(4.10) and (4.13), I find that

$$\begin{aligned} J_{n_{1\,1}}(\tau) &= \frac{1}{2}\left(1 + \mathrm{DOP}_n\right) N_{\mathrm{ASE}}\, r_o(\tau), \\ J_{n_{2\,2}}(\tau) &= \frac{1}{2}\left(1 - \mathrm{DOP}_n\right) N_{\mathrm{ASE}}\, r_o(\tau), \end{aligned} \tag{4.14}$$

and $J_{n_{1\,2}}(\tau) = 0$. Therefore, the autocorrelation functions of the components of the optically filtered noise are

$$\left\langle e_{n_{o,1}}(t)\, e^*_{n_{o,1}}(t')\right\rangle \;=\; \frac{1}{2}(1 + \mathrm{DOP}_n) N_{\mathrm{ASE}}\, r_o\left(t' - t\right) \tag{4.15}$$

and

$$\left\langle e_{n_{o,2}}(t)\, e^*_{n_{o,2}}(t')\right\rangle \;=\; \frac{1}{2}(1 - \mathrm{DOP}_n) N_{\mathrm{ASE}}\, r_o\left(t' - t\right), \tag{4.16}$$

while the cross correlation is

$$\left\langle e_{n_{o,1}}(t)\, e^*_{n_{o,2}}(t')\right\rangle \;=\; 0, \tag{4.17}$$

where N_{ASE} is the total power spectral density of the noise prior to the receiver that was introduced in (4.2).

4.2.3 Moments of the electric current

Since I am assuming that the signal is polarized, I can express the Jones vector of the optically filtered signal $\mathbf{e}_{s_o}(t) = \mathbf{e}_s(t) * h_o(t)$ as $\mathbf{e}_{s_o}(t) = e_{s_o}(t)\,\hat{\mathbf{e}}_{s_o}$, where $e_{s_o}(t)$ is a scalar field and $\hat{\mathbf{e}}_{s_o}$ is a unit Jones vector. The expression for $\mathbf{e}_{s_o}(t)$ in the basis $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2\}$ is

$$\mathbf{e}_{s_o}(t) = e_{s_o}(t)\ (\hat{\mathbf{e}}_{s_o} \cdot \hat{\mathbf{a}}_1^*)\ \hat{\mathbf{a}}_1 + e_{s_o}(t)\ (\hat{\mathbf{e}}_{s_o} \cdot \hat{\mathbf{a}}_2^*)\ \hat{\mathbf{a}}_2. \tag{4.18}$$

Physically, the term $e_{s_o}(t)\ (\hat{\mathbf{e}}_{s_o} \cdot \hat{\mathbf{a}}_1^*)\ \hat{\mathbf{a}}_1$ is the component of the Jones vector of the filtered signal that is in the same polarization state as the polarized part of the noise, while $e_{s_o}(t)\ (\hat{\mathbf{e}}_{s_o} \cdot \hat{\mathbf{a}}_2^*)\ \hat{\mathbf{a}}_2$ is the component of the Jones vector of filtered signal that is orthogonal to the polarized part of the noise.

I now substitute the expressions for $\mathbf{e}_{n_o}(t)$ and $\mathbf{e}_{s_o}(t)$ in (4.12) and (4.18) into (4.1) to obtain

$$i\,(t) \;=\; R\left\{\left|(\hat{\mathbf{e}}_{s_o} \cdot \hat{\mathbf{a}}_1^*)\,e_{s_o}(t) + e_{n_o,1}(t)\right|^2 + \left|(\hat{\mathbf{e}}_{s_o} \cdot \hat{\mathbf{a}}_2^*)\,e_{s_o}(t) + e_{n_o,2}(t)\right|^2\right\} * h_e(t). \tag{4.19}$$

In order to compute the mean $\langle i_n \rangle(t)$ and the variance ${\sigma_i}^2(t)$ of the current at any time t, I use the statistical properties of the optically filtered noise that I described earlier in this section. In Appendix A, I present a detailed derivation of the integral expressions of the mean and the variance of the electric current in (4.22), (4.24), and (4.32).

Combining (4.15)–(4.17) with (4.19), I find that

$$\langle i \rangle(t) \;=\; i_s(t) + \langle i_n \rangle(t), \tag{4.20}$$

where $\langle\cdot\rangle(t)$ is the average over the statistical realizations of the noise at time t, and

$$i_s(t) = R\,|\mathbf{e}_{s_o}(t)|^2 * h_e(t) \tag{4.21}$$

is the electric current due to the signal. Furthermore,

$$\langle i_n\rangle(t) \;=\; \langle i_n\rangle \;=\; R\,N_{\mathrm{ASE}}\,B_o \tag{4.22}$$

is the mean current due to noise, which is time-independent, since the optical noise is a stationary random process. The parameter B_o is the power-equivalent spectral width [24] of the optical filter. If, for example, the noise has a power spectral density of 1 W/Hz, then after passing an optical filter with power-equivalent spectral width B_o (in Hz) the optical power is exactly equal to B_o (in W), regardless of the shape of the optical filter. As a consequence, the power-equivalent spectral width is a measure of the optical bandwidth that is more useful for the study of optical noise statistics in this context than other more traditional bandwidth measures, such as the full-width at half maximum and the root-mean-square width. In (4.22), I have neglected the gain and attenuation in the optical filter and in the electrical circuit of the receiver, since they do not affect the SNR. Therefore, the absolute value of the transfer functions of the optical and the electrical filters have the value 1 at the central frequency of the channel, $|H_o(0)| = |H_e(0)| = 1$. Therefore, $B_o = r_o(0) = \int_{-\infty}^{+\infty} |h_o(\tau)|^2 \mathrm{d}\tau$.

Following a similar procedure used in the derivation of (4.22), I find that the variance of the current at any time t has the form

$$\sigma_i{}^2(t) \;=\; \langle i^2\rangle(t) - \langle i\rangle^2(t) \;=\; \sigma^2_{\mathrm{ASE-ASE}} + \sigma^2_{\mathrm{S-ASE}}(t). \tag{4.23}$$

The first term on the right side of (4.23) is the variance of the current due to the noise beating with itself at the receiver—the noise-noise beating. The variance of the current due to the noise-noise beating is given by

$$\sigma^2_{\mathrm{ASE-ASE}} = \frac{1}{2} R^2 \, N^2_{\mathrm{ASE}} \, \frac{I_{\mathrm{ASE-ASE}}}{\Gamma_{\mathrm{ASE-ASE}}}, \tag{4.24}$$

where

$$\Gamma_{\mathrm{ASE-ASE}} = \frac{1}{1 + \mathrm{DOP}_n^2}, \tag{4.25}$$

and

$$I_{\mathrm{ASE-ASE}} = \int_{-\infty}^{+\infty} |r_o(\tau)|^2 \, r_e(\tau) \mathrm{d}\tau. \tag{4.26}$$

Here, the expression

$$r_e(\tau) = \int_{-\infty}^{+\infty} h_e(\tau') \, h_e(\tau + \tau') \mathrm{d}\tau' \tag{4.27}$$

is the autocorrelation function of the electrical filter. The noise-noise beating factor $\Gamma_{\mathrm{ASE-ASE}}$ is the ratio between the variance of the current due to noise-noise beating in the case that the noise is unpolarized and the actual variance of the current due to noise-noise beating. The second term on the right side of (4.23) is the variance of the current due to the beating between the signal and the noise at the receiver—the signal-noise beating. The variance of the current due to the signal-noise beating is given by

$$\sigma^2_{\mathrm{S-ASE}}(t) = \frac{1}{2} R^2 N_{\mathrm{ASE}} \left[(1 + \mathrm{DOP}_n) \, |\hat{\mathbf{e}}_{s_o} \cdot \hat{\mathbf{a}}_1^*|^2 + (1 - \mathrm{DOP}_n) \, |\hat{\mathbf{e}}_{s_o} \cdot \hat{\mathbf{a}}_2^*|^2 \right] I_{\mathrm{S-ASE}}(t), \tag{4.28}$$

where

$$I_{\mathrm{S-ASE}}(t) = 2 \int_{-\infty}^{+\infty} e_{s_o}(\tau) h_e(t - \tau) \int_{-\infty}^{+\infty} e^*_{s_o}(\tau') h_e(t - \tau') r_o(\tau - \tau') \mathrm{d}\tau' \mathrm{d}\tau. \tag{4.29}$$

The integral expressions in (4.26) and (4.29) may be derived following a procedure similar to the one in [20]. The terms $|\hat{\mathbf{e}}_{s_o} \cdot \hat{\mathbf{a}}_1^*|^2$ and $|\hat{\mathbf{e}}_{s_o} \cdot \hat{\mathbf{a}}_2^*|^2$ are the relative intensities of the components of the field $e_{s_o}(t)\,\hat{\mathbf{e}}_{s_o}$ that are respectively parallel and perpendicular in Jones space to the polarized part of the optically filtered noise. These terms can be represented in Stokes space as

$$|\hat{\mathbf{e}}_{s_o} \cdot \hat{\mathbf{a}}_1^*|^2 = \frac{1}{2}\left(1 + \mathbf{s}_s \cdot \mathbf{s}_n^{(p)}\right), \tag{4.30}$$

and

$$|\hat{\mathbf{e}}_{s_o} \cdot \hat{\mathbf{a}}_2^*|^2 = \frac{1}{2}\left(1 - \mathbf{s}_s \cdot \mathbf{s}_n^{(p)}\right), \tag{4.31}$$

where $\mathbf{s}_s$ and $\mathbf{s}_n^{(p)}$ are the normalized Stokes vectors of the signal and the polarized part of the noise respectively. The dot products on the right-hand sides of (4.30) and (4.31) are the standard dot product between two real three vectors, which can also be mathematically represented by the definition given after (4.2). Substituting (4.30) and (4.31) into (4.28) one finds that

$$\sigma^2_{\mathrm{S-ASE}}(t) = R^2 N_{\mathrm{ASE}}\,\Gamma_{\mathrm{S-ASE}}\,I_{\mathrm{S-ASE}}(t), \tag{4.32}$$

where

$$\Gamma_{\mathrm{S-ASE}} = \frac{1}{2}\left[1 + \mathrm{DOP}_n\left(\mathbf{s}_s \cdot \mathbf{s}_n^{(p)}\right)\right], \tag{4.33}$$

is the signal-noise beating factor, which is the fraction of the noise that beats with the signal.

4.2.4 Clock recovery

To define the sampling time, I recover the clock throughout this dissertation using algorithms based on ones described by Trischitta and Varma [48]. For RZ systems,

the sampling time is set equal to the time at which the phase of the 10 GHz tone of the electrical current is equal to zero or a multiple of 2π. This method is implemented in receivers with the use of a narrowband electrical bandpass filter centered at 10 GHz. In NRZ systems, on the other hand, the 5 GHz tone could, in principle, be used to recover the clock. However, the 5 GHz tone in NRZ systems is not as strong as the 10 GHz tone in RZ systems. To recover the clock in NRZ systems, I delay the bit stream by half bit slot and subtract it from the original stream, the result of which is then squared. The resulting signal has a strong tone at 10 GHz. The sampling time is set equal to the time at which the phase of the tone is equal to $\pi/2$. In this dissertation, I only study systems that operate at 10 Gbit/s. However, these clock recovery algorithms can be applied to systems with different bit rates by rescaling the frequency of the tone that is used to recover the clock.

4.2.5 The Q-factor and the enhancement factor

I now use the expressions for the mean and the variance of the current that I derived in Sec. 4.2.3 to derive a general expression for the Q-factor as a function of the OSNR. I follow a procedure that is similar to the one described in [18], but I use the exact mean and the exact variance of the electric current. I start with the standard time-domain definition of the Q-factor,

$$Q = \frac{\langle i_1 \rangle - \langle i_0 \rangle}{\sigma_1 + \sigma_0} \tag{4.34}$$

as in [18], which is related to the BER by

$$\mathrm{BER} = \mathrm{erfc}(Q/\sqrt{2})/2 \approx \exp(-Q^2/2)/(\sqrt{2\pi}Q), \tag{4.35}$$

provided that the electric currents in the marks and in the spaces are Gaussian distributed.

In this section, I will assume that there are no patterning effects, either in transmission or in the receiver. Consequently, the current of each mark obeys the same probability density function, regardless of the neighboring bits, as does the current of the spaces. In Sec 5.2, I extend these results to describe how one can accurately compute the Q-factor in the presence of pattern dependencies. After clock recovery, as described in Sec. 4.2.4, the detection times of the marks may be written t_{1_j}, $j = 1, \ldots, N_1$, where N_1 is the number of marks. Likewise, the detection times of the spaces may be written t_{0_j}, $j = 1, \ldots, N_0$, where N_0 is the number of spaces. The assumption that there are no patterning effects implies that $i_s(t_{1_1}) = i_s(t_{1_2}) = \ldots = i_s(t_{1_{N_1}}) \equiv i_s(t_1)$, where $i_s(t_1)$ is the common value of all the $i_s(t_{1_j})$. Similarly, one may write $\sigma^2_{\text{S-ASE}}(t_{1_j}) \equiv \sigma^2_{\text{S-ASE}}(t_1)$, $j = 1, \ldots, N_1$, and $i_s(t_{0_j}) \equiv i_s(t_0)$, $\sigma^2_{\text{S-ASE}}(t_{0_j}) \equiv \sigma^2_{\text{S-ASE}}(t_0)$, $j = 1, \ldots, N_0$. Substituting (4.20) and (4.23) into (4.34), I now obtain

$$Q = \frac{[i_s(t_1) + \langle i_n \rangle] - [i_s(t_0) + \langle i_n \rangle]}{\left(\sigma^2_{\text{S-ASE}}(t_1) + \sigma^2_{\text{ASE-ASE}}\right)^{1/2} + \left(\sigma^2_{\text{S-ASE}}(t_0) + \sigma^2_{\text{ASE-ASE}}\right)^{1/2}}, \tag{4.36}$$

where t_1 and t_0 are sampling times of a mark and a space, respectively. To determine the sampling times, I used the clock recovery algorithm that I described in Sec. 4.2.4.

By rearranging (4.36), the Q-factor can be expressed in terms of the SNR of a mark,

$$\text{SNR}_1 = \frac{i_s(t_1)}{\langle i_n \rangle}, \tag{4.37}$$

and the extinction ratio of the electric current in the receiver,

$$\alpha_e = \frac{i_s(t_0)}{i_s(t_1)}, \tag{4.38}$$

as

$$Q = \frac{(1-\alpha_e)\,\mathrm{SNR}_1\,M^{1/2}}{(K_1\,\mathrm{SNR}_1+1)^{1/2}+(K_0\,\alpha_e\,\mathrm{SNR}_1+1)^{1/2}}, \tag{4.39}$$

where

$$K_0 = \frac{\langle i_n\rangle\,\sigma^2_{\mathrm{S-ASE}}(t_0)}{i_s(t_0)\,\sigma^2_{\mathrm{ASE-ASE}}} \tag{4.40}$$

and

$$K_1 = \frac{\langle i_n\rangle\,\sigma^2_{\mathrm{S-ASE}}(t_1)}{i_s(t_1)\,\sigma^2_{\mathrm{ASE-ASE}}}. \tag{4.41}$$

We call K_0 and K_1 the signal-noise beating parameters for the marks and spaces, respectively, because they are directly proportional to the ratio between the variance of the current due to the signal-noise beating and the variance of the current due to the noise-noise beating. The parameter

$$M = \frac{\langle i_n\rangle^2}{\sigma^2_{\mathrm{ASE-ASE}}} \tag{4.42}$$

is the effective number of noise modes. I use this terminology because M is a generalization of the number of noise modes parameter that was introduced by Marcuse [18]. In that work, Marcuse found that only a finite number of noise Fourier components—noise modes—contribute to the electric current due to noise in a receiver with a rectangular optical filter and an integrate-and-dump electrical filter.

The parameters K_0, K_1 and M are unitless, and do not depend on the average power of the signal, nor on the average power of the noise. To separate out the

dependence of these three parameters on the polarization state of the noise I use (4.24) and (4.32) to express K_0 and K_1 as

$$K_0 = 2\,\Gamma_{\mathrm{S-ASE}}\,\Gamma_{\mathrm{ASE-ASE}}\,\kappa_0, \tag{4.43}$$

and

$$K_1 = 2\,\Gamma_{\mathrm{S-ASE}}\,\Gamma_{\mathrm{ASE-ASE}}\,\kappa_1, \tag{4.44}$$

where

$$\kappa_0 = \frac{R\,B_o\,I_{\mathrm{S-ASE}}(t_0)}{i_s(t_0)\,I_{\mathrm{ASE-ASE}}}, \tag{4.45}$$

and

$$\kappa_1 = \frac{R\,B_o\,I_{\mathrm{S-ASE}}(t_1)}{i_s(t_1)\,I_{\mathrm{ASE-ASE}}}. \tag{4.46}$$

If the noise is either unpolarized or completely co-polarized with the signal, then K_0 and K_1 are equal to κ_0 and κ_1 respectively, since then $2\,\Gamma_{\mathrm{S-ASE}}\,\Gamma_{\mathrm{ASE-ASE}} = 1$. Similarly, I obtain $M = \Gamma_{\mathrm{ASE-ASE}}\,\mu$, where

$$\mu = \frac{2\,B_o^2}{I_{\mathrm{ASE-ASE}}} \tag{4.47}$$

is equal to M in the case where the noise is unpolarized. Then, the Q-factor can be expressed as

$$Q = \frac{(1-\alpha_e)\,\mathrm{SNR}_1\,\left(\Gamma_{\mathrm{ASE-ASE}}\,\mu\right)^{1/2}}{\left(2\,\Gamma_{\mathrm{S-ASE}}\,\Gamma_{\mathrm{ASE-ASE}}\,\kappa_1\,\mathrm{SNR}_1+1\right)^{1/2}+\left(2\,\Gamma_{\mathrm{S-ASE}}\,\Gamma_{\mathrm{ASE-ASE}}\,\kappa_0\,\alpha_e\,\mathrm{SNR}_1+1\right)^{1/2}}. \tag{4.48}$$

In this formula, the dependence on the polarization state of the noise is accounted for by the noise-noise beating factor $\Gamma_{\mathrm{ASE-ASE}}$ defined in (4.25) and the signal-noise beating factor $\Gamma_{\mathrm{S-ASE}}$ given in (4.33). These factors have values in the range $1/2 \leq$

$\Gamma_{\mathrm{ASE-ASE}} \leq 1$ and $0 \leq \Gamma_{\mathrm{S-ASE}} \leq 1$. If the noise is unpolarized, $\Gamma_{\mathrm{ASE-ASE}} = 1$ and $\Gamma_{\mathrm{S-ASE}} = 1/2$. If the noise is completely co-polarized with the signal, $\Gamma_{\mathrm{ASE-ASE}} = 1/2$ and $\Gamma_{\mathrm{S-ASE}} = 1$.

The expression for the Q-factor in (4.48) separates the effects of the polarization state of the noise and of the signal from the pulse shape and the receiver characteristics. However, the Q-factor is given as a function of the signal-to-noise ratio of the marks SNR_1 in the receiver, which also depends on the pulse shape and on the characteristics of the receiver. Thus, it is appropriate to express the Q-factor in terms of a quantity that does not depend on the pulse shape nor on the receiver, such as the OSNR. I define the OSNR by

$$\mathrm{OSNR} = \frac{\langle |e_s(t)|^2 \rangle_t}{N_{\mathrm{ASE}}\, B_{\mathrm{OSA}}}, \tag{4.49}$$

where $\langle |e_s(t)|^2 \rangle_t$ is the time-averaged noiseless optical power per channel prior to the optical filter, and B_{OSA} is the power-equivalent spectral width of an optical spectrum analyzer that is used to measure the optical power of the noise. This definition of OSNR is consistent with the definition in [5], [30] and agrees with the OSNR value that can be obtained directly from an optical spectrum analyzer whose resolution bandwidth is large compared to the bandwidth of the signal, provided that the signal power in the resolution bandwidth of the optical spectrum analyzer is much larger than the noise power in the same bandwidth ($\mathrm{OSNR} \gg 1$).

In order to express the Q-factor as a function of the OSNR, I define the enhancement factor ξ as the ratio between the signal-to-noise ratio of the electric current of the marks SNR_1 and the OSNR at the receiver. The enhancement factor can be

expressed as

$$\xi = \frac{\mathrm{SNR}_1}{\mathrm{OSNR}} = \frac{i_s(t_1)}{\langle i_n \rangle} \frac{N_{\mathrm{ASE}}\, B_{\mathrm{OSA}}}{\left\langle |e_s(t)|^2 \right\rangle_t} = \xi' \frac{B_{\mathrm{OSA}}}{B_o}, \tag{4.50}$$

where $\xi' = i_s(t_1) / \left[R \left\langle |e_{\mathrm{in}}(t)|^2 \right\rangle_t \right]$ is the normalized enhancement factor, which is equal to ξ when $B_{\mathrm{OSA}} = B_o$. The enhancement factor quantifies how efficiently the combination of the pulse shape and receiver translates the OSNR into the SNR of the electric current of the marks in the receiver.

Substituting (4.50) into (4.48), I finally obtain an exact expression that relates the Q-factor directly to the OSNR and to the polarization states of the optical noise and of the signal prior to the receiver:

$$Q = \frac{(1-\alpha_e)\,\xi\,\mathrm{OSNR}\,\left(\Gamma_{\mathrm{ASE-ASE}}\,\mu\right)^{1/2}}{\left(2\,\Gamma_{\mathrm{S-ASE}}\,\Gamma_{\mathrm{ASE-ASE}}\,\kappa_1\,\xi\,\mathrm{OSNR}+1\right)^{1/2}+\left(2\,\Gamma_{\mathrm{S-ASE}}\,\Gamma_{\mathrm{ASE-ASE}}\,\kappa_0\,\alpha_e\,\xi\,\mathrm{OSNR}+1\right)^{1/2}}. \tag{4.51}$$

The optical pulse shape prior to the receiver and the shapes and bandwidths of the optical and electrical filters in the receiver are taken into account in the determination of the values for κ_0, κ_1 and μ, as given in (4.45)–(4.47).

4.2.6 Comparison with previous formulae for the Q-factor

In an optical fiber system with the NRZ modulation format that consists of perfectly rectangular pulses with perfect optical extinction ratio ($\alpha_o = 0$), unpolarized optical noise, and a receiver that consists of a rectangular optical filter and an integrate and dump electrical filter, the formula (4.48) for Q becomes

$$Q = \frac{\mathrm{SNR}_1}{\left(\kappa_1\,\mathrm{SNR}_1 + 1\right)^{1/2} + 1} \sqrt{\mu}, \tag{4.52}$$

where $\kappa_1 = 2$, which is the same as the formula for Q in [18]. Using (4.51) and (4.52), I can express the Q-factor in terms of the OSNR as

$$Q = \frac{\xi\,\mathrm{OSNR}}{\left(\kappa_1\,\xi\,\mathrm{OSNR} + 1\right)^{1/2} + 1}\sqrt{\mu}, \tag{4.53}$$

where $\kappa_1 = 2$ and $\xi = 2$, which is the same as the formula for Q in [30] for the same system that I considered in (4.52). In (4.53), $\xi = 2$ because the average optical power in this system is equal to half of the average optical power in the marks.

4.2.7 Numerical efficiency

To efficiently compute the parameters κ_0, κ_1, and μ, one can use Fourier transforms to numerically compute the multiple integrals in (4.26) and (4.29). From the convolution theorem, one obtains

$$I_{\mathrm{ASE-ASE}} = \int_{-\infty}^{+\infty} \left| \mathcal{F}_\tau^{-1}\left\{ \left| \tilde{H}_o \right|^2 \right\} \right|^2 \mathcal{F}_\tau^{-1}\left\{ |H_e|^2 \right\} \mathrm{d}\tau, \tag{4.54}$$

where $\tilde{H}_o(\omega) = H_o(-\omega)$, and $\mathcal{F}_\tau\left[\cdot\right]$ and $\mathcal{F}_\tau^{-1}\left\{\,\cdot\,\right\}$ denote the forward and inverse Fourier transform with respect to τ, while

$$I_{\mathrm{S-ASE}}(t) = 2\,\mathcal{F}_t^{-1}\left\{ H_e\,\mathcal{F}_\tau\left[e_{s_o}(\tau)\mathcal{F}_\tau^{-1}\left\{ \left| \tilde{H}_o \right|^2 \mathcal{F}_{\tau'}\left[e_{s_o}^*(\tau') h_e(t-\tau') \right] \right\} \right] \right\}. \tag{4.55}$$

4.3 Receiver model validation with Monte Carlo simulations

I now present a validation of formula (4.51) for computing the Q-factor from the OSNR by comparison with two sets of Monte Carlo simulations in which the Q-factor

is computed using the standard time domain formula $Q = (\langle i_1 \rangle - \langle i_0 \rangle) / (\sigma_1 + \sigma_0)$. In the first set of validation simulations, I used a back-to-back 10 Gbit/s optical system with unpolarized optical noise, and in the second set of simulations, I used another back-to-back 10 Gbit/s system with partially polarized optical noise. Since my study was focused on the combined effect that the pulse shape and the receiver have on the system performance, I did not include transmission effects here, such as those due to nonlinearity and dispersion.

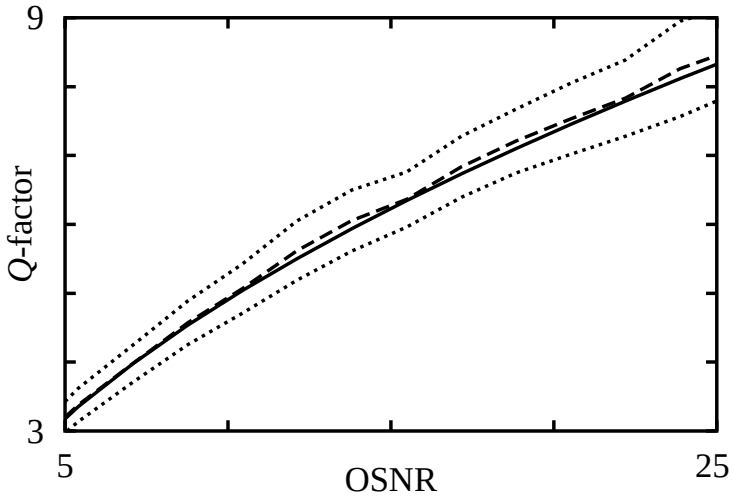

Figure 4.1: Comparison of the formula (4.51) for the Q-factor as a function of the OSNR with the time domain Monte Carlo method of computing the Q-factor for an RZ raised cosine format. The power-equivalent spectral width of the OSA was 25 GHz. For the Monte Carlo simulations, the statistics of the Q-factor were obtained using 100 Q-samples each with 128 bits. The solid line shows the result using (4.51). The dashed line and the two dotted lines show the mean Q-factor for all 100 Q-samples and the confidence interval for a single Q-sample, defined by the mean Q-factor plus and minus one standard deviation computed using the time domain Monte Carlo method.

In Fig. 4.1, I show the Q-factor versus the OSNR for an RZ raised cosine format with a 50% duty cycle and an optical extinction ratio of 18 dB. The electric field

of an RZ raised cosine pulse is given by $e_s(t) = [P_0 \cos^2(\pi t/T)]^{1/2}$, where P_0 is the peak power and T is the bit period. Unpolarized optical noise was added prior to the receiver using a Gaussian noise source that has a constant spectral density within the spectrum of the optical filter. The receiver consisted of a Gaussian-shaped optical filter with a full-width at half maximum (FWHM) of 124 GHz and a fifth-order low-pass electrical Bessel filter with a 3 dB width of 8.5 GHz. The power-equivalent spectral width [24] of the optical spectrum analyzer equaled 25 GHz. The parameters in (4.51) for this system are $\xi = 0.6$, $\alpha_e = -18$ dB, $\kappa_0 = \kappa_1 = 3$, $\mu = 21.23$, $\Gamma_{\mathrm{ASE-ASE}} = 1$, and $\Gamma_{\mathrm{S-ASE}} = 1/2$.

In Fig. 4.1, I show the results using (4.51) with a solid line. My colleagues and I obtained these results using only a single mark and a single space of the transmitted bit string. I show the results for the time-domain Monte Carlo method with a dashed line. We obtained these results by averaging over 100 samples of the Q-factor, where for each sample the means and standard deviations of the marks and spaces were estimated using 128 bits. To determine the sampling time for marks and spaces, we used the clock recovery algorithm that I described in Sec. 4.2.4. The agreement between the two methods is excellent, except for the statistical error in the computation of the Q-factor using the Monte Carlo method. The numerical estimator of the standard deviation of the current due to noise is a biased estimator [49], which contributes to the statistical error in the estimation of the Q-factor using Monte Carlo simulations. The region between the dotted lines in Fig. 4.1 is the confidence interval for the computation of the Q-factor using the time domain Monte Carlo method for a string with 128 bits. The confidence interval, defined by the mean Q-factor plus

and minus one standard deviation of the Q-factor, gives an estimate of the error in the computation of the Q-factor using the time domain Monte Carlo method with a single string of bits. Since we used 100 strings, the actual uncertainty in Fig. 4.1 is ten times smaller. In Fig. 4.1, the time domain Monte Carlo method has a relative statistical uncertainty larger than 15% with a single string when $Q = 6$.

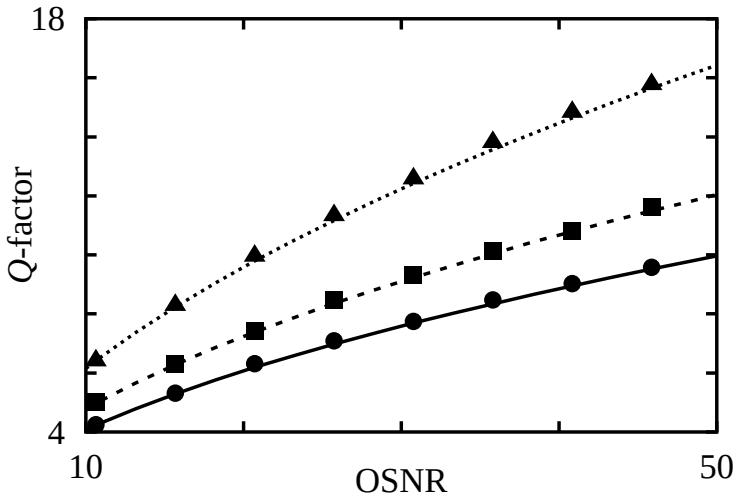

Figure 4.2: Comparison of the formula (4.51) with the time domain Monte Carlo method for computing the Q-factor as a function of the OSNR for the RZ raised cosine format for different noise polarization states with $\mathrm{DOP}_n = 0.5$. These results are for a horizontally polarized optical signal. The power-equivalent spectral width of the optical spectrum analyzer was 25 GHz. The curves show the results obtained using (4.51) and the symbols show the results obtained using Monte Carlo simulations. The solid curve and the circles show the results when the polarized part of the noise is in the horizontal linear polarization state. Similarly, the dashed curve and the squares, and the dotted curve and the triangles show the results when the polarized part of the noise is in the left circular and vertical linear polarization states, respectively.

For the second set of simulations, I used partially polarized noise. In Fig. 4.2, I plot the Q-factor versus the OSNR for a horizontal linearly-polarized RZ raised

cosine signal with an optical extinction ratio of 18 dB. I used a 10 Gbit/s back-to-back system and added partially polarized noise with $\mathrm{DOP}_n = 0.5$ prior to the receiver. The receiver and the optical spectrum analyzer bandwidth were the same as that used for Fig. 4.1. In Fig. 4.2, the curves show the results obtained using (4.51) and the symbols show the results obtained using Monte Carlo simulations. The solid curve and circles show the results when the polarized part of the noise is in the horizontal linear polarization state. In this case the noise is co-polarized with the signal. Similarly, the dashed curve and the squares, and the dotted curve and the triangles show the results when the polarized part of the noise is in the left circular and vertical linear polarized states, respectively. The agreement between (4.51) and Monte Carlo simulations is excellent. When $\mathrm{DOP}_n = 0.5$, the Q-factor varies by about 60% as we vary the polarization state of the noise. This variation occurs because the signal-noise beating factor $\Gamma_{\mathrm{S-ASE}}$ in (4.33) depends on the angle between the Stokes vectors of the signal and the polarized part of the noise. The parameters in (4.51) for this system are the same ones in Fig. 4.1, except that $\Gamma_{\mathrm{ASE-ASE}} = 0.8$ and $\Gamma_{\mathrm{S-ASE}} = 1$ for the solid line, $\Gamma_{\mathrm{S-ASE}} = 0.5$ for the dashed line, and $\Gamma_{\mathrm{S-ASE}} = 0.25$ for the dotted line.

4.4 Receiver model validation with experiments

I now present a validation of formula (4.51) for computing the Q-factor from the OSNR by comparison with two sets of back-to-back 10 Gbit/s experiments. These experiments were carried out primarily by Yu Sun and Hua Jiao in the Optical Fiber Communications Laboratory at the University of Maryland Baltimore County. In

the first set of experiments, the noise is unpolarized ($\mathrm{DOP}_n \approx 0$). In the second set of experiments, for which the noise is partially polarized ($\mathrm{DOP}_n \in [0, 1]$), one finds a strong dependence of the Q-factor on the polarization state of the noise when the noise is partially polarized. In both cases, the experiments agree well with the simulations.

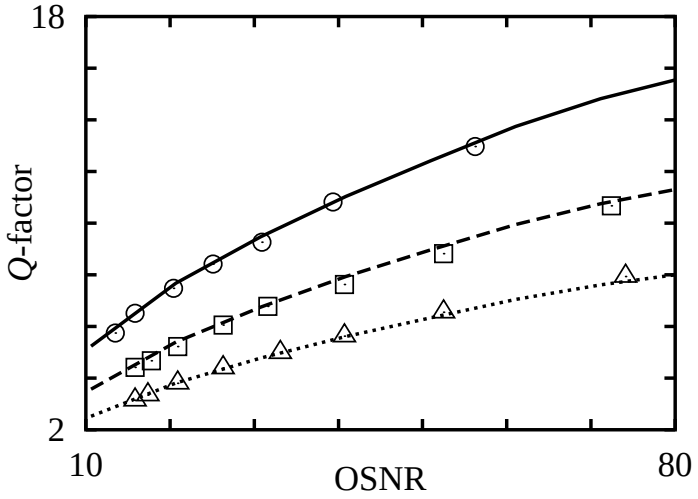

Figure 4.3: Comparison of the Q-factor as a function of the OSNR obtained using (4.51) with experimental results for different modulation formats and receivers. The power-equivalent spectral width of the OSA was 25 GHz. The curves show the results obtained using (4.51) and the experimental results are shown using symbols. The solid curve and circles show the results for an RZ format without the electrical filter. The dashed curve and the squares show the results for an NRZ format with an electrical filter with a 3 dB width of 7 GHz. The dotted curve and the triangles show the results for the NRZ without the electrical filter.

In Fig. 4.3, I plot the Q-factor versus the OSNR obtained using both simulations and experiments for RZ and NRZ signals with unpolarized optical noise ($\mathrm{DOP}_n < 0.05$) that is generated by an erbium-doped fiber amplifier (EDFA) without input power [23]. In the transmitter, we generated a 10 Gbit/s pulse train using an electro-

absorption modulator (EAM). The data was encoded on the pulse train using an electro-optic modulator (EOM). For the NRZ signal, the EAM was bypassed. To avoid pattern dependences in the EOM modulator that was used to encode the data, we used the fixed pattern 01010101. Therefore, these results provide a baseline for future studies that will include pattern dependences. The RZ pulse was Gaussian-shaped with a FWHM of 23 ps, and the NRZ pulses had a rise time of 34 ps. The optical extinction ratio was 18 dB for the RZ signal and 12 dB for the NRZ signal. At the receiver, an optical preamplifier increased the signal and noise power so that the optical noise dominated the electrical noise. The total power into the 20 GHz photodetector was kept fixed at -2 dBm by tuning an attenuator. The FWHM of the Gaussian optical filter was 187 GHz, and either a 7 GHz electrical Bessel filter or no electrical filter was used in the receiver. The power-equivalent spectral width of the optical spectrum analyzer was 25 GHz. A high-speed sampling oscilloscope was used to measure the Q-factor at the same time that the OSNR was measured. In Fig. 4.3, the curves show results obtained using (4.51), and the symbols show the experimental results. The solid curve and circles show the results for the RZ format without the electrical filter. The dashed curve and squares show the results for the NRZ format with the electrical filter, and the dotted curve and triangles show the results for the NRZ format without the electrical filter. In Fig. 4.3, the parameters in (4.51) for the RZ signal without the electrical filter are $\alpha_e = -18$ dB, $\xi = 0.74$ ($\xi' = 5.9$), $\mu = 17.7$, and $\kappa_0 = \kappa_1 = 3.17$. The parameters for the NRZ signal without electrical filter are $\alpha_e = -11.9$ dB, $\xi = 0.24$ ($\xi' = 2$), $\mu = 17.7$, $\kappa_0 = 2.79$, and $\kappa_1 = 2.82$. The parameters for the NRZ signal with the electrical filter are $\alpha_e = -11.3$ dB, $\xi = 0.24$

($\xi' = 1.9$), $\mu = 38.8$, $\kappa_0 = 2.68$, and $\kappa_1 = 2.88$. Since the noise is unpolarized, one obtains $\Gamma_{\mathrm{ASE-ASE}} = 1$ and $\Gamma_{\mathrm{S-ASE}} = 1/2$.

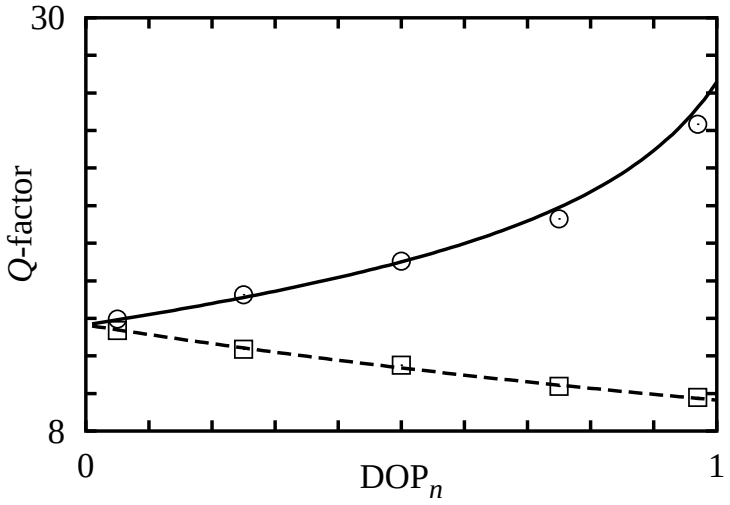

Figure 4.4: Comparison of the Q-factor as a function of the degree of polarization of the noise DOP_n obtained using (4.48) with experiments for different polarization states of the signal and of the noise. The solid curve shows the result obtained using (4.48) and the circles show the experimental results when the Jones vectors of the signal and the polarized part of the noise are orthogonal. The dashed curve shows the result obtained using (4.48), and the squares show the experimental results when the signal is co-polarized with the polarized part of the noise.

In the second set of experiments, we investigated the effect of partially polarized noise on the system performance. In Fig. 4.4, I plot the maximum and the minimum values of the Q-factor as a function of the degree of polarization of the noise at the receiver DOP_n. These simulation and experimental results are for an RZ signal with optical noise added by two erbium-doped fiber amplifier without input power [21]. The RZ pulse was Gaussian-shaped with a FWHM of 23 ps. In the simulations, we used a perfect optical extinction ratio. The receiver consisted of a photodetector, an

optical filter, and an electrical amplifier. We used the method described in [15] to obtain the Q-factor from the BER margin measurements. Bergano, *et al.* [15] showed that the Q-factor obtained from BER margin measurements is well-correlated to the BER. The first optical amplifier generated unpolarized noise, and the noise generated by the second optical amplifier was polarized by passing it through a polarizer and a polarization controller. The degree of polarization of the noise DOP_n was controlled by adjusting two variable attenuators that follow the two optical amplifiers. The direction on the Poincaré sphere of the normalized Stokes vector $\mathbf{s}_n^{(p)}$ of the polarized part of the noise was adjusted by the polarization controller that follows the polarizer. The SNR of the electric current of the marks SNR_1, which was defined in (4.37), was fixed at 10.9 dB. The FWHM of the Gaussian optical filter was 187 GHz, and the electrical bandwidth of the receiver was 10 GHz. In the experiment, for each value of DOP_n, we varied the angle between the Stokes vectors of the signal and the polarized part of the noise and recorded the maximum value $Q_{\max}$, and minimum value $Q_{\min}$ of the Q-factor. In the simulations, I obtained $Q_{\max}$ by choosing the Stokes vectors of the signal and the polarized part of the noise to be antiparallel, $\mathbf{s}_s \cdot \mathbf{s}_n^{(p)} = -1$, so that the corresponding Jones vectors are orthogonal to each other. Similarly, I obtained $Q_{\min}$ by choosing the Stokes vectors $\mathbf{s}_s$ and $\mathbf{s}_n^{(p)}$ to be parallel, so that $\mathbf{s}_s \cdot \mathbf{s}_n^{(p)} = +1$.

In Fig. 4.4, the curves show results obtained using (4.48), and the symbols show the experimental results. The solid curve and circles show the maximum value of the Q-factor versus DOP_n, while the dashed curve and squares show the minimum value of the Q-factor. As the DOP of the noise increases from 0 to 1, one observes a dramatic increase in the range $[Q_{\min}, Q_{\max}]$ of the Q-factor. For this receiver, the

parameters in (4.48) are $\mu = 38.6$, $\kappa_1 = 1.7$, $\kappa_0 = 0$, and $\alpha_e = 0$. When the noise is unpolarized ($\mathrm{DOP}_n = 0$), one has $\Gamma_{\mathrm{ASE-ASE}} = 1$ and $\Gamma_{\mathrm{S-ASE}} = 1/2$. When the noise is completely polarized ($\mathrm{DOP}_n = 1$), one has $\Gamma_{\mathrm{ASE-ASE}} = 1/2$, so that $\Gamma_{\mathrm{S-ASE}} = 1$ when the noise is co-polarized with the signal and $\Gamma_{\mathrm{S-ASE}} = 0$ when the polarization state of the noise is orthogonal to the signal. The results that I show in Figs. 4.2 and 4.4 illustrate the significant impact that partially polarized noise can have on the performance of an optical fiber transmission system. Typical values for the PDL per optical amplifier in optical fiber systems range from 0.1 dB to 0.2 dB [10], which can partially polarize the optical noise in the transmission line. In a prototypical trans-oceanic system described in [8], which has 270 amplifiers with 0.15 dB of PDL per amplifier, the average noise DOP at the receiver is equal to 0.15. However, the noise DOP in that trans-oceanic system exceeds 0.3 with a probability larger than 10^{-3}.

4.5 Modulation format comparisons

In this section, I describe the use of (4.51) to compare the back-to-back performance of three formats—CRZ, RZ, and NRZ—using the same receiver. To separate out the effect of partially polarized noise from the modulation format in the system performance, my colleagues and I only considered unpolarized noise in this study. We found that the CRZ format outperforms both the RZ and the NRZ formats because the CRZ format has a larger enhancement factor. We also found that the RZ format outperforms the NRZ format for the same reason.

In Fig. 4.5, I use (4.51) to plot the Q-factor versus the OSNR for the CRZ, RZ, and NRZ formats. The optical noise is unpolarized. The electric field of a CRZ pulse is given by

$$e_s(t) = \left[P_0 \cos^2(\pi t/T)\right]^{1/2} \exp\left[iA\pi \cos(2\pi t/T)\right], \tag{4.56}$$

where $A = -0.6$ is the chirp parameter, P_0 is the peak power, and T is the bit period [50]. To minimize the width of the pulses prior to the receiver, we transmitted the CRZ signal through dispersive fiber with a total dispersion of -126 ps/nm [50] in the simulations. The receiver consisted of a Gaussian-shaped optical filter with a full-width at half maximum of 124 GHz and a fifth-order electrical Bessel filter with a 3 dB spectral width of 8.5 GHz. The power-equivalent spectral width of the optical spectrum analyzer was 50 GHz. The RZ pulse shape was determined by setting $A = 0$. The rise time of an NRZ pulse was 30 ps. We compared the performance of these formats using both a perfect extinction ratio and an optical extinction ratio of 18 dB. With a finite extinction ratio, for the CRZ and the RZ formats we used the same pulse shapes in the spaces as in the marks but with a lower power, while the NRZ format was constructed by low-pass filtering a two-level step function. In Fig. 4.5, the solid, dashed and dotted curves show the results for the CRZ, RZ, and NRZ formats respectively, all with a perfect extinction ratio. The solid curve with circles, dashed curve with squares, and dotted curve with triangles show the corresponding results with an optical extinction ratio of 18 dB.

In Fig. 4.6, I compare the shapes of an isolated mark for the three formats prior to the receiver. The solid, dashed and dotted curves are results for the CRZ, RZ, and NRZ formats, respectively. The parameters used in (4.51) for the formats that my

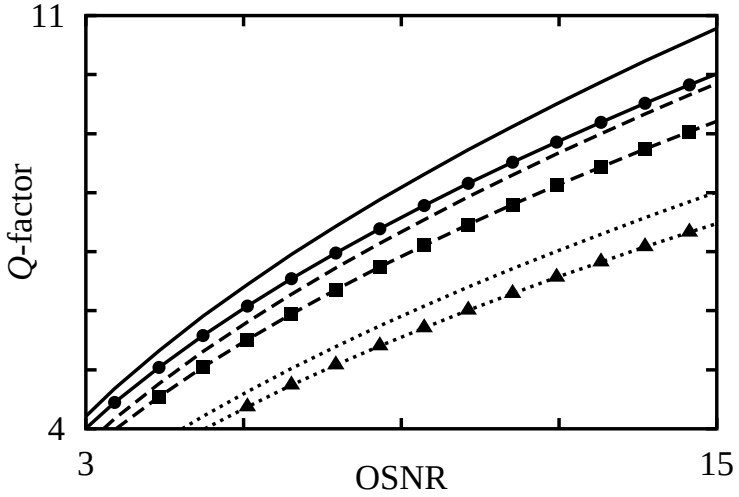

Figure 4.5: A performance comparison of the Q-factor as a function of the OSNR for three modulation formats using (4.51). The parameters of the formats are given in Table 4.1. The power-equivalent spectral width of the optical spectrum analyzer was 50 GHz. The solid, dashed, and dotted curves show the results for the CRZ, RZ and NRZ formats respectively, all with a perfect extinction ratio. The solid curve with circles, the dashed curve with squares, and the dotted curve with triangles show the corresponding results with an optical extinction ratio of 18 dB.

colleagues and I simulated are shown in Table 4.1. The currents in the formula (4.38) for the α_e can be determined from the optical extinction ratio α_o using (4.21). The enhancement factor (4.50) is larger for the CRZ format than for the RZ format and is larger for the RZ format than for the NRZ format, due to the decrease in the pulse duration prior to the receiver, as shown in Fig. 4.6. Thus, for the results shown in Fig. 4.5, and also for those in [20], the primary reason the RZ format has a better back-to-back performance than the NRZ format using the same receiver is because the RZ format has a larger enhancement factor ξ. For the same reason, the CRZ format has a better performance than the RZ format. Fig. 4.5 shows a substantial

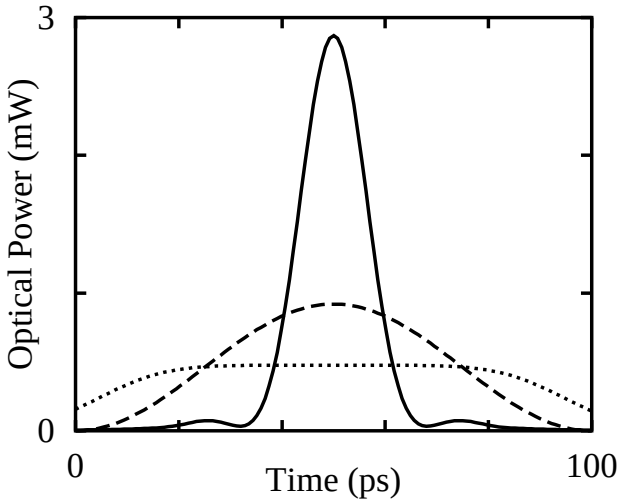

Figure 4.6: The shapes of an isolated mark for the different formats prior to the receiver whose performance comparison is shown in Fig. 4.5. The solid, dashed and dotted curves are results for the CRZ, RZ, and NRZ formats, respectively.

degradation in the system performance for the three systems that we studied when we used an optical extinction ratio of 18 dB, rather than a perfect extinction ratio. A finite extinction ratio penalizes the system by transferring part of the optical intensity from the marks to the spaces, which reduces the enhancement factor and produces signal-noise beating in the spaces according to (4.51). In [18], [19], and [30], $\kappa_1 = 2$ since the electrical filter was approximated by an integrate-and-dump filter. For the receiver that we studied, the use of this approximation would underestimate the signal-noise beating coefficient of the NRZ format by 29%. In addition, the use of an integrate-and-dump filter would make the enhancement factor a format-independent quantity, which could underestimate the enhancement factor of the CRZ format that we studied by 52%.

Format	α_o (dB)	ξ'	ξ	K_1	K_0	M
CRZ	$-\infty$	4.20	0.80	3.48	0	21.23
CRZ	-18	4.13	0.78	3.48	2.99	21.23
RZ	$-\infty$	3.21	0.61	3.05	0	21.23
RZ	-18	3.16	0.60	3.05	3.04	21.23
NRZ	$-\infty$	2.12	0.40	2.83	0	21.23
NRZ	-18	2.05	0.39	2.83	2.74	21.23

Table 4.1: Parameters of the modulation formats used in Figs. 4.5 and 4.6.

Chapter 5

Importance sampling for the polarization-induced penalty

5.1 Computation of outage probability

A major goal in the design of wavelength-division-multiplexed (WDM) optical fiber communications systems is to minimize the probability of channel outages due to polarization effects. System designers commonly allocate a prescribed margin to polarization effects, such as 2 dB, with a fixed probability that the margin will be exceeded, such as 10^{-5}, corresponding to approximately five minutes per year. When this margin is exceeded an outage is said to occur. Because outages are extremely rare, it has been difficult to study them theoretically using standard Monte Carlo simulations and experimentally using laboratory and field experiments.

Wang and Menyuk [8] proposed the reduced Stokes model as a tool to compute the evolution of the optical signal-to-noise ratio (OSNR) and the polarization states of the signal and noise for each channel in long-haul optical fiber communications systems, which I described in Chapter 3. For a given fiber realization, one may calculate the Q-factor from the OSNR and the polarization states of the signal and noise using

the receiver model that I introduced in Chapter 4. First, one computes the Q-factor when PDL and PDG are present and when they are absent for a fixed level of PMD. From that, one may determine the Q-penalty in dB due to these effects, where the Q-penalty is defined to be the difference between the Q-factor in dB without PDL and PDG and the Q-factor in dB when all the polarization effects are included. One then defines the outage probability due to the polarization effects to be the probability that the Q-penalty exceeds an allowed margin, such as 1 or 2 dB. This margin, which is specified during the design of the system, determines the tolerance limit of the operation of the system.

The reduced Stokes model decreases the computational time of simulations of the polarization effects by several orders of magnitude when compared to a full time-domain simulation. Even so, until now, efficient computation of outage probabilities has only been carried out using numerical extrapolation with a Gaussian function [8] to estimate the tails of the probability density function (pdf) of the Q-penalty. With the collaboration of my colleagues, I developed an importance sampling technique [51] to resolve the tails of the pdf of the Q-penalty and thereby efficiently compute outage probabilities as small as 10^{-6} [16] using a fraction of the computational time required by standard Monte Carlo simulations. In addition, we have been able to determine the error that one can make in the computation of the Q-penalty when the polarization states of the signal and noise are not accounted for in the receiver.

In Sec. 5.2, I present a validation of the reduced Stokes model that I described in Chapter 3 in combination with the accurate receiver model that I derived in Chapter 4 by comparison to full time-domain simulations. In Sections 5.3 and 5.4, I show how

the probability of a rare event can be efficiently computed using Monte Carlo simulations with an importance sampling technique that my colleagues and I developed. Using this technique, one can compute outage probabilities using a small fraction of the time required by standard Monte Carlo simulations. We validated this importance sampling technique by comparison with extensive Monte Carlo simulations. In Sec. 5.5, I show how the outage probability due to the polarization effects depends on the polarization-dependent loss (PDL) per optical amplifier, based on simulations with the reduced Stokes model and the accurate receiver model. I also show the relative variation in the results that we obtained using Monte Carlo simulations with importance sampling.

5.2 Reduced Stokes model validation

In order to compute outage probabilities using the reduced Stokes model, my colleagues and I had to first validate the implementation of the reduced model by comparison to a full time-domain model using the Manakov-PMD equation [27] that I described in Chapter 2. Figures 5.1(a) and 5.1(b) show numerical results of the mean of the Q-penalty ΔQ in dB and its standard deviation, respectively, as a function of the PDL per optical amplifier for the full and the reduced models. I define Q in dB as $Q_{\mathrm{dB}} = 20\log_{10}(Q)$, as in [15]. The system that we studied had eight 10 Gbit/s return-to-zero channels spaced 1 nm (124 GHz) apart. The total propagation distance was $8,910$ km, with an amplifier spacing of 33 km. Therefore, there were 270 optical amplifiers in the system. The PDG was equal to 0.06 dB per optical amplifier. For the full model, the fiber's nonlinear coefficient n_2 was 2.6×10^{-20} m^2/W, the fiber's

effective area was 80 μm^2, and the optical signal consisted of a string with 64 bits obtained using the pseudo-random bit string (PRBS) algorithm [52]. The periodic dispersion map consisted of one section of dispersion-shifted fiber whose dispersion was -2 ps/nm-km at 1550 nm and whose length was 264 km, followed by a section of single mode fiber whose dispersion was 16 ps/nm-km and whose length was 33 km. In both sections, the dispersion slope was equal to 0.07 ps/nm^2-km. The residual dispersion in each of the channels was compensated using symmetric pre- and post-dispersion compensation. Since the full simulations require a large amount of computer time, we only compared the mean and the standard deviation of the Q-penalty, and we only used 20 random fiber realizations with the same mean DGD. To compare the two models I used the same fiber realizations in both the full and the reduced model simulations to avoid the statistical uncertainty in the computation of the mean and the standard deviation of the Q-penalty. In addition, we used the same set of fiber realizations for each PDL value. Therefore, the only uncertainty in the comparison comes from the nonlinear polarization rotation in the full time-domain simulations. In [8], a similar comparison was carried out, except that the receiver model did not account for the polarization state of the noise in the receiver. Therefore, the comparison in [8] was essentially a comparison of signal-to-noise ratios in the two models, rather than a comparison of Q-factors. In addition, the random fiber realizations used in the full and the reduced model were independently generated. The use of different fiber realizations in the full and reduced models significantly contributed to the uncertainty in the comparison. In Fig. 5.1, the agreement between the full model and the reduced model with 20 Monte Carlo samples is good. The reduced

model slightly overestimates the standard deviation of the Q-penalty in Fig. 5.1, because the reduced model neglects the slight depolarization that a channel suffers due to PMD. Consequently, with the reduced model, an entire channel can be aligned with the high-loss axis of a PDL element, whereas that is not the case with the full model since the channel will suffer some depolarization due to PMD. We did not compute more than 20 Monte Carlo samples per PDL in the full model simulations because the full model requires an extensive amount of computer time. When we increased the number of realizations in the reduced model to 10^3, the agreement was still very good.

My colleagues and I computed the Q-factor from the full time and frequency domain model using the Manakov-PMD equation [27] in combination with a procedure introduced in [20]. We used a step size of 0.5 km to numerically solve the Manakov-PMD equation. In [20], the BER was computed by averaging the probability of errors in all the marks and spaces of the bit string, since each bit in the string can suffer significant nonlinear polarization rotation and, therefore, suffer a different impairment due to the signal-noise beating. Using the Gaussian approximation, the BER in a string with N_0 spaces and N_1 marks sampled with the sampling time offset t_s is given by

$$\begin{aligned} \mathrm{BER}(t_s, i_{\mathrm{th}}) = \quad & \frac{1}{N_0} \sum_{j=1}^{N_0+N_1} I_0(t_s + jT)\, \mathrm{erfc}\left[\frac{i_{\mathrm{th}} - i(t_s + jT) - \langle i_n \rangle}{\sqrt{2}\sigma_i(t_s + jT)}\right] \\ & + \frac{1}{N_1} \sum_{j=1}^{N_0+N_1} I_1(t_s + jT)\, \mathrm{erfc}\left[\frac{i(t_s + jT) + \langle i_n \rangle - i_{\mathrm{th}}}{\sqrt{2}\sigma_i(t_s + jT)}\right], \end{aligned} \tag{5.1}$$

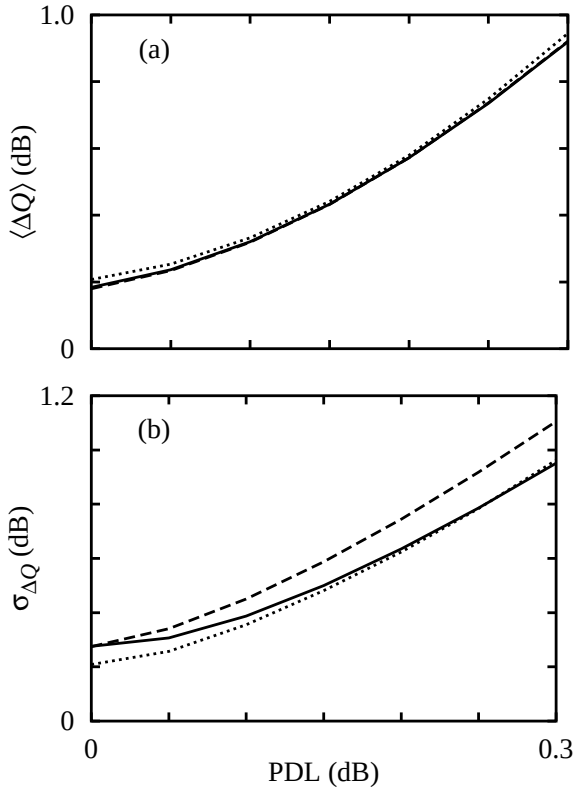

Figure 5.1: Validation of the reduced Stokes model. I show the (a) mean and (b) standard deviation of the Q-penalty ΔQ in dB as a function of the PDL per optical amplifier. The dotted lines are results of full model simulations with 20 samples of fiber realizations. The dashed lines are results of reduced model simulations with 20 samples. The solid lines are results of reduced model simulations with 10^3 samples.

where $\langle i_n \rangle$ and $\sigma_i(t)$ are respectively the mean and the instantaneous standard deviation of the electric current due to optical noise, which were defined in (4.22) and (4.23), respectively. The function $I_0(t)$ is 1 if the bit is a space and 0 otherwise. The function $I_1(t)$ is given by $1 - I_0(t)$. The bit period T is equal to 100 ps. To determine the sampling time for marks and spaces, we used the clock recovery algorithm that I described in Sec. 4.2.4. Using (5.1), we determined the decision threshold i_{th} as the

value that minimizes the BER, which we computed using a zero-finding algorithm. Once the BER was obtained, we computed the corresponding Q-factor by solving the transcendental equation that relates the Q-factor to the BER,

$$\mathrm{BER} = \frac{1}{2}\mathrm{erfc}\left(\frac{Q}{\sqrt{2}}\right). \tag{5.2}$$

Since the term on the right-hand side of (5.2) varies exponentially with Q, we solved Q for a given BER by computing the logarithm of the terms in each side of this equation. Therefore, we computed the Q-factor by searching for the root of the equation

$$\log_{10}(\mathrm{BER}) - \log_{10}\left[\frac{1}{2}\mathrm{erfc}\left(\frac{Q}{\sqrt{2}}\right)\right] = 0 \tag{5.3}$$

using Brent's root search method [53].

5.3 Importance sampling

Importance sampling is a well known technique in statistics that makes a more efficient use of Monte Carlo simulations to compute the probability of rare events. For the problem of computing the outage probability due to the polarization effects for a given channel in a WDM system, the rare events that interest us are those for which the Q-penalty is close to the allowed margin in a given channel. This value is typically large compared to the average or expected value of the Q-penalty. To successfully apply importance sampling to this problem, one needs some knowledge of which random realizations of the fiber PMD produce these large Q-penalties. These random fiber realizations are those for which the Stokes vector of the signal in a given channel is more likely to be aligned with the high-loss axes of the PDL elements than with the low-loss axes.

The basic idea of Monte Carlo simulations with importance sampling is to bias the pdfs of the sources of randomness so that the rare events of interest occur more frequently than they would in standard Monte Carlo simulations and then to weight the samples by the likelihood that the biased random quantities would actually appear in standard Monte Carlo simulations. My colleagues and I observed that the use of a single biased distribution with importance sampling is insufficient to resolve the histogram of the penalty produced by the polarization effects in the configuration space. Thus, it is necessary to combine the statistical results from several biased distributions, since each distribution statistically resolves a different region in the configuration space. However, the biased distributions that do not statistically resolve a given region in the configuration space can significantly degrade the accuracy of the estimation of the probability in that region if they are inadequately combined with the other distributions. This effect is particularly evident in the computation of the penalty because the angles between the polarization state of the channel and the high-loss axes of the PDL elements in each amplifier that we bias are not the only quantities that determine the penalty. The direction of the normalized Stokes vectors of the signal in other channels throughout the transmission line and the angle between the normalized Stokes vector of the signal and the normalized Stokes vector of the polarized part of the noise in a given channel at the receiver also influence the penalty. Using importance sampling, the probability P_I of an event defined by the indicator function $I(\mathbf{x})$, whose value is either 0 or 1, can be mathematically expressed by [54]

$$\hat{P}_I = \sum_{j=1}^{J} \frac{1}{M_j} \sum_{i=1}^{M_j} I(\mathbf{x}_{j,i}) w_j(\mathbf{x}_{j,i}) L_j(\mathbf{x}_{j,i}), \tag{5.4}$$

where

$$L_j(\mathbf{x}_{j,i}) = \frac{p(\mathbf{x}_{j,i})}{p_j^*(\mathbf{x}_{j,i})} \tag{5.5}$$

is the likelihood ratio of the i-th sample $\mathbf{x}_{j,i}$ drawn from the j-th biased distribution, and M_j is the number of samples drawn from the j-th biased distribution $p_j^*(\mathbf{x})$. The term $p(\mathbf{x})$ is the pdf of the j-th unbiased distribution, and J is the number of different biased distributions. The weights $w_j(\mathbf{x})$ allow one to combine different biased distributions. In this application, the random vector $\mathbf{x}$ corresponds to the random realization of the fiber PMD, which is determined by the random mode coupling transformations between the birefringent sections, and the unbiased pdf is known *a priori*. The indicator function I is chosen to compute the probability of having the Q-penalty within a given range, such as a bin in a histogram.

An efficient technique to combine the samples from multiple biased distributions is the balanced heuristic method [54], which my colleagues and I used for the work reported in this dissertation. The balanced heuristic weight assigned to the sample $\mathbf{x}$ drawn from the j-th distribution is given by

$$w_j(\mathbf{x}) = \frac{M_j L_j^{-1}(\mathbf{x})}{\sum_{k=1}^{J} M_k L_k^{-1}(\mathbf{x})}. \tag{5.6}$$

The idea behind the balanced heuristic method is that samples are weighted according to the likelihood that each biased distribution produces samples in that region; distributions that are more likely to put samples there are weighted more heavily. The computation of the balanced heuristic weight for any given sample requires that the likelihood ratio of all the biased distributions be evaluated for that sample. In other words, the likelihood ratio of all J biased distributions have to be evaluated

for the i-th sample drawn from the j-th distribution, even though this sample was obtained using only the biased pdf of the j-th distribution.

A confidence interval for the estimator of the probability P_I of the indicator function $I(\mathbf{x})$ in (5.4) can be defined from the estimator of the variance of $\hat{P}_I$, which is given by

$$\hat{\sigma}^2_{\hat{P}_I} = \sum_{j=1}^{J} \frac{1}{M_j\,(M_j-1)} \sum_{i=1}^{M_j} \left[I(\mathbf{x}_{j,i}) w_j(\mathbf{x}_{j,i}) L_j(\mathbf{x}_{j,i}) - \hat{P}_{I_j} \right]^2, \tag{5.7}$$

where

$$\hat{P}_{I_j} = \frac{1}{M_j} \sum_{i=1}^{M_j} I(\mathbf{x}_{j,i}) w_j(\mathbf{x}_{j,i}) L_j(\mathbf{x}_{j,i}), \tag{5.8}$$

is the contribution of the samples drawn from j-th biased distribution to the estimator $\hat{P}_I$. The confidence interval of the estimator $\hat{P}_I$ is defined to be the interval $\left(\hat{P}_I - \hat{\sigma}_{\hat{P}_I}, \hat{P}_I + \hat{\sigma}_{\hat{P}_I}\right)$. The relative variation equals $\hat{\sigma}_{\hat{P}_I}/\hat{P}_I$. The key to the derivation of (5.7), which we will present in [55], was to realize that the product between the indicator function $I(\mathbf{x})$ and its corresponding weight $w_j(\mathbf{x})L(\mathbf{x})$ is the random variable in the standard expression of the variance [49] for the j-th distribution. Then, the variances produced by all the independent biased distributions are added to obtain the expression for the variance of the estimator $\hat{P}_I$ in (5.7).

5.4 Importance sampling for the polarization-induced penalty

The key issue in applying importance sampling is to choose the biased pdf $p^*(\mathbf{x})$. In the reduced model, the PDL is lumped in the optical amplifiers. So, to compute the

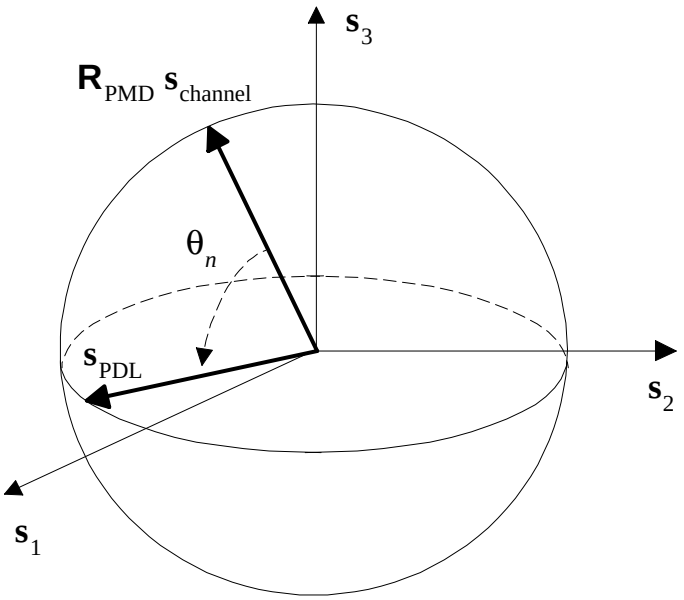

Figure 5.2: Poincaré sphere with the diagram of the importance sampling technique for the PDL-induced penalty. The vector $\mathsf{R}_{\mathrm{PMD}}\,\mathbf{s}_{\mathrm{channel}}$ is the normalized Stokes vector of a given channel prior to the n-th PDL element, $\mathbf{s}_{\mathrm{PDL}}$ is the normalized Stokes vector that is parallel to the high-loss axis of the n-th PDL element, and θ_n is the angle between these two vectors, which I bias towards zero.

outage probability for a given channel, the appropriate pdfs to bias are the pdfs of the angles θ_n between the polarization state of the channel and the polarization state that undergoes the highest loss due to PDL in the n-th optical amplifier. This procedure is illustrated in Fig 5.2, where the polarization transformation matrix $\mathsf{R}_{\mathrm{PMD}}$ of the last birefringent section of the fiber is biased so that the transformed polarization state of a given channel $\mathsf{R}_{\mathrm{PMD}}\,\mathbf{s}_{\mathrm{channel}}$ is more likely to be aligned to the high-loss axis $\mathbf{s}_{\mathrm{PDL}}$ of the PDL elements in each of the optical amplifiers along the transmission. I stress that the method is self-validating. If one biases the samples incorrectly, then

the variance of the results is large.[1] By biasing $\cos\theta_n$ towards one, one increases the likelihood that the Q-penalty of the channel will be large. The angles θ_n are directly determined by the realization of the random mode coupling in the last birefringent section of the fiber that precedes the n-th optical amplifier. Thus, the values of $\cos\theta_n$ play the role of the components of the random vector $\mathbf{x}$ in (5.4). In standard Monte Carlo simulations, in which the PMD is modeled using the coarse step method [27], the cosines of each of the angles θ_n are uniformly distributed in the interval $[-1, 1]$. One can note that an unbiased importance sampling simulation, for which $L(\mathbf{x}) = 1$, is exactly the same as a standard Monte Carlo simulation. However, in biased importance sampling one fixes a biasing parameter α and selects the $\cos\theta_n$ from a biased pdf, $p^*_\alpha(\cos\theta)$. In the work reported in this dissertation, my colleagues and I used the biased pdf

$$p^*_\alpha(\cos\theta) = \frac{\alpha}{1 - e^{-2\alpha}} e^{-\alpha(1-\cos\theta)}, \tag{5.9}$$

which biases $\cos\theta$ towards 1 when α is positive and which corresponds to standard Monte Carlo simulations in the limit $\alpha = 0$. The likelihood ratio for this pdf is given by

$$L_\alpha(\cos\theta) = \frac{1 - e^{-2\alpha}}{2\alpha} e^{\alpha(1-\cos\theta)}. \tag{5.10}$$

Since the unbiased $\cos\theta_n$ are independent random variables, the likelihood ratio for a biased realization of the fiber PMD is equal to the product of the likelihood ratios for each of its biased angles. The actual pdf of the Q-penalty is obtained using (5.4)

[1] There is one important caveat: If two or more disconnected regions in the phase space produce the same penalty, then it is possible in principle to have low variance and for the result to still be incorrect. This outcome does not appear to be physically possible for the problem studied here.

by weighting each fiber sample by its likelihood ratio, which is the likelihood that the biased fiber realization would actually occur in an unbiased simulation. By using the balanced heuristic method to combine several Monte Carlo simulations with different values of the biasing parameter α, one can statistically resolve the pdf of the Q-penalty in any desired range.

To use importance sampling, one must determine the value of the biasing parameter α that enables one to statistically resolve the histogram of the Q-penalty over a range of Q-penalty values whose probability is on the order of a given target probability P_α, such as $P_\alpha = 10^{-6}$. My colleagues and I observed that a biased distribution efficiently resolves the probability of a given bin in the histogram of the Q-penalty when the indicator function for this bin has the value 1 for a large proportion of the biased samples. Therefore, the required value of the biasing parameter α is the one for which the target probability P_α is equal to the likelihood ratio of the biased realization of the fiber PMD evaluated at the expected value of the random variable $\cos\theta$ with biased pdf p^*_α. That is, the bias parameter α satisfies the equation

$$P_\alpha \;=\; L_\alpha^N \left[\langle \cos\theta \rangle\right] \;=\; L_\alpha^N \left(\frac{1 + e^{-2\alpha}}{1 - e^{-2\alpha}} - \frac{1}{\alpha} \right), \tag{5.11}$$

where N is the number of optical amplifiers, and $\langle \cdot \rangle$ is the expectation operator. For example, $P_{0.3} = 1.5 \times 10^{-2}$ and $P_{0.6} = 1.8 \times 10^{-7}$ for $N = 80$. In Fig. 5.3, I show the pdf of $\cos\theta$ for $\alpha = 0$, which produces unbiased samples, together with $\alpha = 0.3$ and $\alpha = 0.6$. Just as in a standard Monte Carlo simulation, as one increases the number of samples with the bias parameter α, one increases the size of the interval about P_α for which the histogram of the Q-penalty is well resolved.

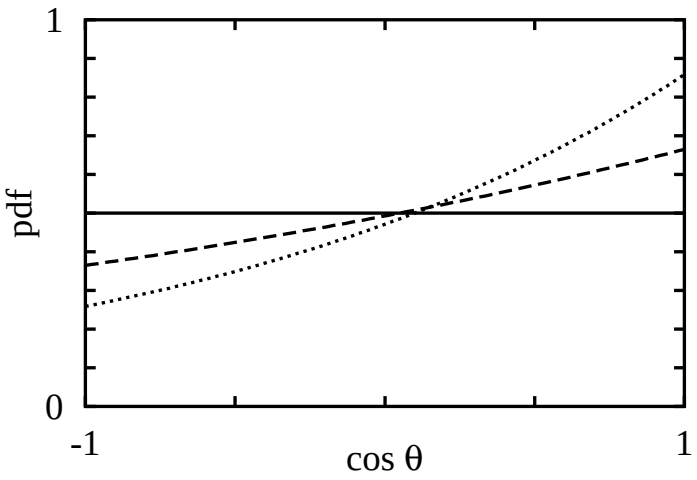

Figure 5.3: Biased pdf of $\cos\theta$ defined in (5.9) for different values of bias strength α. The solid curve shows results for $\alpha = 0$, which corresponds to the unbiased pdf. The dashed curve shows results for $\alpha = 0.3$. The dotted curve shows results for $\alpha = 0.6$.

For the importance sampling technique to be efficient it is required that the performance measure, the Q-penalty, be highly correlated with the random variables that are being biased, the values of $\cos\theta_n$. In other words, given a set of random values of $\cos\theta_n$, a perturbation in any of the values of $\cos\theta_n$ towards 1 should always produce an increase in the Q-penalty. If this requirement is not met, the statistical uncertainty of the results may be so large that importance sampling provides little or no improvement over standard Monte Carlo simulations. For a single channel system with PMD and PDL, small values of the angles θ_n are very highly correlated to large Q penalties. Moreover, the effect of adding gain saturation and PDG to such a system is to always increase the Q-penalty. Consequently, my colleagues and I found that importance sampling is efficient for single channel simulations. However, in a WDM system, the Q-penalty for a given channel also depends on the angles

the polarization states of the other channels make with the polarization state with highest loss due to PDL, as can be illustrated by the following two extreme cases: I consider a given channel, whose polarization state is aligned with the polarization state with highest loss due to PDL. In the first case, I suppose that the polarization states of the other channels are all aligned with the polarization state with lowest loss due to PDL, so that when the total power is partially restored by each of the gain-saturated amplifiers, the given channel loses power relative to the other channels and the largest possible Q-penalty is incurred. In the second case, I suppose that all the channels are co-polarized, so that the given channel does not transfer any power to the other channels and therefore incurs a much smaller Q-penalty. When PDG is included, the variation in the Q-penalty can be even larger for a fixed set of biasing angles θ_n, since in this case the penalty also depends on the degree of polarization (DOP) and on the total polarization state. Therefore, in a WDM system with only a few channels, the correlation between the Q-penalty and the angles θ_n is not as high as in the single channel system and the importance sampling procedure that I described is not as efficient. On the other hand, since PMD tends to randomize the polarization states of the channels, one finds that as the number of channels increases the DOP decreases. Thus, the Q-penalty becomes less dependent on the polarization states of the other channels and on the PDG, and the importance sampling procedure becomes more efficient. In the next section, I will show that this procedure is highly efficient for a WDM system with as few as eight channels.

5.5 Numerical results

In this section, I describe the outage probabilities due to polarization effects for a trans-oceanic WDM system by applying importance sampling to resolve the tails of the pdf of the Q-penalty for the same system that I presented in Sec. 5.2. My colleagues and I validated the method of computing outage probabilities using Monte Carlo simulations with importance sampling by making comparisons with extensive standard Monte Carlo simulations. To obtain these results, we used the reduced model that I described in Chapter 3 and the receiver model that I described in Chapter 4. For the results in this chapter, we obtained the final histograms of the Q-penalty by dividing the range of penalties in dB into 50 bins and by combining all the biased distributions using the balanced heuristic method that I described in Sec. 5.3.

We simulated a 10 Gbit/s return-to-zero (RZ) system with eight WDM channels spaced 1 nm (124 GHz) apart. The total propagation distance was $8,910$ km and the amplifier spacing was 33 km. The PMD was 0.1 ps/km$^{1/2}$. In Fig. 5.4, I plot the outage probability as a function of the allowed Q-penalty margin ΔQ. The PDL was 0.13 dB and the PDG was 0.06 dB per optical amplifier. The solid curve shows the outage probability that is computed using 10^4 Monte Carlo simulations with importance sampling for each biased distribution $p_\alpha(\cdot)$ in (5.9) with: $\alpha = 0$, which corresponds to unbiased simulations, $\alpha = 0.3$, $\alpha = 0.4$, $\alpha = 0.5$ and $\alpha = 0.6$. In Fig. 5.4, a dashed curve shows results of extensive Monte Carlo simulations with a total of 5×10^6 samples. The agreement with the importance sampling results is good up to outage probability at which we could obtain in a statistically significant

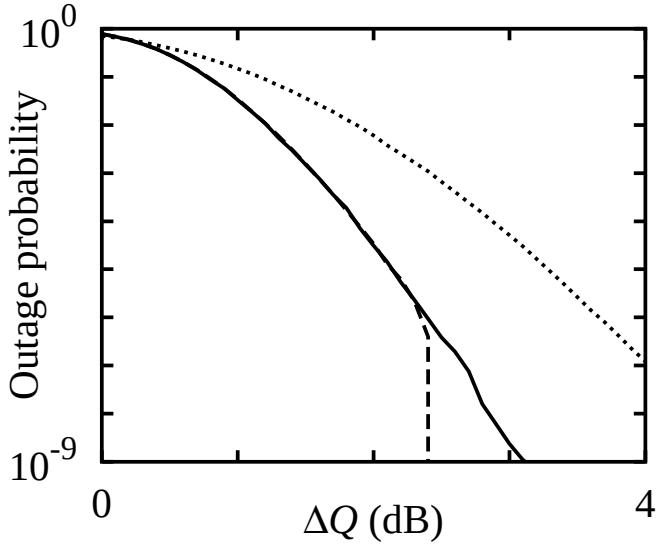

Figure 5.4: The outage probability as a function of the allowed Q-penalty margin ΔQ in dB for an eight channel WDM system. We set PDL = 0.13 dB and PDG = 0.06 dB per optical amplifier. The solid curve shows results of 5×10^4 Monte Carlo simulations with importance sampling. The dashed curve shows results of 5×10^6 standard Monte Carlo simulations. The dotted curve shows results obtained with the same 5×10^4 Monte Carlo simulations with importance sampling, but assuming that the noise is always unpolarized at the receiver.

number of samples. We used the same computer and the same program for both the extensive Monte Carlo simulations and the simulations with importance sampling; the only difference is that we did not bias the fiber birefringence for the standard Monte Carlo method. The results with importance sampling required only 1% of the computational time that was required by extensive Monte Carlo simulations. The dotted curve shows the outage probability obtained with the same Monte Carlo simulations with importance sampling, but using the receiver model that assumes unpolarized noise in the receiver. As a consequence, this model implicitly assumes that half of the noise beats with the signal in the photodetector, which leads to

incorrect results. Thus, we conclude that it is very important to properly account for the polarization states of the signal and the noise in the receiver.

In Fig. 5.5, I plot the outage probability at margins of 1 dB and 2 dB as a function of the PDL per optical amplifier for the same system as in Fig. 5.4. For these results, my colleagues and I used Monte Carlo simulations with importance sampling. The solid curve with circles shows that the PDL per optical amplifier must not exceed 0.11 dB in order for the probability that the Q-penalty exceeds 2 dB be smaller that 10^{-6}. Therefore, if the margin allocated in the design to the penalty produced by the polarization effects is equal to 2 dB with an outage probability of 10^{-6}, the PDL per optical amplifier must not be larger than 0.11 dB. If the margin allocated to the polarization effects is equal to 3 dB, shown with a dashed line with diamonds, the PDL per optical amplifier can be as large as 0.15 dB with an outage probability of 10^{-6}.

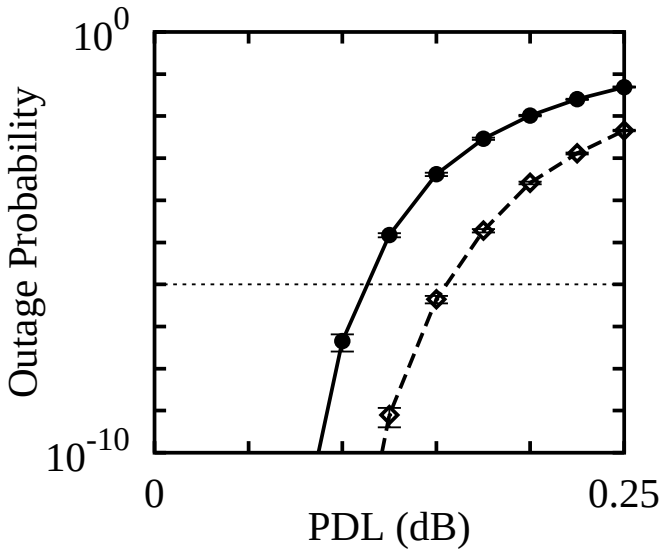

Figure 5.5: The outage probability as a function of the PDL per optical amplifier for two Q-penalty margins ΔQ. The solid curve with circles shows the probability that the Q-penalty exceeds 2 dB. The dashed curve with diamonds shows the probability that the Q-penalty exceeds 3 dB. The error bars show the confidence interval for the curves computed using (5.7). The dotted line shows the 10^{-6} outage probability level.

Chapter 6

Fiber recirculating loop experiments

6.1 Fiber recirculating loops

Optical fiber recirculating loops are useful tools for laboratory studies of the effects that limit the performance of optical fiber communications systems [5]. In this chapter, I present an experimental validation of the reduced Stokes model that I described in Chapter 3 in combination with the receiver model that I derived in Chapter 4 by comparison with dispersion-managed soliton, optical-fiber recirculating loop experiments.

The evolution of the polarization states in fiber recirculating loops is periodic, which contrasts with the random evolution of the polarization states due to PMD in straight-line systems. As a consequence of this periodicity, polarization effects accumulate faster in fiber recirculating loops than in straight-line systems [56], [57]. To overcome this problem, my colleagues and I use a loop-synchronous polarization scrambler in both sets of experiments that I describe in this chapter. Loop-synchronous polarization scramblers are used to rotate the polarization state of the signal after each round trip of the loop to emulate PMD-induced polarization rota-

tions and, thereby, to break the periodicity of the evolution of the polarization states during the transmission [56]. In this set of fiber recirculating loop experiments, which was carried out primarily by Yu Sun at the University of Maryland Baltimore County, we showed that partially polarized noise due to PDL causes an asymmetrical Q-factor distribution in the receiver. We also showed that it is very important to account for the polarization state of the noise at the receiver when computing the performance of fiber recirculating loop systems.

6.2 Experimental Q-factor distribution

In this section, I show that partially polarized noise due to PDL causes an asymmetrical Q-factor distribution at the receiver. This asymmetry occurs because the SNR of the electrical current in the receiver does not completely determine the Q-factor if the noise is partially polarized. Therefore, it is very important to accurately model the contribution of the partially polarized noise at the receiver as I described in Chapter 4.

In Fig 6.1, I show a diagram of the optical fiber recirculating loop that my colleagues and I used. This fiber recirculating loop was a dispersion-managed soliton (DMS) system that consisted of 100 km of dispersion-shifted fiber with $D = -1$ ps/nm-km, followed by 7 km of single-mode fiber with $D = 17$ ps/nm-km. We used a $LiNbO_3$ polarization controller to synchronously scramble the polarization state of the light in the loop once per round trip, which we controlled using a programmable waveform generator. The PDG was equal to 0.05 dB per optical amplifier and the PMD was equal to 0.1 ps/km$^{1/2}$. The PDL per round trip could be varied

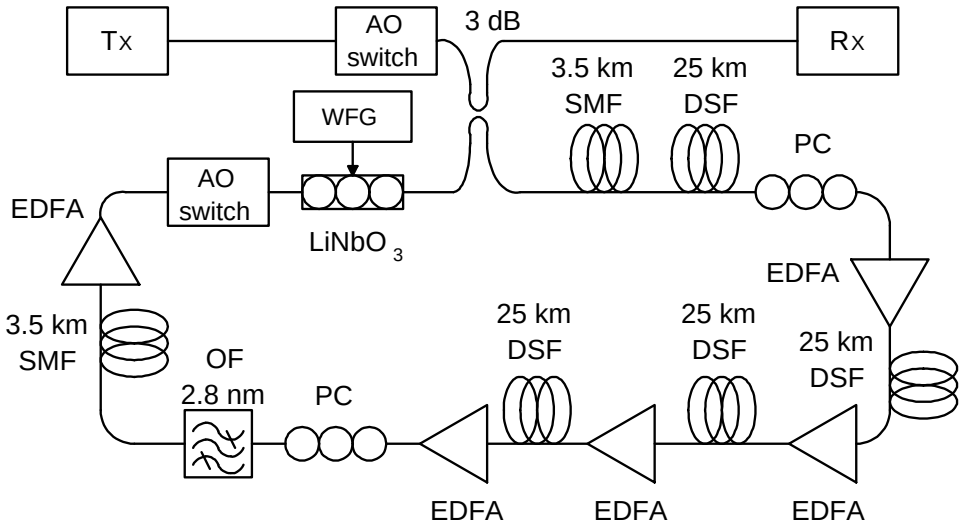

Figure 6.1: Diagram of the fiber recirculating loop experiment.

from 0.13 dB to 0.6 dB by adjusting the manual polarization controllers inside the loop. In the DMS loop, the interaction between dispersion and nonlinearity produces optical pulses that can be transmitted over ultra-long distances without suffering waveform distortions even in the presence of intra-channel PMD [13], [58]. We propagated a 10 Gbit/s return-to-zero (RZ) modulation format through 94 round trips to emulate a long-haul undersea system with a total propagation distance of $10,000$ km. We also employed a polarization scrambler at the transmitter to reduce the effect of PDG. We chose a slow scrambling rate of 12.6 GHz to simultaneously reduce the effects of PDG and the amplitude modulation induced by PDL [14].

In Fig. 6.2, I show the pdf of the Q-factor for the DMS fiber recirculating loop that I showed in Fig. 6.1 with a PDL per round trip of 0.6 dB [22]. In Fig. 6.2, the dotted bars are experimental results. The solid curve shows the result obtained from Monte Carlo simulations using the reduced Stokes model with the accurate receiver model that accounts for the polarization state of the noise at the receiver. The dashed

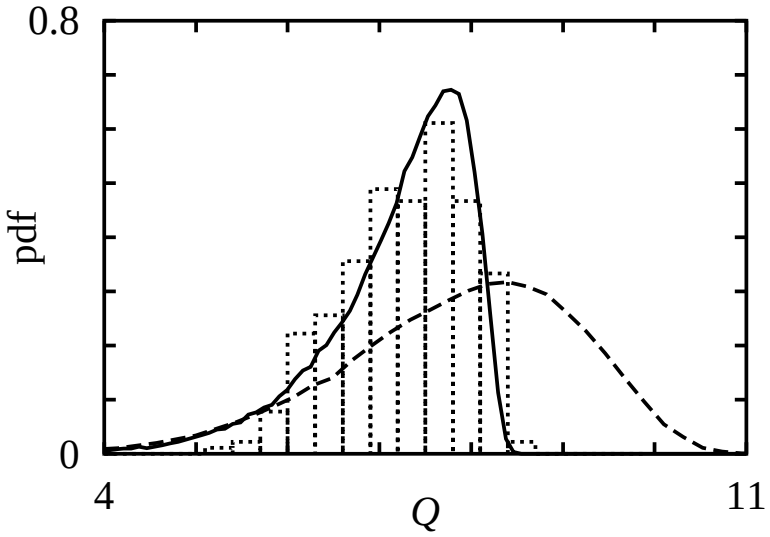

Figure 6.2: The pdf of the Q-factor for the DMS fiber recirculating loop at the University of Maryland Baltimore County with 0.6 dB of PDL per round trip. The dotted bars show the experimental result with 200 measurements with different realizations of the transformations produced by a $LiNbO_3$ polarization controller. The solid curve shows result obtained from a Monte Carlo simulation with 10^5 samples using the reduced Stokes model with the accurate receiver model that accounts for the polarization state of the noise at the receiver. The dashed curve shows the result from a Monte Carlo simulation with 10^5 samples using the reduced Stokes model with the receiver model that assumes unpolarized noise at the receiver.

curve shows the result obtained from Monte Carlo simulations using the reduced Stokes model with the receiver model that assumes that the noise is unpolarized at the receiver. We observed a very good agreement between the experiments and the simulations using the accurate receiver model that accounts for the polarization state of the noise at the receiver that I derived in Chapter 4. The pdf of the Q-factor in this single-channel experiment has a sharp cut-off at about $Q = 8.5$. The receiver model that does not account for the polarization state of the noise predicts a

substantially larger tail for the large Q values. These large Q values were not observed in the experiment because they come from realizations of the $LiNbO_3$ polarization controller that preferentially attenuate the noise that is in the polarization state orthogonal to the polarization state of the signal. In this case, the optical noise at the receiver is primarily co-polarized with the signal, which produces a variance of the current due to the signal-noise beating that is equal to what would be produced in the presence of an arbitrarily large noise component in the polarization state orthogonal to the polarization state of the signal. Therefore, the receiver model that does not account for the polarization state of the noise incorrectly assumes that half of the optical intensity of the received noise produces signal-noise beating. The use of this approximation significantly overestimates the probability of obtaining large Q values for the same reason that it overestimates the penalty due to the polarization effects—the probability of obtaining low Q values—as I showed in Chapter 5.

Chapter 7

Conclusions

Polarization effects can significantly limit the performance of long-haul terrestrial and trans-oceanic optical fiber communications systems. Polarization effects can produce randomly varying signal distortions and power penalties that can lead to system outages on a time scale that varies between milliseconds and hours. Therefore, accurately computing the penalty produced by the polarization effects is crucial in the design of long-haul optical fiber communications systems. Wang and Menyuk [8] proposed the reduced Stokes model as a tool for the computation of the penalty produced by the polarization effects that is much more efficient than using the full time-domain model. Even so, it is not currently feasible to accurately compute outage probabilities as low as 10^{-6} using standard Monte Carlo simulations with the reduced Stokes model. Moreover, the receiver model used by Wang and Menyuk did not account for the effect of partially polarized noise in the receiver.

In this dissertation, I have described two contributions that I made with the collaboration of my colleagues to the field of optical fiber communications systems. These contributions extend and substantially improve the accuracy of the reduced

Stokes model. I described a technique that uses Monte Carlo simulations with importance sampling that we developed to compute the probability density function of the Q-factor and the outage probability for a channel in a long-haul, wavelength-division-multiplexed optical fiber communications system due to the combination of inter-channel polarization mode dispersion (PMD), polarization-dependent loss (PDL), and polarization-dependent gain (PDG). This technique allows one to accurately compute outage probabilities as small as 10^{-6} using the reduced Stokes model at a fraction of the computational cost required by standard Monte Carlo simulations. This technique was validated by comparison to extensive Monte Carlo simulations.

I also described an accurate formula that we developed that relates the Q-factor to the OSNR for amplitude-shift-keyed optical fiber communication systems with arbitrary optical pulse shapes, arbitrary receiver characteristics and arbitrarily polarized noise using an accurate receiver model. Using this formula, we showed how the performance of a system depends on the state of polarization of the noise, which can be significantly affected by PMD in the optical fiber, PDL, and PDG in the optical amplifiers. We also defined the enhancement factor and three other parameters that explicitly quantify the performance of different modulation formats, which can be useful in receiver optimization studies. We validated this formula by comparison to back-to-back 10 Gbit/s experiments and Monte Carlo simulations. We also validated the receiver model that I developed in combination with the reduced Stokes model by comparison to optical fiber recirculating loop experiments. We applied the proposed receiver model to compare the performance of the chirped return-to-zero (CRZ), return-to-zero (RZ), and nonreturn-to-zero (NRZ) modulation formats

with a finite extinction ratio. We observed that the CRZ modulation format outperforms both the RZ and the NRZ formats with the same optical power and receiver characteristics because the CRZ format has the largest enhancement factor for the receiver that we studied. This method should be applicable to the study of other amplitude-shift-keyed formats.

Appendix A

Derivation of integral expressions for the receiver model

In this Appendix, I present a derivation of the integral expressions for the mean and the variance of the electric current in the receiver given in (4.22), (4.24), and (4.32) in Chapter 4. For consistency, I use the same definitions introduced in that chapter. This derivation is based on previous work by Winzer *et al.* [20] that I extended to account for arbitrarily polarized noise.

From (4.19), the electric current $i(t)$ in the receiver due to the signal and the noise can be expressed as the sum of three terms,

$$i\,(t) \;=\; i_s(t) + i_n(t) + i_{sn}(t), \tag{A.1}$$

where

$$i_s\,(t) \;=\; R\;|e_{s_o}(t)|^2 * h_e(t) \tag{A.2}$$

is the electric current due to the noise-free signal,

$$i_n\,(t) \;=\; R\left\{\left|e_{n_{o,1}}(t)\right|^2 + \left|e_{n_{o,2}}(t)\right|^2\right\} * h_e(t) \tag{A.3}$$

is the electric current due to the noise beating with itself, and

$$\begin{aligned} i_{sn}(t) \;=\; & R\,\Big[e_{s_{o,1}}(t)\,e^{*}_{n_{o,1}}(t) + e^{*}_{s_{o,1}}(t)\,e_{n_{o,1}}(t) \\ & + e_{s_{o,2}}(t)\,e^{*}_{n_{o,2}}(t) + e^{*}_{s_{o,2}}(t)\,e_{n_{o,2}}(t)\Big] * h_e(t) \end{aligned} \tag{A.4}$$

is the electric current due to the beating between the signal and the noise. In (A.3) and (A.4), $e_{s_{o,1}}(t)$ is the Jones vector that mathematically represents the the filtered noise field that is co-polarized with the polarization state of the polarized component of the noise, and $e_{s_{o,2}}(t)$ is its orthogonal Jones vector component.

The expression for the variance of the electric current at any time t is given by

$$\sigma_i^2(t) \;=\; \langle i^2\rangle(t) - \langle i\rangle^2(t). \tag{A.5}$$

To compute the mean and the mean square of the current $i(t)$, I recall the statistical properties of the optically filtered noise given in (4.15)–(4.17) to write the autocorrelation functions of the filtered noise as

$$\langle e_{n_{o,1}}(t)\,e^{*}_{n_{o,1}}(t')\rangle \;=\; \frac{1}{2}(1+\mathrm{DOP}_n)N_{\mathrm{ASE}}\,r_o(t'-t) \tag{A.6}$$

and

$$\langle e_{n_{o,2}}(t)\,e^{*}_{n_{o,2}}(t')\rangle \;=\; \frac{1}{2}(1-\mathrm{DOP}_n)N_{\mathrm{ASE}}\,r_o(t'-t)\,, \tag{A.7}$$

where N_{ASE} is the total power spectral density of the noise, while the terms

$$\langle e_{n_{o,1}}(t)\,e^{*}_{n_{o,2}}(t')\rangle \;=\; \langle e_{n_{o,1}}(t)\,e_{n_{o,1}}(t')\rangle \;=\; \langle e_{n_{o,2}}(t)\,e_{n_{o,2}}(t')\rangle \;=\; 0. \tag{A.8}$$

The lack of correlation between the fields $e_{n_{o,1}}(t)$ and $e_{n_{o,2}}(t)$ follows from an appropriate choice of basis $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2\}$ for the Jones vector of the filtered noise. Following the

same procedure used in the derivation of (A.6) and (A.7), I find that

$$\begin{aligned}\langle e_{n_{o,1}}(t^{\mathrm{i}})\, e^{*}_{n_{o,1}}(t^{\mathrm{ii}})\, e_{n_{o,1}}(t^{\mathrm{iii}})\, e^{*}_{n_{o,1}}(t^{\mathrm{iv}})\rangle \;=\; & \frac{1}{4}(1+\mathrm{DOP}_n)^2 N^2_{\mathrm{ASE}} \\ & \times \left[r_o(t^{\mathrm{ii}}-t^{\mathrm{i}})\, r_o(t^{\mathrm{iv}}-t^{\mathrm{iii}}) + r_o(t^{\mathrm{iv}}-t^{\mathrm{i}})\, r_o(t^{\mathrm{iii}}-t^{\mathrm{ii}})\right],\end{aligned} \tag{A.9}$$

and

$$\begin{aligned}\langle e_{n_{o,2}}(t^{\mathrm{i}})\, e^{*}_{n_{o,2}}(t^{\mathrm{ii}})\, e_{n_{o,2}}(t^{\mathrm{iii}})\, e^{*}_{n_{o,2}}(t^{\mathrm{iv}})\rangle \;=\; & \frac{1}{4}(1-\mathrm{DOP}_n)^2 N^2_{\mathrm{ASE}} \\ & \times \left[r_o(t^{\mathrm{ii}}-t^{\mathrm{i}})\, r_o(t^{\mathrm{iv}}-t^{\mathrm{iii}}) + r_o(t^{\mathrm{iv}}-t^{\mathrm{i}})\, r_o(t^{\mathrm{iii}}-t^{\mathrm{ii}})\right].\end{aligned} \tag{A.10}$$

Using (A.1) and (A.6)–(A.8), I find that the mean value of the current is given by

$$\langle i\rangle(t) \;=\; i_s(t) + \langle i_n\rangle(t), \tag{A.11}$$

which implies that

$$\langle i\rangle^2(t) \;=\; i_s^2(t) + 2\, i_s(t)\, \langle i_n\rangle(t) + \langle i_n\rangle^2(t) \tag{A.12}$$

and

$$\langle i^2\rangle(t) \;=\; i_s^2(t) + \langle i_n^2\rangle(t) + \langle i_{sn}^2\rangle(t) + 2\, i_s(t)\, \langle i_n\rangle(t). \tag{A.13}$$

In (A.13), I do not include the terms that are proportional to $\langle i_{sn}\rangle(t)$ because $\langle i_{sn}\rangle(t)$ averages out to zero, since the optical noise field has zero mean. Therefore, I can express the variance of the current in the receiver as

$$\sigma_i^2(t) \;=\; \sigma^2_{\mathrm{ASE-ASE}}(t) + \sigma^2_{\mathrm{S-ASE}}(t), \tag{A.14}$$

where

$$\sigma^2_{\mathrm{ASE-ASE}}(t) = \langle i_n^2\rangle(t) - \langle i_n\rangle^2(t) \tag{A.15}$$

is the variance of the current due to the noise-noise beating, and

$$\sigma^2_{\mathrm{S-ASE}}(t) = \langle i^2_{sn} \rangle (t) \tag{A.16}$$

is the variance of the current due to signal-noise beating.

To obtain closed form integral expressions for the mean and the variance of the electric current due to noise in (A.15) and (A.16), respectively, I use the statistical properties of the noise that I described in (A.6)–(A.10). The convolution of two arbitrary functions $g(t)$ and $h(t)$ is defined by

$$g(t) * h(t) = \int_{-\infty}^{+\infty} g(\tau) h(t-\tau) \mathrm{d}\tau. \tag{A.17}$$

The limits of all of the integral expressions in this Appendix are from $-\infty$ to $+\infty$, which for simplicity I suppress in all the subsequent expressions. Substituting (A.17) into (A.3), I obtain

$$\begin{aligned} \langle i_n \rangle (t) \;=\; & R \int \mathrm{d}t^{\mathrm{i}} \int \mathrm{d}t^{\mathrm{ii}} \int \mathrm{d}t^{\mathrm{iii}} \langle e_{n_1}(t^{\mathrm{i}})\, e^*_{n_1}(t^{\mathrm{ii}}) \rangle\, h_o(t^{\mathrm{iii}} - t^{\mathrm{i}})\, h^*_o(t^{\mathrm{iii}} - t^{\mathrm{ii}})\, h_e(t - t^{\mathrm{iii}}) \\ & + R \int \mathrm{d}t^{\mathrm{i}} \int \mathrm{d}t^{\mathrm{ii}} \int \mathrm{d}t^{\mathrm{iii}} \langle e_{n_2}(t^{\mathrm{i}})\, e^*_{n_2}(t^{\mathrm{ii}}) \rangle\, h_o(t^{\mathrm{iii}} - t^{\mathrm{i}})\, h^*_o(t^{\mathrm{iii}} - t^{\mathrm{ii}})\, h_e(t - t^{\mathrm{iii}}). \end{aligned} \tag{A.18}$$

Substituting (A.6) and (A.7) into (A.18), I find that

$$\langle i_n \rangle \;=\; R\, N_{\mathrm{ASE}} \int \left| h_o(t^{\mathrm{i}}) \right|^2 \mathrm{d}t^{\mathrm{i}} \int h_e(t^{\mathrm{ii}}) \mathrm{d}t^{\mathrm{ii}}. \tag{A.19}$$

Equation (A.19) can be further simplified to

$$\langle i_n \rangle \;=\; R\, N_{\mathrm{ASE}}\, B_o, \tag{A.20}$$

where I define the electrical and the optical filter transfer functions such that $|H_e(0)| = |H_o(0)| = 1$. Therefore, $B_o = r_o(0) = \int |h_o(\tau)|^2 d\tau$ is the power-equivalent spectral width [24] of the optical filter.

Using (A.3), I can express the mean square of $i_n(t)$ as

$$\langle i_n^2 \rangle(t) \;=\; \langle i_{n_1}^2 \rangle(t) \;+ \langle i_{n_2}^2 \rangle(t) \;+ \langle i_{n_{12}}^2 \rangle(t), \tag{A.21}$$

where

$$\langle i_{n_1}^2 \rangle(t) \;=\; \left\langle \left[R \left| e_{n_{o,1}}(t) \right|^2 * h_e(t) \right]^2 \right\rangle, \tag{A.22}$$

$$\langle i_{n_2}^2 \rangle(t) \;=\; \left\langle \left[R \left| e_{n_{o,2}}(t) \right|^2 * h_e(t) \right]^2 \right\rangle, \tag{A.23}$$

and

$$\langle i_{n_{12}}^2 \rangle(t) \;=\; 2 \left\langle \left[R \left| e_{n_{o,1}}(t) \right|^2 * h_e(t) \right] \left[R \left| e_{n_{o,2}}(t) \right|^2 * h_e(t) \right] \right\rangle. \tag{A.24}$$

Substituting (A.17) into (A.22)–(A.24), I obtain

$$\begin{aligned} \langle i_{n_1}^2 \rangle(t) \;=\; & R^2 \int \mathrm{d}t^{\mathrm{i}} \int \mathrm{d}t^{\mathrm{ii}} \int \mathrm{d}t^{\mathrm{iii}} \int \mathrm{d}t^{\mathrm{iv}} \int \mathrm{d}t^{\mathrm{v}} \int \mathrm{d}t^{\mathrm{vi}} \left\langle e_{n_1}(t^{\mathrm{i}})\, e_{n_1}^*(t^{\mathrm{ii}})\, e_{n_1}(t^{\mathrm{iv}})\, e_{n_1}^*(t^{\mathrm{v}}) \right\rangle \\ & \times h_o(t^{\mathrm{iii}} - t^{\mathrm{i}})\, h_o^*(t^{\mathrm{iii}} - t^{\mathrm{ii}})\, h_o(t^{\mathrm{vi}} - t^{\mathrm{iv}})\, h_o^*(t^{\mathrm{vi}} - t^{\mathrm{v}})\, h_e(t - t^{\mathrm{iii}})\, h_e(t - t^{\mathrm{vi}}), \end{aligned} \tag{A.25}$$

and $\langle i_{n_2}^2 \rangle(t)$, which is the same as (A.25) when one replaces $e_{n_1}(t)$ by $e_{n_2}(t)$, and

$$\begin{aligned} \langle i_{n_{12}}^2 \rangle(t) \;=\; & 2\, R^2 \int \mathrm{d}t^{\mathrm{i}} \int \mathrm{d}t^{\mathrm{ii}} \int \mathrm{d}t^{\mathrm{iii}} \int \mathrm{d}t^{\mathrm{iv}} \int \mathrm{d}t^{\mathrm{v}} \int \mathrm{d}t^{\mathrm{vi}} \left\langle e_{n_1}(t^{\mathrm{i}})\, e_{n_1}^*(t^{\mathrm{ii}}) e_{n_2}(t^{\mathrm{iv}})\, e_{n_2}^*(t^{\mathrm{v}}) \right\rangle \\ & \times h_o(t^{\mathrm{iii}} - t^{\mathrm{i}})\, h_o^*(t^{\mathrm{iii}} - t^{\mathrm{ii}})\, h_o(t^{\mathrm{vi}} - t^{\mathrm{iv}})\, h_o^*(t^{\mathrm{vi}} - t^{\mathrm{v}})\, h_e(t - t^{\mathrm{iii}})\, h_e(t - t^{\mathrm{vi}}). \end{aligned} \tag{A.26}$$

Substituting (A.6) and (A.7) into (A.26), I obtain

$$\langle i_{n_{12}}^2 \rangle \;=\; \frac{1}{2}\, R^2\, (1 - \mathrm{DOP}_n^2)\, N_{\mathrm{ASE}}^2\, B_o^2. \tag{A.27}$$

Likewise, substituting (A.9) and (A.10) into (A.25), I obtain

$$\begin{aligned}\langle i_{n_1}^2\rangle(t) \;=\; & \frac{1}{4}\,R^2\,(1+\mathrm{DOP}_n)^2\,N_{\mathrm{ASE}}^2\left[B_o^2 \;+\; \int \mathrm{d}t^{\mathrm{i}} \int \mathrm{d}t^{\mathrm{ii}} \int \mathrm{d}t^{\mathrm{iii}} \int \mathrm{d}t^{\mathrm{iv}}\right. \\ & \left.\times \left|h_o(t^{\mathrm{iii}}-t^{\mathrm{i}})\right|^2 h_o(t^{\mathrm{vi}}-t^{\mathrm{ii}})\,h_o^*(t^{\mathrm{vi}}-t^{\mathrm{i}})\,h_e(t-t^{\mathrm{iii}})\,h_e(t-t^{\mathrm{iv}})\right].\end{aligned} \tag{A.28}$$

and

$$\begin{aligned}\langle i_{n_2}^2\rangle \;=\; & \frac{1}{4}\,R^2\,(1-\mathrm{DOP}_n)^2\,N_{\mathrm{ASE}}^2\left[B_o^2 \;+\; \int \mathrm{d}t^{\mathrm{i}} \int \mathrm{d}t^{\mathrm{ii}} \int \mathrm{d}t^{\mathrm{iii}} \int \mathrm{d}t^{\mathrm{iv}}\right. \\ & \left.\times \left|h_o(t^{\mathrm{iii}}-t^{\mathrm{i}})\right|^2 h_o(t^{\mathrm{vi}}-t^{\mathrm{ii}})\,h_o^*(t^{\mathrm{vi}}-t^{\mathrm{i}})\,h_e(t-t^{\mathrm{iii}})\,h_e(t-t^{\mathrm{iv}})\right].\end{aligned} \tag{A.29}$$

Substituting (A.27)–(A.29) into (A.21), I also obtain

$$\langle i_n^2\rangle \;=\; R^2\,N_{\mathrm{ASE}}^2\left[B_o^2+\frac{1}{2}\left(1+\mathrm{DOP}_n^2\right)\int |r_o(\tau)|^2\,r_e(\tau)\mathrm{d}\tau\right] \tag{A.30}$$

where $r_o(\tau)$ is the autocorrelation function of the optical filter described in (4.4) and $r_e(\tau)$ is the autocorrelation function of the electrical filter described in (4.27). The autocorrelation functions of the filters appear in (A.30) after rearranging the integral expressions in the second term in the right-hand side of both (A.28) and (A.29). Finally, substituting (A.20) and (A.30) into (A.15), I obtain a closed-form integral expression for the variance of the electric current due to the noise-noise beating,

$$\sigma_{\mathrm{ASE-ASE}}^2 \;=\; \frac{1}{2}\,R^2\,N_{\mathrm{ASE}}^2\left(1+\mathrm{DOP}_n^2\right)\int |r_o(\tau)|^2\,r_e(\tau)\mathrm{d}\tau, \tag{A.31}$$

which completes the derivation of (4.24).

In order to compute the variance of the electric current due to the signal-noise beating $\sigma_{\mathrm{S-ASE}}^2(t)$ in (A.16), I compute the mean-square of $i_{sn}(t)$ in (A.4), which is

equal to

$$\begin{aligned}\langle i_{sn}^2(t)\rangle \;=\;& R^2\left\langle\left[e_{s_{o,1}}(t)\,e_{n_{o,1}}^*(t)*h_e(t)\right]\,\left[e_{s_{o,1}}^*(t)\,e_{n_{o,1}}(t)*h_e(t)\right]\right\rangle \\ &+R^2\left\langle\left[e_{s_{o,2}}(t)\,e_{n_{o,2}}^*(t)*h_e(t)\right]\,\left[e_{s_{o,2}}^*(t)\,e_{n_{o,2}}(t)*h_e(t)\right]\right\rangle.\end{aligned} \tag{A.32}$$

Substituting (A.9), (A.10), and (A.17) into (A.32), I obtain

$$\begin{aligned}\sigma_{\mathrm{S-ASE}}^2(t) =\;& \langle i_{sn}^2(t)\rangle \;=\; R^2\left(1+\mathrm{DOP}_n\right)N_{\mathrm{ASE}} \\ &\times\int \mathrm{d}t^{\mathrm{i}}\int \mathrm{d}t^{\mathrm{ii}} e_{s_{1,o}}(t^{\mathrm{i}})\,e_{s_{1,o}}^*(t^{\mathrm{ii}})\,r_o(t^{\mathrm{i}}-t^{\mathrm{ii}})\,h_e(t-t^{\mathrm{i}})\,h_e(t-t^{\mathrm{ii}}) \\ &+R^2\left(1-\mathrm{DOP}_n\right)N_{\mathrm{ASE}} \\ &\times\int \mathrm{d}t^{\mathrm{i}}\int \mathrm{d}t^{\mathrm{ii}} e_{s_{2,o}}(t^{\mathrm{i}})\,e_{s_{2,o}}^*(t^{\mathrm{ii}})\,r_o(t^{\mathrm{i}}-t^{\mathrm{ii}})\,h_e(t-t^{\mathrm{i}})\,h_e(t-t^{\mathrm{ii}}),\end{aligned} \tag{A.33}$$

which is the variance of the electric current due to the signal-noise beating for an arbitrarily polarized signal and an arbitrarily polarized noise. The integral expressions in (A.33) can be rearranged so that

$$\begin{aligned}\sigma_{\mathrm{S-ASE}}^2(t) =\;& R^2\left(1+\mathrm{DOP}_n\right)N_{\mathrm{ASE}} \\ &\times\int e_{s_{1,o}}(\tau)h_e(t-\tau)\int e_{s_{1,o}}^*(\tau')h_e(t-\tau')r_o(\tau-\tau')\mathrm{d}\tau'\mathrm{d}\tau \\ &+R^2\left(1-\mathrm{DOP}_n\right)N_{\mathrm{ASE}} \\ &\times\int e_{s_{2,o}}(\tau)h_e(t-\tau)\int e_{s_{2,o}}^*(\tau')h_e(t-\tau')r_o(\tau-\tau')\mathrm{d}\tau'\mathrm{d}\tau.\end{aligned} \tag{A.34}$$

If one assumes that the signal is polarized, then the unit Jones vector that defines the polarization state of the signal $\hat{\mathbf{e}}_s$ is time-independent, so that $e_{s_{1,o}}(t) =$

$(\hat{\mathbf{e}}_s \cdot \hat{\mathbf{a}}_1)\, e_{s_o}(t)$ and $e_{s_{2,o}}(t) = (\hat{\mathbf{e}}_s \cdot \hat{\mathbf{a}}_2)\, e_{s_o}(t)$ are the optically filtered parts of the noise-free signal decomposed in the $\hat{\mathbf{a}}_1$ and $\hat{\mathbf{a}}_2$ orthogonal polarization states. Using these results, I can rearrange (A.34) as

$$\begin{aligned} \sigma^2_{\mathrm{S-ASE}}(t) = \; & R^2 N_{\mathrm{ASE}} \left[(1 + \mathrm{DOP}_n) \left| \hat{\mathbf{e}}_s \cdot \hat{\mathbf{a}}_1 \right|^2 + (1 - \mathrm{DOP}_n) \left| \hat{\mathbf{e}}_s \cdot \hat{\mathbf{a}}_2 \right|^2 \right] \\ & \times \int e_{s_o}(\tau) h_e(t - \tau) \int e^*_{s_o}(\tau') h_e(t - \tau') r_o(\tau - \tau') \mathrm{d}\tau' \mathrm{d}\tau, \end{aligned} \tag{A.35}$$

which completes the derivation of (4.28).

Bibliography

[1] R. Newman, "Exitation of the Nd^{3+} fluorescence in $CaWO_4$ by recombination radiation in GaAs," *J. Appl.Phys.*, vol. 34, pp. 437–438, 1963.

[2] R. J. Keyes and T. M. Quist, "Injection luminescent pumping of CaF_2:U^{2+} with GaAs diode lasers," *Appl. Phys. Lett.*, vol. 4, pp. 50–52, 1964.

[3] F. P. Kapron, D. B. Keck, and R. D. Maurer, "Radiation losses in glass optical waveguides," *Appl. Phys. Lett.*, vol. 17, pp. 423–425, 1970.

[4] R. J. Mears, L. Reekie, I. M. Jauneey, and D. N. Payne, "Low-noise erbium-doped fibre amplifier operating at 1.45 μm," *Electron. Lett.*, vol. 23, pp. 1026–1028, 1987.

[5] I. P. Kaminow and I. L. Koch, *Optical Fiber Telecommunications*, vol. III-A, San Diego, CA: Academic, 1997.

[6] H. Bulöw, W. Baumert, H. Schmuck, F. Mohr, T. Schulz, F. Küppers, and W. Weiershausen, "Measurement of the maximum speed of PMD fluctuation in installed field fiber," in *Proceedings OFC 1999*, paper WE4, 1999.

[7] 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.

[8] D. Wang and C. R. Menyuk, "Calculation of penalties due to polarization effects in a long-haul WDM system using a Stokes parameter model," *J. Lightwave Technol.*, vol. 19, pp. 487–494, 2001.

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

[10] I. P. Kaminow and I. L. Koch, *Optical Fiber Telecommunications*, vol. III-B, San Diego, CA: Academic, 1997.

[11] 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.

[12] 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.

[13] R.-M. Mu, V. S. Grigoryan, C. R. Menyuk, G. M. Carter, and J. M. Jacob, "Comparison of theory and experiment for dispersion-managed solitons in a recirculating fiber loop," *IEEE J. of Selected Topics in Quantum Electron.*, vol. 6, pp. 248–257, 2000.

[14] M. G. Taylor, "Improvement in Q with low frequency polarization modulation on transoceanic EDFA link," *IEEE Photon. Technol. Lett.*, vol. 6, pp. 860–862, 1994.

[15] N. S. Bergano, F. W. Kerfoot, and C. R. Davidson, "Margin measurements in optical amplifier systems," *IEEE Photon. Technol. Lett.*, vol. 5, pp. 304–306, 1993.

[16] I. T. Lima Jr., A. O. Lima, J. Zweck, and C. R. Menyuk, "Efficient computation of outage probabilities due to polarization effects in a WDM system using a reduced Stokes model and importance sampling," *IEEE Photon. Technol. Lett.*, vol. 15, pp. 45–47, 2003.

[17] P. M. Hahn and M. C. Jeruchim, "Developments in the theory and application of importance sampling," *IEEE Trans. Commun.*, vol. 35, pp. 706–714, 1987.

[18] D. Marcuse, "Derivation of analytical expressions for the bit-error probability in lightwave systems with optical amplifiers," *J. Lightwave Technol.*, vol. 8, pp. 1816–1823, 1990.

[19] P. A. Humblet and M. Azizoğlu, "On the bit error rate of lightwave systems with optical amplifiers," *J. Lightwave Technol.*, vol. 9, pp. 1576–1582, 1991.

[20] P. J. Winzer, M. Pfennigbauer, M. M. Strasser, and W. R. Leeb, "Optimum filter bandwidth for optically preamplified NRZ receivers," *J. Lightwave Technol.*, vol. 19, pp. 1263–1273, 2001.

[21] Y. Sun, I. T. Lima Jr., A. O. Lima, H. Jiao, L. Yan, J. Zweck, C. R. Menyuk, and G. M. Carter, "Effects of partially polarized noise in a receiver," in *Proceedings OFC 2003*, paper MF82, 2003.

[22] Y. Sun, A. O. Lima, I. T. Lima Jr., L. Yan, J. Zweck, C. R. Menyuk, and G. M. Carter, "Accurate Q-factor distributions in optical transmission systems with polarization effects," in *Proceedings OFC 2003*, paper ThJ4, 2003.

[23] H. Jiao, I. T. Lima Jr., A. O. Lima, Y. Sun, J. Zweck, L. Yan, C. R. Menyuk, and G. M. Carter, "Experimental validation of an accurate receiver model for systems with unpolarized noise," in *Proceedings CLEO 2003*, paper CThJ1, 2003.

[24] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*, New York: John Wiley & Sons, 1991.

[25] M. Born and E. Wolf, *Principles of Optics*, Cambridge, UK: Cambridge University Press, 1999.

[26] J. D. Jackson, *Classical Electrodynamics*, New York, NY: John Wiley & Sons, 1999.

[27] 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.

[28] I. P. Kaminow, "Polarization in optical fibers," *IEEE J. Quantum Electron.*, vol. 17, pp. 15–22, 1981.

[29] C. R. Menyuk, "Application of multiple-length-scale methods to the study of optical fiber transmission," *J. Eng. Math.*, vol. 36, pp. 113–136, 1999.

[30] E. Golovchenko, "The challenges of designing long-haul WDM systems," in *Tutorial Sessions OFC 2002*, paper TuL, 2002.

[31] E. A. Golovchenko, A. N. Pilipetskii, N. S. Bergano, C. R. Davidson, F. I. Khatri, R. M. Kimball, and V. J. Mazurczyk, "Modeling transoceanic fiber-optic WDM communication systems," *IEEE J. Select. Topics Quantum Electron.*, vol. 6, pp. 224–226, 2000.

[32] G. Poole and D. L. Favin, "Polarization dispersion measurements based on transmission spectra through a polarizer," *J. Lightwave Technol.*, vol. 12, pp. 917–929, 1994.

[33] G. Biondini, W. L. Kath, and C. R. Menyuk, "Importance sampling for polarization-mode dispersion," *IEEE Photon. Technol. Lett.*, vol. 14, pp. 310–312, 2002.

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

[35] H. Goldstein, *Classical Mechanics*, Reading, MA: Addison-Wesley, 1980.

[36] F. Corsi, A. Galtarossa, and L. Palmieri, "Polarization mode dispersion characterization of single-mode optical fiber using backscattering technique," *J. Lightwave Technol.*, vol. 16, pp. 1832–1843, 1998.

[37] D. Wang, *Polarization effects in dense WDM systems*, Ph.D. Dissertation, University of Maryland Baltimore County, Baltimore, MD, May 2000.

[38] G. P. Agrawal, *Fiber Optic Communication Systems*, New York: Wiley-Interscience, 1997.

[39] L. Kazovsky, S. Benedetto, and A. Willner, *Optical Fiber Communication Systems*, Norwood, MA: Artech House, 1996.

[40] C. R. Menyuk, B. S. Marks, I. T. Lima Jr., J. Zweck, Y. Sun, G. M. Carter, and D. Wang, "Polarization effects in long-haul undersea systems," *Undersea Fibre Communication Systems*, José Chesnoy, ed., San Diego, CA: Academic Press, 2002, Sec. II.D.

[41] G. Arfken, *Mathematical Methods for Physicists*, San Diego, CA: Academic Press, 1985.

[42] A. Carena, V. Curri, R. Gaudino, P. Poggiolini, and S. Benedetto, "A time-domain optical transmission system simulation package accounting for nonlinear and polarization-related effects in fiber," *J. Sel. Areas Commun.*, vol. 15, pp. 751–765, 1997.

[43] R. Holzlöhner, C. R. Menyuk, V. S. Grigoryan, and W. L. Kath, "Accurate calculation of eye diagrams and error rates in long-haul transmission systems using linearization," *J. Lightwave Technol.*, vol. 20, pp. 389–400, 2002.

[44] I. P. Kaminow and I. L. Koch, *Optical Fiber Telecommunications*, vol. III-B, San Diego, CA: Academic, 1997.

[45] I. T. Lima Jr., A. O. Lima, J. Zweck, and C. R. Menyuk, "Performance characterization of chirped return-to-zero modulation format using an accurate receiver model," *IEEE Photon. Technol. Lett.*, vol. 15, pp. 608–610, 2003.

[46] R. Holzlöhner, C. R. Menyuk, W. L. Kath, and V. S. Grigoryan, "Efficient and accurate calculation of eye diagrams and bit-error rates in a single-channel CRZ system," *IEEE Photon. Technol. Lett.*, vol. 14, pp. 1079–1081, 2002.

[47] E. Desurvire, *Erbium Doped Fiber Amplifiers*, New York, NY: John Wiley & Sons, 1994.

[48] P. R. Trischitta and E. L. Varma, *Jitter in Digital Transmission Systems*, Boston: Artech House, 1989.

[49] R. E. Walpole and R. H. Myers, *Probability and Statistics for Engineers and Scientists*, New York, NY: Macmillan Publishing Co., 1993.

[50] R.-M. Mu., T. Yu, V. S. Grigoryan, and C. R. Menyuk, "Dynamics of the chirped return-to-zero modulation format," *J. Lightwave Technol.*, vol. 20, pp. 47–57, 2002.

[51] R. Y. Rubinstein, *Simulation and the Monte Carlo Method*, New York: John Wiley & Sons, 1981.

[52] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems: Modeling, Methodology, and Techniques*: Plenum Publishing Corp., 2001.

[53] R. P. Brent, *Algorithms for Minimization without Derivatives*, Englewood Cliffs, NJ: Pretice Hall, 1973.

[54] E. Veach, *Robust Monte Carlo methods for light transport simulation*, Ph.D. Dissertation, Stanford University, Palo Alto, CA, December 1997.

[55] I. T. Lima Jr., A. O. Lima, G. Biondini, C. R. Menyuk, and W. L. Kath, "A comparative study of single-section polarization-mode dispersion compensators," *J. Lightwave Technol., Special Issue—PMD*, to appear in May 2003.

[56] L.-S. Yan, Q. Yu, Y. Xie, and A. E. Willner, "Experimental demonstration of the system performance degradation due to the combined effects of polarization-dependent loss with polarization-mode dispersion," *IEEE Photon. Technol. Lett.*, vol. 14, pp. 224–226, 2002.

[57] Y. Sun, I. T. Lima Jr., H. Jiao, J. Wen, H. Xu, H. Ereifej, G. Carter, and C. R. Menyuk, "Study of system performance in a 107 km dispersion-managed recirculating loop due to polarization effects," *IEEE Photon. Technol. Lett.*, vol. 13, pp. 966–968, 2001.

[58] P. K. A. Wai, C. R. Menyuk, and H. H. Chen, "Stability of solitons in randomly varying birefringent fibers," *Optics Lett.*, vol. 16, pp. 1231–1233, 1991.