Thesis for the degree of Doctor of Philosophy

Channel-Aware Multilevel Coded Modulation for Coherent Fiber-Optic Communications

Lotfollah Beygi

CHALMERS

Communication Systems Group
Department of Signals and Systems
Chalmers University of Technology
Gothenburg, Sweden 2013

Channel-Aware Multilevel Coded Modulation

for Coherent Fiber-Optic Communications

Lotfollah Beygi
ISBN 978-91-7385-873-1

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny Serie No. 3554
ISSN 0346-718X

Front cover illustration:
First step of four-dimensional set partitioning of
a 16-ary signal set (see [Paper E] for details).

This thesis has been prepared using LaTeX.

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

Printed by Chalmers Reproservice
Gothenburg, Sweden, August 2013

To my wife

Abstract

The past decades have shown an ever-increasing demand for high-rate Internet services, motivating a great effort to increase the spectral efficiency of optical networks. In general, fiber-optic links are non-Gaussian, and in contrast to additive white Gaussian noise (AWGN) channels, there is no standard framework for quantifying fundamental limits or designing capacity-approaching coding schemes for such channels. In this thesis, some steps are taken toward this challenging goal by first developing a channel model for fiber-optic links and, second, using an information-theoretic design framework to investigate joint design of forward error correction and multilevel modulation, so-called coded modulation (CM), techniques for these channels.

We extend the signal statistics of highly nonlinear single-polarization fiber-optic links with negligible dispersion to the polarization-multiplexed case. Taking chromatic dispersion into account, we derive an analytical discrete-time model for single-wavelength, polarization-multiplexed, non-dispersion-managed (non-DM) links. According to this model, for high enough symbol rates, a fiber-optic link can be described as a linear dispersive channel with AWGN and a complex constant scaling.

We exploit the proposed channel model for highly nonlinear fiber-optic links to devise a new channel-aware multilevel CM scheme based on the minimization of the total block error rate. We introduce a CM system with an N-dimensional constellation constructed from the Cartesian product of N identical one-dimensional constellations. The multidimensional scheme shows better trade-off between complexity and performance than a one-dimensional multilevel CM scheme.

By invoking the introduced channel model for the dispersive non-DM links, we present a four-dimensional CM scheme, which shows a better trade-off between digital signal processing complexity and transparent reach than existing methods. This CM scheme together with a probabilistic signal shaping method is used to devise a rate-adaptive scheme with a single low-density parity-check code. The performance evaluation of the proposed CM scheme for a single-channel transmission fiber-optic system justifies the improvement of the system spectral efficiency for a wide range of transparent reaches, observing more than 1 dB performance gain compared to existing methods.

Keywords:
Fiber-optic communications, channel model, nonlinear phase noise, low-complexity detector, coded modulation, multidimensional set partitioning, signal statistics.

Acknowledgements

> The mediocre teacher tells. The good teacher explains. The superior teacher demonstrates. The great teacher inspires.
>
> -William Arthur Ward

To my supervisor, Professor Erik Agrell, thank you for the insightful advices, patient guidance, and generous support throughout the whole period of my PhD journey. I gained immensely from your excellent academic supervision. Putting aside the academic aspect, I was very lucky to have such a considerate and patient adviser who understood and encouraged me in times of difficulty. I am also very grateful to have Professor Magnus Karlsson as my co-supervisor. I benefitted greatly from his profound knowledge, challenging and motivating suggestions that contributed to improving my work and making this PhD thesis much richer.

Interacting with FORCE members from S2 and MC2 departments provided me with a unique research environment to work on such an interdisciplinary field. In particular, I would like to thank Professor Pontus Johannisson for making himself always available for discussions, sharing ideas, and enlightening fundamental concepts of optical fibers.

I would like to give thanks to Professor Erik Ström for providing an enthusiastic research environment in the Communication Systems and Information Theory group and the opportunity of the teaching assistant experience in his Wireless Communications course. His profound intuition and insightful comments enriched my understanding in this area extensively.

I am indebted to many of the current and former PhD students and postdoctoral fellows of the Communication Systems group at Chalmers for being open to discuss my questions and helping me in fixing my problems with IT++ and using the cluster. I am grateful to Arne Svensson, Thomas Eriksson, Giuseppe Durisi, and Henk Wymeersch for very helpful discussions in PhD courses as well as Fredrik Brännström, Tommy Svensson, Alexandre Graell i Amat, and all other people who have helped and inspired me during my doctoral study. For helping me with administrative tasks related to this thesis, I send my gratitude to Agneta and Natasha. They have helped me to the greatest extent possible. I would like to thank Lars Börjesson for the computer and software support. I also thank Nima Jamaly for helping me in using the LaTex template (originally belongs to the Signal Processing Group) of this thesis and Sima

Shahsavari for preparation of the first pages of the appended papers of this template.

I wish to extend my deep appreciation to Professor Joseph M. Kahn from Stanford University for giving me a research visit opportunity to his group. It was a great experience for me to learn from his clarity of thought and fundamental attitude toward research problems. I thank Olivier Rival and his colleagues for hosting me at Alcatel-Lucent in Paris and fruitful discussions related to industrial aspects of fiber-optic communications. I am grateful to Professor Alberto Bononi for providing me the possibility of spending a summer with his group in Parma, during which I enjoyed tapping into his great knowledge, meticulous attitude toward technicalities, and many inspiring ideas in our fruitful discussions together with Paolo Serena. Along my research, I have benefitted extensively from collaborations and discussions with several great researchers at Ericsson in Gothenburg.

Last but not least, my deepest gratitude goes to my family for their unconditional love and support; this work would not have been possible without them.

This work has been supported by the Swedish Research Council (VR) under Grant 2007-6223, the Swedish Foundation for Strategic Research (SSF) under grant RE07-0026, and the Swedish Governmental Agency for Innovation Systems (VINNOVA/CELTIC) under Grant 2007-02930. The numerical simulations in this thesis were performed in part on resources provided by the Swedish National Infrastructure for Computing (SNIC) at C3SE.

List of Publications

This thesis is based on the following appended papers:

Paper A
L. Beygi, E. Agrell, M. Karlsson, and P. Johannisson, "Signal Statistics in Fiber-Optical Channels with Polarization-Multiplexing and Self-Phase Modulation," *IEEE /OSA Journal of Lightwave Technology*, vol. 29, no. 16, pp. 2379–2386, August 2011.

Paper B
L. Beygi, E. Agrell, P. Johannisson, M. Karlsson, and H. Wymeersch, "A Discrete-Time Model for Fiber-Optical Channels," *IEEE Transactions on Communications*, vol. 60, no. 11, pp. 3440–3450, November 2012.

Paper C
L. Beygi, E. Agrell, P. Johannisson, and M. Karlsson, "A Novel Multilevel Coded Modulation Scheme for Fiber Optical Channel with Nonlinear Phase Noise," *in Proceedings of IEEE Global Communications Conference*, Miami, USA, December 2010.

Paper D
L. Beygi, E. Agrell, and M. Karlsson, "On the Dimensionality of Multilevel Coded Modulation in the High SNR Regime," *IEEE Communications Letters*, vol. 14, no. 11, pp. 1056–1058, November 2010.

Paper E
L. Beygi, E. Agrell, J. M. Kahn, and M. Karlsson, "Rate-Adaptive Coded Modulation for Fiber-Optic Communications," submitted to *IEEE/OSA Journal of Lightwave Technology*, June 2013.

List of Additional Related Papers

Publications by the author not included in this thesis:

1. L. Beygi, E. Agrell, M. Karlsson, and B. Makki, "A Novel Rate Allocation Method for Multilevel Coded Modulation," *in Proceedings of IEEE International Symposium on Information Theory*, pp. 1983–1987, June 2010.

2. L. Beygi, E. Agrell, and M. Karlsson, "Optimization of 16-point Ring Constellations in the Presence of Nonlinear Phase Noise," *in Proceedings of Optical Fiber Communication Conference*, Los Angeles, CA, USA, March 2011.

3. L. Beygi, E. Agrell, P. Johannisson, M. Karlsson, and H. Wymeersch, "The Limits of Digital Backpropagation in Nonlinear Coherent Fiber-Optical Links," *in Proceedings of European Conference and Exhibition on Optical Communication*, Amsterdam, The Netherlands, September 2012, P4.14.

4. L. Beygi, E. Agrell, and M. Karlsson, "Adaptive Coded Modulation for Nonlinear Fiber-Optical Channels," *in Proceedings of IEEE Global Communications Conference*, Anaheim, USA, December 2012.

5. D. A. A. Mello, A. N. Barreto, T. C. Lima, T. F. Portela, L. Beygi and J. M. Kahn, "Optical Networking with Variable-Code-Rate Transceivers," submitted to *IEEE/OSA Journal of Lightwave Technology*, April 2013.

6. L. Beygi, E. Agrell, J. M. Kahn, and M. Karlsson, "Coded Modulation for Fiber-Optical Channels," submitted to *IEEE Signal Processing Magazine (special issue)*, May 2013.

7. L. Beygi, C. Häger, E. Agrell, P. Johannisson, M. Karlsson, and A. G. i Amat, "A Low-Complexity Detector for Memoryless Polarization-Multiplexed Fiber-Optical Channels," under revision for publication in *IEEE Communications Letters*, July 2013.

8. L. Beygi, N. V. Irukulapati, E. Agrell, P. Johannisson, M. Karlsson, H. Wymeersch, P. Serena, and A. Bononi, "On nonlinearly-induced noise in optical links with digital backpropagation," submitted to *Optics Express*, June 2013.

Contents

Abstract i

Acknowledgments iii

List of Publications v

List of Additional Related Papers vii

Contents ix

Acronyms xv

Part I: Introduction 1

1 Background 3

1.1 Forward error correction in optical communications 4

1.2 Coded modulation for fiber-optic channels 5

1.3 Organization of the thesis 6

2 Channel Modeling of Fiber-Optic Links 7

2.1 Propagation of light in fiber-optic channels 7

2.1.1 Signal attenuation 9

2.1.2 Amplifier noise 9

2.1.3 Chromatic dispersion 10

2.1.4 Nonlinear Kerr effect 10

2.2 Fiber-optic channel simulation techniques 11

2.2.1 Split-step Fourier method 11

2.2.2 First-order perturbation technique 11

2.2.3 Continuous-time model 12

2.2.4 Discrete-time model 13

2.3 Dispersion-managed fiber-optic links . . . 14
2.3.1 Zero-dispersion model . . . 15
2.3.2 Statistics of the zero-dispersion model . . . 16
2.3.3 Nonlinear phase noise compensation . . . 17
2.4 Non-dispersion-managed fiber-optic links . . . 20
2.4.1 Gaussian noise model . . . 20
2.4.2 Gaussian-noise based channel models . . . 21
2.5 Numerical justification . . . 23
2.5.1 Gaussian assumption . . . 23
2.5.2 Nonlinear Gaussian distortion characterization . . . 25
2.6 Alternative channel models . . . 29
2.6.1 Volterra-series method . . . 29
2.6.2 Nonlinear Fourier transform method . . . 30

3 Channel-Aware Coded Modulation 31
3.1 System constraints . . . 32
3.1.1 Channel capacity . . . 32
3.1.2 Operating transmit power . . . 33
3.1.3 Quality parameters . . . 34
3.2 Design framework . . . 35
3.2.1 Multilevel coded modulation (MLCM) . . . 36
3.2.2 Bit-interleaved coded modulation (BICM) . . . 38
3.2.3 Trellis-coded modulation (TCM) . . . 38
3.2.4 Nonbinary coded modulation . . . 39
3.2.5 Polar nonbinary coded modulation . . . 39
3.3 Coded modulation for dispersion-managed links . . . 40
3.3.1 Concatenated codes . . . 41
3.4 Two- versus four-dimensional schemes . . . 42
3.4.1 Two-dimensional schemes with binary codes . . . 42
3.4.2 Two-dimensional schemes with nonbinary codes . . . 42
3.4.3 Four-dimensional schemes with binary codes . . . 43
3.4.4 Four-dimensional schemes with nonbinary codes . . . 43
3.5 Rate-adaptive coded modulation schemes . . . 43
3.5.1 Multiple codes with different rates . . . 44
3.5.2 Single fixed-rate code . . . 44
3.5.3 Signal shaping . . . 46
3.6 Performance and complexity analysis . . . 48
3.6.1 Latency-constrained comparison . . . 48
3.6.2 Complexity-constrained comparison . . . 49
3.6.3 Performance improvement using probabilistic shaping . . . 50

3.6.4 Hardware complexity 51

4 Conclusions **53**

References **57**

Part II: Publications **69**

Paper A: Signal Statistics in Fiber-Optical Channels with Polarization Multiplexing and Self-Phase Modulation **71**

Abstract 73
1 Introduction 73
2 System model 75
2.1 Fiber-optic channel with low dispersion 75
2.2 Dispersion-managed fiber-optical channel 76
3 Nonlinear phase noise 77
3.1 Distributed amplification 77
3.2 Lumped amplification 78
4 The joint pdf of the received amplitudes and phases of the DP signal . 79
5 Numerical results 81
5.1 Fiber-optic channel with low dispersion 81
5.2 Dispersion-managed channel 83
6 Conclusion 85
Appendix 86
References 87

Paper B: A Discrete-Time Model for Uncompensated Single-Channel Fiber-Optical Links **91**

Abstract 93
1 Introduction 93
2 Continuous-time model 95
3 Discrete-time model 97
4 Statistics of the propagated signal 98
4.1 Signal statistics for the case of strong dispersive effects 98
4.2 Signal statistics for a segment length applicable to SSFM 99
5 Statistics of the received signal 100
6 Numerical results 102
7 Conclusion 105
Appendix A 106
Appendix B 107
Appendix C 108

Appendix D . . . 110
Appendix E . . . 112
References . . . 112

Paper C: A Novel Multilevel Coded Modulation Scheme for Fiber Optical Channel with Nonlinear Phase Noise 117
Abstract . . . 119
1 Introduction . . . 119
2 System model . . . 120
3 SER of a uncoded 16-point ring constellation . . . 122
4 Optimized MLCM scheme for NLPN . . . 124
4.1 Set partitioning in radial direction . . . 124
4.2 Set partitioning in phase direction . . . 126
5 Rate allocation of the MLCM scheme . . . 127
6 Simulation results . . . 128
7 Conclusion . . . 129
References . . . 130

Paper D: On the Dimensionality of Multilevel Coded Modulation in the High SNR Regime 133
Abstract . . . 135
1 Introduction . . . 135
2 System model . . . 136
3 ACG of MLCM systems . . . 136
4 Multidimensional set partitioning . . . 138
5 Complexity and performance comparison . . . 139
6 Conclusion . . . 140
References . . . 140

Paper E: Rate-Adaptive Coded Modulation for Fiber-Optic Communications 143
Abstract . . . 145
1 Introduction . . . 145
2 System model . . . 147
3 Information-theoretic design framework . . . 148
4 Bit-to-symbol mapper . . . 150
5 Polar coded modulation . . . 154
5.1 LDPC coding and decoding . . . 155
5.2 Probabilistic signal shaping . . . 155
6 Complexity analysis . . . 156
7 Numerical results . . . 157

8 Rate adaptation 158
9 Conclusion 161
References 161

Acronyms

1D:	one-dimensional
2D:	two-dimensional
4D:	four-dimensional
ADC:	analog-to-digital convertor
ASE:	amplified spontaneous emission
AWGN:	additive white Gaussian noise
bps:	bits per second
BCH:	Bose–Chaudhuri–Hocquenghem
BER:	bit-error ratio
BICM:	bit-interleaved coded modulation
CM:	coded modulation
CSI:	channel state information
DBP:	digital backpropagation
DCF:	dispersion compensation fiber
DM:	dispersion-managed
DSP:	digital signal processing
EDC:	electronic chromatic dispersion
EDFA:	erbium-doped fiber amplifier
FEC:	forward error correction
GF:	Galois field

HDD:	hard-decision decoding
ISI:	inter-symbol interference
LDPC:	low-density parity-check
LLR:	log-likelihood ratio
MAP:	maximum a posteriori probability
MED:	minimum Euclidean distance
MI:	mutual information
ML:	maximum likelihood
MLCM:	multilevel coded modulation
MSD:	multi-stage decoding
NCG:	net coding gain
NLPN:	nonlinear phase noise
NLSE:	nonlinear Schrödinger equation
OFDM:	orthogonal frequency division multiplexing
PAM:	pulse amplitude modulation
pdf:	probability density function
PM:	polarization-multiplexed
PSK:	phase shift keying
QAM:	quadrature amplitude modulation
RS:	Reed–Solomon
SDD:	soft-decision decoding
SER:	symbol-error ratio
SMF:	single mode fiber
SNR:	signal-to-noise ratio
SSFM:	split-step Fourier method
TCM:	trellis coded modulation
WDM:	wavelength-division-multiplexing

Part I

Introduction

Chapter 1

Background

Digital communication has obtained a vital role in the infrastructure of modern society, from multimedia broadcasting to advanced military communication systems. The foundation of rapidly growing communications technology, one of the greatest engineering achievements of the twentieth century, was established by Shannon in 1948 [1]. He introduced a mathematical proof that data transmission over a noisy channel is possible with an arbitrarily chosen low error rate at finite signal power as long as the transmission rate (coding rate) is not greater than a certain constant, called the capacity of the channel. Conversely, error-free transmission is impossible if one uses a transmission rate that exceeds the capacity of the channel, regardless of the exploited transmission scheme. However, Shannon did not introduce any practical approach to design such a capacity-achieving channel coding scheme. Since then, tremendous efforts have been devoted to devising such schemes for practical use. Traditional channel coding schemes, often called forward error correction (FEC), add some parity bits in order to use them for recovering the information bits after receiving them with some distortion. This technique reduces the required signal-to-noise ratio (SNR) to attain a desired bit-error ratio (BER) in a practical data transmission system. The reduction in the required SNR (or higher sensitivity) comes at the cost of higher computational digital signal processing (DSP) complexity.

Among different available data transmission systems, fiber-optic communication systems have introduced significant changes in the telecommunications industry, serving as a low-loss (0.2 dB/km) transmission medium with a huge bandwidth (larger than 50 THz with the center wavelength at 1.55 μm). These systems were first developed in the 1970s and have played a major role in the drastic development of information technology. Optical fibers have extensively replaced copper wire links in data networks, enabling higher data transmission rates [2, Ch. 1].

The optical signal decays along fiber link and vanishes in the channel noise after a certain transmission distance (greater than 80 km). Generally, for long links, optical amplifiers are used to keep the SNR above a certain recoverable level. The main

Table 1.1: Three generations of FEC schemes in fiber-optic communications

	FEC generation		
	First	**Second**	**Third**
Year	1990s	early 2000s	2004–now
Standard	ITU-T G.709 [10]	ITU-T G.975.1 [11]	ITU-T G.975.1
Coding	HDD RS	concatenated block and convolutional codes	iterative SDD LDPC and turbo product
Overhead	7%	7-11%	12.5–20%
SNR gain over uncoded system at BER $= 10^{-15}$	6 dB	7.5 dB	10.5 dB

limitation of this approach is the amplified spontaneous emission (ASE) noise added in each amplifier, which is modeled as additive white Gaussian noise (AWGN). Signal regeneration, i.e., converting the optical signal to an electrical signal, recovering the transmitted data, and converting it back to optical, is an alternative but very costly solution, especially in long-haul (e.g., submarine) systems. Many network owners prefer to invest in advanced DSP techniques to retain the current networks instead of handling the excessive cost of signal regeneration. In contrast to wireline and wireless channels (without analog relay), amplification noise is added to the transmitted signal during propagation in fiber-optic channels. Moreover, the more severe nonlinear effect of optical links causes interaction between the noise and the transmitted signal, giving rise to a signal-dependent noise-like distortion. As the demand for higher rates continues to increase rapidly at about 60% per year [3], many studies on the channel capacity as well as FEC schemes have been performed for fiber-optic links by taking into account the nonlinear behavior of these channels [4–9].

1.1 Forward error correction in optical communications

FEC schemes have been employed to obtain a reliable data transmission (BERs less than 10^{-15}) in fiber-optic links with ASE and nonlinear distortion, which have evolved over several generations. The three generations of FEC techniques [10, 11] in optical fiber systems are summarized in Table 1.1. The first generation used Reed–Solomon (RS) block codes with hard-decision decoding (HDD) in the mid-1990s, typically with 7% redundancy overhead [10]. Hard-decision decoders receive as input the quantized decisions without any indication of the reliability of this decision [12, Ch. 1]. The second-generation FEC schemes use concatenated inner and outer codes with HDD

from RS and Bose–Chaudhuri–Hocquenghem (BCH) codes [11]. The significant gain obtained by modern capacity-achieving codes for AWGN channels has inspired the devotion of much effort to adopt these codes in fiber-optic links. For example, low-density parity-check (LDPC) codes have up to 5 dB better performance than RS codes with HDD at BERs around 10^{-11} [13]. Therefore, many studies have been performed, addressing the limited analog-to-digital convertor (ADC) resolution, DSP computational complexity, performance improvement, and latency of binary and nonbinary LDPC codes for fiber-optic channels [14–16]. This has led to the third generation of FEC schemes in fiber-optic links, consisting of various types of codes with iterative soft-decision decoding (SDD) [11, 17]. SDD exploits greater than one bit ADC, resulting in more than two quantization levels. This group includes iterative soft-decision modern codes such as LDPC or turbo product codes. Codes with higher overhead (around 20%) were mostly used in long-haul systems [14]. In all these schemes, coding and modulation units operate independently.

1.2 Coded modulation for fiber-optic channels

Much effort was expended to jointly optimize the complexity and performance of FEC schemes. The eventual outcome was to jointly design coding and (multilevel) modulation, often called coded modulation (CM). The superiority of CM schemes over conventional schemes with independent FEC and modulation in providing a better trade-off between DSP complexity and performance was already known in the 1960s [18]. The combination of modulation and convolutional codes with soft Viterbi decoding was introduced by Ungerboeck and Csajka, who introduced trellis coded modulation (TCM) [19, 20]. Independently, Imai and Hirakawa proposed a multilevel coded modulation (MLCM) scheme based on multistage decoding [21]. The idea behind MLCM is to convert a single channel with a multilevel modulation input to parallel binary subchannels. These subchannels corresponding to the particular binary labeling need different error protections. The MLCM scheme introduces unequal error protection for the subchannels; e.g., the subchannel with higher capacity should be protected by a higher rate code. In the absence of channel state information (CSI), no distinction can be found between the subchannels. Therefore, Zehavi introduced bit-interleaved CM (BICM) simply by exploiting an interleaver to obtain the same protection over subchannels and exploit their diversity [22]. The performance of the above three main categories of CM schemes, together with nonbinary coding schemes as well as hybrids of these schemes, have been evaluated for fiber-optic links [16, 23–28].

However, the available CM schemes, originally designed for AWGN channels, need to be redesigned for fiber-optic links to take into account the signal-dependent nonlinear distortion, yielding a channel-aware CM scheme. Moreover, a dynamic or heterogeneous structure of optically switched mesh networks demands adaptive transceivers to

operate with different signal qualities, for which the required error protection provided by a CM scheme needs to vary with the uncoded performance of the link. The so-called rate-adaptive CM schemes provide enough flexibility to adapt the data transmission scheme to the CSI in these networks. In this thesis, we aim to shed light on how an optical transceiver can be designed to account for this requirement of optical networks and provide a better trade-off between (i) DSP complexity, (ii) transparent reach, i.e., the transmission distance of a fiber-optic link with no inline electrical signal regenerators, and (iii) spectral efficiency, i.e., the number of information bits sent in each polarization per symbol period, than existing methods.

1.3 Organization of the thesis

This thesis is organized in two main parts: Introduction and Publications. The first part is subdivided into four chapters, whereas the second part consists of the published or submitted research papers.

Part I is organized as follows: Chapter 2 is dedicated to describing fiber-optic links and the main linear and nonlinear impairments in these channels. In addition, the available mathematical channel models for fiber-optic links are reviewed in this chapter. The statistics of the received signal are studied with weak and strong nonlinearity. The main purpose is to provide an accurate description of random effects (distortions) in a fiber-optic link, which is used in Chapter 3 for designing a channel-aware CM scheme. Finally, the different analytical models are numerically evaluated.

In Chapter 3, we introduce an information-theoretic framework to analyze the CM techniques based on the channel models introduced in Chapter 2 for fiber-optic channels. The framework provides a comprehensive classification of CM techniques in the literature and a design guideline for a CM scheme with a high spectral efficiency. Moreover, four-dimensional (4D) CM schemes constructed for polarization-multiplexed (PM) signals are compared with two-dimensional (2D) schemes, using both binary and nonbinary component codes. To address the quest for adaptive, so-called elastic, optical transceivers, we investigate different rate-adaptive CM schemes. Furthermore, the joint design of 4D CM schemes with probabilistic shaping is introduced to even further increase the spectral efficiency of fiber-optic data transmission with reasonable complexity.

Finally, Chapter 4 concludes the discussion and provides a short description about the contributions made in Part II of this thesis. We also briefly point out the limitations inherent in our studies presented in the included papers.

The second part of this thesis contains the research papers in their published or submitted format. The papers are arranged to correlate with the organization of Part I, rather than chronologically.

Chapter 2

Channel Modeling of Fiber-Optic Links

An optical fiber is a cylindrical dielectric waveguide consisting of a core surrounded by a cladding layer. Light propagates along the axis of this channel by internal reflection. The optical signal will be confined to the core where the refractive index is higher than that of the cladding. Transverse modes exist because of boundary conditions imposed on the light by the fiber [29, Ch. 3]. The allowed modes can be determined by solving Maxwell's equations for the boundary conditions of a given fiber [30, Ch. 2]. Fibers with more than one mode are called multi-mode optical fibers. These types of fibers are usually used in noncoherent data transmission for short-range links. In contrast, the single-mode optical fibers (SMF), with the so-called "fundamental mode," are mostly exploited in coherent communication systems.

Light as an electromagnetic signal is modulated to convey information bits in a fiber-optic channel. Although the loss of a fiber is very low compared to other transmission media (around 0.2 dB/km), the optical transmitted signal experiences significant attenuation due to the long transmission distance. Therefore, for reliable data transmission without costly electrical signal regeneration, optical amplifiers are inserted periodically, typically every 50–100 km in a long-haul fiber link, as illustrated in Fig. 2.1. In this chapter, we will address deterministic and random effects in fiber-optic links by describing the propagation of light in these channels based on the Manakov equation [31]. Then, we introduce two types of optical links, constructed using inline optical or electronic channel compensation. Moreover, we analyze the signal statistics and the design of maximum a posteriori (MAP) detector for fiber-optic channels under certain conditions. Finally, we provide a review of known channel models for fiber-optic links with weak nonlinearity as well as a performance comparison of these models.

2.1 Propagation of light in fiber-optic channels

We consider the electric field $\mathbf{U}(t, z)$ with complex components $(U_{\mathrm{x}}, U_{\mathrm{y}})$, where t is the time coordinate in a co-moving reference frame and z is the propagation distance in

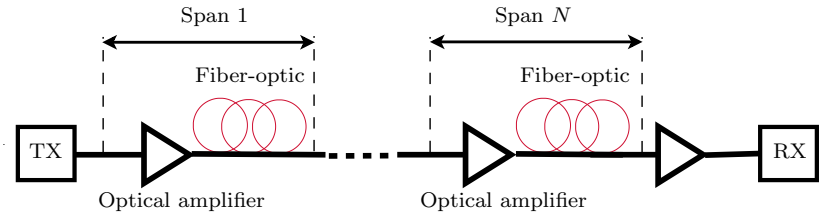


Figure 2.1: A fiber-optic link with N spans, each consisting of a fiber and an optical amplifier.

a PM fiber-optic link. For a monochromatic wave (a single frequency signal with a constant amplitude and phase), the propagation of the signal can be described as [30, Ch. 2]

$$\mathbf{U}(t,z) = \mathbf{U}(0,0)e^{j(\omega t - \beta z)},$$

where ω is the optical carrier frequency and β is the propagation constant (also known as refractive index). The dependence of β on the frequency and power of the signal that is propagating through the fiber gives rise to two main impairments in the fiber-optic link, namely chromatic dispersion and nonlinear Kerr effect, respectively. Before continuing further, we address how these two effects can be described by the Manakov equation[1], which governs the propagation of light in a fiber-optic link. To this end, one may decompose the propagation constant β using Taylor expansion around center frequency ω_0 into a frequency-dependent, a power-dependent, and a constant term as $\beta(\omega) = \beta_{\mathrm{L}}(\omega) + \beta_{\mathrm{NL}} + \frac{j}{2}\alpha$. The constant $\beta_2 = \partial^2\beta_{\mathrm{L}}/\partial\omega^2$ is defined as the group velocity dispersion coefficient. The loss coefficient α represents the signal attenuation caused by Rayleigh scattering and infrared absorption [32, Ch. 2]. The power-dependent term is written as $\beta_{\mathrm{NL}}(\omega) = \gamma P(t,z)$, where $P(t,z) = \mathbf{U}(t,z)\mathbf{U}(t,z)^{\dagger}$ is the instantaneous power of the propagating wave, in which $\dagger$ denotes the conjugate transpose, and γ is called the nonlinear coefficient [30, Ch. 2].

In this thesis, we neglect polarization mode dispersion caused by random imperfections and asymmetries of the fiber cross section. Moreover, we neglect laser phase and amplitude noise. Thus, the propagation of the electric field $\mathbf{U}(t,z)$ can be described by the Manakov equation, with loss and amplification included, as [2, Ch. 3], [32, Ch. 2]

$$\frac{\partial \mathbf{U}(t,z)}{\partial z} + j\frac{\beta_2}{2}\frac{\partial^2 \mathbf{U}(t,z)}{\partial t^2} - j\gamma P(t,z)\mathbf{U}(t,z) + \frac{\alpha - g(z)}{2}\mathbf{U}(t,z) = \mathbf{n}(t,z), \tag{2.1}$$

where $g(z)$ is the amplification factor and ASE noise $\mathbf{n}(t,z) = (n_{\mathrm{x}}(t,z), n_{\mathrm{y}}(t,z))$ is added in each amplifier. Since the available bandwidth in a fiber-optical link is much

[1]The Manakov equation is obtained from the nonlinear Schrödinger equation by averaging over random polarization rotations, which are assumed to be changing fast relative to other propagation effects [31].

larger than the signal bandwidth for optical-to-electrical convertors, the propagating electric field $\mathbf{U}$ may contain many channels with different center wavelengths, often called wavelength-division multiplexing (WDM). A typical channel spacing in commercial WDM systems is 37.5 or 50 GHz.

2.1.1 Signal attenuation

In the absence of chromatic dispersion (the second term), nonlinear Kerr effect (the third term), the amplification factor $g(z)$, and the noise (the last term) in (2.1), the optical signal decays along the fiber link as $\partial \mathbf{U}(t,z)/\partial z = -\alpha \mathbf{U}(t,z)/2$ or equivalently the signal power decays as $\partial P(t,z)/\partial z = -\alpha P(t,z)$. Thus, $P(t,z) = P(t,0)\exp(-\alpha z)$, where $P(t,0)$ is the instantaneous signal power at $z=0$. The loss coefficient α in terms of dB/km is expressed as α (dB/km) $= -10\log_{10}(P(t,z)/P(t,0))/z \approx 3.343\ \alpha$, which is around 0.2 dB/km for the wavelengths around 1.55 μm [30, Ch. 1].

2.1.2 Amplifier noise

Although fiber loss is very low, it makes the signal eventually vanish in the channel noise over long transmission distances. Therefore, as mentioned earlier, optical amplifiers are used periodically in a long fiber link. We call each period of the fiber-optic link a *span*, as illustrated in Fig. 2.1. Optical amplifiers add ASE noise $\mathbf{n}(t,z)$ as given in (2.1). Here, we describe the characteristics of this noise for periodic, often called lumped, as well as distributed amplification. With lumped amplification, each erbium-doped fiber amplifier (EDFA) compensates for the attenuation in each fiber span and adds independent circular white complex Gaussian ASE noise vector $\mathbf{n}(t,z)$ to the PM signal. The autocorrelation of the noise in two polarizations is [32, Ch. 2]

$$\mathbb{E}\{\mathbf{n}(t,z)\mathbf{n}^{\dagger}(t',z')\} = GF_n h\nu_{\text{opt}}\delta(t-t')\sum_{i=1}^{N}\delta(z-iL)\delta(z'-iL), \tag{2.2}$$

where $\delta(.)$ is the Dirac delta function, $G=\exp(\alpha L)$ is the required gain to compensate for the attenuation in each span of length L, $F_n = 2n_{\text{sp}}(1-G^{-1})$ is the noise figure, n_{sp} is ASE noise factor, and $h\nu_{\text{opt}}$ is the photon energy [33, eq. (8.1.15)]. The amplification factor $g(z)$ for lumped amplification is given by $g(z) = \alpha L\sum_{i=1}^{N}\delta(z-iL)$. A typical noise figure value for an EDFA is between 4 and 7 dB. The variance of the noise generated by each amplifier is $\sigma_0^2 = GF_n h\nu_{\text{opt}}W$, where W is the signal bandwidth, and the accumulated noise (excluding the amplifier at the receiver) is $\sigma_{\text{t}}^2 = N\sigma_0^2$ in two polarizations.

The distributed amplification is the asymptotic case of lumped amplification when the number of spans, N, tends to infinity (or equivalently, the span length tends to zero). A link with a total length L_{t} and distributed amplification can be modeled as a link with lumped amplification consisting of N spans of length $L = L_{\text{t}}/N$ when

$N \to \infty$. Thus, it can be shown that the variance of the total accumulated noise $\int_0^{L_\mathrm{t}} n(t,z)\mathrm{d}z$ generated by inline amplifiers in two polarizations is $\sigma_\mathrm{t}^2 = \lim_{N\to\infty} N\sigma_0^2 = \alpha L_\mathrm{t} n_\mathrm{sp} h\nu_\mathrm{opt} W$. The amplification factor $g(z)$ for a distributed amplification is α, implying that signal power is constant during propagation (no decay). Moreover, using $\lim_{N\to\infty} NGF_n/L_\mathrm{t} = 2\alpha n_\mathrm{sp}$ and $\lim_{N\to\infty} L_\mathrm{t}/N \sum_{i=1}^{N} \delta(z - iL_\mathrm{t}/N)\delta(z' - iL_\mathrm{t}/N) = \delta(z-z')$, the autocorrelation of the noise for distributed amplification is obtained from (2.2) as

$$\mathbb{E}\{\mathbf{n}(t,z)\mathbf{n}^\dagger(t',z')\} = 2\alpha n_\mathrm{sp} h\nu_\mathrm{opt}\delta(t-t')\delta(z-z'). \tag{2.3}$$

As discussed, the noise variance for lumped amplification shows an exponential growth with the span length, while it grows linearly with the fiber length for distributed amplification. Thus, distributed amplification shows better performance than lumped for the same system parameters. In practice, distributed amplification can be realized via stimulated Raman scattering, referred to as Raman amplification [30, Ch. 8].

2.1.3 Chromatic dispersion

The chromatic dispersion causes a frequency-dependent phase shift in the frequency domain with no change in the amplitude of the spectrum (the second term of (2.1) with parameter β_2) [34, 35]. In other words, chromatic dispersion makes the different frequencies of a pulse travel at different speeds. Thus, the different spectral components, which are sent at the same time, arrive at the receiver at different times. This imposes pulse broadening in time and consequently inter-symbol interference (ISI) on the transmitted signal consisting of a train of pulses. In principle, chromatic dispersion can be modeled as an all-pass filter

$$H(\omega, z) = e^{-j\omega^2 \frac{\beta_2}{2} z}, \tag{2.4}$$

for a fiber of length z, which can introduce memory into the channel. The equivalent impulse response is $h(t,z) = e^{j(t^2/(2\beta_2 z))}/\sqrt{j2\pi\beta_2 z}$ [35]. The group velocity dispersion parameter is expressed based on the dispersion parameter β_2 as $D = -2\pi c/\lambda^2\beta_2$, where c is the speed of light, with numerical values around 17 ps/(nm $\cdot$ km) for the wavelength 1.5 μm. Finally, the dispersion length is defined as $L_\mathrm{D} = 1/(|\beta_2|W^2)$ [30, p. 55]. Compensation of chromatic dispersion can be performed (gradually) during propagation using inline optical dispersion compensation fibers (DCF) or in bulk at the receiver with electronic dispersion compensation (EDC).

2.1.4 Nonlinear Kerr effect

The third term in (2.1) with the nonlinear parameter γ represents the nonlinear Kerr effect. The origin of this effect is the power-dependent refractive index of the fiber. The Kerr effect shows no change in the amplitude of the signal but a power-dependent

phase shift in the time domain. In the absence of chromatic dispersion, amplification, and ASE noise, (2.1) can be analytically solved at time t, as

$$\mathbf{U}(t,z) = \mathbf{U}(t,0)e^{-\frac{\alpha}{2}z}e^{j\gamma L_{\text{eff}}(z)P(t,0)}, \tag{2.5}$$

where the so-called effective length is $L_{\text{eff}}(z) = (1-\exp(\alpha z))/\alpha$. According to (2.5), the power-dependent phase shift on the specific channel originates from the signal power of the same channel or, in the case of WDM, from other channels. The first one is called self-phase modulation and the latter is known as cross-phase modulation. The third-order distortion (cross-signal products of more than two channels) can also cause this phase shift. In fact, the so-called four-wave mixing appears when these spurious signals resulting from signal mixing fall in the desired channel. If the channel spacing is too close, then four-wave mixing is more likely to occur. We will refer to interference (the exponential term in (2.5)) caused by self-phase modulation as *intrachannel* interference, while the interference originated by cross-phase modulation and four-wave mixing is considered *inter-channel* interference. Moreover, in the presence of noise, the nonlinear Kerr effects can be categorized as signal–signal, signal–noise, and noise–noise interactions. The first term is deterministic, while the other two are random.

2.2 Fiber-optic channel simulation techniques

2.2.1 Split-step Fourier method

Although the Manakov equation accurately describes the propagation of light in a fiber-optic link, in general it is difficult to solve this equation analytically, particularly if neither chromatic dispersion nor the nonlinear Kerr effect is negligible. The split-step Fourier method (SSFM) can be exploited to numerically solve this equation by modeling each fiber span as the concatenation of some short pieces of fiber, called "segments," with linear and nonlinear effects, as shown in Fig. 2.2 [30, eq. (2.4.10)]. The length of each segment, h, should be chosen to be small enough so that the linear and nonlinear effects can act independently[2]. Since the chromatic dispersion filter $h(t, \Delta z)$ is implemented using the fast Fourier transform, it is a low complexity method for the numerical simulation of fiber-optic links.

2.2.2 First-order perturbation technique

In order to find an approximate analytical solution, a so-called first-order perturbation technique is used to decompose the electric field $\mathbf{U}$ into a linear $\mathbf{U}_{\text{L}} = (U_{\text{L,x}}, U_{\text{L,y}})$ and a small perturbative (first-order) term $\mathbf{U}_{\text{p}} = (U_{\text{p,x}}, U_{\text{p,y}})$ as

$$\mathbf{U}(t,z) = \mathbf{U}_{\text{L}}(t,z) + \mathbf{U}_{\text{p}}(t,z). \tag{2.6}$$

[2]For numerical simulations, the segment length is computed by $\Delta z_m = (\kappa L_{\text{N}} L_{\text{D}}^2)^{1/3}$, where m is the segment index, $\kappa = 10^{-4}$, and L_{N} is the nonlinear length [30, p. 55] of segment $m-1$ [36].

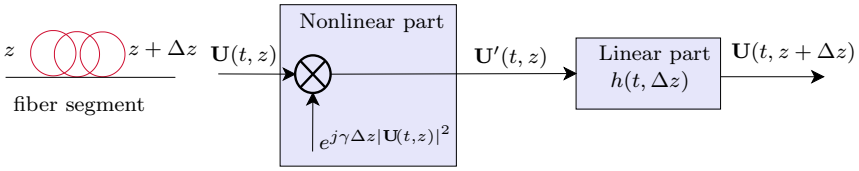


Figure 2.2: A model of a fiber segment as a concatenation of the linear and nonlinear part.

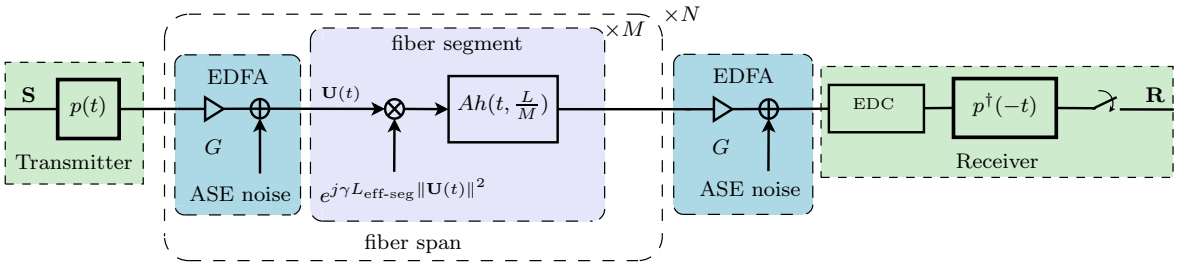


Figure 2.3: A baseband continuous-time model based on the SSFM for a PM fiber-optic link with N spans of fiber, each consisting of M segments, and EDC.

The linear term $\mathbf{U}_{\mathrm{L}}$ solves (2.1) for $\gamma = 0$ and it is assumed that $|U_{\mathrm{p,x}}| \ll |U_{\mathrm{L,x}}|$ and $|U_{\mathrm{p,y}}| \ll |U_{\mathrm{L,y}}|$ to obtain the first-order approximation $\mathbf{U}\mathbf{U}^{\dagger}\mathbf{U} \approx \mathbf{U}_{\mathrm{L}}\mathbf{U}_{\mathrm{L}}^{\dagger}\mathbf{U}_{\mathrm{L}}$.

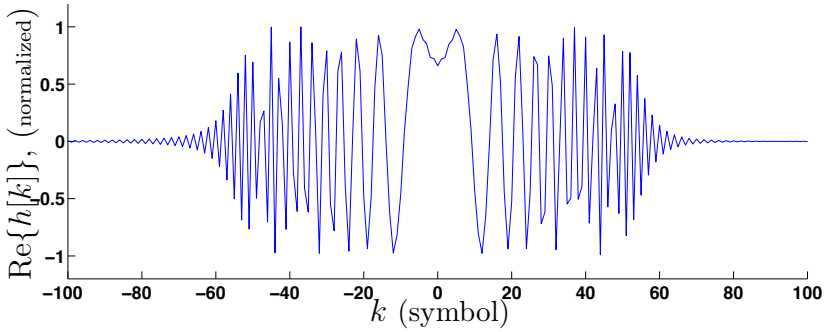
Here, by substituting (2.6) into (2.1), neglecting the amplifier noise, and using the first-order approximation, we obtain

$$\frac{\partial \mathbf{U}_{\mathrm{p}}}{\partial z} + j\frac{\beta_2}{2}\frac{\partial^2 \mathbf{U}_{\mathrm{p}}}{\partial t^2} - j\gamma \mathbf{U}_{\mathrm{L}}\mathbf{U}_{\mathrm{L}}^{\dagger}\mathbf{U}_{\mathrm{L}} + \frac{\alpha - g(z)}{2}\mathbf{U}_{\mathrm{p}} = \mathbf{0}. \tag{2.7}$$

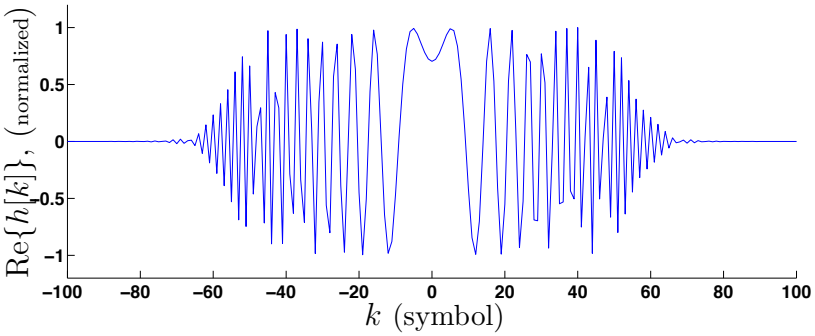
Since the linear term $\mathbf{U}_{\mathrm{L}}$ can be computed analytically, the first-order term $\mathbf{U}_{\mathrm{p}}$ is also obtained analytically by solving (2.7) under certain conditions on the electric field $\mathbf{U}$ [37–39].

2.2.3 Continuous-time model

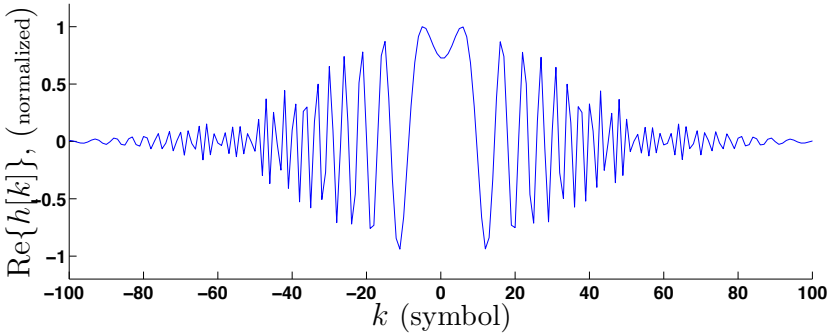
The continuous-time model of the fiber-optic link depicted in Fig. 2.1 is shown in Fig. 2.3 using the SSFM with M segments in each span. The effective length of each segment is $L_{\text{eff-seg}} = L_{\text{eff}}(L/M)$ (see Section 2.1.4). The linear propagation is modeled by a filter with impulse response $h(t, \frac{L}{M})$ performed independently over elements of the electric field vector, and $A \triangleq \exp[-\alpha L/(2M)]$ is the signal attenuation for each segment. The (uncoded) independent symbols $\mathbf{S} = (S_{\mathrm{x}}, S_{\mathrm{y}})$, e.g., PM quadrature phase shift keying (QPSK), are transmitted every T seconds with a pulse-shaping filter $p(t)$ and received as the distorted symbol sequence $\mathbf{R} = (R_{\mathrm{x}}, R_{\mathrm{y}})$ after the EDC, matched filtering, and Nyquist sampler with perfect carrier and timing synchronization. It is assumed that $\mathbb{E}\{|S_{\mathrm{x}}|^2\} = \mathbb{E}\{|S_{\mathrm{y}}|^2\} = PT$, where P is the average of the transmitted power $P(t, 0)$ in one polarization and T is the symbol period (for a sinc pulse, $W =$



(a) Sinc pulse



(b) Raised-cosine pulse with an excess bandwidth of 0.2



(c) Gaussian pulse with a spectral full width at half maximum of $1/(2T)$

Figure 2.4: The chromatic dispersion filter (real part) for a fiber-optic link with 10 spans of 80 km and $D = 17$ ps/(nm $\cdot$ km).

$1/T$). Each EDFA compensates for the attenuation in each fiber span by amplifying the signal with the gain $G = \exp(\alpha L)$ and adds ASE noise. The EDC is the inverse of the chromatic filter $h(t, NL)$.

2.2.4 Discrete-time model

The nonlinear effect prevents transforming the continuous-time model illustrated in Fig. 2.3 into an equivalent discrete-time model, convenient for the analysis of a digital communication system. In order to obtain a tractable discrete-time model, we consider

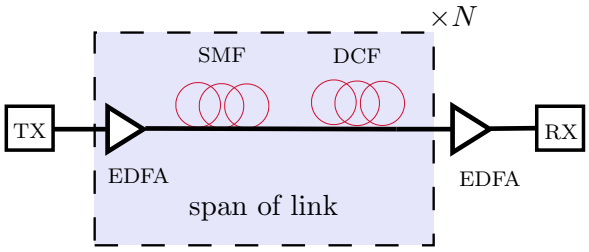
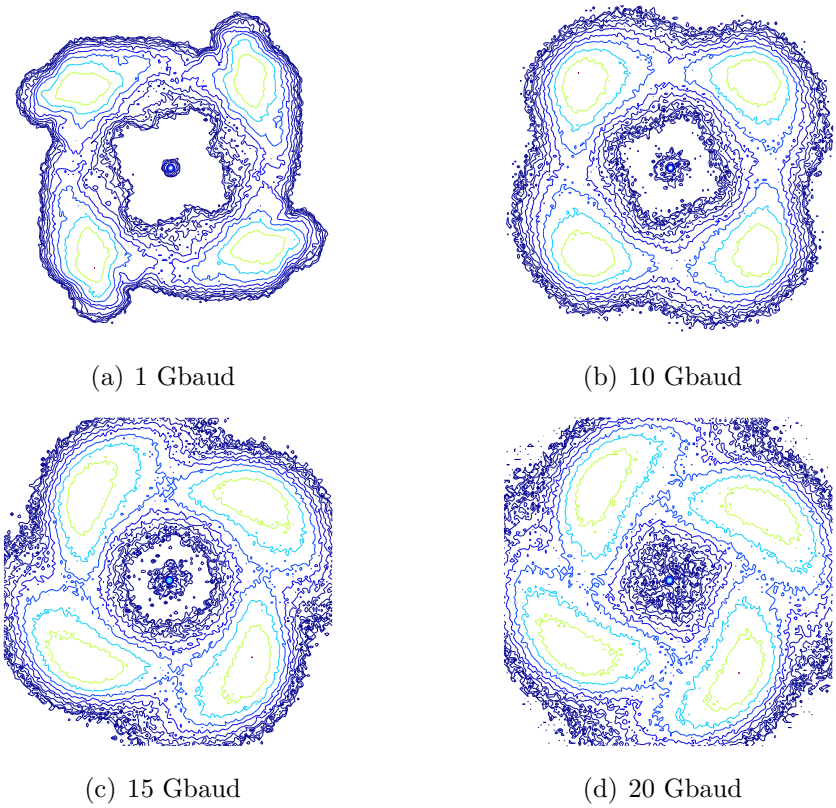


Figure 2.5: A DM fiber-optic link with N spans where each one consists of an amplifier, an SMF fiber, and a DCF fiber.

sinc-shaped pulses. However, the numerical results show that the discrete-time model is not very dependent on the exact pulse shape, e.g., raised-cosine and Gaussian pulses can be used as well. In the continuous-time model considering sinc pulse shape, the transmitted baseband signal $\mathbf{U}(t,0)$ is band-limited to $[-1/2T, 1/2T]$. One may neglect the spectral broadening due to the nonlinear effects provided that the nonlinearity is weak, i.e., the bandwidth of $\mathbf{U}(t,z)$ in (2.5) is assumed to be limited to $1/T$. This assumption helps us to fulfill the Nyquist criterion for sampling the continuous-time signals with a sampling rate of $1/T$ and obtain the discrete-time model. The discrete band-limited chromatic dispersion filter is $h[k]$, obtained by sampling $h\left(t, L/M\right) * \mathrm{sinc}\left(t/T\right)$ at $t = kT$. The real part of this filter coefficients are shown in Fig. 2.4 for three different pulse shapes: sinc, raised-cosine (with an excess bandwidth of 0.2 [40, Ch. 9]), and Gaussian (with a spectral full width at half maximum of $1/(2T)$) pulses. As seen, the ISI caused by the chromatic-dispersion filter with Gaussian pulses has a larger channel memory, $|k| > 70$, than raised-cosine or sinc pulses. This is simply because the bandwidth of the Gaussian pulses is larger.

2.3 Dispersion-managed fiber-optic links

A dispersion-managed (DM) link is characterized by using optical inline DCFs with a negative group velocity dispersion parameter (e.g., $D = -120$ ps/(nm $\cdot$ km)) in each span, to compensate for the chromatic dispersion caused by SMF as depicted in Fig. 2.5. Hence, the channel memory caused by dispersion can be removed, and the received signal is only distorted by the ASE noise added in each amplifier and the channel nonlinear effect. Figure 2.6 illustrates the pdf of the received signal for a DM link with four different symbol rates using numerical Monte-Carlo simulations. In general, the derivation of the statistics of the received signal based on this model is cumbersome. Therefore, we consider the following simplified model for a DM link to give some insight on the channel nonlinear effect, similar to what was done in [41, 42]. As will be shown shortly, the simplified model can explain these pdf plots for low dispersions (either small group velocity dispersion parameter or low symbol rate) very well.



(a) 1 Gbaud (b) 10 Gbaud

(c) 15 Gbaud (d) 20 Gbaud

Figure 2.6: The estimated pdf of the received signal for a DM fiber-optic link with of 25 spans each consisting of an SMF fiber of length 80 km with the dispersion coefficient $D = 17$ ps/(nm $\cdot$ km) and the nonlinear parameter $\gamma = 1.2 \times 10^{-3}$ $(\mathrm{Wm})^{-1}$ together with a DCF fiber with $D = -120$ ps/(nm $\cdot$ km) and a proper length to get chromatic dispersion fully compensated. The results are plotted for four symbol rates at 0 dBm transmit power.

2.3.1 Zero-dispersion model

We consider a fiber-optic link of total length L_t and N spans with no dispersion. The noise is assumed within an optical signal bandwidth, ignoring the Kerr effect induced from out-of-band signal and noise similar to [41]. The system model is first described for lumped amplification and then we extend it to the distributed case. For simplicity of notation, we denote with U and n the electric field and the amplifier noise, respectively, in polarization x, dropping the index x. The analysis is performed for a single-polarization case and the extension of the analytical results for PM case is provided in [Paper A]. By neglecting chromatic dispersion ($\beta_2 = 0$ for both SMF and DCF) and using the solution of (2.1) given in (2.5), we obtain [41, 43, 44]

$$U(t, L_\mathrm{t}) = (U(t, 0) + \sum_{k=1}^{N} n(t, kL_\mathrm{t}/N))e^{j\phi_\mathrm{n}}, \tag{2.8}$$

where $\phi_{\rm n} = \gamma L_{\rm eff}(L) \sum_{k=1}^{N} |U(t,0) + \sum_{i=1}^{k} n(t, kL_{\rm t}/N)|^2$ is called nonlinear phase noise[3] (NLPN). Here, we assumed the same nonlinear coefficient γ for the SMF and DCF.

The NLPN is generally believed to be a major impairment in long-haul optical transmission system [41,46,47]. By definition, the complex signal U is a time-dependent electric field, not a vector representation of the projected received electric field in a signal space. We nevertheless use (2.8) to model the discrete-time system. This is a standard approximation and has been shown numerically [48, 49] to be reasonably accurate, although the theoretical justification is insufficient. One may consider the distributed amplification as a lumped amplification with an infinite number of spans. This gives $\lim_{N\to\infty} NL_{\rm eff}(L_{\rm t}/N) = L_{\rm t}$. The NLPN can be computed for distributed amplification as $\phi_{\rm n} = \gamma \int_{\tau=0}^{L_{\rm t}} |U(t,0) + \int_0^{\tau} n(t,z){\rm d}z|^2 {\rm d}\tau$.

2.3.2 Statistics of the zero-dispersion model

In general, signal statistics are necessary in the design of a maximum likelihood receiver for a data transmission system. As mentioned before, the statistics of the received signal for a fiber-optic channel in general are unknown. In this section, we derive the statistics of a received signal for a zero-dispersion fiber studied in [45, Ch. 5], which is based on the system model described in Section 2.3.1. The joint pdf of the received amplitude and phase given the initial phase of the transmitted signal and the SNR was derived in [45, Ch. 5] for a fiber channel with NLPN caused by distributed or lumped amplification. Mecozzi also derived these statistics for distributed amplification in [50]. A similar model with no dispersion was introduced in [8] for noise with the same bandwidth as the received signal $U(t, L_{\rm t})$ rather than the launched signal $U(t, 0)$ to the fiber-optic link. Some other relevant models were studied in [51, 52]. In [Paper A], for the first time, we provide the statistics of the received signal for a PM fiber-optic link with negligible dispersion.

The characteristic function (i.e., the Fourier transform of the pdf) of NLPN for a single-polarization system has been studied analytically in [43, 50, 53, 54] by taking into account the correlation of the NLPN and the amplitude of the received signal. It was shown in [42, 43, 53] that the NLPN distribution cannot be approximated by a Gaussian distribution. Then we compute the pdf by taking the inverse Fourier transform of its characteristic function [55]. The characteristic functions are computed for distributed and lumped amplification separately [45, p. 157]. The pdf of the NLPN can then be computed by taking the inverse Fourier transform of the characteristic function. Figure 2.7 shows the pdf for two different SNRs, 10 and 20 dB, and two types of amplifications. The pdfs are plotted for lumped amplification with 10, 16, and 40 spans, $L_{\rm t} = 4000$ km, $\gamma = 1.2$ (W km)$^{-1}$, $\alpha = 0.25$ dB/km, SNR = 15 dB, and 42.7 Gbaud. As seen in this figure, the pdf for $N > 32$ spans is very close to the corresponding distributed system.

[3]The NLPN $\phi_{\rm n}$ in [45, p. 157] was normalized by $\gamma L_{\rm t}\sigma_{\rm t}^2$.

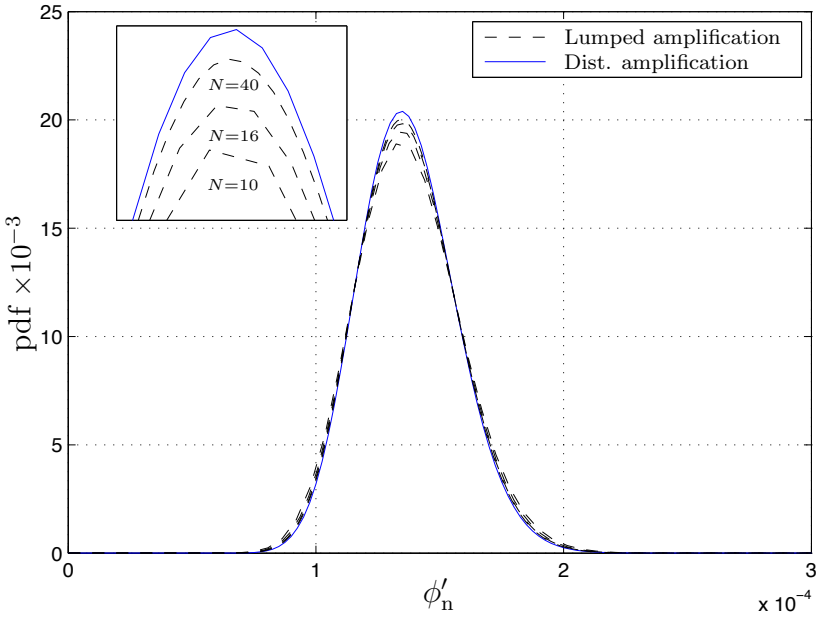


Figure 2.7: pdfs of the normalized NLPN for a fiber link with distributed ($\phi'_{\rm n} = \phi_{\rm n}/L_{\rm t}$) and lumped ($\phi'_{\rm n} = \phi_{\rm n}/(NL_{\rm eff}(L)$, $N = 10,\ 16,$ and 40) amplification.

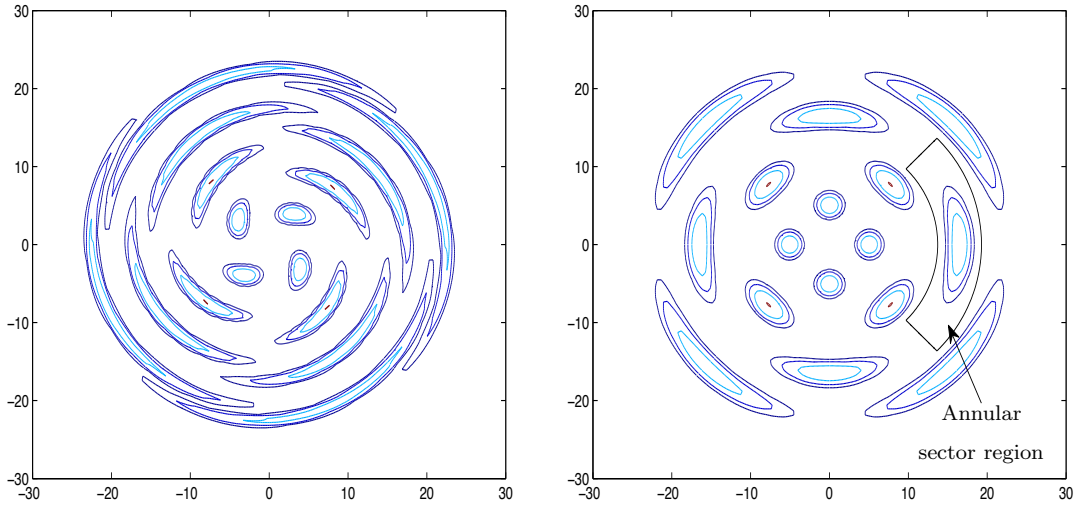
The characteristic functions can be used to derive the statistics of a single-polarization signal after propagation through a fiber channel for distributed and lumped amplifications. Before continuing further, we define the SNR ρ as $|U(t,0)|^2/\sigma_{\rm t}^2$ and $|U(t,0)|^2/(N\sigma_0^2)$ for distributed and lumped amplification, respectively, where σ_0 and $\sigma_{\rm t}$ were defined in Sec. 2.1.2. The normalized received amplitude r is defined as $|U(t,L_{\rm t})|/\sigma_{\rm t}$ and $|U(t,NL)|/(\sigma_0\sqrt{N})$ for distributed and lumped amplifications, respectively. The joint pdf of the received phase θ and the normalized amplitude r of a dispersion-managed fiber channel with distributed amplification is [45, p. 225]

$$f_{\Theta,R}(\theta,r) = \frac{f_R(r)}{2\pi} + \frac{1}{\pi}\sum_{k=1}^{\infty} {\rm Re}\left\{C_k(r)e^{jk(\theta-\theta_0)}\right\}, \tag{2.9}$$

where $f_R(r) = 2re^{-(r^2+\rho)}I_0(2r\sqrt{\rho})$ is the pdf of the Ricean random variable r and θ_0 is the initial transmitted phase. The Fourier series coefficients $C_k(r)$ are given in [45, p. 225]. Figure 2.8(a) shows the joint pdf of the amplitude and phase of the received signal for a 16-point constellation with the transmitted power of 0 dBm and the following channel parameters: $L = 5000$ km, $\gamma = 1.2$ (W km)$^{-1}$, $\alpha = 0.25$ dB/km, and 42.7 Gbaud [56]. These results were extended to PM in [Paper A].

2.3.3 Nonlinear phase noise compensation

In order to minimize the error probability of the system, suitable compensation techniques can be derived by exploiting the statistics of the received signal. The joint pdf



(a) (b)

Figure 2.8: The joint pdf of the amplitude and phase of (a) the received signal, (b) the complex signal $re^{j\theta'}$ after the NLPN compensation, for a 16-point constellation (four rings with four equally-spaced-phase points in each one). Values of contours are $10^{-2}, 10^{-1.5}, \ldots, 1$. [56]

(2.9) was exploited to compensate the NLPN from the received phase at the receiver in a single step [57–60]. However, some approaches have been proposed to mitigate the effect of NLPN by pre-distortion [60–62] at the transmitter.

A linear detector was proposed in [58] based on minimizing the variance of the residual NLPN. Moreover, a minimum mean square error detector was introduced in [45, Ch. 6] for NLPN compensation. Parallel to theoretical methods, some experimental schemes were demonstrated based on the linear minimum mean square error [59] and an optical compensation method in [63]. In this section, we describe the MAP detector for compensating NLPN. The derived pdf of the NLPN and the joint pdf of the amplitude and phase of the received signal can be exploited to show that only the received amplitude is needed to estimate the added NLPN in the channel. This MAP detector introduced in [57] is superior to the detector based on the linear compensator proposed in [64] in terms of performance (at the expense of higher complexity) and slightly better than the minimum mean square detector of [45, Ch. 6] with nearly the same complexity.

The MAP receiver can be derived exactly for an M-PSK signal set in which the phases of the signal alternatives are $\theta_k = (2k-1)\pi/M$, $k = 1, \ldots, M$. Since the simplified model with no dispersion is rotationally invariant and all the symbols have the same amplitude, the MAP detector boils down to the ML detector. One can readily derive the ML decision boundary [40] between the symbols with the phases θ_1 and θ_M, exploiting the joint pdf of the amplitude and phase of the received signal given that

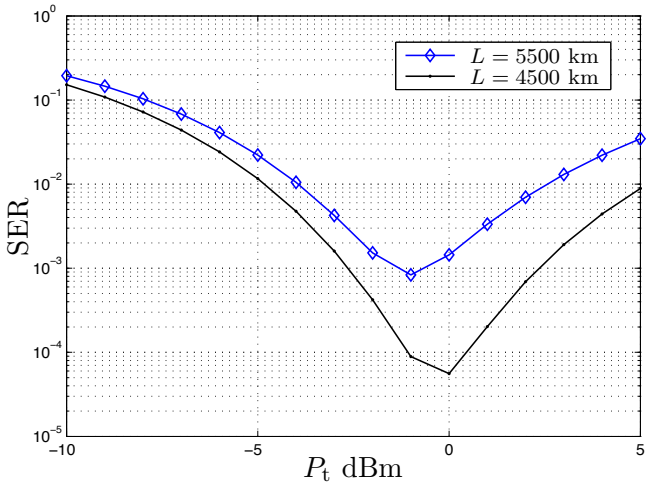


Figure 2.9: The SER of a dispersion-managed fiber link with $L = 4500$ and 5500 km.

the transmitted symbol is one of these two symbols. It was shown in [57] that the ML decision boundaries are rotated by

$$\theta_c(r) = -\angle\, C_1(r), \tag{2.10}$$

where $\angle$ denotes the phase of a complex number. This rotation is a nonlinear (second-order polynomial) function of the received signal amplitude r for a fiber channel with a distributed amplification. By symmetry, the decision boundaries between phases θ_k and θ_{k+1}, for $k = 1, \ldots, M$, are obtained as $\theta_c(r) + 2k\pi/M$, where θ_{M+1} is equivalent to θ_1. Moreover, we can derotate the decision boundaries between symbols to get almost straight line decision boundaries. This approach makes the receiver very simple to implement. In brief, the NLPN compensator computes the phase rotation of decision boundaries based on the received amplitude. This rotation is canceled out by multiplying the received signal $re^{j\theta}$ with $e^{j\theta_c(r)}$.

By using (2.9), (2.10), and together with the approximation $C_k(r) = kC_1(r)$ [57], the joint pdf of the amplitude and phase of the signal $re^{j\theta'} = re^{j(\theta-\theta_c(r))}$ (after the nonlinear compensator) is obtained by

$$f_{R,\Theta'}(r,\theta') \approx \frac{f_R(r)}{2\pi} + \frac{1}{\pi}\sum_{k=1}^{\infty}|C_k(r)|\cos\left(k\left(\theta' - \theta_0\right)\right). \tag{2.11}$$

Figure 2.8(b) shows the joint pdf of the amplitude and phase of the received signal after NLPN compensation ($re^{j\theta'}$) for a 16-point constellation with the transmitted power of 0 dBm and the following channel parameters: $L = 5000$ km, $\gamma = 1.2$ (W km)$^{-1}$, $\alpha = 0.25$ dB/km, and 42.7 Gbaud. As seen in Fig. 2.8(b), the decision boundaries are either straight lines or circular arcs, and the annular sector region (a sector in the area between the two concentric circles) can be used to perform the detection in two steps.

The symbol-error ratios (SER) for this system with the following radii of the rings[4]: $R_1 = 0.28\sqrt{P_\mathrm{t}}, R_2 = 0.66\sqrt{P_\mathrm{t}}, R_3 = 1.06\sqrt{P_\mathrm{t}},$ and $R_4 = 1.53\sqrt{P_\mathrm{t}}$ are shown in Fig. 2.9 for $L = 4500$ and 5500 km. In the high-SNR regime ($P_\mathrm{t} > 0$ dBm), unlike AWGN channels, the SER increases. This behavior is due to NLPN effect, which gives a major degradation at high transmitted powers. Finally, the aforementioned MAP detector can be used for quadrature amplitude modulation (QAM) signals provided that the amplitude is detected first and then the NLPN compensation is performed based on (2.10), leading to a suboptimal *two-stage detector*. In fact, a coded scheme for such a system can be designed provided that the statistics of the signal after compensation are known. For instance, the design of MLCM scheme based on the derived statistics in (2.11) was introduced in [Paper C].

2.4 Non-dispersion-managed fiber-optic links

Numerical and experimental results [65–67] revealed that bulk chromatic dispersion compensation, either optically or electronically, referred to as non-dispersion-managed (non-DM), provides better performance results for the same transmission length than gradual compensation using inline optical DCFs. This fact convinced system designers to use links consisting of SMFs with EDFA without inline optical DCFs, known as non-DM links, which have attracted a global interest in exploiting EDC rather than optical DCF for next-generation optical networks. More precisely, as it is shown in [66, 68, 69] and [Paper B], the accumulated dispersion in a non-DM fiber-optic link, in the absence of DCFs, turns the distribution of the electric field to a Gaussian distribution. Since the nonlinear effect is intimately related to the instantaneous power of the electric field during propagation, having the electric field shaped with a Gaussian distribution helps to mitigate the distortion caused by the nonlinear effects. Thanks to high-speed electronic DSP technology, optical links with symbol rates at 32 Gbaud can be realized with EDC [70]. Hence, a non-DM link with N spans of length L is considered according to Fig. 2.3. Each span consists of an SMF and an EDFA. Here, we address the Gaussian noise model depicted in Fig. 2.10 as an equivalent discrete-time model of a non-DM fiber-optic link.

2.4.1 Gaussian noise model

The distribution of the propagated signal in a non-DM fiber-optic link is shown to be approximately Gaussian, provided that the accumulated chromatic dispersion is large enough [66, 68, 69] (as also shown analytically in [Paper B]). Although the nonlinear effect of each segment (introduced in Fig. 2.3) changes the signal distribution to non-Gaussian, the channel dispersion in the succeeding segments turns it back to Gaussian.

[4]This radii distribution is selected by performing a numerical optimization to minimize the SER for transmitted power $P_\mathrm{t} = -4$ dBm [56].

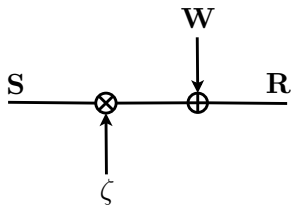


Figure 2.10: The introduced Gaussian noise model (ζ is a complex scalar) [Paper B].

Interestingly, the distortion caused by the nonlinear effect can be modeled as an additive Gaussian noise–like nonlinear distortion, sometimes called nonlinear noise. In other words, the nonlinear effect and the chromatic dispersion convert a part of the transmitted signal to an additive Gaussian noise. Therefore, the received signal can be represented by

$$\mathbf{R} = \zeta \mathbf{S} + \mathbf{W}, \tag{2.12}$$

where ζ is a complex constant that accounts for the signal depletion as well as average phase rotation, and $\mathbf{W}$ represents a PM complex zero-mean circularly symmetric AWGN in each polarization. The ASE noise as well as Gaussian noise–like inter-channel and intrachannel nonlinear interference (see Section 2.1.4) contribute to the noise term $\mathbf{W}$. As we discuss shortly, the complex constant ζ has been considered unity in some of the proposed analytical models, neglecting the signal depletion due to nonlinear effect [66]. This model, which is obtained for a channel with electronic dispersion compensation as depicted in Fig. 2.3, can be considered an approximation with nonlinear compensation techniques such as digital backpropagation (DBP) [71, 72]. Here, we review different channel modeling techniques, introduced for a non-DM link with weak nonlinearity. All these methods leads to a zero-mean additive Gaussian noise for the nonlinear distortion influence. Thus, the main challenge is to compute the variance of the AWGN noise $\mathbf{W}$ as a function of the transmit power and channel parameters.

2.4.2 Gaussian-noise based channel models

With special interest in non-DM links, we review the available analytical channel models derived based on first-order perturbation [37, 73] such as the spectral slicing method [37,73–76], and inter-channel-nonlinearity methods including cross-phase modulation [77, 78] and four-wave mixing [79]. Further, we provide a quick survey of the Volterra-series techniques [80, 81] and the orthogonal frequency division multiplexing (OFDM) [7] method. It is worth mentioning that the common assumption shared by all the above models is that the electric field (complex baseband signal) is zero-mean Gaussian. In addition, the nonlinearity is considered weak and the input signal is assumed to have a specific form [37]. Finally, we investigate the performance of a

single-channel fiber-optic link and the growth of the nonlinear (Gaussian) distortion for a non-DM fiber-optic link with linear EDC and nonlinear DBP.

Spectral slicing method [37–39, 73, 74, 76, 79, 82]

The pioneering work in modeling the nonlinear interference as a (white) noise–like interference was performed in 1993 [74]. The analysis was based on the key assumption that the propagated signal has uncorrelated spectral components. The SNR is computed for an AWGN consisting of ASE noise and nonlinear distortion and used to estimate the channel capacity of a DM fiber-optic link. The results revealed that standard SMF fibers provide better performance compared to dispersion-shifted fibers. The same notion explains the superiority of non-DM over DM fiber-optic links. Almost two decades later, the key assumption in this model was fulfilled vey well by removing inline optical DCFs, which were exploited in DM links (with a reasonably large signal bandwidth), resulting in a non-DM link. This model was studied thoroughly in [73, 76, 82] for a non-DM link. Different derivations of the model based on the first-order perturbation method were provided in [37–39] under a weak nonlinearity assumption. A closed-form expression for the variance of the Gaussian nonlinear distortion was provided in [79] with a PM signal consisting of delta-like pulses using the first-order perturbation for both DM and non-DM links. The model is used for the quantitative understanding of system parameters using different types of modulation formats for which the nonlinear signal–noise interaction sets performance limits.

The main constraints in the derivation of the Gaussian noise model based on the first-order perturbation are high accumulated chromatic dispersion, weak nonlinearity, neglecting signal–noise interaction, and the specific signal form [37]. It is shown in [38] that a white jointly-Gaussian wide-sense stationary assumption for the propagating PM signal fulfills the required conditions. This assumption leads to the random processes in each polarization with uncorrelated spectral lines. Since the channel input signal is indeed non-Gaussian, the model becomes valid only after a certain propagation distance. For example, the Gaussian approximation was showed to be accurate enough after $4L_{\mathrm{D}}$ propagation [Paper B].

Inter-channel-nonlinearity model [83, 84]

The inter-channel interference is the dominant impairment after compensating for the intrachannel interference by DBP. The methods based on this assumption are studied in [83,84]. The perturbation of the sampled field, with the only approximation that the fiber nonlinearity is treated to the first order, is computed analytically. In contrast to the models devised by the spectral slicing method, this model can be evaluated for any desired modulation format of the interfering channels. It turns out that the nonlinear distortion caused by inter-channel interference depends on these modulation formats. The model is derived for a single polarization signal.

Look-up table method [6]

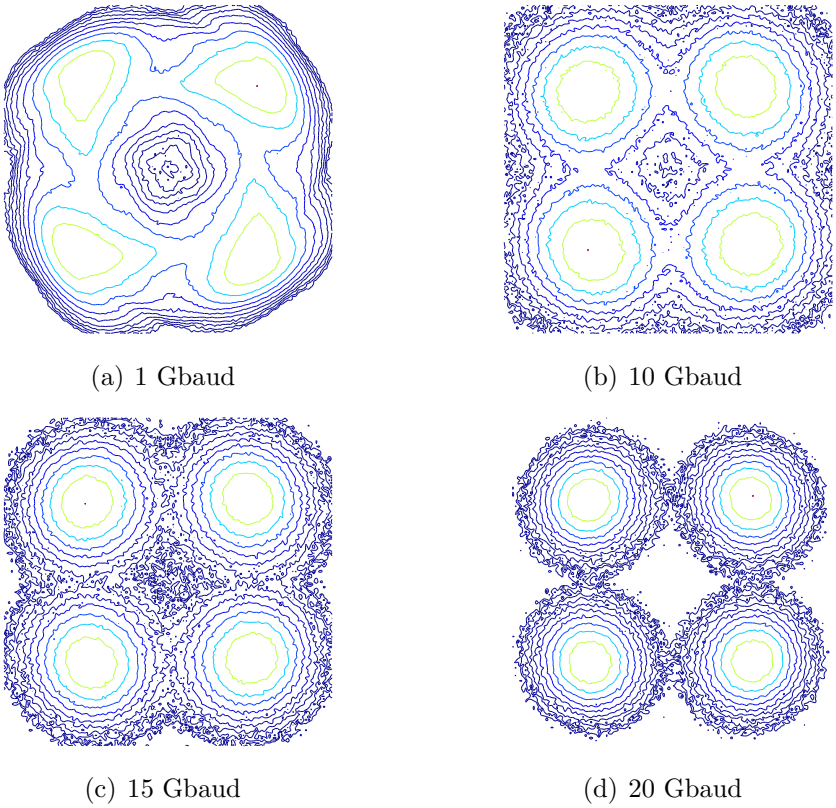
To investigate the channel limit of a nonlinear fiber-optical channel, a model was introduced in [6] based on an empirical method. A database of covariance matrices was computed by a numerical approach, one matrix for each set of channel parameters, which is also dependent on the amplitude of the input symbol. Exhaustive simulations were performed using a continuous-time channel model with enough input realizations. The covariance of the bivariate complex Gaussian is estimated by fitting a Gaussian distribution to the numerical data. The model was derived for a single polarization system. To mitigate the computational complexity, the analysis is performed for a rotational invariant channel with a memoryless discrete-time equivalent model. In a rotational invariant channel, for a certain transmit symbol S_x in one polarization, rotating this symbol by $\Delta\phi$ gives rise to an output symbol with the same signal statistics as before the rotation but now only rotated by the phase $\Delta\phi$ around the origin of the complex plane. More precisely, $f(R_\mathrm{x} \mid S_\mathrm{x}) = f(R_\mathrm{x}e^{j\Delta\phi} \mid S_\mathrm{x}e^{j\Delta\phi})$, where S_x and R_x represent the input and output symbols of a non-DM fiber-optic link for single polarization x (see Fig. 2.3). This property is considered for a nonlinear fiber-optic channel with DBP to consider all points with the same amplitude in a ring constellation statistically equivalent.

2.5 Numerical justification

To investigate the validity of the Gaussian noise model, we provide some numerical simulations. This investigation deals with the distribution as well as the variance of the additive noise $\mathbf{W}$ in a non-DM fiber-optic link.

2.5.1 Gaussian assumption

We evaluate the accuracy of the Gaussian model introduced in (2.12) for two fiber-optical links with PM QPSK and PM 16-QAM signals. In the simulations, a root raised-cosine pulse [40, p. 675] was used with an excess bandwidth of 0.17 and a truncation length of 16 symbols as well as an input sequence consisting of 8192 discrete-time symbols. The input bits to the PM QPSK and PM QAM are generated as independent, uniform random binary digits. The following channel parameters are used for the numerical simulations: the dispersion coefficient $D = 17$ ps/(nm km), the nonlinear coefficient $\gamma = 1.4$ W^{-1}km^{-1}, the optical wavelength $\lambda = 1.55$ μm, the attenuation coefficient $\alpha = 0.2$ dB/km, and the ASE noise figure $F_\mathrm{n} = 5$ dB. For the numerical simulation, we use a link consisting of 90 spans of length 80 km at 32 Gbaud and 30 spans of length 120 km at 42.7 Gbaud. The pulse shaping excess bandwidth and the symbol rates for the numerical simulations are chosen to obtain signal bandwidths of 37.5 and 50 GHz. It is worth mentioning that, as shown in [85],



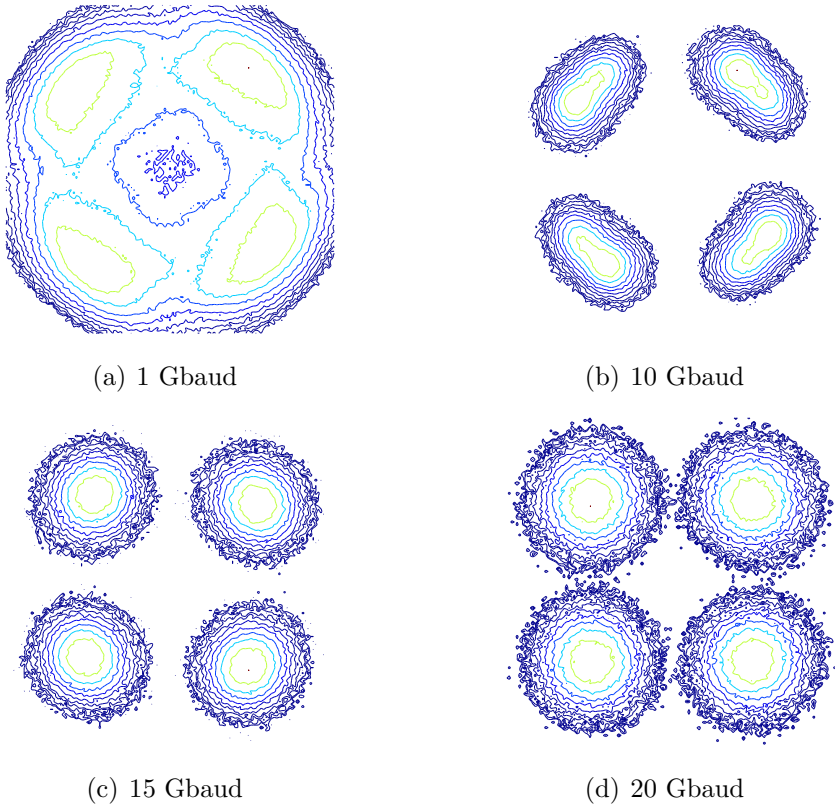
(a) 1 Gbaud (b) 10 Gbaud

(c) 15 Gbaud (d) 20 Gbaud

Figure 2.11: The estimated pdf of the received signal R_{x} for a non-DM fiber-optic link with EDC, 25 spans of length 80 km, $D = 17$ ps/(nm·km), and four symbol rates.

intrachannel effects are dominant for very long optical links, which is why we have simulated optical links with a large number of spans to provide a realistic numerical justification of the first-order approximation derived for a single-channel scenario.

The estimated pdf of the received signal R_{x} for a non-DM fiber-optic link is shown in Figs. 2.11 and 2.12 for EDC and DBP, respectively. The results are illustrated for four symbol rates from 1 to 20 Gbaud. As seen, the higher the symbol rate, the higher accumulated chromatic dispersion, and consequently the better Gaussian approximation is obtained. For the simulated systems, the circular symmetric Gaussian assumption becomes quite accurate at symbol rates above 15 Gbaud and 20 Gbaud for EDC and DBP, respectively. The numerical results show that the discrete-time additive Gaussian noise channel model described by (2.12) can be used as an accurate model for a non-DM fiber-optic link with both EDC and DBP. One may apply nonlinear equalization based on DBP to mitigate the effect of the signal–signal interference. Since the system performance is significantly improved in this case, the Gaussian assumption is not sufficiently accurate in the tails of the noise probability density function [72]. Thus, in contrast to linear EDC, the analytical result derived based on the Gaussian assumption is solely used as a lower bound for the system performance. We have



(a) 1 Gbaud (b) 10 Gbaud

(c) 15 Gbaud (d) 20 Gbaud

Figure 2.12: The estimated pdf of the received signal R_{x} for a non-DM fiber-optic link with DBP, 25 spans of length 80 km, $D = 17$ ps/(nm·km), and four symbol rates.

illustrated this property of the propagated signal $U_{\mathrm{x}}(t, z)$ in a non-DM fiber-optic link in Fig. 2.13. This figure has been plotted for a linear channel without any nonlinear effects. Even without nonlinearity, the tails of the signal distribution at the output of a single span with a length of 80 km and dispersion coefficient D of 17 ps/(nm·km) deviate from a Gaussian distribution with the same variance. However, this deviation vanishes by increasing the number of spans to $N = 10$.

2.5.2 Nonlinear Gaussian distortion characterization

In general for a WDM system, the nonlinear (Gaussian) distortion consists of inter- and intrachannel nonlinear noise–like distortion. Here, a PM non-DM fiber-optic link is modeled using the analytical results introduced in [73] and [Paper B] for a single-channel case in the absence of inter-channel effect. According to (2.12), the AWGN noise vector $\mathbf{W}$ consists of zero-mean noises W_{x} and W_{y} in polarization x and y, respectively. For the same transmit power in the two polarizations, the variance of W_{x} and W_{y} can be computed as $\sigma_W^2 = \mathbb{E}\{\mathbf{W}^\dagger \mathbf{W}\}/(2T) = N\sigma^2 + \sigma_{\mathrm{NL}}^2$. The variance of nonlinear (Gaussian) distortion σ_{NL}^2 is computed as the amount of power depleted from

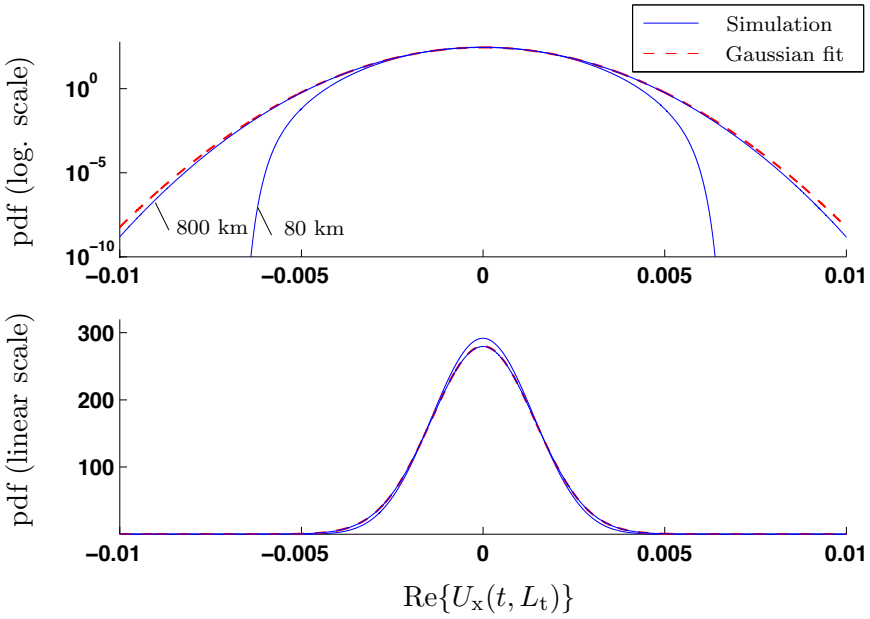


Figure 2.13: The distribution of the real part of the electric field $U_{\mathrm{x}}(t, L_{\mathrm{t}})$ in x polarization for a linear (without nonlinear effect) fiber-optic link with a span length of $L = 80$ km, $D = 17$ ps/(nm $\cdot$ km), and a QPSK signal.

the launched signal due to the channel nonlinear effect. In other words, since the power loss in the fiber-optic link is compensated by inline amplifiers, the attenuation caused by the complex scaling constant with $|\zeta|^2 < 1$ is coming from the fact that the nonlinearity converts a part of the transmitted power to the nonlinear distortion. This can also be motivated according to energy conservation law, as the advent of this nonlinear distortion makes a part of the signal power disappear. Thus, $\sigma^2_{\mathrm{NL}} = (1 - |\zeta|^2)P$, where the amplitude square of the complex constant ζ is given by [72]

$$|\zeta|^2 \approx 1 - 3N^{1+\epsilon}\gamma^2\alpha^{-2}\tanh(\tfrac{\alpha}{4}L_{\mathrm{D}})P^2, \tag{2.13}$$

in which $L_{\mathrm{D}} = T^2/(|\beta_2|)$. Here, the linear growth of the nonlinear distortion with N, introduced in [Paper B], has been changed to $N^{1+\epsilon}$ [72], where

$$\epsilon = \tfrac{3}{10}\log\left(1 + \frac{6}{\alpha L \ \mathrm{asinh}\left(\frac{\pi^2}{2\alpha L_{\mathrm{D}}}\right)}\right).$$

As discussed in [73], this takes into account the partially correlated rather than entirely uncorrelated accumulation of nonlinear noise–like distortions from different spans. The variance σ^2_W is used to compute the a posteriori probabilities of the coded symbols required for the CM decoder (see Section 3). According to [73, eq. (7), (13), and (23)],

Table 2.1: The variance of the additive Gaussian noise σ_W^2 introduced in (2.12) in each polarization with EDC and DBP consisting of the linear (ASE) and nonlinear noise–like distortion [72].

EDC	$a_{\mathrm{NL}}P^3 + N\sigma^2$
DBP	$a_{\mathrm{NL}}(N-1)\sigma^2P^2 + N\sigma^2$
Scale factor	$a_{\mathrm{NL}} = 3N^{1+\epsilon}\gamma^2\alpha^{-2}\tanh(\frac{\alpha}{4}L_{\mathrm{D}})$

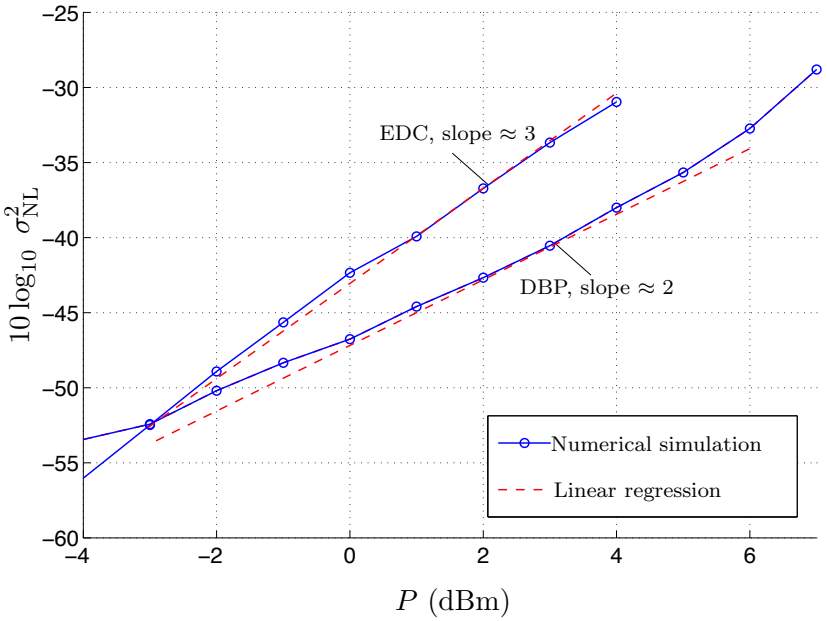


Figure 2.14: The variance of the nonlinear (Gaussian) distortion versus the average transmit power per polarization P for two non-DM fiber-optic links with EDC and DBP. The quadratic and cubic growth are observed for DBP and EDC, respectively.

the variance of the nonlinear (Gaussian) distortion is computed as

$$\sigma_{\mathrm{NL}}^2 \approx \tfrac{32}{27\pi}N^{1+\epsilon}\gamma^2\alpha^{-1}L_{\mathrm{D}}^{-1}\mathrm{asinh}\left(\frac{\pi^2}{2\alpha L_{\mathrm{D}}}\right)P^2. \tag{2.14}$$

Finally, the total SNR, $|\zeta|^2P/(N\sigma^2+\sigma_{\mathrm{NL}}^2)$, and the inverse of the normalized nonlinear (Gaussian) distortion variance $\sigma_{\mathrm{NL}}^2/P = 1-|\zeta|^2$ for the system with 38 Gbaud are plotted versus the transmitted power P in Fig. 2.15. This figure illustrates the cubic growth of the nonlinear (Gaussian) distortion variance with the input power for linear equalization (EDC) and the quadratic growth for nonlinear equalization (DBP). The noise variance σ_W^2 for a non-DM fiber-optic link with linear EDC and nonlinear DBP is given in Table 2.1. As seen, the signal–signal nonlinear interference caused by

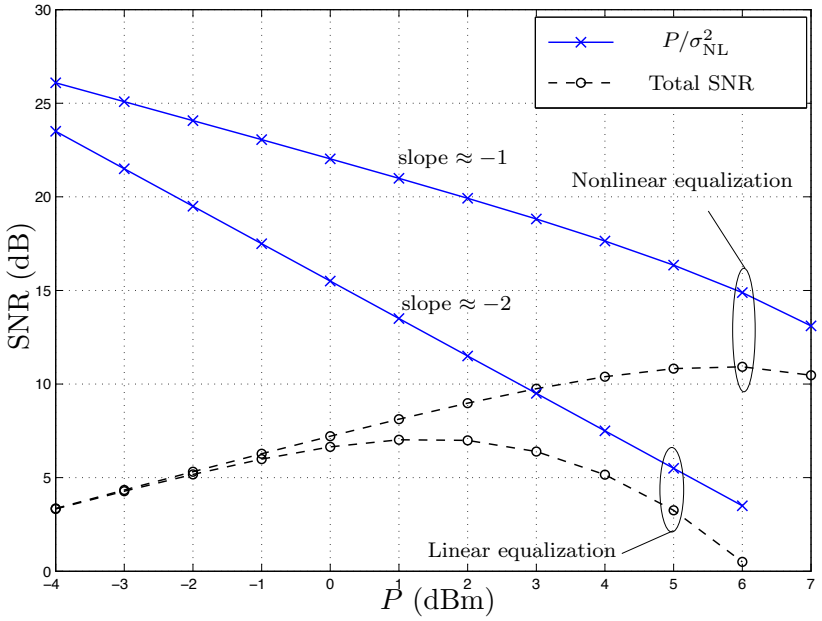


Figure 2.15: The total SNR and the normalized inverse variance of the nonlinear (Gaussian) distortion for a system with a symbol rate of 38 Gbaud and a dispersion coefficient of $D = 17$ ps/(nm km).

the Kerr effect can be removed to mitigate the cubic growth of the nonlinear (Gaussian) distortion variance of a system with EDC to a quadratic growth with DBP. This behavior is intuitively predictable. In fact, within the first-order perturbation assumption [4, 73, 77, 79], the nonlinear distortion comes from the integration of electric field terms $\mathbf{U}\mathbf{U}^\dagger\mathbf{U}$ whose variance scales as P^3. According to (2.1), in presence of ASE noise, the nonlinear distortion of the first-order perturbation comes from terms $(\mathbf{U}+\mathbf{n})(\mathbf{U}+\mathbf{n})^\dagger(\mathbf{U}+\mathbf{n})$. When the solely signal-dependent nonlinear distortion term $\mathbf{U}\mathbf{U}^\dagger\mathbf{U}$ is perfectly compensated by DBP, then the variance of the leftover nonlinear distortion scales as P^2.

The SER versus transmitted power per polarization P of two fiber-optical links with EDC for PM QPSK are shown in Fig. 2.16. The SERs of these two systems with EDC have also been evaluated analytically using (2.13) and [73, eqs. (7), (13), and (23)]. As seen, the analytical models show a good agreement with the numerical simulations for low and moderate transmit powers, almost up to the optimal power, the so-called nonlinear threshold. For high transmit powers, the nonlinear effect is strong such that the chromatic dispersion cannot return the non-Gaussian distribution resulting from the nonlinear effect to Gaussian, so the Gaussian assumption is no longer valid and the analytical models overestimate the actual SERs. Moreover, we investigate the growth of nonlinear distortion with transmit power for two different system configurations: (i) 70 spans of 120 km in length with a QPSK signal and DBP at 32 Gbaud (ii) 30 spans of 120 km in length at 42.7 Gbaud with EDC. As seen in Fig. 2.14, the nonlinear

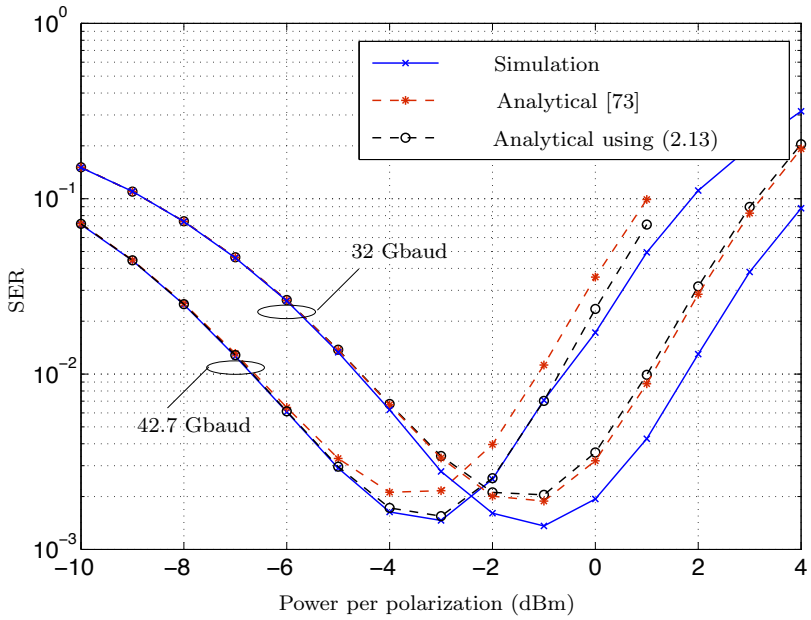


Figure 2.16: The SERs of two fiber-optic links with EDC and PM QPSK versus transmitted power per polarization P, consisting of 90 spans of length 80 km at 42.7 Gbaud and 30 spans of length 120 km at 32 Gbaud. The analytical results using the introduced model in (2.13) as well as the model in [73, eq. (7), (13), and (23)].

distortion shows a quadratic and cubic growth with the transmit power for a non-DM link with EDC an DBP, respectively.

2.6 Alternative channel models

In this section, we review a channel model introduced to predict the deterministic effects of a nonlinear fiber-optic channel. Moreover, we describe a new transmission scheme built on advanced mathematical tools.

2.6.1 Volterra-series method

A two-dimensional (time and wavelength) discrete-time model of deterministic impairments based on the third-order Volterra-series transfer function is introduced in [81]. The authors managed to overcome the well-known triple integral problem associated with the Volterra-series transfer function method and reduce it to a single integral. In contrast to spectral slicing methods, this method takes into account optical filtering, photo detection, and time-varying system parameters. The transmit pulse $p(t)$ is assumed to be Gaussian. Moreover, no random effect such as ASE noise is considered for this model. The Volterra-series method introduces a polynomial expansion that represents the input–output relationship of a fiber-optic link in the frequency domain

output of the fiber computed by taking the Fourier transform of $U_{\mathrm{x}}(t, L_{\mathrm{t}})$. Since the output field can be approximated by the Volterra-series, this model can be used as an analytical tool to compensate the deterministic nonlinear effects such as intrachannel signal–signal interference [86].

2.6.2 Nonlinear Fourier transform method

This model [87,88] can be seen as a new transmission scheme, in which the information is encoded in the nonlinear spectrum of the signal. This approach generalizes the role of the Fourier transform in linear channels in converting a linear convolutional channel into a number of parallel scalar channels, to a nonlinear dispersive channel described by nonlinear Schrödinger equation. This technique takes into account both dispersion and nonlinearity without the need for dispersion or nonlinearity compensation methods. In the absence of ASE noise, the signal gets decorrelated by properly choosing the basis functions in much the same way that the Fourier transform does for linear systems. This approach may be used for the derivation of signal statistics of fiber-optic channels in spectrum domain, which has been studied for distributed amplification.

Chapter 3

Channel-Aware Coded Modulation

The coding theory for non-Gaussian channels, such as nonlinear fiber-optical channels, is not well investigated. As discussed in Chapter 2, the lack of an accurate statistical model for a general fiber-optical channel might be the main reason for the absence of analytical results in the design of forward error correcting (FEC) codes for these channels. Classic FEC schemes such as Reed–Solomon (RS) and Bose–Chaudhuri–Hocquenghem (BCH) codes have been widely used in optical communications. The most common design criterion for classic codes is the maximization of the minimum Hamming distance of the codewords, while the capacity-achieving codes such as turbo [89], LDPC [90], and polar codes [91] were originally proposed by inspiration from information theory. The capacity-achieving codes were optimized for binary-input AWGN, binary symmetric, and binary erasure channels. Recently, many new approaches based on capacity-achieving codes [92] have been proposed for optical data transmission systems. These schemes intend to increase the performance of the system in the moderate SNR regime, where the classic RS codes show poor performance [6, 93, 94]. Moreover, significant efforts have been devoted to evaluating the performance of known codes for fiber-optic channels [24, 95–97].

In this thesis, joint coding and (multilevel) modulation, so-called coded modulation (CM), schemes are considered for optical communications. The CM techniques are known to be superior to conventional approaches with independent FEC and modulation, in the sense of requiring less SNR for the same spectral efficiency [98]. More precisely, compared with well-established FEC schemes for optical communications, CM schemes provide a better trade-off between the DSP complexity and the transparent reach of the system for a given spectral efficiency [24, 25].

Traditional approaches in the design of CM systems were focused on the minimum Euclidean distance (MED) and asymptotic gains [21, 99], while use of techniques from information theory have changed these design criteria [100]. It was proven in [101] that multilevel CM (MLCM) together with multistage decoding (MSD) can approach the Shannon capacity for the AWGN channel [1]. In fact, for AWGN channels, not

only MED of the generated codewords, consisting of a sequence of symbols, plays an important role in the performance of a system with multilevel modulation format, but neighboring coefficients (the number of codewords at a certain Euclidean distance from a specific codeword, averaged over all the codewords) [102] are relevant, especially in a low SNR regime [103–105]. In addition, the stochastic channel impairments can be taken into account in the design of a CM scheme, leading to a channel-aware design method. Channel-aware CM is a joint coding and modulation scheme that considers the impairments of a channel.

In this chapter, we first briefly review the fiber-optic system constraints as well as quality parameters for evaluating the system performance. Then we introduce an information-theoretic framework to analyze the CM techniques for AWGN and fiber-optic channels. The framework is used to provide a design guideline for a CM scheme with a high spectral efficiency and a comprehensive classification of CM techniques in the literature. Furthermore, four-dimensional (4D) CM schemes constructed for PM signals are compared with two-dimensional (2D) schemes, using both binary and nonbinary component codes. In fact, a CM scheme can exploit the four available dimensions of a fiber-optic link, i.e., two polarizations each consisting of an in-phase and a quadrature dimension, with more flexibility than conventional schemes. To address the quest for adaptive, sometimes called elastic, optical transceivers, we investigate different rate-adaptive CM schemes. Finally, the joint design of 4D CM schemes with probabilistic shaping is introduced to even further increase the spectral efficiency of fiber-optic data transmission with a reasonable complexity[5].

3.1 System constraints

The transparent reach, i.e., the transmission distance of a fiber-optic link with no inline electrical signal regenerators, is intimately related to the desired spectral efficiency, as well as to the digital signal processing (DSP) complexity (depicted in Fig. 3.1). For example, to obtain a reliable data transmission scheme for a given spectral efficiency, there is a trade-off between the transparent reach and the DSP complexity. The larger the transparent reach is, the higher DSP complexity is needed, provided that the desired spectral efficiency is achievable for this transparent reach.

3.1.1 Channel capacity

As we discussed in Chapter 1, the main motivation for the research on channel coding was the mathematical theory of Shannon on channel capacity. Despite considerable research efforts, the capacity of fiber-optic channel is still unknown. In fact, the nonlinear nature of optical fibers makes the capacity much more difficult to determine compared

[5]The major part of this chapter has been prepared based on the submission [106].

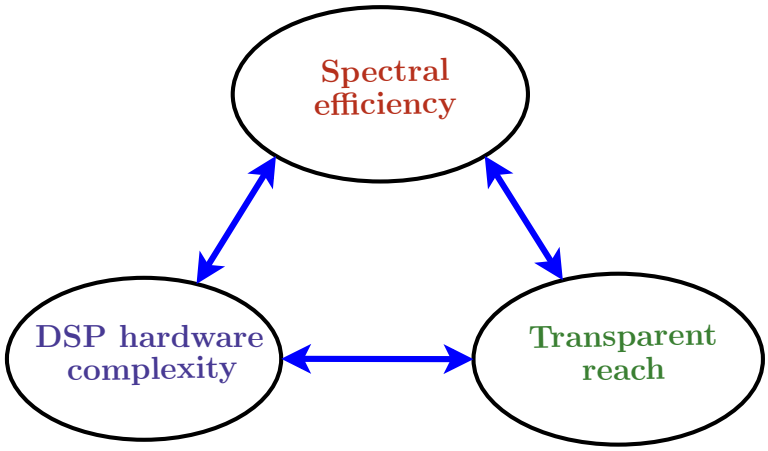


Figure 3.1: Three main factors in the design of a CM scheme for fiber-optic channel.

to the analytical approach introduced by Shannon for linear AWGN channels. The lack of an accurate discrete-time model in the presence of channel nonlinear behavior may be the main bottleneck for capacity studies. More precisely, well-developed signal space design methods for linear channels are not applicable to nonlinear fiber-optic channels. However, the common use of the available signal design techniques, due to their implementation simplicity, has given rise to a widespread belief that the nonlinearity limits capacity of fiber-optic channels. As showed in [9, 107, 108], the channel capacity is a nondecreasing function of SNR for many of the channel realizations of fiber-optic links. Moreover, it has been suggested [109, 110] that the nonlinear effect may help to manipulate the inline added AWGN by techniques such as noise squeezing to change the manifestation of the channel nonlinear effect as a constructive rather than destructive influence [6, 77, 111–113].

In this chapter, we use the nonlinear and Gaussian channel models introduced in Sections 2.3 and 2.4 for DM and non-DM links, respectively, to design CM schemes. Due to a lack of exact analytical capacity result for these channels, we use the AWGN capacity introduced in 3.1.3 as a reference to compare different systems. We emphasize here that although the AWGN capacity is popular as a benchmark, it does not represent the capacity of the nonlinear fiber-optic channel.

3.1.2 Operating transmit power

In the Gaussian noise model for a non-DM optical link with EDC, the variance of the (nonlinear distortion) noise grows as the cube of the transmitted power (see Section 2.5.2). Therefore, the system performance is, as shown in Figs. 2.16 and 3.2, eventually degraded at high transmitted power levels. This nonlinear behavior distinguishes these channels from classical AWGN channels. Clearly, there is an optimum power (shown by red stars in Fig. 3.2), which yields the minimum uncoded SER or the maximum

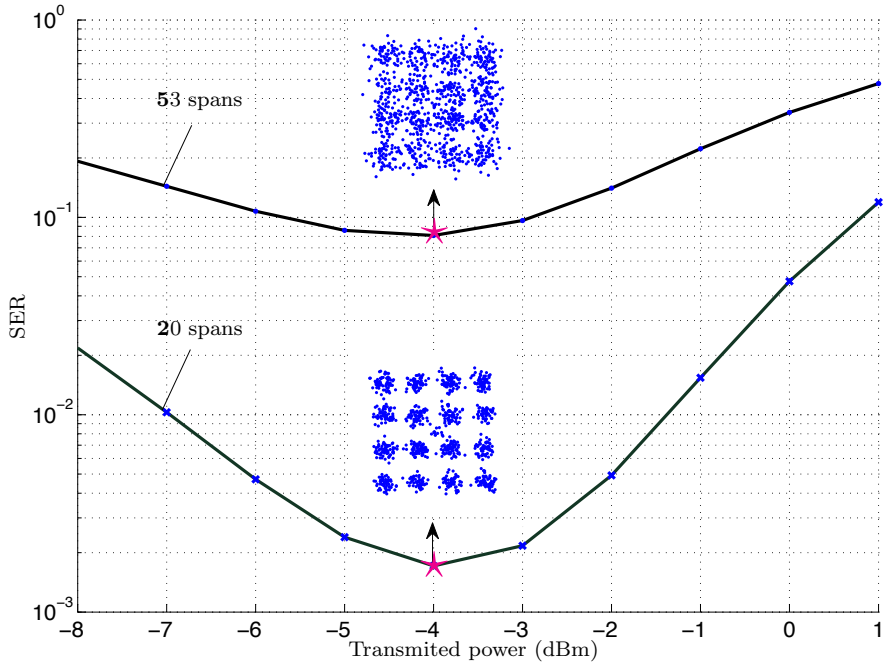


Figure 3.2: The SERs of a nonlinear fiber-optic link with 20 and 53 spans together with the scatter plots of the received signals at the minimum SER, marked by red stars.

SNR after the EDC. This optimum signal power is almost independent of the transparent reach [111], and the system is assumed to operate at the optimal transmit power throughout this chapter. Since a well-designed CM scheme allows reliable data transmission with a higher uncoded SER, an improved system performance by exploiting a CM scheme can be used to increase the transparent reach as shown in Fig. 3.2.

3.1.3 Quality parameters

We will use three quality parameters to evaluate the performance of optical data transmission systems with HDD and SDD, including FEC threshold, net coding gain (NCG), and gap to the AWGN channel capacity. These will be discussed separately below.

FEC threshold

Traditionally, due to the use of independent FEC and modulation together with hard-decision demodulation, the maximum BER of a hard-decision demodulator (the input BER of the FEC decoder), the so-called FEC threshold, for obtaining an information BER of 10^{-15} at the output of the FEC decoder was widely used as a metric for these channels [2, Ch. 5]. In fact, the main goal of system designers was to meet the desired FEC threshold for an uncoded system. The advent of CM schemes in fiber-optic communications with soft-decision decoding enables a different evaluation of these systems. Since the CM decoder uses soft input, the uncoded BER is not relevant in

these systems. Instead, the next two metrics are used to benchmark different systems with the same spectral efficiency η.

Net coding gain

The reduction in the SNR requirement resulting from adding coding at the same information bit rate and the same (low) information BER for both coded and uncoded systems is called the net coding gain (NCG). The code rate[6] of the coded system is $R = \eta_{\text{uncod}}/\eta$, where η_{uncod} and η are the spectral efficiencies of the uncoded and coded systems, respectively. The NCG is precisely defined as the gross coding gain in bit SNR scaled by the code rate of the coded system to compare the two coded and uncoded systems at the same information bit rate [14]. Stating it based on symbol SNR, the NCG of a system at a certain information BER can be expressed as $\text{NCG} = R^2\gamma_{\text{uncod}}/\gamma$, where γ_{uncod} and γ are the required (symbol) SNRs to meet the desired BER for the given uncoded and coded systems, respectively.

Gap to the AWGN channel capacity

For a system with a rate R, there is a minimum SNR γ to obtain a certain BER, which is usually computed by numerical simulations. The gap $\Delta\gamma$ between γ and the minimum SNR obtained using the Shannon formula for an AWGN channel with spectral efficiency η is a useful measure [114] to compare different CM schemes. This gap can be expressed as $\Delta\gamma = \gamma/(2^\eta - 1)$ and it can be related to the NCG as $\Delta\gamma = R^2\gamma_{\text{uncod}}/((2^\eta - 1) \times \text{NCG})$.

3.2 Design framework

In this section, we exploit the concept of equivalent parallel binary subchannels [101] to establish a design framework for CM schemes. This framework is built on two types of decoders, namely multistage and parallel independent decoders, to provide a channel-aware design method. An accurate channel model (such as those described in the previous chapter) is necessary to exploit this design framework. More precisely, this information-theoretic tool requires the signal statistics of the received signal $\mathbf{R}$ from the channel.

Considering the bit-to-symbol mapper shown in Fig. 3.3, the equivalent binary subchannels approach introduced in [101] can be applied to represent the mutual information (MI) between the channel input and the received signal after EDC as $I = \sum_{i=1}^{m} I_i$, where $I_i = I(V_i; \mathbf{R}|V_1, \ldots, V_{i-1})$ is the conditional MI of subchannel i, provided that the transmitted bits of the subchannels $1, \ldots, i-1$ are given. Clearly, the channel

[6] The system redundancy overhead is defined as $\text{OH} = 1/R - 1$.

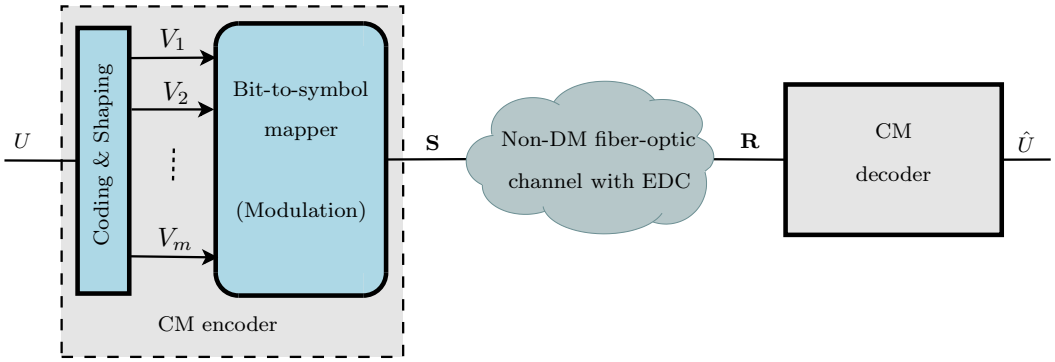
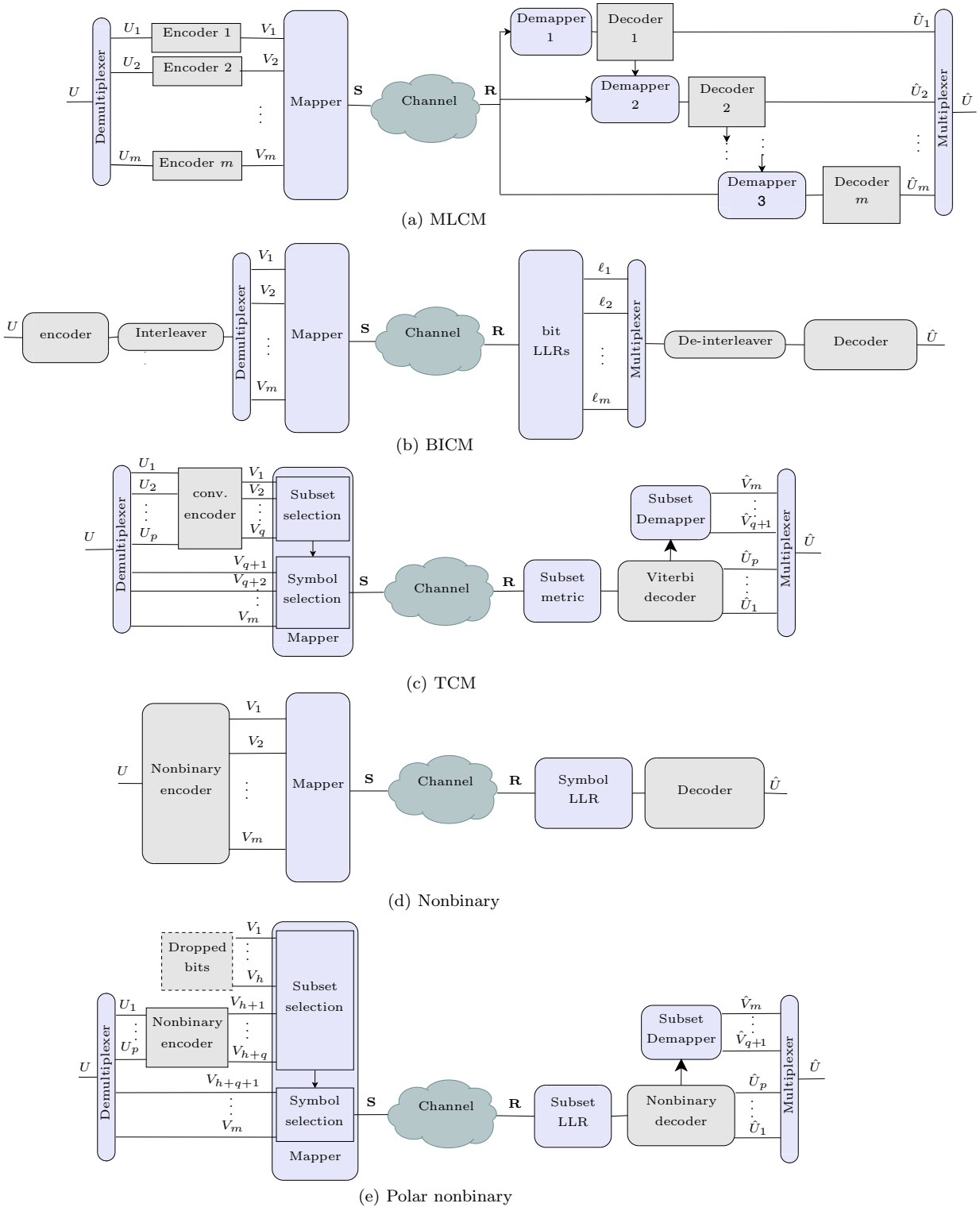


Figure 3.3: A fiber link including a CM encoder at the transmitter, the fiber-optic link with EDC illustrated in Fig. 2.3, and the CM decoder at the receiver.

with input $\mathbf{S}$ and output $\mathbf{R}$ can be modeled as m parallel subchannels with the inputs $V_i,\ i = 1, \ldots, m$ and the output $\mathbf{R}$. An alternative parallel subchannel modeling approach is based on decoding the individual subchannels independently [101, 115], which yields a sum rate of $\hat{I} = \sum_{i=1}^{m} \hat{I}_i$, in which $\hat{I}_i = I(V_i; \mathbf{R})$. It can be shown that $I(V_i; \mathbf{R}) \leqslant I(V_i; \mathbf{R}|V_1, \ldots, V_{i-1})$ [101], implying that $\hat{I} < I$. The gap between $\hat{I}$ and I strongly depends on the selected labeling of the constellation symbols. Caire et al. [115] showed that this gap is surprisingly small with Gray labeling. However, the multistage decoding technique is significantly superior to the parallel independent decoding for a finite length code [101]. The terms (binary) "subchannels" and "layers" are used interchangeably in this chapter. We explain the three main categories of CM schemes, exploiting the equivalent subchannels for AWGN channels, as well as two CM schemes that are constructed from nonbinary component codes separately below. They are all illustrated in Fig. 3.4.

3.2.1 Multilevel coded modulation (MLCM)

For an arbitrary modulation, the binary subchannels have in general different conditional MI I_i. Hence, to approach the channel MI I, an unequal error-protecting technique, as depicted in Fig. 3.4 (a), is applied over the m binary subchannels. To this end, MLCM was exploited consisting of m binary turbo [101] or nonbinary LDPC [116] codes, originally introduced with classic block codes [21,117], each adapted to the conditional MI (I_i for channel i) of the corresponding subchannels. MLCM has been shown to be a capacity-achieving scheme theoretically [104] and through simulations [101,116] for AWGN. An interesting feature of MLCM is the possibility of exploiting a multistage decoder (MSD). As shown in Fig. 3.4 (a), the decoder of the first layer (subchannel) can decode the received bits independently of the other layers, then the second decoder uses the output from the first decoder to decode the bits received in the second layer. This sequential decoding is followed for the rest of the layers. The MSD has lower



(a) MLCM

(b) BICM

(c) TCM

(d) Nonbinary

(e) Polar nonbinary

Figure 3.4: The block diagram of CM schemes: (a) MLCM (b) BICM (c) TCM (d) Nonbinary (e) Polar nonbinary.

complexity than the maximum likelihood detector. An MLCM scheme was tailored in [Paper C] for a memoryless nonlinear fiber-optic channel with RS component codes. In this paper, an unequal error protection scheme in the phase and radial direction of a 16-point ring constellation is exploited to optimize the block error rate of the system. For non-DM fiber-optic channels, two simplified MLCM schemes were introduced in [25, 28] with staircase codes and LDPC codes, respectively. The subchannels are categorized in two groups in [25] and three groups in [28], to reduce the number of component codes.

3.2.2 Bit-interleaved coded modulation (BICM)

Zehavi [22] introduced BICM as shown in Fig. 3.4 (b) simply by adding an interleaver between the encoder and the mapper to distribute the coded bits among different binary subchannels uniformly and exploit the diversity in the subchannels. In the BICM scheme, the subchannels are assumed to be independent and a simplified model using m independent decoders of the binary subchannels is used [101] with the MI $I(V_i; \mathbf{R})$ for subchannel $i = 1, \dots m$, in which each subchannel has no information from the input bits of the other subchannels. Usually, the binary decoder uses the log-likelihood ratios (LLR)[7] of the subchannels after de-interleaving to decode the received bits.

For channels such as wireless fast fading channels, the channel is unknown at the transmitter, and thus, the MIs of the subchannels are also unknown. BICM was originally proposed for fast fading channels to exploit the diversity in binary subchannels [115]. In contrast to wireless channels, the optical channel[8] can be compensated (tracked) by adaptive filtering techniques without noise enhancement [34]. Therefore, one can assume that the channel is almost static and known at the transmitter provided that the CSI is available. Consequently, the TCM and MLCM techniques are more effective than BICM with short-length component codes for fiber-optic channels. Despite this limitation, BICM has been suggested for fiber-optic communications. For example, a comprehensive study of BICM for fiber-optic communications has been performed in [16] with different modulation formats. The extension to multidimensional constellation was studied in [26, 118].

3.2.3 Trellis-coded modulation (TCM)

Ungerboeck [20] introduced a new type of binary labeling based on the set partitioning technique. The subchannels resulting from this labeling have ascending MI values. The early subchannels (with smaller indices) have lower MI values than the subchannels with indices close to m. The original version of TCM, shown in Fig. 3.4 (c), splits the

[7] The LLR of bit v is defined by $\Pr(v = 1|\mathbf{R})/\Pr(v = 0|\mathbf{R})$.
[8] With a negligible polarization-dependent loss.

information bits into two groups of subchannels, where the group with smaller indices, the so-called "subset selection," is protected by a convolutional code, while the second group, denoted as "symbol selection," remains uncoded. Although this scheme can be decoded by MSD, Ungerboeck proposed a maximum likelihood decoder. The Viterbi decoder uses the subset metrics to decode the first group. The second group is decoded by a simple demapper within the decoded subset.

Later, turbo TCM was proposed [119], by replacing the convolutional code with a turbo code to decrease the gap from the Shannon limit for AWGN channels. Furthermore, multidimensional TCM was proposed in [120, 121], which allows a higher spectral efficiency for a given signal constellation than one-dimensional (1D) or 2D TCM methods. In fiber-optic systems, TCM was proposed in [122] with an 8-point cubic polarization shift keying constellation. The simplest 4- and 16-state TCM schemes were applied to 8-point phase shift keying (PSK) and differential PSK in [23]. Finally, the concatenation of 2D TCM with two different outer codes, RS and BCH codes, was studied in [24], which gives NCGs of 8.4 and 9.7 dB, respectively, at a BER of 10^{-13} for the AWGN channel.

3.2.4 Nonbinary coded modulation

The codewords of a nonbinary code are sequences of 2^q-ary symbols, each representing q bits. The code is constructed over a Galois field (GF) of the order 2^q, denoted by GF(2^q). Binary codes can be considered as the simplest case of these codes, defined over GF(2) with two symbols 0 and 1. The binary subchannels can be encoded and decoded jointly using nonbinary codes, at the cost of increased complexity. As shown in Fig. 3.4 (d), the demapper computes symbol LLRs for each soft received symbol, retaining the MI between the subchannels compared to the independent bit LLR calculation in BICM. Different types of nonbinary codes such as classic nonbinary codes, e.g., RS codes with a hard-decision decoding, or modern nonbinary LDPC and turbo TCM codes with a soft-decision decoding, can be used to construct the nonbinary CM schemes. Moderate-length nonbinary LDPC codes have been widely proposed for fiber-optic communications [123], to approach the Shannon limit in AWGN channels. As discussed, the nonbinary scheme can be used with both 2D [123] and 4D [27, 28] constellations.

3.2.5 Polar nonbinary coded modulation

Although many techniques have been suggested to mitigate the computational complexity of nonbinary codes, the decoding complexity in the order of $O(q2^q)$, for a regular nonbinary LDPC code designed over GF(2^q), makes this scheme unrealistic for large constellations [124]. To overcome this problem, a mapper, inspired by the polar coding technique [91], was devised [28] [Paper E] to categorize the binary subchannels into

three groups, namely 'bad,' 'intermediate,' and 'good' subchannels. The 'bad' and 'good' subchannels have MIs near 0 and 1, respectively, while the MIs of 'intermediate' subchannels are between 0 and 1. Then, error protection using nonbinary LDPC coding is performed solely over the 'intermediate' subchannels. As shown in Fig. 3.4 (e), the 'good' subchannels are left uncoded, whereas no information is transmitted on the 'bad' subchannels denoted by dropped bits, which are fixed to zero and known to the receiver. Since the nonbinary encoder performs on 'intermediate' subchannels independently of the constellation size [28], the GF can have a lower order with this design than with the regular nonbinary scheme above, and consequently a CM scheme with a lower complexity is obtained. We notice that the classification into 'intermediate' and 'good' subchannels was used by Ungerboeck in the design of TCM. Here, it is extended by introducing 'bad' subchannels or dropped bits. In fact, increasing the number of dimensions in the bit-to-symbol mapper shown in Fig. 3.3, provides the required freedom in the design of binary labeling to be able to drop some ('bad') subchannels and leave some ('good') uncoded, which leads to a system with lower complexity. The decoding of the 'good' subchannels (with MI near 1) is performed after the decoding of the 'intermediate' bits. This is because the MIs of these subchannels are conditioned on the input bits of 'intermediate' subchannels.

3.3 Coded modulation for dispersion-managed links

The design framework can also be used to design a CM scheme for the simplified DM fiber-optical channels. To this end, one may exploit the statistics of complex received signal for this channel introduced in (2.11). Moreover, it has been shown [57] that a ring constellation consisting of many M-PSK constellations with different amplitudes, sometimes called ring constellation, shows better performance than a square QAM constellation with the same number of symbols, provided that MSD is used [125]. In general, a MAP detector for a ring constellation is not a practical scheme in DM fiber channels. This is due to the non-symmetric decision regions of constellation symbols. Instead of an exact MAP detector, a two-stage detector is proposed in [57], which has a performance close to a MAP detector but with much lower complexity (see Section 2.3.3). In this detector, first the amplitude of the received signal is detected and then in the second stage. The phase of the received signal is determined based on the detected amplitude in the first step.

Here, we consider a ring constellation with 2^{m_r} rings, each consisting of 2^{m_p} equally spaced phase symbols (as shown in Fig. 2.8(b) for a 16-point ring constellation). Moreover, we assume the first m_r bits $V_1, V_2, \ldots, V_{m_r}$ determine the ring amplitude from $\{R_1, R_2, \ldots, R_{2^{m_r}}\}$ and the last m_p bits of the binary labeling, with $m_r + m_p$ bits, choose the phase of the symbol inside the selected ring. To determine the component

code rates of the subchannels in the MLCM scheme, one may use the introduced design framework to allocate the code rates for different layers (subchannels). To this end, the MIs are split into two groups. The first m_r subchannels are represented by the conditional MIs between $V_i, 0 < i \leq m_r$ and the amplitude of the received signal $|R_{\mathrm{x}}|$. Standard set partitioning techniques for a 2^{m_r}-ary pulse amplitude modulation scheme can be used to decrease the decoding complexity (as discussed in [Paper C]). The second m_p subchannels are described by the conditional MIs between $V_i, m_r < i \leq (m_r + m_p)$ and the phase of the received signal $\angle R_{\mathrm{x}}$. Here, one may exploit set partitioning techniques for a 2^{m_p}-ary PSK signal to obtain a low-complexity scheme. Once these MIs are computed, the rate of capacity-achieving codes can be selected for different layers corresponding to computed MI.

A set partitioning scheme with a low decoding complexity is introduced in [Paper C] for an MLCM system in a DM fiber-optic link with no dispersion. The design and performance analysis of this system is performed based on the signal statistics in (2.11). The MLCM scheme is constructed using RS codes with HDD and the 16-point ring constellation of Fig. 2.8(b). Since the exploited RS codes are not capacity-achieving, we use the Lagrange-multiplier approach, which is an optimum method in the high SNR regime [126]. Since the traditional criterion for CM design, which is the maximization of MED, is grossly suboptimal for such a non-Gaussian channel, this design approach significantly improves the performance of the system in comparison to traditional CM methods.

3.3.1 Concatenated codes

Concatenation of inner and outer codes, shown in Fig. 3.5, was introduced as a solution to the problem of finding a coded scheme that has both a rapidly decreasing BER at moderate SNR , known as the waterfall region, and the possibility of reaching extremely low BERs without any error floor [12, Ch. 5]. It was suggested in [14] to exploit the waterfall region of a modern code, e.g., LDPC code, as the inner code to obtain BERs around 10^{-3} at the output of the inner code decoder. Then this BER is suppressed using an outer code constructed based on classic codes with HDD such as RS or BCH codes to meet the desired BER for optical communications, i.e., 10^{-15}. The CM schemes discussed here are suitable as the inner code in this structure.

Figure 3.5: Concatenation of an outer (RS or BCH) and inner (CM scheme) codes.

3.4 Two- versus four-dimensional schemes

A CM scheme can exploit the available four dimensions in the signal space of a fiber-optic link either jointly as a 4D channel or separately as two parallel 2D channels. For the Gaussian noise model introduced in Section 2.4, these parallel channels are independent, as shown in [101], and one can get close to the MI of an AWGN channel using both 1D and 2D schemes. Although a 2D CM scheme can achieve the MI of AWGN channels, a 4D CM scheme has a better trade-off between complexity and performance as discussed later in the performance analysis. In fact, a 4D scheme can provide more flexibility, which facilitates the rate adaptation and probabilistic shaping. To this end, we investigate four different combinations of 2D and 4D CM schemes with two types of component codes.

3.4.1 Two-dimensional schemes with binary codes

In this structure, classic and modern binary codes as well as their concatenation are used together with 2D constellations such as QAM signals. This combination is the most investigated structure in fiber-optic communications and has been realized based on the three traditional CM schemes, i.e., MLCM [25], TCM [24], and BICM [16]. This group of CM schemes is capable of approaching the AWGN capacity provided the block length is large enough. For example, an NCG of 10.8 dB ($\Delta\gamma = 3$ dB) with 20.5% redundancy overhead is achieved with triple-concatenated codes, (4608, 4080)[9] LDPC, (3860, 3824) BCH, and (2040, 1930) BCH using QPSK signals at a BER of 10^{-15} [14].

3.4.2 Two-dimensional schemes with nonbinary codes

This family of CM schemes is built by replacing the binary component codes of the previous family with nonbinary codes such as nonbinary LDPC codes [12, Ch. 14], [127]. Excluding BICM and MLCM, the rest of the CM schemes introduced in Fig. 3.4 belongs to this family. A better performance for short block lengths is achieved at the cost of a higher complexity. As discussed in [123], the correlation between bits of 2D binary labeling is exploited with nonbinary codes. In [123], two 2D CM schemes with nonbinary codes were proposed for a fiber-optic channel. The (1225, 1088) LDPC code over GF(2^3) with 12.6% redundancy overhead provides an NCG of 9.4 dB ($\Delta\gamma = 2.3$ dB) at a BER of 10^{-10}. The improvement over the comparable binary (3136, 2800) LDPC code from the same family is 0.7 dB at a BER of 10^{-7}. Further performance improvement was obtained by a longer (29136, 27315) LDPC code over GF(2^3) with 6.7 % redundancy overhead, which provides an NCG of 9.4 dB ($\Delta\gamma = 2.1$ dB) at a BER of 10^{-10}.

[9] (n, k) denotes a block code with a codeword of length n bits and an input information vector of length k bits.

3.4.3 Four-dimensional schemes with binary codes

Recently, 4D CM schemes adopted from classical communication have been suggested for optical communications based on BICM. Here, we discuss the 4D schemes with binary codes. For example, a 4D BICM scheme with two concatenated codes, an outer (992, 956) RS code and an inner (9252, 7976) LDPC code, can provide an NCG of 10.5 dB ($\Delta\gamma = 2.7$ dB) at a BER of 10^{-13} with an overall redundancy overhead of 20% and QPSK constellation [26]. If a Gray labeling is not used for the 4D constellation, an iterative decoding between the demapper and the binary code decoder is necessary [26]. As an example, for a PM QPSK with four-bit binary labeling, the conditions of the MIs (see Section 3.2) for binary subchannels can be removed provided that the Gray labeling is used, and thus they can be decoded independently, while for a non-Gray labeling they are dependent. This limitation imposes a higher complexity and consequently a larger processing latency than a conventional scheme.

3.4.4 Four-dimensional schemes with nonbinary codes

The main difference between this group and the previous 4D scheme is the use of nonbinary instead of binary codes, as in Fig. 3.4 (d) and (e). This provides improvement in NCG, for example 0.29 dB, 1.17 dB, and 2.17 dB with 16-, 32-, and 64-point 4D constellations, respectively, at the BER of 10^{-7} [27]. The nonbinary scheme in Fig. 3.4 (d) suffers from high complexity for constellations with a large number of symbols. The polar nonbinary CM scheme in Fig. 3.4 (e) decreases the complexity of the nonbinary CM schemes [27, 128] without performance degradation. To this end, the 4D labeling introduced in [120], [Paper D] is modified in [Paper E] to again categorize the binary subchannels into 'bad,' 'intermediate,' and 'good' channels. The polar CM scheme confines the number of 'intermediate' subchannels (the required GF order of the nonbinary block code) to a small number, independent of the constellation size. The mapper devised for polar CM scheme provides the minimum number of 'intermediate' subchannels. This 4D set partitioning using the bits $V_1, \ldots, V_4$ is illustrated in Fig. 3.6 for a PM-QPSK constellation.

3.5 Rate-adaptive coded modulation schemes

To improve the utilization of optical networks with dynamic or heterogeneous structures, the rate of the CM scheme can be adapted to the CSI at the transmitter of each fiber-optic link. Two well-known choices for the CSI are [114] the SNR, which is estimated after EDC, and the inner code BER in a concatenated coding scheme, which is computed by a syndrome-based error estimator. Rate-adaptive schemes have been investigated using multiple codes with different rates [114, 128, 129] or a single fixed-rate code [Paper E].

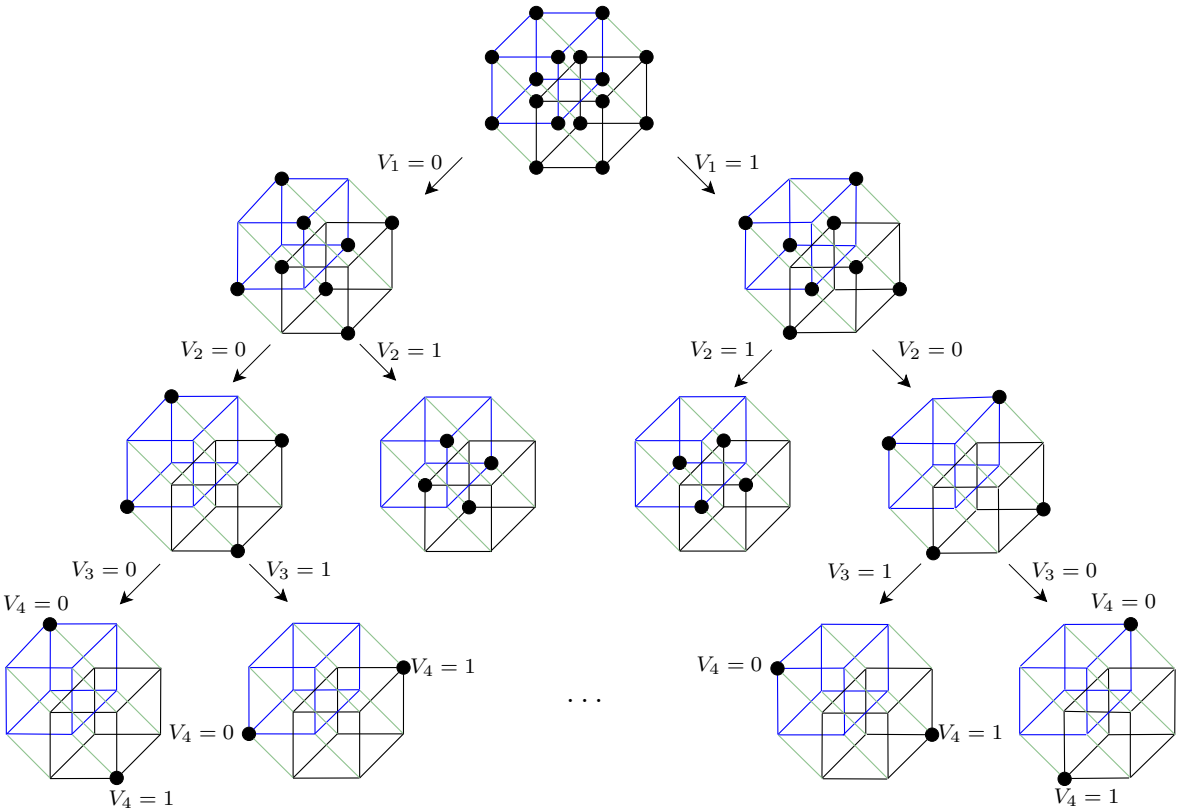


Figure 3.6: 4D set partitioning of a 16-ary 4D constellation representing PM-QPSK ($V_4V_3V_2V_1$ represents the four bits in the binary labeling of the constellation).

3.5.1 Multiple codes with different rates

CM schemes with multiple component codes to adapt the system rate to the desired rate supported by the link have been proposed for optical communications. Every code rate can be constructed either separately or by puncturing or shortening a single mother code. For example, a rate-adaptive nonbinary scheme with six nonbinary LDPC codes was proposed in [128] to provide a transmission bit rate between 100 Gb/s and 300 Gb/s in steps of 26.67 Gb/s at a fixed symbol rate. In a more practical scenario, a rate-adaptive BICM scheme was proposed exploiting six combinations of binary LDPC and RS codes together with three modulation formats [114]. This scheme was capable of obtaining $\Delta\gamma$ smaller than 2.9 and 3.9 dB for long and short non-DM single-channel fiber-optical links, respectively.

3.5.2 Single fixed-rate code

The method based on multiple codes with different rates is demanding in terms of hardware and thus costly to implement. A 4D scheme with a flexible structure can perform rate adaptation with a single component code rather than using a different

Table 3.1: CM schemes for different transparent reach (The number of coded bits is 6 in all cases). System parameters are given in Table 3.2, BER $\leq 10^{-15}$, RS(1022, 1004) outer code and (1920, 1546) LDPC inner code over GF(2^6) [Paper E].

Reach	CM parameters						
(km)	Constellation	DB	UB	SE	OH%	$\Delta\gamma$ (dB)	SNR
17×80	PM 64-QAM	0	6	5.25	14.3	2.67	17.57
31×80	PM 64-QAM	1	5	4.41	36.0	2.13	14.91
44×80	PM 64-QAM	2	4	3.74	61.3	2.60	13.41
55×80	PM 64-QAM	3	3	3.44	75.4	2.33	12.24
91×80	PM 16-QAM	1	1	2.77	44.9	2.41	10.23
112×80	PM 16-QAM	2	0	2.36	69.5	2.84	9.33

Table 3.2: System parameter values

Symbol rate R_s	32 Gbaud
Nonlinear coefficients γ	1.4 $\mathrm{W^{-1}km^{-1}}$
Attenuation coefficients α	0.2 dB/km
Dispersion coefficient D	17 ps/nm/km
Optical center wavelength λ	1550 nm
EDFA noise figure F_n	5 dB
Span length L	80 km

code for each rate. In fact, the new 4D mapper [Paper E], inspired by the polar coding technique [91], can be exploited to devise a rate-adaptive scheme with a single fixed-rate encoder. In contrast to [114], this approach exploits a simple circuitry to provide a rate-adaptive CM scheme. More precisely, the number of bits in the different 'good' and 'bad' groups introduced in the polar CM scheme in Section 3.2.5 are adjusted according to the CSI such that the number of 'intermediate' bits is always the same. Since the mapper is solely a simple look-up table, the rate adaptation is straightforward to implement.

Furthermore, as discussed in the previous section, the probabilistic shaping is performed over 'good' subchannels to increase the system's spectral efficiency or reduce its $\Delta\gamma$. Interestingly, this rate adaptation can be seen as a proper selection of 4D constellations extracted from well-known lattices [98]. Table 3.1 indicates how the number of dropped bits (DB) in the 4D mapper needs to be changed to support spectral efficiencies (SE) from 2.36 to 5.14 per polarization. As seen, for high SNRs, the PM 64-QAM constellation extracted from the 4D cubic lattice is used with 12 bits carried with each 4D symbol, while by decreasing the SNR, simply by changing the mapper and decreasing the number of uncoded bits (UB) with 1, we remove half of the 4D symbols to obtain a constellation with a dense packing, extracted from the so-called D_4 lattice with 2048 symbols, which is the best 4D packing lattice [130, 131]. The

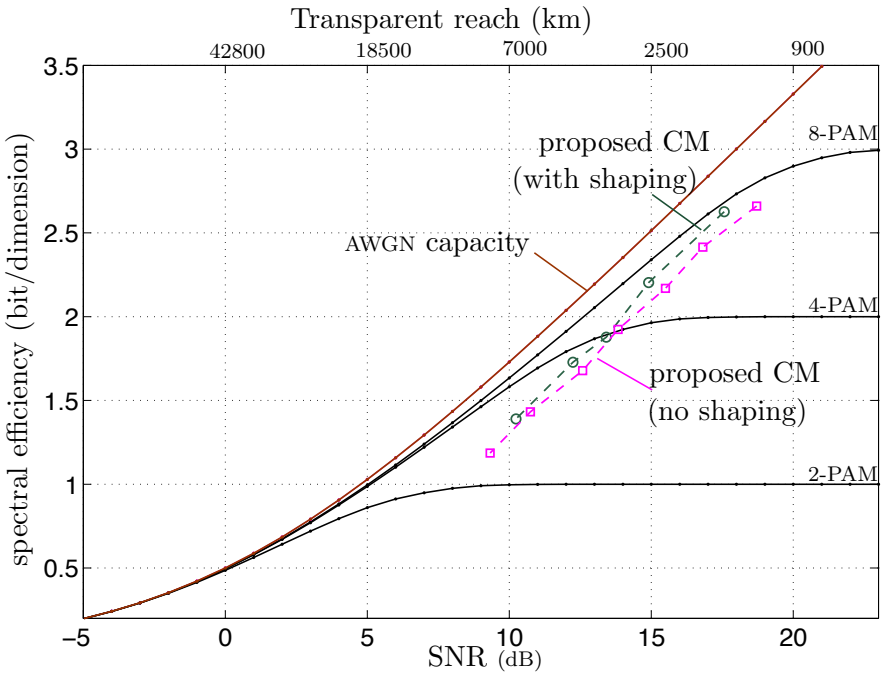


Figure 3.7: The spectral efficiency per dimension versus the transparent reach and the SNR for a non-DM link with EDC. The CM scheme curves are based on the results given in Table 3.1 and the spectral efficiency for the Gaussian noise model is computed by $\log_2(1 + \mathrm{SNR})/2$, where $\mathrm{SNR} = P/\sigma^2$ [Paper E].

number of coded bits, i.e., 'intermediate' subchannels, is fixed to 6, and these bits are coded by a fixed-rate nonbinary LDPC code.

In Fig. 3.7, the AWGN capacity (spectral efficiency) per dimension is illustrated versus the transmission length as well as the SNR for the rate-adaptive CM scheme introduced in [Paper E] over a non-DM fiber-optic link with the parameters given in Table 3.2. The spectral efficiency of the system with standard constellations, 4-PAM and 8-PAM, with or without probabilistic shaping are also plotted in this figure. The results show that the rate-adaptive CM scheme using single nonbinary code with probabilistic shaping can achieve $\Delta\gamma < 3$ dB for transmission lengths from 17×80 to 112×80 km.

3.5.3 Signal shaping

Signal shaping in data transmission systems with AWGN refers to the manipulation of the symbol distribution to make it as close as possible to a Gaussian distribution [98, 132–136]. Two types of shaping methods have been proposed for optical communications: probabilistic [25, 28], [Paper E] and geometric [137] shaping. Probabilistic shaping means changing the symbol probabilities for a standard constellation such as QAM, while geometric shaping implies changing the coordinates of the points

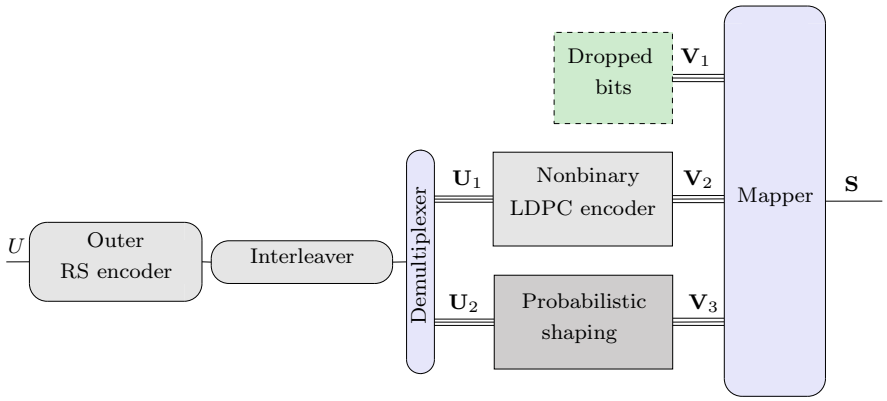


Figure 3.8: The encoder of the polar nonbinary CM scheme [28], [Paper E] with probabilistic shaping, including RS outer code and the three groups of, so-called 'bad' ($\mathbf{V}_1$), 'intermediate' ($\mathbf{V}_2$), and 'good' ($\mathbf{V}_3$), subchannels.

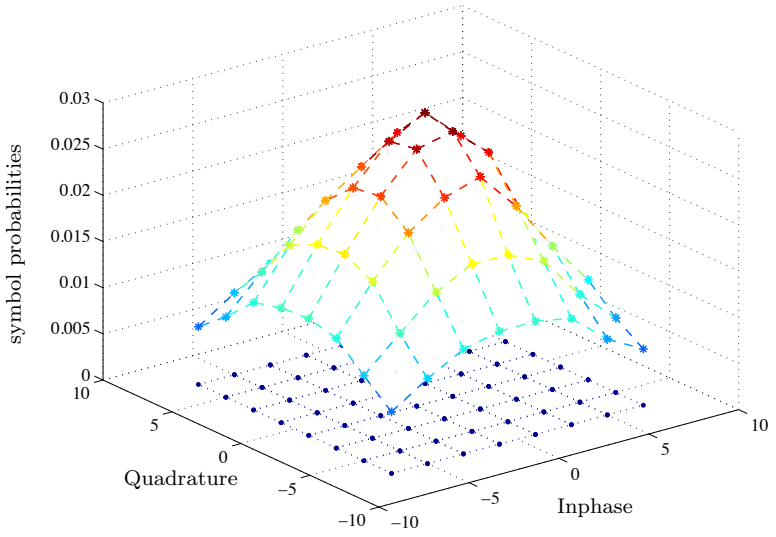


Figure 3.9: The 2D symbol probabilities of the probabilistically-shaped 4D CM scheme for the fiber-optic link with 17 spans (first row in Table 3.1).

in the constellation, which typically results in irregular constellations. In general, geometric shaping requires more complex hardware, such as higher-resolution ADC, than probabilistic shaping. Here, we describe probabilistic shaping and its combination with a CM scheme.

With probabilistic shaping, instead of having a uniform distribution for the input symbols, the symbols close to the origin of the constellation (with small amplitudes) are sent more often than the symbols far from the origin. This reduces the average trans-

mitted power compared with a uniform distribution. Bearing in mind that the variance of the introduced nonlinear distortion is cubic with input power (see Section 2.5.2), the system performance improves by performing probabilistic shaping as shown in Fig. 3.7. Two well-established probabilistic shaping methods, shell mapping [132] and trellis shaping [133], have been applied to fiber-optic communications in [28] and [25], respectively. As shown in Fig. 3.8, the binary shell mapping method [28], [Paper E] is applied to the 'good' subchannels, i.e., $\mathbf{V}_3$. As an example, the 2D symbol probabilities of 64-QAM with shell mapping are plotted in Fig. 3.9 for the system with 17 spans (first row in Table 3.1).

3.6 Performance and complexity analysis

We compare the BER performance for three CM schemes: 2D BICM, 2D nonbinary CM, and 4D polar nonbinary CM schemes, illustrated in Fig. 3.4 (b), (d), (e), respectively. All schemes were simulated with PM 64-QAM and an overall coding redundancy overhead of 21% over a single-channel non-DM fiber-optic link with the system parameters given in Table 3.2. Here, the schemes are compared based on two constraints: latency and complexity.

3.6.1 Latency-constrained comparison

Three systems are simulated with a small information block latency (consisting of inner and outer codes together with an interleaver as in Fig. 3.5) for the following scenarios: (i) a 2D BICM scheme with a (3, 21)-regular quasi-cyclic[10] binary (10752, 9236) LDPC inner code concatenated with a (1016, 980) shortened RS outer code over GF(2^{10}), to bring down the output BER of the inner code from 2.2×10^{-4} to 10^{-15}; (ii) a 2D nonbinary CM scheme with a (3, 9)-regular quasi-cyclic nonbinary (2688, 2309) LDPC inner code over GF(2^6) concatenated with a (970, 930) shortened RS code over GF(2^{10}), to bring down the output BER of the inner code from 1.9×10^{-4} to 10^{-15}; (iii) a 4D polar nonbinary CM scheme with a (3, 9)-regular quasi-cyclic nonbinary (1728, 1162) LDPC inner code over GF(2^6) concatenated with a (963, 949) shortened RS code over GF(2^{10}), to bring down the output BER of the inner code from 1.5×10^{-5} to 10^{-15}. The length of the interleaver between the inner and the outer code is 11 times the inner code length for the 2D BICM and 7 times the inner code length for the 2D nonbinary CM schemes, resulting coded block lengths of $11\times10752=118272$ and $7\times2688\times6=112896$ bits, respectively. The interleaver length is 5 times the inner code length for the 4D polar nonbinary CM scheme, resulting a coded block length of $5\times1728\times12=103680$ bits. Considering transmission of every 12 bits by a 4D symbol at 32 Gbaud, we obtain latencies of 308, 294, and 270 ns for the 2D BICM,

[10] A (γ,ρ)-regular quasi-cyclic LDPC code has γ nonzero elements in each column and ρ nonzero elements in each row of its parity-check matrix [12, Ch. 5].

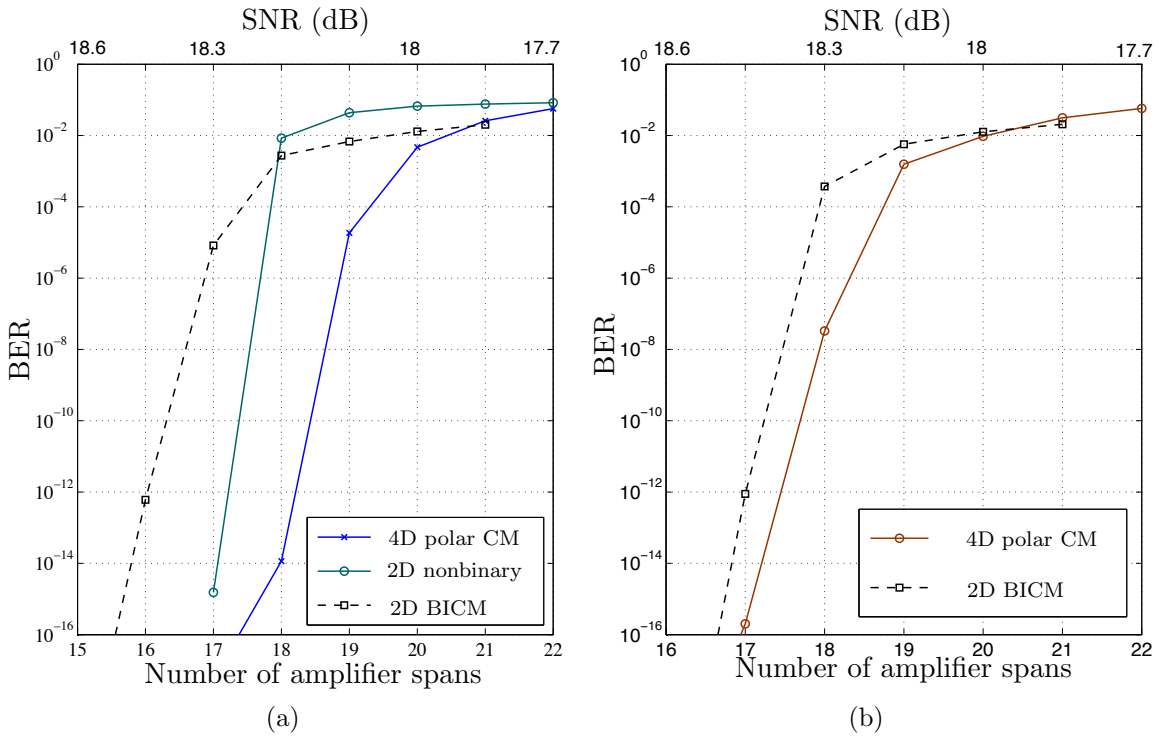


Figure 3.10: (a) The BER of three CM schemes with latency constraint. (b) The BER of 2D and 4D CM schemes with binary and nonbinary LDPC codes, respectively and similar complexity. All the CM schemes use PM 64-QAM with 21% coding redundancy overhead.

2D nonbinary, and polar 4D nonbinary schemes, respectively. According to the BER results shown in Fig. 3.10(a), the polar 4D nonbinary is superior to the 2D BICM and 2D nonbinary CM schemes with nearly the same latencies.

3.6.2 Complexity-constrained comparison

We designed the following 2D and 4D schemes with similar complexities: (i) a 2D BICM scheme consisting of a (3, 21)-regular quasi-cyclic binary (16128, 13844) LDPC inner code concatenated with a (1015, 977) shortened RS outer code over GF(2^{10}), to bring down the output BER of the inner code from 2.3×10^{-4} to 10^{-15}; (ii) a 4D polar nonbinary CM scheme consisting of a (3, 9)-regular quasi-cyclic nonbinary (1152, 778) LDPC inner code over GF(2^6) concatenated with a (1011, 995) shortened RS outer code over GF(2^{10}), to bring down the output BER of the inner code from 2.5×10^{-5} to 10^{-15}. As seen in Fig. 3.10(b), the 4D polar nonbinary scheme performs slightly better. Since the GF order can be kept fixed in this scheme, i.e., GF(2^6), independent of the constellation size, the 4D scheme is superior to the 2D scheme for large constellations.

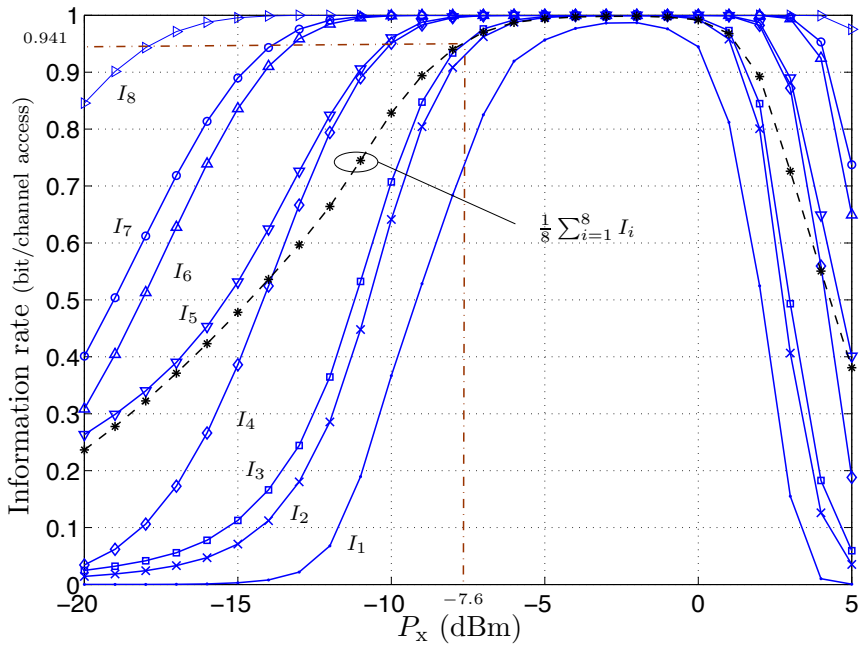


Figure 3.11: The information rate of binary channels for a 4D mapper with PM 16-QAM constellation and a fiber-optical link with 15 spans of length 100 km at 28 Gbaud.

3.6.3 Performance improvement using probabilistic shaping

In this section, we evaluate the performance of the polar CM scheme in Fig. 3.8 for a single-channel fiber-optical system with PM 16-QAM constellation, 32-ary (1024, 928) NB-LDPC code designed using the method introduced in [12, Sec. 14.4.2], and three different probabilistic shaping schemes: no probabilistic shaping and probabilistic shaping of rates 0.9375 and 0.875 [28]. The numerical results are extracted for a non-DM link with 15 spans of 100 km in length at 28 Gbaud. In Fig. 3.11, the information rates of the binary channels of the described fiber-optical link are plotted versus the transmit power $P_x = P/2$ in one of the polarizations (assuming the same transmit power in the other polarization). As shown in this figure, a CM scheme with a total rate of 0.941 was selected for numerical simulations. At this rate, we protected five bits by a 32-ary NB-LDPC code and the rest of the bits can be left uncoded. The bit, symbol, and frame error rates (BER, SER, and FER) of this system are evaluated for three different scenarios: the CM scheme with a total rate of 0.941 but no probabilistic shaping and the CM scheme with probabilistic shaping and total rates of 0.926 and 0.934. The SER is measured by adding an HDD in parallel to the CM decoder. We emphasize that the CM decoder uses soft input symbols and this HDD is only used to measure the uncoded system SER. As seen in Fig. 3.12, exploiting a probabilistic shaping with an overhead of 0.016 can improve the system performance around 0.4 dB at BER $= 10^{-5}$. This improves the uncoded 4D FEC threshold of the system from a SER of 0.058 to 0.072 at the information BER of 10^{-5}.

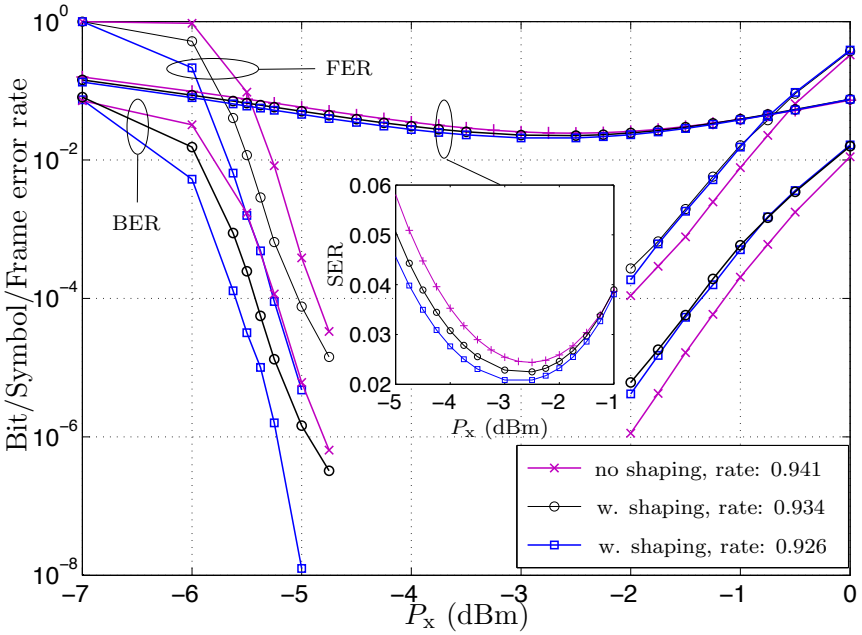


Figure 3.12: The performance of CM with PM 16-QAM constellation, 32-ary (1024, 928) NB-LDPC code, without probabilistic shaping and with two shaping rates. The fiber-optical link has 15 spans of length 100 km at 28 Gbaud.

3.6.4 Hardware complexity

The hardware requirements and electronic processing complexity of CM schemes play a crucial role for fiber-optic communications. Although the semiconductor technology is capable of providing ultra-high-speed ADC and massively parallelized DSP circuits, the system power consumption and hardware cost also need to be taken into account. In particular, since high-resolution ADC and DSP are costly for high-speed data transmission, the performance sensitivity of CM schemes to quantization errors has become an important factor in the design of these schemes [14]. The impact of quantization errors on the performance of a concatenated TCM scheme with two interleaved BCH outer codes was evaluated in [24], and it was shown that 4-bit quantization was sufficient to approach the infinite precision performance to within 0.15 dB.

The complexity of a CM scheme is dominated by its two main components: LLR calculation from soft received symbols and the encoder and decoder of the component codes. Since the complexity of LLR calculations considering well-known approximations [12, Ch. 7] is almost the same for the studied schemes, one may compare the complexity of the CM schemes with different dimensions by taking into account solely the complexity of the component code decoders per dimension [120]. The complexity of LDPC and RS codes have been well studied [12, Ch. 14], [138].

Chapter 4

Conclusions

To efficiently utilize the available resources in an optical network, the trade-offs between spectral efficiency, DSP hardware complexity, and transparent reach need to be considered for the different links in the network. Joint coding and modulation schemes offer more freedom to exploit the available four dimensions in these channels than traditional independent FEC and modulation techniques. As discussed, a CM scheme can enable a link with a larger transmission distance than conventional schemes but with the same (or even lower) complexity, for a wide range of spectral efficiencies. The aim of this work was to design a rate-adaptive CM scheme with a higher spectral efficiency than existing methods, while keeping the system complexity affordable. In order to achieve this goal, an accurate channel model is essential, which was developed as part of this work as well. The main contributions of the thesis are summarized in this chapter.

Summary of Contributions

We reviewed well-established and novel CM schemes for AWGN channels and their adoption in optical communications consisting of MLCM, BICM, TCM, nonbinary, and polar nonbinary schemes using an information-theoretic design framework. MLCM with a small number of component codes is an attractive solution for fiber-optic communications because of its low-complexity multistage decoding. The main bottleneck of nonbinary schemes is the decoder and its high complexity for large constellations. Although BICM has been widely investigated using two- and four-dimensional constellations for fiber-optic channels, channel-aware designed schemes such as turbo TCM or polar nonbinary schemes provide better error performance for short-length component codes.

In order to design a channel-aware CM scheme for a PM DM fiber-optical link, statistics of the received signal are derived in [Paper A] with a negligible chromatic dispersion. This makes it possible, for the first time, to analytically evaluate the perfor-

mance of data transmission systems over PM fiber-optic channels with phase noise and low enough symbol rate, and to optimize the performance of such systems. Moreover, a quantitative approach is proposed to measure the accuracy of the analytically derived pdf for different DM links using the SSFM. The derivations are performed for both lumped and distributed amplifications. These statistics consist of the pdf of the nonlinear phase noise and the joint probability density function of the received amplitudes and phases given the SNR of both polarizations.

Including chromatic dispersion and nonlinear Kerr effect, an analytical approach is introduced in [Paper B] to model a non-DM fiber-optic link in the pseudolinear regime based on the SSFM. According to this model, for high enough symbol rates, a fiber-optic link can be described as a linear dispersive channel with AWGN and a complex scaling. The variance of this AWGN noise and the attenuation are computed analytically as a function of input power and channel parameters. The results illustrate a cubic growth of the noise variance with input power. Moreover, the interchannel effect between the two polarizations and the intrachannel effect between the amplifier noise and the transmitted signal due to the nonlinear Kerr effect are described. In particular, it is found that the channel noise variance in one polarization is affected twice as much by the transmitted power in that polarization than by the transmitted power in the orthogonal polarization. The effect of pulse shaping is also investigated through numerical simulations. The SSFM numerical results justify the accuracy of this model for a symbol rate of 28 Gbaud and above.

Exploiting the signal statistics of highly nonlinear single-polarization fiber-optic links and the information-theoretic design framework, we devised an MLCM scheme with a hard-decision multistage decoder in [Paper C]. More precisely, an unequal error protection in the phase and radial direction is exploited to optimize the performance (block error rate) of the system. It is shown that the new MLCM system can give better performance with lower complexity than independent FEC and modulation. Hence, the MLCM scheme provides the possibility of reliable data transmission in a longer fiber or at a higher spectral efficiency.

To decrease the complexity of the MLCM scheme, an algorithm is proposed in [Paper D] with a multidimensional set partitioning method. This multidimensional MLCM shows better trade-off between performance and complexity for classical codes such as RS and BCH codes. The numerical results illustrate that for practical SNRs, we can design four-dimensional MLCM schemes with lower complexities and a higher power efficiency than the one-dimensional systems.

A better trade-off between DSP complexity and transparent reach of 4D CM schemes makes them superior to 2D schemes. In [Paper E], we addressed this result and used the introduced design framework to devise a new 4D mapper inspired by polar coding to reduce the computational complexity of the nonbinary CM schemes without performance degradation. A distinct contribution of the new CM scheme is in providing a flexible 4D structure, using the 4D mapper and a probabilistic shaping method based

on the shell mapping algorithm. In contrast to the existing rate-adaptive schemes in the literature, the proposed scheme used solely a single nonbinary LDPC code rather than several binary or nonbinary component codes. Finally, simulation results are provided over a non-DM PM single-channel fiber-optic link. According to the numerical results, the proposed scheme can operate within 2.7 dB from the AWGN capacity, showing 1 dB performance improvement compared to existing results with the same system constraints.

Bibliography

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Systems Technical Journal*, vol. 27, pp. 379–423, Jul. 1948.

[2] G. P. Agrawal, *Lightwave Technology: Telecommunication Systems*. Hoboken, NJ: Wiley-Interscience, 2005.

[3] R. Tkach, "Scaling optical communications for the next decade and beyond," *Bell Labs Tech. J.*, vol. 14, no. 4, pp. 3–9, Feb. 2010.

[4] K. S. Turitsyn, S. A. Derevyanko, I. V. Yurkevich, and S. K. Turitsyn, "Information capacity of optical fiber channels with zero average dispersion," *Phys. Rev. Lett.*, vol. 91, no. 20, p. 203901, Nov. 2003.

[5] M. Taghavi, G. Papen, and P. Siegel, "On the multiuser capacity of WDM in a nonlinear optical fiber: Coherent communication," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5008–5022, 2006.

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

[7] W. Shieh and X. Chen, "Information spectral efficiency and launch power density limits due to fiber nonlinearity for coherent optical OFDM systems," *IEEE Photon. Journal*, vol. 3, no. 22, pp. 158–173, Nov. 2011.

[8] M. I. Yousefi and F. R. Kschischang, "On the per-sample capacity of nondispersive optical fibers," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7522–7541, Nov. 2011.

[9] E. Agrell, "On monotonic capacity–cost functions," 2012. [Online]. Available: http://arxiv.org/abs/1108.0391

[10] ITU-T G.709, "Interfaces for optical transport network," Dec. 2009.

[11] ITU-T G.975.1, "Forward error correction for high bit-rate DWDM submarine systems,," Feb. 2004.

[12] W. E. Ryan and S. Lin, *Channel Codes: Classical and Modern.* Cambridge University Press, 2009.

[13] B. Zhou, L. Zhang, J. Kang, Q. Huang, Y. Tai, S. Lin, and M. Xu, "Non-binary LDPC codes vs. Reed-Solomon codes," *in Proc. Information Theory and Applications Workshop*, pp. 175–184, Feb. 2008.

[14] F. Chang, K. Onohara, and T. Mizuochi, "Forward error correction for 100 G transport networks," *IEEE Commun. Mag.*, vol. 48, no. 3, pp. S48–S55, Mar. 2010.

[15] K. Onohara, T. Sugihara, Y. Konishi, Y. Miyata, T. Inoue, S. Kametani, K. Sugihara, K. Kubo, H. Yoshida, and T. Mizuochi, "Soft-decision-based forward error correction for 100 Gb/s transport systems," *IEEE J. Sel. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1258–1267, Sep. 2010.

[16] I. B. Djordjevic, M. Arabaci, and L. L. Minkov, "Next generation FEC for high-capacity communication in optical transport networks," *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3518–3530, Aug. 2009.

[17] B. P. Smith and F. R. Kschischang, "Future prospects for FEC in fiber-optic communications," *IEEE J. Sel. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1245–1257, Sep./Oct. 2010.

[18] J. L. Massey, "Coding and modulation in digital communications," *in Proc. Int. Zurich Seminar on Digital Commun.*, Mar. 1974.

[19] G. Ungerboeck and I. Csajka, "On improving data-link performance by increasing channel alphabet and introducing sequence decoding," *in Proc. IEEE Int. Symp. on Inform. Theory*, Ronneby, Sweden, June 1976.

[20] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 55–67, Jan. 1982.

[21] H. Imai and S. Hirakawa, "A new multilevel coding method using error correcting codes," *IEEE Trans. Inf. Theory*, vol. 23, pp. 371–377, May 1977.

[22] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, no. 5, pp. 873–884, May 1992.

[23] H. Zhao, E. Agrell, and M. Karlsson, "Trellis-coded modulation in PSK and DPSK communications," *in Proc. European Conf. and Exhibition on Optic. Commun.*, Sep. 2006.

[24] M. Magarini, R.-J. Essiambre, B. E. Basch, A. Ashikhmin, G. Kramer, and A. J. de Lind van Wijngaarden, "Concatenated coded modulation for optical communications systems," *IEEE Photon. Technol. Lett.*, vol. 22, no. 16, pp. 1244–1246, Aug. 2010.

[25] B. P. Smith and F. R. Kschischang, "A pragmatic coded modulation scheme for high-spectral-efficiency fiber-optic communications," *J. Lightw. Technol.*, vol. 30, no. 13, pp. 2047–2053, Jul. 2012.

[26] H. Bülow and E. Masalkina, "Coded modulation in optical communications," *in Proc. Optic. Fiber Commun. Conf.*, Mar. 2011.

[27] M. Arabaci, I. Djordjevic, L. Xu, and T. Wang, "Four-dimensional non-binary LDPC-coded modulation schemes for ultra-high-speed optical fiber communication," *IEEE Photon. Technol. Lett.*, vol. 23, no. 18, pp. 1280–1282, Sep. 2011.

[28] L. Beygi, E. Agrell, and M. Karlsson, "Adaptive coded modulation for nonlinear fiber-optical channels," *in Proc. IEEE Global Commun. Conf.*, Dec. 2012, Anaheim, USA.

[29] E. Hecht, *Optics*, 4th ed. Addison Wesley, 2001.

[30] G. P. Agrawal, *Nonlinear fiber optics*, 4th ed. Academic Press, 2007.

[31] D. Wang and C. Menyuk, "Polarization evolution due to the Kerr nonlinearity and chromatic dispersion," *J. Lightw. Technol.*, vol. 17, no. 12, pp. 2520–2529, Dec. 1999.

[32] B. P. Smith, "Error-correcting codes for fibre-optic communication systems," *Ph.D. dissertation, University of Toronto, Canada*, 2011. [Online]. Available: https://tspace.library.utoronto.ca/bitstream/1807/31942/1/Smith_Benjamin_P_201111_PhD_thesis.pdf

[33] G. P. Agrawal, *Fiber-Optic Communication Systems*, 3rd ed. Wiley, 2002.

[34] E. Ip and J. M. Kahn, "Digital equalization of chromatic dispersion and polarization mode dispersion," *J. Lightw. Technol.*, vol. 25, no. 8, pp. 2033–2043, Aug. 2007.

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

[36] E. Ip, "Nonlinear compensation using backpropagation for polarization-multiplexed transmission," *J. Lightw. Technol.*, vol. 28, no. 6, pp. 939 –951, Mar. 2010.

[37] P. Johannisson and M. Karlsson, "Perturbation analysis of nonlinear propagation in a strongly dispersive optical communication system," *J. Lightw. Technol.*, vol. 31, no. 8, pp. 1273–1282, 2013.

[38] A. Bononi and P. Serena, "An alternative derivation of Johannisson's regular perturbation model," 2012. [Online]. Available: http://arXiv:1207.4729

[39] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "A detailed analytical derivation of the GN model of non-linear interference in coherent optical transmission systems," 2012. [Online]. Available: http://arXiv:1209.0394

[40] J. G. Proakis and M. Salehi, *Digital Communications*, 5th ed. McGraw-Hill, 2008.

[41] J. P. Gordon and L. F. Mollenauer, "Phase noise in photonic communications systems using linear amplifiers," *Opt. Lett.*, vol. 15, no. 23, pp. 1351–1353, 1990.

[42] A. Mecozzi, "Probability density functions of the nonlinear phase noise," *Opt. Lett.*, vol. 29, no. 7, pp. 673–675, 2004.

[43] K.-P. Ho, "Probability density of nonlinear phase noise," *J. Opt. Soc. Am. B*, vol. 20, no. 9, pp. 1875–1879, 2003.

[44] ——, "Performance of DPSK signals with quadratic phase noise," *IEEE Trans. Commun.*, vol. 53, no. 8, pp. 1361–1365, Aug. 2005.

[45] ——, *Phase-Modulated Optical Communication Systems*. Springer, 2005.

[46] S. Kumar, "Analysis of nonlinear phase noise in coherent fiber-optic systems based on phase shift keying," *J. Lightw. Technol.*, vol. 27, no. 21, pp. 4722–4733, Nov. 2009.

[47] A. Demir, "Nonlinear phase noise in optical-fiber-communication systems," *J. Lightw. Technol.*, vol. 25, no. 8, pp. 2002–2032, Aug. 2007.

[48] G. Charlet, N. Maaref, J. Renaudier, H. Mardoyan, P. Tran, and S. Bigo, "Transmission of 40 Gb/s QPSK with coherent detection over ultra-long distance improved by nonlinearity mitigation," *in Proc. European Conf. and Exhibition on Optic. Commun.*, 2006, Th4.3.4.

[49] C. Xu and X. Liu, "Postnonlinearity compensation with data-driven phase modulators in phase-shift keying transmission," *Opt. Lett.*, vol. 27, no. 18, pp. 1619–1621, 2002.

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

[51] Y. Yadin, M. Orenstein, and M. Shtaif, "Statistics of nonlinear phase noise in phase modulated fiber-optic communications systems," *in Proc. Optic. Fiber Commun. Conf.*, vol. 1, pp. 1800–1801, Feb. 2004, MF59.

[52] M. I. Yousefi and F. R. Kschischang, "A Fokker-Planck differential equation approach for the zero-dispersion optical fiber channel," *in Proc. IEEE Int. Symp. on Inform. Theory*, pp. 206–210, Jun. 2010.

[53] Y. Yadin, M. Shtaif, and M. Orenstein, "Nonlinear phase noise in phase-modulated WDM fiber-optic communications," *IEEE Photon. Technol. Lett.*, vol. 16, no. 5, pp. 1307–1309, May 2004.

[54] P. Serena, A. Orlandini, and A. Bononi, "Parametric-gain approach to the analysis of single-channel DPSK/DQPSK systems with nonlinear phase noise," *J. Lightw. Technol.*, vol. 24, no. 5, p. 2026, May 2006.

[55] G. Grimmett and D. Stirzaker, *Probability and Random Processes*, 3rd ed. Oxford University Press, 2001.

[56] L. Beygi, E. Agrell, and M. Karlsson, "Optimization of 16-point ring constellations in the presence of nonlinear phase noise," *in Proc. Optic. Fiber Commun. Conf.*, 2011, OThO4.

[57] A. P. T. Lau and J. M. Kahn, "Signal design and detection in presence of nonlinear phase noise," *J. Lightw. Technol.*, vol. 25, no. 10, pp. 3008–3016, Oct. 2007.

[58] K.-P. Ho and J. M. Kahn, "Electronic compensation technique to mitigate nonlinear phase noise," *J. Lightw. Technol.*, vol. 22, no. 3, pp. 779–783, Mar. 2004.

[59] K. Kikuchi, M. Fukase, and S.-Y. Kim, "Electronic post-compensation for nonlinear phase noise in a 1000-km 20-Gbit/s optical QPSK transmission system using the homodyne receiver with digital signal processing," *in Proc. Optic. Fiber Commun. Conf.*, Mar. 2007, OTuA2.

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

[61] K. Roberts, C. Li, L. Strawczynski, M. O'Sullivan, and I. Hardcastle, "Electronic precompensation of optical nonlinearity," *IEEE Photon. Technol. Lett.*, vol. 18, no. 2, pp. 403–405, Jan. 2006.

[62] R. Waegemans, S. Herbst, L. Holbein, P. Watts, P. Bayvel, C. Fürst, and R. I. Killey, "10.7 Gb/s electronic predistortion transmitter using commercial FPGAs and D/A converters implementing real-time DSP for chromatic dispersion and SPM compensation," *Opt. Express*, vol. 17, no. 10, pp. 8630–8640, 2009.

[63] K. Sponsel, C. Stephan, G. Onishchukov, B. Schmauss, and G. Leuchs, "Nonlinear phase noise compensation using a modified nonlinear optical loop mirror," *in Proc. Optic. Fiber Commun. Conf.*, Mar. 2010.

[64] K.-P. Ho and J. M. Kahn, "Electronic compensation technique to mitigate nonlinear phase noise," *J. Lightw. Technol.*, vol. 22, no. 3, pp. 779–783, Mar. 2004.

[65] E. Torrengo, R. Cigliutti, G. Bosco, A. Carena, V. Curri, P. Poggiolini, A. Nespola, D. Zeolla, and F. Forghieri, "Experimental validation of an analytical model for nonlinear propagation in uncompensated optical links," *Opt. Express*, vol. 19, pp. B790–B798, Dec. 2011.

[66] G. Bosco, R. Cigliutti, A. Nespola, A. Carena, V. Curri, F. Forghieri, Y. Yamamoto, T. Sasaki, Y. Jiang, and P. Poggiolini, "Experimental investigation of nonlinear interference accumulation in uncompensated links," *IEEE Photon. Technol. Lett.*, vol. 24, no. 14, pp. 1230–1232, 2012.

[67] F. Vacondio, O. Rival, C. Simonneau, E. Grellier, A. Bononi, L. Lorcy, J.-C. Antona, and S. Bigo, "On nonlinear distortions of highly dispersive optical coherent systems," *Opt. Express*, vol. 20, no. 2, pp. 1022–1032, Jan. 2012.

[68] A. Carena, G. Bosco, V. Curri, P. Poggiolini, M. Taiba, and F. Forghieri, "Statistical characterization of PM-QPSK signals after propagation in uncompensated fiber links," *in Proc. European Conf. and Exhibition on Optic. Commun.*, Sep. 2010, P4.07.

[69] F. Vacondio, C. Simonneau, L. Lorcy, J. C. Antona, A. Bononi, and S. Bigo, "Experimental characterization of Gaussian-distributed nonlinear distorsion," *in Proc. European Conf. and Exhibition on Optic. Commun.*, 2011, We.7.B.1.

[70] G. Raybon, P. Winzer, A. Adamiecki, A. Gnauck, A. Konczykowska, F. Jorge, J.-I. Dupuy, A. Sureka, C. Scholz, R. Delbue, P. Pupalaikis, L. Buhl, C. R. Doerr, S. Chandrasekhar, B. Zhu, and D. Peckham, "8 x 320-Gb/s transmission over 5600 km using all-ETDM 80-Gbaud polarization multiplexed QPSK transmitter and coherent receiver," *in Proc. Optic. Fiber Commun. Conf.*, Mar. 2012, OTu2A.1.

[71] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *J. Lightw. Technol.*, vol. 26, no. 20, pp. 3416–3425, Oct. 2008.

[72] L. Beygi, N. V. Irukulapati, E. Agrell, P. Johannisson, M. Karlsson, H. Wymeersch, P. Serena, and A. Bononi, "On nonlinearly-induced noise in optical links with digital backpropagation," *Opt. Express*, Jun. 2013, submitted.

[73] P. Poggiolini, "The GN model of non-linear propagation in uncompensated coherent optical systems," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3857–3879, Dec. 2012.

[74] A. Splett, C. Kurzke, and K. Petermann, "Ultimate transmission capacity of amplified optical fiber communication systems taking into account fiber nonlinearities," *in Proc. European Conf. and Exhibition on Optic. Commun.*, vol. 2, pp. 41–42, 1993.

[75] H. Louchet, A. Hodzic, and K. Petermann, "Analytical model for the performance evaluation of DWDM transmission systems," *IEEE Photon. Technol. Lett.*, vol. 15, no. 9, pp. 1219–1221, 2003.

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

[77] A. Mecozzi and R.-J. Essiambre, "Nonlinear Shannon limit in pseudolinear coherent systems," *J. Lightw. Technol.*, vol. 30, pp. 2011–2024, Jun. 2012.

[78] E. F. M. Secondini, "Analytical fiber-optic channel model in the presence of cross-phase modulation," *IEEE Photon. Technol. Lett.*, vol. 24, no. 22, pp. 2016–2019, Nov. 2012.

[79] A. Bononi, P. Serena, N. Rossi, E. Grellier, and F. Vacondio, "Modeling nonlinearity in coherent transmissions with dominant intrachannel-four-wave-mixing," *Opt. Express*, vol. 20, no. 7, pp. 7777–7791, Mar. 2012.

[80] J. Tang, "Perturbation analysis of nonlinear propagation in a strongly dispersive optical communication system," *J. Lightw. Technol.*, vol. 24, pp. 2070–2075, 2006.

[81] H. Song and M. Brandt-Pearce, "A 2-D discrete-time model of physical impairments in wavelength-division multiplexing systems," *J. Lightw. Technol.*, vol. 30, no. 5, pp. 713–726, Mar. 2012.

[82] P. Poggiolini, A. Carena, V. Curri, G. Bosco, and F. Forghieri, "Analytical modeling of nonlinear propagation in uncompensated optical transmission links," *IEEE Photon. Technol. Lett.*, vol. 23, no. 11, pp. 742–744, Jun. 2011.

[83] A. Mecozzi and F. Matera, "Polarization scattering by intra-channel collisions," *Opt. Express*, vol. 20, no. 2, pp. 1213–1218, Jan. 2012.

[84] R.-J. Essiambre and A. Mecozzi, "Capacity limits in single-mode fiber and scaling for spatial multiplexing," *in Proc. Optic. Fiber Commun. Conf.*, Mar. 2012.

[85] A. Bononi, N. Rossi, and P. Serena, "Transmission limitations due to fiber nonlinearity," *in Proc. Optic. Fiber Commun. Conf.*, Mar. 2011, OWO7.

[86] A. Vannucci, P. Serena, and A. Bononi, "The RP method: a new tool for the iterative solution of the nonlinear Schrodinger equation," *Lightwave Technology, Journal of*, vol. 20, no. 7, pp. 1102–1112, 2002.

[87] M. I. Yousefi and F. R. Kschischang, "Information transmission using the nonlinear Fourier transform, Part I: Mathematical tools," *Submitted to IEEE Transactions on Information Theory, arXiv:1202.3653v2*, 2012.

[88] M. I. Yousefi, "Information transmission using the nonlinear Fourier transform," *Ph.D. dissertation, University of Toronto, Canada*, 2013. [Online]. Available: https://tspace.library.utoronto.ca/bitstream/1807/35179/6/Yousefi_Mansoor_Isvand_201303_PhD_thesis.pdf

[89] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes." *in Proc. IEEE Int. Conf. on Commun.*, vol. 2, pp. 1064–1070, May 1993.

[90] R. Gallager, "Low-density parity-check codes," *IRE Trans. on Inf. Theory*, vol. 8, no. 1, pp. 21–28, Jan. 1962.

[91] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, Jul. 2009.

[92] G. Bosco and S. Benedetto, "Soft decoding in optical systems: Turbo product codes vs. LDPC codes," in *Optical Communication Theory and Techniques*, E. Forestieri, Ed. Springer US, 2005, pp. 79–86. [Online]. Available: http://link.springer.com/content/pdf/10.1007%2F0-387-23136-6_9.pdf

[93] M. S. Kumar, "Asynchronous BPPM OCDMA systems with trellis-coded modulation," *IEE Proceedings Optoelectronics*, vol. 151, no. 4, pp. 193–201, Aug. 2004.

[94] J. M. Kahn, "Modulation and detection techniques for optical communication systems," *Optical Amplifiers and Their Applications/Coherent Optical Technologies and Applications*, 2006, CThC1.

[95] Y. Cai, J. M. Morris, T. Adali, and C. R. Menyuk, "On turbo code decoder performance in optical-fiber communication systems with dominating ASE noise," *J. Lightw. Technol.*, vol. 21, no. 3, pp. 727–734, Mar. 2003.

[96] I. B. Djordjevic, M. Arabaci, and L. L. Minkov, "Next generation FEC for high-capacity communication in optical transport networks," *J. Lightw. Technol.*, vol. 27, pp. 3518–3530, Aug. 2009.

[97] M.-S. Kao, H.-Y. Chen, and D. Lin, "A product-coded WDM coding system," *IEEE Trans. Commun.*, vol. 44, no. 1, pp. 43–46, Jan. 1996.

[98] G. D. Forney, Jr. and G. Ungerboeck, "Modulation and coding for linear Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2384–2415, Oct. 1998.

[99] A. R. Calderbank, "Multilevel codes and multistage decoding," *IEEE Trans. Inf. Theory*, vol. 37, no. 3, pp. 222–229, Mar. 1989.

[100] D. J. Costello, Jr. and G. D. Forney, Jr., "Channel coding: The road to channel capacity," *Proc. IEEE*, vol. 95, no. 6, pp. 1150–1177, Jun. 2007.

[101] U. Wachsmann, R. F. H. Fischer, and J. B. Huber, "Multilevel codes: theoretical concepts and practical design rules," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1361–1391, Jul. 1999.

[102] T. J. Lunn and A. G. Burr, "Number of neighbours for staged decoding of block coded modulation," *Electron. Lett*, vol. 29, no. 21, pp. 1830–1831, Oct. 1993.

[103] A. G. Burr and T. J. Lunn, "Block-coded modulation optimized for finite error rate on the white Gaussian noise channel," *IEEE Trans. Inf. Theory*, vol. 43, no. 1, 1997.

[104] G. D. Forney, Jr., M. Trott, and S. Y. Chung, "Sphere-bound-achieving coset codes and multilevel coset codes," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 820–850, May 2000.

[105] G. J. Pottie and D. P. Taylor, "Multilevel codes based on partitioning," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 87–98, Jan. 1989.

[106] L. Beygi, E. Agrell, J. M. Kahn, and M. Karlsson, "Coded modulation for fiber-optical channels," *IEEE Signal Process. Mag.*, (special issue), May 2013, submitted.

[107] E. Agrell and M. Karlsson, "Satellite Constellations: Towards the nonlinear channel capacity," *in Proc. IEEE Photonic Conference (IPC), Burlingame, CA*, Sep. 2012.

[108] E. Agrell, "Nonlinear fiber capacity," *in Proc. European Conf. and Exhibition on Optic. Commun.*, London, UK, Sept. 2013, to appear.

[109] K. S. Turitsyn and S. K. Turitsyn, "Nonlinear communication channels with capacity above the linear Shannon limit," *Opt. Lett.*, vol. 37, no. 17, pp. 3600–3602, Sep 2012. [Online]. Available: http://ol.osa.org/abstract.cfm?URI=ol-37-17-3600

[110] E. Ip and J. M. Kahn, "Fibre communications: Time-reversed twin," *Nat. Photon.*, vol. 7, pp. 507–508, June 2013. [Online]. Available: http://dx.doi.org/10.1038/nphoton.2013.155

[111] G. Bosco, P. Poggiolini, A. Carena, V. Curri, and F. Forghieri, "Analytical results on channel capacity in uncompensated optical links with coherent detection," *Opt. Express*, vol. 19, no. 26, pp. B440–B451, Dec. 2011.

[112] A. D. Ellis, J. Zhao, and D. Cotter, "Approaching the non-linear Shannon limit," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 423–433, Feb. 2010.

[113] E. Ip, P. Ji, E. Mateo, Y.-K. Huang, D. Xu, L. ; Qian, N. Bai, and T. Wang, "100G and beyond transmission technologies for evolving optical networks and relevant physical-layer issues," *Proc. IEEE*, vol. 100, pp. 1065–1078, 2012.

[114] G.-H. Gho and J. M. Kahn, "Rate-adaptive modulation and low-density parity-check coding for optical fiber transmission systems," *J. Optical Commun. Netw.*, vol. 4, no. 10, pp. 760–768, Oct. 2012.

[115] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[116] J. Hou, P. Siegel, L. Milstein, and H. Pfister, "Capacity-approaching bandwidth-efficient coded modulation schemes based on low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 9, pp. 2141–2155, Sep. 2003.

[117] H. Imai and S. Hirakawa, "Correction to 'A new multilevel coding method using Error-Correcting codes'," *IEEE Trans. Inf. Theory*, vol. 23, no. 6, pp. 784–784, Nov. 1977.

[118] H. Batshon, I. Djordjevic, L. Xu, and T. Wang, "Multidimensional LDPC-coded modulation for beyond 400 Gb/s per wavelength transmission," *IEEE Photon. Technol. Lett.*, vol. 21, no. 16, pp. 1139–1141, Aug. 2009.

[119] P. Robertson and T. Worz, "Bandwidth-efficient turbo trellis-coded modulation using punctured component codes," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 2, pp. 206–218, Feb. 1998.

[120] L. F. Wei, "Trellis-coded modulation with multidimensional constellations," *IEEE Trans. Inf. Theory*, vol. 33, no. 4, pp. 483–501, Jul. 1987.

[121] S. S. Pietrobon and D. J. Costello, Jr., "Trellis coding with multidimensional QAM signal sets," *IEEE Trans. Inf. Theory*, vol. 39, no. 2, pp. 325–336, Mar. 1993.

[122] S. Benedetto, G. Olmo, and P. Poggiolini, "Trellis coded polarization shift keying modulation for digital optical communications," *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 1591–1602, Feb./Mar./Apr. 1995.

[123] I. Djordjevic and B. Vasic, "Nonbinary LDPC codes for optical communication systems," *IEEE Photon. Technol. Lett.*, vol. 17, no. 10, pp. 2224–2226, Oct. 2005.

[124] D. Declercq and M. Fossorier, "Decoding algorithms for nonbinary LDPC codes over GF(q)," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 633–643, Apr. 2007.

[125] B. Goebel, R.-J. Essiambre, G. Kramer, P. J. Winzer, and N. Hanik, "Calculation of mutual information for partially coherent Gaussian channels with applications to fiber optics," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5720–5736, Sep. 2011.

[126] L. Beygi, E. Agrell, M. Karlsson, and B. Makki, "A novel rate allocation method for multilevel coded modulation," *in Proc. IEEE Int. Symp. on Inform. Theory*, pp. 1983–1987, Jun. 2010.

[127] A. Bennatan and D. Burshtein, "Design and analysis of nonbinary LDPC codes for arbitrary discrete-memoryless channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 549–583, Feb. 2006.

[128] M. Arabaci, I. B. Djordjevic, L. Xu, and T. Wang, "Nonbinary LDPC-Coded modulation for high-speed optical fiber communication without bandwidth expansion," *IEEE Photonics Journal*, vol. 4, no. 3, pp. 728–734, Jun. 2012.

[129] G.-H. Gho and J. M. Kahn, "Rate-adaptive modulation and coding for optical fiber transmission systems," *J. Lightw. Technol.*, vol. 30, no. 12, pp. 1812–1818, Jun. 2012.

[130] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, 3rd ed. Springer-Verlag, 1998.

[131] M. Karlsson and E. Agrell, "Spectrally efficient four-dimensional modulation," *in Proc. Optic. Fiber Commun. Conf.*, Mar. 2012, OTu2C.1.

[132] R. Laroia, N. Farvardin, and S. A. Tretter, "On optimal shaping of multidimensional constellations," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1044–1056, Jul. 1994.

[133] G. D. Forney, Jr., "Trellis shaping," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 281–300, Mar. 1992.

[134] A. K. Khandani and P. Kabal, "Shaping multidimensional signal spaces. I. Optimum shaping, shell mapping," *IEEE Trans. Inf. Theory*, vol. 39, no. 6, pp. 1799–1808, Nov. 1993.

[135] ——, "Shaping multidimensional signal spaces. II. Shell-addressed constellations," *IEEE Trans. Inf. Theory*, vol. 39, no. 6, pp. 1809–1819, Nov. 1993.

[136] F. R. Kschischang and S. Pasupathy, "Optimal shaping properties of the truncated polydisc," *IEEE Trans. Inf. Theory*, vol. 40, no. 3, pp. 892–903, May 1994.

[137] I. B. Djordjevic, H. G. Batshon, L. Xu, , and T. Wang, "Coded polarization-multiplexed iterative polar modulation (PM-IPM) for beyond 400 Gb/s serial optical transmission," *in Proc. Optic. Fiber Commun. Conf.*, Mar. 2010, OMK2.

[138] J. Justesen, "On the complexity of decoding Reed–Solomon codes," *IEEE Trans. Inf. Theory*, vol. 22, no. 2, pp. 237–238, Mar. 1976.