CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Chalmers Publication Library

Thesis for the Degree of Doctor of Philosophy

# Towards the Limits of Nonlinearity Compensation for Fiber-Optic Channels

Naga V. Irukulapati


CHALMERS

Communication Systems Group
Department of Signals and Systems
Chalmers University of Technology
Göteborg, Sweden, 2016

Towards the Limits of Nonlinearity Compensation for Fiber-Optic Channels

Naga V. Irukulapati

©Naga V. Irukulapati, 2016
except where otherwise stated.
All rights reserved.

ISBN 978-91-7597-446-0
Doktorsavhandlingar vid Chalmers tekniska högskola
Series No. 4127
ISSN 0346-718X

Communication Systems Group
Department of Signals and Systems
Chalmers University of Technology
SE–412 96 Göteborg
Sweden
Telephone: +46–(0)31–772 10 00

Printed by Chalmers Reproservice
Göteborg, Sweden, 2016

To my family and friends

"When you really want something to happen, the whole world conspires to help you achieve it."

-Paulo Coelho

"Be the change that you wish to see in the world."

-Mahatma Gandhi

Towards the Limits of Nonlinearity Compensation
for Fiber-Optic Channels

# Abstract

The performance of long-haul coherent optical systems is fundamentally limited by fiber nonlinearity and its interplay with chromatic dispersion and noise. Due to nonlinearity, the signal propagating through the fiber interacts with itself and with the noise generated from the inline amplifiers. This process results in nonlinear inter-symbol interference (NISI) and nonlinear signal–noise interaction (NSNI). The state-of-the-art algorithm for combating these impairments is digital backpropagation (DBP) and is typically used as a benchmark against other detectors. However, DBP compensates only for NISI, while studies have revealed that NSNI limits the capacity of the coherent optical communications. The goal of the thesis is to use a methodical approach to develop a near-optimal nonlinearity compensation algorithm that also accounts for NSNI. This allows us to identify the fundamental performance limits of the fiber-optic channel.

Starting from the maximum a posteriori principle, we develop an algorithm called stochastic digital backpropagation (SDBP) using the framework of factor graphs. In contrast to DBP, SDBP accounts not only for NISI but also for NSNI. To account for the effects of pulse shaping, we propose three variants of SDBP in this thesis. In the first variant, the output of SDBP is processed using a matched filter (MF) followed by sampling, and decisions are taken on a symbol-by-symbol (SBS) basis. In terms of symbol error rate (SER), SBS-SDBP has better performance than DBP. However, residual memory remains after performing the MF as the MF operation need not be the optimal processing for the fiber-optic channel. This is accounted for in the second variant of SDBP, where the Viterbi algorithm is used after the MF to compensate for the residual memory. The SER of this variant is further improved compared to SBS-SDBP. In the third variant of SDBP, we use Gaussian message passing to account for the effect of pulse shaping, instead of using the MF. The SER of this third variant of SDBP is better than SBS-SDBP.

For estimating the achievable throughput in a typical transmission system, mutual information is a better metric than error rate for soft-decision coded optical systems. We show that SDBP can be used as a tool to compute lower bounds on the mutual information, which are tighter than those obtained using DBP.

**Keywords:** Digital backpropagation, fiber-optic channel, factor graphs, mutual information, near-MAP detector, nonlinear compensation, performance limits, stochastic digital backpropagation, Viterbi algorithm.

# Publications

This thesis includes the following papers:

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

Part of this paper is also presented in
N. V. Irukulapati, H. Wymeersch, P. Johannisson, and E. Agrell, "Extending digital backpropagation to account for noise", in *Proc. of European Conference and Exhibition on Optical Communication (ECOC)*, 2013.

[B] N. V. Irukulapati, D. Marsella, P. Johannisson, M. Secondini, H. Wymeersch, E. Agrell, and E. Forestieri, "On maximum likelihood sequence detection for single-channel coherent optical communications", in *Proc. of European Conference and Exhibition on Optical Communication (ECOC)*, 2014.

[C] N. V. Irukulapati, D. Marsella, P. Johannisson, E. Agrell, M. Secondini, and H. Wymeersch, "Stochastic digital backpropagation with residual memory compensation", (invited paper) *IEEE/OSA Journal of Lightwave Technology*, vol. 34, no. 2, pp. 566–572, Jan. 2016.

[D] H. Wymeersch, N. V. Irukulapati, I. Sackey, P. Johannisson, and E. Agrell, "Backward particle message passing", (invited paper) in *Proc, of International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2015.

[E] N. V. Irukulapati, M. Secondini, E. Agrell, P. Johannisson, and H. Wymeersch, "Tighter lower bounds on mutual information for single-channel fiber-optic communications", submitted to *IEEE/OSA Journal of Lightwave Technology*, Jun. 2016.

Other contributions by the author (not included in this thesis):

[F] H. Wymeersch, N. V. Irukulapati, D. Marsella, P. Johannisson, E. Agrell, M. Secondini, "On the Use of factor graphs in optical communications", (invited paper) in *Proc. Optical Fiber Communications Conference (OFC)*, 2015.

[G] L. Beygi, N. V. Irukulapati, E. Agrell, P. Johannisson, M. Karlsson, H. Wymeersch, P. Serena, and A. Bononi, "On nonlinearly-induced noise in single-channel optical links with digital backpropagation", *Optics Express*, vol. 21, no. 22, pp. 26376–26386, Oct. 2013.

[H] T. Ahmad, Y. Ai, P. Muralidharan, N. V. Irukulapati, P. Johannisson, H. Wymeersch, E. Agrell, P. Larsson-Edefors, and M. Karlsson, "Methodology for power-aware coherent receiver design", in *Signal Processing in Photonics Communications (SPPCom) Topical Meeting*, 2013.

[I] D. Sen, H. Wymeersch, N. V. Irukulapati, E. Agrell, P. Johannisson, M. Karlsson, P. A. Andrekson, "MCRB for timing and phase offset for low-rate optical communication with self-phase modulation", in *IEEE Communication Letters*, vol. 17, no. 5, pp. 1004–1007, May 2013.

# Contents

**Abstract** iii

**Publications** v

**Acknowledgment** xi

**Acronyms** xiii

**I Overview** xv

**1 Introduction** 1
- 1.1 The Role of Fibers in Communication Networks . . . 1
- 1.2 Fiber Impairments and their Compensation . . . 2
- 1.3 Goal of the Thesis . . . 2
- 1.4 Organization of the Thesis . . . 3
- 1.5 Notation . . . 3

**2 Fiber-Optic Communication Systems** 5
- 2.1 Signal Propagation in the Fiber . . . 6
  - 2.1.1 Chromatic Dispersion . . . 6
  - 2.1.2 Nonlinear Kerr Effect . . . 9
  - 2.1.3 Power Losses and Noise . . . 9
  - 2.1.4 Nonlinear Signal–Noise Interactions . . . 10
  - 2.1.5 Other Impairments . . . 11
- 2.2 System Model . . . 11
  - 2.2.1 Channel Models . . . 12
- 2.3 Numerical Methods for Signal Propagation . . . 13
- 2.4 Summary . . . 14

**3 Bayesian Inference using Factor Graphs** 15
- 3.1 Bayesian Inference . . . 15
- 3.2 Factor Graphs . . . 16
- 3.3 Message Passing . . . 17

3.3.1 Message Computation . . . 18
3.3.2 Interpretation of Messages . . . 19
3.4 Particle Representation . . . 20
3.5 Application of Factor Graphs for the Fiber-Optic Channel . . . 20
3.5.1 Particle Representation for the Building Blocks of the Fiber-Optic Channel . . . 21
3.5.2 A Particle-Based Digital Communication Receiver . . . 22
3.6 Summary . . . 24

**4 Nonlinear Compensation Techniques** **25**
4.1 Optical Nonlinearity Compensation . . . 25
4.2 Digital Nonlinearity Compensation . . . 26
4.2.1 Digital Backpropagation . . . 26
4.2.2 Maximum Likelihood Sequence Detectors . . . 28
4.3 Stochastic Digital Backpropagation . . . 28
4.3.1 SBS-SDBP . . . 29
4.3.2 VA-SDBP . . . 29
4.3.3 GMP-SDBP . . . 29
4.4 Performance Metrics . . . 30
4.4.1 Error Probability . . . 30
4.4.2 Mutual Information . . . 31
4.5 Comparison of SDBP and DBP . . . 31
4.6 Summary . . . 32

**5 Contributions and Future Work** **33**
5.1 Paper A . . . 33
5.2 Paper B . . . 34
5.3 Paper C . . . 35
5.4 Paper D . . . 35
5.5 Paper E . . . 35
5.6 Future Work . . . 36

**References** **37**

## II Papers 49

**A Stochastic Digital Backpropagation** **A1**
1 Introduction . . . A2
2 System Model . . . A4
2.1 High-Level Description . . . A4
2.2 Signal Propagation in Optical Fibers . . . A5
3 MAP Detection . . . A6
3.1 Factorization of the Joint Distribution . . . A7
3.2 Message Passing: Theory . . . A8
3.3 Message Passing in SDBP . . . A9

4 Near-MAP Detector Implementation . . . A11
4.1 Particle Representations . . . A11
4.2 Symbol-by-Symbol Detector . . . A14
4.3 Remarks . . . A14
5 SDBP for a Simplified Model . . . A15
6 Numerical Simulations and Discussions . . . A16
6.1 SER and Reach Analysis . . . A17
6.2 Estimated Distributions . . . A20
6.3 Influence of PMD on DBP and SDBP . . . A20
6.4 Complexity Analysis of SDBP . . . A22
7 Conclusions . . . A23

**B On Maximum Likelihood Sequence Detectors for Single-channel Coherent Optical Communications B1**
1 Introduction . . . B2
2 System Model . . . B2
3 Detectors . . . B3
4 VA with Cartesian and Polar Gaussian Metric . . . B4
5 Stochastic Digital Backpropagation . . . B4
6 Numerical Simulations . . . B5
7 Conclusions . . . B7

**C Stochastic Digital Backpropagation with Residual Memory Compensation C1**
1 Introduction . . . C2
2 Factor Graphs for Receiver Design . . . C3
3 System Model . . . C4
4 SDBP and Proposed Approaches . . . C5
4.1 Stochastic Digital Backpropagation . . . C5
4.2 Viterbi Algorithm-Stochastic Digital Backpropagation . . . C7
4.3 Decision Directed-Stochastic Digital Backpropagation . . . C8
5 Numerical Simulations and Discussion . . . C9
5.1 Simulation Setup and Performance Metrics . . . C9
5.2 Results and Discussion . . . C10
6 Conclusions . . . C13

**D Backward Particle Message Passing D1**
1 Introduction . . . D2
2 Problem Formulation . . . D3
3 Message Computation . . . D4
3.1 3-way equality . . . D4
3.2 Addition of noise . . . D5
3.3 Summation . . . D5
3.4 Forward message for (non)linear functions . . . D5
3.5 Backward message for (non)linear functions . . . D6
3.6 Backward message for linear functions . . . D6

4 Numerical Examples . . . . . D9
4.1 Illustrating example . . . . . D9
4.2 Coherent optical receiver . . . . . D9
5 Conclusions . . . . . D12

**E Tighter Lower Bounds on Mutual Information for Fiber-Optic Channels E1**
1 Introduction . . . . . E2
2 Mutual Information . . . . . E3
3 Lower bounds on mutual information . . . . . E4
3.1 Lower Bounds using Auxiliary Forward Channel $q(\mathbf{y}|\mathbf{x})$ . . E4
3.2 Lower Bounds using Auxiliary Backward Channel $r(\mathbf{x}|\mathbf{y})$ . E5
3.3 Monte Carlo Estimation of AIR . . . . . E6
4 Computation of AIR for the FOC . . . . . E7
4.1 Computation of AIR using DBP . . . . . E7
4.2 Computation of AIR using SDBP . . . . . E8
5 Numerical Results and Discussion . . . . . E10
5.1 System Parameters . . . . . E10
5.2 Results . . . . . E11
5.3 Discussion . . . . . E13
6 Conclusion . . . . . E14
7 Acknowledgments . . . . . E14

# Acknowledgments

First and foremost, I want to thank Prof. Henk Wymeersch along with Prof. Erik Agrell, Prof. Magnus Karlsson, Prof. Debarati Sen, and others who were involved during my PhD interview process and for accepting me as a PhD student. I would like to express my sincere gratitude to my three supervisors, Henk Wymeersch, Pontus Johannisson, and Erik Agrell for the continuous support, for your patience, motivation, enthusiasm, and immense knowledge. Your guidance helped me at all the times of research and during writing of this thesis.

I owe a debt of gratitude to my main supervisor, Henk, who always encouraged and supported me throughout my PhD. Besides acquiring the knowledge of Bayesian inference, I learnt a lot of skills from you: getting things done without procrastinating (although I did not fully succeed at it yet), presenting to a wide range of audience, focusing and multitasking, giving feedbacks for the papers very quickly, prioritizing and many more. Thank you Pontus for explaining even basic questions related to optical communications with patience. You are a person with background in optics and also have good knowledge in DSP, which is a very big plus during my research. My gratitude is also extended to Erik for his careful attention to detail, for giving feedback that helped me grow professionally and personally. During the last year, I enjoyed working with you in information theory. I also learnt how to formulate problems mathematically and then use the right tools to solve these problems.

I would like to acknowledge Prof. Erik Ström and all current and former members of ComSys for creating such a friendly and inspiring atmosphere in the group. A special thanks to my former and current office mates: Kasra, Wanlu, Wei, Chao for creating a nice atmosphere at work. A special mention goes to Lotfollah and Debarati, who were there for any technical questions and for the helpful discussions during the beginning of the PhD. I would also like to thank Christian Häger, who is open to discuss any basic questions. FORCE group has nice composition of people working in theoretical and practical aspects of fiber-optics, which benefitted me a lot during my PhD. During the last year of my PhD, I enjoyed and benefitted from discussing with Rahul, Kamran, Amina, Cristian, Alireza, Srikar, Tobias Fehenberger, and Alex Alvarado on many aspects of fiber-optics and information theory. Many thanks to the administrative staff at S2, Natasha, Lars, Malin, and specially Agneta. It was great to work with you in organizing workshops, social events, and conferences.

I also want to take this opportunity to thank all my collaborators during these 5 years of PhD. It was a great learning experience working with Prof. Marco

Secondini and Dr. Domenico Marsella. Thanks to all the friends at TeCIP institute at Pisa for a wonderful time during my stay there. My gratitude also goes to Dr. Antonio Napoli, Dr. Bernhard Spinnler, Dr. Danish Rafique, and all colleagues at Coriant for a splendid time. I would also like to thank the group at Parma: Prof. Alberto Bononi, Prof. Paolo Serena, Prof. Giulio Colavolpe, and Dr. Tommasso Foggi for discussions on various topics related to the thesis. I would extend my gratitude to Prof. Hans-Andrea Loeliger and Prof. Frank Kschischang for sharing their expertise in factor graphs. I take this opportunity to thank the students whom I have worked as an advisor during their master thesis projects: Isaac, Yun, Pavithra, and Tauseef. I would like to give a special mention also to Prof. Vijay Kumar, my bachelors thesis guide, who constantly encouraged me to pursue further studies after bachelors at DA-IICT.

During the past years, I had some of the most memorable experiences in my life (paragliding, trekking, skiing, dancing, learning to swim, water rafting, and many more) and I would like to thank all my friends in Gothenburg for such nice experiences. S2 ski trips are very memorable experiences and thanks to the participants for an amazing time: Lennart, Astrid, Keerthi, Abu, Rikard, Tommy, Tilak, Abbas, Sadegh, Pegah, Arne, Ingemar, Lars, Malin, and Katharine. I would like to thank my friends: Ajay, Seshu, Murali, Koushik, Swathi, Srikar, Abu, Sathya, Haritha, Keerthi, Deepthi, Shreyansh for a great time off-work and for gatherings that we had and we shall continue it. I should also thank each and everyone of 'TeamIndia' group, where I started to learn some dance and become part of showcasing Indian culture in Gothenburg. Thank you Ananda, Annelie, Abu, Gokul, Sathya, Fereshteh, Dipti, Ninva, Amanda, Sanjna, Nitin, Frida, Sandra, Raneesh, Anamika, Jagadeep, Rameez, Sarang, Mahsa, and Jasmine.

PhD studies at ComSys, combined with my other activities off-work, gave me an opportunity to take out time for myself, which was very much needed to learn about myself, introspect, try new things professionally and personally. It is great to have friends like you at ComSys, TeamIndia group, my family and friends in Gothenburg, India and elsewhere. This journey would not have been possible without the support of my well wishers in India especially my mother Adeswari, father Nageswara Rao, sister Manasa, cousin Vijaya Bhaskar, cousin Aruna Devi, uncle Vakuleshwar Rao, and friend Sitara. You are the best and are always there for me. I always knew that you believed in me and wanted the best for me. Thank you for teaching me that my job in life was to learn constantly and be happy.

This research was supported by the Swedish Research Council (VR) under grants no. 2010-4236 and no. 2013-5642. Research visits to ETH Zurich and TeCIP Pisa were supported by Ericsson research foundation. The simulations were performed in part on resources provided by the Swedish National Infrastructure for Computing (SNIC) at C3SE.

Naga V. Irukulapati
Gothenburg, August, 2016.

# Acronyms

| | |
|---|---|
| AIR | achievable information rate |
| ASE | amplified spontaneous emission |
| AWGN | additive white Gaussian noise |
| BER | bit error rate |
| FBG | fiber Bragg gratings |
| CG | Cartesian Gaussian |
| CD | chromatic dispersion |
| DBP | digital backpropagation |
| DCF | dispersion-compensating fiber |
| DCM | dispersion-compensating module |
| DGD | differential group delay |
| DM | dispersion managed |
| DSP | digital signal processing |
| EDC | electronic dispersion compensation |
| EDFA | erbium-doped fiber amplifier |
| FEC | forward error correction |
| FG | factor graphs |
| GMP | Gaussian message passing |
| ISI | inter-symbol interference |
| MAP | maximum a posteriori |
| MI | mutual information |
| MF | matched filter |
| MLSD | maximum likelihood sequence detection |
| NDM | non-dispersion managed |
| NISI | nonlinear ISI |
| NLPN | nonlinear phase noise |
| NLSE | nonlinear Schrödinger equation |
| NSNI | nonlinear signal–noise interactions |
| OBC | optical backpropagataion |
| OPC | optical phase conjugation |
| PCTW | phase-conjugated twin waves |
| PSA | phase-sensitive amplifiers |
| PDF | probability density function |

| | |
|---|---|
| PG | polar Gaussian |
| PMD | polarization-mode dispersion |
| PMF | probability mass function |
| PR | particle representation |
| QAM | quadrature amplitude modulation |
| QPSK | quadrature phase-shift keying |
| SBS | symbol by symbol |
| SDBP | stochastic digital backpropagation |
| SER | symbol error rate |
| SMF | single-mode fiber |
| SNR | signal-to-noise ratio |
| SPA | sum–product algorithm |
| SPM | self-phase modulation |
| SSFM | split-step Fourier method |
| VA | Viterbi algorithm |
| WDM | wavelength-division multiplexing |

# Part I

# Overview

# Chapter 1

# Introduction

## 1.1 The Role of Fibers in Communication Networks

Today's information society relies to a large extent on solutions based on broadband communications, with applications such as mobile voice and data services, high-speed internet access, and multimedia broadcast systems [1]. Each of these applications brings its own set of challenges, which can be addressed using electronic, radio-frequency, or optical communication systems. Among the different communication technologies, optical communications generally has the edge over baseband electronic or radio-frequency transmission systems whenever high aggregate bit rates and/or long transmission distances are involved. Both the advantages are deeply rooted in physics. First, the high optical carrier frequencies allow for high-capacity systems at small relative bandwidths. For example, a mere 2.5% bandwidth at a carrier frequency of 193 THz (1.55 $\mu$m wavelength) opens up a 5-THz chunk of contiguous communication bandwidth. Second, fibers exhibit losses of around 0.2 dB/km, which is very little compared to the losses in typical coaxial cables that support a bandwidth of 1 GHz. The latter generally exhibits losses of 2 to 3 orders of magnitude higher than that for the fiber. Other advantages of communicating information through fiber-optic channels are the unregulated spectrum in the optical regime and the absence of electromagnetic interference.

Data rates of optical communication links have grown exponentially since their introduction in the late 1970s. Until very recently, optical communications used binary modulation formats such as on-off keying at the transmitter and power detection at the receiver using a fixed-threshold detection. This receiver does not require any complex hardware; however, the data rates that can be achieved are limited. To increase the data rates of the fiber-optic communications, modulation formats with higher spectral efficiency are of great interest which require coherent detection at the receiver. However these higher-order modulation formats are less tolerant towards some of the channel impairments.

## 1.2 Fiber Impairments and their Compensation

To understand fiber impairments, recall that an optical fiber is a waveguide consisting of a cylindrical core surrounded by a cladding. The refractive index of the core is higher than the cladding so that the light is guided in the optical fiber. The dependence of the refractive index on the frequency and the power gives rise to two dominant impairments in fiber-optic channels, namely chromatic dispersion and the nonlinear Kerr effect. Due to chromatic dispersion, the signal that is sent at the input of the fiber is broadened in time, causing inter-symbol interference. Due to the nonlinearities in the fiber, the signal phase is changed in proportion to the signal power and this power-dependent phase shift causes spectral broadening. In addition to these two impairments, the third dominant impairment in fiber-optic channels is the noise added by the inline amplifiers. When the signal propagates in the fiber, the signal power reduces exponentially due to the fiber loss. Therefore, amplification of the signal is needed, especially, for long-haul communications (distances spanning from a hundred to a few thousands of kilometers). These three phenomena are distinct, occur simultaneously, are distributed along the propagation path, and importantly influence each other, leading to deterministic and stochastic impairments [2]. For low signal powers, the transmission performance is limited by the noise, similar to linear channels, the capacity of the system can be enhanced by increasing signal power. However, for high signal powers, nonlinear effects in the fiber dominate and the transmission performance is not necessarily enhanced by simply increasing signal power. Digital backpropagation (DBP) is often considered to be a universal technique for jointly compensating the linear and the nonlinear impairments [3]. However, DBP compensates only for the deterministic nonlinear inter-symbol interference (NISI) and does not account for the stochastic nonlinear signal–noise interactions (NSNI). In fact, for single-channel fiber-optic communications, NSNI is widely believed to be limiting the capacity [2, 4–6].

## 1.3 Goal of the Thesis

This thesis is focused on finding a near-optimal receiver for single-channel fiber-optic communication systems. Specifically, this thesis aims to answer the following questions:

- Can we derive algorithms that account also for NSNI, besides NISI, in a methodical approach?
- How does such an algorithm perform with respect to metrics such as symbol error rate (SER) and mutual information (MI)?

In paper A, starting from the maximum a posteriori principle, a detector named stochastic digital backpropagation (SDBP) is developed using the framework of factor graphs and message passing. Just as in DBP, the output of SDBP is processed using a matched filter (MF) and the decisions are taken on a symbol-by-symbol (SBS) basis. Hence, we call this algorithm SBS-SDBP. As an MF need not

be an optimal choice for nonlinear channels [7], there maybe some residual memory left after SBS-SDBP. To quantify the loss resulting from SBS decisions after MF in SBS-SDBP, a comparison is made with a detector that uses the Viterbi algorithm (VA) after DBP [8], and the results are presented in paper B. In paper C, the VA is used after MF and sampling at the output of SDBP, and we call this algorithm VA-SDBP. VA-SDBP has significantly improved performance compared to SBS-SDBP. Since the MF operation is suboptimal, information may have been lost and using the combination of the VA with the MF is also suboptimal. Hence, we approach the problem of accounting for the pulse shaping by using the factor graphs and sum-product algorithm framework itself, from which SDBP was derived. This resulted in paper D, where instead of an MF, a Gaussian message passing (GMP) is performed on the output of SDBP, and we call this algorithm GMP-SDBP. In papers A–D, SER is used as a performance metric and lower SER than DBP was observed. In paper E, mutual information is used as a performance metric and it is observed that using SDBP, tighter lower bounds can be obtained than existing lower bounds on mutual information computed using DBP.

## 1.4 Organization of the Thesis

A PhD thesis in Sweden can be written either as a monograph or as a collection of papers and the latter is followed for this thesis. The thesis in this type is divided into two parts: the first part provides the readers the background necessary for understanding the second part, the research papers. The intended audience of the thesis are graduate students and researchers currently working or planning to work in optical communications, who have some background in digital communications.

In the first part of the thesis, introduction material needed to understand the concepts behind the papers of part II are presented. Specifically, in Chapter 2, we introduce the fiber-optic channel, starting with the signal propagation model, impairments existing in this channel, and the numerical approach used to simulate the signal propagation in a fiber. Bayesian inference using factor graphs is introduced in Chapter 3. In Chapter 4, nonlinear compensation techniques is described mainly emphasizing DBP, followed by some basic principles needed to understand the detectors: SBS-SDBP, VA-SDBP, and GMP-SDBP. Performance metrics used in the thesis, SER and MI, are also explained in Chapter 4. In Chapter 5, a summary of the contributions of the papers in part II of the thesis are highlighted with possible future directions of the current research.

## 1.5 Notation

The following notation is used in the introductory part of the thesis and in the appended papers.

- Lower case bold letters (e.g., $\mathbf{x}$) are used to denote vectors, including sequences of symbols and vector representations of continuous-time signals (e.g., through oversampling). With a slight abuse of notation, $\mathbf{r}(t)$ is used

to represent a vector of dual-polarization single-wavelength continuous-time signal and $\mathbf{r}$ is used to represent samples of $\mathbf{r}(t)$, where the data for each polarization is combined. In paper A, lower case bold letters (e.g., $\mathbf{x}$) are used for vector representation of the continuous-time signals and underlined lower case bold letters (e.g., $\underline{\mathbf{y}}$) for a vector of discrete-time symbols. However, in papers B–D, we did not differentiate between these two classes of signals, and have used lower case bold letters (e.g., $\mathbf{x}$) to represent both these classes.

- Hermitian conjugate of the vector $\mathbf{v}$ is denoted by $\mathbf{v}^{\mathrm{H}}$.
- The transpose of the vector $\mathbf{v}$ is denoted by $\mathbf{v}^{\mathrm{T}}$.
- The cardinality of a set $\mathcal{A}$ is denoted by $|\mathcal{A}|$.
- Random variables are denoted by capital letters (for e.g., $X$) and their realizations by lower-case letters (for e.g., $x$).
- A multivariate Gaussian probability density function (PDF) of a variable $\mathbf{r}$ with mean $\mathbf{z}$ and covariance matrix $\Sigma$ is denoted by $\mathcal{N}(\mathbf{r}; \mathbf{z}, \Sigma)$.
- The indicator function with proposition $\mathcal{P}$ is given by $\mathbb{1}(\mathcal{P})$ and the Dirac delta function is denoted by $\delta(\cdot)$.
- Expectation is denoted by $\mathbb{E}\{.\}$
- Messages in sum-product algorithm associated with edge/variable $\mathbf{Q}$, evaluated in $\mathbf{q}$ is denoted by $\mu_{\mathbf{Q}}(\mathbf{q})$ (or $\mu(\mathbf{q})$ when the variable is clear from the context). The direction of the message will be represented by arrows: $\overleftarrow{\mu}_{\mathbf{Q}}(\mathbf{q})$ and $\overrightarrow{\mu}_{\mathbf{Q}}(\mathbf{q})$.

# Chapter 2

# Fiber-Optic Communication Systems

An optical fiber is a waveguide consisting of a cylindrical core surrounded by a cladding. The refractive index of the core is higher than the cladding so that the light is guided in the optical fiber. A waveguide mode is a configuration of the electric field that propagates without changing its spatial distribution, apart from an amplitude change and a phase shift. A single-mode fiber (SMF), commonly used for transmission in long-haul communications, supports only one propagating mode. In fiber-optic channels, the physical dimensions that can be used for modulation and multiplexing are time, quadrature (amplitude/phase), frequency, polarization, and space (for example by using multiple modes in a multi-mode fiber) [9]. In polarization-multiplexed signals, the spectral efficiency is increased by transmitting two different signals at the same wavelength but in two orthogonal polarizations. For example, for a symbol rate of 28 Gbaud using polarization-multiplexed 16-QAM modulation format, a raw data rate of $28 \times 4 \times 2 = 224$ Gb/s per wavelength can be achieved.

A single-mode optical fiber is an exceptionally transparent medium. Unlike typical coaxial cables, where losses are on the order of several tens of dB/km for a bandwidth of around 1 GHz, a modern telecom fiber features attenuation coefficients below 0.2 dB/km across a bandwidth of many THz. Nevertheless, as the signal propagates in the fiber, the signal power is reduced due to the fiber loss and for long-haul communications, this attenuation calls for amplification of the signal.

This chapter is organized as follows. In Sec. 2.1, starting with the equations governing signal propagation in a fiber, impairments arising in the fiber are described. The system model will be described along with the assumptions in Sec. 2.2, detailing different blocks of the model. Existing analytical channel models will be presented in Sec. 2.2, followed by numerical methods for describing the signal propagation in the fiber in Sec. 2.3.

## 2.1 Signal Propagation in the Fiber

The propagation of light in an optical fiber is modeled using the Manakov equation with loss included [10]

$$\frac{\partial \mathbf{a}}{\partial z} = i\gamma \|\mathbf{a}\|^2 \mathbf{a} - i\frac{\beta_2}{2}\frac{\partial^2 \mathbf{a}}{\partial t^2} - \frac{\alpha}{2}\mathbf{a}, \tag{2.1}$$

where $\mathbf{a} \triangleq [\mathbf{a}_x \ \mathbf{a}_y]^T$ is the complex envelope of the two polarization components of the optical field, $\gamma$ is the nonlinear coefficient, $\|\mathbf{a}\|^2 = \mathbf{a}^H\mathbf{a}$ represents the optical power, where $H$ is the hermitian conjugate, $\beta_2$ is the group velocity dispersion coefficient, $\alpha$ is the power attenuation factor, $z$ is the distance of propagation, and $t$ is the time coordinate in a reference frame moving with the signal group velocity. The nonlinear Schrödinger equation (NLSE) is the corresponding modeling equation for the single-polarization case.

### 2.1.1 Chromatic Dispersion

If the group velocity[1] is different for different frequency components of the wave, the medium is said to be dispersive and the effect is known as group velocity dispersion or chromatic dispersion (CD). The CD broadens the pulse in the time domain leading to inter-symbol interference (ISI) as depicted in Fig. 2.1. An important parameter is the dispersion length, $L_\mathrm{D}$, which is the propagation distance after which the dispersive effects become important and is given by $L_\mathrm{D} = 1/(|\beta_2|W^2)$ [11, p. 55], where $W$ is the bandwidth of the transmitted signal. The dispersion parameter, $D$, is related to $\beta_2$ as $D = -2\pi c\beta_2/\lambda^2$, where $c$ is the speed of light and $\lambda$ is the wavelength.

When $\gamma = 0$ and $\alpha = 0$ in (2.1), a closed-form solution is given by $\tilde{\mathbf{a}}(z,\omega) = \tilde{\mathbf{a}}(0,\omega)\exp(i\beta_2\omega^2 z/2)$, where $\tilde{\mathbf{a}}(z,\omega)$ is the spectrum of $\mathbf{a}(z,t)$. Hence CD can be modeled as an all-pass filter. CD does not change the amplitude of the spectrum but causes a frequency-dependent phase shift in the frequency domain [10, 12] as shown in Fig. 2.2.

Dispersion can be compensated for in the optical domain either using dispersion-compensating fibers (DCFs) or fiber Bragg gratings (FBG) [13]. The DCFs have the opposite sign of $\beta_2$ compared to the SMF and also have higher nonlinear coefficient than the SMF [10, ch. 9]. An FBG has no nonlinearities and has low insertion loss. When the dispersion is compensated optically within the fiber-optic channel, the system is known as a dispersion-managed (DM) link. Otherwise the link is said to be non-dispersion-managed (NDM). In NDM links, CD is typically compensated through digital signal processing (DSP) in the receiver using an electronic dispersion compensation (EDC) block. This EDC is a filter with a frequency response equal to $\exp(-i\beta_2\omega^2 z/2)$.

[1] The group velocity is the velocity with which the complex envelope of the wave propagates through the fiber.

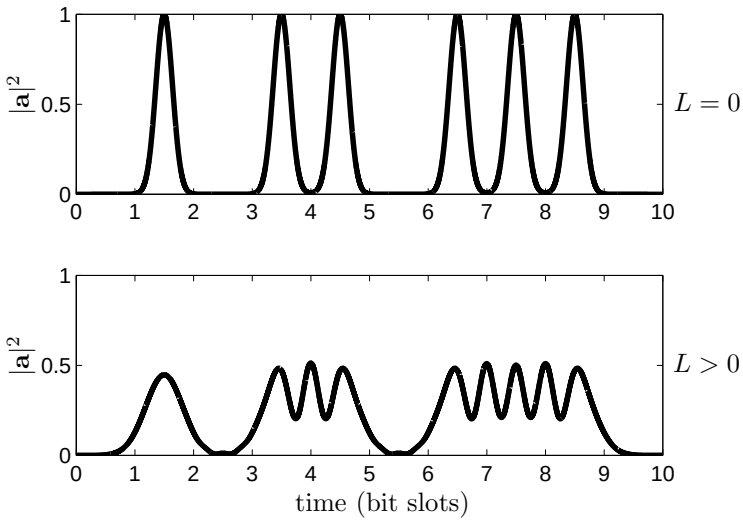



**Figure 2.1:** Effect of the CD in the time domain. In the top (resp., bottom) figure, a pulse at the input (resp., output) of a fiber can be seen. One can see that pulses broaden in the time domain and start to interfere.

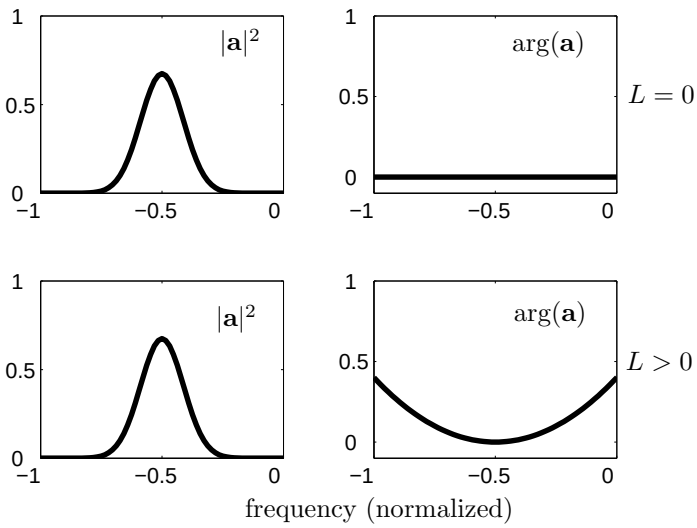



**Figure 2.2:** Spectrum of a single pulse affected by the CD. In the top (resp., bottom) row of the figures, a pulse at the input (resp., output) of a fiber is shown. One can see that the amplitude is not changed and only a quadratic phase modulation occurs. In the time domain, this corresponds to a broadening of the pulses, as in Fig. 2.1.

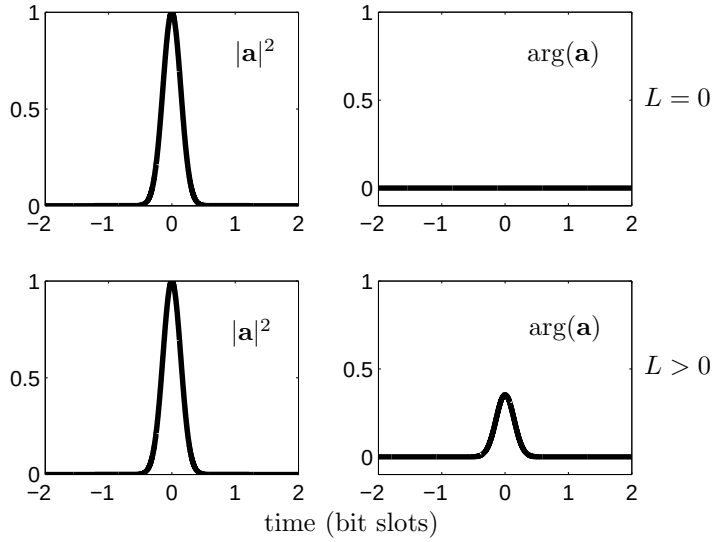



**Figure 2.3:** Effect of the Kerr nonlinearity in the time domain. In the top (resp., bottom) row of the figure, a pulse at the input (resp., output) of a fiber can be seen. The amplitude of the pulse is not changed but a phase shift is introduced.

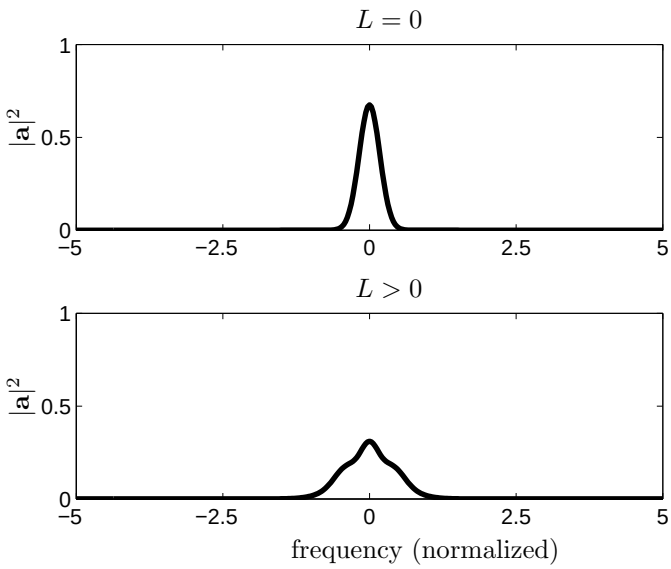



**Figure 2.4:** Spectrum of a single pulse affected by the Kerr nonlinearity. In the top (resp., bottom) figure, a pulse at the input (resp., output) of a fiber can be seen. Due to the phase shift in the time domain (Fig. 2.3), the spectrum is broadened.

### 2.1.2 Nonlinear Kerr Effect

The term with the nonlinear (NL) parameter $\gamma$ in the Manakov equation (2.1) represents the effect due to Kerr nonlinearity [14]. This arises due to the power-dependent refractive index of the fiber. The solution of the Manakov equation (2.1) setting $\alpha = 0$ and $\beta_2 = 0$ is

$$\mathbf{a}(L,t) = \mathbf{a}(0,t)\exp[i\gamma\|\mathbf{a}(0,t)\|^2 L_{\text{eff}}], \tag{2.2}$$

where the signal phase is changed in proportion to the signal power and this effect is called self-phase modulation (SPM). Here $L$ is the length of the fiber. The nonlinear phase shift is denoted by $\phi_{\text{NL}} \triangleq \gamma\|\mathbf{a}(0,t)\|^2 L_{\text{eff}}$. The effective fiber length, $L_{\text{eff}} = [1-\exp(-\alpha L)]/\alpha$, is an indication of the fiber length along which the nonlinearities are effective. The amplitude of the time-domain signal is not changed but a power-dependent phase shift is introduced due to the SPM as shown in Fig. 2.3. As a result in the frequency domain, the spectrum is broadened as depicted in Fig. 2.4.

#### Numerical Example

Similar to the dispersion length, the nonlinear length is defined as $L_{\text{NL}} = 1/(\gamma P_0)$, where $P_0$ is the initial peak power. For a particular system, comparing $L_{\text{D}}$ and $L_{\text{NL}}$ for a fiber of length $L$ helps in determining whether the dispersion or the nonlinearity will be the dominant effect. If $L \gtrsim L_{\text{D}}$ and $L \ll L_{\text{NL}}$, then the CD dominates over the nonlinearities. As an example, consider a standard telecommunication fiber at wavelength 1550 nm, dispersion parameter $D = 17$ ps/(nm km), and bandwidth $W = 28$ Gbaud, then $L_{\text{D}} \approx 60$ km. If the nonlinear parameter $\gamma = 1.3$ 1/(W km) and $P_0 = 0$ dBm, then $L_{\text{NL}} \approx 750$ km. If the fiber length is $L = 80$ km, then CD is the dominant effect.

### 2.1.3 Power Losses and Noise

By setting $\gamma = 0$ and $\beta_2 = 0$ in the Manakov equation (2.1), a closed-form solution is given by $\mathbf{a}(z,t) = \mathbf{a}(0,t)\exp(-\alpha z/2)$. That is, when the signal propagates in the fiber, the signal power reduces exponentially due to the fiber loss. Over sufficiently long distances $z$, the signal-to-noise ratio (SNR) of the detected signal will be too low, leading to a high bit-error rate. Therefore amplification of the signal is needed for long-haul communications. Optical amplification can be done in a distributed manner using Raman amplification or in lumped components using erbium-doped fiber amplifiers (EDFAs) [15]. In this thesis, EDFAs are used for optical amplification. Unlike radio frequency amplifiers, most of the optical amplifiers exhibit constant gain across the spectrum and do not distort the optical signals. Instead, the main degrading effect of the optical amplification is the generation of amplified spontaneous emission (ASE) noise.[2] The ASE noise can be modeled as additive white Gaussian noise (AWGN) [10]. The one-sided power

[2] Light that is coupled into the erbium-doped fiber is amplified through stimulated emission: incident photons stimulate the excited ions to return to the ground state and emit a photon

spectral density per polarization is given by $S_{\mathrm{sp}}(\nu) = (G-1)n_{\mathrm{sp}}h\nu$, where $G = \exp(\alpha L)$ is the required gain needed to compensate for the attenuation in the fiber of length $L$, $\nu$ is the optical frequency with $h\nu$ being the photon energy, and $n_{\mathrm{sp}}$ is the spontaneous-emission factor.

The impact of noise can be reduced by increasing the power launched into the fiber. However, unlike in wireless communications, as the input power is increased beyond the so-called nonlinear regime, nonlinear distortions in the fiber due to Kerr effect increase, worsening the signal quality. Thus, the power regime can be divided into a linear regime, where the system is dominated by the noise, and a nonlinear regime, where system performance is limited due to nonlinearities. This leads to an optimal launch power suitable for system operation. This optimal launch power depends on a number of factors such as modulation order, total transmission distance, nonlinear parameter $\gamma$, EDFA noise figure, type of compensation algorithm used at the receiver. When the system is operating in the linear regime, the ASE field PDF can be modeled as an AWGN process at the end of the link. On the other hand, in the nonlinear regime, the received ASE field PDF changes to a bean-like non-Gaussian shape, while the ASE power spectral density changes from white to colored [8, 16, 17].

### 2.1.4 Nonlinear Signal–Noise Interactions

Once deterministic linear and nonlinear effects are accounted for, it is known that the transmission performance of a fiber-optic channel is limited by NSNI. In fact, for single-channel fiber-optic systems, NSNI is widely believed to be limiting the capacity [2, 4–6, 18–25]. The impact of NSNI has been studied extensively through numerical simulations and through experiments [2, 4, 6, 17, 19, 23, 24, 26]. One popular way to assess the impact of NSNI is to compare a system where noise is distributed with a system where an equivalent noise is loaded at the receiver as a single AWGN process [4, 6, 17]. These studies revealed that the strength of NSNI depends on the total reach, type of dispersion management, symbol rate, and different nonlinear compensation used at the receiver. In [6], it is shown that NSNI increases with spans as the level of cumulated ASE is increased, and as a result, NSNI decreases the total transmission reach. There is a significant penalty due to NSNI for DM links in comparison to NDM links [17, 18, 26]. This is due to the fact that the cumulative dispersion allows averaging of the fiber nonlinearity, and thereby reducing NSNI [27]. Regarding the impact of NSNI with respect to the nonlinear equalization, it is shown in [6] that NSNI reduces the efficiency of a DBP especially in long-haul communications. In [28], DBP is split between transmitter and receiver thereby reducing the impact of NSNI. In [24, 29, 30], a comparison of nonlinearity compensation in optical and digital domain is performed and the impact of NSNI has been studied with respect to the optimal placement of these equalizers. It is often argued that NSNI cannot be compensated for in DSP due

of identical frequency, phase and polarization. However, ions also return to their ground state spontaneously, thereby emitting photons of random phase and polarization; this spontaneous emission becomes amplified upon propagation along the fiber link, a process that is known as amplified spontaneous emission.

to the nondeterministic nature of ASE noise [18, 20, 31] and as a result, none of the DBP methods account for NSNI.

As it can be seen from the amount of recent literature on this topic, reducing the impact of NSNI is an active research area and the thesis contributes to this research by proposing a DSP based algorithm (SDBP) to compensate also for NSNI in single-channel fiber optical systems.

### 2.1.5 Other Impairments

The other stochastic effect that degrades the performance of single-channel fiber-optic system is polarization-mode dispersion (PMD). Different polarizations of light travel at slightly different speeds, leading to random spreading of the optical pulses, and this effect is known as PMD. It is worth noting that in SDBP, ASE is the only non-deterministic impairment that is currently compensated for. However, other non-deterministic impairments, in particular PMD, may deteriorate the performance of both DBP and SDBP. This deterioration becomes significant when the total differential group delay (DGD) of the system approaches the symbol period [3, 5, 32]. To quantify the performance degradation of DBP and SDBP in the presence of PMD, simulations were performed and reported in paper A. It was noted in paper A that a full integration of PMD compensation with SDBP would constitute an entire research project in itself. Instead, two simple strategies are considered in the simulations: one where no PMD compensation is performed, and one where the receiver is assumed to operate under perfect knowledge[3] of the rotation matrices and the amount of DGD. In the latter case, the PMD is compensated after (S)DBP. Besides PMD, polarization-dependent loss, third order dispersion, and laser phase noise [10] were also not considered in the thesis.

## 2.2 System Model

The system model is shown in Fig. 2.5 and consists of a data sequence of $K$ symbols, $\mathbf{s} = [\mathbf{s}[1], \mathbf{s}[2], \cdots, \mathbf{s}[K]] \in \Omega^K$, where $\Omega$ is the set of symbols in the constellation, a pulse shaper, a fiber-optic link with $N$ spans, and a receiver with a compensation algorithm followed by a decision unit. Each span[4] of the fiber-optic link consists of an SMF followed by an EDFA, and an optional dispersion compensating module (DCM) followed by an EDFA for the DM links. In this work, we have considered either DCF or FBG as DCM. At the receiver, the signal is sent into a compensating unit, where the impairments are compensated for. In this work, either DBP or SDBP is applied for compensation of impairments, which will be explained in Ch. 4. This signal is then sent to the decision unit where the

[3] Several algorithms such as the constant modulus algorithm exist to compensate for PMD.

[4] In paper A, a single-stage amplification is used, i.e., EDFA after SMF is not used and instead EDFA after DCM compensates for the losses of both SMF and DCM. In papers B-E, dual-stage amplification as in Fig. 2.5 is used. For the same total reach, dual-stage amplification offers better performance as accumulated noise is lower in comparison to single-stage amplification. This is because large amplifier spacing implies amplifiers of high gain, resulting in a prohibitive increase in the ASE noise [33].

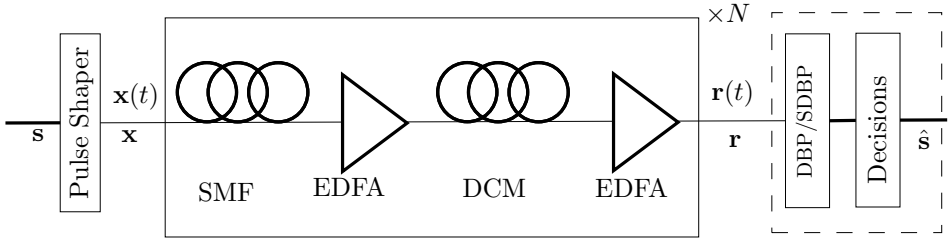



**Figure 2.5:** A fiber-optic link with $N$ spans where each span consists of an SMF followed by an EDFA, and an optional DCM and an EDFA for DM links. The transmitter consists of a pulse shaper and the receiver consists of a compensation algorithm (DBP/SDBP) and a decision block. The transmitted data is denoted by $\mathbf{s}$, decoded data by $\hat{\mathbf{s}}$, and the received signal by $\mathbf{r}$.

symbols are decoded. The problem of interest is to estimate the data sequence $\mathbf{s}$ given the received signal $\mathbf{r}(t)$. As a closed-form expression for the input-output relationship of the fiber-optic channel is not present, many approximate channel models have been proposed in literature, which will be described next. In this thesis, a "channel" refers to the fiber-optic link of Fig. 2.5 for a single-wavelength system with a dual-polarization[5] complex envelope signal.

Coherent detection with perfect timing, phase, and frequency synchronization, as well as perfect polarization tracking at the receiver, is assumed. A long-haul single-channel system is considered in the theory and the simulations. However, the theory can be extended to wavelength-division multiplexed systems.

### 2.2.1 Channel Models

As described earlier, the signal propagation in a fiber is modeled through the Manakov equation (2.1) and the NLSE for a single-polarization signal. The Manakov equation does not lend itself to an analytic solution except for some specific cases, but approximate analytical solutions exist for the fiber-optic channel of Fig. 2.5, including dispersion, nonlinearity, losses in the fiber, and noise from the amplifiers. Linearization of the Manakov equation (2.1) is often used to find approximate analytical solutions of the equation. Most linearization techniques can be classified broadly into two categories: perturbation-based techniques [16, 34–36] and techniques based on a Volterra series transfer function [37, 38].

Research in this field can also be classified into channels with non-dispersive effects (excluding CD) and channels with highly dispersive effects. In the former case, analytical expressions for the PDF have been derived [39–44]. For the highly dispersive channel case, the distribution after EDC is assumed to be close to Gaussian [45–51]. For detailed classification of different channel models, see [52, Table I] and references therein. It is important to point out that there exist no channel models that account for the dispersion, nonlinearity, and noise from

[5] In papers D and E, a single-polarization signal is used.

amplifiers for a DM link, and therefore the PDF of the received signal is still unknown. In the next section, a numerical method for signal propagation will be described.

## 2.3 Numerical Methods for Signal Propagation

The Manakov equation (or the NLSE for the single-polarization case) does not generally have an exact analytical solution and approximate analytical solutions are not available for all cases. Numerical approaches such as the split-step Fourier method (SSFM) [53] are often used to describe, understand, and solve for the signal propagation in dispersive and nonlinear media [10]. In this section, the main idea behind the SSFM will be discussed.

The Manakov equation (2.1) can be re-written as

$$\frac{\partial \mathbf{a}}{\partial z} = (\hat{D} + \hat{N})\mathbf{a} \tag{2.3}$$

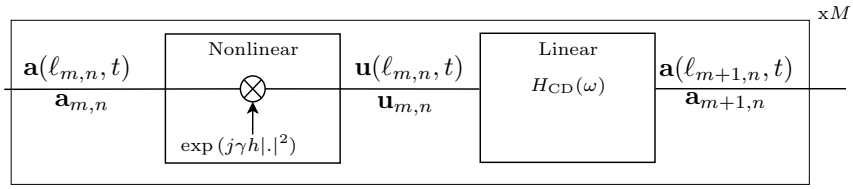
where $\hat{D}$ is a linear differential operator accounting for dispersion and losses in the fiber and $\hat{N}$ is a nonlinear operator. Even though nonlinearity and dispersion act together in the fiber, the SSFM assumes that when the optical field is propagated over a small distance $h$, the dispersion and the nonlinear effects act independently. Using this assumption, an approximate solution is obtained by propagating the signal from $z$ to $z+h$ in a two-step process. In the first step, by setting $\hat{D} = 0$ in (2.3), one can account for the nonlinear effects. In the second step, by setting $\hat{N} = 0$ in (2.3), dispersion and losses are accounted for. Mathematically, given the field at $z$, the field at $z+h$ can be approximated as $\mathbf{a}(z+h,t) \approx \exp(h\hat{D})\exp(h\hat{N})\mathbf{a}(z,t)$, where the exponential operator $\exp(h\hat{D})$ can be evaluated in Fourier domain [11, eq. 2.4.5]. The interpretation and implementation of this equation is as follows. In the first step, the phase change due to the Kerr effect is applied in the time domain using $\mathbf{a}(z+h,t) = \mathbf{a}(z,t)\exp(j\gamma h\|\mathbf{a}(z,t)\|^2)$. In the second step, phase changes due to CD and power losses in the fiber are introduced in the frequency domain as $\tilde{\mathbf{a}}(z+h,\omega) = \tilde{\mathbf{a}}(z,\omega)\exp((j\beta_2\omega^2 - \alpha)h/2)$.

The SSFM explained above is known in the literature as asymmetric SSFM. In symmetric SSFM,[6] the signal propagation over a segment from $z$ to $z+h$ is performed such that the nonlinearity is placed in the middle of the segment rather than at the segment boundary as

$$\mathbf{a}(z+h,t) \approx \exp\left(\frac{h}{2}\hat{D}\right)\exp(h\hat{N})\exp\left(\frac{h}{2}\hat{D}\right)\mathbf{a}(z,t). \tag{2.4}$$

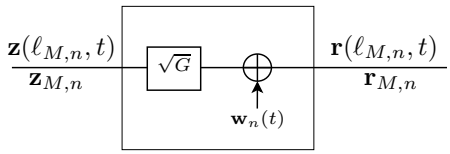
This equation is then applied repeatedly over the length of the fiber $L$, divided into $M$ segments each of length $h$, i.e., $L = Mh$. Even though these methods are

[6] For the theoretical analysis in the paper A, asymmetric SSFM is used. However, the principles suggested in papers do not change even if we use symmetric SSFM. Simulations in all appended papers were carried out using symmetric SSFM. The difference between symmetric and asymmetric SSFM vanishes as $M$ grows large.




Asymmetric SSFM view of fiber for span $n$

(a)




EDFA

(b)

**Figure 2.6:** (a) Approximate model for the fiber (SMF/DCF) using SSFM with nonlinear and linear segments, where the frequency response of CD filter is $H_{\mathrm{CD}}(\omega) = \exp(i\beta_2\omega^2 z/2)$ as defined in Sec. 2.1.1. (b) EDFA with gain $G = \exp(\alpha L)$ and noise.

straightforward to implement, the selection of the step size $h$ is crucial and involves a complexity-accuracy tradeoff [54]. In papers A–E, the SSFM is applied with an adaptive step size of $(\epsilon L_{\mathrm{NL}} L_{\mathrm{D}}^2)^{1/3}$, with $\epsilon = 10^{-4}$ based on [54]. In Fig. 2.6(a), the asymmetric SSFM view of the fiber (SMF/DCF) for span $n \in \{1,\ldots,N\}$ is shown. Focusing on the segment $m \in \{1,\ldots,M\}$ of span $n$, and introducing $\ell_{m,n} = (m-1)h + (n-1)L$, the input signal will be denoted by $\mathbf{a}(\ell_{m,n},t)$ and the output signal after each segment by $\mathbf{a}(\ell_{m+1,n},t)$. In Fig. 2.6(b), mathematical model for the EDFA with gain and noise is shown. When the fiber and EDFA are concatenated, as shown in Fig. 2.6, $\mathbf{z}_{M,n} = \mathbf{a}_{m+1,n}$ when $m = M$. At the end of each span, $\mathbf{r}_{M,n\neq N} = \mathbf{a}_{1,n+1}$, while at the end of the link $\mathbf{r}_{M,N} = \mathbf{r}$. This structure will be used as a building block for factor graphs, as will be explained in the next chapter.

## 2.4 Summary

Starting from the Manakov equation that governs the signal propagation in the nonlinear dispersive medium, impairments existing in the fiber were described. Then the system model used in the papers was described along with the assumptions in this work. The SSFM used to numerically model the fiber was described in Sec. 2.3. It can be noted that the basic building blocks of the fiber-optic channel of Fig. 2.5 can be divided into different sub blocks: pulse shaper operation, CD and NL operations of the fiber, and additive noise operation of the EDFA block. In Chapter 4, these building blocks will be used for the design of optimal receiver.

# Chapter 3

# Bayesian Inference using Factor Graphs

## 3.1 Bayesian Inference

Given the received signal $\mathbf{r}(t)$ of Fig. 2.5, the aim of the detector is to estimate the transmitted data $\mathbf{s}$. Optimal detectors in terms of minimizing the symbol error rate can be built based on the maximum a posteriori (MAP) principle, which is used in this work. Mathematically, MAP detection involves the optimization

$$\hat{\mathbf{s}} = \arg \max_{\mathbf{s} \in \Omega^K} p(\mathbf{s}|\mathbf{r}), \tag{3.1}$$

where $p(\mathbf{s}|\mathbf{r})$ is a shorthand notation for $p_{\mathbf{S}|\mathbf{R}}(\mathbf{s}|\mathbf{r})$.

In the case of fiber-optic channel, finding a closed form expression for $p(\mathbf{s}|\mathbf{r})$ is difficult except in some simplified cases. However, the joint distribution of the input and all intermediates states of the system is generally available, and the determination of $p(\mathbf{s}|\mathbf{r})$ can be seen as a marginalization of this joint distribution.

For instance, assume we have a joint distribution with four random variables $\mathbf{S}$, $\mathbf{X}$, $\mathbf{Z}$, $\mathbf{R}$ and we would like to find $p(\mathbf{s}|\mathbf{r})$, assuming that the underlying structure behind these variables is governed by the Markov property.[1] In particular, let the random variables $\mathbf{S}, \mathbf{X}, \mathbf{Z}, \mathbf{R}$ form a Markov chain as in Fig. 3.1, i.e., $p(\mathbf{r}|\mathbf{s}, \mathbf{x}, \mathbf{z}) = p(\mathbf{r}|\mathbf{z})$ and $p(\mathbf{z}|\mathbf{s}, \mathbf{x}) = p(\mathbf{z}|\mathbf{x})$. Using Bayes' rule and the Markov property, the joint distribution can be factorized as

$$\begin{aligned} p(\mathbf{s}, \mathbf{x}, \mathbf{z}, \mathbf{r}) &= p(\mathbf{r}|\mathbf{s}, \mathbf{x}, \mathbf{z}) p(\mathbf{s}, \mathbf{x}, \mathbf{z}) \\ &= p(\mathbf{r}|\mathbf{z}) p(\mathbf{z}|\mathbf{s}, \mathbf{x}) p(\mathbf{s}, \mathbf{x}) \\ &= p(\mathbf{r}|\mathbf{z}) p(\mathbf{z}|\mathbf{x}) p(\mathbf{x}|\mathbf{s}) p(\mathbf{s}). \end{aligned} \tag{3.2}$$

[1] A stochastic process has the Markov property if the conditional probability distribution of future states of the process (conditional on both past and present values) depends only upon the present state, not on the sequence of events that preceded it.

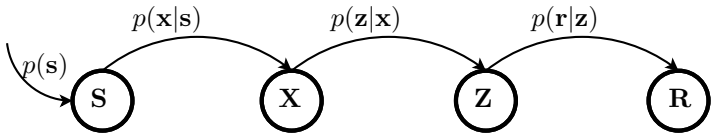



**Figure 3.1:** The random variables $\mathbf{S}, \mathbf{X}, \mathbf{Z}, \mathbf{R}$ forming a Markov chain.

Now for a given observation $\mathbf{r}$, $p(\mathbf{s}|\mathbf{r})$ can be written as a marginal of $p(\mathbf{s}, \mathbf{x}, \mathbf{z}|\mathbf{r})$

$$p(\mathbf{s}|\mathbf{r}) = \sum_{\mathbf{x},\mathbf{z}} p(\mathbf{s}, \mathbf{x}, \mathbf{z}|\mathbf{r}), \tag{3.3}$$

where $p(\mathbf{s}, \mathbf{x}, \mathbf{z}|\mathbf{r}) = p(\mathbf{s}, \mathbf{x}, \mathbf{z}, \mathbf{r})/p(\mathbf{r})$. To find this marginal in a systematic and computationally efficient, the factorization (3.2) can be used. Factor graphs (FGs) and the sum-product algorithm (SPA) are tools to compute these marginal posteriors rigorously and efficiently.

## 3.2 Factor Graphs

As seen in the previous section, the global function of several variables is factorized into several local functions, each involving a small subset of variables. In many applications, as in our case, the global function represents the joint probability density and the corresponding local functions are various conditional distributions. FGs visualize this factorization and the interaction of the various random variables that are involved in a particular problem. One of the key features of FGs is that they support a variety of summary propagation or message-passing algorithms (e.g., the SPA, the min-sum algorithm, and other variations) that can be used for Bayesian inference.

Factor graphs are a generalization of other graphs proposed in the literature. FGs are strongly connected with coding theory, and the foundations of graphical models usage in coding dates back to Gallager, who in his PhD thesis in 1963 visualized a code as a graph [55]. Forney in 1973 introduced a trellis diagram as a way to show the time evolution of a finite-state machine [56]. Tanner graphs were introduced in 1981 as a way to describe a family of codes [57]. The work of Pearl in 1988 on probability propagation (or belief propagation) in Bayesian networks has attracted much attention in artificial intelligence and statistics. Applications of these graph-theoretical models beyond coding were described by Wiberg in his PhD thesis in 1996 [58]. Wiberg also introduced SPA as a message-passing algorithm over a graph [59]. A large number of existing algorithms in the fields of coding, signal processing, and computer science can be viewed as instances of the SPA. The algorithms derived in this way are often viewed as special cases or as obvious approximations of existing well-known algorithms. For example, the decoding algorithm for low-density parity check codes, the Viterbi algorithm,

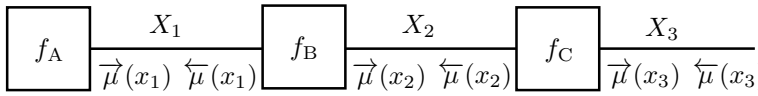



**Figure 3.2:** A simple factor graph for variables $X_1, X_2, X_3$ where every factor has at most two variables. The leftward and rightward messages are also shown in the figure.

Kalman filtering, and the fast Fourier transform can be seen as an instance of the SPA over an appropriately chosen FG. New algorithms for complex detection and estimation problems can also be derived as instances of the SPA [60, 61].

Let us now take a simple example to show how an FG is drawn. Consider a joint distribution $f : \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3 \to \mathbb{R}$, which is factorized into 3 factors,

$$f(x_1, x_2, x_3) = f_A(x_1) f_B(x_1, x_2) f_C(x_2, x_3), \tag{3.4}$$

where an instance (a realization) of a random variable, say $x_i$, belongs to the set $\mathcal{X}_i$. Here $f$ is the global function and $f_A, f_B, f_C$ are the non-negative valued local functions. From this factorization, an FG can be drawn according to the rules stated below and is shown in Fig. 3.2. A Forney-style FG will be used in this work and it generally contains nodes ($f_A$, $f_B$, $f_C$ of Fig. 3.2), edges ($X_1$ and $X_2$ of Fig. 3.2) and half-edges (if connected to only one node, as in $X_3$ of Fig. 3.2) and is drawn according to following rules[2]:

- a node is created for every factor,
- an edge (or half-edge) is drawn for every variable,
- node $f_A$ is connected to edge $X_i$ iff variable $X_i$ appears in factor $f_A$.

## 3.3 Message Passing

To find the marginals, messages, which are functions of the corresponding variables, are to be exchanged over the edges of the FG through a message-passing algorithm: the SPA. The SPA for our Forney-style FG is summarized as follows. The message out of a factor node $h(X_m, X_n)$ along the edge $X_m$ is the product of $h(X_m, X_n)$ and the message towards $h$ from edge $X_n$ summed over all possible values of $X_n$. Here $h$ represents $f_A$, $f_B$, $f_C$ of Fig. 3.2. We denote messages associated with $x_m$ by $\mu_{X_m}(x_m)$ (or $\mu(x_m)$ when the variable is clear from the context), and the direction of the message will be represented by arrows: $\overleftarrow{\mu}_{X_m}(x_m)$ and $\overrightarrow{\mu}_{X_m}(x_m)$. In Fig. 3.2, the leftward message $\overleftarrow{\mu}_{X_2}(x_2)$ is from the node $f_C$ towards $f_B$ and $\overrightarrow{\mu}_{X_2}(x_2)$ is other way around. The detailed SPA is given in [59, p. 39] but in our case, where a factor has at most two variables and a variable can be in

[2]If a variable appears in more than two factors, special measures have to be taken and the above rules have to be modified slightly. This will not be described here as the problem of our interest deals with variables appearing in at most two factors.

at most two factors, the FG is linear (a FG without any branches and loops) as shown in Fig. 3.2. We note that messages can be normalized without affecting the normalized marginal. These normalized messages can be interpreted as probability mass functions (PMFs) (resp., PDFs) when the variables are discrete (resp., continuous).

### 3.3.1 Message Computation

- *Initialization*: All edges $X_k$ connected to a single node, such as $X_3$ in Fig. 3.2, transmit message $\mu_{X_k}(x_k) = 1$, $\forall x_k \in \mathcal{X}_k$. Nodes $h$ connected to a single edge $X_m$, such as $f_{\mathrm{A}}$ of Fig. 3.2, transmit message $\mu_{X_m}(x_m) = h(x_m)$, $\forall x_m \in \mathcal{X}_m$.

- *Message computation*: When a node $h$ has received the incoming message, the outgoing message on the remaining edge, say $X_m$, is computed as

$$\mu_{X_m}(x_m) = \sum_{x_n} h(x_m, x_n)\mu_{X_n}(x_n), \forall x_m \in \mathcal{X}_m. \tag{3.5}$$

  Note that the summation is over all possible values of $x_n$.

- *Termination*: The marginal for variable $x_k$ can be obtained once the two messages on the corresponding edge are available using[3]

$$g_{X_k}(x_k) = \overleftarrow{\mu}_{X_k}(x_k)\overrightarrow{\mu}_{X_k}(x_k), \forall x_k \in \mathcal{X}_k. \tag{3.6}$$

Summation should be replaced by integration for continuous variables.

For example in Fig. 3.2, suppose the aim is to find the marginal distribution of variable $X_1$. The first step is to start with message $\overleftarrow{\mu}_{X_3}(x_3)$, which is a constant. Then the leftward messages for variables $X_2$ and $X_1$ are computed using (3.5) to get $\overleftarrow{\mu}_{X_1}(x_1)$. During the initialization stage, the rightward message for $X_1$, $\overrightarrow{\mu}_{X_1}(x_1)$ can also be found. In the last step, the leftward and rightward messages for variable $X_1$ are multiplied to get the marginal of $X_1$.

When the variables $X_i$ are continuous (resp., discrete), the messages are scaled probability density (resp., mass) functions. For such variables, SPA rules often lead to intractable integrals and therefore the representation of messages is an important issue in such works. Many different approaches exist how to solve this problem such as the following.

- Considering a grid and evaluating the message at each of the grid points, which leads to a vector that represents the message.

- Approximating the message with parameterized distributions such as a mixture of Gaussians.

- Approximating the message by a list of samples or particles.

In the next section, the interpretation of the messages and the message passing rules will be detailed assuming that the variables are discrete.

[3]The details behind getting marginals when two messages are multiplied is not explained here but is described in the literature [60, 61].

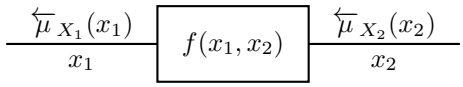



**Figure 3.3:** A simple factor graph with random variables $x_1$ and $x_2$ related by the factor $f(x_1, x_2)$. Also shown are the messages corresponding to these random variables.

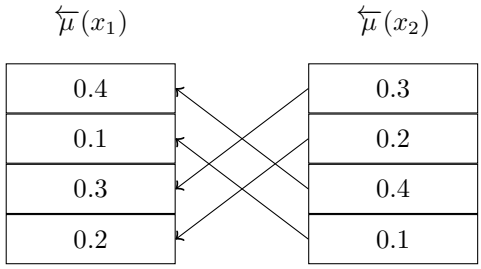



**Figure 3.4:** The messages, which are PMFs of variables $x_1$ and $x_2$, are shown. The arrows represent the permutations of the bijective function $\phi$.

### 3.3.2 Interpretation of Messages

Consider two discrete random variables $x_1$ and $x_2$ related by a joint distribution function $f(x_1, x_2)$. The messages, $\overleftarrow{\mu}_{X_1}(x_1)$ and $\overleftarrow{\mu}_{X_2}(x_2)$ in Fig. 3.3, are then PMFs of the corresponding random variables and thus can be interpreted as vectors. The message passing rules can be determined (and even interpreted) using a matrix-vector multiplication.

For example, assume $x_1$ and $x_2$ to be two quaternary random variables and suppose $x_2 \in \{\kappa, \psi, \chi, \zeta\}$ is an instance of the random variable $X_2$ that has a PMF: $(0.3, 0.2, 0.4, 0.1)$ and the message, $\overleftarrow{\mu}_{X_2}(x_2)$, is this PMF represented as a vector. Assume $x_1 \in \{a, b, c, d\}$ and $x_2$ are related by a bijective function $\phi$ as $x_2 = \phi(x_1)$ with $\phi^{-1}(x_2) = [c, d, a, b]$ as shown in Fig. 3.4. According to the SPA, the message $\overleftarrow{\mu}_{X_1}(x_1)$ can be computed as

$$\begin{aligned}\overleftarrow{\mu}_{X_1}(x_1) &= \sum f(x_1, x_2)\overleftarrow{\mu}_{X_2}(x_2) = \sum \delta(x_2 - \phi(x_1))\overleftarrow{\mu}_{X_2}(x_2) \\ &= \overleftarrow{\mu}_{X_2}(\phi(x_1)),\end{aligned} \tag{3.7}$$

or $\overleftarrow{\mu}_{X_1}(\phi^{-1}(x_2)) = \overleftarrow{\mu}_{X_2}(x_2)$ i.e., $\overleftarrow{\mu}_{X_1}([c, d, a, b]) = \overleftarrow{\mu}_{X_2}([\kappa, \psi, \chi, \zeta])$. This means that the output message, $\overleftarrow{\mu}_{X_1}(x_1)$, is a permutation of the input message. Thus the message $\overleftarrow{\mu}_{X_1}(x_1) \propto (0.4, 0.1, 0.3, 0.2)$. In this example, the message computation can be described as $\overleftarrow{\mu}_{X_1}(x_1) = A^T\overleftarrow{\mu}_{X_2}(x_2)$, where $A$ is a permutation matrix based on the relation $f(x_1, x_2) = \delta(x_2 - \phi(x_1))$ and Fig. 3.4.

When the variable is discrete, the extension to non-bijective mappings is straightforward. In this case, the matrix $A$ will not be invertible, but messages can still be found using a matrix-vector multiplication. When $x_1$ and $x_2$ are continuous random variables and can take on real values, then instead of the PMFs, we will have

the PDFs as messages. In particular, an expression of the form $\int \delta(y-x)\mu_X(x)\mathrm{d}x$ can be simplified immediately to $\mu_X(y)$. However, when the message is of the form $\int \delta(y-\phi(x))\mu_X(x)\mathrm{d}x$, then the random variables need to be transformed (for example by introducing $z=\phi(x)$) before getting the final message.

## 3.4 Particle Representation

As mentioned at the end of Sec. 3.2, the SPA rules often lead to intractable integrals. In this section, a brief introduction to the particle representation (PR), used in SDBP, will be given.

Given a probability density function $p_X$, a PR, denoted by $\mathrm{PR}\{p_X\}$, is a list of values[4] $\{x^{(k)}\}_{k=1}^{N_p}$, with the property that for any integrable function $f$

$$\frac{1}{N_p}\sum_{k=1}^{N_p} f(x^{(k)}) \to \int f(x)p_X(x)\mathrm{d}x, \quad N_p \to \infty. \tag{3.8}$$

One way to obtain a PR is to draw $N_p$ i.i.d. samples from $p_X(x)$, though many other methods exist [59, ch. 3]. Note that a PR can easily be extended to high-dimensional variables. PR can be interpreted as follows: in the context of (3.8), $p_X(x)$ can be approximated as $p_X(x) \approx 1/N_p \sum_{k=1}^{N_p} \delta(x-x^{(k)})$. In other words $\mathrm{PR}\{p_X\}$ can be considered as a uniform probability mass function, so that $X$ is considered to be a uniform discrete random variable that can take on values in the set $\{x^{(k)}\}_{k=1}^{N_p}$.

## 3.5 Application of Factor Graphs for the Fiber-Optic Channel

For the fiber-optic channel of Fig. 2.5, the joint distribution consists of variables involved in each segment and span of the SSFM representation of the fiber and the input and output variables of the EDFAs. These internal variables are defined as

$$\mathcal{H} = \{\{\mathbf{a}_{m,n}, \mathbf{u}_{m,n}\}_{1\le m\le M}, \mathbf{z}_{M,n}, \mathbf{r}_{M,n\neq N}\}_{1\le n\le N},$$

where we note that the observation $\mathbf{r}_{M,N}=\mathbf{r}$ is not part of $\mathcal{H}$. Harnessing conditional independence among these internal states, the joint distribution $p(\mathbf{s},\mathcal{H},\mathbf{r})$ is factorized as

$$\begin{aligned} &p(\mathbf{s}, \{\{\mathbf{a}_{m,n}, \mathbf{u}_{m,n}\}_{1\le m\le M}, \mathbf{z}_{M,n}, \mathbf{r}_{M,n\neq N}\}_{1\le n\le N}, \mathbf{r}) \\ &= p(\mathbf{s})\; p(\mathbf{a}_{1,1}|\mathbf{s}) \prod_{n=1}^{N}\prod_{m=1}^{M} p(\mathbf{u}_{m,n}|\mathbf{a}_{m,n}) p(\mathbf{a}_{m+1,n}|\mathbf{u}_{m,n}) p(\mathbf{r}_{M,n}|\mathbf{z}_{M,n}), \end{aligned} \tag{3.9}$$

[4] Variations exist where the values are weighted by using importance sampling. We do not apply weighting in this context.

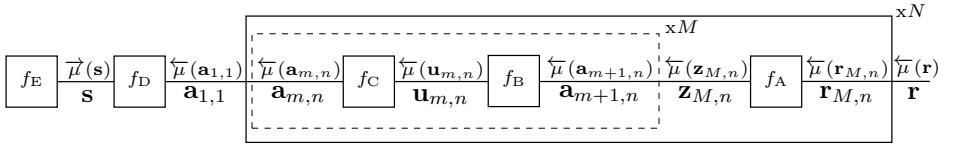



**Figure 3.5:** Factor graph for the system in Fig. 2.5 where $f_{\mathrm{A}}$ corresponds to the EDFA block, $f_{\mathrm{B}}$ corresponds to the CD block (for either an SMF or for a DCF), $f_{\mathrm{C}}$ corresponds to NL block (either for SMF or DCM), $f_{\mathrm{D}}$ corresponds to the pulse shape and $f_E$ represents prior knowledge about the symbols.

where $\mathbf{z}_{M,n} = \mathbf{a}_{M+1,n}$ for last segment of SSFM, i.e., when the fiber and EDFA are concatenated. An FG can be drawn as shown in Fig. 3.5 with the factors defined as

$$f_{\mathrm{A}} : p(\mathbf{r}_{M,n}|\mathbf{z}_{M,n}) = \mathcal{N}(\mathbf{r}_{M,n}; \sqrt{G}\mathbf{z}_{M,n}, \mathbf{\Sigma}), \tag{3.10}$$

$$f_{\mathrm{B}} : p(\mathbf{a}_{m+1,n}|\mathbf{u}_{m,n}) = \delta(\mathbf{a}_{m+1,n} - \sqrt{(1/G)}\mathbf{u}_{m,n} * \mathbf{h}), \tag{3.11}$$

$$= \delta(\mathbf{a}_{m+1,n} - \mathbf{B}\mathbf{u}_{m,n}), \tag{3.12}$$

$$f_{\mathrm{C}} : p(\mathbf{u}_{m,n}|\mathbf{a}_{m,n}) = \delta(\mathbf{u}_{m,n} - \mathbf{a}_{m,n}\exp(j\gamma h\|\mathbf{a}_{m,n}\|^2)) \tag{3.13}$$

$$= \delta(\mathbf{u}_{m,n} - \phi(\mathbf{a}_{m,n})), \tag{3.14}$$

$$f_{\mathrm{D}} : p(\mathbf{a}_{1,1}|\mathbf{s}) = \delta(\mathbf{a}_{1,1} - \mathbf{s} * \mathbf{p}) \tag{3.15}$$

$$= \delta(\mathbf{a}_{1,1} - \mathbf{P}\mathbf{s}), \tag{3.16}$$

where $\mathbf{h}$ is the CD time domain response of $H_{\mathrm{CD}}(\omega)$, (3.10) is due to the fact that $\mathbf{r}_{M,n}$ is obtained by scaling $\mathbf{z}_{M,n}$ with $\sqrt{G}$ and adding Gaussian noise of zero mean and covariance matrix $\mathbf{\Sigma}$. The pulse shape is represented by $\mathbf{p}$ in (3.15) and the transmitted signal is represented by $\mathbf{a}_{1,1} \triangleq \mathbf{a}(\ell_{1,1}, t) = \sum_{k=1}^{K} \mathbf{s}[k]\mathbf{p}(t - kT)$, where $T$ is the symbol period. The pulse shaping operation can be represented as a matrix multiplication, as $\mathbf{a}_{1,1} = \mathbf{P}\mathbf{s}$. Similarly, the CD operation can be represented as a matrix operation for an invertible $\mathbf{B}$. To simplify the notation, we have defined the NL operator to be $\phi()$. To keep the FG framework compact, single-stage representation is shown in the FG. However, for DM links, depending on SMF/DCF, $\beta_2$, $\gamma$, $G$ need to be adjusted accordingly in $f_{\mathrm{B}}$, $f_{\mathrm{C}}$, $f_{\mathrm{A}}$, respectively.

### 3.5.1 Particle Representation for the Building Blocks of the Fiber-Optic Channel

Given the message $\overleftarrow{\mu}_{\mathbf{X}_2}(\mathbf{x}_2)$ of Fig. 3.3 in particle representation, $\{\mathbf{x}_2^{(k)}\}_{k=1}^{N_p}$, the message $\overleftarrow{\mu}_{\mathbf{X}_1}(\mathbf{x}_1)$ in particle format will be computed below using SPA of (3.5) for each of the local functions of the FOC, $f_{\mathrm{A}}$, $f_{\mathrm{B}}$, $f_{\mathrm{C}}$, and $f_{\mathrm{D}}$.

1. Addition of noise in EDFA: Given the incoming message $\{\mathbf{r}_{M,n}^{(k)}\}_{k=1}^{N_p}$, the message $\overleftarrow{\mu}_{\mathbf{Z}_{M,n}}(\mathbf{z}_{M,n})$ can be computed by generating $\mathbf{n}_{M,n}^{(k)} \sim \mathcal{N}(0, \Sigma)$ and setting $\mathbf{z}_{M,n}^{(k)} = (1/\sqrt{G})\mathbf{r}_{M,n}^{(k)} + \mathbf{n}_{M,n}^{(k)}$.

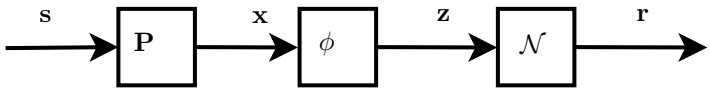



**Figure 3.6:** A simplified FG of Fig. 3.5 with $M = 1$, $N = 1$, ignoring CD, data-bearing symbols $\mathbf{s}$, a transmitted signal $\mathbf{x} = \mathbf{P}\mathbf{s}$, an invertible non-linearity $\mathbf{z} = \phi(\mathbf{y})$, and addition of noise, leading to an observation $\mathbf{r}$. Note that $\mathbf{r}$ is observed, while $\mathbf{s}$, $\mathbf{y}$, $\mathbf{z}$ are hidden. Note also that $\mathbf{s} \in \mathbb{R}^M$, while $\mathbf{y}, \mathbf{z}, \mathbf{r} \in \mathbb{R}^N$, $N \gg M$.

2. NL: Given the incoming message $\{\mathbf{u}_{M,n}^{(k)}\}_{k=1}^{N_p}$, the message $\overleftarrow{\mu}_{\mathbf{a}_{M,n}}(\mathbf{a}_{M,n})$ can be computed using $\mathbf{a}_{M,n}^{(k)} = \phi^{-1}(\mathbf{u}_{M,n}^{(k)})$, i.e., $\overleftarrow{\mu}_{\mathbf{a}_{M,n}}(\phi^{-1}(\mathbf{u}_{M,n})) = \overleftarrow{\mu}_{\mathbf{u}_{M,n}}(\mathbf{u}_{M,n})$. If $\phi^{-1}(\cdot)$ is applied to the signal $\mathbf{u}_{M,n}$, then the resulting signal $\mathbf{a}_{M,n}$ will have the same probability density as $\mathbf{u}_{M,n}$.

3. Pulse shaper: One way of accounting for the effect of the pulse shaping operation is through filtering matched to the pulse shape. A particle $\mathbf{s}^{(k)}$ can be computed from a particle $\mathbf{a}_{1,1}^{(k)}$ as $\mathbf{s}^{(k)} = \left(\mathbf{P}^H\mathbf{P}\right)^{\#}\mathbf{P}^H\mathbf{a}_{1,1}^{(k)} = T_s\mathbf{P}^H\mathbf{a}_{1,1}^{(k)}$, which corresponds to discrete-time matched filtering. In Sec. 4.3.3, an alternative to an MF operation will be presented.

4. CD: As $B$ is an invertible operation, the message can be computed using similar approach as for the NL operation.

While MF is an optimal operation to perform for linear channels affected by AWGN, it may be suboptimal for nonlinear channels. In Sec. 4.3, other ways of dealing with this suboptimal MF operation will be explained for the FOC. In the next section, a simple digital communication receiver will be shown, where particle representation will be used to detect the transmitted signal.

### 3.5.2 A Particle-Based Digital Communication Receiver

As a simple example of the concepts presented above, consider a simplified communication system whose FG is shown in Fig. 3.6, where $\mathbf{s}$ is a sequence of $M$ numbers drawn from $\{-1, +1\}$. The variable $\mathbf{x}$ is further affected by an invertible non-linear function, giving rise to $\mathbf{z} = \phi(\mathbf{x})$, after which Gaussian noise is added with zero mean and covariance matrix $\Sigma_{\text{noise}}$. Given an observation $\mathbf{r}$, backward particle-based message passing now operates as follows:

1. Generate $N_p$ copies of $\mathbf{r}$, denoted by $\mathbf{r}^{(1)}, \mathbf{r}^{(2)}, \ldots, \mathbf{r}^{(N_p)}$. These form an unweighted particle representation of $\overleftarrow{\mu}_{\mathbf{R}}(\mathbf{r})$.

2. Compute a particle representation of $\overleftarrow{\mu}_{\mathbf{Z}}(\mathbf{z})$ by generating $\mathbf{z}^{(k)} = \mathbf{r}^{(k)} + n^{(k)}$, where $n_k \sim \mathcal{N}(n; 0, \Sigma_{\text{noise}})$.

3. Compute a particle representation of $\overleftarrow{\mu}_{\mathbf{X}}(\mathbf{x})$ by generating $\mathbf{x}^{(k)} = \phi^{-1}(\mathbf{z}^{(k)})$.

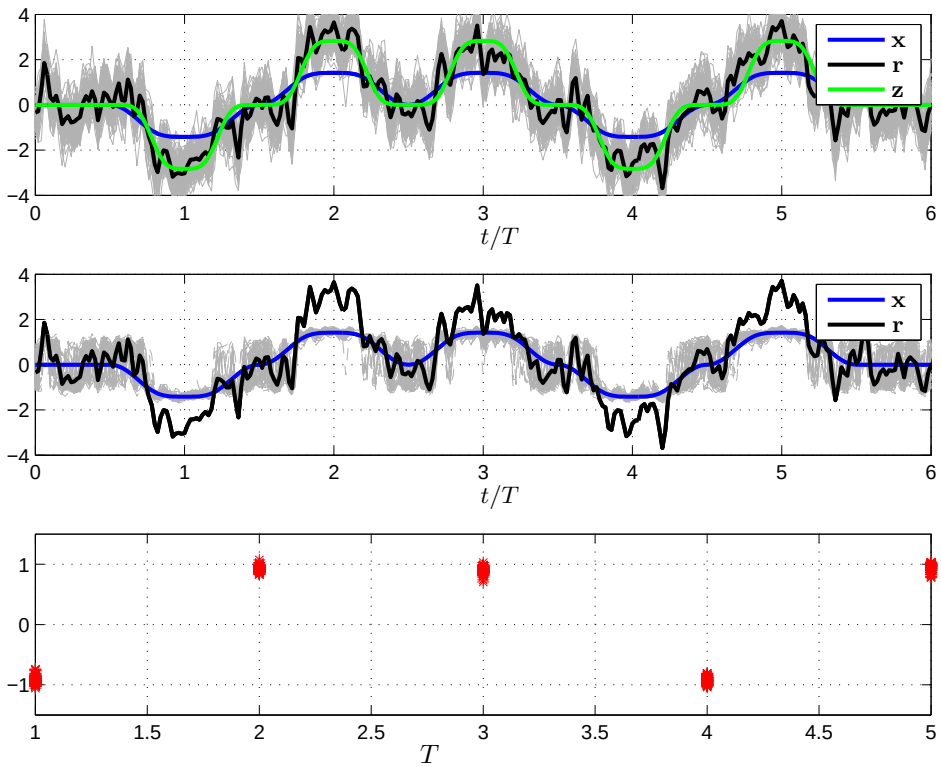



**Figure 3.7:** Particle messages in a digital communication system. The transmitted symbols were $\mathbf{s} = [-1 \ +1 \ +1 \ -1 \ +1]^T$. The top figure shows $\mathbf{x}$ (the signal after pulse shaping), $\mathbf{z}$ (after passing through the nonlinearity), and $\mathbf{r}$ (after adding noise). The grey waveforms in the background correspond to the particle representation of $\overleftarrow{\mu}_{\mathbf{Z}}(\mathbf{z})$. The middle figure shows the particle representation of $\overleftarrow{\mu}_{\mathbf{X}}(\mathbf{x})$ (in gray), while the bottom figure shows the particle representation of $\overleftarrow{\mu}_{\mathbf{S}}(\mathbf{s})$. Observe that the $M = 5$ symbols can be correctly recovered from the particles.

4. Compute a particle representation of $\overleftarrow{\mu}_{\mathbf{S}}(\mathbf{s})$ by generating $\mathbf{s}^{(k)} = T_s \mathbf{P}^H \mathbf{x}^{(k)}$, which is equivalent to applying a matched filter to each particle $\mathbf{x}^{(k)}$.

5. From the particles $\mathbf{s}^{(k)}$, we can make optimal decisions regarding $\mathbf{s}$ using MAP decision of (3.1).

An example of such a scheme is shown in Fig. 3.7, for the case where $M = 5$, $\phi(\mathbf{x}) = \mathbf{x}^3$, defined as applying the third power to each entry in $\mathbf{x}$, which is invertible over $\mathbb{R}$, and low-pass noise.

## 3.6 Summary

The principles behind our proposed detector of SDBP are presented, starting with Bayesian inference, factor graphs, the sum-product algorithm, and particle representation of the messages. We introduced the notion of global and local functions and how the factorization can be pictorially represented using an FG. Then SPA was introduced, which can be applied to the FG, and a marginal can be found. We noted that the messages represent scaled probability density functions and since a closed-form expression of the messages is difficult, we used a particle representation for the messages. In particular, messages are approximated with a list of samples in SDBP. Message passing for the different building blocks of FOC was then explained using particle representation. A simple digital communication receiver was built based on particle representation before applying these concepts for the FOC in the next chapter.

# Chapter 4

# Nonlinear Compensation Techniques

Nonlinear compensation techniques can be broadly classified into two categories: electronic compensation through DSP and optical compensation. Different techniques have been surveyed in [31, Table I].

## 4.1 Optical Nonlinearity Compensation

Techniques in this category can be further categorized into the use of solitons, optical phase conjugation (OPC), optical backpropagataion (OBP), phase-conjugated twin waves (PCTW), and phase-sensitive amplifiers (PSAs). The use of optical solitons is the first method suggested to mitigate the nonlinear distortions due to Kerr effect [33, 62]. A soliton is an optical pulse that is formed when the phase shift induced by the self-phase modulation exactly counterbalances the phase shift induced by the CD, leading to a pulse propagation unaltered in the fiber. However, solitons were limited to low-order modulation formats such as on-off keying and suffer from pulse interactions. Communications based on Solitons is regaining attention in the context of eigenvalue communication and nonlinear Fourier transform [63–66].

The second method for compensating for the impairments in fiber-optic channels is OPC. The idea behind OPC is to reverse the nonlinear distortions generated in the first half of the link by performing a phase conjugation in the optical domain on the signal at the mid-point of a fiber-optic link. Even before coherent detection was possible, the use of phase conjugation was suggested as a way to compensate for the dispersion in optical fibers [67]. Based on this work, it was demonstrated that an OPC can also cancel out the influence of the nonlinear Kerr effect [68]. In order for this approach to work efficiently, nonlinearity, dispersion, and power profiles need to be symmetric with respect to the OPC location. Recently, OPC has regained attention as a promising candidate for reducing the

impact of NSNI [21, 24, 29, 30]. In [29], a comparison of DBP and OPC was performed and the study has concluded that ideal OPC has 1.5 dB in SNR improvement compared to DBP. A combination of OPC and DBP has been suggested an an alternative to complement the shortcomings in each approach and thereby reaping the benefits of both optical and digital worlds [24].

The third approach for compensating for nonlinearities in the optical domain is the use of materials with a negative nonlinear coefficient. A medium with a negative nonlinear index (e.g., semiconductors) to reverse the effects of transmission without OPC was first suggested by Paré in 1996 [69]. OBP can also be seen in this category [70, 71]. A basic version of OBP consists of an OPC, a DCF or an FBG, and highly nonlinear fiber. An extension of this basic algorithm of OBP was investigated by introducing a single optical device which can compensate for dispersion and nonlinearity [71].

Two other notable techniques include PCTW [72] and the use of PSAs [73]. In the basic version of PCTW, a phase-conjugated copy of the signal is sent on the orthogonal polarization along with the signal of interest. As a result, the spectral efficiency is reduced by 50% using PCTW. Extensions of this basic version of PCTW to utilize other dimensions than polarizations, and also a way to improve spectral efficiency has been suggested [74, Ch. 4]. The investigation of using PSAs for nonlinear distortion is in early stages and currently the results are available for a single span, single-channel systems. A detailed comparison of OPC, PCTW, and PSA has been reported in [74, Ch. 4].

## 4.2 Digital Nonlinearity Compensation

On the other end, compensation for the impairments has been done using electrical signal processing. Techniques in this category can be broadly categorized into predistortion [75–80], DBP either at the transmitter or at the receiver or a combination of both [7,8,20,23,28,30,31,81–105], techniques based on Volterra kernels [37,38], and techniques based on maximum a posteriori criteria [7,8,106–108].

One of the first studies were done in early 1990s where dispersion compensation was suggested using electronic signal processing at the transmitter [75]. With the advancements in DSP, dispersion was compensated for using electronic precompensation at the transmitter or using a coherent or intradyne detection with coherent DSP equalization at the receiver [76–79]. Starting 2005, many studies have been focused on nonlinearity mitigation mainly using predistortion at the transmitter [80]. In these approaches, predistorted signals are calculated by backward propagating the desired signal from the receiver to the transmitter. The idea is that the pre-distortions in the transmitted signal cancel the distortions accumulated in the real fiber propagation.

### 4.2.1 Digital Backpropagation

In the absence of noise, the transmitted signal can be found by solving the inverse of the Manakov equation (2.1) by propagating the output signal with inverse

parameters ($-\beta_2$, $-\gamma$, $-\alpha$) to invert the channel effects and get $\mathbf{a}(0,t)$, and this technique is called DBP [7, 8, 20, 28, 30, 31, 81–92]. In [81, 82], DBP was studied as a transmitter-side precompensation algorithm, whereas in [83, 84], DBP was performed using receiver DSP. To backpropagate the signal through a section of virtual fiber which extends from $z+h$ to $z$, methods such as noniterative asymmetric SSFM [83] and iterative symmetric SSFM [84,87] have been used. In spite of the high computational complexity, DBP has been proposed as a universal technique[1] for jointly compensating for the linear and nonlinear impairments and is often used to benchmark against other detectors [85–90]. Many variations of DBP were suggested for reducing the complexity including weighted DBP [85], perturbation DBP [86,88], folded DBP [91], and filtered DBP [89,92]. All these techniques aim to reduce the step-size requirements in DBP by improving the performance [83,84]. A survey of these approaches can be found in [20, 31].

In a single-channel system, deterministic signal–signal interactions, NISI, are perfectly compensated for by DBP. However, DBP does not account for the NSNI, and hence is optimal when noise from the optical amplifiers of the fiber-optic channels (Fig. 2.5) is ignored. As discussed in Sec. 2.1.4, reducing the impact of NSNI after compensating for nonlinearities has been one of the active research areas. In [28], NSNI is reduced by splitting DBP between transmitter and receiver. In [30], it has been shown that splitting nonlinear compensation between transmitter and receiver has better performance than doing compensation only at the receiver or at the transmitter. A comparison of DBP and OPC was performed in [29] and it was observed that an ideal OPC has 1.5 dB in SNR improvement than DBP. It has been pointed out that this gain for OPC techniques is due to the fact that DBP does not account for the NSNI. In [30], multiple OPC modules were used and the NSNI is shown to be reduced in comparison to using only one OPC module.

DBP has been used as a benchmark not only for single-channel systems but also for so-called superchannels, which are optical channels consisting of closely spaced carriers that are treated as a single channel propagating through a link [23,93–105]. These studies reveal that there is gain in jointly performing DBP over multiple channels. However, adding an extra channel in the DBP beyond a certain number of channels results in only a incremental increase in gains in comparison to the significant computational complexity [23, 99, 105]. When the DBP is performed over all the channels of the superchannel, all the deterministic effects including inter-channel effects are compensated for and the performance gains are limited by the signal–noise interactions [23].

For WDM systems, DBP is used only for the center channel (channel of interest), thereby reducing the intra-channel effects [6, 23, 97]. This is because DBP requires the knowledge of the parameters of the forward propagation, which are not available for all channels of a WDM system. Recently, a combination of DBP and OPC has been suggested in [24], which relaxes both the required DBP accuracy and the OPC placement, and offers a performance benefit.

[1] Independent of modulation formats and applicable both for single-carrier and OFDM formats.

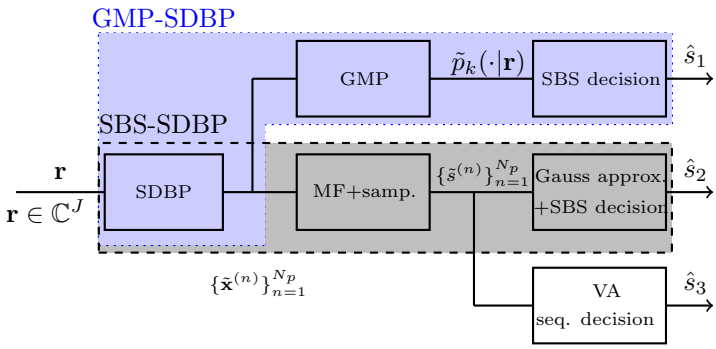



**Figure 4.1:** The received signal $\mathbf{r}$ from the fiber-optic channel of Fig. 2.5 processed through three variations of SDBP.

### 4.2.2 Maximum Likelihood Sequence Detectors

Receivers can also be designed based on maximum likelihood sequence detection techniques that are optimal in terms of minimizing the SER [7,8,106–108]. A look-up table based implementation was used in [106,108,109] to mitigate the pattern dependent nonlinearities, and is thus limited to the size of the look-up table. In [8], the Viterbi algorithm is used after DBP with different metrics to find the optimum nonlinearity compensator. Although the results are very promising, this approach is guided by a heuristic approach [8] and applicable for low inter-symbol interference scenarios. Recently, an optimal detector based on MAP criteria was proposed for single span system [7]. However, except for SDBP, presented in this thesis, there exists no general framework that accounts also for signal–noise interactions besides compensating for the intra-channel effects in the single-channel fiber-optic links.

## 4.3 Stochastic Digital Backpropagation

In all three variations of SDBP, the message passing is done as follows. Starting with the received signal $\mathbf{r}$, the message $\overleftarrow{\mu}_{\mathbf{R}}(\mathbf{r})$ (in particle form) is passed through the inverse of each of the NL, CD steps of SSFM, and EDFAs, to get the message $\overleftarrow{\mu}_{\mathbf{X}}(\mathbf{x})$ in particle form $\{\tilde{\mathbf{x}}^{(n)}\}_{n=1}^{N_p}$. A pseudo-code for the implementation of the SDBP algorithm is provided in paper C. Given $\overleftarrow{\mu}_{\mathbf{X}}(\mathbf{x})$ in particle form, the task is to compute $\overleftarrow{\mu}_{\mathbf{S}}(\mathbf{s})$ by accounting for the pulse shaping operation. This step involves going from a high dimensional space (oversampled waveforms) to a low dimensional space (data symbols). As the message $\overleftarrow{\mu}_{\mathbf{X}}(\mathbf{x})$ is in particle format, there exists no known method that can account for the effects of the pulse shaper operation. At this stage, three different approaches as illustrated in Fig. 4.1 are used and are detailed next.

### 4.3.1 SBS-SDBP

In the first approach, proposed in paper A, to account for the effects of pulse shape, we take the particles $\{\tilde{\mathbf{x}}^{(n)}\}_{n=1}^{N_p}$ and pass each particle waveform through MF and sampling. Then, a set of $N_p$ particles are obtained corresponding to each symbol duration, which are approximated with a Gaussian distribution, to get $\tilde{p}_k(s_k|\mathbf{r})$, for $k = 1, \ldots, K$. Based on this distribution, decisions on the transmitted symbols is done on a symbol-by-symbol basis according to MAP rule

$$\hat{s}_k = \arg\max_{s_k \in \Omega} \tilde{p}_k(s_k|\mathbf{r}). \tag{4.1}$$

As the decisions are taken on a symbol level, we call this approach symbol-by-symbol SDBP (SBS-SDBP).

For the fiber-optic channel, MF followed by sampling is a heuristic approach and is optimal only when the message before accounting for the effect of pulse shaping follows the statistics of an AWGN channel. To quantify the loss using this suboptimal approach, we compared SBS-SDBP with a sequence detector based on the VA after DBP proposed in [8]. Results of this comparison are presented in paper B of this thesis. It was observed that although SBS-SDBP has better performance than DBP, SBS-SDBP had worse performance than the algorithm using VA after DBP [110].

### 4.3.2 VA-SDBP

As SBS-SDBP has residual memory left after the MF, the VA was implemented after MF and sampling as shown in Fig. 4.1, and we call this algorithm VA-SDBP. The results of VA-SDBP is shown in paper C. The particles after MF and sampling, $\{\tilde{\mathbf{s}}^{(n)}\}_{n=1}^{N_p}$, can be viewed as samples from a distribution $q_{\mathrm{c}}(\mathbf{s})$ (defined for $\mathbf{s} \in \mathbb{R}^{4K}$), for which $q_{\mathrm{d}}(\mathbf{s})$ (defined for $\mathbf{s} \in \Omega^K$) provides an approximation of $p(\mathbf{s}|\mathbf{y})$. In paper C, we exploited the residual memory by making a decision regarding $\mathbf{s}$ based on the entire distribution $q_{\mathrm{c}}(\mathbf{s})$, rather than its marginals, leading to the following detector

$$\hat{\mathbf{s}} = \arg\max_{\mathbf{s} \in \Omega^K} q_{\mathrm{d}}(\mathbf{s}), \tag{4.2}$$

where $q_{\mathrm{d}}(\mathbf{s}) \propto q_{\mathrm{c}}(\mathbf{s})$. The state in the VA is a function of $L$ past transmitted symbols and the branch metric is a function of the state and the current symbol. The states and branch metrics are obtained from the particles $\{\tilde{\mathbf{s}}^{(n)}\}_{n=1}^{N_p}$, and are explained in Sec. IV. B of paper C.

### 4.3.3 GMP-SDBP

It is important to note that $q_{\mathrm{d}}(\mathbf{s})$ of (4.2) is only an approximation of $p(\mathbf{s}|\mathbf{r})$ and need not be identical to $p(\mathbf{s}|\mathbf{r})$, as the use of the MF followed by sampling at the symbol rate is a heuristic and may lead to information loss. Hence, performing SBS decisions on the marginals of $q_{\mathrm{c}}(\mathbf{s})$ as in paper C may not lead to optimal

performance (in terms of minimizing the probability of error, either symbol-wise or sequence-wise). In paper D, alternatives to an MF were explored performance improvements were seen compared to VA-SDBP.

As explained in Sec. 3.5, the relation between $\mathbf{s}$ and $\mathbf{x}$ in in Fig. 4.1 can be modeled as $\mathbf{x} = \mathbf{P}\mathbf{s}$, where $\mathbf{P}$ represents the pulse shape including up-sampling. In this approach, to obtain $\overleftarrow{\mu}_{\mathbf{X}}(\mathbf{x})$, we start with a Gaussian approximation[2] of $\{\mathbf{x}^{(k)}\}$ with mean $m_{\mathbf{X}}$ and covariance $\Sigma_{\mathbf{X}}$. We then apply the linear GMP according to [111, Table III], leading to $m_{\mathbf{S}} = \left(\mathbf{P}^{\mathrm{T}}\Sigma_{\mathbf{X}}^{-1}\mathbf{P}\right)^{\#}\mathbf{P}^{\mathrm{T}}\Sigma_{\mathbf{X}}^{-1}m_{\mathbf{X}}$ and $\Sigma_{\mathbf{S}}^{-1} = \mathbf{P}^{\mathrm{T}}\Sigma_{\mathbf{X}}^{-1}\mathbf{P}$ followed by SBS decisions. Note that in our application, the number of particles $N_p = 500$ is much smaller than the dimension of $\mathbf{x} = 2048$. To avoid singular estimates of $\Sigma_{\mathbf{X}}$, we used an approach known as tapering as explained in paper D.

SBS-SDBP can be seen as a special case of GMP-SDBP in the following way. In SBS-SDBP, each $\mathbf{s}^{(k)}$ is projected onto $\mathbf{x} = \mathbf{P}\mathbf{s}$, leading to particles $\mathbf{s}^{(k)} = \left(\mathbf{P}^{\mathrm{T}}\mathbf{P}\right)^{\#}\mathbf{P}^{\mathrm{T}}\mathbf{x}^{(k)}$, corresponding to matched filtering followed by symbol rate sampling of each $\mathbf{x}^{(k)}$. For each symbol, the corresponding particles are approximated with a multivariate Gaussian distribution based on which a symbol-by-symbol decision is made.

## 4.4 Performance Metrics

The most widely used performance measures in fiber-optic communications are the bit-error-rate (BER) (or the SER), the Q-factor[3], and the MI. The first two measures are especially relevant and easy to measure for uncoded systems whereas the MI is especially relevant for coded systems. SER is used as a performance metric papers A–D and MI is used as a metric in paper E.

### 4.4.1 Error Probability

For systems affected by AWGN and for simplified fiber-optic channel models, closed-form expressions of the error-rate can be obtained [112–119] such as only ASE [112], ASE with polarization effects for binary amplitude-shift-keying [114], ASE noise with filtering and CD [116], and considering some nonlinearity sources along with ASE and CD [119]. For a fiber-optic channel with nonlinearities, dispersion, and noise, BER and SER are computed through Monte-Carlo simulations by taking samples corresponding to different realizations accounting for all possible data-dependent patterns.

When the pre-forward error correction (FEC) BER threshold is used as a performance metric, an estimate of information rate is obtained by multiplying the code rate, using a high-performance FEC code that can guarantee reliable communication after decoding, with raw transmission data rate. It is shown in [120] that

[2] We separate real and imaginary parts of $\mathbf{x}^{(k)}$ to capture correlations between them.

[3] Not to be confused with Q function used in digital communications. The Q-factor is a measure of the eye opening and is often used as a measure of signal-quality. For certain scenarios, there exists closed-form expressions relating Q-factor and BER [114, Eq. (56)], [17].

pre-FEC BER fails at predicting the post-FEC BER when binary soft-decision FEC is employed. Under some conditions, MI and generalized MI are better metrics than pre-FEC BER for predicting the post-FEC BER, and thereby the information rate [120–126].

### 4.4.2 Mutual Information

Shannon proved that reliable communication through a noisy channel is possible with channel coding, as long as the information rate is less than the channel capacity [127]. Reliable communication means that coding schemes exist that can make the probability of error arbitrarily small [128, 129]. The evaluation of the channel capacity for the fiber-optic channel accounting for the dispersion, nonlinearity, and noise is still an open problem due to the unavailability of an exact and mathematically tractable channel law, which is given by the conditional distribution of the output of the channel given the input of the channel. Therefore, accurately predicting the capacity of the fiber-optic channel has been the focus of much recent research [2, 52, 130–133].

For a fixed input distribution, the MI gives a lower bound on the channel capacity. The most commonly used approach to lower-bound the MI is to approximate the channel and compute the rate based on the receiver that is optimal for this approximate channel [124, 134–139]. This rate is achievable by that receiver is often referred to as an achievable information rate (AIR) [136, 140, 141].

## 4.5 Comparison of SDBP and DBP

We have simulated all three variations of SDBP across different symbol rates (14 GBd, 28 GBd, and 56 GBd), modulation formats (QPSK, 16-QAM), different DCM modules (DCF, FBG), and NDM. We have observed that in all these cases, the performance of SDBP in terms of SER is better than DBP. There is a different optimal power obtained in each of these algorithms, which is higher than that obtained from DBP. SER of VA-SDBP was better than SBS-SDBP algorithm. SER of GMP-SDBP was better than SBS-SDBP algorithm. The gains over DBP of all three variations of SDBP are higher for DM links than NDM links. This indicates that for NDM systems, the loss of performance when using DBP is smaller. This result corroborates the result from [17, 19, 26, 142] which quantifies the gains in handling NSNI for both DM and NDM links. Gains of SDBP can be explained as follows. The larger the deviation of the particle clouds of the the signal passed through an MF and a sampler, from a circular symmetric Gaussian, the higher are the expected gains in SDBP compared to DBP. For a DM link, we have observed that the particle clouds are less circularly Gaussian and hence SDBP performs better than DBP.

The gains of SDBP for DCF and FBG as DCM modules were observed to follow similar trend. The gains of 16-QAM format were seen higher than that of QPSK. As the symbol rate increases, the gains of SBS-SDBP over DBP has decreased. However, this trend was not seen in the case of VA-SDBP. The effects

of NSNI changes depending on whether an EDC is used to study the impact as in [17, 19, 26, 142] or whether a nonlinear equalizer like DBP is used. Part of the reason is because the intra-channel effects are compensated for when DBP is used, which is not the case when EDC is used. Therefore, different effects dominate depending on whether EDC or DBP or SDBP is used as compensating algorithm.

As the input power is increased, SER of SDBP also increases but at a different optimal power than DBP. PDFs associated with the particles after MF are shown in paper A for 28 Gbaud, 44 spans of 80 km SMF each, 16-QAM, and FBG link. It was shown that a multivariate Gaussian distribution is often a good approximation especially for low input powers. For some symbols (e.g., at constellation point $3 + j$), the histogram-based and Gaussian PDFs do not fit, which means that distributions other than multivariate Gaussian are needed to achieve the optimal performance.

The complexity of SDBP scales as $(N_p \times N \times M \times \mathcal{C}_{\mathrm{DBP,M}})$, where $\mathcal{C}_{\mathrm{DBP,M}}$ is the complexity of the DBP algorithm per segment of a fiber span. The complexity of SDBP can be reduced by reducing $N_p$, $M$, $\mathcal{C}_{\mathrm{DBP,M}}$. The performance as a function of $N_p$ and $M$ is presented in paper A for SBS-SDBP but the conclusions hold also for VA-SDBP and GMP-SDBP. There is a decoding complexity that is different in all these three variations of SDBP. For SBS-SDBP, computation of mean and covariance for each symbol has to be performed and then a MAP estimate has to be taken for each symbol. The computational cost for decoding in SBS-SDBP, $\mathcal{C}_{\mathrm{dec}}$, scales as $K \times N_p + K \times |\Omega|$. For VA-SDBP, the main complexity in decoding comes from the VA and the complexity scales as $|\Omega|^{L}_{\mathrm{mem}}$, where $L_{\mathrm{mem}}$ is the number of symbols in the state of the VA accounting for residual memory. For GMP-SDBP, the main complexity is in finding the inverse of the covariance matrix.

## 4.6 Summary

Nonlinearity compensation algorithms in both optical and digital domain were presented, with an emphasis on DBP. SDBP with three different variations, SBS-SDBP, VA-SDBP, and GMP-SDBP, were then explained. Performance metrics used in this thesis was explained, and finally a comparison of SDBP with DBP was presented.

# Chapter 5

# Contributions and Future Work

## 5.1 Paper A

Starting with a high-level overview of the existing nonlinear compensation algorithms, DBP was presented, which is seen as a universal technique for jointly compensating linear and nonlinear impairments. Then we moved on to describe the principles behind our proposed detector of paper A, starting with Bayesian inference, factor graphs, the sum-product algorithm, and particle representation of the messages. We introduced the notion of global and local functions and how the factorization can be pictorially represented using an FG. Then SPA was introduced, which can be applied to the FG, and a marginal can be found. We noted that the messages represent scaled probability density functions and since a closed-form expression of the messages is difficult, we used a particle representation for the messages. In particular, messages are approximated with a list of samples in paper A.

In paper A, we applied the FG framework for the fiber-optic channel of Fig. 2.5. The joint distribution $f$ introduced in Sec. 3.2 is the joint distribution of the variables involved in each segment and span of the SSFM representation of the fiber and the EDFAs. The local functions correspond to the CD and nonlinear blocks within each segment, an EDFA within each span, and a pulse shaper at the transmitter. Each of these correspond to one local function.

In this work, we extend the MAP-based detector for a single channel [143] system to account also for dispersive effects. The proposed detector is based on the MAP criterion and compensates not only for linear and nonlinear effects but also takes the noise from the amplifiers into account. As a consequence, nonlinear signal–noise interactions can be handled using the proposed detector. This allows us to (i) get closer to the fundamental performance limits of the fiber-optic channel and (ii) identify regimes where DBP is not optimal. Our proposed near-MAP detector turns out to be a generalization of DBP, and hence we call the

method *stochastic digital backpropagation*. In this thesis, we use the word SDBP algorithm to mean reversing the effect of only the fiber-link of Fig. 4.1. Given the distribution in particle form, the reversing of the pulse shape operation had to be done at the output of SDBP. In this paper, we used a matched filtering following by sampling operation for this operation, and hence this approach is named as symbol-by-symbol SDBP in this thesis. However, in paper A, the usage of SDBP was including the matched filtering operation also. We then studied the influence of PMD on DBP and SDBP and observed performance deterioration for both DBP and SDBP with increasing DGD but SDBP maintains a performance gain over DBP which decreases as DGD value increases. Complexity analysis of SDBP in terms of DBP complexity was studied. The effect of number particles in SDBP with respect to SER performance was also reported.

Contributions: NVI derived and analyzed the SDBP detector, carried out simulations, and wrote the paper. HW formulated the problem and contributed to the analysis. PJ provided optical communications expertise. All authors provided mathematical expertise, contributed to the interpretation of the results, reviewed, and revised the paper.

## 5.2 Paper B

In this paper, we compared SBS-SDBP with algorithms proposed in [8] which has a VA after DBP with two different metrics. The first metric, Cartesian Gaussian (CG), is derived according to the linearized regular perturbation model. In this model, the output signal is affected by nonlinear ISI and colored Gaussian noise. Thus, conditional on the transmitted symbols $\mathbf{s}$, the in-phase and quadrature components of the output samples $\mathbf{r}$ can be modeled as correlated Gaussian variables. The second metric, polar Gaussian (PG), is based on a more accurate model and obtained by accounting for the presence of nonlinear phase noise. In particular, conditional on the transmitted symbols $\mathbf{s}$, the amplitude and phase of the received samples are correlated Gaussian variables. The MLSD rule can be implemented by a VA (with CG or a PG metric), by using a suitable training sequence in order to estimate and store in a look-up table the required conditional expectations and covariance matrices for each transition of the trellis diagram. The SER of DBP-PG was lower than DBP-CG as was observed in [8]. For DM links, SBS-SDBP has worse performance than the VA after DBP as in the latter case, correlation of the samples is accounted for by using the VA. This study suggests that the correlation of the samples is an important aspect when accounting for the nonlinear signal–noise interactions. The VA can be improved by computing joint detection for both polarizations while SDBP can be improved by accounting for correlations among samples. The VA has exponential complexity with respect to the modulation order, whereas, the complexity for SDBP is essentially independent of the modulation format used.

Contributions: NVI and DM formulated the problem, carried out simulations, and wrote the paper. NVI, DM, HW, and MS performed the mathematical analysis.

PJ and MS provided optical communications expertise. All authors contributed to the interpretation of the results, reviewed, and revised the paper.

## 5.3 Paper C

The decisions in SBS-SDBP are taken on a symbol-by-symbol basis, ignoring any residual memory, which may be present due to nonoptimal processing in SBS-SDBP. In this paper, we extend SDBP to account for memory between symbols. In particular, two different methods, a VA and a decision directed approach, are applied at the output of MF and sampling stage as seen in Fig. 2.5. SER for memory-based SDBP is significantly lower than the previously proposed SBS-SDBP. For inline dispersion-managed links, the VA-SDBP has up to 10 and 14 times lower SER than DBP for QPSK and 16-QAM, respectively.

Contributions: NVI formulated the problem, carried out simulations, and wrote the paper. NVI, DM, and HW performed mathematical analysis. All authors contributed to the interpretation of the results, reviewed, and revised the paper.

## 5.4 Paper D

As MF operation is not guided by FG and SPA principles, we wanted to go back to this problem of reversing the effect of pulse shaping and derive an algorithm using SPA rule. Given the message $\overleftarrow{\mu}_{Q_2}(q_2)$ of Fig.3.3 in particle form, the problem is to find $\overleftarrow{\mu}_{Q_1}(q_1)$. This problem is addressed in this paper for 5 different building blocks out of which the important one that will be useful for SDBP corresponds to $f(q_1, q_2)$ being a pulse shaping operation. We used the linear Gaussian message passing idea from [60, Table III], and hence we call this approach GMP-SDBP in this thesis. In this approach, the message $\overleftarrow{\mu}_{Q_2}(q_2)$ before reversing pulse shape is approximated to be Gaussian. Then the mean and covariance for $\overleftarrow{\mu}_{Q_1}(q_1)$ is found using [60, Table III]. SER of GMP-SDBP was lower than SBS-SDBP.

Contributions: NVI and HW formulated the problem. IAS implemented the algorithm. IAS and NVI carried out simulations. HW wrote the paper. All authors provided mathematical expertise, contributed to the interpretation of the results, reviewed, and revised the paper.

## 5.5 Paper E

In papers A–D, SER was used as a performance metric. However, MI is shown to be a better metric than the pre-FEC BER for estimating the post-FEC BER in soft-decision FEC systems. Moreover, for the FOC, the state-of-the-art estimates are based on DBP. In this paper, we wanted to find out if tighter lower bounds on the MI can be obtained using SDBP instead of DBP as nonlinearity compensation. In other words, if we change MI as a performance metric than SER, can we see

a performance improvement using SDBP. To use output of SDBP for computing lower bounds on MI, we used the concept of auxiliary backward channel, which is used for the first time for the FOC scenario. We computed lower bounds on MI for SBS-SDBP and GMP-SDBP in comparison to the state-of-the-art method of using DBP. Through simulations, it was also found that up to 0.8 bit/symbol higher AIR is obtained using GMP-SDBP compared to DBP. This means that in comparison to the DBP approach, tighter lower bounds can be obtained using the SDBP approach.

Contributions: NVI and MS formulated the problem. NVI implemented the algorithm, carried out simulations, and wrote the paper. NVI, EA, and MS analyzed the model. All authors provided mathematical expertise, contributed to the interpretation of the results, reviewed, and revised the paper.

## 5.6 Future Work

Possible extensions on SDBP include comparing all three variants of SDBP with DBP-CG and DBP-PG in terms of SER and MI. This analysis can be firstly done for single-channel and extended to a super-channel scenario. The reason for the gains of SDBP over DBP, and thereby the impact of NSNI after DBP and SDBP, across different symbol rates, modulation formats is an interesting future direction. Experimental validation of SDBP needs to be performed to see the gains compared to the simulations.

On a theoretical level, one interesting direction is to explore ways of incorporating other stochastic impairments, such as PMD, into the FG framework and develop an optimal detector. The complexity of SDBP is currently very high and one direction is to reduce the complexity of these algorithms. Approximation of particles is currently performed using a Gaussian distribution, and as noted in paper A, distributions other than Gaussian should be considered especially at high input powers.

# References


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

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

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

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

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

[6] P. Serena, "Nonlinear signal-noise interaction in optical links with nonlinear equalization," *Journal of Lightwave Technology*, vol. 34, no. 6, pp. 1476–1483, 2016.

[7] G. Liga, A. Alvarado, E. Agrell, M. Secondini, R. I. Killey, and P. Bayvel, "Optimum detection in presence of nonlinear distortions with memory," in *Proceedings of European Conference on Optical Communication (ECOC)*, 2015, p. P.4.13.

[8] D. Marsella, M. Secondini, and E. Forestieri, "Maximum likelihood sequence detection for mitigating nonlinear effects," *Journal of Lightwave Technology*, vol. 32, no. 5, pp. 908–916, 2014.

[9] P. J. Winzer, "Modulation and multiplexing in optical communication systems," *LEOS Newsletter*, pp. 4–10, 2009.

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

[11] ——, *Nonlinear Fiber Optics*, 4th ed. Academic Press, 2006.

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

[13] K. Hill and G. Meltz, "Fiber Bragg grating technology fundamentals and overview," *Journal of Lightwave Technology*, vol. 15, no. 8, pp. 1263–1276, 1997.

[14] R. H. Stolen and A. Ashkin, "Optical Kerr effect in glass waveguide," *Applied Physics Letters*, vol. 22, no. 6, pp. 294–296, 1973.

[15] R. Olshansky, "Noise figure for erbium-doped optical fibre amplifiers," *Electronics Letters*, vol. 24, no. 22, pp. 1363–1365, 1988.

[16] M. Secondini, E. Forestieri, and C. R. Menyuk, "A combined regular-logarithmic perturbation method for signal-noise interaction in amplified optical systems," *Journal of Lightwave Technology*, vol. 27, no. 16, pp. 3358–3369, 2009.

[17] A. Bononi, P. Serena, and N. Rossi, "Nonlinear signal-noise interactions in dispersion-managed links with various modulation formats," *Optical Fiber Technology*, vol. 16, no. 2, pp. 73–85, 2010.

[18] L. Beygi, N. V. Irukulapati, E. Agrell, P. Johannisson, M. Karlsson, H. Wymeersch, P. Serena, and A. Bononi, "On nonlinearly-induced noise in single-channel optical links with digital backpropagation," *Optics Express*, vol. 21, no. 22, pp. 26 376–26 386, 2013.

[19] D. Foursa, O. Sinkin, A. Lucero, J.-X. Cai, G. Mohs, and A. Pilipetskii, "Nonlinear interaction between signal and amplified spontaneous emission in coherent systems," in *Proceedings of Optical Fiber Communication Conference (OFC)*, 2013, p. JTh2A.35.

[20] L. B. Du, D. Rafique, A. Napoli, B. Spinnler, A. D. Ellis, M. Kuschnerov, and A. J. Lowery, "Digital fiber nonlinearity compensation: toward 1-Tb/s transport," *IEEE Signal Processing Magazine*, vol. 31, no. 2, pp. 46–56, 2014.

[21] S. T. Le, M. E. McCarthy, S. K. Turitsyn, I. Phillips, D. Lavery, T. Xu, P. Bayvel, and A. D. Ellis, "Optical and digital phase conjugation techniques for fiber nonlinearity compensation," in *Opto-Electronics and Communications Conference (OECC)*, 2015.

[22] E. Temprana, N. Alic, B. P.-P. Kuo, and S. Radic, "Beating the nonlinear capacity limit," *Physics Today*, pp. 30–37, 2016.

[23] R. Dar and P. J. Winzer, "On the limits of digital back-propagation in fully loaded WDM systems," *Photonics Technology Letters*, vol. 28, no. 11, pp. 1253–1256, 2016.

[24] J. C. Cartledge, A. D. Ellis, A. Shiner, A. I. A. El-Rahman, M. E. McCarthy, M. Reimer, A. Borowiec, and A. Kashi, "Signal processing techniques for reducing the impact of fiber nonlinearities on system performance," in *Proceedings of Optical Fiber Communication Conference (OFC)*, 2016, p. Th4F.5.

[25] M. A. Z. Al-Khateeb, M. E. McCarthy, C. Sánchez, and A. D. Ellis, "Effect of second order signal-noise interactions in nonlinearity compensated optical transmission systems," *Optics Letters*, vol. 41, no. 8, pp. 1849–1852, 2016.

[26] N. Rossi, P. Serena, and A. Bononi, "Symbol-rate dependence of dominant nonlinearity and reach in coherent WDM links," *Journal of Lightwave Technology*, vol. 33, no. 14, pp. 3132–3143, 2015.

[27] R.-J. Essiambre, G. Raybon, and B. Mikkelsen, "Pseudo-linear transmission of high-speed TDM signals," *Optical Fiber Telecommunications IV-B*, vol. IV, pp. 232–304, 2002.

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

[29] A. D. Ellis, S. T. Le, M. E. McCarthy, and S. K. Turitsyn, "The impact of parametric noise amplification on long haul transmission throughput," in *International Conference on Transparent Optical Networks (ICTON)*, 2015.

[30] A. D. Ellis, M. E. McCarthy, M. A. Z. Al-Khateeb, and S. Sygletos, "Capacity limits of systems employing multiple optical phase conjugators." *Optics Express*, vol. 23, no. 16, pp. 20 381–20 393, 2015.

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

[32] C. B. Czegledi, G. Liga, D. Lavery, M. Karlsson, E. Agrell, S. J. Savory, and P. Bayvel, "Polarization-mode dispersion aware digital backpropagation," in *Proceedings of European Conference on Optical Communication (ECOC)*, 2016.

[33] J. Gordon and L. Mollenauer, "Effects of fiber nonlinearities and amplifier spacing on ultra-long distance transmission," *Journal of Lightwave Technology*, vol. 9, no. 2, pp. 170–173, 1991.

[34] P. Serena, A. Bononi, J.-C. Antona, and S. Bigo, "Parametric gain in the strongly nonlinear regime and its impact on 10-Gb/s NRZ systems with forward-error correction," *Journal of Lightwave Technology*, vol. 23, no. 8, pp. 2352–2363, 2005.

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

[36] E. Ciaramella and E. Forestieri, "Analytical approximation of nonlinear distortions," *Photonics Technology Letters*, vol. 17, no. 1, pp. 91–93, 2005.

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

[38] H. B. Song and M. Brandt-Pearce, "A 2-D discrete-time model of physical impairments in wavelength-division multiplexing systems," *Journal of Lightwave Technology*, vol. 30, no. 5, pp. 713–726, 2012.

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

[40] K.-P. Ho, *Phase-Modulated Optical Communication Systems*. New York: Springer-Verlag, 2005.

[41] A. Demir, "Nonlinear phase noise in optical-fiber-communication systems," *Journal of Lightwave Technology*, vol. 25, no. 8, pp. 2002–2032, 2007.

[42] L. Beygi, E. Agrell, M. Karlsson, and P. Johannisson, "Signal statistics in fiber-optical channels with polarization multiplexing and self-phase modulation," *Journal of Lightwave Technology*, vol. 29, no. 16, pp. 2379–2386, 2011.

[43] A. G. Green, P. P. Mitra, and L. G. L. Wegener, "Effect of chromatic dispersion on nonlinear phase noise," *Optics Letters*, vol. 28, no. 24, pp. 2455–2457, 2003.

[44] P. Serena, A. Orlandini, and A. Bononi, "Parametric-gain approach to the analysis of single-channel DPSK/DQPSK systems with nonlinear phase noise," *Journal of Lightwave Technology*, vol. 24, no. 5, pp. 2026–2037, 2006.

[45] G. Bosco, A. Carena, R. Cigliutti, V. Curri, P. Poggiolini, and F. Forghieri, "Performance prediction for WDM PM-QPSK transmission over uncompensated links," in *Proceedings of Optical Fiber Communication Conference (OFC)*, 2011, p. OThO7.

[46] A. Carena, G. Bosco, V. Curri, P. Poggiolini, M. T. Taiba, and F. Forghieri, "Statistical characterization of PM-QPSK signals after propagation in uncompensated fiber links," in *Proceedings of European Conference on Optical Communication (ECOC)*, 2010, p. P4.07.

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

[48] L. Beygi, E. Agrell, P. Johannisson, M. Karlsson, and H. Wymeersch, "A discrete-time model for uncompensated single-channel fiber-optical links," *IEEE Transactions on Communications*, vol. 60, no. 11, pp. 3440–3450, 2012.

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

[50] P. Poggiolini, "The GN model of non-linear propagation in uncompensated coherent optical systems," *Journal of Lightwave Technology*, vol. 30, no. 24, pp. 3857–3879, 2012.

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

[52] E. Agrell, G. Durisi, and P. Johannisson, "Information-theory-friendly models for fiber-optic channels: A primer," in *Information Theory Workshop (ITW)*, 2015.

[53] R. A. Fisher and W. K. Bischel, "Numerical studies of the interplay between self-phase modulation and dispersion for intense plane-wave laser pulses," *Journal of Applied Physics*, vol. 46, no. 11, pp. 4921–4934, 1975.

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

[55] R. G. Gallager, *Low Density Parity Check Codes*. MIT Press, 1963.

[56] G. D. Forney, Jr, "The Viterbi algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–278, 1973.

[57] R. Tanner, "A recursive approach to low complexity codes," *IEEE Transactions on Information Theory*, vol. 27, no. 5, pp. 533–547, 1981.

[58] N. Wiberg, "Codes and Decoding on General Graphs," PhD thesis, Linköping University, Sweden, 1996.

[59] H. Wymeersch, *Iterative Receiver Design*. Cambridge University Press, 2007.

[60] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Signal Processing Magazine*, vol. 21, no. 1, pp. 28–41, 2004.

[61] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, 2001.

[62] A. Hasegawa, "An historical review of application of optical solitons for high speed communications," *Chaos*, vol. 10, no. 3, pp. 475–485, 2000.

[63] S. Wahls and H. V. Poor, "Introducing the fast nonlinear Fourier transform," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2013.

[64] M. I. Yousefi and F. R. Kschischang, "Information Transmission Using the Nonlinear Fourier Transform, Part I: Mathematical Tools," *IEEE Transactions on Information Theory*, vol. 60, no. 7, pp. 4312–4328, 2014.

[65] S. T. Le, J. E. Prilepsky, and S. K. Turitsyn, "Nonlinear inverse synthesis technique for optical links with lumped amplification," *Optics Express*, vol. 23, no. 7, pp. 8317–8328, 2015.

[66] S. Hari, M. I. Yousefi, and F. R. Kschischang, "Multieigenvalue Communication," *Journal of Lightwave Technology*, vol. 34, no. 13, pp. 3110–3117, 2016.

[67] A. Yariv, D. Fekete, and D. M. Pepper, "Compensation for channel dispersion by nonlinear optical phase conjugation." *Optics Letters*, vol. 4, no. 2, pp. 52–54, 1979.

[68] R. A. Fisher, B. R. Suydam, and D. Yevick, "Optical phase conjugation for time-domain undoing of dispersive self-phase-modulation effects," *Optics Letters*, vol. 8, no. 12, pp. 611–613, 1983.

[69] C. Paré, A. Villeneuve, P.-A. Bélanger, and N. J. Doran, "Compensating for dispersion and the nonlinear Kerr effect without phase conjugation," *Optics Letters*, vol. 21, no. 7, pp. 459–461, 1996.

[70] S. Kumar and D. Yang, "Optical backpropagation for fiber-optic communications using highly nonlinear fibers." *Optics Letters*, vol. 36, no. 7, pp. 1038–1040, 2011.

[71] X. Liang, S. Kumar, and J. Shao, "Ideal optical backpropagation of scalar NLSE using dispersion-decreasing fibers for WDM transmission," *Optics Express*, vol. 21, no. 23, pp. 52–54, 2013.

[72] X. Liu, A. R. Chraplyvy, P. J. Winzer, R. W. Tkach, and S. Chandrasekhar, "Phase-conjugated twin waves for communication beyond the Kerr nonlinearity limit," *Nature Photonics*, vol. 7, no. 7, pp. 560–568, 2013.

[73] S. L. I. Olsson, B. Corcoran, C. Lundstrom, T. A. Eriksson, M. Karlsson, and P. A. Andrekson, "Phase-sensitive amplified transmission links for improved sensitivity and nonlinearity tolerance," *Journal of Lightwave Technology*, vol. 33, no. 3, pp. 710–721, 2015.

[74] H. Eliasson, "Mitigation of nonlinear fiber distortion using phase conjugation," Chalmers University of Technology, Tech. Rep., 2016.

[75] J. Winters and R. Gitlin, "Electrical signal processing techniques in long-haul, fiber-optic systems," in *International Conference on Communications (ICC)*, 1990, pp. 397–403.

[76] H. F. Haunstein, K. Sticht, A. Dittrich, M. Lorang, W. Sauer-Greff, , and R. Urbansky, "Implementation of near optimum electrical equalization at 10 Gbit/s," in *Proceedings of European Conference on Optical Communication (ECOC)*, 2000.

[77] M. Taylor, "Coherent detection method using DSP for demodulation of signal and subsequent equalization of propagation impairments," *Photonics Technology Letters*, vol. 16, no. 2, pp. 674–676, 2004.

[78] R. Killey, P. Watts, V. Mikhailov, M. Glick, and P. Bayvel, "Electronic dispersion compensation by signal predistortion using digital Processing and a dual-drive Mach-Zehnder Modulator," *Photonics Technology Letters*, vol. 17, no. 3, pp. 714–716, 2005.

[79] H. Bülow, F. Buchali, and A. Klekamp, "Electronic dispersion compensation," *Journal of Lightwave Technology*, vol. 26, no. 1, pp. 158–167, 2008.

[80] R. Killey, "Dispersion and nonlinearity compensation using electronic predistortion techniques," in *Proc. IEE Seminar on Fibre Communications Electronic Signal Processing*, 2005.

[81] R.-J. Essiambre and P. J. Winzer, "Fibre nonlinearities in electronically predistorted transmission," in *Proceedings of European Conference on Optical Communication (ECOC)*, 2005, p. Tu 3.2.2.

[82] K. Roberts, C. Li, L. Strawczynski, M. O'Sullivan, and I. Hardcastle, "Electronic precompensation of optical nonlinearity," *Photonics Technology Letters*, vol. 18, no. 2, pp. 403–405, 2006.

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

[84] X. Li, X. Chen, G. Goldfarb, E. Mateo, I. Kim, F. Yaman, and G. Li, "Electronic post-compensation of WDM transmission impairments using coherent detection and digital signal processing," *Optics Express*, vol. 16, no. 2, pp. 880–888, 2008.

[85] D. Rafique, M. Mussolin, M. Forzati, J. Mårtensson, M. N. Chugtai, and A. D. Ellis, "Compensation of intra-channel nonlinear fibre impairments using simplified digital back-propagation algorithm," *Optics Express*, vol. 19, no. 10, pp. 9453–9460, may 2011.

[86] T. Hoshida, L. Dou, W. Yan, L. Li, and Z. Tao, "Advanced and feasible signal processing algorithm for nonlinear mitigation," in *Proceedings of Optical Fiber Communication Conference (OFC)*, 2013, p. OTh3C.3.

[87] E. M. Ip, E. Mateo, and T. Wang, "Reduced-complexity nonlinear compensation based on equivalent-span digital backpropagation," in *International Conference on Optical Internet (COIN)*, 2012, pp. 28–29.

[88] W. Yan, Z. Tao, L. Dou, and L. Li, "Low complexity digital perturbation back-propagation," in *Proceedings of European Conference on Optical Communication (ECOC)*, 2011, p. Tu.3.A.2.

[89] L. B. Du and A. J. Lowery, "Improved single channel backpropagation for intra-channel fiber nonlinearity compensation in long-haul optical communication systems," *Optics Express*, vol. 18, no. 16, pp. 17 075–17 088, 2010.

[90] D. S. Millar, S. Makovejs, C. Behrens, S. Hellerbrand, R. I. Killey, P. Bayvel, and S. J. Savory, "Mitigation of fiber nonlinearity using a digital coherent receiver," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 5, pp. 1217–1226, 2010.

[91] L. Zhu and G. Li, "Nonlinearity compensation using dispersion-folded digital backward propagation," *Optics Express*, vol. 20, no. 13, pp. 14 362–14 370, jun 2012.

[92] Y. Gao, J. H. Ke, J. C. Cartledge, K. P. Zhong, and S. S.-H. Yam, "Implication of parameter values on low-pass filter assisted digital back propagation for DP 16-QAM," *Photonics Technology Letters*, vol. 25, no. 10, pp. 917–920, 2013.

[93] S. Chandrasekhar, X. L. X. Liu, B. Zhu, and D. W. Peckham, "Transmission of a 1.2-Tb/s 24-carrier no-guard-interval coherent OFDM superchannel over 7200-km of ultra-large-area fiber," in *Proceedings of European Conference on Optical Communication (ECOC)*, 2009.

[94] S. Randel, P. J. Winzer, A. Sureka, S. Chandrasekhar, N. Fontaine, R. Delbue, R. Ryf, P. Pupalaikis, and X. Liu, "Fiber nonlinearity compensation by digital backpropagation of an entire 1.2-Tb/s superchannel using a full-field spectrally-sliced receiver," in *Proceedings of European Conference on Optical Communication (ECOC)*, 2013.

[95] T. Fehenberger and N. Hanik, "Digital back-propagation of a superchannel: achievable rates and adaption of the GN model," in *Proceedings of European Conference on Optical Communication (ECOC)*, 2014, p. We.3.3.6.

[96] X. Liu and S. Chandrasekhar, "Superchannel for next-generation optical networks," in *Proceedings of Optical Fiber Communication Conference (OFC)*, 2014, p. W1H.5.

[97] J.-X. Cai, H. Zhang, H. G. Batshon, M. Mazurczyk, O. V. Sinkin, D. G. Foursa, A. N. Pilipetskii, G. Mohs, and N. S. Bergano, "200 Gb/s and dual wavelength 400 Gb/s transmission over transpacific distance at 6.0 b/s/Hz spectral efficiency," *Journal of Lightwave Technology*, vol. 32, no. 4, pp. 832–839, 2014.

[98] E. P. da Silva, K. J. Larsen, and D. Zibar, "Impairment mitigation in superchannels with digital backpropagation and MLSD," *Optics Express*, vol. 23, no. 23, pp. 29 493–29 501, 2015.

[99] R. Maher, T. Xu, L. Galdino, M. Sato, A. Alvarado, K. Shi, S. J. Savory, B. C. Thomsen, R. I. Killey, and P. Bayvel, "Spectrally shaped DP-16QAM

super-channel transmission with multi-channel digital back-propagation," *Scientific Reports*, vol. 5, p. 8214, 2015.

[100] D. S. Millar, R. Maher, D. Lavery, T. Koike-Akino, M. Pajovic, A. Alvarado, M. Paskov, K. Kojima, K. Parsons, B. C. Thomsen, S. J. Savory, and P. Bayvel, "Design of a 1 Tb/s superchannel coherent receiver," *Journal of Lightwave Technology*, vol. 34, no. 6, pp. 1453–1463, 2016.

[101] T. Rahman, D. Rafique, B. Spinnler, S. Calabro, E. de Man, U. Feiste, A. Napoli, M. Bohn, G. Khanna, N. Hanik, E. Pincemin, C. Le Bouette, J. Jauffrit, S. Bordais, C. Andre, C. Dourthe, B. Raguenes, C. M. Okonkwo, A. M. J. Koonen, and H. de Waardt, "Long-haul transmission of PM-16QAM-, PM-32QAM-, and PM-64QAM-based terabit superchannels over a field deployed legacy fiber," *Journal of Lightwave Technology*, vol. 34, no. 13, pp. 3071–3079, 2016.

[102] F. P. Guiomar, S. B. Amado, R. M. Ferreira, J. D. Reis, S. M. Rossi, A. Chiuchiarelli, J. R. F. De Oliveira, A. L. Teixeira, and A. N. Pinto, "Multicarrier Digital Backpropagation for 400G Optical Superchannels," *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1896–1907, 2016.

[103] P. Bayvel, R. Maher, T. Xu, G. Liga, N. A. Shevchenko, D. Lavery, A. Alvarado, and R. I. Killey, "Maximizing the optical network capacity," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2062, 2016.

[104] E. Giacoumidis, S. Mhatli, T. Nguyen, S. T. Le, I. Aldaya, M. E. McCarthy, A. D. Ellis, and B. J. Eggleton, "Comparison of DSP-based nonlinear equalizers for intra-channel nonlinearity compensation in coherent optical OFDM," *Optics Letters*, vol. 41, no. 11, pp. 2509–2512, 2016.

[105] L. Galdino, M. Tan, A. Alvarado, D. Lavery, P. Rosa, R. Maher, J. D. Ania-Castanon, P. Harper, S. Makovejs, B. C. Thomsen, and P. Bayvel, "Amplification schemes and multi-channel DBP for unrepeatered transmission," *Journal of Lightwave Technology*, vol. 34, no. 9, pp. 2221–2227, 2016.

[106] Y. Cai, D. G. Foursa, C. R. Davidson, J. Cai, O. Sinkin, M. Nissov, and A. Pilipetskii, "Experimental demonstration of coherent MAP detection for nonlinearity mitigation in long-haul transmissions," in *Proceedings of Optical Fiber Communication Conference (OFC)*, 2010, p. OTuE1.

[107] T. Koike-Akino, K. Parsons, K. Kojima, C. Duan, T. Yoshida, T. Sugihara, and T. Mizuochi, "Fractionally-spaced statistical equalizer for fiber nonlinearity mitigation in digital coherent optical systems," in *Proceedings of Optical Fiber Communication Conference (OFC)*, 2012, p. OTh3C.3.

[108] A. Rezania and J. C. Cartledge, "Transmission performance of 448 Gb/s single-carrier and 1.2 Tb/s three-carrier superchannel using dual-polarization 16-QAM with fixed LUT based MAP detection," *Journal of Lightwave Technology*, vol. 33, no. 23, pp. 4738–4745, 2015.

[109] J. Zhao and A. D. Ellis, "MAP detection for impairment compensation in coherent WDM systems," *Optics Express*, vol. 17, no. 16, pp. 13 395–13 401, 2009.

[110] N. V. Irukulapati, D. Marsella, P. Johannisson, M. Secondini, H. Wymeersch, E. Agrell, and E. Forestieri, "On maximum likelihood sequence detectors for single-channel coherent optical communications," in *Proceedings of European Conference on Optical Communication (ECOC)*, 2014.

[111] H.-A. Loeliger, J. Dauwels, J. Hu, S. Korl, L. Ping, and F. R. Kschischang, "The factor graph approach to model-based signal processing," *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1295–1322, 2007.

[112] S. Personick, "Receiver design for optical fiber systems," *Proceedings of the IEEE*, vol. 65, no. 12, pp. 1670–1678, 1977.

[113] G. L. Cariolaro, "Error probability in digital fiber optic communication systems," *IEEE Transactions on Information Theory*, vol. 24, no. 2, pp. 213–221, 1978.

[114] D. Marcuse, "Derivation of analytical expressions for the bit-error probability in lightwave systems with optical amplifiers," *Journal of Lightwave Technology*, vol. 8, no. 12, pp. 1816–1823, 1990.

[115] P. Humblet and M. Azizoglu, "On the bit error rate of lightwave systems with optical amplifiers," *Journal of Lightwave Technology*, vol. 9, no. 11, pp. 1576–1582, 1991.

[116] E. Forestieri, "Evaluating the error probability in lightwave systems with chromatic dispersion, arbitrary pulse shape and pre- and postdetection filtering," *Journal of Lightwave Technology*, vol. 18, no. 11, pp. 1493–1503, 2000.

[117] R. Holzlöhner, V. S. Grigoryan, C. R. Menyuk, and W. L. Kath, "Accurate calculation of eye diagrams and bit error rates in optical transmission systems using linearization," *Journal of Lightwave Technology*, vol. 20, no. 3, pp. 389–400, 2002.

[118] J. Zweck, I. T. Lima Jr., Y. Sun, A. O. Lima, C. R. Menyuk, and G. M. Carter, "Modeling receivers in optical communication systems with polarization effects," *Optics and Photonics News*, vol. 14, no. 11, pp. 30–35, 2003.

[119] O. V. Sinkin, V. S. Grigoryan, R. Holzlöhner, A. Kalra, J. Zweck, and C. R. Menyuk, "Calculation of error probability in WDM RZ systems in presence of bit-pattern-dependent nonlinearity and of noise," in *Proceedings of Optical Fiber Communication Conference (OFC)*, 2004.

[120] A. Alvarado, E. Agrell, D. Lavery, R. Maher, and P. Bayvel, "Replacing the soft-decision FEC limit paradigm in the design of optical communication

systems," *Journal of Lightwave Technology*, vol. 33, no. 20, pp. 4338–4352, 2015.

[121] L. Wan, S. Tsai, and M. Almgren, "A fading-insensitive performance metric for a unified link quality model," in *Wireless Communications and Networking Conference (WCNC)*, 2006.

[122] M. Franceschini, G. Ferrari, and R. Raheli, "Does the performance of LDPC codes depend on the channel?" *IEEE Transactions on Communications*, vol. 54, no. 12, pp. 2129–2132, 2006.

[123] K. Brueninghaus, D. Astely, T. Salzer, S. Visuri, A. Alexiou, S. Karger, and G.-A. Seraji, "Link performance models for system level simulations of broadband radio access systems," in *International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2005.

[124] A. Leven, F. Vacondio, L. Schmalen, S. T. Brink, and W. Idler, "Estimation of soft FEC performance in optical transmission experiments," *Photonics Technology Letters*, vol. 23, no. 20, pp. 1547–1549, 2011.

[125] L. Schmalen, A. Alvarado, and R. Rios-Müller, "Predicting the performance of nonbinary forward error correction in optical transmission experiments," in *Proceedings of Optical Fiber Communication Conference (OFC)*, 2016.

[126] T. Koike-akino, D. S. Millar, K. Parsons, and K. Kojima, "GMI-maximizing constellation design with Grassmann projection for parametric shaping," in *Proceedings of Optical Fiber Communication Conference (OFC)*, 2016.

[127] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3/4, pp. 379–423/623–656, 1948.

[128] R. Gallager, "A Simple Derivation of the Coding Theorem and Some Applications," *IEEE Transactions on Information Theory*, vol. 11, no. 1, pp. 3–18, 1965.

[129] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, U.S.A: Wiley, 2006.

[130] P. P. Mitra and J. B. Stark, "Nonlinear limits to the information capacity of optical fibre communications." *Nature*, vol. 411, no. 6841, pp. 1027–1030, 2001.

[131] A. Ellis, J. Zhao, and D. Cotter, "Approaching the non-linear Shannon limit," *Journal of Lightwave Technology*, vol. 28, no. 4, pp. 423–433, 2010.

[132] E. Agrell, A. Alvarado, G. Durisi, and M. Karlsson, "Capacity of a nonlinear optical channel with finite memory," *Journal of Lightwave Technology*, vol. 32, no. 16, pp. 2862–2876, 2014.

[133] G. Kramer, M. I. Yousefi, and F. R. Kschischang, "Upper bound on the capacity of a cascade of nonlinear and noisy channels," in *Information Theory Workshop (ITW)*, 2015.

[134] I. B. Djordjevic, B. Vasic, M. Ivkovic, and I. Gabitov, "Achievable information rates for high-speed long-haul optical transmission," *Journal of Lightwave Technology*, vol. 23, no. 11, pp. 3755–3763, 2005.

[135] G. Colavolpe, T. Foggi, A. Modenini, and A. Piemontese, "Faster-than-Nyquist and beyond: how to improve spectral efficiency by accepting interference." *Optics Express*, vol. 19, no. 27, pp. 26 600–26 609, 2011.

[136] M. Secondini, E. Forestieri, and G. Prati, "Achievable information rate in nonlinear WDM fiber-optic systems with arbitrary modulation formats and dispersion maps," *Journal of Lightwave Technology*, vol. 31, no. 23, pp. 3839–3852, 2013.

[137] T. Fehenberger, A. Alvarado, P. Bayvel, and N. Hanik, "On achievable rates for long-haul fiber-optic communications," *Optics Express*, vol. 23, no. 7, pp. 9183–9191, 2015.

[138] T. A. Eriksson, T. Fehenberger, P. A. Andrekson, M. Karlsson, N. Hanik, and E. Agrell, "Impact of 4D channel distribution on the achievable rates in coherent optical communication experiments," *Journal of Lightwave Technology*, vol. 34, no. 9, pp. 2256–2266, 2016.

[139] G. Liga, A. Alvarado, E. Agrell, and P. Bayvel, "Information rates of next-generation long-haul optical fiber systems using coded modulation," pp. 1–11, 2016. [Online]. Available: http://arxiv.org/abs/1606.01689

[140] A. Ganti, A. Lapidoth, and I. E. Telatar, "Mismatched decoding revisited: general alphabets, channels with memory, and the wide-band limit," *IEEE Transactions on Information Theory*, vol. 46, no. 7, pp. 2315–2328, 2000.

[141] D. M. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavčić, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Transactions on Information Theory*, vol. 52, no. 8, pp. 3498–3508, 2006.

[142] A. Bononi, N. Rossi, and P. Serena, "Transmission limitations due to fiber nonlinearity," in *Proceedings of Optical Fiber Communication Conference (OFC)*, 2011, p. OWO7.

[143] N. Jiang, Y. Gong, J. Karout, H. Wymeersch, P. Johannisson, M. Karlsson, E. Agrell, and P. Andrekson, "Stochastic backpropagation for coherent optical communications," in *Proceedings of European Conference on Optical Communication (ECOC)*, 2011, p. We.10.P1.81.