THESIS FOR THE DEGREE OF LICENTIATE OF ENGINEERING

Influence of Fiber Nonlinearity on the Capacity of Optical Channel Models

Kamran Keykhosravi

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Electrical Engineering
Chalmers University of Technology
Göteborg, Sweden, 2017

Influence of Fiber Nonlinearity on the Capacity of Optical Channel Models

Kamran Keykhosravi

Technical Report No. R009/2017
ISSN 1403-266X

This thesis has been prepared using LaTeX.

Department of Electrical Engineering
Chalmers University of Technology
SE-412 96 Göteborg, Sweden
Phone: +46 (0)31 772 1746
www.chalmers.se

Printed by Chalmers Reproservice
Göteborg, Sweden, September 2017

Abstract

The majority of today's global Internet traffic is conveyed through optical fibers. The ever-increasing data demands has pushed the optical systems to evolve from using regenerators and direct-direction receivers to a coherent multi-wavelength network. The future services like cloud computing and virtual reality will demand more bandwidth, so much so that the so called capacity-crunch is anticipated to happen in near future. Therefore, studying the capacity of the optical system is needed to better utilizing the existing fiber network.

The capacity of the dispersive and nonlinear optical fiber described by the nonlinear Schrödinger equation is an open problem. There is a number of lower bounds on the capacity which are mainly obtained based on the mismatched decoding principle or by analyzing simplified channels. These lower bounds either fall to zero at high powers or saturate. The question whether the fiber-optical capacity has the same behavior as the lower bounds at high power is still open as the only known upper bound increases with the power unboundedly.

In this thesis, we investigate the influence of the simplifying assumption used in some optical channel models on the capacity. To do so, the capacity of three different memoryless simplified models of the fiber-optical channel are studied. The results show that in the high-power regime the capacities of these models grow with different pre-logs, which indicates the profound impact of the simplifying assumptions on the capacity of these channels.

Next, we turn our attention to the demodulation process which is usually done by matched filtering and sampling. It is shown that by deploying a proper demodulation scheme the performance of optical systems can be improved substantially. Specifically, a two-user simplified memoryless WDM network is studied, where the effects of nonlinear distortion are considered in the model. It is shown that unlike matched filtering and sampling, with the optimal demodulator, the symbol error rate decreases to zero at high power.

Keywords: Fiber optics, perturbation, nonlinearity mitigation, channel capacity, information theory, achievable rate.

List of Publications

This thesis is based on the following publications:

[A] K. Keykhosravi, G. Durisi, and E. Agrell, "Bounds on the Capacity of Memoryless Simplified Optical Channel Models," submitted to *IEEE Trans. Inform. Theory.*, June 2017.

[B] K. Keykhosravi, E. Agrell, and G. Durisi, "Demodulation and Detection Schemes for a Memoryless Optical WDM Channel," submitted to *IEEE Trans. Commun.*, Aug. 2017.

Other publications by the author, not included in this thesis, are:

[C] K. Keykhosravi and E. Agrell, "A novel demodulation scheme for a memoryless optical interference channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT),* Aachen, Germany, June 2017.

[D] K. Keykhosravi, G. Durisi, and E. Agrell, "A tighter upper bound on the capacity of the nondispersive optical fiber channel," in *Proc. European Conference on Optical Communication (ECOC),* Gothenburg, Sep. 2017.

[E] K. Keykhosravi, H. Rastegarfar, and E. Agrell, "Multicast scheduling for optical data center switches with tunability constraints," in *Proc. IEEE International Conference on Computing, Networking and Communications (ICNC),* Silicon Valley, Jan 2017, pp. 308–312.

[F] K. Keykhosravi, H. Rastegarfar, and E. Agrell, "Multicast Scheduling of Wavelength-Tunable, Multiqueue Optical Data Center Switches," submitted to *J. Opt. Commun. Netw.*

[G] H. Rastegarfar, K. Keykhosravi, K. Szczerba, E. Agrell, L. LaComb, and M. Glick, "Optical circuit granularity impact in TCP-dominant hybrid data center networks," in *Proc. IEEE International Conference on Computing, Networking and Communications (ICNC),* Silicon Valley, Jan 2017, pp. 318–322.

Acknowledgments

First and foremost, I would like to thank my supervisor Prof. Erik Agrell, who introduced me to the world of optics, and my cosupervisor Giuseppe Durisi, who helped me to gain a deeper understanding about information theory. I have had many fruitful discussions with them and I have developed many new skills. I am grateful for the time they have spent to guide me and give me valuable feedbacks.

Many thanks to Agneta and Natasha for directing the administrative work. A special thanks goes to my office-mate and friend, Alireza, for his guidance and support. Also, I would like to thank Houman Rastegarfar for introducing me to the field of optical network. Thanks to all the seniors for creating such a friendly environment in the communication group. I am grateful to be a part of the E2-force group, where I have learned about many different aspects of optical communications from Cristian, Li, Alireza, Arni, Andreas, and Morteza. Also, I am thankful for the guidance that I have received from the FORCE group, specially from Prof. Magnus Karlsson and Prof. Pontus Johannisson. I would also like to thank all of my colleagues, former and current, in communication system group with whom I have had many valuable memories, in particular, Mohamad Ali, Naga, Wei, Behrooz, Amina, Mouhamed, Rahul, Keerthi, Anver, Jesper, Sven, Chao, Johan, Nil, Gabriel, Christopher, Katharina, Jessica, and Erik. I have used their guidance, experiences, and expertise in many instances during these years. Furthermore, I would like thank all my Iranian friends both in E2 and MC2 for their companionship in the past couple of years; in particular, Abbas, Sadegh, Parastoo, Abolfazl, Samar, Fatemeh, Mehrzad, Ramin, Navid, Pegah, Nastaran, Saeed, Maryam, Elham, Amir, Bita, Ahad, Sobhan, Mohammad, Mehdi, Ebrahim, Roham, Aidin & Maryam.

It would not have been possible for me to pursue my studies without the support of my parents and my brother, to whom I am indebted for their sacrifices. Last but not least, I would like to thank my loving wife Parisa who was always on my side and supported me.

Kamran Keykhosravi
Gothenburg, April 2017

Acronyms

ASE:	amplified spontaneous emission
AWGN:	additive white Gaussian noise
DBP:	digital backpropagation
EDFA:	erbium-doped fiber amplifier
FWM:	four-wave mixing
LP:	logarithmic perturbation
LPC:	logarithmic perturbative channel
MAP:	maximum a posteriori
MNC:	memoryless NLS channel
NLS:	nonlinear Schrödinger
OPC:	optical phase conjugation
pdf:	probability distribution function
ROADM:	reconfigurable optical add-drop multiplexers
RP:	regular perturbation
RPC:	regular perturbative channel
SMF:	single-mode fiber
SNR:	signal-to-noise ratio
SPM:	self-phase modulation
SSF:	split-step Fourier
WDM:	wavelength-division multiplexing
XPM:	cross-phase modulation

Contents

Part I

Overview

CHAPTER 1

Introduction

1.1 Overview

Today's modern society relies on fast and reliable communication systems that connect the world together. The massive amount of data produced every day is transmitted almost exclusively through the optical fibers that connect the cities, countries, and continents. The two most significant advantages of optical fibers compared to other communication media are providing enormous bandwidth and having extremely low loss, which allows the fiber to transfer data at high rates to far destinations. Comparing to coaxial cable, the loss is 2–3 orders of magnitude less and the bandwidth is about 4000 times more. Fibers are also immune to the electro-magnetic and radio-wave interference, resistant to lighting strikes, and lighter and smaller than copper cables.

The optical-fiber communication system was born by manufacturing low-loss fibers in 1970 [1]. The history of the optical communication is divided to three eras [2, Sec. 2]. In the first era, from 1977–1993, on–off keying and direct-detection was used to transmit the data through the fiber. Also, at this stage, to compensate for the fiber loss regenerators were deployed along the fiber, which recover the data in the electric domain and retransmit it over the fiber. With the introduction of the optical amplifiers in the late 1980s [3] the second era began. Transmitting multiple wavelength-devision multiplexed (WDM) channels over a single fiber became possible as all the channels could be amplified with a single optical amplifier. Finally, we are now in a new phase where using modern digital signal processors makes it possible to move from direct detection to coherent transmission systems.

With all the aforementioned advances, the optical communication systems were able to

cope with the increasing data demand in the past five decades which is mainly attributed to the exponential growth of the Internet. The call for higher data rates is expected to grow even faster in the future as the bandwidth-hungry applications like Internet of things, cloud computing, and virtual reality are beginning to emerge. It is anticipated that the current optical systems will meet their limit in the near future, an incident that is referred to as the capacity crunch.

To avoid the capacity crunch, it is essential to exploit all the potentials of the optical fiber. In today's fiber-optical transmissions, the coding, modulation, and demodulation algorithms are adopted from wireless systems. However, the optical fiber is a nonlinear medium as the intensity of the light changes the refractive index of the fiber core. Since the transmission system is designed for the linear channel, the nonlinearity is treated as noise. This so called nonlinear noise increases with the intensity of the signal such that the performance of the system degrades after an optimal input power, causing the achievable rates with commonly used transmission schemes to have a maximum at low powers. An interesting open question is whether this peaky behavior of the achievable rate is fundamental or is merely because of deploying suboptimal algorithms. Information-theoretic analysis of the capacity is a standard way to obtain insights about the answer.

After about five decades from the birth of optical fibers and seven decades from the dawn of information theory [4], the capacity of the fiber-optical channel remains an open problem. The main question is whether the capacity increases to infinity or is bounded. To answer this question, either an ever-increasing lower bound or a bounded upper bound on the capacity is required. However, no such bounds are established as yet. All the lower bounds either vanish or saturate at high powers where the only known upper bound [5] increases indefinitely with power. To obtain approximations of the capacity, many simplified models have been developed and lower bounds were established on their capacity (see [6, 7] for two recent reviews).

Contributions: In this thesis we investigate the influence of the simplifying assumptions, used in developing simple optical channels, on capacity. We show that these assumptions have a profound influence on the capacity of optical models at high power, and therefore the capacity results based on these models are not accurate in the high-power regime. Next, the matched filtering and sampling which is the common demodulation method in optical systems is revisited. Multiple demodulation and detection schemes are proposed for a two-user memoryless WDM network. A significant gain in performance was observed with the optimal demodulation method, with which the symbol error rate decreases to zero with power, whereas with matched filtering and sampling the symbol error rate goes to approximately one in the high-power regime.

1.2 Thesis Outline

This thesis is divided into two parts, where the first one includes some background and general information about the topic and the second comprises two papers whose contributions were summarized in the abstract. The rest of the first part is organized as follows. First, we study the NLS equation that governs the propagation of a signal through a fiber in Chapter 2. The main goal of Chapter 2 is to describe the continuous-time optical channels to justify the models used in the appended papers. In Chapter 3 different notions of capacity are investigated for continuous- and discrete-time channels. Some optical channel models are studied in Chapter 4. Specifically, the channel models used in the papers are described in Sections 4.3 and 4.4. A review on the nonlinearity mitigation methods is provided in Chapter 5. Finally, a summary of the contributions of the papers is provided in Section 6.

1.3 Notation

The following notation is used in the introduction. All logarithms are in base two. Vectors are denoted by underlined letters. $\|\cdot\|$ denotes the Euclidean norm. We use boldface letters to denote random quantities. $h(\,\cdot\,)$ represents the differential entropy functions. The mutual information between two random variables $\mathbf{x}$ and $\mathbf{y}$ is denoted by $I(\mathbf{x};\mathbf{y})$. The imaginary unit is represented by $j = \sqrt{-1}$.

CHAPTER 2

Fiber-Optical Transmission Systems

The purpose of this chapter is to connect the discrete-time optical channel models used in the appended papers to the underling physical channel. Also, it provides the reader with some general information about optical channel modeling.

Optical core networks, which are information highways connecting cities, countries, and continents, carry the vast amount of Internet data. To connect two nodes via these networks, reconfigurable optical add-drop multiplexers (ROADMs) are used to add or drop a channel at each switching point. Fig. 2.1 shows a simple schematic of how WDM channels are added and dropped at different nodes of the network. Therefore, through propagation of a signal, the neighboring channels, which introduce interference to the signal, may change at each switch. The first step to analyze this enormous and complicated network is to study the point-to-point fiber channel, which is the main concern of this thesis.

The standard single-mode fiber has an extremely low loss. However, for long transmission distances (> 100 km), it should be compensated for. Optical amplifiers boost the signal energy while adding noise to it. Apart from the loss and noise, there are two main impairments associated with fibers, namely, dispersion and nonlinearity. They are mainly caused by the dependence of the refractive index on wavelength and electromagnetic field, respectively.

This chapter begins by describing the signal propagation through the fiber in Sec. 2.1. The effects of optical amplification is added to the model in Sec. 2.2. A brief introduction to a numerical simulation of the fiber channel is presented in Sec. 2.3. Finally, in Sec. 2.4, WDM systems are described briefly.

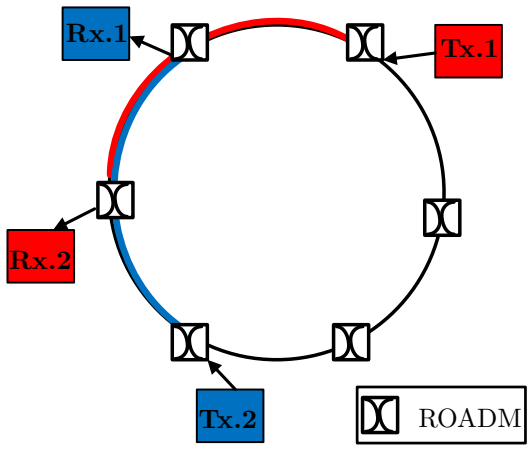

Figure 2.1: An schematic of an optical network, where switching is performed by ROADMs.

2.1 Signal Propagation in Fibers

An optical signal can be described by a slowly varying envelope of the light's electromagnetic field. Therefore, its propagation through the fiber is governed by Maxwell's equations. Under some assumptions, which are discussed later, these equations lead to the Manakov equation

$$\frac{\partial \underline{\mathbf{a}}}{\partial z} + \underbrace{j\frac{\beta_2}{2}\frac{\partial^2 \underline{\mathbf{a}}}{\partial t^2}}_{\text{Dispersion}} - \underbrace{j\gamma \left\|\underline{\mathbf{a}}\right\|^2 \underline{\mathbf{a}}}_{\text{Nonlinearity}} + \underbrace{\frac{\alpha}{2}\underline{\mathbf{a}}}_{\text{Attenuation}} = 0. \tag{2.1}$$

Here, $\underline{\mathbf{a}} = [\mathbf{a}_x(z,t), \mathbf{a}_y(z,t)]$, where $\mathbf{a}_x(z,t)$ and $\mathbf{a}_y(z,t)$ are complex baseband signals propagating in the x and y polarizations at time t and location z, respectively. β_2 is the group-velocity dispersion parameter and α is the attenuation constant. Parameter $\gamma = 2\pi n_2/(\lambda A_{\text{eff}})$ is the nonlinear coefficient, where n_2 is the nonlinear refractive index, A_{eff} is the effective area of the fiber core, and λ is the wavelength in vacuum.

We note that (2.1) describes the signal propagation in a single mode fiber (SMF) without any amplification and hence without noise. The effects of optical amplifiers are added to this equation in Sec. 2.2. If the signal is transmitted in only one polarization, (2.1) changes to the nonlinear Schrödinger (NLS) equation by replacing the random vector $\underline{\mathbf{a}}$ by the random variable $\mathbf{a}$ and also the vector norm $\|\cdot\|$ by the absolute value $|\cdot|$. As we mainly consider single polarization in this thesis, we use the NLS equation to

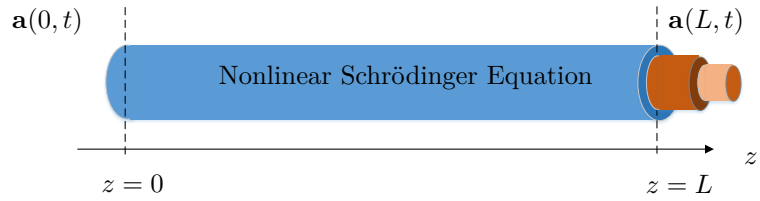

Figure 2.2: An schematic representation of signal propagation through the optical fiber governed by NLS equation.

describe the signal propagation in fibers. It reads

$$\frac{\partial \mathbf{a}}{\partial z} + j\frac{\beta_2}{2}\frac{\partial^2 \mathbf{a}}{\partial t^2} - j\gamma|\mathbf{a}|^2\mathbf{a} + \frac{\alpha}{2}\mathbf{a} = 0. \tag{2.2}$$

We note that the NLS equation is an excellent model for fibers and its validity has been checked by many experiments. However, it also should be noted that many assumptions have been considered in the derivation of this equation; here we list the most important ones.

- The signal's spectral width $\Delta\omega$ is much smaller than the central frequency ω_0: $\Delta\omega/\omega_0 \ll 1$
- The wavelength region $0.5 - 2\mu$m is utilized for transmission
- The nonlinear effect is small
- The fiber-loss is low
- The refractive index of the core and cladding is invariant of spatial coordinates
- The nonlinear response is instantaneous
- The higher order nonlinearities and dispersion are weak

Therefore, the NLS equation may be inaccurate at very high powers or at very low or high wavelengths. Also, it does not cover all the impairments of the fiber. More information about these assumption can be found in [8, Ch. 2].

No analytical solution for NLS equation has been found. However, some numerical methods can be used to solve this equation with arbitrary accuracy; we discuss one of these methods in detail in Sec. 2.3.

2.1.1 Chromatic Dispersion

In a SMF, each frequency component of a signal propagates with a different velocity through the fiber. Therefore, some components of the signal arrive at the receiver sooner than others; which results in the broadening of the signal in time. Specifically, a pulse with spectral width $\Delta\omega$ broadens in time after propagating through a fiber with length L by approximately $L\beta_2\Delta\omega$.

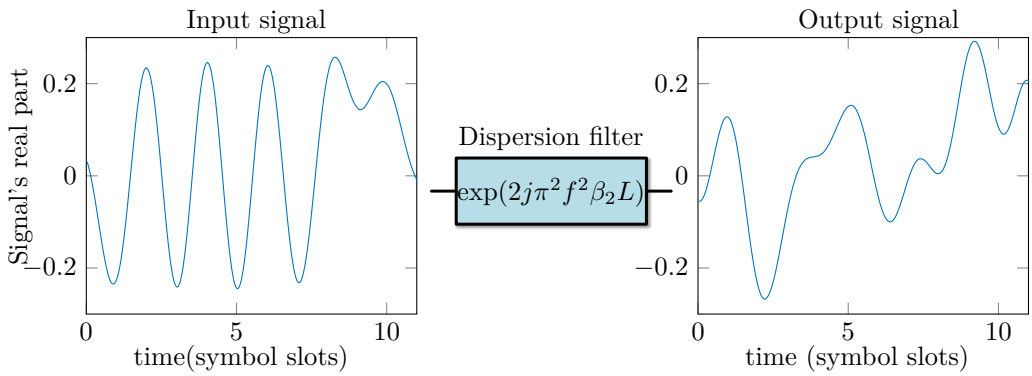

Figure 2.3: Effect of dispersion on a signal after a distance $L > 0$.

The effect of the dispersion on the transmitted signal can be studied by neglecting the nonlinearity ($\gamma = 0$) and the attenuation ($\alpha = 0$). Doing so, we obtain from (2.2)

$$\frac{\partial \mathbf{a}}{\partial z} + j\frac{\beta_2}{2}\frac{\partial^2 \mathbf{a}}{\partial t^2} = 0. \tag{2.3}$$

To solve (2.3), one can apply the Fourier transform to both sides of the equality to obtain

$$\frac{\partial \mathbf{a}(z,f)}{\partial z} = 2j\pi^2 f^2 \beta_2 \mathbf{a}(z,f). \tag{2.4}$$

Solving (2.4), the signal at position z can be expressed in the frequency domain as

$$\mathbf{a}(z,f) = \mathbf{a}(0,f)e^{2j\pi^2 f^2 \beta_2 z}. \tag{2.5}$$

Therefore, the chromatic dispersion can be modeled by a linear all-pass filter with frequency response $H(z,f) = \exp(2j\pi^2 f^2 \beta_2 z)$. The corresponding impulse response is

$$h(z,t) = \frac{\exp(jt^2/(2\beta^2 z))}{\sqrt{2j\pi\beta_2 z}}. \tag{2.6}$$

If not compensated for, dispersion severely deteriorates the performance of the optical systems. Fig. 2.3 illustrates the effect of dispersion on a signal after propagation. It can be seen that the signals becomes completely deformed. Two categories of dispersion compensation techniques exist. The first one is to use dispersion compensating fibers to compensate the dispersion in the optical domain. The second is done by utilizing digital signal processors at the receiver to perform the compensation in the electrical domain.

2.1.2 Kerr Nonlinearity

The presence of Kerr nonlinearity is the main difference between the optical communication systems and the linear wireless ones. Its origin comes from the dependence of the

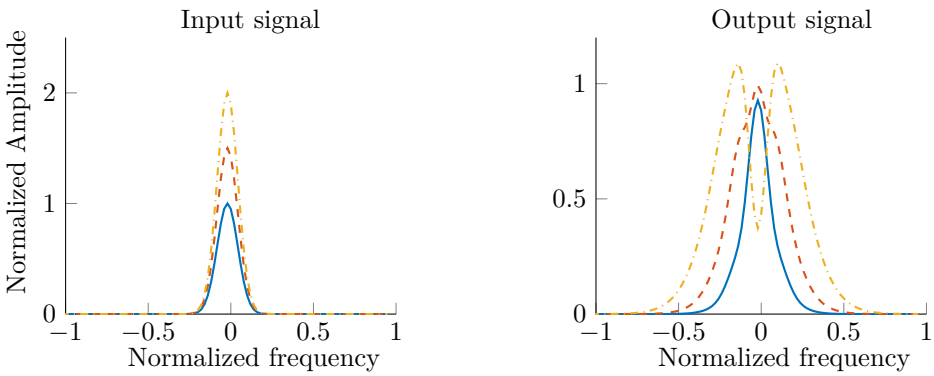

Figure 2.4: Effect of nonlinearity on three pulses with different amplitudes in the frequency domain after propagation.

refractive index on the electromagnetic field. Since the fiber core has a very small area, the intensity of the optical power launched into the fiber becomes very high. This high intensity changes the physical features of the glass like its refractive index. The effects of Kerr nonlinearity can be studied by neglecting dispersion $\beta_2 = 0$ and loss $\alpha = 0$ in (2.2) to obtain

$$\frac{\partial \mathbf{a}}{\partial z} = j\gamma|\mathbf{a}|^2\mathbf{a}. \tag{2.7}$$

This equation has the solution

$$\mathbf{a}(z,t) = \mathbf{a}(0,t)\exp(j\gamma|\mathbf{a}(0,t)|^2 z). \tag{2.8}$$

Therefore, the nonlinearity appears as a phase distortion that is a function of the signal power. The effects of the nonlinearity distortion can be studied in the frequency domain. Fig. 2.4 shows the evolution of three Gaussian pulsed in the frequency domain after propagation. It can be seen that the spectrum of the pulse broadens and this spectrum broadening increases with the amplitude of the signal. Many methods for compensating nonlinearity exist, some of which are presented in Sec. 5.

2.1.3 Fiber Loss

The fiber core is an extremely transparent glass. Its loss is approximately 0.2 dB/km at wavelength 1.55μm which is mainly due to a fundamental loss mechanism, Rayleigh scattering, and exists in all fibers [9, Sec. 2.5.3]. Neglecting dispersion and nonlinearity, the solution to (2.2) becomes

$$\mathbf{a}(z,t) = \mathbf{a}(0,t)\exp(-\alpha z/2) \tag{2.9}$$

which indicates that the power decreases with the rate $\exp(-\alpha z)$. The power loss is often stated in terms of dB/km which can be calculated from α by $\alpha_{\text{dB}} = \alpha 10\log_{10}(e)$.

2.1.4 Beyond the NLS Equation

As mentioned before, although the NLS equation covers the important features of signal propagation over fiber, it does not describe all the impairments. Here we briefly list some of these additional phenomena.

- *Stimulated Raman scattering* is an important nonlinear effect that occurs at high powers. It transfers a portion of an optical field's energy to a lower frequency. This feature of the fiber is utilized to build distributed Raman amplifiers where an optical beam is transmitted at higher frequencies than the signal to boost its energy continuously during propagation. This effect can also severely deteriorate the performance of WDM systems in the high-power regime by introducing inter-channel interference.
- *Stimulated Brillouin scattering* is a nonlinear process similar to the Raman effect. In this process some of the energy of an optical field is transferred to lower frequencies and propagates backward through the fiber.
- *Polarization effects:* In the derivation of the Manakov equation it is assumed that the state of the polarization is preserved completely during propagation. However, this is not the case in practice, where the refractive index for the x and y polarizations are not exactly the same. Therefore, the signal travels with different speeds in each polarization and causes a phenomenon called polarization mode dispersion.

2.2 Optical Amplifications

For long-haul transmission, in order to compensate for the loss, the signal needs to be amplified. Traditionally this process was performed by installing costly regenerators that brought the optical signal to the electrical domain and then retransmitted it. Now this task is done by optical amplifiers, which are more economical. Moreover, optical amplification is the key enabling technology to WDM systems. A single amplifier can boost the energy of optical signals in a wide spectrum of wavelengths, while with the old systems, one regenerator is needed for each WDM channel. The optical amplification is based on *stimulated emission*, where a copy of a photon is generated by an excited atom. This process is always accompanied with *spontaneous emission*, where an excited atom emits a photon randomly. Therefore, optical amplification always comes at the price of increased noise level. By considering the optical amplification, the signal propagation through the fiber can be described by the statistic NLS equation as

$$\frac{\partial \mathbf{a}}{\partial z} + j\frac{\beta_2}{2}\frac{\partial^2 \mathbf{a}}{\partial t^2} - j\gamma|\mathbf{a}|^2\mathbf{a} + \frac{\alpha - g(z)}{2}\mathbf{a} = \mathbf{n}. \tag{2.10}$$

Here, $\mathbf{n}$ is the amplified spontaneous emission (ASE) noise, which is circularly symmetric white Gaussian, and $g(z)$ is the gain profile. Next, we present two commonly used amplification schemes and study the noise and $g(z)$ for them.

Table 2.1: Amplifier parameters

Parameter	Symbol	Value
Amplifier Gain	G	$e^{\alpha L}$
Span length	L	80–100 km
Noise figure	F_n	4–7 dB
Number of spans	K	≥ 1
Planck's constant	h	$6.626 \cdot 10^{-34}$ Js
Carrier frequency	ν	$1.911 \cdot 10^{14} - 1.961 \cdot 10^{14}$ Hz

2.2.1 Lumped Amplifiers

In this amplification scheme the fiber is divided to multiple spans and an erbium-doped fiber amplifier (EDFA) is used at the end of each span to boost the signal's energy. The noise, $\mathbf{n}(z,t)$ has the autocorrelation of

$$\mathrm{E}[\mathbf{n}(z,t)\mathbf{n}(z',t')^*] = \frac{1}{2}GF_n h\nu\delta(t-t')\delta(z-z')\sum_{i=1}^{K}\delta(z-iL). \tag{2.11}$$

Here, $\delta(\,\cdot\,)$ is the Dirac delta function and $h\nu$ is the optical photon energy. The rest of the parameters are presented in Table 2.1. The gain profile can be described as

$$g(z) = \alpha L\sum_{i=1}^{K}\delta(z-iL). \tag{2.12}$$

If we assume that the signal bandwidth, W, is constant during propagation, the total noise variance is

$$P_N = \frac{1}{2}KGF_n h\nu W \tag{2.13}$$

.

2.2.2 Distributed Amplifiers

Unlike lumped amplification, in the distributed one the transmission loss is continuously compensated for along propagation. Therefore, the signal power level remains almost constant. In this scheme, a pump wave is transmitted at higher frequencies than that of the signal. The pump copropagates with the signal and through the Raman effect, it gives the signal a portion of its energy. The gain profile for the ideal Raman amplification is $g(z) = \alpha$ and the noise autocorrelation is

$$\mathrm{E}[\mathbf{n}(z,t)\mathbf{n}(z',t')^*] = \alpha n_{sp} h\nu\delta(z-z')\delta(t-t') \tag{2.14}$$

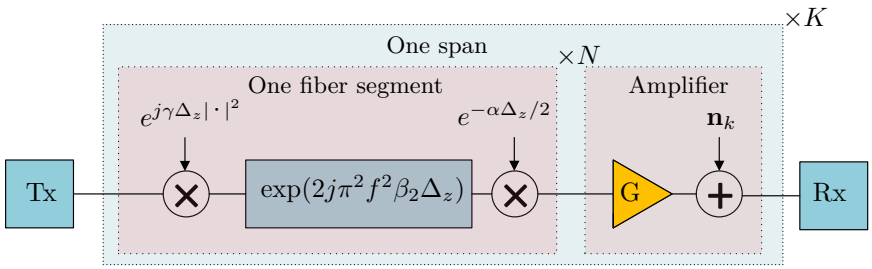

Figure 2.5: Split-step Fourier method for lumped amplification with K spans. Each fiber span is divided to N segments. $G = e^{N\alpha\Delta_z/2}$ is the amplification gain.

where n_{sp} is the ASE noise factor. The total noise variance is

$$P_N = \alpha L_t n_{sp} h\nu W \tag{2.15}$$

where L_t is the total fiber length. We note that, in practice, the gain profile is not a constant and it decreases as the distance from the amplifier increases.

2.3 Split-step Fourier Method

For special cases of the input signal such as solitons [8, Ch. 5], the lossless NLS equation ((2.2) with $\alpha = 0$) has an analytical solution. However, a general solution for the NLS equation in (2.2) has not been found. The evolution of the signal can be tracked by means of numerical methods, of which the most famous one is the split-step Fourier (SSF) method.

2.3.1 Lumped Amplification

In the SSF method, for lumped amplified systems, a fiber span with length L is split into N small segments with length $\Delta z = L/N$. It is assumed that the linear and nonlinear operators in the NLS equation act independently in each segment. The output of the kth segment, $\mathbf{a}_k = \mathbf{a}(k\Delta z, t)$, can be obtained by applying three operators to the output of its previous segment, $\mathbf{a}_{k-1}$: *i)* a nonlinear operator as in (2.8) *ii)* a linear operator as in (2.5), and *iii)* an attenuation operator as in (2.9), i.e.,

$$\mathbf{a}_k = e^{-\alpha\Delta z/2}\left[h(\Delta z, t) * \left[\mathbf{a}_{k-1}\exp(j\gamma\Delta z|\mathbf{a}_{k-1}|^2)\right]\right]. \tag{2.16}$$

Here, $h(z,t)$ is the impulse response of the dispersion filter defined in (2.6).

In practice, to implement the SSF method, the input signal is sampled much faster than the Nyquist rate so that the effects of spectrum broadening, caused by the channel nonlinearity, is correctly captured. Moreover, to decrease the complexity of the method,

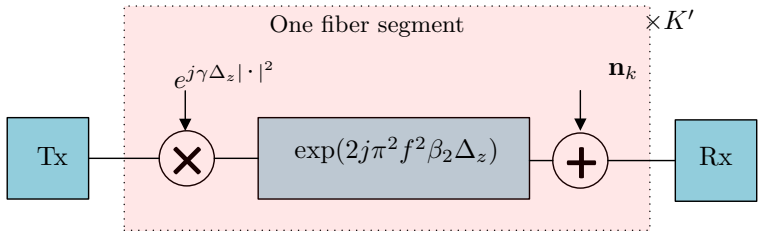

Figure 2.6: Split-step Fourier method for distributed amplification with K' segments.

the convolution in (2.16) is calculated in the frequency domain using fast Fourier transform.

For lumped amplification, the effect of amplifier can be shown by multiplying the signal by $\exp(\alpha L/2)$ and adding the ASE noise to the signal. A model of the SSF method for lumped amplification is provided in Fig. 2.5.

If dispersion is ignored, we have $|\mathbf{a}_k|^2 = \exp(-k\alpha\Delta z)|\mathbf{a}_0|$. Therefore, the output of the first span can be described by

$$\mathbf{a}_N = \mathbf{a}_0 e^{-\alpha L/2} \prod_{k=0}^{N-1} \exp\left(j\gamma\Delta z e^{-k\alpha\Delta z}|\mathbf{a}_0|^2\right) \tag{2.17}$$

$$= \mathbf{a}_0 e^{-\alpha L/2} \exp\left(j\gamma|\mathbf{a}_0|^2\Delta z \frac{1-e^{-\alpha L}}{1-e^{-\alpha\Delta z}}\right). \tag{2.18}$$

Using $\lim_{\Delta z\to 0} \Delta z/(1-\exp(-\alpha\Delta z)) = 1/\alpha$, the right-hand side of (2.18) at the limit of $\Delta z \to 0$ becomes

$$\mathbf{a}_0 e^{-\alpha L/2} \exp(j\gamma L_{\text{eff}}|\mathbf{a}_0|^2) \tag{2.19}$$

where L_{eff} is the effective length of the fiber with length z and is defined as

$$L_{\text{eff}} = \frac{1-\exp(-\alpha z)}{\alpha}. \tag{2.20}$$

Note that $L_{\text{eff}} \leq z$, which indicates that the nonlinear phase shift is smaller than in the lossless case in (2.8). Similar channel model as in (2.19) is obtained in Sec. 4.2.

2.3.2 Distributed Amplification

For ideal distributed amplification, the fiber loss is compensated for completely and ASE noise is added to the signal continuously during propagation. One step of the SSF method with distributed amplification can be described as

$$\mathbf{a}_{k+1} = h(\Delta z, t) * [\mathbf{a}_k \exp(j\gamma\Delta z|\mathbf{a}_k|^2)] + \mathbf{n}_k \tag{2.21}$$

where $\mathbf{n}_k$ is the amplifier noise and has the autocorrelation of

$$\mathrm{E}[\mathbf{n}_k(z,t)\mathbf{n}_k(z',t')^*] = \frac{\alpha n_{sp} h\nu}{K'}\delta(z-z')\delta(t-t'). \tag{2.22}$$

This method is illustrated in Fig. 2.6.

2.4 WDM Systems and Impairments Related to It

In WDM systems, multiple modulated signals at different carrier wavelengths are transmitted through a single fiber. The number of channels is usually limited by the bandwidth of the optical amplifier used in these systems, which for EDFAs is roughly 5 THz. Therefore, around 100 channels with bandwidth of 50 GHz can be transmitted. Spectrum guard bands are applied to reduce the crosstalk between channels at the cost of reducing spectral efficiency. The origin of the cross-talk is the dependence of the refractive index at each wavelength on the signal amplitudes at the other wavelengths.

To model the effect of the cross-talk, we consider a WDM channel with two wavelengths. The signal propagation through these channels in the fiber can be described by the pair of coupled equations [10, p. 274].

$$\frac{\partial \mathbf{a}_1}{\partial z} + j\frac{\beta_{21}}{2}\frac{\partial^2 \mathbf{a}_1}{\partial t^2} - j\gamma_1(|\mathbf{a}_1|^2 + 2|\mathbf{a}_2|^2)\mathbf{a}_1 + \frac{\alpha}{2}\mathbf{a}_1 = 0 \tag{2.23}$$

$$\frac{\partial \mathbf{a}_2}{\partial z} + j\frac{\beta_{22}}{2}\frac{\partial^2 \mathbf{a}_2}{\partial t^2} - j\gamma_2(|\mathbf{a}_2|^2 + 2|\mathbf{a}_1|^2)\mathbf{a}_2 + \frac{\alpha}{2}\mathbf{a}_2 + jd\frac{\partial \mathbf{a}_2}{\partial t} = 0. \tag{2.24}$$

Here, β_{2i} is the group velocity dispersion at wavelength $i \in \{1,2\}$ and γ_i is the nonlinear parameter. $\mathbf{a}_i$ corresponds to the ith baseband signal. The attenuation constant α is considered to be equal for both channels. The parameter d is the group velocity mismatch, that is

$$d = \frac{1}{v_{g2}} - \frac{1}{v_{g1}} \tag{2.25}$$

where v_{gi} is the group velocity. In (2.23) and (2.24), the time is measured according to a reference frame moving with the first signal. The term $j\gamma_1|\mathbf{a}_1|^2$ in (2.23) corresponds to the self-phase modulation (SPM), meaning that the amplitude of the signal modulates the signal's phase. The term $j\gamma_1|\mathbf{a}_2|^2$ in (2.23) is known as cross-phase modulation (XPM), which illustrates the interference between the copropagating signals. It can be seen that the phase shift of a signal from XPM is twice the amount of SPM.

The phase distortions caused by SPM and XPM, although they do not change the signal's amplitude, expand the signal's spectrum. The only exception is the perfect rectangular pulse shape, where ϕ_1^{nl} and ϕ_2^{nl} are constant during one symbol period.

Another important impairment of WDM systems is the four-wave mixing (FWM), where three signal at different wavelengths ω_1, ω_2, and ω_3 combine with each other to create a signal at $\omega_4 = \omega_1 + \omega_2 - \omega_3$. With many signals propagating in WDM, FWM can

be a major source of distortion. This effect can be mitigated by proper channel spacing. Other sources of distortion like Brillouin and Raman scattering can also deteriorate the performance of WDM systems.

CHAPTER 3

Channel Capacity

In this chapter, we study different notions of channel capacity. Differential entropy and mutual information are defined in Sec. 3.1. Sec. 3.2 provides the definition of the capacity of the discrete-time channels. The capacity and the spectral efficiency of the continuous-time AWGN channel is defined in Sec. 3.3. Sec. 3.4 reviews some results on the capacity of the fiber-optical channel.

3.1 Differential Entropy and Mutual Information

Consider two random vectors $\underline{\mathbf{x}}$ and $\underline{\mathbf{y}}$ that take values in $\mathbb{R}^n$ and are distributed according to the joint probability distribution function (pdf) $f_{\underline{\mathbf{x}},\underline{\mathbf{y}}}(\underline{x},\underline{y})$. The differential entropy of $\underline{\mathbf{x}}$ is defined as [11, Sec. 8]

$$h(\underline{\mathbf{x}}) = -\int_{\mathbb{R}^n} f_{\underline{\mathbf{x}}}(\underline{x}) \log\big(f_{\underline{\mathbf{x}}}(\underline{x})\big)\, \mathrm{d}\underline{x}. \tag{3.1}$$

Similarly, the joint differential entropy of $\underline{\mathbf{x}}$ and $\underline{\mathbf{y}}$ is

$$h(\underline{\mathbf{x}},\underline{\mathbf{y}}) = -\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f_{\underline{\mathbf{x}},\underline{\mathbf{y}}}(\underline{x},\underline{y}) \log\Big(f_{\underline{\mathbf{x}},\underline{\mathbf{y}}}(\underline{x},\underline{y})\Big)\, \mathrm{d}\underline{x}\, \mathrm{d}\underline{y}. \tag{3.2}$$

Moreover the mutual information between the random vectors $\underline{\mathbf{x}}$ and $\underline{\mathbf{y}}$ is defined as

$$I(\underline{\mathbf{x}};\underline{\mathbf{y}}) = h(\underline{\mathbf{y}}) + h(\underline{\mathbf{x}}) - h\big(\underline{\mathbf{x}},\underline{\mathbf{y}}\big)\,. \tag{3.3}$$

Let now $\underline{\mathbf{x}}$ take value in $\mathbb{C}^n$. We have

$$h(\underline{\mathbf{x}}) = h(\underline{\mathbf{x}}^r, \underline{\mathbf{x}}^i) \tag{3.4}$$

where $\underline{\mathbf{x}}^r$ and $\underline{\mathbf{x}}^i$ are the random vectors representing the real and imaginary parts of $\underline{\mathbf{x}}$, resp. Also, the mutual information between two complex random vectors can be defined by (3.3).

3.2 Capacity of Discrete-Time Channels

A discrete-time memoryless channel can be described by a conditional pdf $f_{\mathbf{y}|\mathbf{x}}(y \mid x)$, where $\mathbf{x}$ and $\mathbf{y}$ are complex random variables. The capacity, $\mathcal{C}^{\mathrm{DMC}}$, of this channel is the maximum rate in bits per channel use at which the information can be transfered through the channel with arbitrary low error probability. By Shannon's channel coding theorem [4], the capacity under the average power constraint $P > 0$, can be calculated as

$$\mathcal{C}^{\mathrm{DMC}} = \sup_{f_{\mathbf{x}}(x)} I(\mathbf{x};\mathbf{y}) \qquad \text{bits per channel use} \tag{3.5}$$

where the supremum is taken over all input distributions $f_{\underline{\mathbf{x}}}(\underline{x})$ such that $\mathrm{E}\big[|\mathbf{x}|^2\big] \leq P$.

Similarly, for a discrete-time block-memoryless communication channel described by the conditional pdf $f_{\underline{\mathbf{y}}|\underline{\mathbf{x}}}(\underline{y} \mid \underline{x})$, the capacity, $\mathcal{C}^{\mathrm{DBC}}$, under the average power constraint $P > 0$, is

$$\mathcal{C}^{\mathrm{DBC}} = \sup_{f_{\underline{\mathbf{x}}}(\underline{x})} I(\underline{\mathbf{x}},\underline{\mathbf{y}}) \qquad \text{bits per block .} \tag{3.6}$$

Here, the supremum is taken over all input distributions $f_{\underline{\mathbf{x}}}(\underline{x})$ such that $\mathrm{E}\big[||\underline{\mathbf{x}}||^2\big] \leq nP$, where n is the input vector length. Alternatively, we have

$$\mathcal{C}^{\mathrm{DBC}} = \frac{1}{n} \sup_{f_{\underline{\mathbf{x}}}(\underline{x})} I(\underline{\mathbf{x}},\underline{\mathbf{y}}) \qquad \text{bits per channel use.} \tag{3.7}$$

3.3 Capacity of the Continuous-Time AWGN Channel

The capacity of a continuous-time channel in bits per second is the maximum of the average number of information bits that can be transmitted through the channel during one second at an arbitrarily low error probability. For the AWGN channel, the base-band input–output relation can be written as $\mathbf{y}(t) = x(t) + \mathbf{n}(t)$, where $\mathbf{n}(t)$ is a complex circularly-symmetric white Gaussian process and $x(t)$ is the complex input signal that is band-limited to B Hz. The capacity, $\mathcal{C}^{\mathrm{AWGN}}$, of this channel under the power constraint $\lim_{T\to\infty} 1/T \int_0^T |x(t)|^2 \,\mathrm{d}t \leq P$ is [4]

$$\mathcal{C}^{\mathrm{AWGN}} = B \log\left(1 + \frac{P}{N_0}\right) \qquad \text{bits per second.} \tag{3.8}$$

For a continuous-time channel, spectral efficiency is measured in terms of bits per seconds per hertz by dividing its capacity to its bandwidth, B. For the AWGN channel

the spectral efficiency, SE^{AWGN}, is

$$SE^{\mathrm{AWGN}} = \log\left(1 + \frac{P}{N_0}\right) \qquad \text{bits per second per hertz.} \tag{3.9}$$

3.4 Known Results on the Capacity of Fiber-Optical Channel Models

The fiber-optical system described by NLS equation is a continuous-time channel with memory. There are two main difficulties in studying the spectral efficiency of this channel. The first is to obtain a discrete-time model. When linear modulation and demodulation are applied, because of the nonlinear distortion, the output of the demodulator will be a nonlinear function of all the input symbols, which is hard to formalize. Some efforts to obtain an approximation of the discrete-time channel have been made, which have lead to the perturbative models (see Sec. 4.1). The second is that because of nonlinearity, the spectrum of the signal changes as it propagates through the fiber, which makes it difficult to define the channel bandwidth.

A number of lower bounds on the capacity of a variety of optical channel models have been proposed, many of which are based on mismatched decoding [12](see for example [13–18]). On the other hand, only one upper bound is known as yet [5], which indicates that the capacity, in bits per channel use, of a discrete-time NLS channel cannot exceed that of the AWGN one. A recent review on the capacity results is available in [6].

There are different optical communication systems studied in the literature. An optical system can be lumped- or distributed-amplified and single- or dual-polarized. Also, WDM systems or a single channel can be considered. Furthermore, different models can be obtained using various simplifying assumptions. All these models may be studied with or without memory (dispersion). Here, we categorize some capacity results based on different types of channel models.

SSF: The SSF method has been used in the literature with two different approaches to obtain bounds on the capacity. It has been used as a channel model, which maps an input complex vector to an output vector after some iterative operations [5, 19] or it is used to obtain channel statistics by simulation [20] to evaluate lower bounds on mutual information [17, 21, 22]. In [5], it has been proved that the capacity, in bits per channel use, of the SSF channel is upper-bounded by $\log(1 + \mathrm{SNR})$, where SNR is the signal-to-noise ratio. The asymptotic behavior of a modified SSF channel is studied in [19]. In [17] different mismatched lower bounds based on different receivers have been evaluated using SSF simulation. An auxiliary backward channel is used in [21] to obtain a lower bound that is evaluated by the SSF method.

Memoryless NLS channel: In [23] a lower bound on the capacity of the single-user memoryless NLS channel has been developed. The asymptotic behavior of this lower bound is studied to show that the capacity grows unboundedly with power with a minimum pre-log of one-half. The capacity of the same channel is evaluated by simulation

in [24].

Perturbative models: Since analyzing the capacity of NLS channel is difficult, perturbative models are considered in the literature to obtain an approximation of the capacity (see Sec. 4.1 for more information about perturbation theory). In [25] the influence of FWM is studied on the capacity of a simplified perturbative WDM channel. In [26] the Volterra series is used to derive a WDM channel model. It is shown that if joint processing is possible, the effects of nonlinear distortion on the capacity region is minimal. In [16] a perturbative model is used to develop a closed-form mismatched lower bound on the capacity. To illustrate the effects of XPM mitigation techniques on the capacity, lower bounds on the capacity of a perturbative model is evaluated in [15, 27–29].

Finally, here we list some other capacity results. The capacity of a Gaussian noise model has been investigated in [30]. In [31, 32] lower bounds on the capacity are established through simulation of the fiber-optical channel, using multiple ring constellation. A modification of the Gaussian noise model [33–36] is studied in [37] to obtains lower bounds on its capacity. The peaky behavior of the lower bounds is not observed in this study as the lower bound saturates at high power. In [38], it has been proved that the capacity of discrete-time static point-to-point channels, including optical models, cannot decrease with power.

CHAPTER 4

Optical Channel Models

In this chapter, the channel models considered in the appended papers are reviewed. Moreover, we investigate the input–output relation of a discrete-time channel obtained by applying linear modulation and matched filtering and sampling demodulation to a nonlinear optical channel.

4.1 Perturbation Theory

For many nonlinear systems, a linearization around a working point can provide an approximation of the system's output. With the fiber-optical channel, this approach leads to perturbative channel models. The main assumption here is that the effects of nonlinear distortion are weak. Specifically, the solution to the NLS equation, with initial condition $\mathbf{a}(0,t)$, is approximated by

$$\mathbf{a}(z,t) \approx \mathbf{a}^L(z,t) + \Delta\mathbf{a}(z,t) \tag{4.1}$$

where $\mathbf{a}^L(z,t)$ is the solution to the NLS equation with $\gamma = 0$ and $\Delta\mathbf{a}(z,t)$ is a nonlinear perturbation. This approach of approximating the solution is commonly referred to as the regular perturbation (RP) method.

There are three main approaches to calculating $\Delta\mathbf{a}(z,t)$. The first one is to insert $\mathbf{a}^L(z,t)+\Delta\mathbf{a}(z,t)$ into the NLS equation and to neglect the nonlinear terms that include $\Delta\mathbf{a}(z,t)$ [39–41]. The second is to write the solution as a power series of the nonlinear parameter γ, i.e.,

$$\mathbf{a}(z,t) = \sum_{i=0}^{\infty} \gamma^i \mathbf{a}_i(z,t) \tag{4.2}$$

and then inserting (4.2) into the NLS equation to find out the signals $\mathbf{a}_i(z,t)$ [42,43]. Note that $\mathbf{a}_0(z,t) = \mathbf{a}^L(z,t)$. The third method is to use Volterra series, which approximates the NLS channel by a transfer function in a multi-dimensional frequency domain [26,44]. All these methods result in roughly similar channel models. In [43] it has been shown that the order n solution of the second method coincides with the order $2n+1$ result of Volterra series.

Now, we consider the channel model derived in [42, Sec. 3], where the second method is used to approximate the NLS equation. Let $\mathbf{b}_i(z,t) = \mathbf{a}_i(z,t)\exp(\alpha z/2)$ for all z. The first-order perturbative term, which is a cubic function of the optical field, can be calculated by

$$\mathbf{b}_1(z,t) = j\int_0^z (|\mathbf{b}_0(z,t)|^2\mathbf{b}_0(z,t)) \otimes h(z-\zeta,t)e^{-\alpha\zeta}\,\mathrm{d}\zeta. \tag{4.3}$$

where $h(\,\cdot\,,\cdot\,)$ is defined in (2.6). Here, we consider the memoryless case when $\beta_2 = 0$, which is the case in Paper A. If the channel memory is neglected, the linear solution, $\mathbf{b}_0(z,t)$, can be obtained by setting $\beta_2 = \gamma = 0$ in (2.2) as $\mathbf{b}_0(z,t) = \mathbf{a}(0,t)$. The first perturbative term $\mathbf{b}_1$ in the memoryless case can be calculated by substituting $h(z,t) = \delta(t)$ into (4.3) to obtain

$$\mathbf{b}_1(z,t) = j|\mathbf{a}(0,t)|^2\mathbf{a}(0,t)\int_0^z e^{-\alpha\zeta}\,\mathrm{d}\zeta \tag{4.4}$$

$$= jL_{\text{eff}}|\mathbf{a}(0,t)|^2\mathbf{a}(0,t) \tag{4.5}$$

where L_{eff} is defined in (2.20). Therefore, the first-order perturbation theory approximates the solution of the memoryless NLS equation as

$$\mathbf{a}(z,t) \approx \big(\mathbf{a}(0,t) + j\gamma L_{\text{eff}}|\mathbf{a}(0,t)|^2\mathbf{a}(0,t)\big)\, e^{-\alpha z/2}. \tag{4.6}$$

Another approach to obtain perturbative channel models is to use the logarithmic perturbation (LP) method to obtain a faster convergence [42, 45]. This method can be viewed as applying the regular perturbation technique to the logarithm of the field. The normalized solution $\mathbf{b}(z,t) = e^{\alpha z/2}\mathbf{a}(z,t)$ is written as

$$\mathbf{b} = \mathbf{b}_0 \exp\left(j\gamma\sum_{i=0}^{\infty}\gamma^i\mathbf{b}_i^{\text{lp}}\right) \tag{4.7}$$

where $\mathbf{b}_0 = \mathbf{a}^L\exp(\alpha z/2)$ and $\mathbf{b}_i^{\text{lp}}$ can be found by substituting (4.7) into the NLS equation. Doing so, it is obtained that [45, Eq. (10b)]

$$\mathbf{b}_0^{\text{lp}} = \frac{1}{\mathbf{b}_0}\int_0^z \big(|\mathbf{b}_0|^2\mathbf{b}_0\big) \otimes h(z-\zeta,t)e^{-\alpha\zeta}\,\mathrm{d}z \tag{4.8}$$

For the memoryless case, $h(z,t) = \delta(t)$, we obtain

$$\mathbf{b}_0(z,t) = \mathbf{a}(0,t) \tag{4.9}$$

$$\mathbf{b}_0^{\text{lp}}(z,t) = L_{\text{eff}}|\mathbf{a}(0,t)|^2. \tag{4.10}$$

Therefore, the first-order logarithmic perturbation results in the approximate solution of

$$\mathbf{a}(z,t) \approx \mathbf{a}(0,t)e^{-\alpha z/2}e^{j\gamma L_{\text{eff}}|\mathbf{a}(0,t)|^2}. \tag{4.11}$$

Observe that (4.6) and (4.11) are equal up to a first-order linearizion. A combination of regular and logarithmic approach is presented in [46] to obtain a more complete approximation.

4.2 A Memoryless Optical Channel

In this section we ignore the effects of dispersion to obtain a memoryless channel model. If the dispersion is set to zero $\beta_2 = 0$ in (2.2), we get

$$\frac{\partial \mathbf{a}}{\partial z} - j\gamma|\mathbf{a}|^2\mathbf{a} + \frac{\alpha}{2}\mathbf{a} = 0. \tag{4.12}$$

This equation can be solved analytically to obtain [8, Sec. 4.1.1]

$$\mathbf{a}(z,t) = \mathbf{a}(0,t)e^{-\alpha z/2}e^{j\gamma L_{\text{eff}}|\mathbf{a}(0,t)|^2} \tag{4.13}$$

where L_{eff} is defined in (2.20). To obtain a channel model for memoryless lumped amplified systems, we can use (4.13) to describe the signal propagation over a single span, before amplification. At the end of each span, amplifiers multiply the signal by $\exp(\alpha L/2)$ and add a Gaussian noise $\mathbf{n}$, described in the previous section. We denote the signal at the end of the kth span and after amplification by $\mathbf{a}_k = \mathbf{a}(kL,t)$ for $k = 0,\dots,K$. The output of the memoryless channel model, $\mathbf{a}_K$, can be obtained by iterating the following step for $k = 0,\dots,K-1$

$$\mathbf{a}_{k+1} = \mathbf{a}_k e^{j\gamma L_{\text{eff}}|\mathbf{a}_k|^2} + \mathbf{n}_k \tag{4.14}$$

where $\mathbf{n}_k$ is the noise added by the kth amplifier.

A similar channel model can be developed for a distributed amplification system by letting the number of spans, K', go to infinity for a fixed system length, L. The channel output, $\mathbf{a}_{K'}$ can be obtained by iterating the following equation for $k = 0,\dots,K'$

$$\mathbf{a}_{k+1} = \mathbf{a}_k e^{j\gamma L/K'|a_k|^2} + \mathbf{n}_k \tag{4.15}$$

where $\mathbf{n}_k$ describes the amplifier ASE noise discussed in (2.22). This model accurately describes the nondispersive NLS channel when $K' \to \infty$. A discrete-time version of (4.15) is studied in Paper A.

4.3 Channel Models in Paper A

As will be illustrated in Sec. 4.5.1, the effects of nonlinearity prevents the development of a tractable discrete-time model attained by linear filtering and sampling at the receiver. To

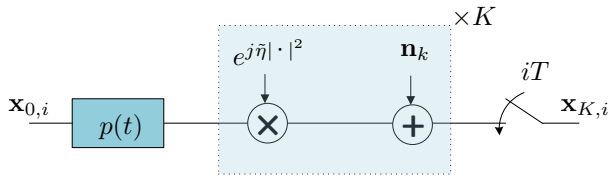

Figure 4.1: Discrete-time memoryless optical channel model with K spans of lumped amplification.

obtain discrete-time models from nonlinear continuous channels, in Paper A, we assume that samples are taken at the symbol rate at the receiver. It is assumed that the noise is band-limited because of the use of inline optical filters. Using these assumptions, the input–output relation of the discrete-time channels resembles that of the continuous-time ones. This method of developing discrete-time optical channels has been used extensively throughout the literature (see [24, 47, 48] for example). The importance of designing demodulation methods that are compatible with the nonlinear nature of the optical fiber is highlighted in Paper B.

In Paper A, three discrete-time optical channels were considered, which have been developed using different methods. All these models present the same fiber-optical channel under different assumptions. Although the simplifying assumptions used in these methods are only valid in the low-power regime, it is not uncommon to use these models to draw conclusions at high powers. We analyze the capacity of these models for the special case of zero dispersion to study the effects of the simplifying assumptions on the capacity. The results indicate that although these models represent the same physical channel, their capacity is drastically different in the high-power regime. Specifically, their capacities grow according to different pre-logs. Here, the capacity pre-log is defined as $\lim_{P\to\infty} \mathcal{C}/\log P$, where P is the average input power and $\mathcal{C}$ is the capacity of the channel.

All the three models in Paper A describe the same physical optical system and share the same set of parameters. A distributed-amplified optical system with length L is considered. The loss is assumed to be completely compensated for by the amplifiers. The discrete-time model is obtained from the continuous-time channel using the sampling receiver. The three considered channels can be found in the following list.

1. *Regular perturbative channel (RPC)* is a discrete-time distributed-amplified system with length L, which is based on the regular perturbation method. The signal–noise interaction is neglected and all the noise is added at the receiver. Therefore, the channel model can be obtained from (4.6) by letting $\alpha \to 0$ and $L_{\text{eff}} = L$ as

$$\mathbf{y} = \mathbf{x} + j\eta|\mathbf{x}|^2\mathbf{x} + \mathbf{n} \tag{4.16}$$

where the nonlinearity is quantified by $\eta = L\gamma$. The amplification noise is de-

noted by $\mathbf{n}$, which is complex circularly symmetric Gaussian with zero mean and variance P_N defined in (2.15).

2. *Logarithmic perturbative channel (LPC)* can be obtained from the LP channel in (4.11). It reads
$$\mathbf{y} = \mathbf{x}\exp(j\eta|\mathbf{x}|^2) + \mathbf{n} \tag{4.17}$$
where η and $\mathbf{n}$ are the same as those used in RPC.
3. *Memoryless NLS Channel (MNC)* is a discrete-time model of the continuous-time channel presented in (4.15). It can be described as a concatenation of K segments. Denoting the input of the MNC by $\mathbf{x}_0$, its output $\mathbf{x}_K$ can be obtained by iterating the following equation for $k = 0, \ldots, K-1$
$$\mathbf{x}_{k+1} = \mathbf{x}_k \exp(j\tilde{\eta}|\mathbf{x}_k|^2) + \mathbf{n}_k. \tag{4.18}$$
Here, $\tilde{\eta} = \eta/K$ and $\mathbf{n}_k$ are independent complex circularly symmetric Gaussian noise with zero mean and variance P_N/K. Fig. 4.1 illustrates the input–output relation in this model, where $\mathbf{x}_{0,i}$ and $\mathbf{x}_{K,i}$ denote the ith input and output symbol, respectively. This channel describes the signal propagation in the memoryless NLS channel without any simplification and has been studied extensively in the literature (see [47–49] for some examples).

4.4 Channel Model in Paper B

In Paper B, we consider a simplified model of a two-user WDM network, which has been introduced in Sec. 2.4. If the dispersion and group velocity mismatch in (2.23)–(2.24) are neglected ($d = \beta_{21} = \beta_{22} = 0$), these equations have the analytical solution
$$\mathbf{a}_1(L,t) = \mathbf{a}_1(0,t)\exp(-\frac{\alpha}{2})\exp\big(j\phi_1^{\mathrm{nl}}\big) \tag{4.19}$$
$$\mathbf{a}_2(L,t) = \mathbf{a}_2(0,t)\exp(-\frac{\alpha}{2})\exp\big(j\phi_2^{\mathrm{nl}}\big) \tag{4.20}$$
where
$$\phi_1^{\mathrm{nl}} = \gamma_1 L_{\mathrm{eff}}\left(|\mathbf{a}_1(0,t)|^2 + 2|\mathbf{a}_2(0,t)|^2\right) \tag{4.21}$$
$$\phi_2^{\mathrm{nl}} = \gamma_2 L_{\mathrm{eff}}(\underbrace{|\mathbf{a}_2(0,t)|^2}_{\mathrm{SPM}} + \underbrace{2|\mathbf{a}_1(0,t)|^2}_{\mathrm{XPM}}) \tag{4.22}$$
where L_{eff} is defined in (2.20). In Paper B, we consider a lumped amplified system with K spans, where a simplified model is obtained from (4.19) and (4.20) by multiplying the phase shifts, ϕ_1^{nl} and ϕ_2^{nl}, by K and adding all the amplification noise at the receiver.

A discrete-time network based on this model is studied in [50, 51]. We show that by using nonrectangular pulse shaping and proper demodulation technique, the XPM distortion can be effectively eliminated. However, when matched filtering and sampling is used as a demodulation scheme, because of SPM and XPM distortions, the error probability increases with power in the high-power regime.

4.5 From Continuous- to Discrete-Time Channel Models

In this section, we study two discrete-time channels based on applying the linear modulation and the matched filtering and sampling demodulation on *i)* the AWGN channel and *ii)* a nonlinear optical channel.

4.5.1 Matched Filtering and Sampling for a Linear Channel

Consider the complex-valued linear additive white Gaussian noise (AWGN) channel, where the input–output relation is described by

$$\mathbf{y}(t) = x(t) + \mathbf{n}(t). \tag{4.23}$$

Here, $\mathbf{n}(t)$ is a zero-mean circularly symmetric complex Gaussian noise with power spectral density N_0. With linear modulation, we have

$$x(t) = \sum_n x_n p(t - nT) \tag{4.24}$$

where x_n is a complex number and represents the information symbols generated by the source. The symbol period is given by T. Moreover, $p(t)$ is a real-valued function of time t and represents the pulse shape which is assumed to satisfy the orthonormality condition, i.e.,

$$\int_{-\infty}^{\infty} p(t)p(t - lT)\,\mathrm{d}t = \begin{cases} 1 & l = 0 \\ 0 & l \neq 0 \end{cases} \tag{4.25}$$

for all integers l. The transmitted power can be computed as

$$P = \lim_{T' \to \infty} \frac{1}{2T'} \int_{-T'}^{T'} |x(t)|^2\,\mathrm{d}t = \lim_{N \to \infty} \frac{1}{2N} \sum_{n=-N}^{N} |x_n|^2. \tag{4.26}$$

To detect the transmitted symbols, the receiver performs demodulation to obtain a complex number per each time interval T. It is well-known that matched filtering and sampling is the optimal demodulation method for the linear AWGN channel. The output of this method is obtained by calculating the inner product between the received signal and the time-shifted pulse shape, i.e.,

$$y_n = \int_{-\infty}^{\infty} y(t)p(t - nT)\,\mathrm{d}t \tag{4.27}$$

for all integers n. The output of this demodulator, y_n, also can be obtained by passing the signal through a filter with impulse response $p(-t)$ and sampling at time instances $t = nT$.

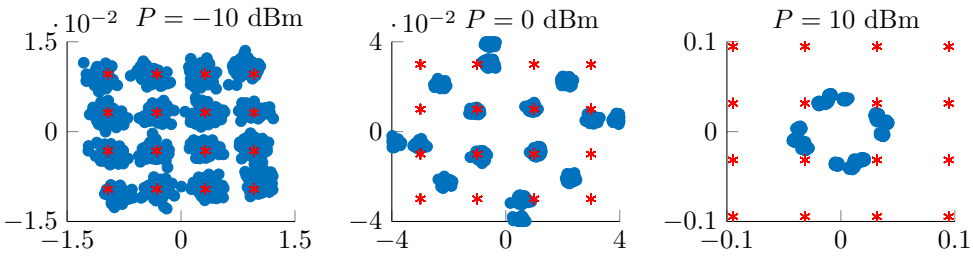

Figure 4.2: Scatter plots of the output of the matched filtering and sampling demodulator in the presence of Kerr nonlinearity described in (4.28).

4.5.2 Matched Filtering and Sampling for a Nonlinear Channel

Matched filtering and sampling provides sufficient statistics for the detection process [52, Thm. 26.4.1], i.e., given the set $\{y_k\}$, the received signal $y(t)$ becomes independent of the transmitted signal. However, this is not the case for nonlinear channels such as a fiber. Here, we study the impact of the fiber nonlinearity on the performance of matched filtering and sampling. We assume that the pulse shape $p(t)$ is zero outside the interval $[0, T)$ and only consider the effects of nonlinearity on the signal. In this case, the output of the demodulator for the first symbol is (see (2.8))

$$y_0 = x_0 \int_0^T p^2(t) e^{j\gamma L |x_0|^2 p^2(t)} \, \mathrm{d}t + \mathbf{n}. \tag{4.28}$$

where $\mathbf{n}$ is a complex Gaussian noise. The value of y_0 depends on the shape of $p(t)$. If $p(t)$ is a rectangular pulse shape, i.e., $p(t) = 1$, $0 \leq t \leq T$, we obtain from (4.28) that $y_0 = x_0 \exp(j\gamma L |x_0|^2)$. Therefore, the affects of nonlinearity appears as a phase shift in the discrete channel. In general, if $p(t)$ is not rectangular, the nonlinearity effects both the phase and the amplitude of the received signal. At high powers, where $|x_0|^2 \gg 1$, the phase of the integrand in (4.28) changes quickly with time, which scales down the result of the integral, i.e., $|y_0| < |x_0|$.

Fig. 4.2 is the scatter plot of the output of matched filtering and sampling demodulator described in (4.28). It can be seen that at the input power of -10 dBm, the effects of nonlinearity is minor. The nonlinear distortion manifests itself as a phase noise in the discrete-time channel at $P = 0$ dBm. By increasing the power to 10 dBm, it can be seen that the nonlinearity affects both the amplitude and phase of the output. Therefore, matched filtering and sampling is not a suitable demodulation method for the fiber channel at high powers.

CHAPTER 5

Nonlinearity Mitigation Methods

In this chapter, we review relevant results on nonlinearity mitigation methods in optical systems. We note that the demodulation technique presented in Paper B can be viewed as a nonlinearity mitigation method as it eliminates the effects of SPM and XPM nonlinear distortions.

The fiber nonlinearity is the central impairment of optical transmission systems that limits the achievable data rate. Here, we list a number of fiber-nonlinearity mitigation techniques that are implemented in the optical or digital domains, and also, we add our contribution at the end.

The inverse scattering transform [53,54] is among the first and yet most powerful nonlinearity mitigation techniques. It is based on the fact that some pulses, which are called solitons, can propagate through the lossless NLS channel without any distortion. Specifically, during the propagation of a soliton, the nonlinear effects are effectively canceled by the linear ones. Developing a communication system based on soliton transmission has received a great deal of attention in the recent years [55–58]. The performance of this approach can be significantly affected by random transmission effects such as polarization-mode dispersion and the fiber loss in lumped amplified systems [59].

Digital backpropagation (DBP) is another important nonlinearity compensation approach which is based on the SSF method. If the effects of noise are ignored, the deterministic distortion of the NLS channel can be completely compensated for using the SSF method with inverse parameters $(-\beta_2, -\gamma, -\alpha)$ at the receiver, or transmitter, or both [60–65]. DBP only compensates for the deterministic signal–signal distortions. To mitigate the effects of signal–noise interaction, stochastic DBP can be implemented [66].

Optical phase conjugation (OPC) was first suggested in [67] as a method for dispersion

compensation, where the phase of the signal is inverted at the middle of the transmission line. Here, we provide a brief justification of this method. The origin of chromatic dispersion is the difference between the group velocities of the high and low frequency components of the signal. After phase conjugation, the high-frequency components change their place with the low-frequency ones and therefore they travel with the same average speed, which mitigates the dispersion. In [68] it was pointed out that OPC can also be used as an effective way of compensating nonlinearity. Since then, OPC has received attention as an effective approach to reduce the effects of signal–noise interaction [69–72].

Nonlinearity-tailored detection techniques use an approximation of the optical input–output joint pdf in order to implement detectors based on the maximum a posteriori or the maximum likelihood criteria [73–76]. In this approach generally the signal is sampled at the receiver after passing a low-pass filter and the sample vector is fed to a Viterbi algorithm for detection. In [77] it is shown that by setting the bandwidth of the receiver filter larger than that of transmitted signal and oversampling the received signal, higher rates can be achieved compared to matched filtering.

XPM compensation by equalization: The XPM distortion varies slowly with time since it is the aggregation of the interference of many channels. Exploiting this property, the XPM can be mitigated by deploying adaptive or turbo equalization [78–80].

Our contribution in Paper B is to introduce a novel SPM and XPM mitigation method by using a novel demodulation technique. This method exploits the temporal correlation of the aforementioned distortions during one time slot, which has not been considered as yet. In one time slot, the distortion generated by SPM and XPM are correlated and dependent on the pulse shape. This property can be exploited to differentiate these distortions from noise and effectively mitigate their effect.

CHAPTER 6

Summaries of the Appended Papers

In this section we provide a summary of the two appended papers.

6.1 Paper A

In Paper A we investigate the accuracy of two perturbative models by studying their capacity for the zero-dispersion case. These models are the *Regular perturbative channel (RPC)* and the *Logarithmic perturbative channel (LPC)*. We also study the *Memoryless NLS Channel (MNC)* that is described in Section 4.3. Comparing their capacity, three different capacity pre-logs are established for these models: 3 for RPC, 1 for LPC, and 1/2 for MNC. This shows that the two perturbative channels, RPC and LPC, are grossly inaccurate in the high-power regime. Therefore, care should be exercised in interpreting the high-power results that have been established using these models.

6.2 Paper B

In Paper B we show that the effects of XPM can be effectively compensated for by exploiting the time coherence of the XPM distortion during one symbol period. This method of XPM mitigation is introduced for a two-user simplified memoryless network. A new demodulation scheme named maximum matching is introduced. Also, the optimal receiver, based on maximum a posteriori detection (MAP) is developed. The performance of these two receivers was compared with the matched filtering and sampling method, using the symbol error rate as a metric. Unlike with matched filtering and sampling,

the symbol error rate with maximum matching and MAP receiver goes to zero at high powers.

Bibliography

[1] F. Kapron, D. Keck, and R. Maurer, "Radiation losses in glass optical waveguides," *Appl. Phys. Lett.*, vol. 17, no. 10, pp. 423–425, Nov. 1970.

[2] E. Agrell, M. Karlsson, A. Chraplyvy, D. J. Richardson, P. M. Krummrich, P. Winzer, K. Roberts, J. K. Fischer, S. J. Savory, B. J. Eggleton *et al.*, "Roadmap of optical communications," *Journal of Optics*, vol. 18, no. 6, p. 063002, 2016.

[3] R. J. Mears, L. Reekie, I. M. Jauncey, and D. N. Payne, "Low-noise erbium-doped fibre amplifier operating at 1.54 μm," *Electron. Lett.*, vol. 23, no. 19, pp. 1026–1028, Sep. 1987.

[4] C. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, Jul.–Oct. 1948.

[5] G. Kramer, M. I. Yousefi, and F. R. Kschischang, "Upper bound on the capacity of a cascade of nonlinear and noisy channels," in *IEEE Info. Theory Workshop (ITW)*, Jerusalem, Israel, Apr-May 2015.

[6] M. Secondini and E. Forestieri, "Scope and limitations of the nonlinear Shannon limit," *J. Lightw. Technol.*, vol. 35, no. 4, pp. 893–902, Apr. 2017.

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

[8] G. P. Agrawal, *Nonlinear Fiber Optics*, 4th ed. New York: Academic Press, 2007.

[9] ——, *Fiber-optic communication systems*, 3rd ed. New York, NY: John Wiley, 2002.

[10] ——, *Nonlinear Fiber Optics*, 3rd ed. San Diego, Calif.: Academic Press, 2001.

[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ: Wiley, 2006.

[12] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai (Shitz), "On information rates for mismatched decoders," *IEEE Trans. Inform. Theory*, vol. 40, no. 6, pp. 1953–1967, Nov. 1994.

[13] L. Wegener, B. Povinelli, A. Green, P. Mitra, J. Stark, and P. Littlewood, "The effect of propagation nonlinearities on the information capacity of WDM optical fiber systems: Cross-phase modulation and four-wave mixing," *Physica D: Nonlinear Phenomena*, vol. 189, no. 1, pp. 81–99, Feb. 2004.

[14] 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, Jun. 2001.

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

[16] M. Secondini, E. Forestieri, and G. Prati, "Achievable information rate in nonlinear WDM fiber-optic systems with arbitrary modulation formats and dispersion maps," *J. Lightw. Technol.*, vol. 31, no. 23, pp. 3839–3852, Dec. 2013.

[17] E. Forestieri and M. Secondini, "The nonlinear fiber-optic channel: Modeling and achievable information rate," in *Progress In Electromagnetics Research Symposium (PIERS)*, Prague, Czech Republic, Jul. 2015.

[18] M. Secondini and E. Forestieri, "The limits of the nonlinear Shannon limit," in *Proc. Optical Fiber Communication Conf. (OFC)*, Anaheim, CA, Mar. 2016.

[19] M. I. Yousefi, "The asymptotic capacity of the optical fiber," *arXiv preprint:1610.06458*, 2016.

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

[21] N. V. Irukulapati, M. Secondini, E. Agrell, P. Johannisson, and H. Wymeersch, "Tighter lower bounds on mutual information for fiber-optic channels," *arXiv preprint:1606.09176*, 2016.

[22] I. B. Djordjevic, B. Vasic, M. Ivkovic, and I. Gabitov, "Achievable information rates for high-speed long-haul optical transmission," *J. Lightw. Technol.*, vol. 23, no. 11, p. 3755, Nov. 2005.

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

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

[25] E. Agrell and M. Karlsson, "Influence of behavioral models on multiuser channel capacity," *J. Lightw. Technol.*, vol. 33, no. 17, pp. 3507–3515, Sep. 2015.

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

[27] D. Marsella, M. Secondini, E. Agrell, and E. Forestieri, "A simple strategy for mitigating XPM in nonlinear WDM optical systems," in *Proc. Optical Fiber Communication Conf. (OFC)*, Los Angeles, CA, 2015.

[28] M. Secondini and E. Forestieri, "On XPM mitigation in WDM fiber-optic systems," *IEEE Photon. Technol. Lett.*, vol. 26, no. 22, pp. 2252–2255, Nov. 2014.

[29] R. Dar, M. Shtaif, and M. Feder, "New bounds on the capacity of the nonlinear fiber-optic channel," *Opt. Lett.*, vol. 39, no. 2, pp. 398–401, Jan. 2014.

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

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

[32] R.-J. Essiambre, G. J. Foschini, G. Kramer, and P. J. Winzer, "Capacity limits of information transport in fiber-optic networks," *Phys. Rev. Lett.*, vol. 101, no. 16, p. 163901, Oct. 2008.

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

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

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

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

[37] E. Agrell, A. Alvarado, G. Durisi, and M. Karlsson, "Capacity of a nonlinear optical channel with finite memory," *J. Lightw. Technol.*, vol. 32, no. 16, pp. 2862–2876, Aug. 2014.

[38] E. Agrell, "Conditions for a monotonic channel capacity," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 738–748, Mar. 2015.

[39] R. Holzlohner, V. Grigoryan, C. Menyuk, and W. Kath, "Accurate calculation of eye diagrams and bit error rates in optical transmission systems using linearization," *J. Lightw. Technol.*, vol. 20, no. 3, pp. 389–400, Mar. 2002.

[40] A. Mecozzi and R.-J. Essiambre, "Nonlinear Shannon limit in pseudolinear coherent systems," *J. Lightw. Technol.*, vol. 30, no. 12, pp. 2011–2024, Jun. 2012.

[41] Z. Tao, L. Dou, W. Yan, L. Li, T. Hoshida, and J. C. Rasmussen, "Multiplier-free intrachannel nonlinearity compensating algorithm operating at symbol rate," *J. Lightw. Technol.*, vol. 29, no. 17, pp. 2570–2576, Sep. 2011.

[42] E. Forestieri and M. Secondini, "Solving the nonlinear Schrödinger equation," in *Optical Communication Theory and Techniques*. New York, NY: Springer, 2005.

[43] A. Vannucci, P. Serena, and A. Bononi, "The RP method: A new tool for the iterative solution of the nonlinear Schrödinger equation," *J. Lightw. Technol.*, vol. 20, no. 7, pp. 1102–1112, Jul. 2002.

[44] B. Xu and M. Brandt-Pearce, "Comparison of FWM-and XPM-induced crosstalk using the Volterra series transfer function method," *J. Lightw. Technol.*, vol. 21, no. 1, p. 40, Jan. 2003.

[45] E. Ciaramella and E. Forestieri, "Analytical approximation of nonlinear distortions," *IEEE Photon. Technol. Lett.*, vol. 17, no. 1, pp. 91–93, Jan. 2005.

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

[47] L. Beygi, E. Agrell, M. Karlsson, and P. Johannisson, "Signal statistics in fiber-optical channels with polarization multiplexing and self-phase modulation," *J. Lightw. Technol.*, vol. 29, no. 16, pp. 2379–2386, Aug. 2011.

[48] C. Häger, L. Beygi, E. Agrell, P. Johannisson, M. Karlsson, and A. Graell i Amat, "A low-complexity detector for memoryless polarization-multiplexed fiber-optical channels," *IEEE Commun. Lett.*, vol. 18, no. 2, pp. 368–371, Feb. 2014.

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

[50] H. Ghozlan and G. Kramer, "Interference focusing for mitigating cross-phase modulation in a simplified optical fiber model," in *Proc. IEEE Int. Symp. Inform. Theory*, Austin, Tx., Jun. 2010, pp. 2033–2037.

[51] ——, "Interference focusing for simplified optical fiber models with dispersion," in *Proc. IEEE Int. Symp. Inform. Theory*, Saint-Petersburg, Russia, Jul-Aug 2011, pp. 376–379.

[52] A. Lapidoth, *A foundation in digital communication.* Cambridge University Press, 2009.

[53] P. D. Lax, "Integrals of nonlinear equations of evolution and solitary waves," *Communications on pure and applied mathematics*, vol. 21, no. 5, pp. 467–490, Sep. 1968.

[54] C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura, "Korteweg-devries equation and generalizations. vi. methods for exact solution," *Communications on pure and applied mathematics*, vol. 27, no. 1, pp. 97–133, Jan. 1974.

[55] M. I. Yousefi and F. R. Kschischang, "Information transmission using the nonlinear Fourier transform, part I: Mathematical tools," *IEEE Trans. Inform. Theory*, vol. 60, no. 7, pp. 4312–4328, Jul. 2014.

[56] ——, "Information transmission using the nonlinear Fourier transform, part II: Numerical methods," *IEEE Trans. Inform. Theory*, vol. 60, no. 7, pp. 4329–4345, Jul. 2014.

[57] ——, "Information transmission using the nonlinear Fourier transform, part III: Spectrum modulation," *IEEE Trans. Inform. Theory*, vol. 60, no. 7, pp. 4346–4369, Jul. 2014.

[58] M. I. Yousefi and X. Yangzhang, "Linear and nonlinear frequency-division multiplexing," in *Proc. European Conference on Optical Communication (ECOC)*, Dusseldorf, Germany, Sep. 2016.

[59] R. Dar and P. J. Winzer, "Nonlinear interference mitigation: Methods and potential gain," *J. Lightw. Technol.*, vol. 35, no. 4, pp. 903–930, Apr. 2017.

[60] R.-J. Essiambre and P. J. Winzer, "Fibre nonlinearities in electronically predistorted transmission," in *Proc. European Conference on Optical Communication (ECOC)*, vol. 2, Glasgow, UK, Sep. 2005, paper Tu3.2.2.

[61] G. Goldfarb, M. G. Taylor, and G. Li, "Experimental demonstration of fiber impairment compensation using the split-step finite-impulse-response filtering method," *IEEE Photon. Technol. Lett.*, vol. 20, no. 22, pp. 1887–1889, Nov. 2008.

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

[63] R. Dar and P. J. Winzer, "On the limits of digital back-propagation in fully loaded WDM systems," *IEEE Photon. Technol. Lett.*, vol. 28, no. 11, pp. 1253–1256, Jun. 2016.

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

[65] 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 Process. Mag.*, vol. 31, no. 2, pp. 46–56, Mar. 2014.

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

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

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

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

[70] A. Ellis, S. Le, M. McCarthy, and S. Turitsyn, "The impact of parametric noise amplification on long haul transmission throughput," in *International Conference on Transparent Optical Networks (ICTON)*, Budapest, Hungary, Jul. 2015.

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

[72] P. Kaewplung and K. Kikuchi, "Simultaneous cancellation of fiber loss, dispersion, and effect in ultralong-haul optical fiber transmission by midway optical phase conjugation incorporated with distributed Raman amplification," *J. Lightw. Technol.*, vol. 25, no. 10, pp. 3035–3050, Oct. 2007.

[73] D. Marsella, M. Secondini, and E. Forestieri, "Maximum likelihood sequence detection for mitigating nonlinear effects," *J. Lightw. Technol.*, vol. 32, no. 5, pp. 908–916, Mar. 2014.

[74] Y. Cai, D. G. Foursa, C. R. Davidson, J.-X. Cai, O. Sinkin, M. Nissov, and A. Pilipetskii, "Experimental demonstration of coherent MAP detection for nonlinearity mitigation in long-haul transmissions," in *Proc. Optical Fiber Communication Conf. (OFC)*, San Diego, CA, Mar. 2010.

[75] G. Bosco, I. N. Cano, P. Poggiolini, L. Li, and M. Chen, "MLSE-based DQPSK transmission in 43 Gb/s DWDM long-haul dispersion-managed optical systems," *J. Lightw. Technol.*, vol. 28, no. 10, pp. 1573–1581, Mar. 2010.

[76] 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," *J. Lightw. Technol.*, vol. 33, no. 23, pp. 4738–4745, Dec. 2015.

[77] G. Liga, A. Alvarado, E. Agrell, M. Secondini, R. I. Killey, and P. Bayvel, "Optimum detection in presence of nonlinear distortions with memory," in *Proc. European Conference on Optical Communication (ECOC)*, Valencia, Spain, Sep.–Oct. 2015.

[78] L. Li, Z. Tao, L. Liu, W. Yan, S. Oda, T. Hoshida, and J. C. Rasmussen, "Nonlinear polarization crosstalk canceller for dual-polarization digital coherent receivers," in *Proc. Optical Fiber Communication Conf. (OFC)*, San Diego, CA, Mar. 2010.

[79] I. B. Djordjevic, L. L. Minkov, L. Xu, and T. Wang, "Suppression of fiber nonlinearities and PMD in coded-modulation schemes with coherent detection by using turbo equalization," *J. Opt. Commun. Netw.*, vol. 1, no. 6, pp. 555–564, Nov. 2009.

[80] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Inter-channel nonlinear interference noise in WDM systems: modeling and mitigation," *J. Lightw. Technol.*, vol. 33, no. 5, pp. 1044–1053, Mar. 2015.