Digital Signal Processing Techniques for Fiber Nonlinearity Compensation in Coherent Optical Communication Systems

by

©O. S. Sunish Kumar, M. Tech

A Dissertation submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Electrical and Computer Engineering
Faculty of Engineering and Applied Science

Memorial University of Newfoundland

June 1, 2020

St. John's Newfoundland

Abstract

The capacity of long-haul coherent optical communication systems is limited by the detrimental effects of fiber Kerr nonlinearity. The power-dependent nature of the Kerr nonlinearity restricts the maximum launch power into the fiber. That results in the reduction of the optical signal-to-noise ratio at the receiver; thereby, the maximum transmission reach is limited. Over the last few decades, several digital signal processing (DSP) techniques have been proposed to mitigate the effects of fiber nonlinearity, for example, digital back-propagation (DBP), perturbation based nonlinearity compensation (PB-NLC), and phase-conjugated twin wave (PCTW). However, low-complexity and spectrally efficient DSP-based fiber nonlinearity mitigation schemes for long-haul transmission systems are yet to be developed.

In this thesis, we focus on the computationally efficient DSP-based techniques that can help to combat various sources of fiber nonlinearity in long-haul coherent optical communication systems. With this aim, we propose a linear time/polarization coded digital phase conjugation (DPC) technique for the mitigation of fiber nonlinearity that doubles the spectral efficiency obtained in the PCTW technique. In addition, we propose to investigate the impact of random polarization effects, like polarization-dependent loss and polarization mode dispersion, on the performance of the linear-coded DPC techniques. We also propose a joint technique that combines single-channel DBP with the PCTW technique. We show that the proposed scheme is

computationally efficient and achieves similar performance as multi-channel DBP in wavelength division multiplexed superchannel systems.

The regular perturbation (RP) series used to analytically approximate the solution of the nonlinear Schrödinger equation (NLSE) has a serious energy divergence problem when truncated to the first-order. Recent results on the transmission of high data-rate optical signals reveal that the nonlinearity compensation performance of the first-order PB-NLC technique decreases as the product of the transmission distance and launch power increases. The enhanced RP (ERP) method can improve the accuracy of the first-order RP approximation by partially solving the energy divergence problem. On this ground, we propose an ERP-based nonlinearity compensation technique to compensate for the fiber nonlinearity in a polarization-division multiplexed dispersion unmanaged optical communication system. Another possible solution to improve the accuracy of the PB-NLC technique is to increase the order of the RP solution. Based on this idea, we propose to extend the first-order solution of the NLSE to the second-order to improve the nonlinearity compensation performance of the PB-NLC technique. Following that, we investigate a few simplifying assumptions to reduce the implementation complexity of the proposed second-order PB-NLC technique.

Acknowledgments

First, I would like to express the deepest gratitude towards my research supervisors Dr. Octavia A. Dobre and Dr. Ramachandran Venkatesan, for giving me the opportunity to do a Ph.D. in the area of optical communications engineering at the Memorial University of Newfoundland under their guidance. I am very much thankful for their constant encouragement, inspiring guidance, and insightful discussions during the period of my Ph.D. study at the Memorial University of Newfoundland. I would like to thank Dr. Sarah Kate Wilson, professor at Santa Clara University, for the numerous discussions that covered topics related to my research work, which immensely helped me to gain a broad insight into the research and improve myself. I would like to express my sincere thanks to my committee member Dr. Paul Gillard for his encouragement and fruitful discussions, which helped me to solve some hard problems during my research study.

It will be a great disservice if I do not acknowledge the significant influence of Dr. Abdelkerim Amari on my research career. His research knowledge, meticulous technical feedback, and personal and emotional support helped me a lot with the successful completion of the research work. I would like to express my sincere gratitude towards Dr. Balagangadhar Bathula for his strong encouragement, which helped me to think about the research career after the completion of my master's degree in optical communications engineering. I would like to thank Dr. Animesh Yadav

for his support and enlightening discussions during my research work at Memorial University. I am very much thankful to my best friend and colleague, Dr. Xiang Lin, for his wholehearted personal and technical support during my research life. I would like to express my sincere gratitude towards the former members of our optical research group, particularly, Dr. Oluyemi Omomukuyo, Dr. Deyuan Chang, Dr. Stanley Johnson, Dr. Shu Zhang, and Dr. Ali Al-Bermani for their technical support and friendship. I would like to thank our former optical project managers, Ms. Lori Hogan and Ms. Suchita Yadav, for their support and help. I would also like to thank all the Electrical and Computer Engineering department staff for providing me all the conveniences and assistance during my studies.

Last but foremost, I express gratitude and love towards my mother, Ms. Sarada Muttumannil Ayyappan, my father, Mr. Soman Kumbikkappuzha Gopalan, my wife Ms. Remya Velanparambil Ravindran, and my cute son Mstr. Adidev Orappanpara Sunish. I am eternally indebted to them for their wholehearted love, prayers, and invaluable sacrifices.

Table of Contents

List of Tables

List of Figures

List of Acronyms and Abbreviations

ASE	Amplified Spontaneous Emission
CD	Chromatic Dispersion
CDR	Conjugate Data Repetition
CMA	Constant Modulus Algorithm
CNLSE	Coupled Nonlinear Schrödinger Equation
CO-OFDM	Coherent Optical Orthogonal Frequency Division Multiplexing
CS	Coherent Superposition
DBP	Digital Back-Propagation
DFT	Discrete Fourier Transform
DGD	Differential Group Delay
DPC	Digital Phase Conjugation
DSP	Digital Signal Processing
EDC	Electronic Dispersion Compensation

EDFA	Erbium-Doped Fiber Amplifier
FEC	Forward Error-Correction
FFT	Fast Fourier Transform
FO	First-Order
FO-PB-NLC	FO Perturbation Theory-Based NLC
IFWM	Intra-channel Four-Wave Mixing
IXPM	Intra-channel Cross-Phase Modulation
LUT	Look-Up Table
MVNLE	Modified VNLE
MZM	Mach-Zehnder Modulator
NLC	Nonlinearity Compensation
NLIN	Nonlinear Interference Noise
NLPN	Nonlinear Phase Noise
NLSE	Nonlinear Schrödinger Equation
NSNI	Nonlinear Signal-Noise Interaction
OPC	Optical Phase Conjugation
OSNR	Optical Signal-to-Noise Ratio
PCP	Phase-Conjugated Pilots
PCSC	Phase-Conjugated Subcarrier Coding

PCTW	Phase-Conjugated Twin Wave
PDF	Probability Density Function
PDL	Polarization-Dependent Loss
PDM	Polarization Division Multiplexing
PMD	Polarization-Mode Dispersion
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase-Shift-Keying
RDE	Radius Directed Equalizer
SE	Spectral Efficiency
SOP	State-Of-Polarization
SPM	Self-Phase Modulation
SSFM	Split-Step Fourier Method
SSMF	Standard Single-Mode Fiber
VNLE	Volterra series-based Nonlinear Equalizer
VSTF	Volterra Series Transfer Function
WDM	Wavelength Division Multiplexing
WVSNE	Weighted VNLE
XPolM	Cross-Polarization Modulation

Chapter 1

Introduction

1.1 Background and Motivation

The invention of Erbium-doped fiber amplifiers (EDFAs), and cost-effective manufacturing of standard single-mode fibers (SSMFs) with less than 0.2 dB/km attenuation marked the beginning of commercially viable fiber-optic communication systems. Since the deployment, the optical communication systems underwent several technical evolutions to fulfill the requirements of high-speed communications. Over the past few decades, there has been an exponential increase in the global demand for communication capacity. Most of the growth has occurred in the last few years when data started dominating the network traffic. According to Cisco's recent white paper [1], the global network traffic will dramatically increase over the next three years. The emergence of bandwidth-hungry applications, such as cloud services and virtual reality, has fueled the global network traffic increase to a large extent [2]- [5]. Furthermore, human-centered applications like video gaming and the exchange of multimedia content via smartphones are among the most bandwidth-consuming applications. That leads to a strong requirement for an increase in the access network capacity, and

consequently, for the core and metro network capacities to meet such ever-increasing traffic demands [5].

The deployment of wavelength division multiplexing (WDM) technology, which enables the multiplexing of several optical signals in the same fiber, dramatically increases the fiber capacity [6]- [8]. Over the past few years, WDM systems with 2.5 Gb/s line rate, employing on-off keying evolved into 10 Gb/s systems with 50 GHz channel spacing [6]. Later, 40 Gb/s system mainly based on quadrature phase-shift-keying (QPSK) modulation and non-coherent detection, were deployed [6]. Following that, a 100 Gb/s transmission system was introduced employing the polarization division multiplexed (PDM) QPSK and coherent detection [7].

To further increase the transmission capacity up to 400 Gb/s or 1 Tb/s, the super-channel technique represents a potential candidate due to its high spectral efficiency and low cost [8]- [11]. The main idea of the superchannel approach is to split the WDM channel into several subcarriers with smaller bandwidths and separated by a small guard-band. These subcarriers are routed through optical add-drop multiplexers and wavelength selective switches as a single entity. In comparison with single-carrier 400 Gb/s and 1 Tb/s transmission systems, the superchannel approach offers multiple advantages [11]. It has lower requirements in terms of optical signal-to-noise ratio (OSNR) and analog-to-digital converters/digital-to-analog converters bandwidth [11]. In comparison with single-carrier 400 Gb/s and 1 Tb/s transmission systems, the superchannel systems exhibit better transmission performance. However, the transmission impairments in the optical fiber degrade the performance of the high data-rate optical superchannel systems. For example, the chromatic dispersion (CD) introduces a frequency-dependent phase shift to the signals and acts as a primary limiting factor in provisioning a reliable long-haul optical communication link [12]. It is noteworthy that there are various optical and electrical techniques available to combat

the adverse effect of the CD in a long-haul optical communication system [12]- [16].

Fig. 1.1 shows the schematic diagram of the modern WDM optical fiber communication system. The data in different wavelength channels are combined using an optical multiplexer in each polarization tributaries. Then, the two orthogonal polarization signals are combined using a polarization combiner and are transmitted over the optical fiber transmission link. The optical transmission link consists of several spans of SSMF having an EDFA after each span to compensate for the optical fiber loss [17]. At the receiver, after the polarization splitting, each wavelength channel is demultiplexed and coherently detected to recover the data.

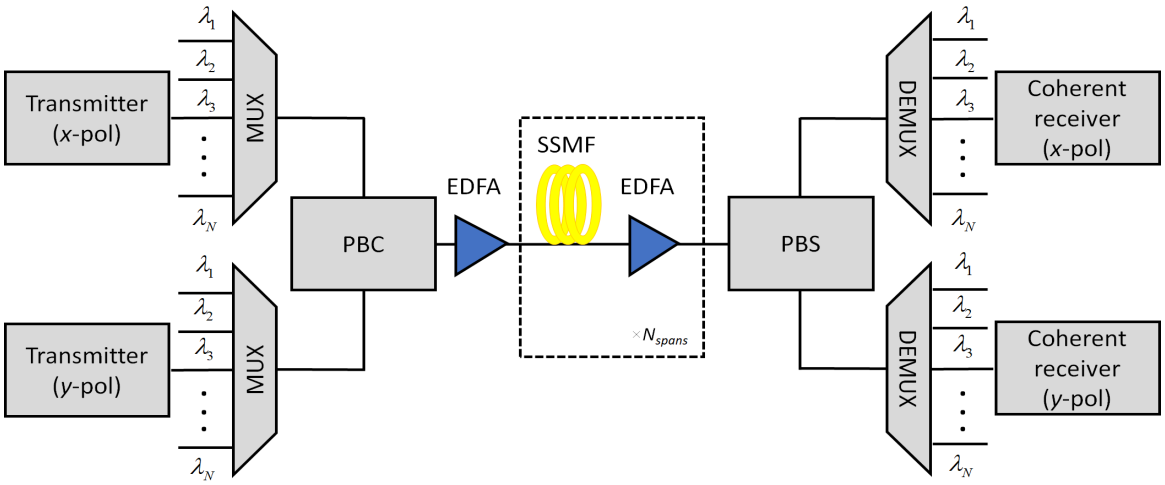


Figure 1.1: Schematic diagram of modern PDM-WDM fiber optic transmission link. MUX: multiplexer, DEMUX: de-multiplexer, PBC: polarization beam combiner, EDFA: erbium-doped fiber amplifier, SSMF: standard single-mode fiber, PBS: polarization beam splitter.

The optical fiber link capacity can be doubled by employing the PDM transmission scheme [18]. However, this technology imparts several challenges to the design of the optical communication system. The polarization-dependent loss (PDL) and polarization-mode dispersion (PMD) are the two prevalent linear impairments pertaining to the long-haul optical communication systems [18], [19]. It is worth mentioning that PDL and PMD are not confined to the fiber; however, they present in

the optical components, such as EDFA, optical multiplexer/demultiplexer, optical splitter/combiner, etc. [19]. In contrast to CD, the PDL and PMD are considered dynamic, rapidly changing with various environmental conditions, such as temperature fluctuations, vibration, or stress [19].

The PDL refers to the energy loss that is preferential to one of the two orthogonal polarization states [17], [19]. That will eventually lead to the OSNR difference and the signal cross-talk between the two polarization tributaries of the SSMF [19]. It is important to note that the total PDL of a link with concatenated optical components with individual PDL value is different from the sum of every single PDL-element contribution [19]. That is because the polarization-sensitive axes of each component are not always aligned to each other; therefore, the resultant accumulated PDL at the end of the transmission link depends on the relative orientations of the PDL axes of the individual components. Since the PDL value of each element depends on the environmental conditions and there exists a randomized PDL-axis alignment between individual components, the accumulated PDL of the whole link is usually estimated using a statistical modeling technique [19]. The seminal work in [20] suggests that the total accumulated PDL for a typical optical fiber communication link is found to be Maxwellian distributed "when expressed in decibels" and the accumulation of the mean PDL grows linearly with the transmission length.

Due to random imperfections and asymmetries, the refractive index of the optical fiber changes with the polarization. That leads to slightly different propagation velocities for the two distinct polarization eigenstates [18]. The state-of-polarization (SOP) corresponding to these eigenstates are usually labeled as the "fast" and "slow" axes. The segments of the fiber exhibiting such fast and slow axes are referred to as the birefringent segments. The birefringence introduces a delay between the two polarization states, termed as differential group delay (DGD). The PMD is caused by the DGD

generated by the concatenated birefringent segments of the optical fiber. The birefringent magnitude and the orientation of the axes of each segment vary randomly along the length of the fiber. As a result, the instantaneous value of DGD exhibited by the concatenated series of fiber birefringent segments changes randomly with time [21]. The study in [21] indicates that the DGD at the end of a concatenated birefringent fiber segments can be well approximated by a Maxwellian probability density function, which is controlled by only one parameter, the average DGD. This average DGD is referred to as the PMD value of the fiber [21].

In a realistic optical communication system, the polarization-dependent effects become quite complicated due to the interplay between the PDL and PMD [22]. That will impair the communication system more than either effect alone. In general, the PDL is wavelength-independent, and the PMD is wavelength-dependent. Therefore, adding some PMD to PDL will result in a wavelength-dependent PDL; similarly, the addition of PDL to PMD can result in a DGD different from what one would expect from PMD alone [22]. In the design of the PDM transmission system, there are different technologies adopted to track the polarization state in the presence of the complex polarization-dependent effects. A widely used technique is to employ a multi-input multi-output equalizer with adaptive tap coefficients. That can be implemented as data-aided, decision-directed, or blind. For example, the constant modulus algorithm (CMA) is a prominent blind equalization technique to track the polarization states for the m-PSK family of constellations [23]. Similarly, the radius directed equalizer is an adaptation of CMA to quadrature amplitude modulation (QAM) systems to track the polarization state of the signal adaptively [24].

Another significant impairment that limits the transmission performance of the long-haul optical communication system is the fiber nonlinearity [5]. The nonlinearity effects in the optical fiber are due to an electro-optic effect, referred to as the Kerr

effect, which arises from the dependence of the optical fiber refractive index on the transmit signal power [5], [17]. That makes the optical fiber channel different from other transmission media used for the information transfer. The modern high data-rate optical transmission systems use multilevel modulation formats, which require higher SNR. However, the optical intensity-dependent nonlinear Kerr effect significantly degrades the transmission performance as the fiber launch power increases [17]. In a linear transmission medium, the information signals are usually perturbed by the additive noise, which generally results in channel capacities monotonically increasing with transmit power, thereby a corresponding increase in the SNR. However, the detrimental effects of Kerr-induced signal nonlinear distortions grow at a faster rate than the SNR capacity gain at higher launch powers. That in turn leads the channel capacity to be a nonmonotonic function of the transmit launch power with a maximum value at a particular launch power termed as optimum launch power. The achievable transmission rate decreases rapidly beyond the optimal power point as the launch power increases due to the corresponding increase in the Kerr-induced signal nonlinear distortions [17].

In a single-channel optical communication system, the intra-channel Kerr nonlinearity is considered a dominant impairment, which limits the transmission performance. The intra-channel nonlinear interactions can be categorized into three types: (i) self-phase modulation (SPM), (ii) intra-channel cross-phase modulation (IXPM), and (iii) intra-channel four-wave mixing (IFWM). The SPM results in a phase modulation induced by the refractive index, which increases with the increase in the input power level. That leads to a frequency shift, known as frequency chirping [17], which interacts with the CD and causes spectral broadening of the optical pulse. The IXPM is the result of the refractive index change proportional to the intensity of the neighboring pulses in the same channel [17]. The IXPM yields a timing jitter between the

co-propagating pulses, which leads to performance degradation of the optical transmission system. The IFWM is caused by the nonlinear interaction between two or more pulses of the same channel [17]. That generates echo or ghost pulses in the time domain, and thus, results in interference between the signal pulses of the same channel. It is important to mention that the SPM, IXPM, and IFWM are considered as deterministic nonlinear impairments, and these effects can be compensated well using digital nonlinearity compensation (NLC) techniques.

In WDM systems, the significant portion of the nonlinear distortion comes from the nonlinear interaction between the channel under consideration and the co-propagating signals in neighboring channels, referred to as inter-channel nonlinearity effects [17]. The inter-channel nonlinear effects can be classified into three types: (i) XPM, (ii) cross-polarization modulation (XPolM), and (iii) FWM. The XPM effect is due to the refractive index change proportional to the intensity of the pulse in the co-propagating channel. The XPM induces frequency chirping and pulse overlapping between channels [17]. The XPolM occurs in PDM transmission systems when the SOP of a transmitted channel depends on the SOP of other co-propagating channels. This is due to the random propagation of SOP inside the optical fiber caused by PMD [17] and causes channel crosstalk for dual-polarization systems [17]. The FWM in the multichannel system is generated by the nonlinear interaction between two or more co-propagating wavelength channels [17]. The FWM results in significant performance degradation due to crosstalk among different wavelength channels [17]. In contrast to the intra-channel deterministic nonlinear effects, the elimination of the inter-channel distortions is considered impractical. The reasons are three-fold: (i) the information in the co-propagating WDM channels is unknown, (ii) the unpredictable add/drop of the WDM channels taking place in the optical network, and (iii) the SOP of different WDM channels evolve randomly and differently in the presence of PMD. For these

reasons, the inter-channel nonlinear interactions are considered as stochastic effects and commonly referred to as nonlinear interference noise (NLIN) [25].

Another category of the stochastic nonlinearity effect is due to the Kerr-induced signal-amplified spontaneous emission (ASE) noise nonlinear interaction, referred to as the nonlinear signal-noise interaction (NSNI) or the Gordon–Mollenauer effect [26]-[28], which leads to the generation of the stochastic nonlinear phase noise (NLPN). The NLPN limits the transmission performance of the optical communication system and cannot be compensated by using traditional NLC methods. However, the NSNI plays a fundamental role in the optical communication systems by contributing to answering the questions regarding the fundamental limits of the performance improvement provided by the NLC techniques [26]- [28].

The stimulated Raman scattering (SRS) and stimulated Brillouin scattering (SBS) are another category of nonlinear effects in the optical fiber. SRS and SBS are inelastic scattering processes in which the optical medium absorbs part of the optical wave power. The SRS can severely affect the WDM transmission systems by transferring energy from one channel to the co-propagating channels [17]. In any molecular medium, the spontaneous Raman scattering can transfer a small amount of power from one optical field to another through a process called Raman scattering [17]. In this process, the optical field to which the power is transferred is downshifted by an amount determined by the vibrational mode of the molecular medium. The incident photon of the input pump field is scattered quantum-mechanically to a lower-frequency photon by one of the molecules, and thereby the molecule takes a transition to the higher energy vibrational states [17]. This frequency-shifted radiation is referred to as the Stokes wave [17]. It was observed in 1962 that, in a medium, the energy of the Stokes wave grows rapidly upon the incident of an intense pump field, which is due to the nonlinear phenomenon of the SRS effect [17]. The effect of SRS has been studied

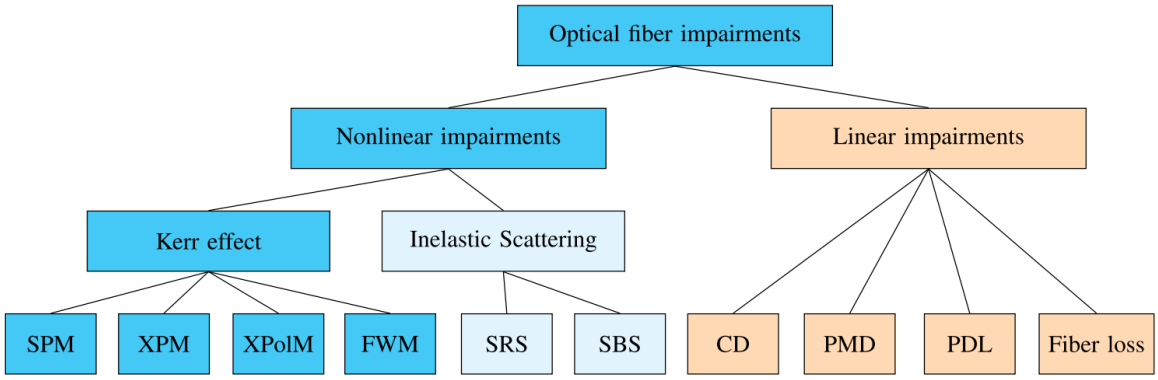


Figure 1.2: Linear and nonlinear impairments in the optical fiber channel, taken from [5].

extensively in the literature for a variety of molecular media, including silica [17]. Unlike SRS, the SBS in optical fibers can occur at much lower input power [17]. Once the input power reaches the Brillouin threshold, the backward propagating Stokes wave is generated, which carries most of the input energy [17]. The nonlinear phenomenon of the SBS was first observed in 1964 and extensively studied in the literature [17]. Similar to the SRS effect, the SBS also generates a frequency downshifted Stokes wave [17]. It is important to mention that these inelastic scattering processes can be neglected because they manifest only at input powers higher than the typical values used in optical communication systems [17].

From above discussions it is clear that the optical signal propagating through the SSMF is degraded due to the linear and nonlinear impairments in the optical fiber. Fig. 1.2 categorizes various linear and nonlinear impairments affecting the quality of optical signal propagation inside the optical fiber.

In recent years, the spectral efficiency of the PDM optical communication systems has dramatically improved by the introduction of a promising detection technique referred to as coherent detection. In coherent detection, the symbol decisions are made using both in-phase and quadrature signals; thereby, information is encoded on all the available degrees of freedom [29]. At the coherent receiver, after down-conversion to

the electrical domain and sampling at the Nyquist rate, the digitized signal waveform retains the full information of the electric field [29]. That enables the compensation of various impairments, including fiber nonlinearity, using digital signal processing (DSP) techniques [30]. The coherent detection also enables the implementation of the advanced forward error-correction (FEC) coding techniques and the adaptive DSP algorithms to combat time-varying transmission impairments [31]. Several DSP-based NLC techniques have been proposed in the last decade to deal with fiber nonlinearity effects. A concise review of such popular techniques is given in Chapter 2.

This research is mainly motivated by the growing demand for the high data rate coherent optical communication systems over trans-oceanic long-haul transmission links. The use of higher-order modulation formats in such systems necessitates the development of effective DSP techniques to compensate for fiber nonlinearity effects. Although a significant amount of research has been carried out by the optical research community to deal with the detrimental effects of fiber nonlinearity, low-complexity, and commercially viable DSP techniques are yet to be developed. Additionally, the growing interest of the industrial research community to move towards the nonlinear regime of optical fiber to realize ultra high capacity coherent optical systems motivates our search on the effective DSP solutions to combat fiber nonlinearity.

1.2 Thesis Contributions

The main contributions of this Ph.D. thesis can be listed as follows:

- We propose two linearly coded digital phase conjugation (DPC) schemes to solve the spectral efficiency problem of the phase-conjugated twin wave (PCTW) and conjugate data repetition (CDR) techniques. We investigate the performance of the proposed schemes with and without considering the dispersion symmetry in

the transmission link. Also, we investigate the impact of polarization-dependent effect, such as PDL, on the performance of the proposed linearly coded DPC schemes. The results of the detailed study are discussed in Chapter 3.

- We propose a joint technique by combining the PCTW technique with the single-channel digital back-propagation (DBP) to solve the limitations of the individual implementations of such techniques in compensating the fiber nonlinearity effects. We further show that the joint technique can provide similar performance as the multi-channel DBP in a WDM transmission system. The details of the proposed technique and results are given in Chapter 4.

- We show that the enhanced regular perturbation (ERP) technique can solve the energy divergence problem of the regular perturbation-based approach in a dispersion unmanaged transmission system. Then, we propose to use the ERP approach, instead of the regular perturbation method, in the perturbation-based nonlinearity compensation technique (PB-NLC). We demonstrate that the proposed technique improves the NLC performance with a negligible increase in the computational complexity when compared to the conventional PB-NLC technique. Chapter 5 is devoted to this work.

- With the help of rigorous mathematical analysis, we derive the second-order (SO) perturbative nonlinear distortion field for a dispersion unmanaged transmission system with a Gaussian shape assumption for the input pulse shape. We develop a SO perturbative distortion field-based predistortion technique to compensate for the fiber nonlinearity effect in a coherent optical communication system. Then, we adopt a few simplifying assumptions to reduce the implementation complexity of the proposed predistortion technique. We carry out the complexity analysis in detail and show that the proposed technique comes with

a reduced implementation complexity when compared to the benchmark DBP technique. This research is discussed in detail in Chapter 6.

1.3 Thesis Organization

The rest of this thesis document is organized as follows. Chapter 2 provides brief discussions on the origin of the nonlinear susceptibility and the Kerr effect, the mechanism of the optical pulse propagation in the optical fiber medium, the numerical and analytical methods commonly used to solve the pulse propagation equation, and various DSP techniques available in the literature to deal with the detrimental effects of fiber nonlinearity. Chapter 3 presents the proposed linearly coded DPC techniques to compensate for fiber nonlinearity and double the spectral efficiency of the existing PCTW technique. This chapter also includes the investigation of the impact of PDL and PMD on the proposed linear coding techniques. Chapter 4 discusses the proposed joint technique, which combines SC-DBP with the PCTW technique to compensate for the fiber nonlinearity in a coherent optical communication system. Chapter 5 presents the proposed ERP-based nonlinearity compensation technique to compensate for the fiber nonlinearity in a polarization-division multiplexed dispersion unmanaged optical communication system. Chapter 6 explains the extension of the first-order perturbation theory to the SO and the developed SO perturbation theory-based predistortion technique to compensate for the fiber nonlinearity effect. Finally, Chapter 7 provides a brief review of this work and presents possible future extensions of the current work.

1.4 Notation

Lower case italic typeface letters are used for the time-domain representation, whereas the frequency-domain is represented by upper case italic typeface letters. Matrices and vectors are represented by upper case bold typeface letters and lower case bold typeface letters, respectively.

1.5 List of Publications

1.5.1 Journal Publications

[J1] **O. S. Sunish Kumar**, A. Amari, O. A. Dobre, and R. Venkatesan, "Second-order perturbation theory-based digital predistortion for fiber nonlinearity compensation," *IEEE Trans. Commun.*, 2020 (Under review).

[J2] **O. S. Sunish Kumar**, A. Amari, O. A. Dobre, and R. Venkatesan, "Enhanced regular perturbation-based nonlinearity compensation technique for optical transmission systems," *IEEE Photon. J.*, vol. 11, no. 4, Aug. 2019.

[J3] **O. S. Sunish Kumar**, A. Amari, O. A. Dobre, and R. Venkatesan, "PDL impact on linearly coded digital phase conjugation techniques in CO-OFDM systems," *IEEE Photon. Technol. Lett.*, vol. 30, no. 9, pp. 769-772, May 2018.

[J4] A. Amari, O. A. Dobre, R. Venkatesan, **O. S. Sunish Kumar**, P. Ciblat, and Y. Jaouën, "A survey on fiber nonlinearity compensation for 400 Gbps and beyond optical communication systems," *IEEE Commun. Surv. Tutorials*, vol. 19, no. 4, pp. 3097-3113, 4th Quart. 2017.

1.5.2 Conference Proceedings

[C1] T. T. Nguyen, P. Harper, **O. S. Sunish Kumar**, and A. Ellis, "Nonlinear tolerance enhancement based on perturbation theory for optical phase conjugation systems," in *Proc. OFC*, San Diego, CA, USA, Mar. 2020, paper Th2A.52.

[C2] **O. S. Sunish Kumar**, A. Amari, O. A. Dobre, R. Venkatesan, and S. K. Wilson, "A joint technique for nonlinearity compensation in CO-OFDM superchannel systems," in *Proc. ACPC*, Guangzhou, China, Nov. 2017, paper Su4B.4.

[C3] **O. S. Sunish Kumar**, O. A. Dobre, R. Venkatesan, S. K. Wilson, O. Omomukuyo, A. Amari, and D. Chang, "A spectrally-efficient linear polarization coding scheme for fiber nonlinearity compensation in CO-OFDM systems," in *Proc. SPIE Opt.*, San Francisco, CA, USA, Jan. 2017, pp. 152-161.

[C4] **O. S. Sunish Kumar**, S. K. Wilson, O. A. Dobre, R. Venkatesan, O. Omomukuyo, A. Al-Bermani, "The impact of the non-linearity and finite extinction ratio of the MZM on the performance of CO-OFDM," in *Proc. NECEC*, St. John's, NL, Canada, Nov. 2015.

Chapter 2

Preliminaries

2.1 Optical Fiber Communication Systems

The modern high capacity communication network uses optical fibers to transmit information from one point to another as modulated light pulses. The optical fiber is a dielectric cylindrical waveguide made up of low-cost material silica [32]. The basic phenomenon responsible for guiding the light pulses inside the optical fiber is total internal reflection [32]. The uncladded optical fibers manufactured in the early 1920s were not suitable for information transfer at long distances [32]. In the 1950s, the use of the cladding layer was a starting point for the field of fiber optics and led to a significant improvement in the fiber characteristics [32]. During the 1960s, the transmission of images through the glass fibers demonstrated the significant development of the field of fiber optics [17], [32]. The fibers manufactured during that time had a drawback of extreme power loss (loss>1000 dB/km) while the optical light is transmitted. Further research efforts in the 1970s reduced the loss of the silica-based fibers below 20 dB/km [32]. During the late 1970s, advancements in fiber fabrication technology significantly reduced the fiber loss down to 0.2 dB/km in the 1550 nm

wavelength range [32]. That eventually led to a breakthrough in the field of optical fiber communication systems [32]. During the 1990s, the optical fibers were doped with rare-earth elements such as Erbium. That led to the development of the optical amplifiers and lasers. The invention of EDFA to compensate for the power losses initiates an implicit revolution in the design of the WDM optical fiber communication systems [8], [32]. It is important to mention that the mathematical derivations given in sections, from Sections 2.2 to 2.4, are mostly taken from [32] and [35].

2.2 Characteristics of Optical Fiber Channel

The optical fiber consists of a glass core surrounded by a cladding layer to confine the modulated light inside the core region. The refractive index n_2 of the cladding layer is slightly lower than that of the core index n_1 to facilitate the total internal reflection [32]. Such fibers are generally referred to as step-index fiber. The step-index optical fibers can be categorized by two parameters, namely core-cladding index difference Δ and the V parameter, given as:

$$\Delta = \frac{n_1 - n_2}{n_1} \tag{2.1}$$

and

$$V = k_0 a \sqrt{n_1^2 - n_2^2}, \tag{2.2}$$

respectively, where $k_0 = \frac{2\pi}{\lambda}$, λ is the wavelength of light and a is the core radius. The V parameter determines the number of propagation modes in the optical fiber. For example, if $V < 2.405$, then the step-index fiber supports only one mode. Such fibers are termed as single-mode fibers [32]. It is important to mention that, in this thesis, we consider nonlinearity effects in the single-mode fibers since they are used to realize

the long-haul high data rate optical communication systems.

2.2.1 Fiber Losses

While the optical signal propagates through the optical fiber channel, the fiber losses attenuate the signal power as a function of the transmission distance. The transmitted signal power considering the fiber attenuation can be represented as [32]:

$$P_T = P_0 \exp\left(-\alpha L\right), \tag{2.3}$$

where P_0 is the power launched at the input of the fiber, α is the fiber attenuation parameter, and L is the transmission distance. It is worth mentioning that the fiber attenuation parameter is usually expressed in units of dB/km, which can be represented by using (2.3), as follows:

$$\alpha_{\mathrm{dB/km}} = -\frac{10}{L}\log_{10}\left(\frac{P_T}{P_0}\right) = 10\log_{10}\left(e\right)\alpha = 4.343\alpha. \tag{2.4}$$

The attenuation in optical fiber is caused by several mechanisms, including absorption, scattering, and geometric effects [32]. The material impurity in the silica core fiber causes the absorption of the light energy. The hydroxyl ion (OH) absorption is one of the main absorptions in the case of glass fibers [32]. The OH absorption causes the multiple absorption peaks in the wavelength range from visible to the infrared band, as shown in Fig. 2.1 [33]. The main source of the scattering loss in optical fiber is due to the Rayleigh scattering [33]. During the fabrication of the optical fiber, the variation of the refractive index is caused by the microscopic variations of fiber material component density, randomly distributed material defects, and inhomogeneous material structure [33]. The scale of this index variation is much smaller than the

wavelength of interest. The energy scattering when the propagating light interacts with such small index variation causes the Rayleigh scattering [33]. The significance of the Rayleigh scattering in optical fiber is reduced as the wavelength increases, as shown in Fig. 2.1.

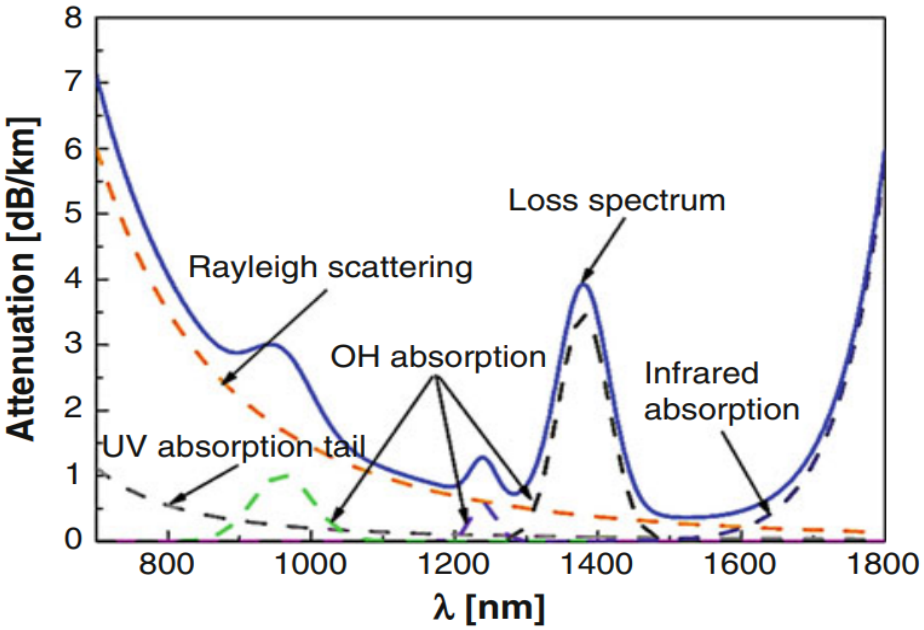


Figure 2.1: Loss spectrum of a single-mode silica fiber, taken from [33].

The physical bending of the fiber is the main source of the geometric effect causing the signal power attenuation. There are two types of bending loss, including macroscopic and microscopic loss. The macroscopic bending loss is produced whenever the optical fiber is subjected to a significant amount of bending beyond a critical value of curvature [17], [32]. The microscopic bending loss is comparatively weaker and is caused by the strain or stress distributed along the length of the fiber [17], [32].

2.2.2 Chromatic Dispersion

In general, the response of a medium to the incident electromagnetic wave depends on the optical frequency of ω. This property of the medium is referred to as the

chromatic dispersion (CD) [32], [34]. The CD effect in optical fiber manifests through the frequency dependence of the refractive index $n(w)$. In the case of the short optical pulses used in the optical communication system, the CD plays a critical role since the optical pulse consists of several frequency components. In the presence of CD, the different spectral components associated with the pulse travel at different speeds given by $c/n(w)$, where c is the speed of light. That will induce pulse broadening [32]. In the weakly nonlinear regime, the CD-induced pulse broadening is the dominant impairment and causes severe penalties to the optical communication systems. In the presence of strong nonlinearity, the interplay between CD and nonlinearity is quite complicated, which results in quantitatively different behavior for the pulse distortions in the optical fiber channel [32].

The effect of CD can be quantitatively accounted by expanding the mode-propagation constant β in a Taylor series about the frequency w_0 at which the pulse spectrum is centered as [32]:

$$\beta(w) = n(w)\frac{w}{c} = \beta_0 + \beta_1(w - w_0) + \frac{1}{2}\beta_2(w - w_0)^2 + ..., \tag{2.5}$$

where

$$\beta_m = \left(\frac{d^m\beta}{dw^m}\right)\Bigg|_{w=w_0} \qquad m = 1, 2, ..., \infty. \tag{2.6}$$

From (2.5), β_2 can be represented as:

$$\beta_2 = \frac{1}{c}\left(2\frac{dn}{dw} + w\frac{d^2n}{dw^2}\right). \tag{2.7}$$

In reality, the envelope of the optical pulse moves at the group velocity, and the parameter β_2 causes the dispersion of the group velocity, which leads to the pulse broadening [32]. This phenomenon is referred to as the group-velocity dispersion

(GVD), and the parameter β_2 is the GVD parameter [32].

In the optical fiber, a part of the optical signal propagates through the cladding, referred to as the dielectric waveguiding, which causes a slight reduction in the effective mode index $n(w)$ of the core. That results in the waveguide dispersion and must be added to the material dispersion, as shown in Fig. 2.2 [34].

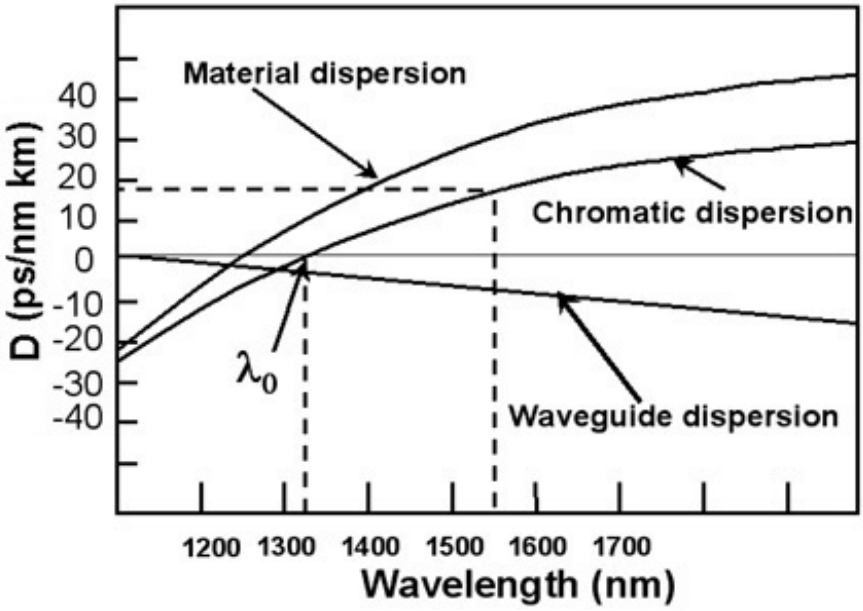


Figure 2.2: Variation of the dispersion parameter D as a function of the wavelength of the light, taken from [34].

In general, the contribution of the waveguide dispersion is relatively small when compared to the material dispersion except near the zero-dispersion wavelength λ_0. For standard fibers, the effect of the waveguide dispersion is to shift λ_0 slightly towards the longer wavelengths, such as $\lambda_0 \approx 1.31\,\mu$m. It is important to mention that in Fig. 2.2, the quantity used to plot in the y-axis is the dispersion parameter D, which is related to β_2 as [34]:

$$D = -\frac{2\pi c}{\lambda^2}\beta_2 \approx \frac{\lambda d^2 n}{c d\lambda^2}. \tag{2.8}$$

It is worth mentioning that the waveguide dispersion is dependent on the fiber

design parameters such as the core radius a, and the core-cladding index difference Δ. The zero-dispersion wavelength λ_0 can be shifted to the vicinity of $1.55\,\mu$m by using this feature of the waveguide dispersion. It is also worthy of mentioning that the fiber loss is minimum at $1.55\,\mu$m wavelength range [34]. Such dispersion-shifted fibers are commonly used for communication systems.

The nonlinearity in optical fiber exhibits a different response behavior depending on the sign of the GVD parameter. For example, for wavelengths $\lambda < \lambda_0$, the fiber exhibits normal dispersion regime where $\beta_2 > 0$. In this regime, the low-frequency components travel faster than the high-frequency components of the same optical pulse. The fiber is said to be in an anomalous dispersion regime when $\beta_2 < 0$. In silica-based optical fibers, the anomalous dispersion regime occurs when the light wavelength exceeds the zero-dispersion wavelength, i.e., $\lambda > \lambda_0$. The anomalous dispersion regime is of particular interest in optical communication systems since it supports the existence of solitons through a balance between the dispersive and nonlinear effects [32].

The CD-induced mismatch in the group velocities of the optical pulses at different wavelengths causes a different speed for those pulses in the optical fiber [32]. This feature leads to an important effect referred to as the walk-off effect [32]. More specifically, when the fast-moving pulse completely walks through the slower moving pulse, the nonlinear interaction between two optical pulses ceases to occur [32]. That is governed by a walk-off parameter defined as [32]:

$$d_{12} = \beta_1(\lambda_1) - \beta_1(\lambda_2) = v_g^{-1}(\lambda_1) - v_g^{-1}(\lambda_2), \tag{2.9}$$

where λ_1 and λ_2 are the center wavelengths of the two pulses and $v_g = \frac{1}{\beta_1}$ is the group velocity and is calculated using (2.6). For pulses of width τ, the walk-off length can

be defined as [32]:

$$L_w = \frac{\tau}{|d_{12}|}. \tag{2.10}$$

2.3 Nonlinear Susceptibility and the Kerr Effect in Optical Fibers

2.3.1 Nonlinear Susceptibility

For intense electromagnetic fields, the response of the optical fiber is nonlinear. The origin of the nonlinear response is due to the anharmonic motion of bound electrons under the influence of the applied electromagnetic field [35]. The electric field of the incident light interacts with the electron and makes it oscillate in accordance with Coulomb's law [35]. The oscillating charge resembles an antenna and radiates the electromagnetic energy at the same frequency as the incident field with a different phase shift. The dynamics of the displaced electron under the influence of the applied electric field is a fundamental field of study in quantum mechanics [35].

On the other hand, in the classical electron oscillator model, the electron is modeled as a charged cloud surrounding the nucleus, as shown in Fig. 2.3(a). Fig. 2.3(b) shows that the electron charge cloud is displaced when an electric field E_x is applied [35]. According to Newton's law, the equation of motion for the center of the electron charge cloud can be represented as [35]:

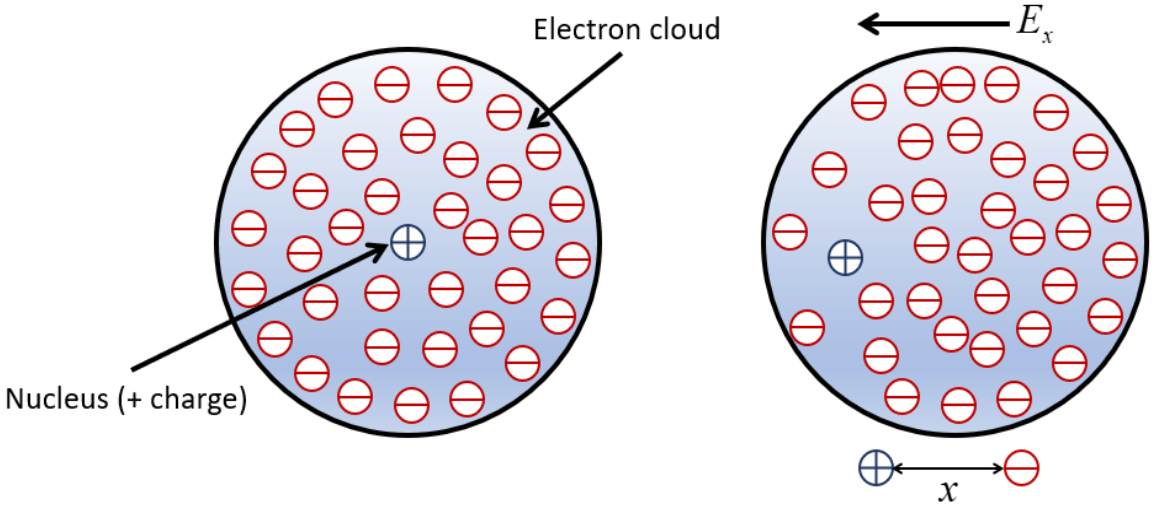
$$m\frac{d^2x}{dt^2} = F_{\text{ext}} = q_e E_x, \tag{2.11}$$

where m is the electron mass, $x(t)$ is the displacement, and q_e is the electron charge. There exists a force of attraction between the nucleus and the electron charge cloud when the electron charge cloud moves away for the equilibrium position [35]. For a

small displacement $x(t)$, the restoration force can be approximated as [35]:

$$F_{\text{restoration}} = -Kx, \tag{2.12}$$

where K is a constant.



(a) (b)

Figure 2.3: Classical electron oscillator model. (a) in equilibrium and (b) in the presence of an external field.

The negative sign in (2.12) indicates that the restoration force acts in a direction opposite to the external force [35]. This situation is similar to the case of a simple pendulum pushed away from the equilibrium position by an external force [35]. The restoration force due to the gravitation pulls back the pendulum to the equilibrium position. The net force acting on the electron can be represented as [35]:

$$F_{\text{net}} = F_{\text{ext}} + F_{\text{restoration}} = q_e E_x - Kx. \tag{2.13}$$

Next, by combining (2.11) and (2.13), we obtain:

$$m\frac{d^2x}{dt^2} = F_{\text{net}} = q_e E_x - Kx \tag{2.14}$$

or

$$\frac{d^2x}{dt^2} + w_0^2 x = \left(\frac{q_e}{m}\right) E_x, \tag{2.15}$$

where $w_0 = \sqrt{K/m}$ is the natural frequency of oscillation. Assume that the applied electric field is of the form as given below:

$$E_x = E_0 \exp\left(-jwt\right). \tag{2.16}$$

The displacement $x(t)$ due to the applied field also change harmonically in the steady state and try a trial solution as [35]:

$$x(t) = B \exp\left(-jwt\right). \tag{2.17}$$

Substituting (2.17) and (2.16) in (2.15), we get:

$$B = \frac{E_0 q_e}{m(w_0^2 - w^2)}, \tag{2.18}$$

and

$$x(t) = \frac{q_e}{m(w_0^2 - w^2)} E_x. \tag{2.19}$$

The dipole moment of an atom is given as [35]:

$$p_x = q_e x(t) = \frac{q_e^2}{m(w_0^2 - w^2)} E_x. \tag{2.20}$$

In general:

$$\mathbf{p} = \frac{q_e^2}{m(w_0^2 - w^2)}\mathbf{E}. \tag{2.21}$$

If we define $\mathbf{P} = N\mathbf{p}$ as the dielectric polarization, (2.21) can be written as:

$$\mathbf{P} = \frac{Nq_e^2}{m(w_0^2 - w^2)}\mathbf{E}. \tag{2.22}$$

where N is the number of atoms per unit volume.

For the weak incident electromagnetic field, we can relate the dielectric polarization $\mathbf{P}$ to the electric field intensity $\mathbf{E}$ as [35]:

$$\mathbf{P} = \varepsilon_0 \chi^{(1)}\mathbf{E}, \tag{2.23}$$

where ε_0 is the electric permittivity of free space and $\varepsilon_0\chi^{(1)}$ is the first-order susceptibility or linear susceptibility, and:

$$\chi^{(1)} = \frac{Nq_e^2}{m(w_0^2 - w^2)\varepsilon_0}. \tag{2.24}$$

It is important to mention that if the medium is not isotropic, the susceptibility depends on direction as well and (2.23) is modified as [35]:

$$P_j = \chi_{jx}^{(1)}E_x + \chi_{jy}^{(1)}E_y + \chi_{jz}^{(1)}E_z, \qquad j = x, y, z \tag{2.25}$$

or

$$\mathbf{P} = \varepsilon_0 \chi^{(1)}.\mathbf{E}, \tag{2.26}$$

where $\chi^{(1)}$ is a 3×3 matrix and $\cdot$ denotes dot product.

If the incident electromagnetic field is intense, the relation between the restoration

force and the displacement is nonlinear, and thereby, the electron cloud oscillation is not harmonic [35]. In this case, the relation between the dielectric polarization $\mathbf{P}$ and the electric dipoles are nonlinear, which can be generalized as [35]:

$$\mathbf{P} = \varepsilon_0 \left(\chi^{(1)} \cdot \mathbf{E} + \chi^{(2)} : \mathbf{EE} + \chi^{(3)} \vdots \mathbf{EEE} + \ldots \right), \tag{2.27}$$

where $\chi^{(j)}$ $(j = 1, 2, \ldots)$ is the jth order susceptibility. $\chi^{(j)}$ is also a tensor of rank $j+1$. $\chi^{(1)}$ is the first-order susceptibility and is related to the linear refractive index. The second-order susceptibility $\chi^{(2)}$ generates the second-harmonic and the sum-frequency terms [35]. However, since SiO_2 is a symmetric molecule, $\chi^{(2)}$ vanishes for the silica glasses [35]. As a result, the optical fibers do not exhibit the $\chi^{(2)}$-induced nonlinear effects [35].

2.3.2 The Kerr Effect

The nonlinearity effects in optical fiber originate from the third-order susceptibility $\chi^{(3)}$. One of the primary sources of the nonlinearity effect in optical fiber is the $\chi^{(3)}$-induced nonlinear refraction, the Kerr effect, a phenomenon referring to the light intensity-dependent refractive index [35]. Assume that the electromagnetic field incident on the optical fiber core has only E_x and H_y components. Then, the tensor equation in (2.27) can be simplified for a centrally symmetric dielectric material as [35]:

$$P_x = \varepsilon_0 \chi^{(1)}_{xx} E_x + \varepsilon_0 \chi^{(3)}_{xxxx} E_x^3, \tag{2.28}$$

where $\chi^{(3)}_{xxxx}$ is a component of the fourth-rank tensor $\chi^{(3)}$. Suppose, the incident optical field is a monochromatic wave given as:

$$E_x = E_0 \exp\left(-jwt\right). \tag{2.29}$$

To find E_x^3, we should first find the real part of E_x, i.e.,

$$\mathrm{Re}[E_x] = \frac{1}{2}\left[E_0 \exp\left(-jwt\right) + E_0^* \exp\left(jwt\right)\right], \tag{2.30}$$

$$\begin{aligned}\left\{\mathrm{Re}[E_x]\right\}^3 = \frac{1}{8}\Big\{&E_0^3 \exp\left(-j3wt\right) + E_0^{*3} \exp\left(j3wt\right) \\ &+3\left|E_0\right|^2\left[E_0 \exp\left(-jwt\right) + E_0^* \exp\left(jwt\right)\right]\Big\}. \end{aligned}\tag{2.31}$$

In the absence of the special phase-matching techniques, the third harmonic terms in (2.31) can be neglected [35].

Let the polarization at frequency w be:

$$P_x = P_0 \exp\left(-jwt\right), \tag{2.32}$$

Then,

$$\mathrm{Re}[P_x] = \frac{1}{2}\left[P_0 \exp\left(-jwt\right) + P_0^* \exp\left(jwt\right)\right]. \tag{2.33}$$

From (2.28), we have:

$$\mathrm{Re}[P_x] = \varepsilon_0 \chi_{xx}^{(1)} \mathrm{Re}[E_x] + \varepsilon_0 \chi_{xxxx}^{(3)} \mathrm{Re}[E_x]^3, \tag{2.34}$$

where the imaginary parts of the susceptibility are ignored. Substituting (2.30) and (2.31) into (2.34), collecting the terms that are proportional to $\exp\left(-jwt\right)$, and comparing it with (2.33), we obtain:

$$P_0 = \varepsilon_0 \left(\chi_{xx}^{(1)} + \frac{3\left|E_0\right|^2}{4}\chi_{xxxx}^{(3)}\right) E_0 = \varepsilon_0 \chi_{\mathrm{eff}} E_0, \tag{2.35}$$

where χ_{eff} is the effective susceptibility that includes both linear and nonlinear susceptibilities [35].

We can represent the electric field density D as [35]:

$$D = \varepsilon_0 E_0 + P_0. \tag{2.36}$$

Substituting (2.35) in (2.36), we can write:

$$D = \varepsilon_0 \left[1 + \chi_{xx}^{(1)} + \frac{3\left|E_0\right|^2}{4}\chi_{xxxx}^{(3)}\right] E_0. \tag{2.37}$$

In general, we can represent the electric flux density as [35]:

$$D = \varepsilon_0 \varepsilon_r E_0, \tag{2.38}$$

where ε_r is the relative permittivity. From (2.37) and (2.38), we can represent ε_r as:

$$\varepsilon_r = 1 + \chi_{xx}^{(1)} + \frac{3\left|E_0\right|^2}{4}\chi_{xxxx}^{(3)}. \tag{2.39}$$

Since the relative permittivity ε_r and refractive index n are related by $n^2 = \varepsilon_r$, we can write:

$$\begin{aligned} n^2 &= 1 + \chi_{xx}^{(1)} + \frac{3\left|E_0\right|^2}{4}\chi_{xxxx}^{(3)} \\ &= n_0^2 + \frac{3\left|E_0\right|^2}{4}\chi_{xxxx}^{(3)}, \end{aligned} \tag{2.40}$$

where n_0 is the linear refractive index and the second term of (2.40) represents the nonlinear contribution to the refractive index.

From (2.40), we can represent:

$$n = n_0 \left(1 + \frac{3\left|E_0\right|^2}{4n_0^2}\chi^{(3)}_{xxxx}\right)^{1/2}$$
$$\cong n_0 + n_2\left|E_0\right|^2, \tag{2.41}$$

where

$$n_2 = \frac{3\chi^{(3)}_{xxxx}}{8n_0}. \tag{2.42}$$

In (2.42), the term n_2 is called the Kerr coefficient [35]. For silica based fiber, the typical value of n_2 varies between 1.2×10^{-20} m^2/W $- 3.2 \times 10^{-20}$ m^2/W [35]. From (2.41), it is clear that the nonlinear part of the refractive index n is proportional to the optical intensity $\left|E_0\right|^2$. This effect is referred to as the Kerr effect [35].

2.4 Pulse Propagation in Optical Fibers

The wave propagation in dispersive nonlinear media is governed by the fundamental theory of electromagnetic wave propagation underpinned by the Maxwell's equations [32]. Using Maxwell's equations, one can easily show that:

$$\nabla \times \nabla \times \mathbf{E} = -\frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} - \mu_0 \frac{\partial^2 \mathbf{P}}{\partial t^2}, \tag{2.43}$$

where μ_0 is the permeability of the free space.

For the Kerr effect-based nonlinearity in optical fiber, the induced dielectric polarization $\mathbf{P}(\mathbf{r}, t)$ consists of two parts such as [32]:

$$\mathbf{P}(\mathbf{r}, t) = \mathbf{P}_L(\mathbf{r}, t) + \mathbf{P}_{\mathrm{NL}}(\mathbf{r}, t), \tag{2.44}$$

where $\mathbf{P}_L$ and $\mathbf{P}_{NL}$ are the linear and nonlinear parts which are given as:

$$\mathbf{P}_{\mathrm{L}}(\mathbf{r},t) = \varepsilon_0 \int_{-\infty}^{\infty} \chi^{(1)}(t-t^{'}).\mathbf{E}(\mathbf{r},t^{'})\ dt^{'}, \tag{2.45}$$

and

$$\mathbf{P}_{\mathrm{NL}}(\mathbf{r},t) = \varepsilon_0 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \chi^{(3)}(t-t_1,t-t_2,t-t_3)$$

$$\vdots\mathbf{E}(\mathbf{r},t_1)\mathbf{E}(\mathbf{r},t_2)\mathbf{E}(\mathbf{r},t_3)\ dt_1dt_2dt_3. \tag{2.46}$$

2.4.1 Nonlinear Schrödinger Equation

From (2.43) and (2.44), we can represent the wave equation as:

$$\nabla^2\mathbf{E} - \frac{1}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2} = \mu_0\frac{\partial^2\mathbf{P}_{\mathrm{L}}}{\partial t^2} + \mu_0\frac{\partial^2\mathbf{P}_{\mathrm{NL}}}{\partial t^2}, \tag{2.47}$$

where $\mathbf{P}_{\mathrm{L}}$ and $\mathbf{P}_{\mathrm{NL}}$ are given by (2.45) and (2.46), respectively. To solve the wave equation in (2.47), we adopt several simplifying assumptions including [32]:

1. The nonlinear part of the dielectric polarization $\mathbf{P}_{\mathrm{NL}}$ is treated as a small perturbation to the linear part $\mathbf{P}_{\mathrm{L}}$;
2. A scalar approach is adopted, i.e., the optical field is assumed to maintain the polarization along the length of the fiber; and
3. A quasi-monochromatic assumption for the optical field.

By adopting the slowly varying envelop approximation, we can represent the electric field $\mathbf{E}(\mathbf{r},t)$ as [32]:

$$\mathbf{E}(\mathbf{r},t) = \frac{1}{2}\hat{x}\left[E_0(\mathbf{r},t)\exp\left(-jw_0t\right) + E_0^*(\mathbf{r},t)\exp\left(jw_0t\right)\right], \tag{2.48}$$

where $\hat{x}$ is the polarization unit vector and w_0 is the center frequency of the optical pulse spectrum. The dielectric polarization components $\mathbf{P}_{\mathrm{L}}$ and $\mathbf{P}_{\mathrm{NL}}$ can also be expressed in similar form as [32]:

$$\mathbf{P}_{\mathrm{L}}(\mathbf{r},t) = \frac{1}{2}\hat{x}\left[P_{\mathrm{L}}(\mathbf{r},t)\exp\left(-jw_0t\right) + P^{*}_{\mathrm{L}}(\mathbf{r},t)\exp\left(jw_0t\right)\right], \tag{2.49}$$

and

$$\mathbf{P}_{\mathrm{NL}}(\mathbf{r},t) = \frac{1}{2}\hat{x}\left[P_{\mathrm{NL}}(\mathbf{r},t)\exp\left(-jw_0t\right) + P^{*}_{\mathrm{NL}}(\mathbf{r},t)\exp\left(jw_0t\right)\right]. \tag{2.50}$$

The linear component $\mathbf{P}_{\mathrm{L}}$ can be calculated by substituting (2.49) in (2.45) as [32]:

$$\begin{aligned}\mathbf{P}_{\mathrm{L}}(\mathbf{r},t) &= \varepsilon_0 \int_{-\infty}^{\infty} \chi^{(1)}_{xx}(t-t')E_0(\mathbf{r},t')\exp\left(jw_0(t-t')\right)\ dt' \\ &= \frac{\varepsilon_0}{2\pi}\int_{-\infty}^{\infty} \tilde{\chi}^{(1)}_{xx}(w)\tilde{E}_0(\mathbf{r},w-w_0)\exp\left(-j(w-w_0)t\right)\ dw,\end{aligned} \tag{2.51}$$

where $\tilde{E}_0(\mathbf{r},w)$ is the Fourier transform of $E_0(\mathbf{r},w)$.

The nonlinear component $\mathbf{P}_{\mathrm{NL}}$ is obtained by substituting (2.50) in (2.46) and after some simplifications we obtain:

$$\mathbf{P}_{\mathrm{NL}}(\mathbf{r},t) = \varepsilon_0\chi^{(3)}\vdots\mathbf{E}(\mathbf{r},t)\mathbf{E}(\mathbf{r},t)\mathbf{E}(\mathbf{r},t). \tag{2.52}$$

When (2.48) is substituted in (2.52) and following the analysis given in Subsection 2.3.1, we can approximate $\mathbf{P}_{\mathrm{NL}}$ as [32]:

$$\mathbf{P}_{\mathrm{NL}}(\mathbf{r},t) \approx \varepsilon_0\varepsilon_{\mathrm{NL}}E_0(\mathbf{r},t), \tag{2.53}$$

where

$$\varepsilon_{\mathbf{NL}} = \frac{3\chi^{(3)}_{xxxx}}{4}\left|E_0(\mathbf{r},t)\right|^2. \tag{2.54}$$

For simplicity, we adopt a frequency-domain analysis to derive the wave equation for the slowly varying amplitude $E_0(\mathbf{r}, t)$. Substituting (2.48)-(2.50) in (2.47), the Fourier transform $\tilde{E}_0(\mathbf{r}, w - w_0)$ can be defined as:

$$\tilde{E}_0(\mathbf{r}, w - w_0) = \int_{-\infty}^{\infty} E_0(\mathbf{r}, t) \exp\left(j(w - w_0)t\right) \, dt, \tag{2.55}$$

which is found to satisfy the Helmholtz equation as given below [32]:

$$\nabla^2 \tilde{E} + \varepsilon(w) k_0^2 \tilde{E} = 0, \tag{2.56}$$

where $k_0 = \frac{w}{c}$ and

$$\varepsilon(w) = 1 + \tilde{\chi}_{xx}^{(1)}(w) + \varepsilon_{\mathrm{NL}}. \tag{2.57}$$

Equation (2.56) can be solved using the method of separation of variables. Assuming the solution of the form:

$$\tilde{E}_0(\mathbf{r}, w - w_0) = F(x, y) Q(z, w - w_0) \exp\left(j\beta_0 z\right), \tag{2.58}$$

where $Q(z, w)$ is a slowly varying function of z and β_0 is the wave number. From (2.56), we can write two equations for $F(x, y)$ and $Q(z, w)$ as:

$$\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} + \left[\varepsilon(w) k_0^2 - \tilde{\beta}^2\right] F = 0, \tag{2.59}$$

and

$$2j\beta_0 \frac{\partial Q}{\partial z} + \left(\tilde{\beta}^2 - \beta_0^2\right) Q = 0. \tag{2.60}$$

The dielectric constant $\varepsilon(w)$ can be approximated as:

$$\varepsilon = (n_0 + \Delta n)^2 \approx n_0^2 + 2n_0 \Delta n, \tag{2.61}$$

where Δn is a small perturbation given by:

$$\Delta n = n_2 \left|E_0\right|^2 + \frac{j\tilde{\alpha}}{2k_0}, \tag{2.62}$$

where $\tilde{\alpha}$ is the absorption coefficient.

The FO perturbation theory can be used to solve (2.59). After solving using the FO perturbation theory, the value of $\tilde{\beta}$ can be represented as [32]:

$$\tilde{\beta}(w) = \beta(w) + \Delta\beta, \tag{2.63}$$

where

$$\Delta\beta = \frac{k_0 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Delta n \left|F(x,y)\right|^2 \, dxdy}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left|F(x,y)\right|^2 \, dxdy}. \tag{2.64}$$

Similarly, the electric field $\mathbf{E}(\mathbf{r},t)$ can be written as [32]:

$$\mathbf{E}(\mathbf{r},t) = \frac{1}{2}\hat{x}\left[F(x,y)q(z,t)\exp\left(j(\beta_0 z - w_0 t)\right)\right. \\ \left. + F^*(x,y)q^*(z,t)\exp\left(-j(\beta_0 z - w_0 t)\right)\right], \tag{2.65}$$

where $q(z,t)$ is the slowly varying pulse envelop [32].

The Fourier transform $Q(z, w-w_0)$ of $q(z,t)$ satisfies (2.60), which can be represented as [32]:

$$\frac{\partial Q}{\partial z} = j\left[\beta(w) + \Delta\beta - \beta_0\right]Q. \tag{2.66}$$

The propagation equation for $q(z,t)$ is obtained by taking the inverse Fourier transform of (2.66). In (2.66), the exact functional form of the mode-propagation constant $\beta(w)$ is rarely known and therefore, it is useful to expand $\beta(w)$ in Taylor series as in (2.5) [32]. Because of the quasi-monochromatic assumption for the optical field, the cubic and higher-order terms in (2.5) can be neglected [32]. Next, substitute (2.5) in

(2.66) and take the inverse Fourier transform, we obtain:

$$\frac{\partial}{\partial z}q(z,t) = -j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}q(z,t) + j\Delta\beta q(z,t). \tag{2.67}$$

The term $\Delta\beta$ in (2.67) includes the effect of loss and the nonlinearity [32]. Next, evaluate $\Delta\beta$ using (2.62) and (2.64), and substitute in (2.67), we obtain the nonlinear Schrödinger equation (NLSE) as:

$$\frac{\partial}{\partial z}q(z,t) + \frac{\alpha}{2}q(z,t) + j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}q(z,t) = j\gamma\left|q(z,t)\right|^2 q(z,t), \tag{2.68}$$

where α is the attenuation, β_2 is the group velocity dispersion, z is the transmission distance, and γ is the nonlinearity coefficient [32], which is given as:

$$\gamma = \frac{n_2 w_0}{cA_{\text{eff}}}, \tag{2.69}$$

where A_{eff} is the effective area of the fiber and is given as:

$$A_{\text{eff}} = \frac{\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|F(x,y)\right|^2\,dxdy\right)^2}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|F(x,y)\right|^4\,dxdy}. \tag{2.70}$$

The evaluation of A_{eff} requires the use of modal distribution $F(x,y)$ for the fundamental fiber mode [32]. It is worth mentioning that the typical value of A_{eff} varies in the range $20-100\ \mu\text{m}^2$ in the 1550 nm region depending on the fiber design [32]. As a result, the nonlinearity coefficient γ takes values in the range $1-10\ \text{W}^{-1}/\text{km}$ if $n_2 \approx 2.6\times10^{-20}\ \text{m}^2/\text{W}$ [32].

The NLSE in (2.68) can be further simplified by introducing a normalized field $u(z,t)$ referred to the delayed time frame corresponding to the group velocity v_g. Thus, by applying the transformation $q(z,t) \triangleq u(z,t)\exp(-\frac{\alpha}{2}z)$, the (2.68) can be modified

as [32]:

$$\frac{\partial}{\partial z}u(z,t) + j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}u(z,t) = j\gamma\left|u(z,t)\right|^2 u(z,t)\exp(-\alpha z). \tag{2.71}$$

2.4.2 Manakov Equation

For dual-polarization transmission systems, the coupled NLSE (CNLSE) provides an accurate model for the nonlinear pulse propagation in optical fiber [32]. It includes the PMD effects in the fiber along with the dispersive and nonlinearity effects. In general, similar to NLSE, the CNLSE also needs to be solved numerically [32]. However, the different length scales associated with PMD, GVD, and nonlinearity effects make the numerical evaluation cumbersome. The dispersive and nonlinearity effects vary on a length scale from 10 to 100 km, while the birefringence in optical fiber varies on a length scale of 10 to 100 m [32]. Therefore, the step size used in the numerical evaluation of CNLSE must be less than 1 m [32]. That increases the computation time. Based on this fact, we adopt an approximation method to solve the CNLSE. It is observed that the birefringence fluctuation changes the SOP of the optical field on a short length scale that the field covers the entire Poincaré sphere after a few kilometers [32]. As a result, the nonlinearity terms in CNLSE can be averaged over the birefringence fluctuations [32]. The resultant propagation equation is referred to as the Manakov equation and can be represented as:

$$\frac{\partial}{\partial z}\mathbf{u}(z,t) + j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}\mathbf{u}(z,t) = j\frac{8}{9}\gamma\left|\mathbf{u}(z,t)\right|^2\mathbf{u}(z,t)\exp\left(-\alpha z\right), \tag{2.72}$$

where $\mathbf{u}(z,t) = [u_x(z,t) \;\; u_y(z,t)]^\dagger$ and the superscript $\dagger$ represents the transpose [32]. It is clear from (2.72) that the rapid random variations in the SOP of the optical field reduce the effect of the nonlinearity parameter γ by a factor of $\frac{8}{9}$.

For WDM systems, the optical field $\mathbf{u}(z,t)$ in (2.72) consists of the sum of all the

multiplexed channels, which can be represented as:

$$\mathbf{u}(z,t) = \sum_{n} \mathbf{u}_n(z,t) \exp\left(j w_n t\right). \tag{2.73}$$

The Manakov equation for the WDM system consists of both intra- and inter-channel nonlinearity effects such as SPM, XPM, and FWM [32]. By substituting (2.73) in (2.72), we can represent the propagation equation for the ith channel as:

$$\begin{aligned}\frac{\partial}{\partial z}u_{x/y,i} + j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}u_{x/y,i} = j\frac{8}{9}\gamma \exp\left(-\alpha z\right) \Bigg(& \underbrace{\left(\left|u_{x/y,i}\right|^2 + \left|u_{y/x,i}\right|^2\right) u_{x/y,i}}_{\text{SPM}} \\ & \underbrace{\underbrace{\sum_{m\neq i}\left(2\left|u_{x/y,m}\right|^2 + \left|u_{y/x,m}\right|^2\right) u_{x/y,i}}_{\text{Coherent XPM}} + \underbrace{\sum_{m\neq i}\left(u^*_{y/x,m}u_{x/y,m}\right) u_{y/x,i}}_{\text{Incoherent XPM}}}_{\text{XPM}} \\ & + \underbrace{\sum_{p',q',r'}\sum_{k\neq l,m}\sum_{l\neq i}\sum_{m\neq i} u^*_{p',k}u_{q',l}u_{r',m}}_{\text{FWM}} \Bigg), \end{aligned} \tag{2.74}$$

where $p^{'}, q^{'}$, and $r^{'}$ takes the x and y polarization tributaries. For simplicity, the space and time variables z, t are not shown in (2.74).

2.5 Mathematical Tools for Solving the Propagation Equation

2.5.1 Numerical Approach

The NLSE in (2.71) cannot be solved analytically, except for some special cases [32]. Numerical approaches are typically used to solve the propagation equation in (2.71).

The most commonly used numerical method to solve the pulse propagation problem in nonlinear and dispersive media is the split-step Fourier method (SSFM) [32].

2.5.1.1 The Split-step Fourier Method

The NLSE in (2.71) can be rearranged by separating the linear and nonlinear parts as [32]:

$$\frac{\partial}{\partial z}u(z,t) = \left(\hat{D} + \hat{N}\right) u(z,t) \tag{2.75}$$

$$\hat{D} = -j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2} \tag{2.76}$$

$$\hat{N} = j\gamma \left|u(z,t)\right|^2 \exp(-\alpha z), \tag{2.77}$$

where $\hat{D}$ and $\hat{N}$ are the linear and nonlinear operators [32].

In the SSFM, the nonlinear and dispersive signal propagation in the optical fiber is iteratively modeled by dividing the fiber spans into small segments, each having a length of h. More specifically, the signal propagation from z to $z + h$ is carried out in two separate steps [32]. First the linear operator $\hat{D}$ is set to zero and only the nonlinearity is taken into account. Second, the nonlinear operator $\hat{N}$ is set to zero and only the dispersion is taken into account.

Mathematically:

$$u(z+h,t) \approx \exp\left(h\hat{D}\right) \exp\left(h\hat{N}\right) u(z,t). \tag{2.78}$$

In (2.78), the step size h is chosen small enough such that the nonlinear and linear effects in each segment can be modeled as acting independently [32].

The accuracy of the SSFM can be improved by using a symmetric SSFM, which

can be represented as follows [32]:

$$u(z+h,t) \approx \exp\left(\frac{h}{2}\hat{D}\right) \exp\left(\int_{z}^{z+h} \hat{N}(z')dz'\right) \exp\left(\frac{h}{2}\hat{D}\right) u(z,t). \tag{2.79}$$

In this method, the nonlinearity is included in the middle of the fiber segment rather than at the segment boundaries, as shown in Fig. 2.4 [32].

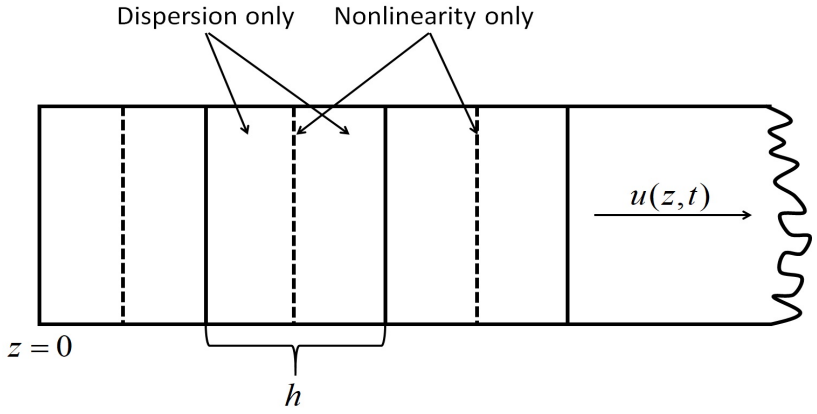


Figure 2.4: Schematic illustration of the symmetric SSFM.

The implementation of the symmetric SSFM is relatively straightforward [32]. The fiber span is divided into a large number of segments of size h. The optical signal is propagated from segment to segment as formulated in (2.79). More specifically, the optical field envelop $u(z,t)$ is first propagated for a distance $h/2$ with only dispersion effect. At the midplane, the optical field is multiplied by the nonlinear term that represents the nonlinearity effect for the whole segment of length h. Finally, the optical field is propagated the remaining distance of length $h/2$ with only dispersion effect to obtain the optical field $u(z+h,t)$ [32].

2.5.2 Analytical Approach

The direct numerical solution of the NLSE using the SSFM was initially adopted as the key design tool by the optical communication research community [32]. However, the implementation complexity of the SSFM-based nonlinear fiber propagation modeling was found to be impractically high [32]. That led to increased interest in research for the simplified versions of the NLSE for which an approximate analytical solution is available [32]. The most commonly used analytical approximation methods to solve the NLSE are Volterra series-based method and the RP theory-based method.

2.5.2.1 Volterra Series-based Analysis

In contrast to SSFM, solutions of the NLSE can be analytically approximated using the Volterra series, which is a well-established tool in nonlinear systems theory [36]-[42]. In a Volterra series-based approach, the input-output relationship of a nonlinear fiber channel can be represented by a series of nonlinear kernel functions, referred to as Volterra series transfer functions (VSTFs) [36]- [42]. The VSTF provides the relationship between the Fourier transforms of the input to the fiber $X(w)$ and the output at the end of the fiber link $Y(w)$, as [36]- [42]:

$$Y(w) = H_1(w)X(w) + \sum_{n=2}^{\infty} \int \cdots \int H_n(w_1, \cdots, w_{n-1}, w) \times X(w_1) \cdots X(w_{n-1}) X(w - w_1 - \cdots - w_{n-1})\, dw_1 \cdots dw_{n-1}, \quad (2.80)$$

where $H_n(w_1, \cdots, w_{n-1}, w)$ is the nth order VSTF kernel in the frequency-domain. Due to the symmetries in the silica-based optical fiber the even-ordered Volterra kernels are zero.

For a PDM transmission system, the VSTF kernels up to third-order can be repre-

sented as [36]- [42]:

$$U_x(z,w) = H_1(z,w)U_x(w) + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H_3(z,w_1,w_2,w) \times \Big[U_x(w_1)U_x^*(w_2) + U_y(w_1)U_y^*(w_2)\Big] \times U_x(w-w_1+w_2)\,dw_1dw_2, \quad (2.81)$$

where $U_x(w) \triangleq U_x(z=0,w)$ and VSTF kernels can be given as:

$$H_1(z,w) = \exp\left(-\frac{\alpha}{2} - \frac{jw^2\beta_2}{2}\right)z, \quad (2.82)$$

$$H_3(z,w_1,w_2,w) = j\frac{8}{9}\frac{\gamma}{4\pi^2}H_1(z,w)\frac{1-\exp\left(-(\alpha+j\beta_2(w-w_1)(w_1-w_2))z\right)}{\alpha+j\beta_2(w-w_1)(w_1-w_2)}. \quad (2.83)$$

Substituting (2.82) and (2.83) in (2.81) and detaching $H_1(z,w)$ from $H_3(z,w_1,w_2,w)$, we can modify (2.81) as [36]- [42]:

$$U_x(z,w) = H_1(z,w)U_x(w) + H_1(z,w)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_3(z,w_1,w_2,w) \times \Big[U_x(w_1)U_x^*(w_2) + U_y(w_1)U_y^*(w_2)\Big] \times U_x(w-w_1+w_2)\,dw_1dw_2, \quad (2.84)$$

where K_3 is defined as:

$$K_3(z,w_1,w_2,w) = j\frac{8}{9}\frac{\gamma}{4\pi^2}\frac{1-\exp\left(-(\alpha+j\beta_2(w-w_1)(w_1-w_2))z\right)}{\alpha+j\beta_2(w-w_1)(w_1-w_2)}. \quad (2.85)$$

It is understood from (2.84) that the pulse propagation in SSMF is modeled by the combination of a linear kernel and a third-order nonlinear kernel. The linear kernel $H_1(z,w)$ models the dispersion and attenuation effect in the optical fiber; whereas, the

third-order nonlinear kernel $H_3(z, w_1, w_2, w)$ models the beating of various frequency components present in the input signal. Nonetheless, the computational complexity of the Volterra series-based modeling may approach that of the SSFM, when applied to long-haul optical fiber links [36]- [42].

2.5.2.2 Perturbation Theory-based Analysis

In contrast to Volterra series-based approach, the regular perturbation (RP) theory-based method provides an approximate analytical solution of the NLSE in a computationally efficient way. The RP method provides a recursive closed-form solution for the NLSE that gives a good insight into the nature of the interaction between the accumulated CD and the Kerr nonlinearity in the optical fiber channel. It is important to note that the derivation given in this subsection is adopted from [43], [44], and [46].

In the RP method, the optical field $u(z,t)$ is expressed in a power series of the nonlinearity coefficient γ as $u(z,t) = \sum_{k'=0}^{\infty} \gamma^{k'} u_{k'}(z,t)$, where k' is the order of the perturbative solution. Then, substituting $u(z,t)$ in (2.71), we get:

$$\sum_{k'=0}^{\infty} \gamma^{k'} \frac{\partial}{\partial z} u_{k'}(z,t) = -\sum_{k'=0}^{\infty} \gamma^{k'} j \frac{\beta_2}{2} \frac{\partial^2}{\partial t^2} u_{k'}(z,t) + j\gamma \sum_{m=0}^{\infty} \sum_{l=0}^{\infty} \sum_{n=0}^{\infty} \gamma^{m+l+n} \times u_m(z,t) u_l^*(z,t) u_n(z,t) \exp(-\alpha z). \quad (2.86)$$

From (2.86), a system of recursive linear differential equations is obtained by equating the terms that multiply equal powers of γ on both sides of the equal sign. The differential equation governing the k'^{th}- order solution can be represented as:

$$\frac{\partial}{\partial z} u_{k'}(z,t) = -j \frac{\beta_2}{2} \frac{\partial^2}{\partial t^2} u_{k'}(z,t) + j \sum \sum_{m+l+n=k'-1} \sum u_m(z,t) u_l^*(z,t) u_n(z,t) \exp(-\alpha z). \quad (2.87)$$

Zeroth-order (or Linear) Solution

The differential equation governing the zeroth-order (or linear) solution is obtained by substituting $k' = 0$ in (2.87), which can be represented as:

$$\frac{\partial}{\partial z}u_0(z,t) = -j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}u_0(z,t). \tag{2.88}$$

By solving (2.88), the zeroth-order solution at a transmission length $z = L$ is obtained as:

$$u_0(L,t) = u(0,t) \otimes h_L(t), \tag{2.89}$$

where $\otimes$ is convolution operation, $h_L(t) = \mathcal{F}^{-1}\{\exp(-j\frac{w^2\beta_2 L}{2})\} = \frac{1}{\sqrt{-2\pi j\beta_2 z}}\exp\left(\frac{-jt^2}{2\beta_2 z}\right)$ at the angular frequency w, and $\mathcal{F}^{-1}\{.\}$ is the inverse Fourier transform operation.

First-order Solution

By substituting $k' = 1$ in (2.8), the differential equation governing the first-order (FO) solution can be represented as:

$$\frac{\partial}{\partial z}u_1(z,t) = -j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}u_1(z,t) + j\left|u_0(z,t)\right|^2 u_0(z,t)\exp(-\alpha z). \tag{2.90}$$

The FO distortion field at a transmission distance $z = L$ is obtained by solving (2.90) and assuming an ideal dispersion compensation at $z = L$, we get:

$$u_1(L,t) = j\gamma\int_0^L \left(h_z(t) \otimes[\left|u_0(z,t)\right|^2 u_0(z,t)]\right)\exp(-\alpha z)dz. \tag{2.91}$$

By taking the Fourier transform of (2.91), we get the distortion field in frequency-domain at a transmission distance $z = L$ as:

$$U_1(L, w) = j\gamma \int_0^L F(z, w) \exp(-j\frac{w^2\beta_2 z}{2}) \exp(-\alpha z) dz, \tag{2.92}$$

where $F(z, w)$ is given as:

$$F(z, w) = \int_{-\infty}^{\infty} |u_0(z, t)|^2 u_0(z, t) \exp(-jwt)\ dt. \tag{2.93}$$

The input field to the optical fiber can be expressed as:

$$u(z = 0, t) = \sqrt{P_0} \sum a_k g(z = 0, t - kT), \tag{2.94}$$

where P_0 is the peak launch power, a_k is the data symbol of the kth pulse, $g(z, t)$ is the pulse temporal waveform at transmission distance z, and T is the symbol duration. Therefore, the product $|u_0(z, t)|^2 u_0(z, t)$ in (2.93) can be represented as:

$$\begin{aligned} |u_0(z, t)|^2 u_0(z, t) &= P_0^{3/2} \sum_m \sum_l \sum_n a_m g(z, t - mT) a_l^* g^*(z, t - lT) a_n g(z, t - nT) \\ &= P_0^{3/2} \sum_m \sum_l \sum_n a_m g(z) a_l^* g^*(z) a_n g(z). \end{aligned} \tag{2.95}$$

Substituting (2.95) in (2.93), we obtain [44]:

$$
\begin{aligned}
F(z,w) &= P_0^{3/2} \int_{-\infty}^{\infty} \left(\sum_m \sum_l \sum_n a_m g(z) a_l^* g^*(z) a_n g(z) \right) \exp\left(-jwt\right)\, dt \\
&= P_0^{3/2} \sum_m \sum_l \sum_n a_m a_l^* a_n \left(G_m(z,w) \otimes G_l^*(z,-w) \otimes G_n(z,w) \right) \\
&= P_0^{3/2} \sum_m \sum_l \sum_n a_m a_l^* a_n \exp\left(j \frac{w^2 \beta_2 z}{2} \right) \exp\left(-jw\left(T_m - T_l + T_n\right)\right) \\
&\times \int\int G(0, w_1 + w) G^*(0, w_1 + w - w_2) G(0, w - w_2) \\
&\times \exp\left(j\beta_2 z w_1 w_2\right) \exp\left(-j\left(w_1(T_m - T_l) + w_2(T_l - T_n)\right)\right)\, dw_1 dw_2, \qquad (2.96)
\end{aligned}
$$

where $G(z,w) = G(0,w)\exp\left(j\frac{w^2\beta_2 z}{2}\right)$ with $G(0,w)$ is the Fourier transform of $g(0,t)$. Substituting (2.96) in (2.92), we get FO perturbation kernel term in frequency-domain as:

$$
\begin{aligned}
U_1(L,w) &= j\gamma P_0^{3/2} \sum_m \sum_l \sum_n a_m a_l^* a_n \exp\left(-jw\left(T_m - T_l + T_n\right)\right) \\
&\times \int_0^L \exp\left(-\alpha z\right) \left(\int\int G(0, w_1 + w) G^*(0, w_1 + w - w_2) G(0, w - w_2) \right. \\
&\left. \times \exp\left(-j\left(w_1(T_m - T_l) + w_2(T_l - T_n)\right)\right) \exp\left(j w_1 w_2 \beta_2 z\right)\, dw_1 dw_2 \right) dz. \qquad (2.97)
\end{aligned}
$$

Next, by taking the inverse Fourier transform of (2.97), the time-domain FO kernel term can be represented as:

$$
\begin{aligned}
u_1(L,(t + T_m - T_l + T_n)) &= j\gamma P_0^{3/2} \sum_m \sum_l \sum_n a_m a_l^* a_n \int_0^L \exp\left(-\alpha z\right) \left(\int\int\int \right. \\
G(0, w_1 + w) G^*(0, w_1 + w - w_2) G(0, w - w_2) &\exp\left(-j\left(w_1(T_m - T_l) + w_2(T_l - T_n)\right)\right) \\
&\left. \times \exp\left(j w_1 w_2 \beta_2 z\right) \exp\left(jwt\right)\, dw_1 dw_2 dw \right) dz. \qquad (2.98)
\end{aligned}
$$

Assuming the Gaussian shape for input pulses, i.e., $G(0,w)=\sqrt{2\pi\tau^2}\exp(-\frac{w^2\tau^2}{2})$, with τ as the pulse width, the product of the triplet pulses in (2.98) can be represented as:

$$G(0,w_1+w)G^*(0,w_1+w-w_2)G(0,w-w_2)=\left(\sqrt{2\pi\tau^2}\right)^3\exp(-\frac{3\tau^2w^2}{2}) \times\exp(-\tau^2[w_1^2+w_2^2+2(w_1-w_2)w-w_1w_2]). \quad (2.99)$$

It is important to mention that with the Gaussian shape assumption for the input pulse shape, the FO nonlinearity coefficients can be calculated using analytic expressions, which involve the exponential integral function [45]. This will explain in detail in subsection 2.7.3.

Substituting (2.99) in (2.98) and integrating w.r.t w, we get:

$$u_1(L,(t+T_m-T_l+T_n))=j\gamma P_0^{3/2}\frac{2\pi\tau^2}{\sqrt{3}}\exp\left(-\frac{t^2}{6\tau^2}\right)\sum_m\sum_l\sum_n a_m a_l^* a_n\int_0^L\int\int \times\exp(-\alpha z)\exp\left(-\frac{1}{3}\tau^2(w_1^2+w_2^2+w_1w_2)\right)\exp\left(-j(\frac{2}{3}(w_1-w_2)t+w_1(T_m-T_l) +w_2(T_l-T_n)-w_1w_2\beta_2 z)\,dw_1dw_2dz\right). \quad (2.100)$$

Next, collecting the terms corresponding to w_1 and integrating over w and w_1, we obtain:

$$u_1(L,(t+T_m-T_l+T_n))=j\gamma P_0^{3/2}\sqrt{\pi\tau^2}\exp\left(-\frac{t^2}{6\tau^2}\right)\sum_m\sum_l\sum_n a_m a_l^* a_n\int_0^L\int \times\exp(-\alpha z)\exp\left(-\frac{1}{4}\tau^2w_2^2\right)\exp\left(jw_2\left[t+\frac{(T_m-T_l)-w_2\beta_2 z}{2}+(T_n-T_l)\right]\right) \times\exp\left(-\frac{3\left[\frac{2}{3}t+(T_m-T_l)-w_2\beta_2 z\right]^2}{4\tau^2}\right)dw_2dz. \quad (2.101)$$

Collecting terms corresponding to w_2 and integrating, we get the time-domain perturbation kernel as:

$$u_1(L,(t+T_m-T_l+T_n)) = j\gamma P_0^{3/2}\sum_m\sum_l\sum_n a_m a_l^* a_n \exp\left(-\frac{t^2}{6\tau^2}\right) \times \int_0^L \frac{\tau^2\exp(-\alpha z)}{\sqrt{\tau^4+2j\tau^2\beta_2 z+3(\beta_2 z)^2}} \times \exp\left(-\frac{\tau^2}{\tau^4+2j\tau^2\beta_2 z+3(\beta_2 z)^2}\left[t+\frac{(T_m-T_l)}{2}+(T_n-T_l)\right]^2\right) \times \exp\left(-\frac{3\left[\tau^4+2j\tau^2\beta_2 z\right]}{4\tau^2\left[\tau^4+2j\tau^2\beta_2 z+3(\beta_2 z)^2\right]}\left(\frac{2}{3}t+(T_l-T_m)\right)^2\right) \times \exp\left(j\frac{3\left(\frac{2}{3}t+(T_l-T_m)\right)\beta_2 z}{\tau^4+2j\tau^2\beta_2 z+3(\beta_2 z)^2}\left[t+\frac{(T_m-T_l)}{2}+(T_n-T_l)\right]\right) dz. \quad (2.102)$$

After carrying out some algebraic simplifications in (2.102), we obtain the FO distortion field (or FO ghost pulse) as:

$$u_1(L,t+kT) = j\gamma P_0^{3/2}\sum_m\sum_l\sum_n a_m a_l^* a_n \exp\left(-\frac{t^2}{6\tau^2}\right)\int_0^L \frac{\exp(-\alpha z)}{\sqrt{1+2j\beta_2 z/\tau^2+3(\beta_2 z/\tau^2)^2}} \times \exp\left\{\begin{array}{l} -\frac{3\left[\frac{2}{3}t+(m-l)T\right]\left[\frac{2}{3}t+(n-l)T\right]}{\tau^2(1+3j\beta_2 z/\tau^2)} \\ -\frac{(n-m)^2T^2}{\tau^2[1+2j\beta_2 z/\tau^2+3(\beta_2 z/\tau^2)^2]}\end{array}\right\} dz, \quad (2.103)$$

where T_m, T_l, and T_n can be represented as mT, lT, and nT, respectively, $k = m+n-l$, m, n, l are the symbol indices, $a_{m/l/n}$ is the symbol complex amplitude, and τ is the pulse width.

In the PDM transmission system, the electric field input to the optical fiber is a column vector $\mathbf{u}(z,t) = [u_x(z,t)\ u_y(z,t)]^\dagger$, with x, y representing the horizontal and vertical polarization, respectively, and the superscript $\dagger$ as the transpose. The propagation of the vector field $\mathbf{u}(z,t)$ through the optical fiber can be represented using the Manakov equation, where the nonlinear effective length is much longer than

the fiber birefringent beating length, as [47]:

$$\frac{\partial}{\partial z}\mathbf{u} + j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}\mathbf{u} = j\frac{8}{9}\gamma(\mathbf{u}^{*\dagger}\mathbf{u}\mathbf{I})\mathbf{u}\exp\left(-\alpha z\right), \tag{2.104}$$

where $\mathbf{I}$ is the identity matrix. It is important to note that (2.104) is the same as the Manakov equation given in (2.72), in which (2.104) is written in a more compact form by omitting the space and time variables z, t for the sake of simplicity. After solving (2.104), the zeroth- and FO solutions for the output field can be represented as:

$$u_{0,x/y}(L,t) = u_{x/y}(0,t) \otimes h_L(t), \tag{2.105}$$

and

$$u_{1,x/y}(L,t) = j\gamma \int_0^L \left(h_z(t) \otimes [\left| u_{0,x/y}(z,t) \right|^2 u_{0,x/y}(z,t)] \right) \exp(-\alpha z) dz. \tag{2.106}$$

Following the steps from (2.92)-(2.103), we get the FO ghost pulse for the PDM transmission system as:

$$\begin{aligned} u_{1,x/y}(L, t+kT) = j\frac{8}{9}\gamma P_0^{3/2} \sum_m \sum_l \sum_n \left[a_{m,x/y} a^*_{l,x/y} a_{n,x/y} + a_{m,y/x} a^*_{l,y/x} a_{n,x/y} \right] \\ \times \exp(-\frac{t^2}{6\tau^2}) \int_0^L \frac{\exp(-\alpha z)}{\sqrt{1 + 2j\beta_2 z/\tau^2 + 3(\beta_2 z/\tau^2)^2}} \\ \times \exp\left\{ \begin{array}{c} -\frac{3\left[\frac{2}{3}t + (m-l)T\right]\left[\frac{2}{3}t + (n-l)T\right]}{\tau^2(1+3j\beta_2 z/\tau^2)} \\ -\frac{(n-m)^2 T^2}{\tau^2[1+2j\beta_2 z/\tau^2+3(\beta_2 z/\tau^2)^2]} \end{array} \right\} dz. \end{aligned} \tag{2.107}$$

2.6 Kerr-induced Fiber Nonlinearity Effects

As stated in Section 2.3, the nonlinearity effects in the optical fiber are due to the Kerr effect, which arises from the dependence of the optical fiber refractive index on

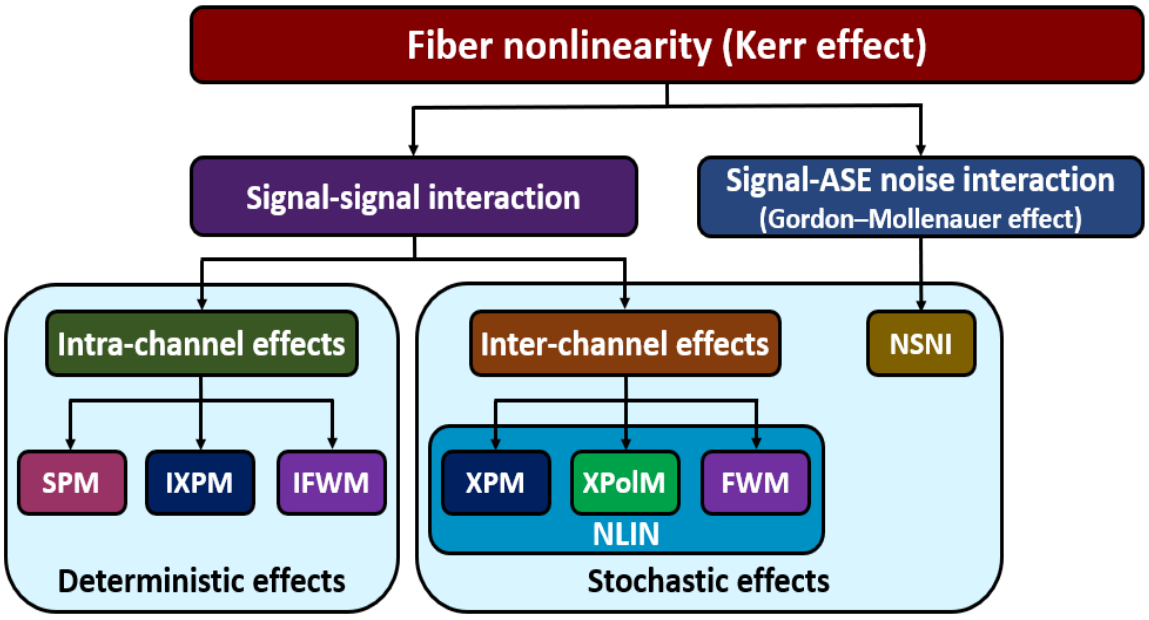


Figure 2.5: Fiber nonlinearity effects. SPM: self-phase modulation, (I)XPM: (intra-channel) cross-phase modulation, XPolM: cross-polarization modulation, (I)FWM: (intra-channel) four wave mixing, NLIN: nonlinear interference noise, ASE: amplified spontaneous emission, and NSNI: nonlinear signal-noise interaction.

the transmit signal power [32]. Fig. 2.5 illustrates different types of Kerr nonlinearity effects in the optical fiber. The nonlinearity effects can be divided into two types: signal-signal and signal-ASE noise nonlinear interaction [5]. In single-channel systems, the intra-channel signal-signal nonlinear interactions can be categorized into three types: (i) self-phase modulation (SPM), (ii) intra-channel cross-phase modulation (IXPM), and (iii) intra-channel four-wave mixing (IFWM). The SPM results in a phase modulation induced by the refractive index, which increases with the increase in the input power level. This leads to a frequency shift, known as frequency chirping [5], which interacts with the CD and causes the spectral broadening of the optical pulse. The IXPM is the result of the refractive index change proportional to the intensity of the neighboring pulses in the same channel [5]. The IXPM yields a timing jitter between the co-propagating pulses, which leads to performance degradation. The

IFWM is caused by the nonlinear interaction between two or more pulses of the same channel [5]. That generates echo or ghost pulses in the time domain, and thus, results in interference between the signal pulses of the same channel. It is important to mention that the SPM, IXPM, and IFWM are considered as deterministic nonlinear impairments, and these effects can be compensated well using digital NLC techniques [5].

In WDM systems, the inter-channel signal-signal nonlinear effects can also be classified into three types: (i) XPM, (ii) cross-polarization modulation (XPolM), and (iii) FWM. The XPM effect is due to the refractive index change proportional to the intensity of the pulse in the co-propagating channel. The XPM induces frequency chirping and pulses overlapping between channels [5]. The XPM effect consists of two parts, the coherent and incoherent XPM, as given (2.74). The coherent XPM is the same as the usual XPM effect, which involves the interaction of two WDM channels. The incoherent XPM results in polarization cross-talk, and the cross-talk coefficient is determined by the polarization cross-product $u^*_{y/x,m}u_{x/y,m}$ of the interfering channel [5]. The XPolM occurs in PDM transmission systems when the SOP of a transmitted channel depends on the SOP of other co-propagating channels. That is due to the random propagation of SOP inside the optical fiber caused by PMD [5] and causes channel cross-talk for dual-polarization systems [5]. The FWM in the multichannel system is generated by the nonlinear interaction between two or more co-propagating wavelength channels [5]. The FWM results in significant performance degradation due to cross-talk among different wavelength channels [5].

Another nonlinearity effect is the Kerr-induced signal-ASE noise interaction, referred to as the Gordon–Mollenauer effect [26], [28], which leads to the generation of the stochastic nonlinear signal-noise interaction (NSNI) [26], [28]. The modeling of the impact of NSNI is important because it can contribute to answering the questions

regarding the fundamental limits of the performance improvement provided by the NLC techniques [26], [28].

Table 2.1 summarizes the variation of the Kerr-induced signal-signal fiber nonlinearity effects as a function of the bit rate and channel spacing in a WDM optical transmission system [5]. Note that, in Table 2.1, the upward-pointing arrow represents the increase in value, whereas a downward pointing arrow shows a decrease in value.

Table 2.1: Kerr-induced signal-signal fiber nonlinearity effects versus bit rate and channel spacing, taken from [5].

Type	SPM	XPM	XPolM	FWM
Bit rate ↗	↗	↗	↗	no effect
Channel spacing ↘	no effect	↗	↗	↗

2.7 Fiber Nonlinearity Compensation Techniques

The impairments due to the Kerr effect degrade the transmission performance of optical systems using higher-order modulation formats. Therefore, NLC is a hot research topic to increase fiber capacity. Several techniques have been reported in the literature to combat the effects of nonlinearity. NLC techniques can be implemented either in optical or digital domains. The possible locations of the commonly used NLC techniques in the optical transmission link are provided in Fig. 2.6. Such techniques are applied either at the transmitter side or in the optical link, or at the receiver side. Due to the introduction of coherent detection, DSP algorithms can be used to combat fiber impairments and, in particular, nonlinear distortions. Usually, DSP algorithms are implemented either at the transmitter side or at the receiver side or a

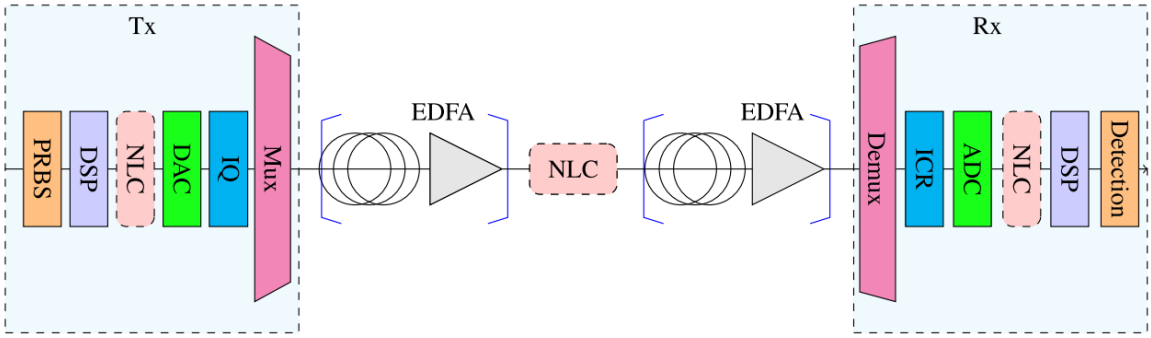


Figure 2.6: Transmission diagram: possible NLC locations. PRBS: pseudo-random binary sequences, DSP: digital signal processing, NLC: nonlinearity compensation, DAC: digital-to-analog converter, IQ: in-phase and quadrature modulator, Mux: multiplexer, Demux: de-multiplexer, ICR: intergrated coherent receiver, ADC: analog-to-digital converter, taken from [5].

combination of both like split nonlinearity compensation techniques [5]. Implementing DSP algorithms in the optical fiber link requires optical-to-electrical/electrical-to-optical conversions at the locations where the DSP modules are deployed. That increases the signal latency in the optical fiber link and the overall cost of the link.

Digital NLC techniques represent a key technology and a cost-effective approach to increase the data rate, being adopted for the next generation WDM optical transmission systems [5]. In this section, we provide a concise review of the most popular NLC techniques in the literature.

2.7.1 Digital Back-propagation

Digital back-propagation (DBP) is a digital technique proposed to compensate for the fiber nonlinearity. This technique can be implemented either at the transmitter side, as a pre-distortion, or at the receiver side as a post-compensation [5]. The DBP is based on the SSFM, which represents an effective numerical technique to solve the signal propagation equation. The idea of the DBP technique is to digitally model a fictitious fiber with exactly opposite characteristics when compared to the real fiber

used for the transmission [5]. The fiber link is divided into several steps with a small length, and at each step, it is modeled as a concatenation of linear and nonlinear sections. There are different ways of DBP implementations proposed in the literature, depending on the implementation order of the linear and nonlinear sections [5]. The most commonly used approach is the application of linear compensation first because nonlinear effects are more important at high input powers, which is the case at the end of the fictitious fiber [5]. The implementation of the DBP technique at the receiver side is shown in Fig. 2.7, where N_s is the number of steps.

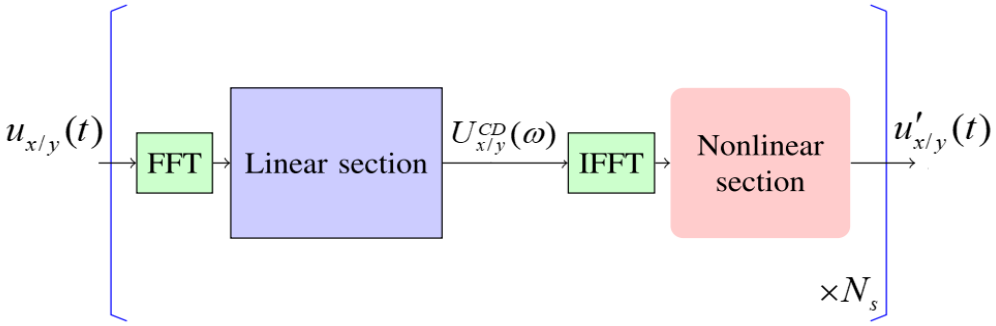


Figure 2.7: DBP implementation principle. (I)FFT: (inverse) fast Fourier transform, taken from [5].

Using SSFM, the output of the linear section which compensates for the CD, is given by:

$$U^{\mathrm{CD}}_{x/y}(z,w) = U_{x/y}(z,w)\exp(-jh(\frac{\alpha}{2}+\frac{\beta_2}{2}\omega^2)), \tag{2.108}$$

where h is the length of each step, w is the frequency variable and z is the transmission distance. Mainly, this operation corresponds to the multiplication of the received signal by an exponential term. This term represents the inverse of the signal phase change due to CD. After that, the nonlinear section is performed in time domain using the expression:

$$u'_{x/y}(z,t) = u^{\mathrm{CD}}_{x/y}(z,t)\exp(-j\varphi\gamma' h(|u^{\mathrm{CD}}_x|^2+|u^{\mathrm{CD}}_y|^2)), \tag{2.109}$$

where $0 < \varphi < 1$ is a real-valued optimization parameter. The exponential term in (2.109) introduces the phase change because of the Kerr effect.

Single-channel DBP (SC-DBP) has been more widely researched because it is typically considered to be more realistic with the current hardware limitations [5]. Since a single-wavelength channel is back-propagated, SC-DBP only compensates for intra-channel nonlinearity, e.g., SPM. In WDM superchannel systems, one way of overcoming the inter-channel nonlinear distortions like XPM, XPolM, and FWM, induced by the co-propagating subchannels, is to use a multi channel-DBP (MC-DBP). The MC-DBP back-propagates the entire WDM channel [5]. However, the implementation of the MC-DBP is limited to point-to-point links, and its computational complexity is considered impractical due to the need for massively parallel processing computer systems to implement [5]. Some new approaches have been proposed to reduce the complexity of DBP based on SSFM, such as weighted DBP [5] and correlated DBP [5]; however, they are still complex for real-time implementation. DBP compensates for all deterministic impairments and is considered as the benchmark to evaluate other NLC techniques.

2.7.2 Volterra Series-based Nonlinear Equalizer

The VSTF can be effectively used to model the fiber nonlinearity effects [5]. The VSTF is a powerful tool for solving the Manakov equation (NLSE in case of single-polarization transmission), as shown in [5]. After modeling the optical channel based on VSTF, the p-th order theory developed in [40] is used to derive the inverse VSTF (IVSTF) kernels as a function of the VSTF ones. IVSTF kernels characterize the nonlinear equalizer which compensates for the fiber nonlinearity and CD. Like DBP, Volterra series based nonlinear equalizer (VNLE) attempts to construct the inverse of the channel. One of the features of the VNLE is that the compensation operation

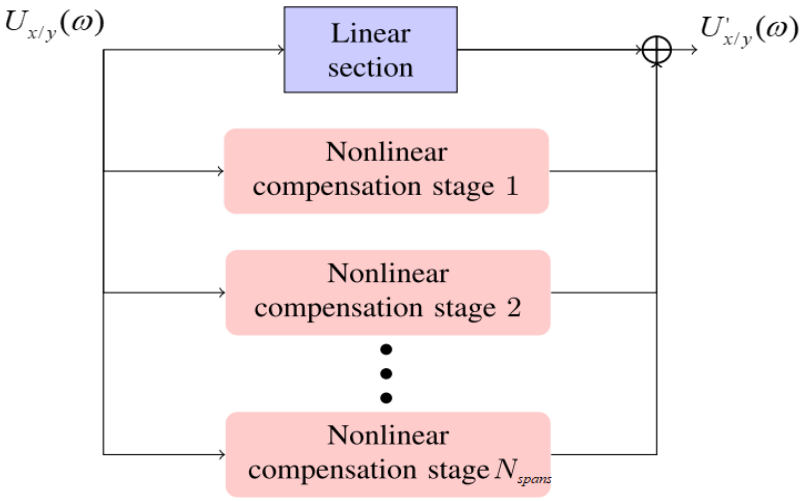


Figure 2.8: Parallel implementation of VNLE, taken from [5].

can be performed in parallel [5]. That reduces the computational complexity when compared to DBP [5]. The principle of the parallel implementation of VNLE is depicted in Fig. 2.8, where N_{spans} corresponds to the number of fiber spans.

For each polarization, the compensation operation can be divided into two parts; one is linear, and the other is nonlinear. The linear part consists of CD compensation, and the nonlinear part compensates for the nonlinear phase change, which is proportional to the instantaneous signal power. In this technique, the nonlinearity compensation for each span can be carried out in parallel. Finally, the output of the VNLE is obtained by combining the linear and nonlinear compensation sections. To improve the compensation performance and reduce the implementation complexity, several approaches have been proposed, such as modified VNLE (MVNLE) and weighted Volterra series nonlinear equalizer (W-VSNE) [5]. It is worth mentioning that the VNLE, MVNLE, and WVSNE are based on the third-order Volterra series. Recently, a fifth-order VNLE has also been proposed; however, it involves much higher implementation complexity when compared with the third-order case [5].

2.7.3 First-order Perturbation Theory-based NLC

The FO perturbation theory-based NLC (FO-PB-NLC) technique relies on some simplifying assumptions in deriving the approximate FO nonlinear distortion field using (2.107), including [46]:

- The accumulated CD is fully compensated electronically at the receiver.
- The input pulses are Gaussian shaped.

Based on FO perturbation theory, three input Gaussian pulses $\sqrt{P_0}a_{m/l/n,x/y}\exp(-(t-T_{m/n/l})^2/2\tau^2)$, at three time instants T_m, T_l, T_n generate a ghost pulse due to the nonlinear interaction as shown in (2.107). Fig. 2.9 shows a schematic representation of the triplet pulses involved in the calculation of the FO nonlinear distortion field using (2.107).

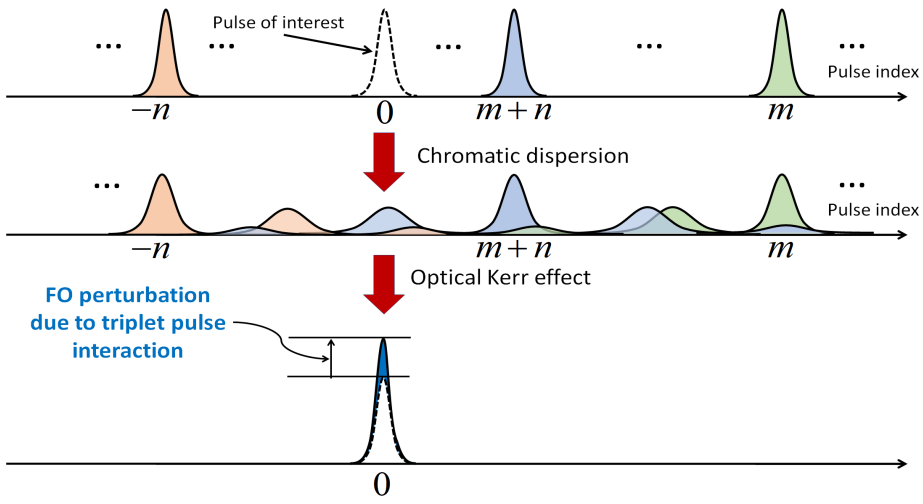


Figure 2.9: Triplet pulses involving in the FO distortion field calculation.

Without loss of generality, in the predistortion technique, the perturbation of the symbol at index $k = 0$, i.e., $l = m + n$ is calculated [46]. It is worth mentioning that the nonlinear distortion calculation using (2.107) at any other index, for example $k = m + n - l$, is the same as the calculation at $k = 0$ [46]. The predistortion is

assumed to operate at the symbol rate; therefore, the perturbation value at $t = 0$ is calculated [46]. Accordingly, (2.107) can be further simplified as:

$$u_{1,x/y}(L,t) = j\frac{8}{9}\gamma P_0^{3/2}\sum_m\sum_n \Big[a_{m,x/y}a^*_{m+n,x/y}a_{n,x/y} + a_{m,y/x}a^*_{m+n,y/x}a_{n,x/y}\Big]\mathbf{C}^{FO}_{m,n}, \quad (2.110)$$

where $*$ represents the complex conjugate operation and $\mathbf{C}^{FO}_{m,n}$ is the FO perturbation coefficient matrix, which is given as:

$$\mathbf{C}^{FO}_{m,n} = \int_0^L \frac{\exp(-\alpha z)}{\sqrt{1+2j\beta_2 z/\tau^2+3(\beta_2 z/\tau^2)^2}}\exp\left(-3\frac{mnT^2}{\tau^2(1+3j\beta_2 z/\tau^2)} - \frac{(m-n)^2T^2}{\tau^2[1+2j\beta_2 z/\tau^2+3(\beta_2 z/\tau^2)^2]}\right)dz. \quad (2.111)$$

In a typical dispersion unmanaged optical transmission system, the CD-induced pulse spreading is usually much larger than the symbol duration, i.e., $\beta_2 z \gg \tau^2$ [46]. With the large CD assumption and following a similar procedure as in [46], the FO perturbation coefficient matrix $\mathbf{C}^{FO}_{m,n}$ can be further simplified as:

$$\mathbf{C}^{FO}_{m,n} = \begin{cases} \frac{\tau^2}{\sqrt{3}|\beta_2|}\int_0^L dz\frac{1}{\sqrt{\tau^4/(3\beta_2^2)+z^2}}, & m=n=0 \\ \frac{\tau^2}{\sqrt{3}|\beta_2|}\frac{1}{2}E_1(\frac{(n-m)^2T^2\tau^2}{3|\beta_2|^2L^2}), & m \text{ or } n = 0 \\ \frac{\tau^2}{\sqrt{3}|\beta_2|}E_1(-j\frac{mnT^2}{\beta_2 L}), & m\neq n\neq 0, \end{cases} \quad (2.112)$$

where $E_1(x) = \int_x^\infty \frac{e^{-t}}{t}dt$ is the exponential integral function [46].

In the FO-PB-NLC technique, the perturbation coefficient matrix $\mathbf{C}^{FO}_{m,n}$ is calculated offline and stored in a look-up table (LUT). The basic idea of the predistortion technique is to calculate the FO nonlinear distortion field $u_{1,x/y}$ firstly using (2.110)

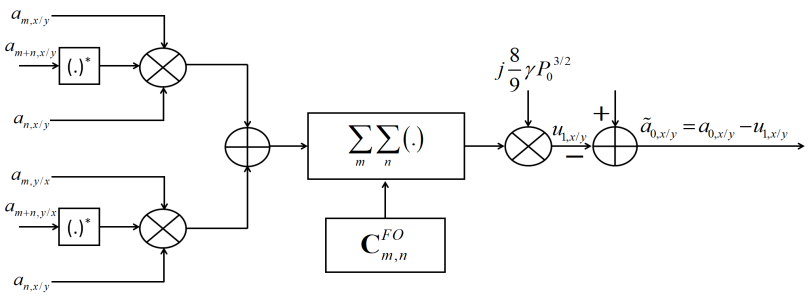


Figure 2.10: The block diagram of the FO-PB-NLC technique.

and then to subtract it from the symbol under consideration (i.e., the symbol at the zeroth index) $a_{0,x/y}$ to generate the predistorted symbol $\tilde{a}_{0,x/y}$, as shown in Fig. 2.10.

2.7.4 Optical/Digital Phase Conjugation

The phase conjugation technique for fiber nonlinearity mitigation can be applied in the optical or electrical domain [48], [49]. In the optical phase conjugation (OPC) technique, the spectrum of the signal is optically inverted in the middle of the optical link. The main idea of this technique is to cancel the nonlinear phase shift generated in the first segment of the fiber link using the nonlinearity in the second segment of the link [48]. However, the OPC technique requires the physical modification of the transmission link and the precise placement of the spectral inversion device. That significantly affects the flexibility of the optical network and make its implementation difficult.

Recently, a digital phase conjugation (DPC) based technique, referred to as the phase-conjugated twin wave (PCTW), was proposed for the mitigation of the FO nonlinear distortions in PDM optical transmission systems. The PCTW technique can be used for the effective mitigation of the nonlinear distortions in PDM coherent optical systems, at the expense of halving the spectral efficiency [49]. The basic principle of the PCTW-based nonlinear distortion cancellation in a PDM optical transmission system is illustrated in Fig. 2.11. The symbols E_x, E_y represent the transmitted elec-

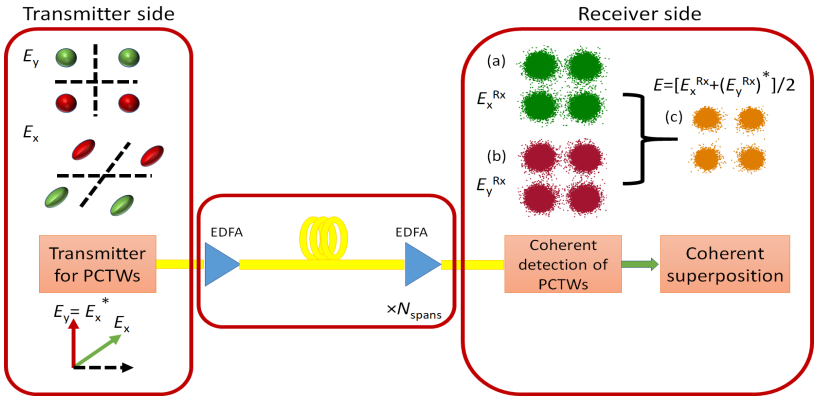


Figure 2.11: Illustration showing the cancellation of nonlinear distortions via PCTWs. EDFA: erbium doped fiber amplifier.

tric fields and $E_x^{R_x}$, $E_y^{R_y}$represent the received electric fields. Fig. 2.11 also shows the received signal constellations on the x and y polarizations after transmission through a 2800 km SSMF. The coherent superposition of these received symbols on the two orthogonal polarizations leads to the cancellation of the nonlinear distortions, and the resultant constellation is shown in Fig. 2.11. Evidently, the constellation quality is much improved through the coherent superposition of the two PCTWs. It is reported in the literature that the performance gain obtained through the PCTW technique in the linear regime is $\sim$ 3 dB [49]. That is because the variance of the linear noise resulting from the ASE noise is halved, by coherently mixing two PCTWs. On the other hand, the performance gain in the nonlinear regime can be well beyond 5 dB owing to the first-order cancellation of the nonlinear distortions through the coherent superposition of two PCTWs [50].

Recently, a time-domain implementation of the generalized PCTW was reported in [51], and this method is referred to as conjugate data repetition (CDR). In the CDR technique, each time-domain signal datum is followed by its conjugate pairs, and the received signals in the adjacent time slots are coherently superimposed at the receiver. Since the nonlinearity interference coefficients change slowly in a highly

dispersive channel, the nonlinear distortions generated by conjugate repetition data can be self-canceled by superimposing [51]. In [51], a theoretical explanation based on time-domain perturbation analysis has been provided for the nonlinear distortion cancellation with the CDR technique. It is also shown that the PCTW and the CDR techniques have similar performance, with a limitation of halving the overall capacity of the coherent optical communication system. A variant of the PCTW technique for coherent optical orthogonal frequency division multiplexing (CO-OFDM) system exploring the Hermitian symmetry has been proposed in [52]. That also comes at the cost of 50% spectral efficiency [52].

2.7.4.1 DPC Techniques with Improved Spectral Efficiency

In [53], a spectrally efficient DPC technique for nonlinearity compensation in the CO-OFDM system has been proposed. This technique is based on the transmission of the phase-conjugated pilots (PCPs) and the coherent superposition at the receiver. In this scheme, a portion of the OFDM subcarriers is transmitted as PCPs of the other subcarriers. These phase conjugate pairs are used at the receiver to estimate and compensate for the nonlinear distortions introduced by the channel. This scheme allows the spectral redundancy to be adjusted (up to 50%) according to the required performance gain. That can be achieved through the proper selection of the number of PCPs in each OFDM band. In [54], a novel technique is proposed based on the joint processing of two pairs of PCTWs, referred to as dual-PCTW, to avoid the loss of spectral efficiency associated with the use of PCTWs. However, this technique does not provide any significant performance improvement without 50% pre-EDC in the transmission link.

Recently, in [55], a frequency domain coding technique, termed as phase-conjugated subcarrier coding (PCSC), combined with electronic dispersion pre-compensation

(pre-EDC) has been demonstrated for nonlinearity mitigation in CO-OFDM system. This technique extends the idea of the dual-PCTW concept to process the neighboring OFDM subcarriers jointly. The benefit of this nonlinearity mitigation technique comes from the fact that the nonlinear distortions on neighboring OFDM subcarriers are highly correlated [55]. The PCSC scheme can be effectively applied for nonlinearity compensation without spectral efficiency loss. However, the PCSC technique does not provide any significant performance improvement without 50% pre-EDC in the transmission link [55]. That is essentially a limiting factor for the implementation of the PCSC scheme in a dynamically routed optical network because it would be very expensive to achieve the optimal pre-EDC in such network links.

In summary, an overview of the type and location of the popular NLC techniques, along with the type of the fiber nonlinearity which they compensate for, is listed in Table 2.2.

Table 2.2: Fiber nonlinearity compensation techniques.

Technique	**Type**	**Location**	**Fiber nonlinearity compensated**	**Transmission system**	**References**
SC-DBP	Digital	Tx/Rx	Intra-subcarrier	Nyquist/OFDM	[59]- [64]
MC-DBP	Digital	Tx/Rx	Intra-/inter subcarrier	Nyquist/OFDM	[65]- [69]
VNLE	Digital	Tx/Rx	Intra-subcarrier	Nyquist/OFDM	[38]- [42]
PB-NLC	Digital	Tx/Rx	Intra-subcarrier/XPM	Nyquist/OFDM	[43]- [47]
O/DPC	Optical/Digital	Link/Rx	Nonlinear phase	Nyquist/OFDM	[48]- [58]

In this thesis, we study the impact of the fiber Kerr nonlinearity-induced performance degradation in long-haul coherent optical communication systems. We consider both CO-OFDM and single-carrier coherent optical transmission systems. Based on the background study detailed in this chapter, we develop four different digital NLC techniques to deal with the detrimental effects of fiber nonlinearity. The first two tech-

niques are developed for the CO-OFDM systems; whereas, the other two techniques are designed for the single-carrier systems. More specifically, Chapter 3 and Chapter 4 provides the DSP techniques that are developed for the CO-OFDM systems; whereas, Chapter 5 and Chapter 6 considers the single-carrier systems.

In Chapter 3, two DPC techniques, referred to as modified-16-quadrature amplitude modulated CDR (MOD-16-QAM-CDR) and MOD-16-QAM-PCTW, are developed to solve the spectral efficiency problem associated with the PCTW technique. Also, in Chapter 3, we investigate the impact of the PDL on the performance of the MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques. In Chapter 4, we discuss the development of a joint NLC technique for the WDM CO-OFDM system by combining the SC-DBP with the PCTW technique, referred to as the SC-DBP-PCTW technique. The SC-DBP-PCTW technique effectively deals with the intra- and inter-channel nonlinearity effects in a WDM system and increase the transmission reach when compared to the individual implementation of the SC-DBP and PCTW techniques. We carry out the complexity analysis and show that the SC-DBP-PCTW technique has a reduced implementation complexity when compared to the MC-DBP. In Chapter 5, we solve the energy divergence problem of the regular perturbation (RP) theory used in the PB-NLC technique by using the enhanced RP (ERP) technique. The resulting NLC technique is referred to as the ERP-NLC technique. The ERP-NLC technique improves transmission performance with a negligible increase in the computational complexity when compared to the PB-NLC technique. In Chapter 6, we extend the FO perturbation theory to the second-order (SO) and develop an NLC technique based on the SO perturbation theory, referred to as the SO-PB-NLC. The complexity analysis shows that the performance enhancement of the SO-PB-NLC technique comes with a reduced implementation complexity when compared to the DBP technique with one step per span.

Chapter 3

Digital Phase Conjugation Schemes for Fiber Nonlinearity Compensation in CO-OFDM Systems

3.1 Preamble

This chapter is compiled from the materials extracted from the manuscripts titled *"A spectrally-efficient linear polarization coding scheme for fiber nonlinearity compensation in CO-OFDM systems"* which appeared in the proceedings of the *SPIE Opto, Jan. 2017* [58] and *"PDL impact on linearly coded digital phase conjugation techniques in CO-OFDM systems"* published in *IEEE Photonics Technology Letters, May 2018* [77].

3.2 Introduction

The next generation of the wavelength division multiplexed (WDM) coherent optical orthogonal frequency division multiplexing (CO-OFDM) systems are required to operate the optical communication links at 400 Gbps/1 Tbps transmission rates. One possible solution to achieve such a transmission rate is the use of subcarrier multiplexing, known as superchannel, along with the higher-order modulation formats, which provide high spectral efficiency (SE) and low cost. The use of higher-order modulation formats and the reduced guard-band between the sub-channels make the superchannels vulnerable to the fiber Kerr nonlinearity. Over the last decade, several digital nonlinear compensation (NLC) techniques have been investigated to compensate for the fiber nonlinearity in CO-OFDM superchannel systems [59].

The phase-conjugated twin wave (PCTW) is an effective technique proposed for the mitigation of the nonlinear distortions in a polarization multiplexed optical transmission system [49]. The PCTW technique transmits the mutually phase-conjugated twin waves on the two orthogonal polarizations and coherently superimpose them at the receiver. Alternatively, the conjugate data repetition (CDR) technique proposed in [51] transmits the phase-conjugated twin waves on the adjacent time slots of the same polarization and coherently superimpose them at the receiver. However, the NLC performance of both PCTW and CDR techniques comes at the expense of halving the overall SE of the link. In [70], the 16-quadrature amplitude modulated (QAM)-PCTW technique has been proposed and compared with the polarization multiplexed 4-QAM transmission system. The results indicate that the combined use of 16-QAM and PCTW performed worse in both weakly and highly nonlinear regimes than the polarization multiplexed 4-QAM due to the much lower receiver sensitivity of 16-QAM. The phase-conjugated subcarrier coding (PCSC) technique proposed in [55] can be effectively applied for NLC without spectral efficiency loss. However,

the PCSC technique does not provide any significant performance improvement without a 50% electronic dispersion pre-compensation (pre-EDC) in the transmission link. That is essentially a limiting factor for the implementation of the PCSC scheme in a dynamically routed optical network because it would be very expensive to achieve the optimal pre-EDC in such network links [56], [57].

To address the issue of the spectral redundancy associated with the PCTW and CDR techniques and the requirement of pre-EDC for the PCSC technique, we propose a scheme that linearly combines the data symbols on the two adjacent subcarriers of the OFDM symbol, one at full amplitude and another at half amplitude. The phase-conjugated pairs of such linearly coded signals are then transmitted on the same subcarriers of the two OFDM symbols on the two orthogonal time/polarization states. At the receiver, the coherent superposition of the recovered phase-conjugated signal pairs is carried out, to cancel the first-order nonlinear distortions.

The polarization effect, such as polarization-dependent loss (PDL), degrades the transmission performance of the optical systems operating at high bit-rates [20], [21]. The PDL causes a signal power/optical signal-to-noise ratio (OSNR) imbalance between the two polarizations of a PDM signal. Only a few studies consider the effect of PDL on the performance of digital NLC techniques. In [71] and [72], an investigation of the impact of polarization effects on the performance of digital back-propagation (DBP) and perturbation-based NLC is carried out. However, no investigation of the impact of polarization effects on the DPC techniques is considered in the literature. In this chapter, we also investigate the impact of PDL on the performance of the proposed DPC techniques.

The main contributions of this chapter are as follows:

- We develop two DPC techniques to solve the spectral redundancy problem of the PCTW and CDR techniques.

- We show that the proposed DPC schemes provide favorable NLC performance without the pre-EDC technique.

- We investigate the PDL impact on the performance of the proposed DPC techniques and show that the scheme which transmits the phase-conjugated pairs on the orthogonal time slots is more tolerant irrespective of the PDL model.

3.3 Proposed DPC Approaches

This section describes the proposed DPC approaches. The objective of the study is to address the issue of halving the SE associated with the PCTW and CDR techniques and the requirement of pre-EDC for the PCSC technique. On this road, we propose two DPC approaches: one uses two orthogonal time slots of the same polarization to transmit the linearly-coded signal and its phase conjugate, while the other uses orthogonal polarizations. The former approach is referred to as a modified-16-QAM-CDR (MOD-16-QAM-CDR) technique, while the latter is termed a MOD-16-QAM-PCTW technique.

3.3.1 The MOD-16-QAM-CDR Technique

In the proposed MOD-16-QAM-CDR technique, the data symbols on the adjacent subcarriers of the OFDM symbol are linearly combined, one at full amplitude and the other at half amplitude, as shown in Fig. 3.1. This technique is an alternate way of generating sixteen constellation points from two quadrature phase-shift keying (QPSK) symbols. The constellation symbols on the rightmost scatter plot in Fig. 3.1 shows the mapping of the lower amplitude QPSK constellation symbols corresponding to a constellation symbol on the higher amplitude QPSK constellation. It is important to note that the generation of sixteen constellation symbols in this way provides an

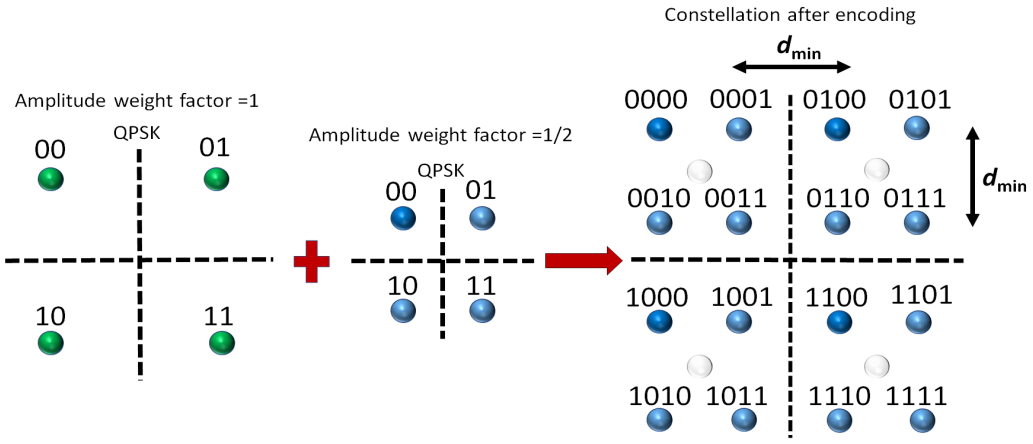


Figure 3.1: Generation of sixteen constellation symbols by combining two QPSK symbols.

equal distance '$d_{\min}$' between any pair of adjacent constellation points in the signal space diagram. That reduces the average probability of symbol error after detection.

On the other hand, if the two component QPSK symbols are combined with any other arbitrary weighted amplitude values, then the distance between the pair of constellation points in the resultant signal space diagram is not uniform (not equal) and leads to the increased average probability of symbol error. Fig. 3.2 shows the received constellations for different arbitrary amplitude values for the second QPSK constellation symbols. It is clear from Fig. 3.2 that the distance between the pair of constellation points in Fig. 3.2(a), (c) and (d) is not uniform (not equal) when compared to the case of half the amplitude for the second QPSK symbol (i.e., in Fig. 3.2(b)).

At the encoder, each pair of the neighboring OFDM subcarriers (with the indices of $2k-1$ and $2k$, where k is an integer number) is encoded, as shown in Fig. 3.3, as:

$$\begin{aligned} \boldsymbol{S}_{x/y,t}(k) &= \boldsymbol{A}_{x/y}(2k-1) + \boldsymbol{A}_{x/y}(2k)/2 \\ \boldsymbol{S}_{x/y,t+T}(k) &= \boldsymbol{A}^{*}_{x/y}(2k-1) + \boldsymbol{A}^{*}_{x/y}(2k)/2, \end{aligned} \tag{3.1}$$

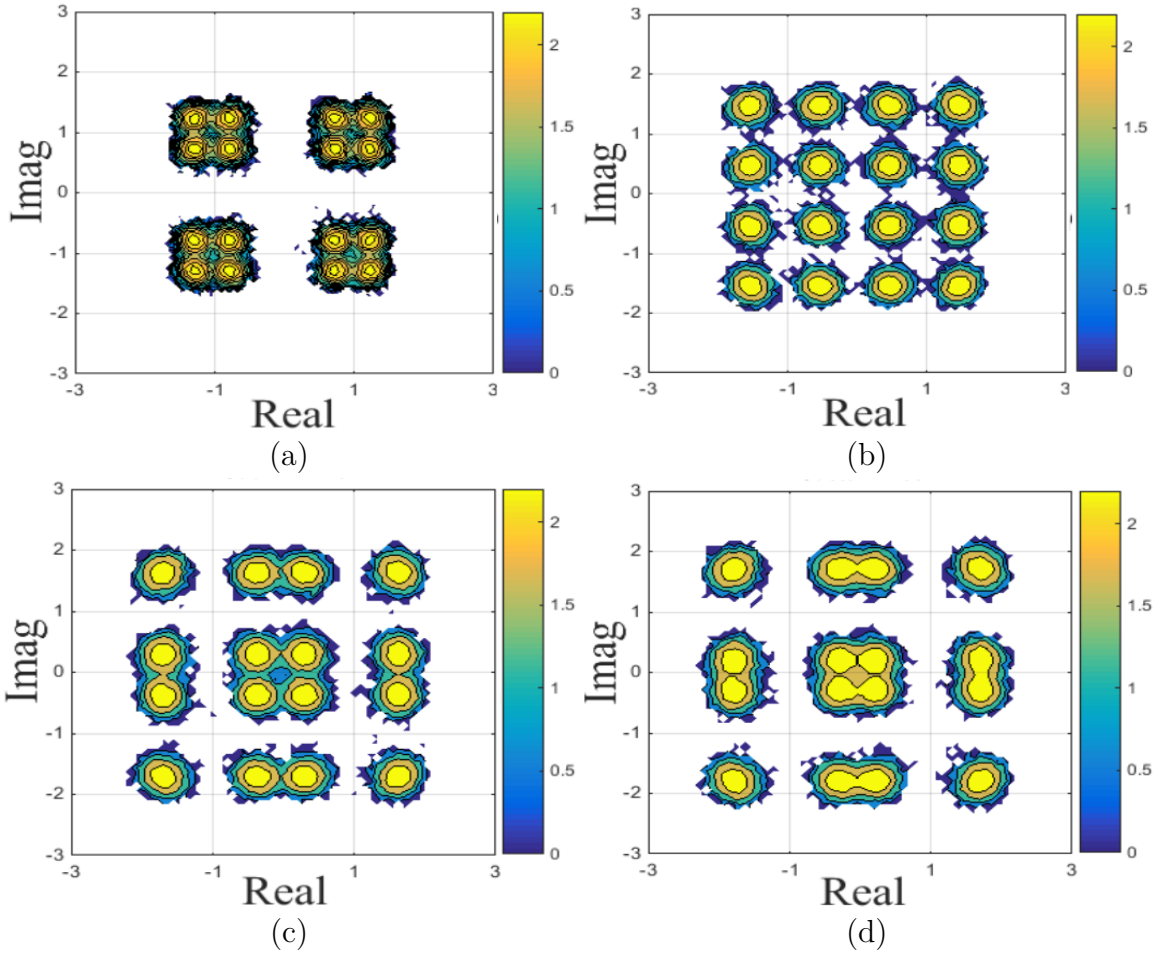


Figure 3.2: The received constellations for different amplitude values for the second QPSK symbol: (a) 1/4, (b) 1/2, (c) 2/3, and (d) 3/4.

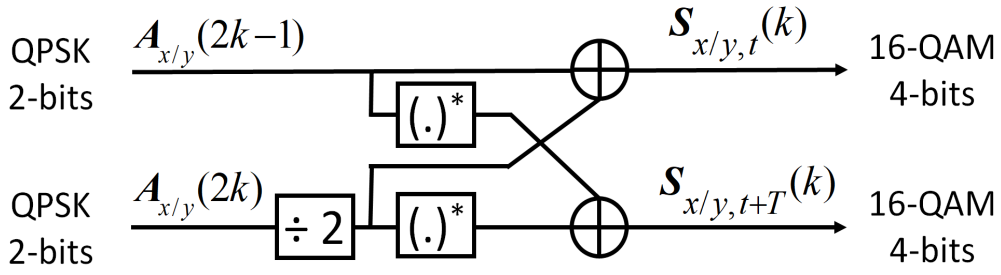


Figure 3.3: MOD-16-QAM-CDR encoder.

where $k = 1, 2, ..., N/2$, is the subcarrier number, N is the number of subcarriers, $t = (2n+1)T$, where $n = 0, 1, 2...$, is the time variable, T is the OFDM symbol duration, $\boldsymbol{A}_{x/y}$ is the OFDM symbol before the encoding process, $\boldsymbol{S}_{x/y,t}$ and $\boldsymbol{S}_{x/y,t+T}$ are

the OFDM symbols after the encoding process on the two polarizations, and the subscripts x, y represents the two orthogonal polarization states of the fiber. Please note that the symbols of $\boldsymbol{A}_{x/y}$ are drawn from the alphabet $\{1+j,\ 1-j,\ -1+j,\ -1-j\}$.

The MOD-16-QAM-CDR technique can be considered as a one-to-one mapping scheme in which the encoder output is essentially a code word from a finite code alphabet. After encoding, the linearly coded symbols, $\boldsymbol{S}_{x/y,t}$ and $\boldsymbol{S}_{x/y,t+T}$ are modulated onto the electric fields and transmitted through the fiber. The corresponding transmitted vector field can be represented as $[\boldsymbol{E}_{x/y,t}(0,w)\ \boldsymbol{E}_{x/y,t+T}(0,w)]$, where ω is the frequency.

At the receiver, the acquired vector field corresponding to two orthogonal polarization tributaries can be represented as $[\boldsymbol{E}_{x/y,t}(L,w)\ \boldsymbol{E}_{x/y,t+T}(L,w)]$, where L is the transmission distance. After analog-to-digital conversion, the information symbols on the subcarriers corresponding to the two adjacent time slots are coherently superimposed, as shown in Fig. 3.4 as:

$$\boldsymbol{R}_{x/y}(k) = (\boldsymbol{B}_{x/y,t}(k) + \boldsymbol{B}^{*}_{x/y,t+T}(k))/2, \tag{3.2}$$

where $\boldsymbol{B}_{x/y}$ and $\boldsymbol{R}_{x/y}$ are the OFDM symbols before and after the decoding process, respectively.

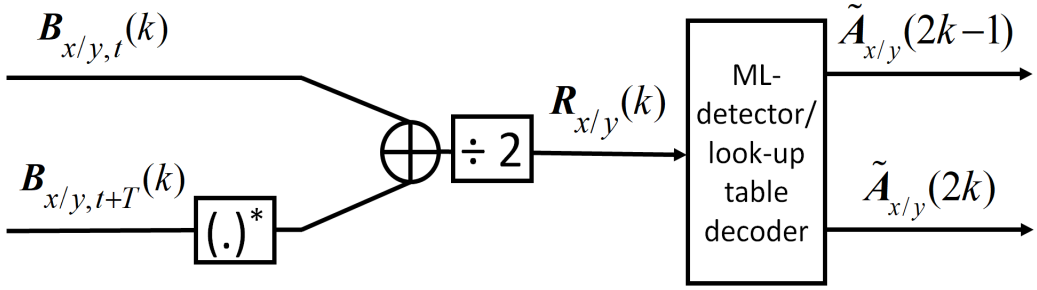


Figure 3.4: MOD-16-QAM-CDR decoder.

After the coherent superposition (CS), the recovered symbol $\boldsymbol{R}_{x/y}$ is passed onto the maximum-likelihood (ML)/ look-up table (LUT) decoder block, as shown in Fig. 3.4. The ML-detector calculates the distance metric on a symbol-by-symbol basis and selects the code word corresponding to the minimum distance value from the set of all possible linear time-codes. Then, a LUT at the receiver is used for decoding the actual data symbols $\tilde{\boldsymbol{A}}_{x/y}(2k-1)$ and $\tilde{\boldsymbol{A}}_{x/y}(2k)$, as shown in Fig. 3.4.

In the proposed MOD-16-QAM-CDR technique, the linearly coded signals are transmitted as phase conjugate pairs on the two adjacent time slots of the same polarization. Thus, the nonlinear distortion field added onto two signal pairs is highly cross-correlated. That leads to the first-order cancellation of the nonlinear distortion fields upon CS at the receiver. One disadvantage is that the modified signal constellation after the linear coding consists of sixteen points with equal probabilities for all the points, as shown in Fig. 3.1. This feature drops the performance of the proposed MOD-16-QAM-CDR technique in the linear (or weakly nonlinear) transmission regime when compared to the PDM 4-quadrature amplitude modulation (QAM) with linear compensation. In the linear (or weakly nonlinear) transmission regime, the performance is limited by the OSNR penalty. However, the numerical simulation results show that the proposed MOD-16-QAM-CDR technique outperforms the PDM-4-QAM in the highly nonlinear transmission regime, where the penalties due to the nonlinearity dominate over the OSNR penalty. This performance gain comes from the nature of the nonlinear distortion cancellation through the CS of the proposed MOD-16-QAM-CDR technique.

3.3.2 The MOD-16-QAM-PCTW Technique

In the MOD-16-QAM-PCTW technique, each pair of the neighboring OFDM subcarriers is encoded, as shown in Fig. 3.5, as:

$$\begin{aligned} \boldsymbol{S}_{x,t}(k) &= \boldsymbol{A}(2k-1) + \boldsymbol{A}(2k)/2 \\ \boldsymbol{S}_{y,t}(k) &= \boldsymbol{A}^*(2k-1) + \boldsymbol{A}^*(2k)/2, \end{aligned} \tag{3.3}$$

where $k = 1, 2, ..., N/2$, is the subcarrier number, $\boldsymbol{A}$ and $\boldsymbol{S}_{x/y}$ are the OFDM symbols before and after the encoding process, and the subscripts x and y represent the two orthogonal polarization states of the fiber. After encoding, the linearly coded symbols, $\boldsymbol{S}_x$ and $\boldsymbol{S}_y$ are modulated onto the electric fields corresponding to the two polarizations of the optical signal and transmitted through the fiber. The corresponding transmitted vector field can be represented as $[E_x(0,w)\ E_y(0,w)]^\dagger$, where w is the frequency.

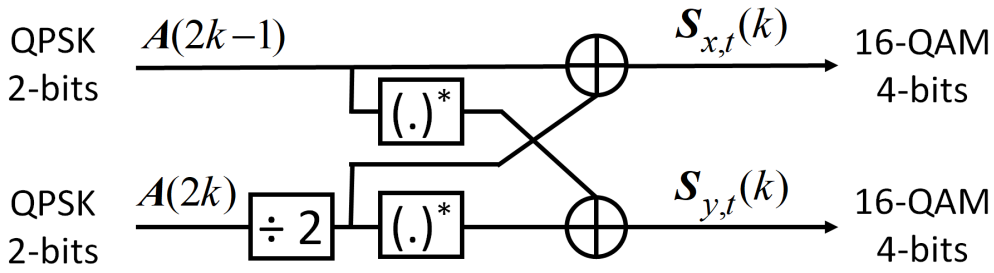


Figure 3.5: MOD-16-QAM-PCTW encoder.

At the receiver, after analog-to-digital conversion, the information symbols on the subcarriers corresponding to the two orthogonal polarizations are coherently superimposed, as shown in Fig. 3.6, as:

$$\boldsymbol{R}_t(k) = (\boldsymbol{B}_{x,t}(k) + \boldsymbol{B}^*_{y,t}(k))/2, \tag{3.4}$$

where $\boldsymbol{B}_{x/y}$ and $\boldsymbol{R}$ are the OFDM symbols before and after the decoding process, respectively. After the CS, as in the MOD-16-QAM-CDR technique, the recovered symbol $\boldsymbol{R}$ is passed onto the ML-detector/LUT decoder block and the actual data symbols $\tilde{\boldsymbol{A}}(2k-1)$ and $\tilde{\boldsymbol{A}}(2k)$ are decoded, as illustrated in Fig. 3.6.

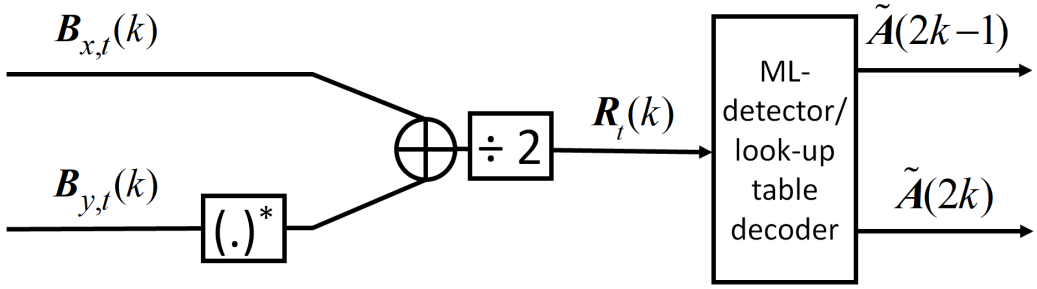


Figure 3.6: MOD-16-QAM-PCTW decoder.

3.4 First-order Perturbation Theory for MOD-16-QAM-CDR/MOD-16-QAM-PCTW Techniques

In this section, we derive the nonlinear distortion field added to the OFDM subcarrier symbols using the first-order perturbation theory. Then, we show that the CS of the MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques can cancel the first-order perturbative nonlinear distortion field. The propagation of an OFDM signal through the optical communication system can be represented as in Fig. 3.7. At the transmitter, the electrical baseband OFDM signal $x(t)$ is up-converted to the optical domain by using a Mach-Zehnder modulator (MZM), and the corresponding signal can be represented as $u(t)$. Since the optical field inside the fiber varies as a function of time and space, let $\hat{u}(z,t)$ be the Spatio-temporal complex envelop of the optical field at time t and distance z along the fiber [44]. Therefore, we can represent $u(t) = \hat{u}(0,t)$ as the input optical field to the fiber at $z = 0$. According to the regular

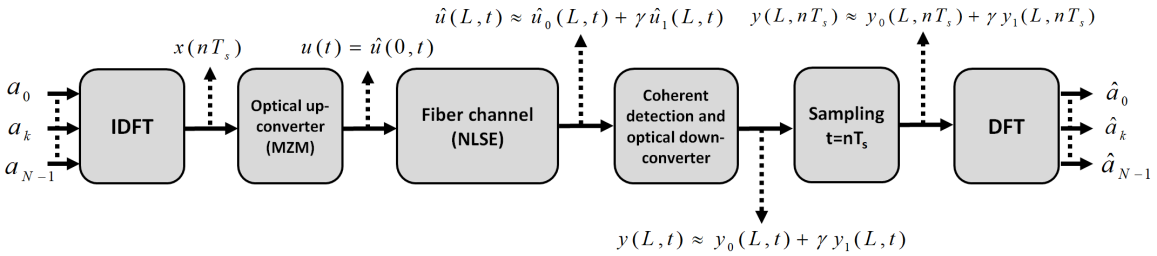


Figure 3.7: The block diagram showing the propagation of an OFDM signal through the optical communication system. (I)DFT: (inverse) discrete Fourier transform, N: number of subcarriers, MZM: Mach-Zehnder modulator, NLSE: nonlinear Schrödinger equation, $a_0, ..., a_{N-1}$ represents the transmitted OFDM subcarriers, $\hat{u}_0(L,t)$, $\hat{u}_1(L,t)$ represents the zeroth-order and the first-order solution of the optical field and $y_0(L,t)$, $y_1(L,t)$ are the corresponding baseband representations, $y_0(L,nT_s)$ and $y_1(L,nT_s)$ represents the sampled version of the baseband signal, where $n = 0, 1, ..., N-1$ and $\hat{a}_0, ..., \hat{a}_{N-1}$ represents the received OFDM subcarriers.

perturbation (RP) theory,* the signal received at the output of the fiber at a distance $z = L$ can be approximated to the first-order as $\hat{u}(L,t) \approx \hat{u}_0(L,t) + \gamma\hat{u}_1(L,t)$, where $\hat{u}_0(L,t)$ is the linear solution and $\hat{u}_1(L,t)$ is the first-order solution, and where γ is the nonlinearity coefficient [44]. The received optical signal is coherently detected and down-converted to the baseband, and the corresponding signal can be represented as $y(L,t) \approx y_0(L,t) + \gamma y_1(L,t)$. The signal is then sampled at $t = nT_s$, where T_s is the sampling interval and the discrete Fourier transform (DFT) is performed to recover the individual subcarriers. The baseband OFDM signal in time-domain can be represented as:

$$x(nT_s) = \frac{1}{N}\sum_{k=0}^{N-1} a_k \exp(j2\pi k(nT_s)/NT_s), \tag{3.5}$$

where N is the number of subcarriers, a_k is the complex data symbol, k is the subcarrier index and T_s is the sampling period. The signal $x(t)$ is then up-converted to

*In this analysis, we consider the RP solution only up to the first-order and consider only a single-channel and single-polarization for the simplicity of the analysis. The impact of the linear phase noise, frequency/timing offset, and ASE noise from the amplifier are not considered in this study. Further, we assume that the optical up-conversion using MZM and the down-conversion to baseband are ideal.

the optical domain and the corresponding optical field can be represented as:

$$u(t) = \exp(j2\pi f_c t).\frac{1}{N}\sum_{k=0}^{N-1} a_k \exp(j2\pi kn/N), \tag{3.6}$$

where f_c is the optical carrier frequency.

The propagation of a spatio-temporal complex envelope $\hat{u}(z,t)$ in a SSMF is governed by the NLSE as [32], [44]:

$$\frac{\partial \hat{u}(z,t)}{\partial z} = j\frac{\beta_2}{2}\frac{\partial^2 \hat{u}(z,t)}{\partial t^2} - j\gamma \exp(-\alpha z)\left|\hat{u}(z,t)\right|^2 \hat{u}(z,t), \tag{3.7}$$

where z is the propagation distance, α is the attenuation coefficient, $\beta_2 = -\lambda^2 D/2\pi c$ is the group-velocity dispersion parameter, λ is the optical carrier wavelength, c is the speed of light, D is the dispersion parameter at λ and γ is the Kerr nonlinearity coefficient, respectively.

3.4.1 Linear Channel Response (Zeroth-order Solution)

In this section, we describe the linear distortions added to the transmitted OFDM signal with the assumption that the nonlinearities are absent (i.e., $\gamma = 0$). As per the RP analysis detailed in the Appendix A, the linear transfer function of the fiber for a subcarrier frequency $w_k = 2\pi k/NT_s$, after substituting the value of β_2, can be represented in frequency-domain as:

$$H(z, w_k) \triangleq \exp(j\varphi_D(w_k))z, \tag{3.8}$$

where $\varphi_D(w_k) = \pi c D \left(\frac{w_k}{w_c}\right)^2$. Therefore, the linear solution of (3.7) at the output of the fiber at a length $z = L$ can be represented in time-domain as:

$$\hat{u}_0(L,t) = \hat{u}(0,t) \otimes h(L,t), \tag{3.9}$$

where subscript 0 represents the zeroth-order solution, $h(L,t)$ is the inverse Fourier transform of $H(L,w_k)$ and the symbol $\otimes$ stands for the convolution.
The signal is then coherently detected and down-converted to the baseband as $y_0(L,t)$. After sampled at $t = nT_s$, the complex symbol at the kth subcarrier is obtained by taking the DFT of the sampled signal $y_0(L,nT_s)$ and according to (3.8) as:

$$\begin{aligned}\hat{a}_{0,k} &= \sum_{n=0}^{N-1} y_0(L,nT_s)\exp(-j2\pi kn/N) \\ &= a_k \exp(j\varphi_D(w_k)L).\end{aligned} \tag{3.10}$$

The expression in (3.10) shows that the response of the linear fiber channel for the kth subcarrier is modeled as a multiplication with a phase-shift $\varphi_D(w_k)$, as far as each individual subcarrier is concerned [73].

3.4.2 Nonlinear Channel Response (First-order Solution)

This section describes the first-order solution of (3.7) in the presence of nonlinearities (i.e., $\gamma \neq 0$). Continuing the perturbation analysis given in the Appendix A, the first-order solution for (3.7) at the output of the fiber at a length $z = L$ can be represented in time-domain as:

$$\hat{u}_1(L,t) = \hat{g}(L,t) \otimes h(L,t), \tag{3.11}$$

where $\hat{g}(L,t)$ is a distorted optical field at the output of the fiber and is related to the linear solution $\hat{u}_0(L,t)$ (see (A.15) in Appendix A).

After down-conversion to the baseband and sampled at $t = nT_s$, the first-order nonlinear distortion term added to the symbol at the kth subcarrier is obtained after taking the DFT of the sampled signal $y_1(L, nT_s)$ multiplied by γ and following (A.16) to (A.19) in the Appendix A as:

$$\begin{aligned}\hat{a}_{1,k} &= \gamma \sum_{n=0}^{N-1} y_1(L, nT_s) \exp(-j2\pi kn/N) \\ &= \gamma G(L, w_k) \exp(-j\frac{\beta_2}{2} w_k^2 z),\end{aligned} \tag{3.12}$$

where $G(L, w_k)$ is given by (A.18) in Appendix A. After performing some algebra as shown in the Appendix A, the first-order distortion term can be represented as:

$$\hat{a}_{1,k} = (-j) a_k \exp(j\varphi_D(w_k)L)\varphi_{\mathrm{NL}} + \Delta a_{\mathrm{IFWM}}, \tag{3.13}$$

where

$$\varphi_{\mathrm{NL}} = \gamma \left[|a_k|^2 \, \Xi_{0,0}^{FO} + 2 \sum_{n \neq 0} |a_{k+n}|^2 \, \Xi_{0,n}^{FO} \right] \tag{3.14}$$

and

$$\Delta a_{\mathrm{IFWM}} = (-j)\gamma \exp(j\varphi_D(k)L) \sum_{m \neq 0} \sum_{n \neq 0} a_{k+m} a_{k+n}^* a_{k+m+n} \boldsymbol{\Xi}_{m,n}^{FO}. \tag{3.15}$$

In (3.15), $\boldsymbol{\Xi}_{m,n}^{FO}$ is the first-order nonlinearity coefficient matrix and is given as:

$$\boldsymbol{\Xi}_{m,n}^{FO} = \frac{1}{L} \int_0^L \exp(-A(z') - j\frac{1}{2} C(z')(w_m - w_k)(w_m - w_n)) dz', \tag{3.16}$$

where FO stands for the first-order, $A(z') = \int_0^{z'} \alpha(\xi)\, d\xi$, and $C(z') = \int_0^{z'} \beta_2(\xi)\, d\xi$. According to the perturbation theory, one can represent the received kth subcarrier

with the nonlinear distortions up to the first-order as:

$$\begin{aligned}\hat{a}_k &\approx \hat{a}_{0,k} + \hat{a}_{1,k} \\ &\approx a_k \exp(j\varphi_D(w_k)L)\,[1 + (-j)\varphi_{\mathrm{NL}}] + \Delta a_{\mathrm{IFWM}}. \end{aligned} \tag{3.17}$$

After substituting the approximation $[1 + (-j)\varphi_{\mathrm{NL}}] \approx \exp(-j\varphi_{\mathrm{NL}})$, (3.17) can be represented as:

$$\hat{a}_k \approx a_k \exp(j\varphi_D(w_k)L) \exp(-j\varphi_{\mathrm{NL}}) + \Delta a_{\mathrm{IFWM}}. \tag{3.18}$$

It is important to mention that (3.18) is similar to the *additive-multiplicative* perturbation model for the fiber nonlinearities given in [73]. Also, it is observed from (3.18) that the nonlinear distortion due to SPM and IXPM (i.e., φ_{NL}) just results in a constant phase-rotation and can be compensated by the carrier phase recovery at the receiver. Assuming an ideal dispersion compensation for each subcarrier frequency, (3.18) can be further simplified as:

$$\hat{a}_k \approx a_k + \Delta a_{\mathrm{IFWM}}, \tag{3.19}$$

where

$$\Delta a_{\mathrm{IFWM}} = -j\gamma \sum_{m\neq 0} \sum_{n\neq 0} a_{k+m} a^*_{k+n} a_{k+m+n} \Xi^{FO}_{m,n}. \tag{3.20}$$

When an anti-symmetric dispersion map and a symmetric power map is applied in the link such that $C(z') = -C(L-z')$ and $A(z') = A(L-z')$, then $\Xi^{FO}_{m,n}$ becomes real valued, i.e., $\Xi^{FO}_{m,n} \approx \left(\Xi^{FO}_{m,n}\right)^*$. Note that the symmetric power map in the transmission link can be considered as a loose requirement owing to the low loss profile of the silica optical fibers [70]. Since in the MOD-16-QAM-CDR/MOD-16-QAM-PCTW

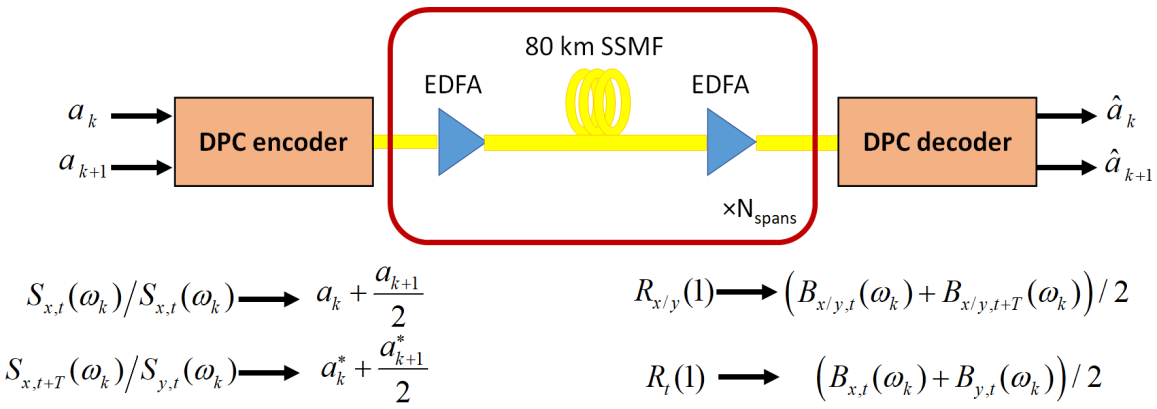


Figure 3.8: Illustration showing the MOD-16-QAM-CDR/MOD-16-QAM-PCTW technique for the first subcarrier of the transmitted OFDM symbol.

technique, the signal and its phase conjugate are transmitted on the two orthogonal time slots/polarizations, the nonlinear distortions added onto two twin signals are essentially anti-correlated. That eventually leads to the cancellation of the nonlinear distortions upon CS of the received twin signals at the receiver. In summary, by transmitting the complex conjugate of a signal on orthogonal time slot/polarization and superimposing the received twin signals at the receiver results in the cancellation of the nonlinear distortions to the first-order. This idea can be illustrated using an example, as shown in Fig. 3.8.

The data symbols a_k and a_{k+1} are linearly combined, one at full amplitude and the other at the half amplitude and transmitted as phase-conjugate pairs on the two orthogonal time slots/polarizations. Upon transmission through the polarization division multiplexed dispersive and nonlinear fiber channel, the perturbative nonlinear distortions are added to the twin signals on the orthogonal dimensions (time slots or polarizations) as $\left(a_k + \hat{a}_{1,k} + \left(\frac{a_{k+1}+\hat{a}_{1,k+1}}{2}\right)\right)$ and $\left(a_k^* + \hat{a}_{1,k} + \left(\frac{a_{k+1}^*+\hat{a}_{1,k+1}}{2}\right)\right)$, where $\hat{a}_{1,k/k+1}$ represents the first-order nonlinear distortion field added to the symbols a_k and a_{k+1} and is given by (3.20).

At the receiver, the CS yields:

$$\begin{aligned}\text{CS} &= \frac{\left[\left(a_k + \hat{a}_{1,k} + \left(\frac{a_{k+1}+\hat{a}_{1,k+1}}{2}\right)\right) + \left(\left(a_k^* + \hat{a}_{1,k} + \left(\frac{a_{k+1}^*+\hat{a}_{1,k+1}}{2}\right)\right)\right)^*\right]}{2} \\ &= \frac{\left[\left(a_k + \hat{a}_{1,k} + \left(\frac{a_{k+1}+\hat{a}_{1,k+1}}{2}\right)\right) + \left(a_k + (\hat{a}_{1,k})^* + \left(\frac{a_{k+1}+\left(\hat{a}_{1,k+1}\right)^*}{2}\right)\right)\right]}{2} \\ &= \left(a_k + \frac{a_{k+1}}{2}\right) + (\hat{a}_{1,k} + (\hat{a}_{1,k})^*) + \left(\frac{\hat{a}_{1,k+1} + (\hat{a}_{1,k+1})^*}{2}\right). \end{aligned} \tag{3.21}$$

When there is an ideal dispersion symmetry condition satisfied in the transmission link, then the perturbative nonlinear distortions added to the two data symbols are essentially anti-correlated, i.e., $(\hat{a}_{1,k})^* = -\hat{a}_{1,k}$ and $(\hat{a}_{1,k+1})^* = -\hat{a}_{1,k+1}$, and it is straightforward to obtain the linear coded symbol $\left(a_k + \frac{a_{k+1}}{2}\right)$ at the subcarrier frequency w_k of the received OFDM symbol in the two adjacent time slots/orthogonal polarizations.

It is important to note that the CS completely cancels the perturbative nonlinear distortions added to the transmitted signal fields, provided a dispersion symmetry condition is satisfied in the transmission link. That brings the performance gain for the proposed MOD-16-QAM-CDR/MOD-16-QAM-PCTW technique in the high nonlinear transmission regime when compared to the recently proposed PCSC technique. On the other hand, if there is no dispersion symmetry in the transmission link, then the (imaginary) amplitudes of the nonlinear distortion terms are unequal, i.e, $|(\hat{a}_{1,k})^*| \neq -|\hat{a}_{1,k}|$ and $|(\hat{a}_{1,k+1})^*| \neq -|\hat{a}_{1,k+1}|$. However, the CS yields the nonlinear distortion terms on the two orthogonal time slots/polarization states to be subtracted from each other. The resultant residual nonlinear distortion terms are smaller in magnitude and impart a considerably low performance penalty to the proposed MOD-16-QAM-CDR/MOD-16-QAM-PCTW technique. This unique feature helps the MOD-16-QAM-CDR/MOD-16-QAM-PCTW technique achieve significant

performance improvement in the absence of the dispersion symmetry in the transmission link when compared to the PCSC technique. The recovered symbol $\left(a_k + \frac{a_{k+1}}{2}\right)$ after the CS is input to the ML-detector for symbol detection. The ML-detector calculates the Euclidean distance between the received coded symbol and all the possible combinations of the linearly-coded symbols, on a symbol-by-symbol basis and selects the one with minimum distance. Then, with the help of a LUT, the data symbol which corresponds to the selected linearly-coded symbol is determined.

3.5 Numerical Simulation of the MOD-16-QAM-CDR/MOD-16-QAM-PCTW Techniques

3.5.1 Simulation Setup

The simulation setup used to study the performance of the proposed DPC approaches is shown in Fig. 3.9. The transmission system consists of the WDM CO-OFDM superchannel employing the MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques. Insets (a) and (b) show the encoder and decoder, respectively, for both MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques. The superchannel comprises four OFDM sub-bands with a frequency spacing of 37.5 GHz. The baud rate is 32 Gbaud. The OFDM symbol consists of 3300 data-carrying subcarriers, and the inverse fast Fourier transform (FFT) size is 4096 [58]. In each OFDM symbol, four pilot subcarriers are inserted for the common phase error compensation, and a cyclic prefix of 3% is added. Therefore, the net data rate is 401.33 Gb/s. The long-haul fiber link consists of 40 spans of SSMF, each having a length of 80 km, an attenuation coefficient of 0.2 dB/km, a nonlinear parameter of 1.22/(W.km), a dispersion parameter of 16 ps/nm/km and a PMD coefficient of 0.1 ps/$\sqrt{\text{km}}$. An erbium-doped fiber amplifier

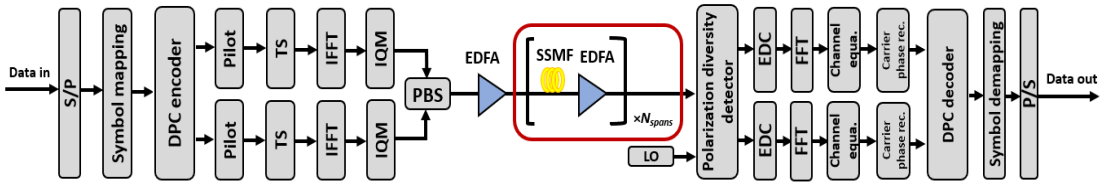


Figure 3.9: The simulation setup for one channel of the CO-OFDM superchannel system employing MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques. S/P: serial-to-parallel, TS: training symbol, IFFT: inverse fast Fourier transform, EDC: electronic dispersion compensation, IQM: inphase/quadrature phase modulator, PBS: polarization beam splitter, EDFA: erbium doped fiber amplifier, SSMF: standard single mode fiber, LO: local oscillator, P/S: parallel-to-serial.

compensates the optical power loss for each span with 16 dB gain and a 5.5 dB noise figure. The transmitter and receiver lasers have the same linewidth of 100 kHz. The ASE noise is added inline to ensure that the nonlinear interaction between the signal and noise is correctly captured [58].

At the receiver, after the polarization diversity detector, the dispersion compensation is performed using the overlapped frequency domain equalizer with the overlap-and-save algorithm [55]. The time-interleaved training symbols (two symbols every 100 symbols) are used to estimate the channel and use a one-tap equalizer for channel equalization. The common phase error estimation is based on the insertion of the pilot subcarriers [55]. After the OFDM processing, the coherent superposition of the received information symbols on the two polarizations is performed. Then, an ML-detector/ LUT block is used for symbol detection/decoding. Finally, the decoded symbols are demapped in the binary form.

3.5.2 Simulation Results

The performance of the CO-OFDM system with proposed MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques and the PCSC technique are compared in Fig. 3.10. The cases with and without pre-EDC are presented to show the advantage of the

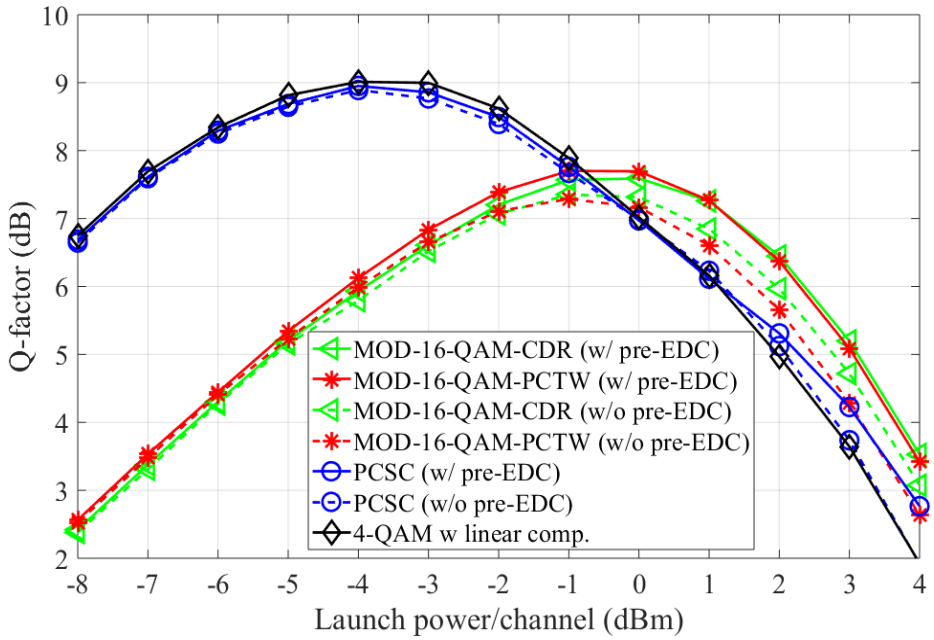


Figure 3.10: The simulation results for the MOD-16-QAM-CDR/MOD-16-QAM-PCTW, PCSC, and 4-QAM techniques for a transmission distance of 2800 km.

proposed MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques in the absence of dispersion symmetry. In this figure, the performance of 4-QAM with linear compensation, providing the same SE, is also presented. The results show that the proposed MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques outperform the 4-QAM with linear compensation and PCSC techniques in the highly nonlinear regime. It is important to mention that for long-haul transport systems, the highly nonlinear regime is attractive since it enables the longer transmission range. The proposed techniques show a Q-factor improvement of 1.5 dB and 1.2 dB with pre-EDC, when compared to the 4-QAM with linear compensation and PCSC technique, respectively, at a launch power of 2 dBm. This performance improvement is obtained through the ability of the proposed techniques to cancel the first-order nonlinear distortion fields through the coherent superposition when there is pre-EDC is applied at the transmitter. In the absence of pre-EDC, the coherent superposition causes the distortion fields to be subtracted from each other and retains a residual distortion term, which provides

a comparatively low penalty to the system performance. This feature of the MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques yields a considerable performance improvement without pre-EDC when compared to the PCSC technique.

On the other hand, the performance of the proposed MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques is highly limited in the linear (or weakly nonlinear) regime compared to the 4-QAM with linear compensation and PCSC schemes. That is due to the increased constellation set after the linear coding, and thereby, the performance is limited because of the OSNR penalty. Besides, without pre-EDC, the PCSC technique does not provide any improvement in the performance of the system, and this observation confirms the results given in [55].

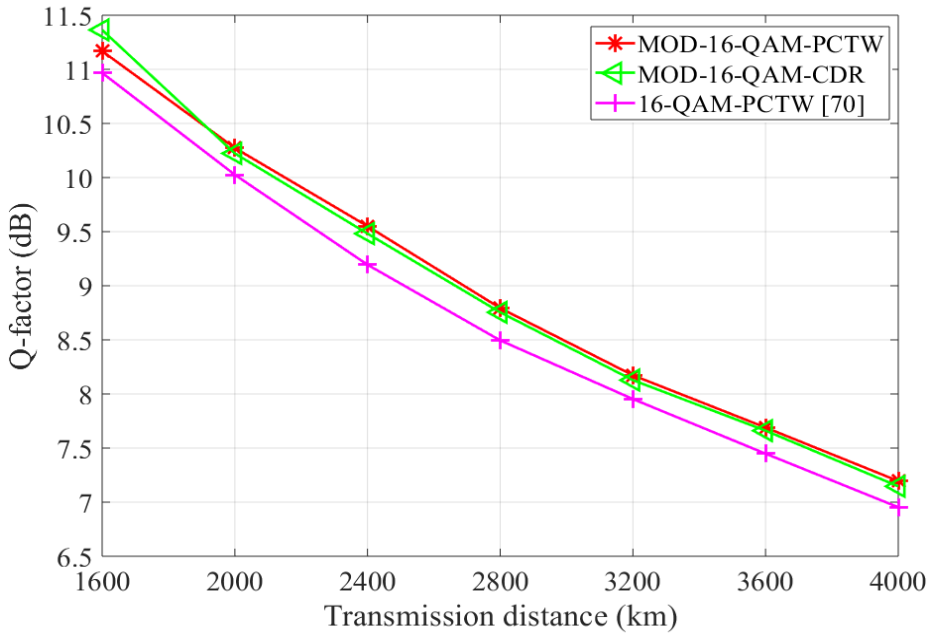


Figure 3.11: The Q-factor vs. transmission distance for the proposed MOD-16-QAM-CDR/MOD-16-QAM-PCTW schemes and the 16-QAM-PCTW technique proposed in [70] at a launch power of 2 dBm.

The performance comparison of the proposed MOD-16-QAM-CDR/MOD-16-QAM-PCTW schemes and the 16-QAM-PCTW technique proposed in [70] is shown in Fig. 3.11. The result indicates that the proposed MOD-16-QAM-CDR/MOD-16-QAM-

PCTW schemes have improved Q-factor performance when compared to the 16-QAM-PCTW technique proposed in [70] for all the transmission distance considered. For example, the Q-factor value of the proposed MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques is improved by ~ 0.3 dB when compared to the 16-QAM-PCTW technique at a transmission distance of 2800 km.

3.6 Complexity Analysis

In this section, we compare the computational complexity of the proposed MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques with that of the benchmark digital NLC technique such as single-channel (SC)-DBP, based on the number of multiplications per subcarrier. We assume that the linear dispersion compensation is implemented using the overlap-and-save algorithm. We further assume that the FFT block size, N_{FFT} of the DBP algorithm is the same as the FFT size used in the OFDM modulation and the DBP algorithm is implemented with one sample per symbol.

Table 3.1: Complexity analysis.

Algorithm	**Complexity expression**	**No. of multiplications**
SC-DBP	$N_{spans}N_{steps}(8\log_2(N_{FFT})+21)$	4680
MOD-16-QAM-CDR/MOD-16-QAM-PCTW	$8\log_2(N_{FFT})+4M+9$	169
Linear comp.	$8(\log_2(N_{FFT})+1)$	96

Table 3.1 provides a comparison of the estimated complexity for the DBP, MOD-16-QAM-CDR/MOD-16-QAM-PCTW and the linear compensation techniques, in

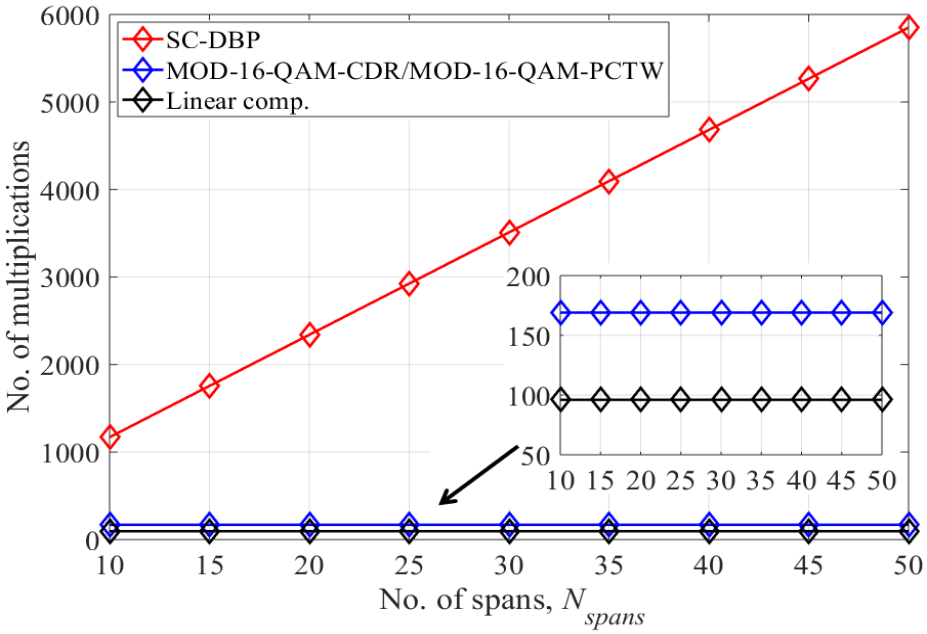


Figure 3.12: Number of multiplications as a function of number of fiber spans, N_{spans}.

terms of the number of multiplications, with parameters $N_{\text{FFT}} = 4096$, $N_{spans} = 40$, $N_{steps} = 1$, and the constellation cardinality $M = 16$.

Fig. 3.12 shows the complexity of the DBP, MOD-16-QAM-CDR/MOD-16-QAM-PCTW, and linear compensation as a function of the number of spans, N_{spans}. The parameters used are the same as the one considered for the estimation of the complexity in Table 3.1. It is observed that the computational complexity of the proposed MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques is independent of the fiber length and only slightly higher than the complexity of the linear compensation case.

Fig. 3.13 shows the computation time in terms of the central processing unit (CPU) running time for the proposed MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques. The results indicate that the computation time of the proposed MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques is significantly lower than that of the SC-DBP technique and only slightly higher than that of the linear compensation case. It is worth mentioning that the computational complexity of the DBP technique

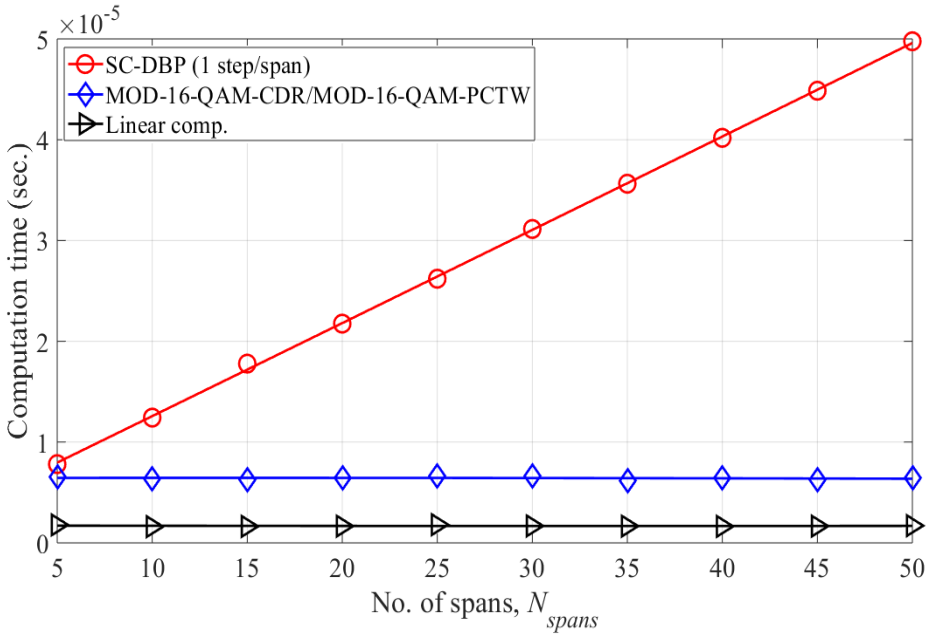


Figure 3.13: The CPU running time as a function of number of fiber spans, N_{spans}.

linearly increases with an increase in the N_{steps} per span. For example, the computational complexity of the DBP technique with $N_{steps} = 16$ is 16 times larger than the case with $N_{steps} = 1$.

3.7 PDL Impact on the MOD-16-QAM-CDR/MOD-16-QAM-PCTW Techniques

In this section, we investigate the impact of the PDL on the performance of the MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques. The PDL-induced signal power imbalance between the two polarizations disrupts the cross-correlation property of the nonlinear impairments, which can affect the distortion cancellation through the CS of the MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques. We carry out the investigation with both aligned- and statistical-PDL models.

3.7.1 The Aligned- and Statistical-PDL Models

The PDL is caused by the polarization dependence on the transmission properties of optical components, where one polarization component of the signal suffers more loss than the other. The input/output field relation of a PDL element, rotated with respect to the signal SOP by an angle θ, can be expressed using the Jones matrix representation as [74]:

$$\begin{bmatrix} u_x(t) \\ u_y(t) \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} v_x(t) \\ v_y(t) \end{bmatrix}, \tag{3.22}$$

where $[v_x(t)\ v_y(t)]^\dagger$ and $[u_x(t)\ u_y(t)]^\dagger$ represent the input and output optical fields, respectively, with the superscript $\dagger$ as the transpose. The parameter $0 < \alpha < 1$ is the PDL coefficient defined as the ratio between the minimum and maximum transmission intensities; this is related to the PDL measured in decibels as $\rho = -20\log\alpha$ [75].

The PDL impact has been studied in coherent optical systems using two different models: the aligned- and statistical-PDL models [71], [76]. In the aligned-PDL model, the signal SOP and the PDL axes of the optical components are aligned with the same rotation angle θ. Fig. 3.14(a)-(b) shows two cases for the aligned-PDL model with $\theta = 0^0$ and 45^0, respectively [77]. At the rotation angle $\theta = 0^0$, the OSNR of one of the polarization is degraded when compared to the other. On the other hand, for $\theta = 45^0$, the PDL causes the same OSNR degradation for both polarization components along with the signal cross-talk due to the loss of orthogonality. In [76], it has been shown that pathological cases of the aligned-PDL elements, such as $\theta = 0^0$ and 45^0, are the worst cases of PDL in linear and nonlinear regimes, respectively.

In the statistical-PDL model, the rotation angle θ varies uniformly within $[0, 2\pi)$, as shown in Fig. 3.14(c). That induces the random signal power and OSNR fluctu-

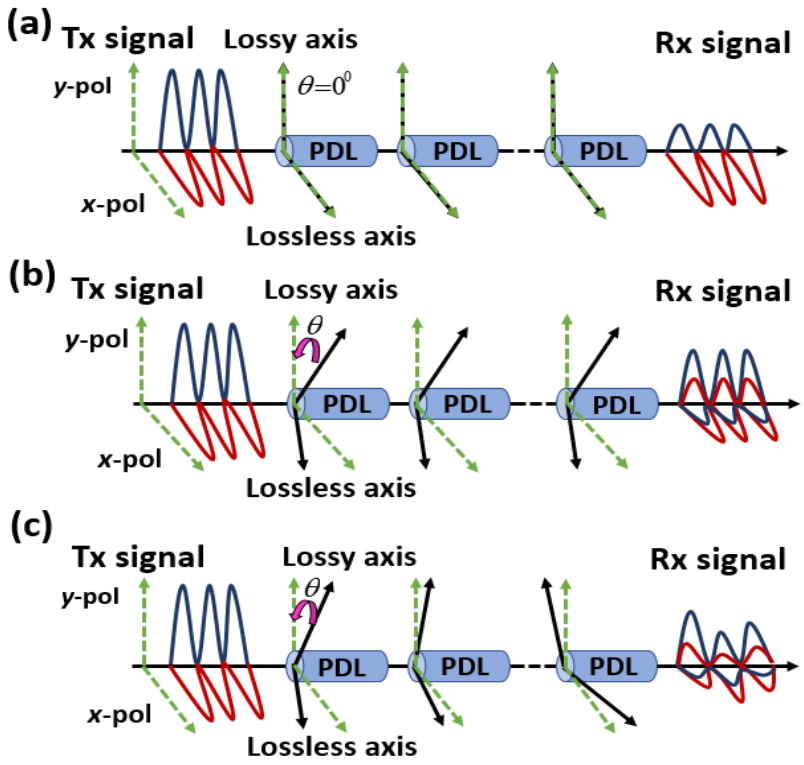


Figure 3.14: Schematic diagram of the impact of PDL on the PDM signal. (a) aligned-PDL with $\theta = 0^0$, (b) aligned-PDL with $\theta = 45^0$, and (c) statistical-PDL with random rotation angle θ.

ations between the two polarizations. In this case, the total cumulated PDL has a Maxwellian distribution with the root mean square (rms) value $\rho_{rms} = \rho\sqrt{N}$, where N_{spans} is the number of spans [78].

The simulation setup used to study the PDL impact on the DPC approaches is shown in Fig. 3.15. We consider a 5-section PDL emulator, which closely approximates a real system [71]. In this setup, the signal interacts with the PDL element after propagating through eight spans of SSMF. Five such loops realize a 5-section PDL emulator. The PDL along the transmission link mainly comes from the lumped optical elements. In a realistic transmission link, such optical elements are placed after several fiber spans. Therefore, placing a PDL element after eight spans of fiber is sufficient to study its impact on the performances of the DPC approaches [71]. A

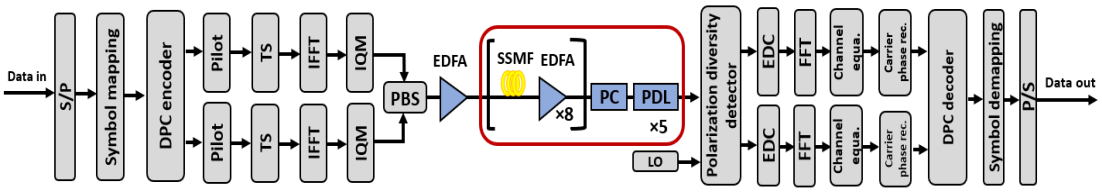


Figure 3.15: The simulation setup for one channel with a 5-section PDL emulator along the link.

polarization controller is placed before the PDL element to control the signal SOP after each round trip. Insets (a) and (b) show the encoder and decoder, respectively, for both MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques.

3.7.2 Performance Evaluation with Aligned-PDL

The DPC is a generalized technique in which one can use orthogonal polarization states or time slots to transmit the phase conjugate pairs [58]. The motivation behind the MOD-16-QAM-PCTW technique is to solve the issue of halving the SE associated with the PCTW technique [58]. Its effectiveness strongly depends on the cross-correlation between the nonlinear distortions added onto the transmitted phase conjugate pairs on the two polarizations. However, the polarization cross-talk induced signal power imbalance between the two polarizations may disrupt this cross-correlation property and significantly degrade the performance of the MOD-16-QAM-PCTW technique. For this reason, we have proposed the MOD-16-QAM-CDR approach, in which we transmit the linearly coded phase conjugate pairs on adjacent time slots of the same polarization.

In Fig. 3.16, the Q-factor performance of the DPC approaches is shown for different values of the fiber launch power at a fixed aligned-PDL. In this case, we select a PDL value of 3.6 dB. That corresponds to the rms value of the cumulated PDL in the

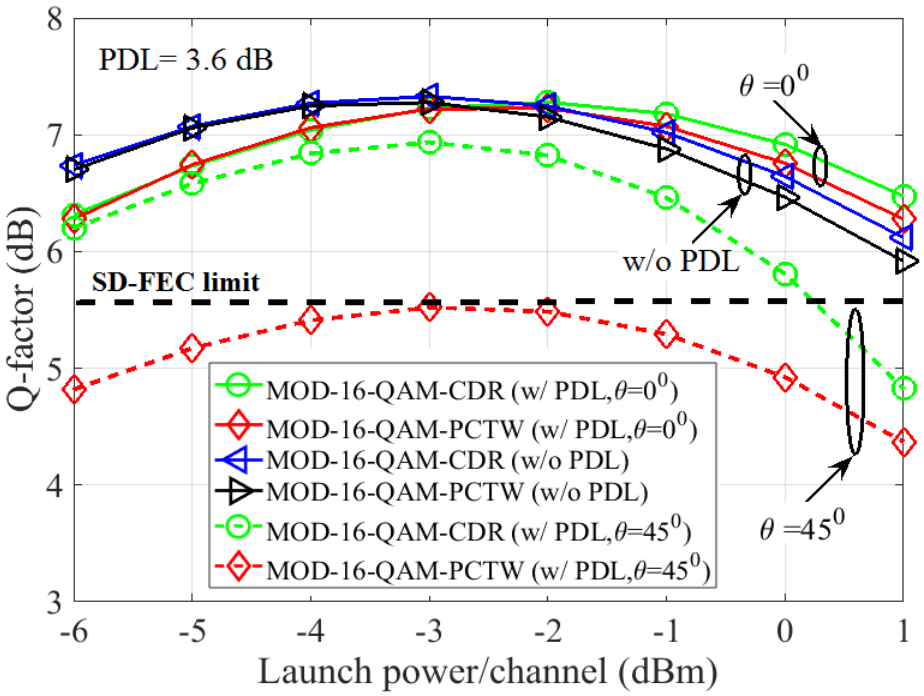


Figure 3.16: Q-factor as a function of launch power in the presence and absence of the PDL.

evaluation with the statistical-PDL model, as it will be discussed in Subsection 3.7.3. We start analyzing the linear regime, i.e., the initial increasing part of the Q-factor curves. For both DPC approaches, the performance in the presence of PDL when $\theta = 0^0$ is lower when compared to the absence of PDL. This degradation in performance can be explained by the impact of the PDL-induced OSNR imbalance between the two polarizations. For $\theta = 45^0$, the performance of the MOD-16-QAM-PCTW is significantly reduced when compared to the MOD-16-QAM-CDR. That is due to the signal cross-talk induced power fluctuations on the two polarizations along with the OSNR degradation. At high input powers, where the performance is limited by the nonlinear distortions, the DPC approaches with $\theta = 0^0$ perform slightly better than the case without PDL. That can be explained by the decrease of the higher-order nonlinear distortions. In fact, in the presence of PDL, the signal in one polarization is attenuated more than in the other. That leads to the reduction of higher-order

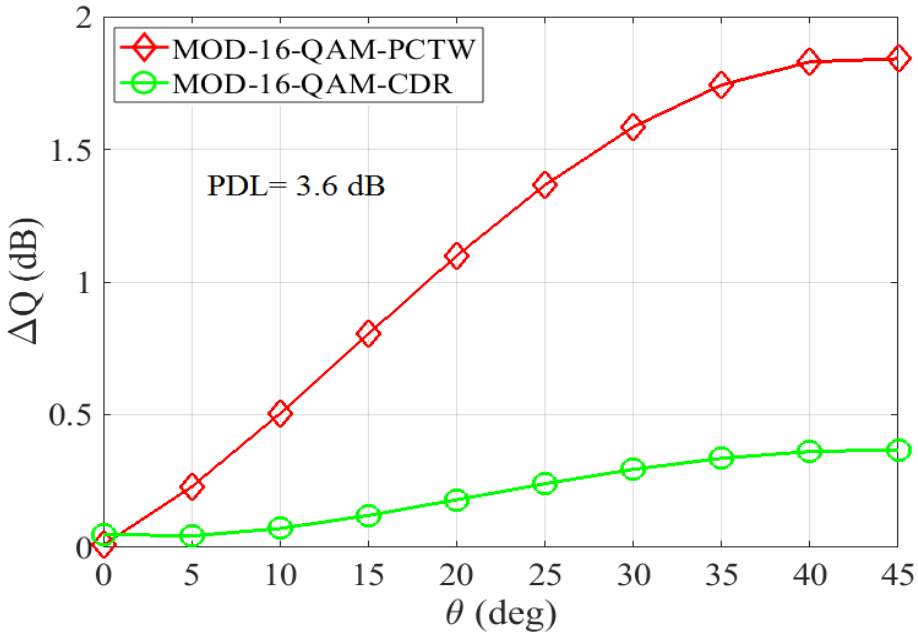


Figure 3.17: Q-factor penalty (ΔQ) for different rotation angles, θ. The optical launch power is fixed at the optimum value of -3 dBm.

nonlinear distortions after the coherent superposition. Note that the higher-order nonlinear distortions are not canceled by the DPC approaches [70]. It is observed that the performance of the MOD-16-QAM-PCTW technique with $\theta = 45^0$ degrades when compared to the MOD-16-QAM-CDR technique in both linear and nonlinear regimes. That is because the polarization cross-talk, due to the loss of orthogonality, causes signal power fluctuations on the two polarizations [72].

In Fig. 3.17, the performance of both DPC approaches is presented as a function of the rotation angle θ in terms of the Q-factor penalty. The Q-factor penalty is defined as $\Delta Q = Q_{opt} - Q$, where Q_{opt} is the Q-factor at the optimum launch power when PDL is not considered. It is seen that the Q-factor penalty is maximum at the rotation angle $\theta = 45^0$ and minimum at $\theta = 0^0$. We also observe that the cross-talk induced Q-factor penalty at $\theta = 45^0$ for the MOD-16-QAM-PCTW technique is 1.84 dB, while it is only about 0.35 dB for the MOD-16-QAM-CDR technique. It should be noted that further increasing the angle from $\theta = 45^0$ to 90^0 would result in the

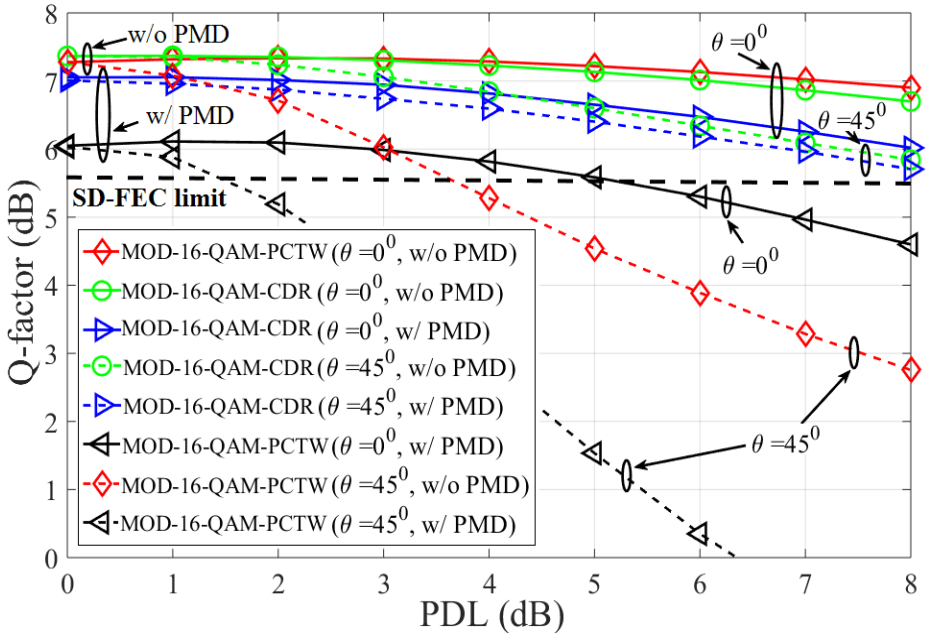


Figure 3.18: Q-factor as a function of PMD and PDL with rotation angles $\theta = 0^0$ and 45^0 at the optimum launch power of -3 dBm. SD-FEC: soft-decision forward error correction.

mirror image of the plot with a minimum Q-factor penalty at $\theta = 90^0$ because of the lower cross-talk induced power fluctuations. Therefore, we provide results for angles ranging from $\theta = 0^0$ to 45^0 only.

In Fig. 3.18, we consider two simulation scenarios to investigate the performances of the DPC approaches: one is with PDL alone and the other is with the PMD and PDL. In both cases, the performances are shown for the two worst case aligned-PDL scenarios with the rotation angles $\theta = 0^0$ and 45^0, respectively. The optical launch power is set at the optimum value of -3 dBm per channel. When considering only the PDL, the performance of the MOD-16-QAM-PCTW technique with $\theta = 45^0$ monotonically decreases as the PDL value increases. The MOD-16-QAM-CDR technique shows an improved performance, above the soft-decision forward error correction (SD-FEC) limit for the two considered worst case scenarios of the aligned-PDL. The PMD effect is included in the transmission fiber by choosing a typical mean differential

group delay of 20 ps, as in [79]. It is observed that the performances of the DPC approaches are significantly affected for the case considering both PMD and PDL. The interplay between PMD and PDL distorts a communication system more than either effect alone.

3.7.3 Performance Evaluation with Statistical-PDL

The Q-factor distribution presents a more realistic impact of the PDL on the performances of DPC approaches. Fig. 3.19 shows the estimated probability density function (PDF) of the Q-factor in the presence and absence of PDL for both MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques.

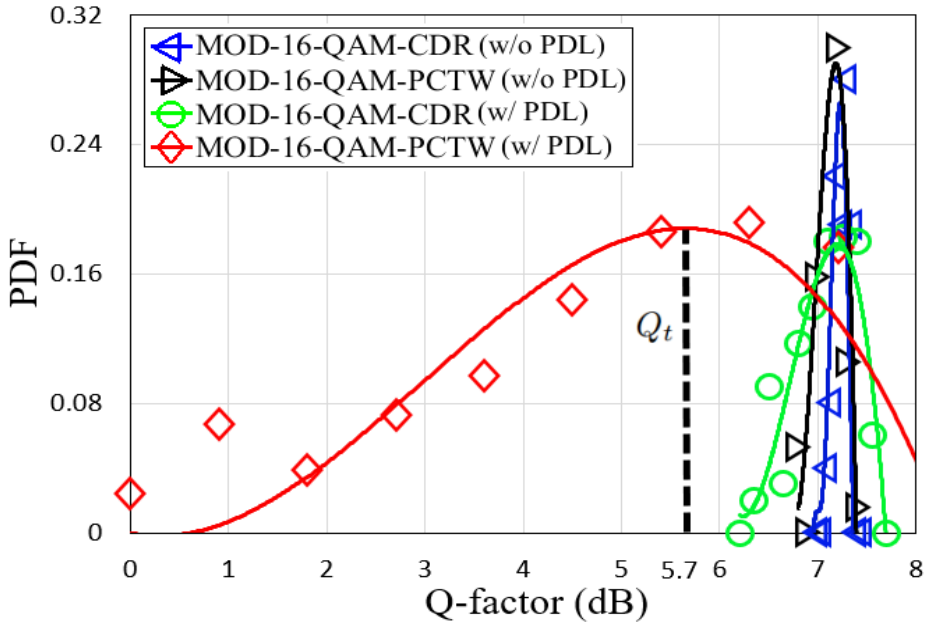


Figure 3.19: Q-factor PDF for MOD-16-QAM-PCTW and MOD-16-QAM-CDR at $\rho_{\text{rms}} = 3.6$ dB and optical launch power $= -3$ dBm.

We carried out Monte Carlo simulation to estimate the Q-factor PDF by using 500 random seeds of the signal SOP and the PDL orientation angle θ in the limit $[0, 2\pi)$. We select a typical PDL value of $\rho = 1.6$ dB [71]. That gives an rms cumulated PDL

value for a 5-section PDL emulator of 3.6 dB (i.e., $1.6 \times \sqrt{5}$). The fiber launch power is fixed at the optimum value of -3 dBm. The results indicate that without PDL, the Q-factor distributions for both DPC approaches are very narrow. On the other hand, in the presence of PDL, the Q-factor distribution of MOD-16-QAM-PCTW significantly enlarges, which leads to an increased outage probability. We define the outage probability as the probability that the random Q-factor is less than a particular threshold value, i.e., $\mathbb{P}\mathrm{r}[Q < Q_t]$, where Q_t is the threshold. For instance, assuming a Q_t value of 5.7 dB corresponding to the SD-FEC limit [81], the outage probability for MOD-16-QAM-PCTW in the presence of PDL is 0.63. However, it is observed that the outage probability for MOD-16-QAM-CDR approaches zero in the presence of PDL. That indicates that the approach which uses the orthogonal time slots of the same polarization is only slightly affected by the PDL-induced distortions.

3.8 Conclusion

In this chapter, we discussed the background of the problem associated with the recently proposed DPC approaches, such as PCTW and PCSC. We found that the performance improvement of the PCTW technique comes with halving the SE of the PDM coherent optical system. The PCSC technique can be effectively applied without SE loss; however, it does not provide any performance improvement without applying pre-EDC in the transmission link. On this ground, we proposed two linear coding techniques, referred to as MOD-16-QAM-CDR/MOD-16-QAM-PCTW. The proposed techniques can be used for the nonlinearity mitigation without halving the SE of the PDM coherent optical systems. They also show considerable performance improvement in the absence of pre-EDC. Furthermore, we investigated the performance penalties induced by the PDL on the MOD-16-QAM-CDR/MOD-16-QAM-PCTW

techniques with both aligned- and statistical-PDL models. In the investigation with the aligned-PDL, MOD-16-QAM-CDR shows a superior PDL tolerance when compared to MOD-16-QAM-PCTW. The Q-factor performance of the former is above the SD-FEC limit for the pathological cases of all aligned-PDL with $\theta = 0^0$ and 45^0. The investigation with the statistical-PDL model also indicates that MOD-16-QAM-CDR outperforms the MOD-16-QAM-PCTW technique, with the former providing an outage probability approaching zero for an rms PDL value of 3.6 dB. We concluded that while MOD-16-QAM-PCTW is severely affected by the PDL-induced distortions, MOD-16-QAM-CDR still provides good performance under such conditions.

Chapter 4

A Joint Technique for Fiber Nonlinearity Compensation in CO-OFDM Superchannel Systems

4.1 Preamble

This chapter is compiled from the materials extracted from the manuscript titled *"A joint technique for nonlinearity compensation in CO-OFDM superchannel systems"* which appeared in the proceedings of the *Asia Communications and Photonics Conference, Nov. 2017* [92].

4.2 Introduction

In recent years, optical communication networks have experienced an exponentially rising capacity demand [1]. The key technology drivers are the widespread use of cloud services, online gaming, internet of things, etc. To meet the ever-increasing

capacity demands, subcarrier-multiplexing, known as superchannel [11], combined with highly spectrally efficient modulation formats, represents the potential candidate. In the superchannel approach, the wavelength division multiplexed (WDM) channel is split into several subchannels with smaller bandwidths and separated by small guard-bands. The higher-order modulation formats that are applied to each subchannel to reach the desired data rate require a high optical signal-to-noise ratio. That leads to the requirement of high input power. The use of smaller guard-bands and high input power in superchannel systems results in substantial intra- and inter-channel nonlinear effects. Several digital nonlinearity compensation (NLC) techniques have been proposed in the last decade to deal with the nonlinear effects. Single-channel (SC) digital back-propagation (DBP) is a widely investigated technique to compensate for intra-channel nonlinear effects [31]. However, the reported performance gains are limited to $\sim$ 1 dB when applied to WDM superchannel systems [31]. The intra- and inter-channel deterministic nonlinear effects can be effectively compensated by applying a multi-channel (MC) DBP [65]. In contrast to SC-DBP, MC-DBP back-propagates the entire WDM superchannel. On the other hand, the implementation of MC-DBP is impractical in a dynamic optical network due to several factors, including [68]:

- The unavailability of the information from the neighboring traffic channels.
- The large computational complexity due to several linear and nonlinear computation steps per fiber span.

The recently proposed phase-conjugated twin wave (PCTW) can be implemented with minimal digital signal processing, providing a simple and effective solution for both intra- and inter-channel nonlinearity mitigation [49]. However, the results given in [50] indicate that the PCTW technique does not enable highly spectrally efficient

transmission systems. For example, the transmission performance of polarization division multiplexed (PDM) 16-quadrature amplitude modulation (16-QAM)-PCTW scheme is lower than that of the PDM quadrature phase-shift keying for the same single-fiber capacity.

In this chapter, we propose a joint technique that combines SC-DBP with the PCTW technique. This scheme, which is referred to as SC-DBP-PCTW, is to exploit the individual advantages of both techniques in compensating fiber nonlinearity effects; SC-DBP compensates for the intra-channel nonlinear effects, while the PCTW compensates both intra- and inter-channel first-order nonlinear effects. Thus, the joint SC-DBP-PCTW technique realizes a two-stage compensation for the intra-channel nonlinear effects and a first-order cancellation for the inter-channel nonlinear effects. The main contributions of this chapter are as follows:

- We develop a joint SC-DBP-PCTW technique which can compensate for both intra- and inter-channel nonlinearity effects.
- We show that the proposed technique has a similar performance as the MC-DBP technique with 16 steps/span.
- We also show that the proposed technique has a low computational complexity when compared to the MC-DBP technique with 16 steps/span.

4.3 The Joint SC-DBP-PCTW Technique

The concept of the joint SC-DBP-PCTW technique is depicted in Fig. 4.1. At the transmitter, the orthogonal frequency division multiplexed (OFDM) subcarriers on

each polarization is encoded as:

$$S_x(k) = A(k)$$
$$S_y(k) = S_x^*(k) = A^*(k), \tag{4.1}$$

where A and S represent the OFDM symbols before and after the encoder, $k = 1, 2, ..., N$, where N is the subcarrier number, x and y represent the horizontal and vertical polarizations, and $*$ stands for the complex conjugation operation. After encoding, the PCTWs are transmitted through standard single mode fiber (SSMF) with N_{span} number of fiber spans.

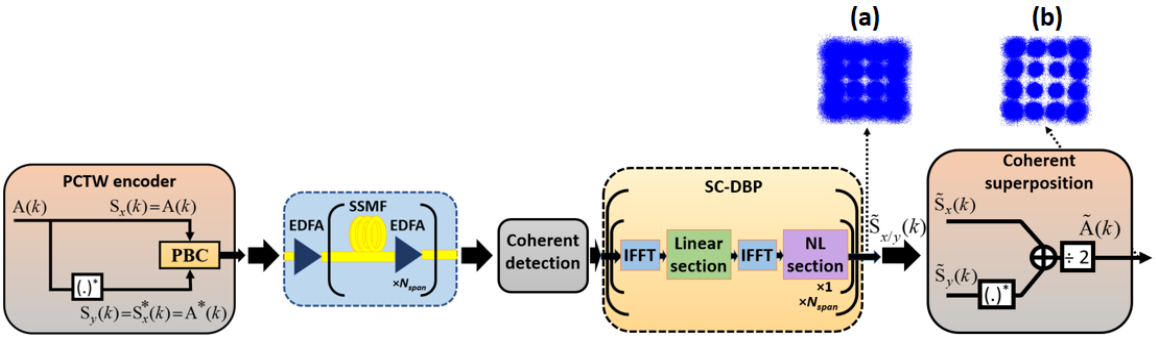


Figure 4.1: Illustration showing the joint SC-DBP-PCTW technique for one channel. E_x and E_y represent the transmitted electric fields in the x and y polarizations, respectively; $\tilde{E}_x$ and $\tilde{E}_y$ are the received electric fields after SC-DBP; and $\tilde{E}$ represents the recovered field after the coherent superposition, $*$ stands for the complex conjugation operation. N_{span}: number of fiber spans, EDFA: erbium doped fiber amplifier, SSMF: standard single mode fiber.

At the receiver, after coherent detection, SC-DBP of the selected channel is carried out with 1 step/span. Note that the implementation of SC-DBP involves a concatenation of linear and nonlinear sections; linear section performs dispersion compensation in frequency-domain, while nonlinear section compensates nonlinear phase shift due to Kerr effect in time-domain. The fast Fourier transform (FFT) and inverse FFT (IFFT) are used to switch between frequency and time domains. The SC-DBP is

followed by coherent superposition of the PCTW technique, which can be represented as:

$$\tilde{A}(k) = (\tilde{S}_x(k) + \tilde{S}_y^*(k))/2, \tag{4.2}$$

where $\tilde{S}$ and $\tilde{A}$ are the OFDM symbols before and after the coherent superposition. Insets (a) and (b) show the signal constellations after the SC-DBP and the coherent superposition of the PCTW technique. Evidently, the constellation quality is much improved after the coherent superposition. The performance improvement of the joint technique comes from the individual abilities of the two constituent techniques in combating the impact of nonlinearities. The SC-DBP compensates for the deterministic intra-channel nonlinear distortions, while the PCTW technique compensates both intra- and inter-channel first-order nonlinear distortions.

4.4 Numerical Simulation of the SC-DBP-PCTW Technique

4.4.1 Simulation Setup

Fig. 4.2 shows the simulation setup for the joint SC-DBP-PCTW technique. The transmission system consists of a WDM superchannel with four 37.5 GHz spaced 32 Gbaud 16-QAM-CO-OFDM signals employing the PCTW technique. The OFDM symbol consists of 3300 data-carrying subcarriers, and an IFFT of size 4096 is carried out to convert the signal into a time-domain. There are four pilot subcarriers in each OFDM symbol, and the cyclic prefix is 3%. Therefore, the net data rate is 401.33 Gb/s. The long-haul fiber link consists of 25 spans of standard single-mode fiber (SSMF), each having a length of 80 km, the attenuation coefficient of 0.2 dB/km, the nonlinearity coefficient of 1.22/(W.km), the dispersion coefficient of 16 ps/nm/km,

and the polarization mode dispersion coefficient of 0.1 ps/$\sqrt{\text{km}}$. The optical power loss for each span is compensated by an erbium-doped fiber amplifier (EDFA) with a 16 dB gain and a 4 dB noise figure. The transmitter and receiver lasers have the same linewidth of 100 kHz. At the receiver, after the polarization diversity detector, the SC-DBP with a 1 step/span is carried out. The channel equalization and carrier phase recovery are carried out as in [80]. After that, the coherent superposition of the PCTW technique is performed. Finally, the recovered symbols are demapped in the binary form.

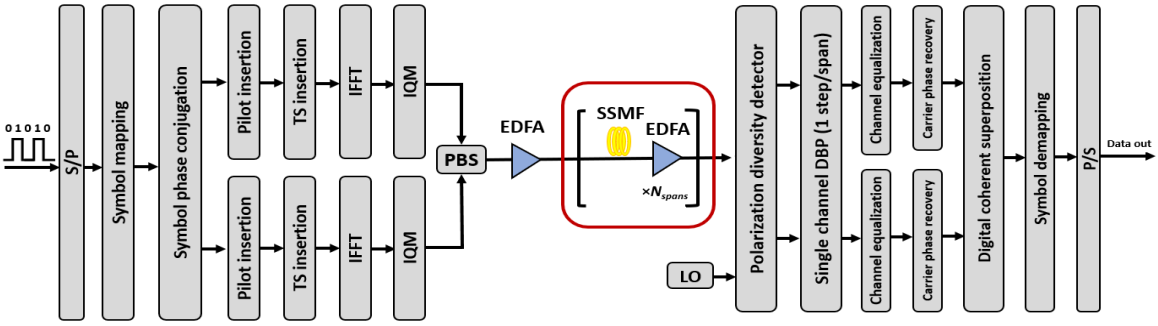


Figure 4.2: Simulation setup for the proposed SC-DBP-PCTW technique for one channel. S/P: serial-to-parallel, TS: training symbol, (I)FFT: (inverse) fast Fourier transform, IQM: inphase/quadrature phase modulator, PBS: polarization beam splitter, LO: local oscillator, P/S: parallel-to-serial.

4.4.2 Simulation Results

We evaluate the performance of the proposed SC-DBP-PCTW scheme, which is compared with the MC-DBP, PCTW, SC-DBP, and LDC techniques. It is evident from Fig. 4.3 that the proposed scheme improves the Q-factor performances by 3 dB, 2.3 dB, and 0.5 dB when compared to the LDC, SC-DBP, and PCTW schemes, respectively. It is interesting to note that the Q-factor performance of the proposed SC-DBP-PCTW scheme is similar to that of the MC-DBP with 16 steps/span, showing the effectiveness of the proposed technique in improving the performance-complexity trade-off.

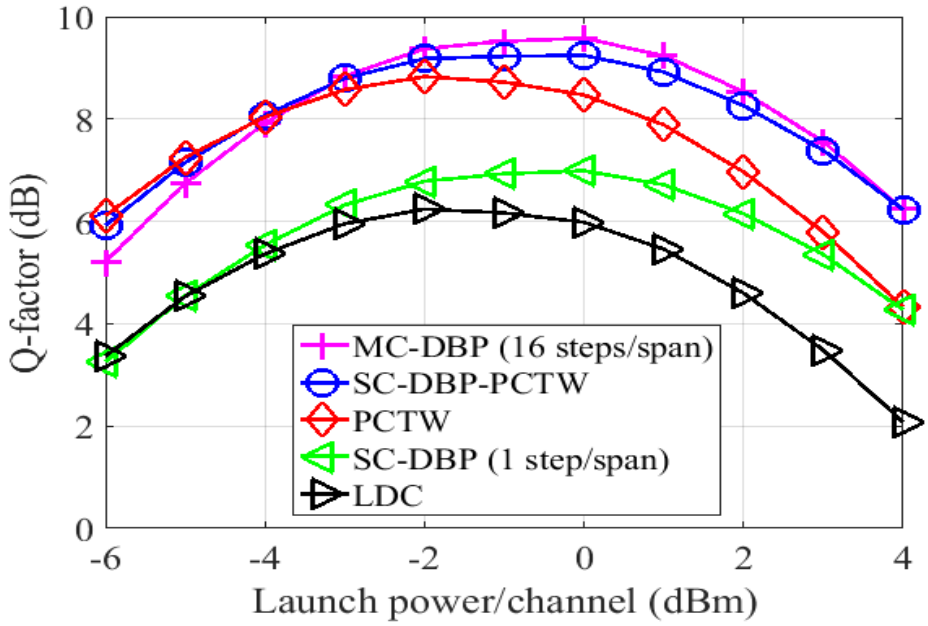


Figure 4.3: Q-factor versus launched power for the 401.33 Gbps CO-OFDM super-channel system with 16-QAM modulation for the MC-DBP (16 steps/span), SC-DBP-PCTW, PCTW, SC-DBP (1 step/span) and LDC techniques.

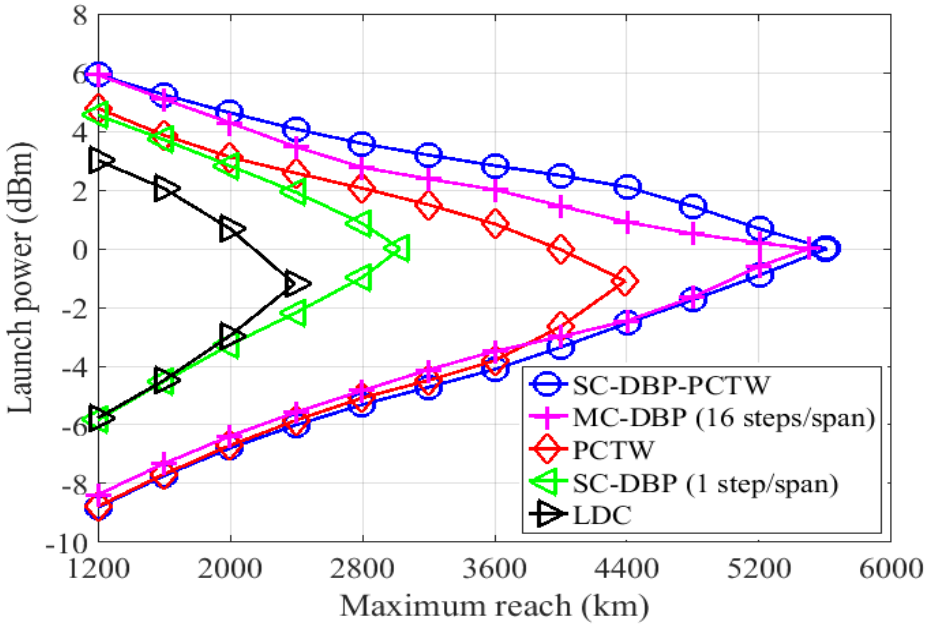


Figure 4.4: Estimated maximum signal reach for the MC-DBP (16 steps/span), SC-DBP-PCTW, PCTW, SC-DBP (1 step/span) and LDC techniques at 20% OH SD-FEC limit.

Fig. 4.4 shows an estimate of the maximum reach, including input power optimization for each propagation distance, at the 20% overhead (OH) soft-decision (SD) forward error correction (FEC) limit with a bit error rate value of 2.7×10^{-2} [81]. It is observed that the maximum reach for the LDC, SC-DBP, PCTW, MC-DBP and SC-DBP-PCTW is 2380 km, 3030 km, 4380 km, 5580 km and 5600 km, respectively. That indicates that the SC-DBP-PCTW scheme provides more than double transmission reach when compared to the LDC case and a similar reach as that of MC-DBP with 16 steps/span. It also shows a $\sim 85\%$ and $\sim 28\%$ reach increase when compared to the SC-DBP and PCTW schemes, respectively. It should be noted that the implementation of the PCTW technique halves the spectral efficiency, and thereby the performance improvement of the proposed technique comes with a cost of spectral efficiency loss.

4.5 Complexity Analysis

In this section, we compare the computational complexity of the proposed SC-DBP-PCTW technique with that of the LDC, SC-DBP, PCTW, and MC-DBP schemes in terms of the number of multiplications per subcarrier. Table 4.1 shows the expressions for the number of multiplications per subcarrier for the considered algorithms with $N_{\mathrm{FFT}} = 4096$ and $N_{span} = 25$.

It is observed that the joint scheme has a complexity less than that of the sum of the individual complexities of SC-DBP and PCTW techniques. The implementation of SC-DBP involves the LDC followed by a nonlinear compensation section, whereas, for the PCTW scheme, the LDC is followed by a coherent superposition operation. Thus, the technique combining SC-DBP with the PCTW scheme has a slightly increased complexity, when compared to its individual implementations. These addi-

tional complexities are from the nonlinear compensation section or from the coherent superposition operation.

Table 4.1: Complexity expression.

Algorithm	**Complexity expression (No. of multiplications)**
MC-DBP (16 steps/span)	$16N_{span}(8\log_2(N_{FFT})+21)$
SC-DBP-PCTW	$N_{span}(8\log_2(N_{FFT})+21)+1$
SC-DBP (1 step/span)	$N_{span}(8\log_2(N_{FFT})+21)$
PCTW	$8N_{span}(\log_2(N_{FFT})+1)+1$
LDC	$8N_{span}(\log_2(N_{FFT})+1)$

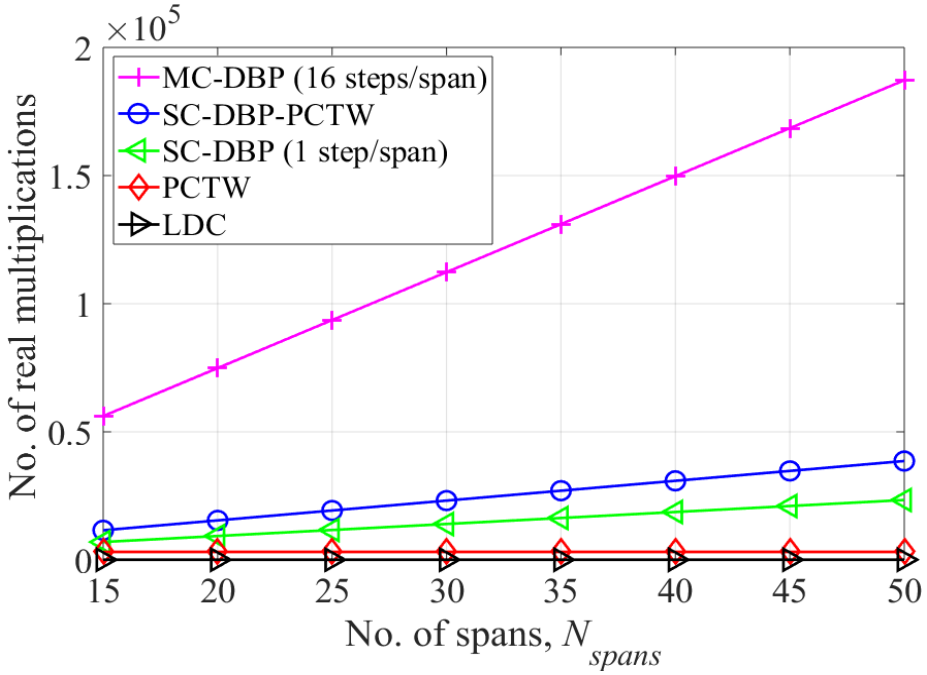


Figure 4.5: The computational complexity for the MC-DBP (16 steps/span), SC-DBP-PCTW, SC-DBP (1 step/span), PCTW and LDC techniques.

Fig. 4.5 shows the number of multiplications for the considered algorithms as a function of the number of spans, N_{span}. The results indicate that the complexity of the proposed SC-DBP-PCTW scheme is significantly lower than that of MC-DBP and only slightly higher than SC-DBP with a 1 step/span. It is worth mentioning that the complexities of PCTW and LDC are very close to each other. However, the slight increase in the complexity of PCTW, when compared to LDC, comes from the coherent superposition operation.

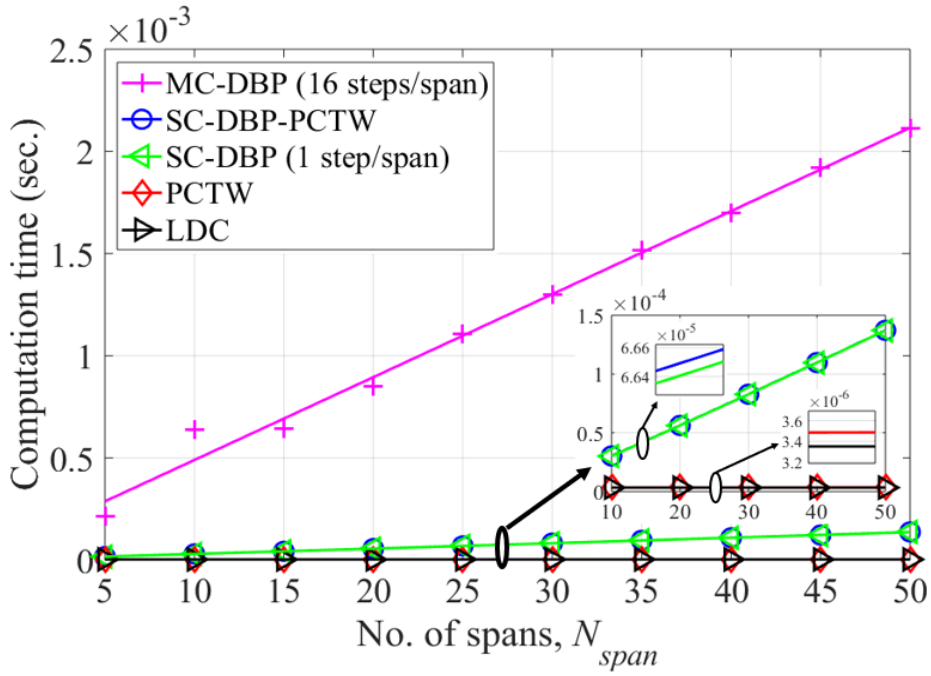


Figure 4.6: The CPU running time for the MC-DBP (16 steps/span), SC-DBP-PCTW, SC-DBP (1 step/span), PCTW and LDC techniques.

Fig. 4.6 shows the central processing unit (CPU) running time as a function of the number of fiber spans for the MC-DBP (16 steps/span), SC-DBP-PCTW, SC-DBP (1 step/span), PCTW and LDC techniques. The results show that the CPU running time of the proposed SC-DBP-PCTW scheme is significantly lower than that of MC-DBP with 16 steps per span. Also, it is observed from the inset of Fig. 4.6 that the computation time of the proposed SC-DBP-PCTW technique is only slightly higher

than the SC-DBP with a 1 step/span.

4.6 Conclusion

In this chapter, we discussed the background of the problem associated with the individual implementations of the SC-DBP, MC-DBP, and PCTW techniques to compensate for the intra- and inter-channel nonlinear effects. Based on the background study, we proposed a low-complexity joint technique for fiber nonlinearity compensation, which combines the SC-DBP and PCTW techniques. The proposed scheme provides a performance gain higher than applying the SC-DBP and PCTW techniques individually in a 401.33 Gbps 16-QAM-CO-OFDM superchannel system, at a transmission distance of 2000 km. It also almost doubles the transmission reach when compared to the LDC case and provides about 28% increase compared to the PCTW technique. In addition to that, the proposed technique shows similar performance as the MC-DBP with 16 steps/span.

Chapter 5

Enhanced Regular Perturbation-Based Fiber Nonlinearity Compensation Technique for Optical Transmission Systems

5.1 Preamble

This chapter is compiled from the materials extracted from the manuscript titled *"Enhanced regular perturbation-based nonlinearity compensation technique for optical transmission systems"* published in the *IEEE Photonics Journal, Aug. 2019* [47].

5.2 Introduction

The intra-channel fiber nonlinearity effect is considered a dominant impairment in a dispersion unmanaged optical communication systems [82]. However, the deterministic nature of intra-channel nonlinearity allows its electronic compensation either at the transmitter as a pre-compensation or at the receiver as post-compensation [83]. That is enabled by the introduction of coherent detection and advances in digital signal processing technology. Digital back-propagation (DBP) is a widely investigated nonlinearity compensation (NLC) technique to combat the detrimental effects of fiber nonlinearity [5]. DBP can compensate both dispersion and deterministic intra-channel nonlinearity based on a numerical solution of the nonlinear Schrödinger equation (NLSE) using the split-step Fourier method (SSFM) [5]. SSFM involves a large number of linear and nonlinear computation steps per fiber span, thereby the practical implementation of DBP is limited [83]. That led to increased interest in research for the NLC techniques based on the simplified versions of the NLSE for which an approximate analytical solution is available. As a result, a Volterra series-based nonlinear equalizer (VNLE) has been proposed in [36], [38] to compensate for fiber nonlinearity. However, when applied to long-haul optical fiber links, the computational complexity of the VNLE may approach that of the DBP technique [83].

In contrast to the VNLE, the first-order regular perturbation (RP) theory-based NLC, referred to as PB-NLC, has been proposed in the literature to compensate for the fiber nonlinearity effects [84]- [88]. The PB-NLC technique exhibits reduced computational complexity in comparison with DBP and VNLE. The first-order RP theory was initially used to model the intra-channel nonlinear distortion effects between short and highly dispersive Gaussian pulses propagating in single-polarization optical fiber links [43]. This technique was later extended to dual-polarization systems and applied as a low-complexity digital NLC in [46]. It is important to note that the first-order

RP theory adopted in the PB-NLC technique has a serious energy divergence problem when the fiber launch power increases [44]. That is due to the inaccuracy of the first-order RP series approximation for the nonlinear phase-shift. To solve this issue, an enhanced RP (ERP) method was proposed in [44], to model the nonlinear signal propagation in the optical fiber. The ERP method employs a change of variable technique to eliminate the average accumulated nonlinear phase, around which the phase of the received signal field swings, before applying the RP method [44]. The ERP-based technique was initially proposed to model the nonlinearity in dispersion-managed systems [44]. Later, in [89], the ERP approach was adopted for an alternative framework to derive the well-known Gaussian noise reference formula in time-domain to model the nonlinear signal propagation in dispersion unmanaged systems. The ERP method improves the accuracy of the first-order RP solution at the power levels of interest in dispersion unmanaged long-haul transmission systems.

In this chapter, we propose to use an ERP-based method to compensate for the intra-channel nonlinearity, referred to as the ERP-NLC technique. We also introduce a technique, which is a variation of the ERP-NLC, by simple phase-rotation (PR) of the nonlinear coefficient matrix of the PB-NLC technique, referred to as the PR-PB-NLC.

The main contributions of this chapter are as follows:

- We provide a generalized description to show that the ERP technique can solve the energy divergence problem of the RP-based approach in a dispersion unmanaged transmission system.

- We derive the first-order ERP distortion field in time-domain with Gaussian shape assumption for the input pulses.

- We develop the nonlinear coefficient matrix of the PR-PB-NLC technique by

considering only a part of the first-order ERP distortion field.

We carried out numerical simulations for a single or five-channel polarization division multiplexed 16-quadrature-amplitude-modulation (QAM) optical transmission system. The results show that the proposed ERP-NLC technique provides significant performance improvement in terms of the Q-factor and transmission reach, with only a negligible increase in the computational complexity when compared to the PR-PB-NLC, PB-NLC, and electronic dispersion compensation (EDC) techniques. Furthermore, we show that the performance improvement of the PR-PB-NLC technique also comes with a negligible additional computational complexity when compared to the PB-NLC technique.

5.3 Principle of ERP-based NLC

5.3.1 The ERP-based Approximation of the NLSE

In this section, we provide a generalized description to show the effectiveness of the ERP method in solving the energy divergence problem of the RP-based approximation for a dispersion unmanaged transmission system. It is important to note that the demonstration with a zero-dispersion fiber given in [44] can be considered as a special case of our generalized description. The NLSE that describes the evolution of the optical field envelope through an optical fiber is represented as [32]:

$$\frac{\partial}{\partial z}q(z,t') + \frac{\alpha}{2}q(z,t') + j\frac{\beta_2}{2}\frac{\partial^2}{\partial t'^2}q(z,t') = j\gamma\left|q(z,t')\right|^2 q(z,t'), \tag{5.1}$$

where $q(z,t')$ is the optical field, t' is the time variable, z is the transmission distance, α is the attenuation, β_2 is the group velocity dispersion, and γ is the nonlinearity coefficient.

The NLSE in (5.1) can be further simplified by applying the transformation $q(z,t') \triangleq u(z,t)\exp(-\frac{\alpha}{2}z)$, referred to the delayed time frame $t = t' - (z/v_g)$ corresponding to the group velocity v_g, and separating the linear and nonlinear parts as [44], [96]:

$$\frac{\partial}{\partial z}u(z,t) = \left(\hat{D} + \hat{N}\right)u(z,t) \tag{5.2}$$

$$\hat{D} = -j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2} \tag{5.3}$$

$$\hat{N} = j\gamma\left|u(z,t)\right|^2\exp(-\alpha z), \tag{5.4}$$

where $\hat{D}$ and $\hat{N}$ are the linear and nonlinear operators [96]. The simplified NLSE in (5.2) can be numerically solved using the symmetric SSFM as given in [96]. For a special case of $z = z'$ (i.e., the first computation step) the symmetric SSFM yields the solution:

$$u^{SSFM}(z',t) = \exp\left(\frac{z'}{2}\hat{D}\right)\exp\left(\int_0^{z'}\hat{N}(z)dz\right)\exp\left(\frac{z'}{2}\hat{D}\right)u(0,t), \tag{5.5}$$

where z' is the step size. It is important to note that the SSFM (in the absence of fiber loss) alternates between a linear unitary transformation and a time-localized nonlinear signal-dependent phase-rotation, and hence it is energy-preserving [90].

Alternatively, (5.2) can be analytically solved using the first-order RP method [44]. The RP-based approach is an iterative method which provides a closed-form approximate solution of the NLSE. The first-order RP approximation to the optical field after a transmission distance $z = z'$ (the step size in the SSFM) is given as:

$$u^{RP}(z',t) = u_0(z=z',t) + j\gamma\int_0^{z'}\exp(-\alpha z)\left(h_z(t)\otimes\left[\left|u_0(z,t)\right|^2u_0(z,t)\right]\right)dz, \tag{5.6}$$

where $u_0(z,t) = [h_z(t) \otimes u(0,t)]$, is the linear (zeroth-order) solution, with $\otimes$ as the convolution operation, $h_z(t) = \mathcal{F}^{-1}\{\exp(-j\frac{w^2\beta_2 z}{2})\}$ at the angular frequency w, and $\mathcal{F}^{-1}\{.\}$ as the inverse Fourier transform (IFT) operation.

By closely inspecting (5.5), it can be seen that, in the absence of fiber loss, the signal energy at a transmission distance $z = z'$ is $\int_{-\infty}^{\infty} |u(0,t)|^2 \, dt$, which is the same as the input energy at $z = 0$. On the other hand, the first-order RP series approximation in (5.6) overestimates the signal energy at $z = z'$, and the relative error grows with increasing the launch power. To mitigate this discrepancy, the ERP method was proposed in [44]. In the ERP method, a change of variable is applied in (5.2) to eliminate the accumulated nonlinear phase before applying the first-order RP method. The first step is to postulate the solution of (5.2) as [44]:

$$u(z,t) \triangleq \tilde{u}(z,t) \exp(-j\gamma P_0 L_{\text{eff}}), \tag{5.7}$$

where $L_{\text{eff}} \triangleq \int_0^z \exp(-\frac{\alpha}{2}\xi) d\xi$ is the fiber effective length and P_0 is the peak input power. Substituting (5.7) in (5.2) factors out the accumulated nonlinear phase from the solution. As a result, (5.2) with the field $\tilde{u}(z,t)$ and substituting the expressions for $\hat{D}$ and $\hat{N}$, we obtain [44]:

$$\frac{\partial}{\partial z}\tilde{u}(z,t) = -j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}\tilde{u}(z,t) + j\gamma[|\tilde{u}(z,t)|^2 - P_0]\tilde{u}(z,t). \tag{5.8}$$

The next step is to solve (5.8) using the first-order RP method. Accordingly, from (5.7) and (5.8), the zeroth-order solution of the optical field at a transmission distance $z = z'$ is obtained as:

$$u_0^{ERP}(z',t) = \tilde{u}_0(z = z', t) \exp(-j\gamma P_0 L_{\text{eff}}), \tag{5.9}$$

where $\tilde{u}_0(z,t) = [h_z(t) \otimes \tilde{u}(0,t)]$. Similarly, the first-order ERP solution can be represented as:

$$u_1^{ERP}(z',t) = j\gamma \int_0^{z'} \exp(-\alpha z) \left(h_z(t) \otimes \left[\left(|\tilde{u}_0(z,t)|^2 - P_0 \right) \tilde{u}_0(z,t) \right] \right) dz \times \exp(-j\gamma P_0 L_{\text{eff}}). \quad (5.10)$$

Combining (5.9) and (5.10), the analytical approximation to the optical field at a transmission distance $z = z'$ based on the first-order ERP series can be represented as:

$$u^{ERP}(z',t) \approx \tilde{u}_0(z',t) \exp(-j\gamma P_0 L_{\text{eff}}) + j\gamma \int_0^{z'} \exp(-\alpha z) \times \left(h_z(t) \otimes \left[\left(|\tilde{u}_0(z,t)|^2 - P_0 \right) \tilde{u}_0(z,t) \right] \right) dz \exp(-j\gamma P_0 L_{\text{eff}}). \quad (5.11)$$

From (5.11), it can be seen that the signal energy at a transmission distance $z = z'$ is close to $\int_{-\infty}^{\infty} |u(0,t)|^2 \, dt$, especially when the input field magnitude approaches its peak value $\sqrt{P_0}$.

5.3.2 The ERP-NLC Technique

The time-domain nonlinear distortion field based on the first-order ERP method is obtained by solving (5.10) with the Gaussian pulse shape assumption for the input pulses. It is important to mention that the Gaussian shape assumption for the input pulse shape is adopted for the simplicity of analysis. This assumption allows the calculation of the FO nonlinearity coefficients using analytic closed-form expression involving the exponential integral function. Following the analysis given in Appendix B.1, the time-domain first-order distortion field at a transmission distance $z = L$ can

be represented as:

$$u_1^{ERP}(L,t+kT) = j\gamma P^{3/2}\exp(-j\gamma P_0 L_{\text{eff}})\sum_m\sum_n\sum_l a_m a_n a_l^* \times \exp(-\frac{t^2}{6\tau^2})\int_0^L \frac{\exp(-\alpha z)}{\sqrt{1+2j\beta_2 z/\tau^2+3(\beta_2 z/\tau^2)^2}} \underbrace{\times \exp\left\{-\frac{3\left[\frac{2}{3}t+(m-l)T\right]\left[\frac{2}{3}t+(n-l)T\right]}{\tau^2\left(1+3j\beta_2 z/\tau^2\right)} - \frac{(n-m)^2T^2}{\tau^2\left[1+2j\beta_2 z/\tau^2+3(\beta_2 z/\tau^2)^2\right]}\right\}dz}_{\text{Term 1}} \underbrace{- j\gamma P_0 a_k \exp(-\frac{t^2}{2\tau^2})\int_0^L \exp(-\alpha z)\exp(-j\gamma P_0 L_{\text{eff}})dz,}_{\text{Term 2}} \quad (5.12)$$

where $k = m+n-l$, m, n, l are the symbol indices, P is the launch power, τ is the pulse width, and T is the symbol interval.

Equation (5.12) calculates the time-domain first-order ERP distortion field at $k = m+n-l$ caused by the nonlinear interaction between three pulses located at the time indices m, n, and l. Since the ERP technique is a modification to the RP method, we followed a similar mathematical analysis in [46] to derive the nonlinear distortion field. In the first term of (5.12), we obtained a modified expression with a time-invariant phase-rotation term $\exp(-j\gamma P_0 L_{\text{eff}})$ when compared to the PB-NLC technique. On the other hand, the second term of (5.12) is independently obtained in our analysis and is proportional to the complex amplitude of the symbol at time index k.

The basic idea of the ERP-NLC pre-compensation technique is to calculate the nonlinear distortion field using (5.12), and then to subtract it from the transmitted field to generate the pre-distorted waveform. In general, the integrals in (5.12) cannot be solved analytically due to the presence of the attenuation term. Therefore, we adopt the conventional RP method by ignoring $\exp(-\alpha z)$, as given in [43], to obtain the

closed-form solution. Without loss of generality, in the implementation, we focus on the perturbation of the symbol at index $k = 0$, i.e., $l = m+n$. That will simplify (5.12) by replacing the triple summation with a double summation. It is worth mentioning that the nonlinear distortion field calculation at any other index, for example $k = m+n-l$, using (5.12) is the same as the calculation at $k = 0$. The pre-compensation is assumed to operate at the symbol rate; therefore, only the perturbation value at $t = 0$ is calculated. In a typical dispersion unmanaged optical transmission system, the chromatic dispersion-induced pulse spreading is usually much larger than the symbol duration, i.e., $\beta_2 z \gg \tau^2$ [46]. With the large chromatic dispersion assumption and following a similar procedure as in [46], the nonlinear distortion field for the zeroth symbol at $t = 0$ can be obtained as:

$$u_1^{ERP}(L,t) = j\gamma \exp(-j\gamma P_0 L_{\text{eff}}) \times \left[\frac{\tau^2}{\sqrt{3}\,|\beta_2|} P^{3/2} \sum_m \sum_n a_m a_{m+n}^* a_n \mathbf{C}_{m,n}^{FO} - P_0 a_0 L_{\text{eff}} \right]. \quad (5.13)$$

In (5.13), $\mathbf{C}_{m,n}^{FO}$ is the first-order perturbation coefficient matrix, which is represented as:

$$\mathbf{C}_{m,n}^{FO} = \begin{cases} \int_0^L dz \frac{1}{\sqrt{\tau^4/(3\beta_2^2)+z^2}}, & m = n = 0 \\ \frac{1}{2} E_1\left(\frac{(n-m)^2 T^2 \tau^2}{3|\beta_2|^2 L^2}\right), & m \text{ or } n = 0 \text{ and } m \neq n \\ E_1\left(-j\frac{mnT^2}{\beta_2 L}\right), & m \neq n \neq 0, \end{cases} \quad (5.14)$$

where $E_1(x) = \int_x^\infty \frac{e^{-t}}{t} dt$ is the exponential integral function.

The nonlinear distortion field in (5.13) can be extended to dual-polarization using the Manakov equation for the nonlinear signal propagation, as shown in Appendix B.2. Using (B.13) one can show that the six Gaussian input pulses $\sqrt{P} a_{m/l/n,x/y} \exp(-(t - T_{m/l/n})^2/2\tau^2)$ at three time instants T_m, T_l, T_n for the two polarizations generate the

nonlinear distortion field for the zeroth symbol, i.e., $l = m + n$, as:

$$u_{1,x/y}^{ERP}(L,t) = j\frac{8}{9}\gamma \exp(-j\gamma P_0 L_{\text{eff}}) \left[\frac{\tau^2}{\sqrt{3}\,|\beta_2|} P^{3/2} \sum_m \sum_n (a_{n,x/y} a_{m+n,x/y}^* a_{m,x/y} + a_{n,y/x} a_{m+n,y/x}^* a_{m,x/y}) \mathbf{C}_{m,n}^{FO} - P_0 a_{0,x/y} L_{\text{eff}} \right], \quad (5.15)$$

where $a_{m/(m+n)/n,x/y}$ and $a_{0,x/y}$ are the symbol complex amplitudes. It is important to note that the peak power P_0 is selected as $\frac{3}{2}P$ in the implementation of the ERP-NLC technique, as per the analysis given in [89].

The first-order ERP-based nonlinear distortion field in (5.15) consists of a time-invariant phase rotation term $\exp(-j\gamma P_0 L_{\text{eff}})$ and a time-variant term proportional to the complex amplitude of the symbol at index 0, when compared to the RP-based distortion field in [46]. It is noteworthy that the perturbation coefficient matrix $\mathbf{C}_{m,n}^{FO}$ and the phase rotation term $\exp(-j\gamma P_0 L_{\text{eff}})$ are calculated offline and stored in look-up tables (LUTs). As a result, the performance improvement of the proposed ERP-NLC technique comes with a negligible additional complexity when compared to the PB-NLC technique.

Besides the ERP-NLC technique, we consider a method by simple PR of the perturbation coefficient matrix of the PB-NLC technique by using only the first part of (5.15). We refer to this technique as PR-PB-NLC. It is worth mentioning that the PR-PB-NLC is similar to an intra-channel NLC technique proposed in [103]. The method given in [103] modifies the perturbation coefficient matrix by multiplying with a PR term similar to the PR-PB-NLC technique. It is important to note that the PR method in [103] selects the optimum phase by sweeping the phase angle in the range 0 to 1 rad. On the other hand, the PR-PB-NLC technique calculates the rotation phase angle, which is proportional to the fiber effective length, as $\gamma P_0 L_{\text{eff}}$. The PR-PB-NLC can be considered as a variation of the ERP-NLC technique by using

only the first part of (5.15). It should be noted that the additional computational complexity of the PR-PB-NLC is negligibly small when compared to the PB-NLC technique. We have included the result for the PR-PB-NLC along with the ERP-NLC technique to compare the performance. We evaluate the implementation complexity of the proposed ERP-NLC and PR-PB-NLC techniques in terms of the number of real-valued multiplications per symbol in Section 5.5.

Fig. 5.1 shows the magnitude of $\mathbf{C}_{m,n}^{FO}$ relative to the largest coefficient $C_{0,0}^{FO}$, at a transmission distance of 1200 km. The nonlinear distortion calculation in (5.15) contains infinite terms when m and n approach infinity. In the implementation, we truncate them when the perturbation coefficient $\mathbf{C}_{m,n}^{FO}$ is less than a threshold value given as $20\log_{10}\left(\left|\mathbf{C}_{m,n}^{FO}\right| / \left|C_{0,0}^{FO}\right|\right) < -40\,\text{dB}$ [46].

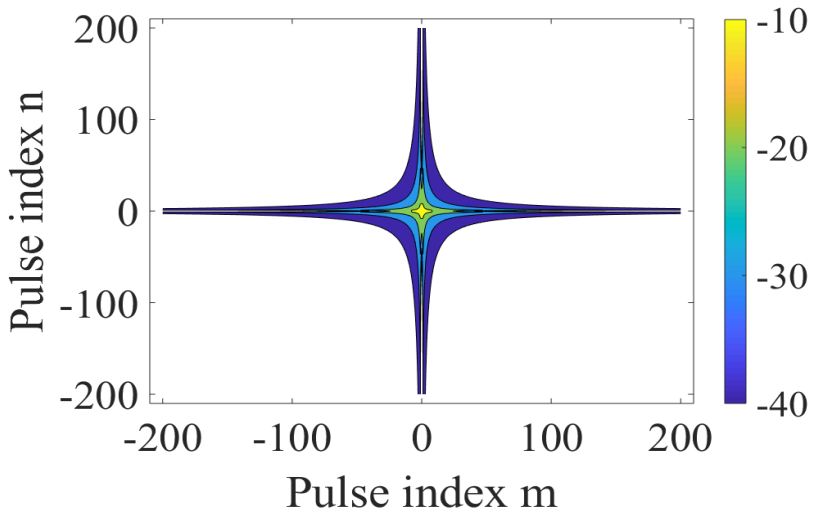


Figure 5.1: The magnitude of $C_{m,n}$ relative to the largest coefficient $C_{0,0}$, at a transmission distance of 1200 km.

Fig. 5.2 shows the block diagram of the ERP-NLC technique. In the pre-compensation technique, the perturbative nonlinear distortion caused by the intra-channel nonlinearity is calculated, first using (5.15) and then subtracted from the transmitted field, as shown in Fig. 5.2.

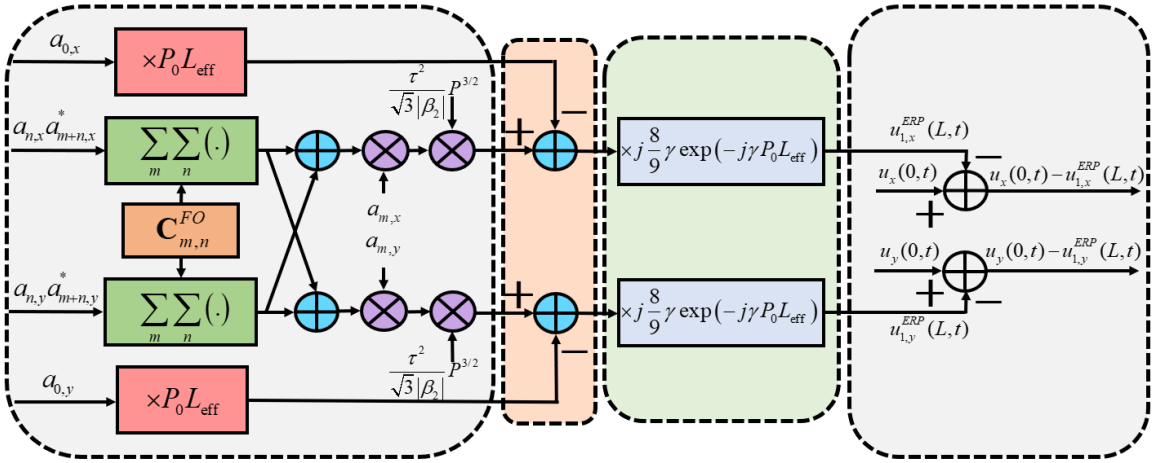


Figure 5.2: The block diagram of the ERP-NLC technique.

5.4 Simulation Setup

Fig. 5.3 shows the simulation setup for the ERP-NLC technique. At the transmitter, after the 16-QAM symbol mapping, the first-order ERP-NLC is carried out as a pre-compensation at one sample/symbol. Then, a root-raised-cosine filter with a roll-off factor 0.1 is applied in each polarization for the pulse shaping. The data transmission rate is 32 Gbaud. After digital-to-analog conversion and low-pass filtering, the pre-compensated signal is converted to the optical domain using an in-phase/quadrature-phase modulator.

The long-haul transmission link consists of several spans of standard single-mode fiber with the span length of 80 km, the attenuation coefficient of 0.2 dB.km^{-1}, the nonlinear parameter of 1.22 W^{-1}.km^{-1}, the dispersion parameter of 16 ps.nm^{-1}.km^{-1}, and the polarization mode dispersion coefficient of 0.1 ps.km$^{-1/2}$. The optical power loss in each fiber span is compensated by an erbium doped fiber amplifier with 16 dB gain and 5.5 dB noise figure. At the receiver, the signal is coherently detected using a polarization diversity detector. After analog-to-digital conversion and root-raised-

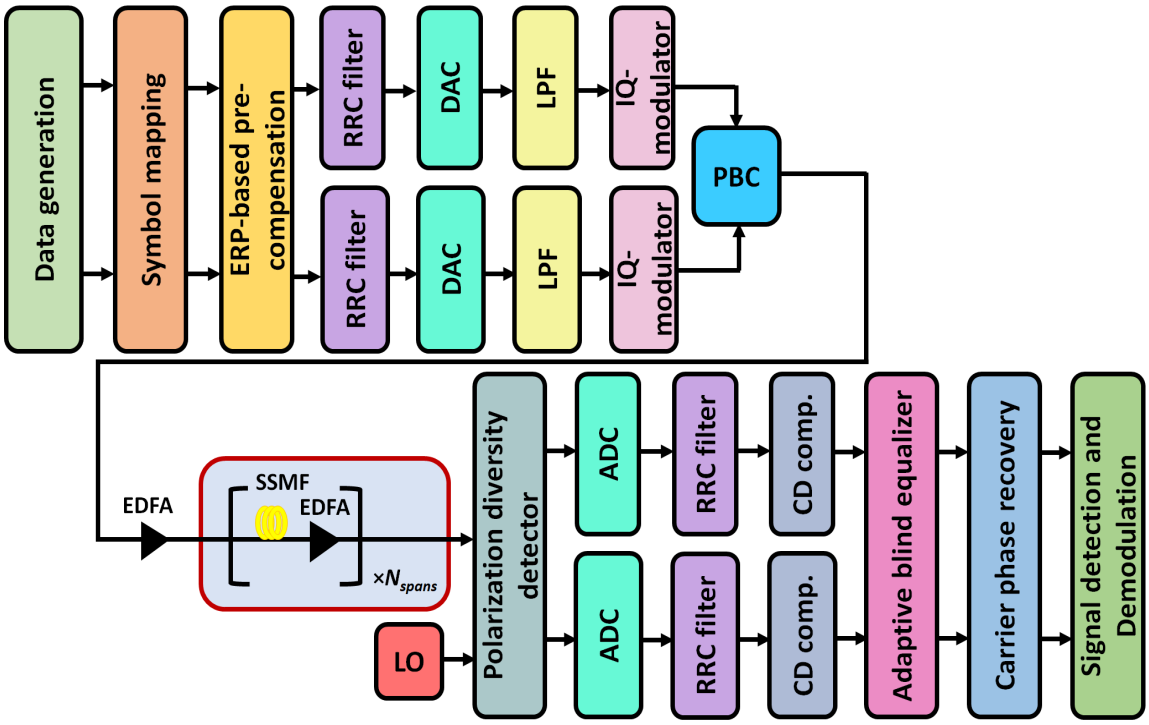


Figure 5.3: Simulation setup for ERP-NLC technique (single-channel). RRC: root-raised-cosine, DAC: digital-to-analog converter, LPF: low pass filter, IQ: in-phase/quadrature phase, PBC: polarization beam combiner, N_{spans}: number of spans, EDFA: erbium doped fiber amplifier, ADC: analog-to-digital converter, CD: chromatic dispersion.

cosine filtering, dispersion compensation is performed. Then, an adaptive equalization is carried out for the state-of-polarization recovery. After that, the carrier phase is recovered using the Viterbi-Viterbi algorithm. Finally, symbol detection and demodulation is applied to recover the transmitted information bits. We adopt the Q-factor to evaluate the system performance which is directly derived from the bit-error rate, as given in [104].

5.5 Simulation Results

We carried out numerical simulation for the single-/five-channel polarization-division multiplexed 16-QAM optical transmission system to evaluate the performance of the

proposed ERP-NLC technique.

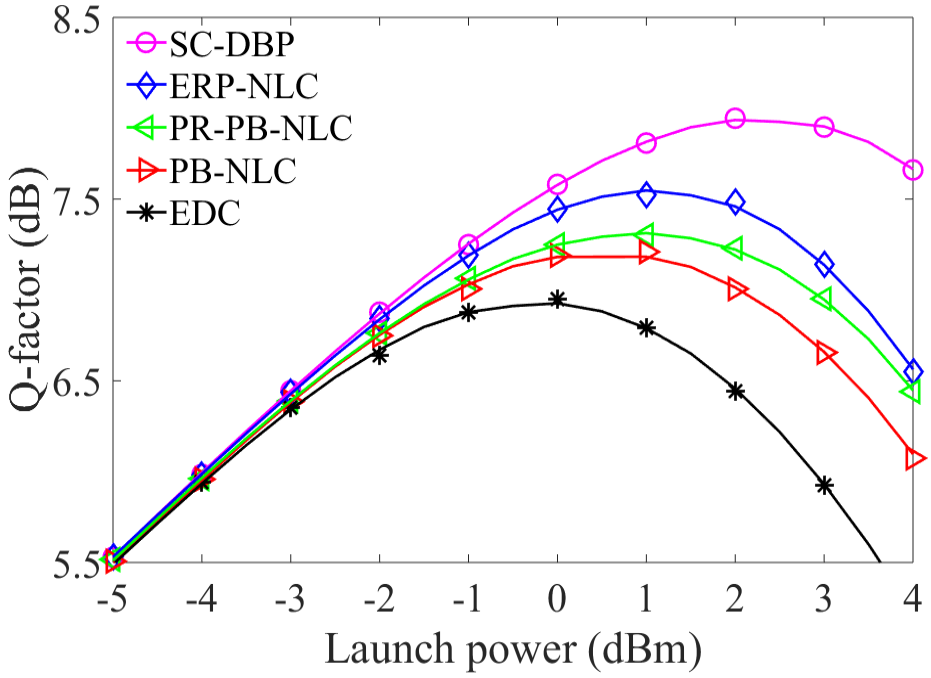


Figure 5.4: Q-factor versus launch power for the SC-DBP, ERP-NLC, PR-PB-NLC, PB-NLC, and EDC techniques after the propagation over 2800 km.

We compare the performance of the ERP-NLC technique with the single-channel (SC)-DBP, PR-PB-NLC, PB-NLC, and EDC techniques. The SC-DBP technique is implemented with one step per span. It is worth mentioning that increasing the number of steps per span increases the compensation performance of the SC-DBP technique with a corresponding increase in the computational complexity. For example, the SC-DBP with 16 steps/span can increase the optimum Q-factor by $\sim$ 1.5 dB for an SC system with a 16 times increase in the computational complexity when compared to the one step/span implementation [91]. Fig. 5.4 shows the Q-factor as function of launch power for a 256 Gb/s SC transmission system. It is evident from Fig. 5.4 that the proposed ERP-NLC technique improves the Q-factor performance by $\sim$0.6 dB and $\sim$0.3 dB when compared to the EDC and the PB-NLC techniques, respectively, at a transmission distance of 2800 km. It is interesting to note that

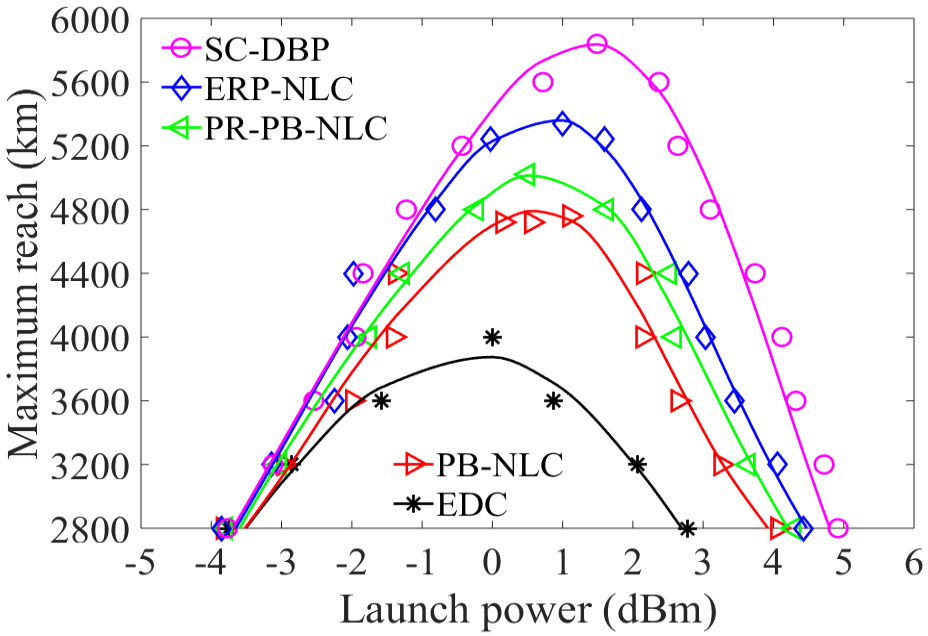


Figure 5.5: Maximum reach as a function of the launch power at 20% OH-SD-FEC limit for the SC-DBP, ERP-NLC, PR-PB-NLC, PB-NLC, and EDC techniques.

the PR-PB-NLC technique improves the Q-factor by ∼0.35 dB and ∼0.1 dB when compared to the EDC and PB-NLC techniques, respectively.

In Fig. 5.5, we plot the maximum reach as a function of the launch power at a 20% overhead (OH) soft-decision (SD) forward error correction (FEC) limit with a BER value of 2.8×10^{-2} [92]. It is observed that the maximum transmission reach for the SC-DBP, ERP-NLC, PR-PB-NLC, PB-NLC, and EDC is 5840 km, 5340 km, 5020 km, 4760 km, and 4000 km, respectively. These results indicate that the ERP-NLC technique provides an extended transmission reach by 33.5% and 12.2% when compared to the EDC and the PB-NLC techniques, respectively. Besides the ERP-NLC technique, the PR-PB-NLC yields an extended transmission reach of 5.5% when compared to the PB-NLC technique. That is attributed to the fact that the fiber has attenuation, in reality, and so, the optimum perturbation coefficient should be different from $C_{m,n}$ of the PB-NLC technique [46]. The PR of the perturbation coefficient matrix in (2.14) partially solves this problem through the parameter L_{eff}

in the calculated rotation phase angle of the PR-PB-NLC technique. Notably, the ERP-NLC technique improves the transmission reach by 6.4% when compared to the PR-PB-NLC technique.

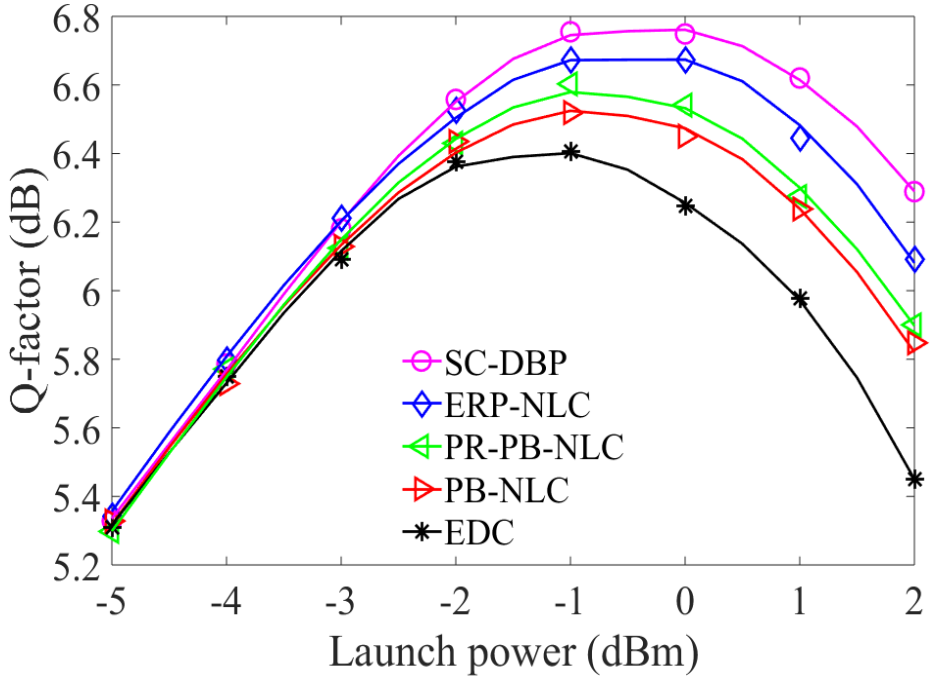


Figure 5.6: Q-factor versus launch power for the central WDM channel of the SC-DBP, ERP-NLC, PR-PB-NLC, PB-NLC, and EDC techniques after the propagation over 2800 km.

The performance of a 1.28 Tb/s five-channel WDM transmission system is shown in Fig. 5.6 and Fig. 5.7. The channel spacing is 37.5 GHz. The central WDM channel is arbitrarily selected for performance evaluation. In Fig. 5.6, results show that the proposed ERP-NLC technique improves the Q-factor performance by $\sim$0.3 dB and $\sim$0.2 dB when compared to the EDC and PB-NLC techniques, respectively, at a transmission distance of 2800 km. On the other hand, the PR-PB-NLC technique shows a Q-factor improvement of $\sim$0.2 dB and $\sim$0.08 dB when compared to the EDC and PB-NLC techniques, respectively.

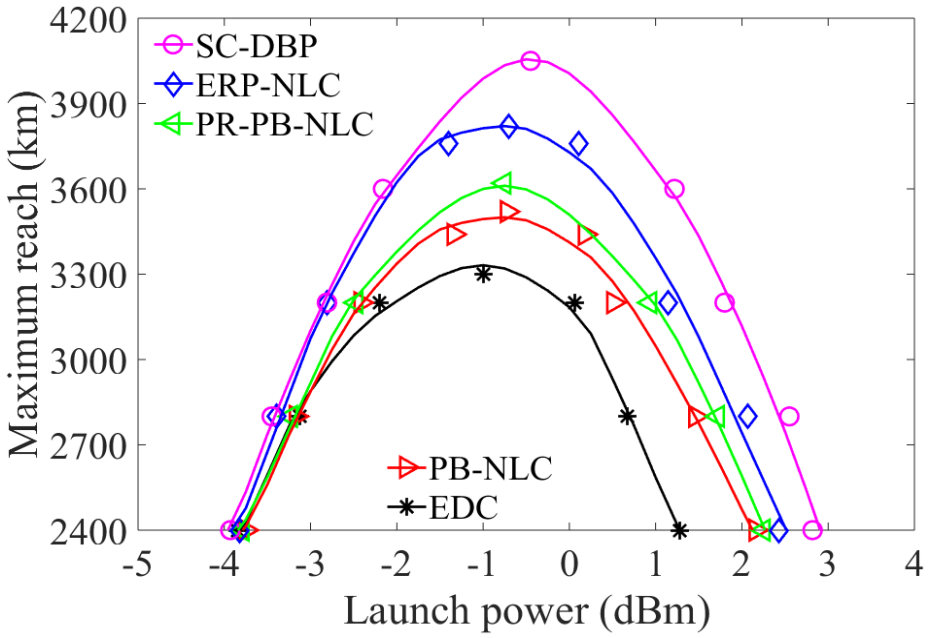


Figure 5.7: Maximum reach as a function of the launch power at 20% OH-SD-FEC limit for the central channel in a five-channel WDM system.

Results given in Fig. 5.7 indicate that the maximum reach at 20% OH-SD-FEC limit for the SC-DBP, ERP-NLC, PR-PB-NLC, PB-NLC, and EDC techniques is 4050 km, 3820 km, 3620 km, 3520 km, and 3300 km respectively. Accordingly, the proposed ERP-NLC technique provides an extended transmission reach of 16%, 8.5%, and 5.5% when compared to the EDC, PB-NLC, and PR-PB-NLC techniques, respectively. It is important to note that the PR-PB-NLC technique improves the transmission reach by only 2.8% when compared to the PB-NLC technique. It is observed that the Q-factor improvement for the WDM system is less when compared to the SC system. That can be easily explained, as the inter-channel nonlinear distortions, such as cross-phase modulation and cross-polarization modulation, are the dominant impairments in a dispersion unmanaged WDM system, which cannot be compensated for by the intra-channel NLC techniques [83]. Further performance improvement can be achieved by including the inter-channel effects in the ERP-NLC technique. For WDM systems with many channels, the strong walk-off and phase-mismatch between the widely

separated channels reduce the nonlinearity effects on transmission beyond an effective bandwidth [91]. In such cases, a mean-field approach can be used for NLC, which neglects the time and z-variations of the channels outside an effective bandwidth [91]. The effective bandwidth is chosen as a trade-off between implementation complexity and compensation performance [91].

5.6 Complexity Evaluation

In this section, the computational complexity evaluation is performed for the SC-DBP, ERP-NLC, PR-PB-NLC, PB-NLC, and EDC techniques; the real-valued multiplications per symbol is considered as the performance metric. SC-DBP is implemented with one step per span. The nonlinearity coefficient matrix of the ERP/RP-based NLC techniques is truncated at a threshold of -40 dB. Number of real-valued multiplications per symbol for the SC-DBP technique is given by $2(4N_{spans}N_{\mathrm{FFT}}\log_2(N_{\mathrm{FFT}}) + 10.5N_{spans}N_{\mathrm{FFT}})$, where N_{spans} is the number of fiber spans and N_{FFT} is the fast Fourier transform size [5]. For the PB-NLC technique, the expression for the number of real-valued multiplications per symbol is given as $2(20M + 3)$, where M is the number of significant perturbation coefficients in $\mathbf{C}_{m,n}^{FO}$ [73]. In the PR-PB-NLC technique, the phase rotation term $\exp(-j\gamma P_0 L_{\mathrm{eff}})$ is calculated offline and stored in a LUT. Accordingly, the PR-PB-NLC technique has only one additional complex-valued multiplication per symbol when compared to the PB-NLC technique. Therefore, the number of real-valued multiplications per symbol for the PR-PB-NLC technique is given as $2(20M + 7)$. The proposed ERP-NLC technique consists of an additional time-invariant phase rotation term $\exp(-j\gamma P_0 L_{\mathrm{eff}})$ and a time-variant term $P_0 a_{0,x/y} L_{\mathrm{eff}}$ when compared to the PB-NLC technique. It is important to note that the phase rotation term $\exp(-j\gamma P_0 L_{\mathrm{eff}})$ is calculated offline and stored in a LUT, as

in the PR-PB-NLC technique. That yields only one complex-valued multiplication per symbol in the implementation of the ERP-NLC technique. Similarly, the term $P_0 a_{0,x/y} L_{\text{eff}}$ contributes two real-valued multiplications and one complex-valued multiplication per symbol to the computational complexity of the ERP-NLC technique. As a result, the number of real-valued multiplications per symbol for the ERP-NLC technique is given as $2(20M + 9)$. For the EDC technique, the number of real-valued multiplications per symbol is given as $2(4N_{\text{FFT}} \log_2(N_{\text{FFT}}) + 4N_{\text{FFT}})/N_s$ [5]. It is important to note that the factor 2 in the complexity expressions accounts for the dual-polarization transmission.

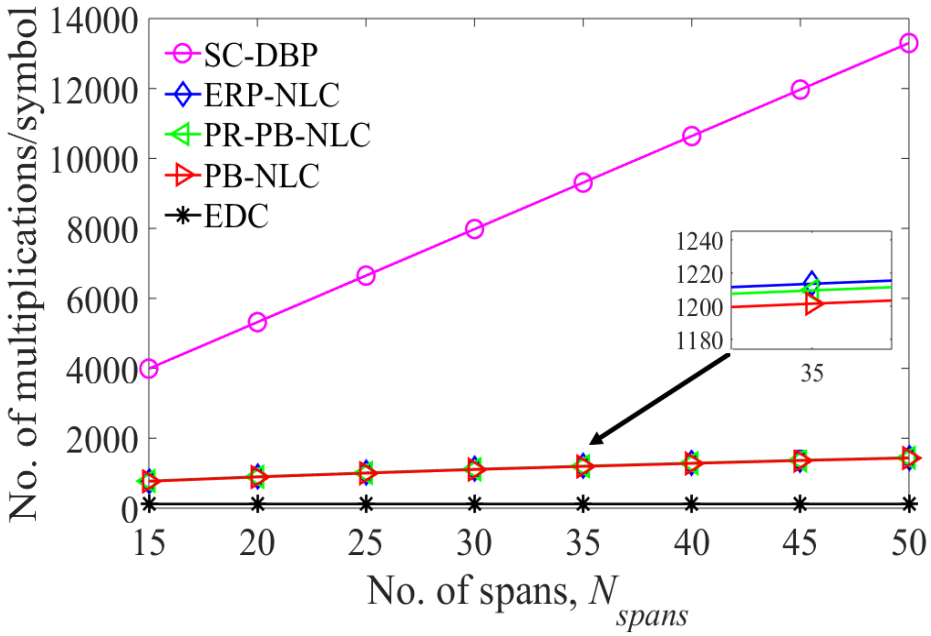


Figure 5.8: The number of real-valued multiplications/symbol for the SC-DBP, ERP-NLC, PR-PB-NLC, PB-NLC, and EDC techniques as a function of the number of spans.

Fig. 5.8 shows the number of real-valued multiplications per symbol as a function of the number of fiber spans, N_{spans} for the considered techniques. Results show that the complexity of the SC-DBP technique increases rapidly as the number of fiber span increases. On the other hand, for the ERP-NLC, PR-PB-NLC, and PB-NLC tech-

niques, the complexity increases only slightly as the number of fiber spans increases. That is due to a slight increase in the number of coefficients in the nonlinear coefficient matrix $\mathbf{C}_{m,n}^{FO}$, satisfying the truncation threshold, as the number of fiber spans increases. It is interesting to note that the additional computational complexity of the ERP-NLC and PR-PB-NLC techniques, when compared to the PB-NLC technique, is negligible.

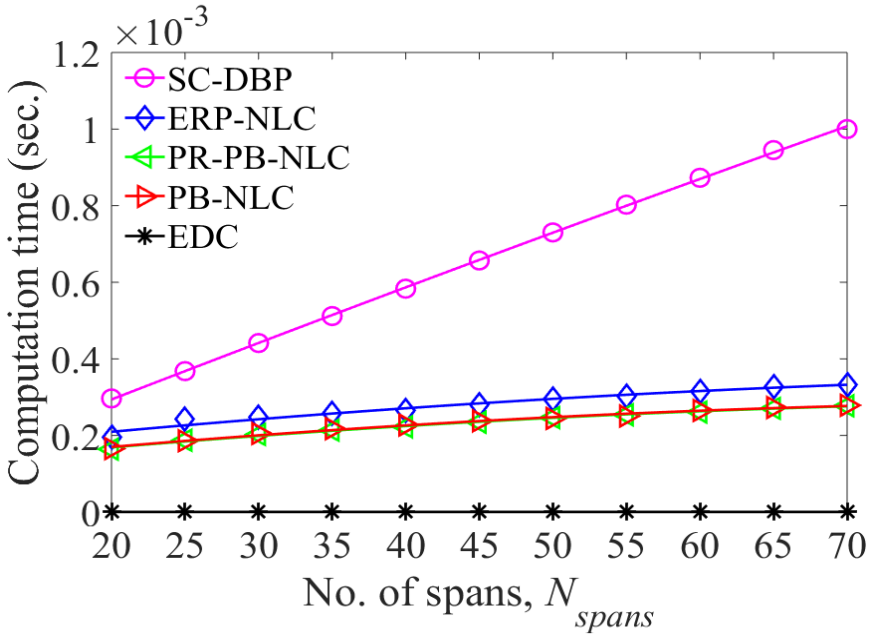


Figure 5.9: The CPU running time for the SC-DBP, ERP-NLC, PR-PB-NLC, PB-NLC, and EDC techniques as a function of the number of spans.

Fig. 5.9 shows the central processing unit (CPU) running time for the SC-DBP, ERP-NLC, PR-PB-NLC, PB-NLC, and EDC techniques as a function of the number of spans. The results indicate that the computation time of the proposed ERP-NLC technique is only slightly higher than that of the PB-NLC technique and significantly lower than that of the SC-DBP technique.

5.7 Conclusion

In this chapter, we have proposed an ERP-based NLC technique, referred to as ERP-NLC. We have shown through numerical simulations that this technique extends the transmission reach by 33.5%, 12.2%, and 6.4% when compared to the EDC, PB-NLC, and PR-PB-NLC techniques, respectively, for a 256 Gb/s single-channel transmission system. We have also demonstrated that, for a 1.28 Tb/s five-channel WDM transmission system, ERP-NLC improves the transmission reach by 16%, 8.5%, and 5.5% when compared to the EDC, PB-NLC, and PR-PB-NLC techniques, respectively. The complexity evaluation using the number of real-valued multiplications per symbol indicates that the additional complexity of the proposed ERP-NLC technique is negligibly small when compared to the PB-NLC technique.

Chapter 6

Second-Order Perturbation Theory-Based Digital Predistortion for Fiber Nonlinearity Compensation

6.1 Preamble

This chapter is compiled from the materials extracted from the preprint titled *"Intra-channel nonlinearity compensation based on second-order perturbation theory," arXiv: 2005.01191v1 [cs.IT], May 2020* [105].

6.2 Introduction

In recent years, the increased usage of bandwidth-intensive applications such as virtual reality and cloud services, as well as Internet-of-Things dramatically increased

the network traffic in the core communication network [2]- [4]. That necessitates the development of high data-rate optical communication systems to handle such traffic surges. The modern high data-rate optical transmission systems use multilevel modulation formats, which require a higher optical signal-to-noise ratio (OSNR). However, the optical intensity-dependent nonlinear Kerr effect significantly degrades the transmission performance as the fiber launch power increases [32], [93]. In a dispersion unmanaged optical communication system, the signal-to-signal intra-channel Kerr nonlinearity is considered a dominant impairment, which limits the transmission performance [93]. It is worth mentioning that the signal-to-signal intra-channel nonlinearity can be compensated in principle due to its deterministic nature [94], [95].

It was shown a few years ago that digital compensation of the intra-channel fiber nonlinearity impairment could be achieved using coherent detection and digital signal processing [5], [58]. Digital back-propagation (DBP) is an extensively investigated fiber nonlinearity compensation (NLC) technique, which uses the numerical solution of the nonlinear Schrödinger equation (NLSE) [96]- [97]. However, the implementation complexity of the DBP technique is impractically high when applied to the long-haul optical transmission links [98]- [102]. Alternatively, the NLSE can be analytically solved using the perturbation theory [44]. In this approach, the solution of NLSE can be expanded as an infinite power series of the fiber nonlinearity coefficient [44]. Such an iterative method provides a closed-form approximation of the nonlinearly distorted signal field, which imparts a good insight into the nature of the interaction between CD and Kerr nonlinearity [44]. The first-order (FO) perturbation theory (the perturbation series approximation truncated to FO) was initially used to model the intra-channel nonlinearity distortion between highly dispersive and ultra-short Gaussian pulses propagating in the optical fiber link [43]. These results were later

adopted in [46] to design an FO perturbation theory-based NLC* (FO-PB-NLC) to deal with the detrimental effects of fiber nonlinearity. The main advantage of the perturbation theory-based approach is the possibility of a single-stage implementation for the entire fiber link [46], [87]. It also facilitates one sample per symbol processing, which relaxes the stringent requirement on the electronic hardware speed [46], [87].

On the other hand, the compensation performance of the FO-PB-NLC technique decreases as the launch power increases [47]. This is attributed to the fact that the FO perturbation series approximation becomes inaccurate to model the nonlinear phase shift as the launch power increases [47]. The use of higher-order modulation formats in the high data-rate optical communication system increases the transmit launch power, and thereby, the higher-order perturbation terms become significant in such transmission systems [47].

In this chapter, we propose the extension of the FO-PB-NLC technique to the second-order (SO), referred to as the SO-PB-NLC, to improve the NLC performance. The main contributions of this chapter are summarized as follows:

- We present a rigorous mathematical analysis to derive the expression for the SO nonlinear distortion field in time-domain with a Gaussian shape assumption for the input pulse shape.

- We investigate simplifying assumptions to make the expression for the SO nonlinear distortion field less complex.

- We design and implement a digital predistorter using the simplified SO nonlinear distortion field to compensate for the fiber nonlinearity.

- We carry out a complexity analysis and show that the implementation com-

*In the perturbation theory-based approach, NLC is often referred to as either predistortion or post-compensation method. In our work on the perturbation theory-based technique, NLC refers to the predistortion method to compensate for the fiber nonlinearity.

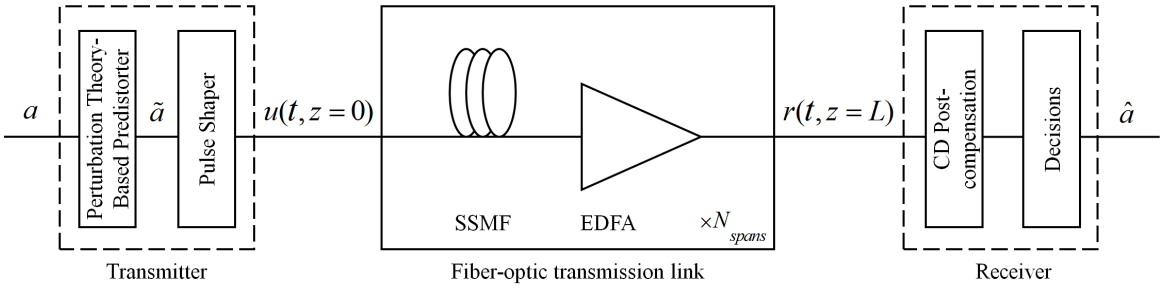


Figure 6.1: System model comprising a transmitter with a perturbation theory-based predistorter and a pulse shaper, fiber-optic transmission link with N_{spans} spans of SSMF and EDFA, and a receiver with a CD post-compensator followed by a decision unit. SSMF: standard single-mode fiber, EDFA: erbium-doped fiber amplifier, CD: chromatic dispersion.

plexity of the digital predistorter based on the simplified SO distortion field is significantly less when compared to the DBP technique.

- We show that the SO-PB-NLC technique provides an extended transmission reach by 14% over the FO-PB-NLC technique, and is only a bit lower in performance than DBP which has a high implementation complexity.

6.3 System Model

6.3.1 High-level Description

The system model, shown in Fig. 6.1, comprises a perturbation theory-based predistorter and a pulse shaper at the transmitter, a fiber-optic transmission link with N_{spans} spans of standard single-mode fiber (SSMF), and the receiver consisting of a CD post-compensator followed by a decision unit. In each fiber span, an erbium-doped fiber amplifier (EDFA) is employed for the periodic amplification of the optical signal to compensate for the fiber attenuation.

A sequence of K symbols $\mathbf{a} = [a_1, a_2, ..., a_K] \in \Omega^K$, with Ω as the symbol alphabet,

is predistorted first to compensate for the intra-channel fiber nonlinearity. Then, the predistorted signal $\tilde{a}$ is passed through a pulse shaping filter $g(\acute{t})$, where $\acute{t}$ is the time variable. The resultant signal can be represented as $u(\acute{t}, z = 0) = \sum_{k=1}^{K} a_k g(\acute{t} - kT)$, where z is the space variable and T is the symbol duration. After pulse shaping, the signal is up-converted to the optical domain and transmitted over the fiber-optic transmission link. At the receiver, after down-conversion to the electrical domain, the baseband signal field can be represented as $r(\acute{t}, z = L)$, where L is the transmission length. Then, the accumulated* CD is compensated in the electrical domain. Finally, we employ a symbol-by-symbol maximum likelihood detection to carry out the symbol decisions.

6.3.2 Optical Fiber Channel: Signal Propagation Model

In this subsection, we describe the model of signal propagation in the optical fiber channel. The derivation shown in this subsection is adopted from [96]. In a single-mode optical fiber channel, the propagation of the optical field complex envelop $q(z, \acute{t})$ can be modeled by using the NLSE (noiseless) as:

$$\frac{\partial}{\partial z} q(z, t') + \frac{\alpha}{2} q(z, t') + j \frac{\beta_2}{2} \frac{\partial^2}{\partial t'^2} q(z, t') = j\gamma \left| q(z, t') \right|^2 q(z, t'), \tag{6.1}$$

where α is the attenuation, β_2 is the group velocity dispersion, γ is the nonlinearity coefficient, and z is the transmission distance. The NLSE can be further simplified by introducing a normalized field $u(z, t)$ referred to the delayed time frame $t = t' - (z/v_g)$ corresponding to the group velocity v_g. Thus, by applying the transformation $q(z, t') \triangleq u(z, t) \exp(-\frac{\alpha}{2} z)$, (2.1) can be modified as:

*In a typical dispersion unmanaged optical transmission system, the accumulated CD is compensated in electrical domain using frequency-domain equalizers employing overlap-add/overlap-save algorithm [106].

$$\frac{\partial}{\partial z}u(z,t) + j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}u(z,t) = j\gamma\left|u(z,t)\right|^2 u(z,t)\exp(-\alpha z). \tag{6.2}$$

The propagation model in (6.2) can be rearranged by separating the linear and nonlinear parts as:

$$\frac{\partial}{\partial z}u(z,t) = \left(\hat{D} + \hat{N}\right)u(z,t) \tag{6.3}$$

$$\hat{D} = -j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2} \tag{6.4}$$

$$\hat{N} = j\gamma\left|u(z,t)\right|^2\exp(-\alpha z), \tag{6.5}$$

where $\hat{D}$ and $\hat{N}$ are the linear and nonlinear operators.

6.4 Theory of the Second-order Perturbation-based Predistortion

In this section, we first discuss the theory of the SO perturbative correction to the nonlinear distortion field, which is the foundation of the SO-PB-NLC technique. Then, we illustrate the implementation of the SO-PB-NLC technique in the context of a single-polarization and single-channel transmission system.

The differential equation governing the SO distortion field can be represented as:

$$\frac{\partial}{\partial z}u_2(z,t) = \underbrace{-j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}u_2(z,t)}_{\text{Linear part}} + \underbrace{\left(\begin{array}{c}\underbrace{j2\left|u_0(z,t)\right|^2\tilde{u}_1(z,t)\exp(-\alpha z)}_{\text{Term 1}} \\ \underbrace{+ju_0^2(z,t)\tilde{u}_1^*(z,t)\exp(-\alpha z)}_{\text{Term 2}}\end{array}\right)}_{\text{Nonlinear part}}, \tag{6.6}$$

where $\tilde{u}_1$ is the FO field distorted by CD in the incremental length of z while evolving

along the optical fiber. It is important to mention that the dispersed FO ghost pulse is considered in the calculation of the SO distortion field.

The equation (6.6) represents the evolution of the SO distortion field along the dispersive and nonlinear optical fiber channel. That is similar to the evolution equation of the noise field, as given in (19) of [107]. In (6.6), the linear part causes the dispersion of the SO nonlinear distortion field u_2 when it evolves through the fiber. The nonlinear part has two terms: Term 1 and Term 2. Term 1 represents the intra-channel cross-phase modulation (IXPM) between the zeroth-order and the FO distortion fields, whereas Term 2 is the intra-channel four-wave mixing (IFWM) term.

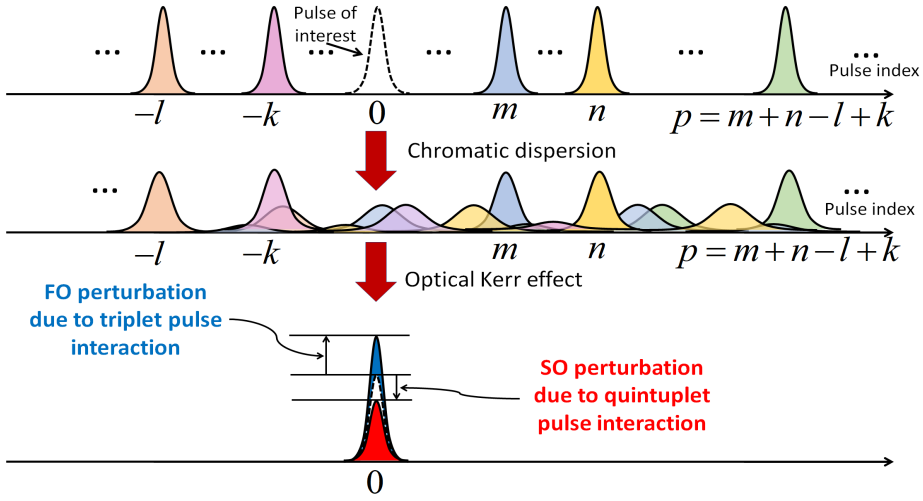


Figure 6.2: The quintuplet pulses involved in the SO distortion field calculation.

In contrast to the triplet pulses-induced nonlinear ghost pulse generation in the FO perturbation theory, the SO distortion field u_2 is generated by nonlinear interaction between quintuplet pulses, as shown in Fig. 6.2. In Fig. 6.2, the triplet pulses located at arbitrary time indices m, n, and l generate the FO ghost pulse at the time index $m+n-l$. That will further interact nonlinearly with the zeroth-order pulses at the time indices k and $p=m+n-l+k$, where p is the phase-matching condition, to induce the SO ghost pulse at zeroth time index. Accordingly, the SO ghost pulse

$$\mathbf{C}^{SO,\,Term\,1}_{m,n,l,k} = -\tau^4 \int_0^L \int_0^z \frac{\exp(-\alpha(z+s))}{\sqrt{-\overline{A}(z,s)\overline{B}(s)}} \exp\left\{ \frac{2T^2}{\tau^2 \overline{A}(z,s)\overline{B}(s)} \left[\overline{C}_{m,n,l,k}\tau^8 \right.\right.$$
$$+ j\beta_2(\overline{D}_{m,n,l,k}s - \overline{E}_{m,n,l,k}z)\tau^6 + \frac{3}{2}\beta_2^2(\overline{F}_{m,n,l,k}s^2 - \overline{G}_{m,n,l,k}sz + \overline{H}_{m,n,l,k}z^2)\tau^4$$
$$\left.\left. + \frac{3}{2}j\beta_2^3 sz(\overline{I}_{m,n,l,k}s - \overline{J}_{m,n,l,k}z)\tau^2 + \frac{1}{4}s^2z^2\beta_2^4\overline{K}_{m,n,l,k} \right] \right\} ds\, dz, \quad (6.7)$$

where

$$\overline{A}(z,s) = \left(j\tau^6 - 3\beta_2\left(s + 2/3z\right)\tau^4 - 6j\beta_2^2\left(s - 7/6z\right)z\tau^2 - 5sz^2\beta_2^3 \right), \quad (6.8)$$
$$\overline{B}(s) = \left(j\tau^2 + \beta_2 s \right), \quad (6.9)$$
$$\overline{C}_{m,n,l,k} = \left(7/4l^2 + \left(-k/2 - 2m - 2n\right)l + 3/4m^2 + \left(k/2 + n\right)m \right.$$
$$\left. +1/2nk + 1/2k^2 + 3/4n^2 \right), \quad (6.10)$$
$$\overline{D}_{m,n,l,k} = \left(-l^2 + \left(-k + m/2 + n/2\right)l + 1/2m^2 + \left(k - n/2\right)m \right.$$
$$\left. +nk + 1/2n^2 + k^2 \right), \quad (6.11)$$
$$\overline{E}_{m,n,l,k} = \left(-5l^2 + \left(7/2k + 6(m+n)\right)l - 2m^2 + 7/2\left(-3/7k - 6/7n\right)m \right.$$
$$\left. -1/2\left(3k + 4n\right)n \right), \quad (6.12)$$
$$\overline{F}_{m,n,l,k} = \left(1/2l^2 - \left(m + n + k\right)l + 1/2m^2 + \left(n + k\right)m + nk + 1/2n^2 + k^2 \right), \quad (6.13)$$
$$\overline{G}_{m,n,l,k} = \left(-10/3l^2 + \left(10/3k + 4(m+n)\right)l - 2/3m^2 + 10/3\left(-3/5k - n\right)m \right.$$
$$\left. -2/3\left(3k + n\right)n \right), \quad (6.14)$$
$$\overline{H}_{m,n,l,k} = \left(7/2l^2 + 7/2\left(-4/3m - 4/3n\right)l + 2nm + 13/6(m + n^2) \right), \quad (6.15)$$
$$\overline{I}_{m,n,l,k} = \left(k\left(l - m - n\right) \right), \quad (6.16)$$
$$\overline{J}_{m,n,l,k} = \left(10/3l^2 + 1/3\left(-13(m+n)\right)l + m^2 + 11/3nm + n^2 \right), \quad (6.17)$$
$$\overline{K}_{m,n,l,k} = \left(\left(l - m - n\right)^2 \right). \quad (6.18)$$

---

is generated by the nonlinear interaction between the FO ghost pulse and two other linearly dispersed pulses.

For simplicity of analysis, we consider Term 1 and Term 2 of the nonlinear part in (6.6) separately, and finally, combine them.

Lemma 1. *By considering Term 1 of the nonlinear part in (6.6), the coefficient of the nonlinear interaction between five input Gaussian pulses* $\sqrt{P_0}a_{m/n/l/k/p}\exp(-(t - T_{m/n/l/k/p})^2/2\tau^2)$ *at five time indices* T_m, T_n, T_l, T_k, T_p *with the assumption of a sym-*

bol rate operation (i.e., $t = 0$) and substituting the phase-matching condition $p = m + n - l + k$, can be expressed as (6.7).

Proof: Please refer to Appendix C.1.

It is important to mention that the Gaussian shape assumption for the input pulse shape is adopted to simplify the mathematical analysis. In the perturbative analysis, the nonlinear perturbation coefficients are calculated using the overlap integrals that relate the symbol under consideration to other symbols that take part in the nonlinear interaction. The overlap integrals cannot be calculated analytically for non-Gaussian pulse shapes, such as root-raised cosine (RRC) or Nyquist pulses [108]. It requires additional approximations, such as the stationary-phase approximation, to circumvent the difficulty in explicitly evaluating overlap integrals. The results in [109] indicate that the Gaussian pulse shape assumption in the perturbative analysis is reasonably valid for systems using RRC pulse shape to demonstrate the proof of concept.

Lemma 2. *By considering Term 2 of the nonlinear part in (6.6), the coefficient of the nonlinear interaction between five input Gaussian pulses $\sqrt{P_0}a_{m/n/l/k/p}\exp(-(t - T_{m/n/l/k/p})^2/2\tau^2)$ at five time indices T_m, T_n, T_l, T_k, T_p with the assumption of a symbol rate operation and substituting the phase-matching condition $p = m + n - l + k$, can be expressed as (6.19).*

Proof: Please refer to Appendix C.2.

The 4-dimensional (D) matrices $\mathbf{C}_{m,n,l,k}^{SO,Term\,1}$ and $\mathbf{C}_{m,n,l,k}^{SO,Term\,2}$ represent the coefficients of nonlinear interaction between the quintuplet pulses shown in Fig. 6.2.

Theorem 1. *The five input Gaussian pulses $\sqrt{P_0}a_{m/n/l/k/p}\exp(-(t-T_{m/n/l/k/p})^2/2\tau^2)$ at five time indices T_m, T_n, T_l, T_k, and T_p, where $p = m + n - l + k$ is the phase-matching condition, generate the SO ghost pulse at the zeroth time index; with the assumption of a symbol rate operation, this can be expressed as:*

$$\mathbf{C}_{m,n,l,k}^{SO,\,Term\,2} = \sqrt{3}\tau^4 \int_0^L \int_0^z \frac{\sqrt{\widehat{A}(z,s)}\exp(-\alpha(z+s))}{\sqrt{\widehat{B}(z,s)\widehat{C}(z)}\left(\sqrt{-\overline{B}(s)\widehat{D}(z,s)}\right)^*} \exp\left\{\frac{-jT^2}{2\tau^2\widehat{B}(z,s)\widehat{E}(z)\widehat{F}(s)}\right.$$
$$\left[\widehat{G}_{m,n,l,k}\tau^8 + 2j\beta_2(\widehat{H}_{m,n,l,k}z + \widehat{I}_{m,n,l,k}s)\tau^6 - 3\beta_2^2(\widehat{J}_{m,n,l,k}z^2 - \widehat{K}_{m,n,l,k}sz + \widehat{L}_{m,n,l,k}s^2)\tau^4\right.$$
$$\left.\left. - 2j\beta_2^3 sz(\widehat{M}_{m,n,l,k}z + \widehat{N}_{m,n,l,k}s)\tau^2 \; - s^2z^2\beta_2^4\widehat{O}_{m,n,l,k}\right]\right\} ds\,dz, \quad (6.19)$$

where

$$\widehat{A}(z,s) = \left(j\tau^4 + jsz\beta_2^2 + 3\tau^2\beta_2\,(s-z)\right), \quad (6.20)$$
$$\widehat{B}(z,s) = \left(\tau^4 - 3j\,(s - 7/3z)\,\beta_2\tau^2 + 5sz\beta_2^2\right), \quad (6.21)$$
$$\widehat{C}(z) = \left(j\tau^2 + \beta_2 z\right), \quad (6.22)$$
$$\widehat{D}(z,s) = \left(\tau^4 + sz\beta_2^2 + j\beta\tau^2\beta_2\,(s-z)\right), \quad (6.23)$$
$$\widehat{E}(z) = \left(j\beta_2 z - \tau^2\right), \quad (6.24)$$
$$\widehat{F}(s) = \left(j\tau^2 - \beta_2 s\right), \quad (6.25)$$
$$\widehat{G}_{m,n,l,k} = \left(-\,(m+n-l+k)^2 - k^2 - 6l^2 + 6\,(m+n)\,l - 2(m^2+mn+n^2)\right), \quad (6.26)$$
$$\widehat{H}_{m,n,l,k} = \left(-3\,(m+n-l+k)^2 + (k + 2(2l-m-n))\,(m+n-l+k) - 3k^2\right.$$
$$\left. + (2(2l-m-n))\,k - 2l^2 + 2\,(m+n)\,l - 2(m^2 - mn + n^2)\right), \quad (6.27)$$
$$\widehat{I}_{m,n,l,k} = \left((m+n-l+k)^2 + k^2 - 3\,(-n+l)\,(l-m)\right), \quad (6.28)$$
$$\widehat{J}_{m,n,l,k} = \left((m+n-l+k)^2 + (2k - 8/3l + 4/3m + 4/3n)\,(m+n-l+k) + k^2\right.$$
$$\left. -4/3\,(2l-m-n)\,k + 10/3l^2 - 10/3\,(m+n)\,l + 2m^2 + 2n^2 - 2/3mn\right), \quad (6.29)$$
$$\widehat{K}_{m,n,l,k} = \left(4/3\,(m+n-l+k)^2 + 4/3\,(k - 2l + m + n)\,(m+n-l+k)\right.$$
$$\left. +4/3\,(n-l+k)\,(m-l+k)\right), \quad (6.30)$$
$$\widehat{L}_{m,n,l,k} = \left((m+n-l+k)^2 + k^2\right), \quad (6.31)$$
$$\widehat{M}_{m,n,l,k} = \left((m+n-l+k)^2 + 2\,(k - 2l + m + n)\,(m+n-l+k)\right.$$
$$\left. +k^2 - 2\,(2l-m-n)\,k + 5\,(-n+l)\,(l-m)\right), \quad (6.32)$$
$$\widehat{N}_{m,n,l,k} = \left((m+n-l+k)^2 - 3\,(m+n-l+k)\,k + k^2\right), \quad (6.33)$$
$$\widehat{O}_{m,n,l,k} = \left((m+n-l+2k)^2\right). \quad (6.34)$$

$$u_2(L,t) = \gamma^2 P_0^{5/2} \sum_m \sum_n \sum_l \sum_k \left[2a_m a_l^* a_n a_k a_{m+n-l+k}^* \mathbf{C}_{m,n,l,k}^{SO,\,Term\,1}\right.$$
$$\left. + a_m^* a_l a_n^* a_k a_{m+n-l+k} \mathbf{C}_{m,n,l,k}^{SO,\,Term\,2}\right], \quad (6.35)$$

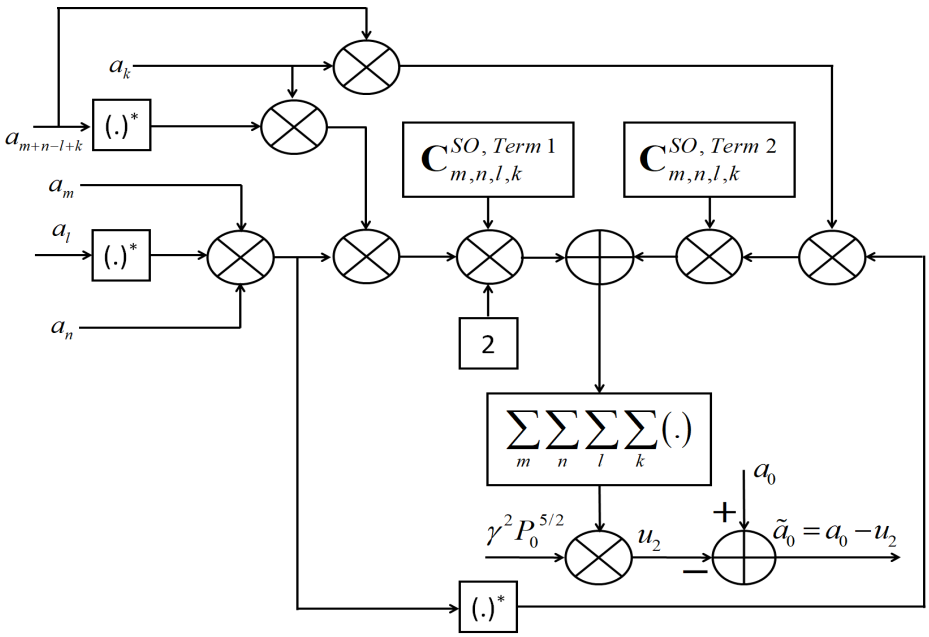


Figure 6.3: The block diagram of the SO-PB-NLC technique using (6.35).

where $\mathbf{C}^{SO,Term\,1}_{m,n,l,k}$ *and* $\mathbf{C}^{SO,Term\,2}_{m,n,l,k}$ *are given by (6.7) and (6.19), respectively.*

Proof: Please refer to Appendix C.3.

The SO predistortion technique using (6.35) is also based on the same simplifying assumptions considered for the FO-PB-NLC technique, such as the full electronic compensation of the CD effect at the receiver and the Gaussian shape assumption for the input pulse shape [46]. Fig. 6.3 shows the block diagram of the SO-PB-NLC technique for a single-polarization and single-channel coherent optical communication system using (6.35). The SO nonlinearity coefficient matrices $\mathbf{C}^{SO,Term\,1}_{m,n,k}$ and $\mathbf{C}^{SO,Term\,2}_{m,n,k}$ are calculated offline and stored in LUTs. Then, the SO distortion field u_2 is calculated using (6.35) and subtracted from the zeroth index symbol a_0 to generate the predistorted symbol $\widetilde{a}_0$, as shown in Fig. 6.3.

6.5 Simplified SO Distortion Field Expression

As discussed previously, the SO-PB-NLC technique uses (6.35) to calculate the SO nonlinear distortion field. It considers the nonlinear interaction of the quintuplet pulses located at all possible arbitrary time indices. However, (6.35) is practically unrealizable as the possible combinations of the dispersed symbols with symbol indices m, n, l, and k approach infinity. Consequently, we put a cap on the maximum number of the perturbation terms in the calculation of (6.35) by introducing a truncation threshold for the 4-D nonlinearity coefficient matrices $\mathbf{C}_{m,n,l,k}^{SO,\,Term\,1}$ and $\mathbf{C}_{m,n,l,k}^{SO,\,Term\,2}$. The truncation threshold can be defined as the threshold at which the magnitude of $\mathbf{C}_{m,n,l,k}^{SO,\,Term\,1/Term\,2}$ is less than the maximum magnitude $C_{0,0,0,0}^{SO,\,Term\,1/Term\,2}$ by a factor μ, i.e., $20\log_{10}\left(\left|\mathbf{C}_{m,n,l,k}^{SO,\,Term\,1/Term\,2}\right| / \left|C_{0,0,0,0}^{SO,\,Term\,1/Term\,2}\right|\right) < \mu$ [46].

Although we truncate at the threshold μ, the 4-D nonlinearity coefficient matrices may still contain a large number of terms. To further reduce the number of coefficient terms, we neglect the FO fields generated at the time indices other than that of the symbol under consideration (i.e., the pulse at zeroth index) for which the SO distortion field is calculated. That can be achieved by substituting the phase-matching condition $l = m + n$ in (6.7) and (6.19). Figs. 6.4 and 6.5 show the nonlinear interaction between the quintuplet pulses in Term 1 and Term 2 of (6.6), respectively, for the phase-matching condition $l = m + n$. In Fig. 6.4, the phase-matching condition $l = m + n$ leads to $|p| = |k|$. That results in a two-pulse collision between the FO ghost pulse and one other linearly dispersed pulse. On the other hand, in Fig. 6.5, the phase-matching condition $l = m+n$ leads to $p = -k$ and $-k = p$, which increases the chance of constructive/destructive interference caused by the three-pulse collision between the FO ghost pulse and the linearly dispersed pulses. We numerically evaluate the performance of the SO predistorter in the presence and absence of this simplifying assumption in Section 6.7.

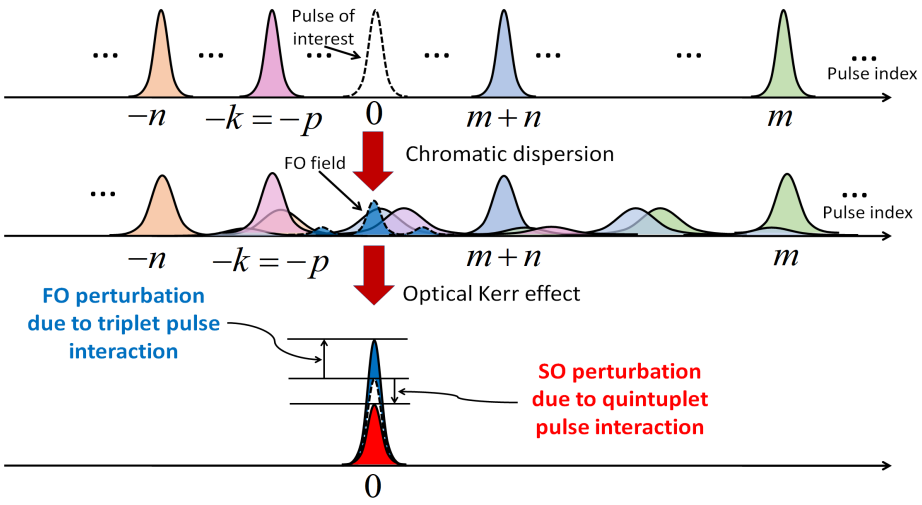


Figure 6.4: The quintuplet pulses involved in Term 1 of (6.6) and their nonlinear interaction for the phase-matching conditions $l = m + n$ and $|p| = |k|$.

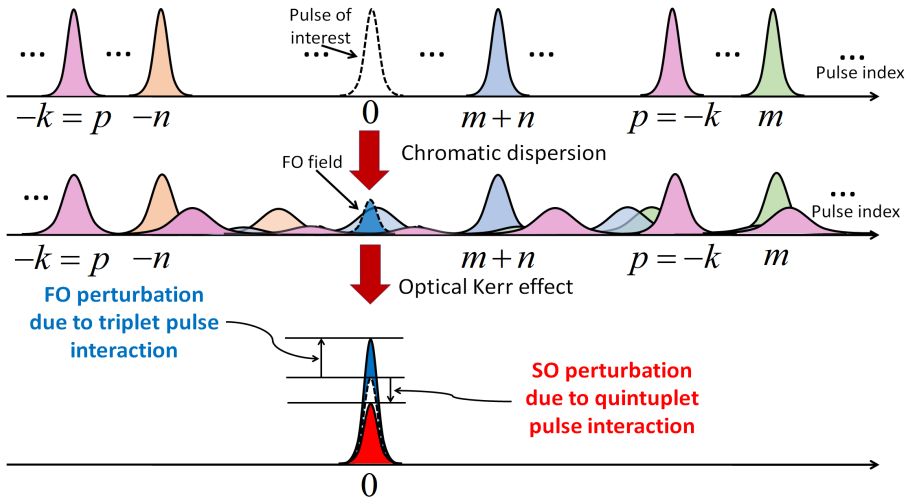


Figure 6.5: The quintuplet pulses involved in Term 2 of (6.6) and their nonlinear interaction for the phase-matching conditions $l = m + n$, $p = -k$, and $-k = p$.

Corollary 1. *By substituting the phase-matching condition $l = m + n$ in (6.35), the expression for the SO distortion field generated at the zeroth time index can be represented as:*

$$\widetilde{\mathbf{C}}_{m,n,k}^{SO,\,Term\,1} = -\tau^4 \int_0^L \int_0^z \frac{\exp(-\alpha(z+s))}{\sqrt{-\overline{A}(z,s)\overline{B}(s)}} \times \exp\left\{ \frac{T^2}{\overline{\overline{A}(z,s)\overline{B}(s)}} \left[\check{A}_{m,n,k}\tau^6 + 2j\beta_2(\check{B}_{m,n,k}z + \check{C}_{m,n,k}s)\tau^4 + 3\beta_2^2(\check{D}_{m,n,k}z^2 - \check{E}_{m,n,k}sz + k^2s^2)\tau^2 - 5jmnsz^2\beta_2^3 \right] \right\} ds\, dz, \quad (6.36)$$

where

$$\check{A}_{m,n,k} = \left(k^2 + m^2 + nm + n^2\right), \quad (6.37)$$
$$\check{B}_{m,n,k} = \left(m^2 + (-2k+n)\,m - 2kn + n^2\right), \quad (6.38)$$
$$\check{C}_{m,n,k} = \left(k^2 - 3/2nm\right), \quad (6.39)$$
$$\check{D}_{m,n,k} = \left(m^2 - 1/3nm + n^2\right), \quad (6.40)$$
$$\check{E}_{m,n,k} = \left(4/3\left(\left(k - 3/2n\right)m + kn\right)\right). \quad (6.41)$$

$$\widetilde{\mathbf{C}}_{m,n,k}^{SO,\,Term\,2} = \sqrt{3}\tau^4 \int_0^L \int_0^z \frac{\sqrt{\widehat{A}(z,s)}\exp(-\alpha(z+s))}{\sqrt{\widehat{B}(z,s)\widehat{C}(z)}\left(\sqrt{-\overline{B}(s)\widehat{D}(z,s)}\right)^*} \times \exp\left\{ \frac{-jT^2}{\tau^2\widehat{B}(z,s)\widehat{F}(s)} \left[\breve{A}_{m,n,k}\tau^6 - j\beta_2(\breve{A}_{m,n,k}z + \breve{B}_{m,n,k}s)\tau^4 + \beta_2^2 s(\breve{C}_{m,n,k}z + 3k^2s)\tau^2 + 2jk^2s^2z\beta_2^3 \right] \right\} ds\, dz, \quad (6.42)$$

where

$$\breve{A}_{m,n,k} = \left(-6k^2 + 4\left(m+n\right)k - 3m^2 + nm - 3n^2\right), \quad (6.43)$$
$$\breve{B}_{m,n,k} = \left(2k^2 - 3nm\right), \quad (6.44)$$
$$\breve{C}_{m,n,k} = \left(-4k^2 + 4k\left(m+n\right) - 5nm\right). \quad (6.45)$$

$$\widetilde{u}_2(L,t) = \gamma^2 P_0^{5/2} \sum_m \sum_n \sum_k \left[2a_m a_{m+n}^* a_n a_k a_k^* \widetilde{\mathbf{C}}_{m,n,k}^{SO,\,Term\,1} + a_m^* a_{m+n} a_n^* a_k^2 \widetilde{\mathbf{C}}_{m,n,k}^{SO,\,Term\,2} \right], \quad (6.46)$$

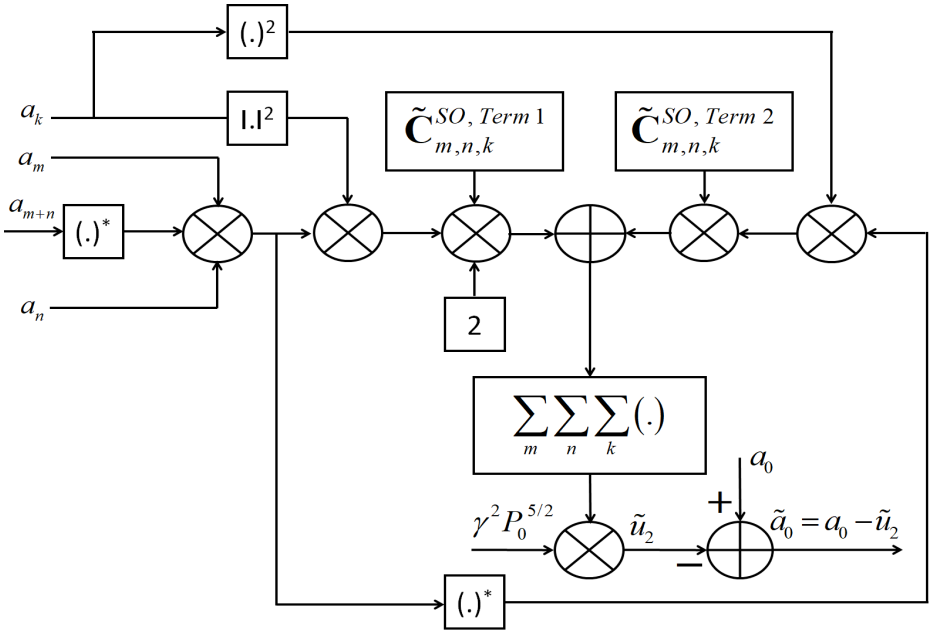


Figure 6.6: The block diagram of the SO-PB-NLC technique using (6.46).

where $\widetilde{\mathbf{C}}^{SO,\,Term\,1}_{m,n,k}$ *and* $\widetilde{\mathbf{C}}^{SO,\,Term\,2}_{m,n,k}$ *are given by (6.36) and (6.42), respectively.*

Proof: Please refer to Appendix C.4.

Fig. 6.6 shows the block diagram of the simplified version of the SO-PB-NLC technique using (6.46). The SO nonlinearity coefficient matrices $\widetilde{\mathbf{C}}^{SO,\,Term\,1}_{m,n,k}$ and $\widetilde{\mathbf{C}}^{SO,\,Term\,2}_{m,n,k}$ are calculated offline and stored in LUTs. The SO distortion field $\widetilde{u}_2$ is calculated using (6.46) and subtracted from the zeroth index symbol a_0 to generate the predistorted symbol $\widetilde{a}_0$, as shown in Fig. 6.6.

6.6 Extension to Dual-polarization

The propagation of the polarization multiplexed signal in the SSMF is governed by the Manakov equation, as given in (2.62). The differential equation governing the SO

distortion field for the polarization multiplexed signal can be represented as:

$$\frac{\partial}{\partial z}u_{2,x/y}(z,t) = \underbrace{-j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}u_{2,x/y}(z,t)}_{\text{Linear part}} + \underbrace{\left(\begin{array}{c} \underbrace{j2\left(\left|u_{0,x/y}(z,t)\right|^2+\left|u_{0,y/x}(z,t)\right|^2\right)\tilde{u}_{1,x/y}(z,t)\exp(-\alpha z)}_{\text{Term 1}} \\ \underbrace{+j\left(u_{0,x/y}^2(z,t)+u_{0,y/x}^2(z,t)\right)\tilde{u}_{1,x/y}^*(z,t)\exp(-\alpha z)}_{\text{Term 2}} \end{array}\right)}_{\text{Nonlinear part}}. \quad (6.47)$$

Theorem 2. *For the case of the transmission of a polarization multiplexed optical signal through the SSMF, the five input Gaussian pulses* $\sqrt{P_0}a_{m/n/l/k/p,x/y}\exp(-(t-T_{m/n/l/k/p})^2/2\tau^2)$ *at five time indices* T_m, T_n, T_l, T_k, *and* T_p, *where* $p=m+n-l+k$ *is the phase-matching condition, generate the SO ghost pulse at the zeroth time index; with the assumption of a symbol rate operation, this can be expressed as:*

$$u_{2,x/y}(L,t) = \frac{64}{81}\gamma^2 P_0^{5/2}\sum_m\sum_n\sum_l\sum_k\Big[2\left(a_{m,x/y}a_{l,x/y}^*+a_{m,y/x}a_{l,y/x}^*\right)a_{n,x/y}$$
$$\times\left(a_{k,x/y}a_{m+n-l+k,x/y}^*+a_{k,y/x}a_{m+n-l+k,y/x}^*\right)\mathbf{C}_{m,n,l,k}^{SO,Term\,1}+\left(a_{m,x/y}^*a_{l,x/y}+a_{m,y/x}^*a_{l,y/x}\right)$$
$$\times a_{n,x/y}^*\left(a_{k,x/y}a_{m+n-l+k,x/y}+a_{k,y/x}a_{m+n-l+k,y/x}\right)\mathbf{C}_{m,n,l,k}^{SO,Term\,2}\Big], \quad (6.48)$$

where $\mathbf{C}_{m,n,l,k}^{SO,Term\,1}$ *and* $\mathbf{C}_{m,n,l,k}^{SO,Term\,2}$ *are given by (6.7) and (6.19), respectively.*

Proof: Please refer to Appendix C.5.

It is important to note that, in the dual-polarization system, the coefficient of nonlinear interaction between the pulses in the co-propagating orthogonal polarization tributaries is the same as the coefficient of nonlinear interaction for the pulses in the same polarization. In other words, the nonlinearity coefficient matrices $\mathbf{C}_{m,n,l,k}^{SO,Term\,1}$ and $\mathbf{C}_{m,n,l,k}^{SO,Term\,2}$ are same for both single-polarization and dual-polarization systems. Fig.

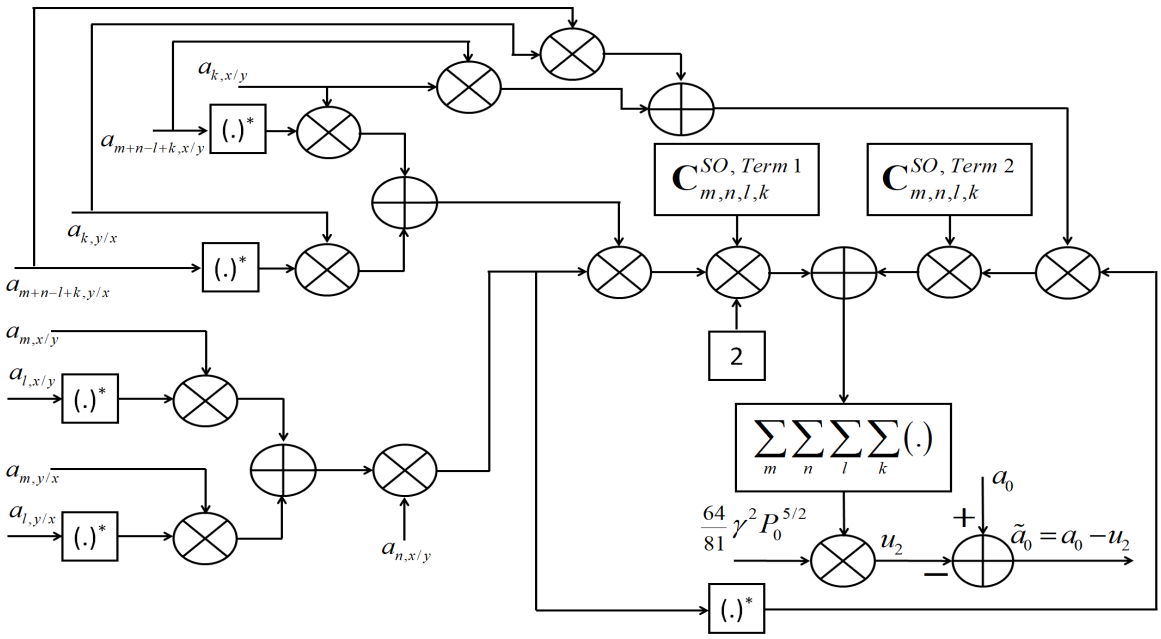


Figure 6.7: The block diagram of the SO-PB-NLC technique using (6.48).

6.7 shows the block diagram of the SO-PB-NLC technique for the dual-polarization system using (6.48). As in the case of the single-polarization system, the nonlinearity coefficient matrices $\mathbf{C}^{SO,Term\,1}_{m,n,l,k}$ and $\mathbf{C}^{SO,Term\,2}_{m,n,l,k}$ are stored in LUTs. Then, the SO distortion field is calculated using (6.48) and subtracted from the zeroth index symbol to generate the predistorted symbol.

Corollary 2. *By substituting the phase-matching condition $l = m + n$ in (6.48), the expression for the SO distortion field generated at the zeroth time index can be represented as:*

$$\tilde{u}_{2,x/y}(L,t) = \frac{64}{81}\gamma^2 P_0^{5/2} \sum_m \sum_n \sum_k \Big[2\left(a_{m,x/y}a^*_{m+n,x/y} + a_{m,y/x}a^*_{m+n,y/x}\right) a_{n,x/y} \times \left(a_{k,x/y}a^*_{k,x/y} + a_{k,y/x}a^*_{k,y/x}\right) \widetilde{\mathbf{C}}^{SO,Term\,1}_{m,n,k} + \left(a^*_{m,x/y}a_{m+n,x/y} + a^*_{m,y/x}a_{m+n,y/x}\right) \times a^*_{n,x/y}\left(a_{k,x/y}a_{k,x/y} + a_{k,y/x}a_{k,y/x}\right) \widetilde{\mathbf{C}}^{SO,Term\,2}_{m,n,k} \Big], \quad (6.49)$$

where $\widetilde{\mathbf{C}}^{SO,Term\,1}_{m,n,k}$ and $\widetilde{\mathbf{C}}^{SO,Term\,2}_{m,n,k}$ are given by (6.36) and (6.42), respectively.

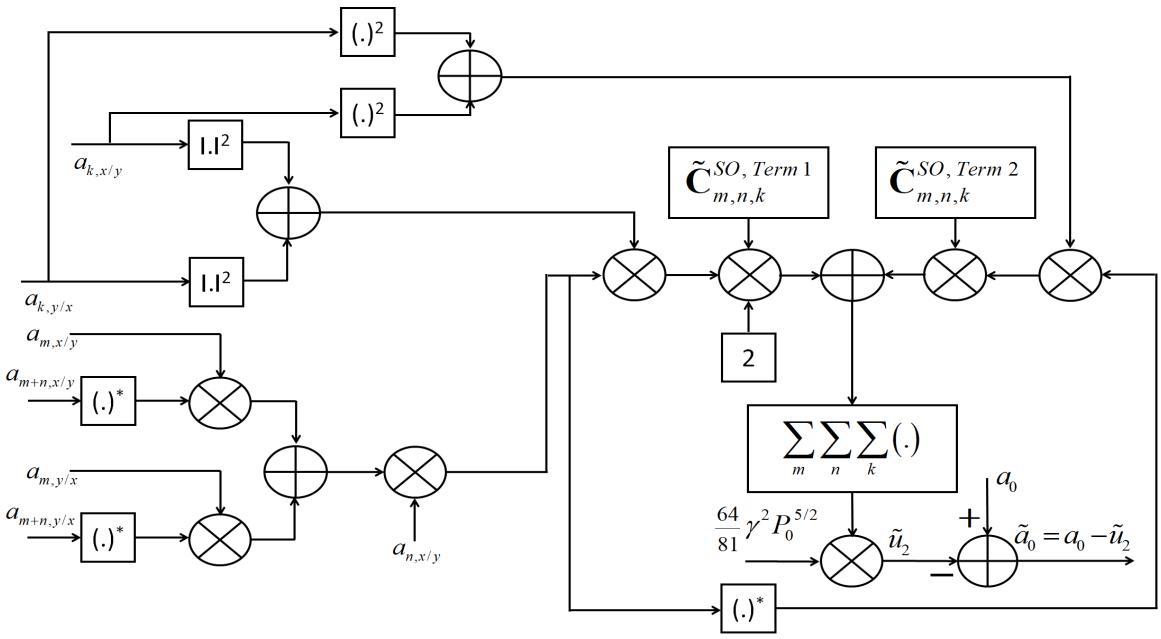


Figure 6.8: The block diagram of the SO-PB-NLC technique using (6.49).

Proof: Please refer to Appendix C.6.

Fig. 6.8 shows the block diagram of the simplified version of the SO-PB-NLC technique using (6.49) for dual-polarization systems. The 3-D nonlinearity coefficient matrices $\tilde{\mathbf{C}}^{SO,Term\,1}_{m,n,k}$ and $\tilde{\mathbf{C}}^{SO,Term\,2}_{m,n,k}$ are calculated offline and stored in LUTs. Then, the nonlinear distortion field is calculated using (6.49), which is followed by the subtraction of the calculated field from the zeroth index symbol to generate the predistorted symbol.

It is worth mentioning that the IFWM terms of the noise evolution equation in [107] (please see (19) in [107]) are neglected since their magnitudes are smaller when compared to the IXPM term in a dispersion unmanaged transmission system. Accordingly, we evaluate the contribution of Term 2 to the SO nonlinear distortion field given in (6.49).

Fig. 6.9 shows the histogram plot of the magnitude of the nonlinearity coefficients in $\tilde{\mathbf{C}}^{SO,Term\,1}_{m,n,k}$ and $\tilde{\mathbf{C}}^{SO,Term\,2}_{m,n,k}$ for a transmission distance of 2800 km. It is worthy to

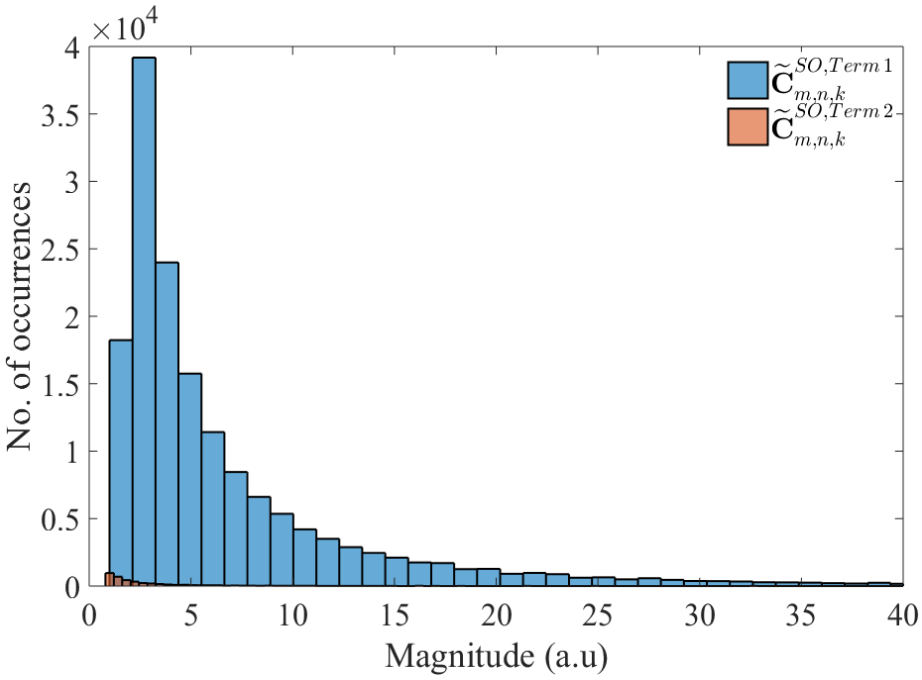


Figure 6.9: The number of occurrences of the magnitude of the nonlinearity coefficients in $\widetilde{\mathbf{C}}_{m,n,k}^{SO,\,Term\,1}$ and $\widetilde{\mathbf{C}}_{m,n,k}^{SO,\,Term\,2}$.

note that the nonlinearity coefficient matrices $\widetilde{\mathbf{C}}_{m,n,k}^{SO,\,Term\,1}$ and $\widetilde{\mathbf{C}}_{m,n,k}^{SO,\,Term\,2}$ are the same for both single-polarization and dual-polarization transmission systems. From Fig. 6.9, it is observed that the number of coefficients in $\widetilde{\mathbf{C}}_{m,n,k}^{SO,\,Term\,2}$ above the truncation threshold μ is significantly lower than that in $\widetilde{\mathbf{C}}_{m,n,k}^{SO,\,Term\,1}$. Also, the magnitudes of the coefficients in $\widetilde{\mathbf{C}}_{m,n,k}^{SO,\,Term\,2}$ are close to zero with a significantly lower variance when compared to the coefficient magnitudes in $\widetilde{\mathbf{C}}_{m,n,k}^{SO,\,Term\,1}$. That may be due to the constructive/destructive interference caused by the three-pulse collision between the FO ghost pulse and the linearly dispersed pulses (please see Fig. 6.5). We investigate the performance of the predistorter with and without considering Term 2 of (6.6) through numerical simulations in Section 6.7.

In the implementation of the SO-PB-NLC technique, we adopt a quantization method proposed in [110] to reduce the computational complexity further. It is important to mention that the nonlinearity coefficients in $\widetilde{\mathbf{C}}_{m,n,k}^{SO,\,Term\,1}$ and $\widetilde{\mathbf{C}}_{m,n,k}^{SO,\,Term\,2}$

are very similar, in particular for those with large indexes. Based on this fact, we ignore the coefficient difference of ± 0.5, which will dramatically reduce the number of nonlinearity coefficients satisfying the thresholding condition. That will significantly reduce the implementation complexity of the SO-PB-NLC technique.

6.7 Numerical Simulations and Discussions

The SO-PB-NLC technique is applied as a predistortion at the transmitter. After the RRC pulse shaping, the predistorted signal is up-converted to the optical domain and transmitted over the long-haul optical fiber link. The simulation parameters used for the study are listed in Table 6.1. The modulation format used is 16-QAM. The data transmission rate is 32 Gbaud. It is assumed that the polarization state, carrier phase, and symbol timing are perfectly known at the receiver [64]. The amplified spontaneous emission (ASE) noise of EDFA is added to the signal after each fiber span to capture the nonlinear interaction between the signal and the ASE noise [64].

Table 6.1: Simulation Parameters [46], [55], [73].

Parameter	Value
RRC filter roll-off factor	0.1
μ	-40 dB
Fiber span length	80 km
α	0.2 dB/km
β_2	-20.47 ps^2/km
γ	1.22 (1/W)/km
Polarization mode dispersion coefficient	0.1 ps/$\sqrt{\text{km}}$
Noise figure of EDFA	5.5 dB

6.7.1 Simulation Results

Fig. 6.10 shows the SNR as a function of the launch power at a transmission distance of 2800 km for the SO-PB-NLC technique with 4-D and 3-D nonlinearity coefficient matrices based on (6.35) and (6.46), respectively. Fig. 6.10 also shows the plot of the SNR gain as a function of the launch power. The SNR gain is defined as the difference between the SNR obtained for the SO-PB-NLC technique with the 4-D and 3-D nonlinearity coefficient matrices and the SNR obtained for the case of electronic dispersion compensation (EDC) only, i.e., $\triangle \text{SNR}_{\text{SO-PB-NLC (4-D/3-D NL coeff.)}} = \text{SNR}_{\text{SO-PB-NLC (4-D/3-D NL coeff.)}} - \text{SNR}_{\text{EDC}}$. The rationale to include the plot for SNR gain in Fig. 6.10 is to show the quantitative values of the SNR improvement for the SO-PB-NLC technique with the 4-D and 3-D nonlinearity coefficient matrices, as a function of the launch power, when compared to the EDC case.

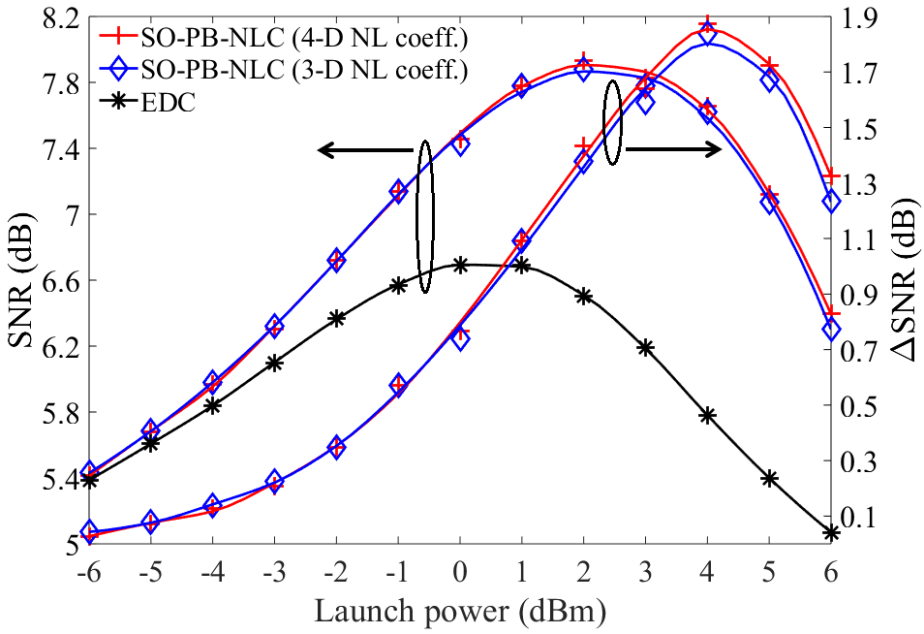


Figure 6.10: SNR as a function of the launch power for SO-PB-NLC technique with 4-D and 3-D nonlinearity coefficient matrix.

It is observed from Fig. 6.10 that the SNR gain for the SO-PB-NLC technique

with the 4-D nonlinearity coefficient matrix is less than 0.05 dB when compared to the case with a 3-D nonlinearity coefficient matrix, and the peak gain is observed at 4 dBm launch power. From this observation, we can ascertain that neglecting the FO ghost pulses generated at the arbitrary time indices $m + n - l \neq 0$ only slightly affects the compensation performance of the SO-PB-NLC technique. That is because the CD-induced pulse overlap between the FO ghost pulse at arbitrary time indices $m + n - l \neq 0$ and the zeroth-order pulses is significantly less, and thereby, the magnitude of the corresponding nonlinearity coefficient is negligibly small.

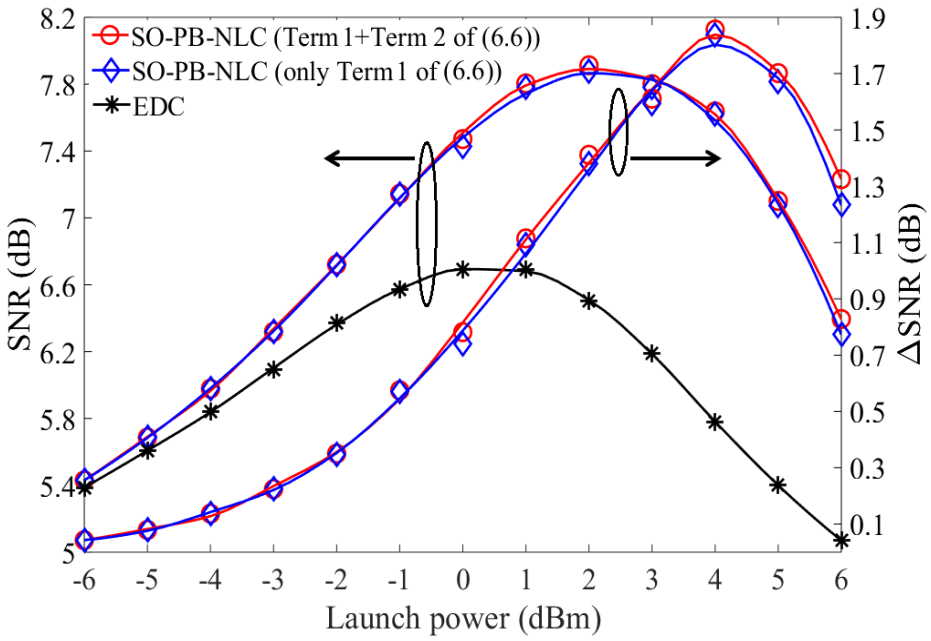


Figure 6.11: SNR as a function of the launch power for SO-PB-NLC technique with the cases of considering Term 1+Term 2 and only Term 1 of (6.6).

Fig. 6.11 shows the SNR as a function of launch power for the SO-PB-NLC technique with 3-D nonlinearity coefficient matrices. In Fig. 6.11, two cases of the SO-PB-NLC technique are shown: one is by considering both Term 1 and Term 2 of (6.6), and the other is considering only Term 1 of (6.6). The result shows that the SNR gain improves by less than 0.03 dB when Term 2 of (6.6) is also considered, in addition

to considering Term 1 of (6.6). This result confirms the observation given in Section 6.6 on the number of occurrences and the magnitude variance of the nonlinearity coefficients satisfying the truncation threshold condition in the nonlinearity coefficient matrix $\widetilde{\mathbf{C}}_{m,n,k}^{SO,Term\,2}$.

From Fig. 6.10 and Fig. 6.11, we can conclude that the SNR gain obtained is negligible when we consider the 4-D nonlinearity coefficient matrix or Term 2 of (6.6). Therefore, we select the implementation of the SO-PB-NLC technique with a 3-D nonlinearity coefficient matrix and by taking into account only Term 1 of (6.6) for further numerical investigations. The detailed evaluation of the computational complexity of the SO-PB-NLC technique will be discussed later in this section.

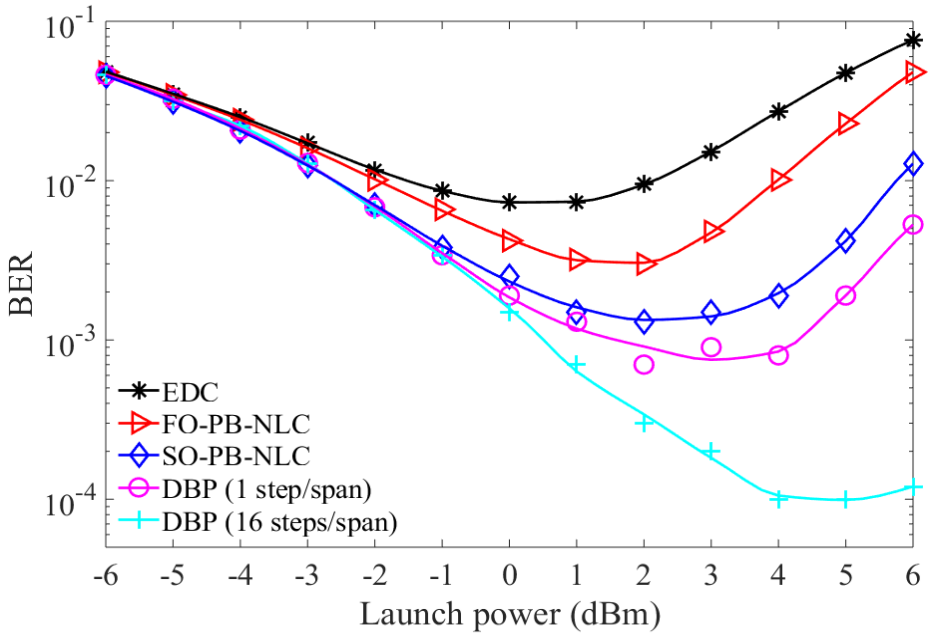


Figure 6.12: BER as a function of the launch power for EDC, FO-PB-NLC, SO-PB-NLC, and DBP (1 and 16 steps/span) techniques at a transmission distance of 2800 km for a single-polarization optical transmission system.

Fig. 6.12 shows the bit error rate (BER) as a function of the launch power for the SO-PB-NLC, FO-PB-NLC, and EDC techniques for a single-polarization and single-channel optical transmission system. The BER performance of the benchmark DBP

technique implemented with 1 and 16 steps/span are also included for comparison. The transmission distance considered is 2800 km.

We observe from Fig. 6.12 that the BER performance of the SO-PB-NLC technique is significantly better than that of the FO-PB-NLC and EDC techniques. Another observation is that the BER performance of the DBP with 1 and 16 steps/span is higher than that of the proposed SO-PB-NLC technique. That is because the DBP is a numerical method that uses the SSFM, and so it compensates for the nonlinearity effects span-by-span [96]. On the other hand, the PB-NLC techniques use an analytical approximation for the solution of the NLSE with the assumption that the fiber link has only one span [46]. That is a general assumption considered in the design of the PB-NLC techniques. It is important to mention that the single span assumption of the PB-NLC techniques allows the compensation of the nonlinearity effect in a single computation step, thus reducing the computational effort required [46].

Fig. 6.13 presents the plot of the maximum system reach as a function of the launch power for DBP (1 and 16 steps/span), SO-PB-NLC, FO-PB-NLC, and EDC techniques at 7% overhead (OH) hard-decision (HD) forward error correction (FEC) limit with a BER value of 4.3×10^{-3} [111] for a single-polarization and single-channel optical transmission system. It is observed that the maximum transmission reach for DBP (16 steps/span), DBP (1 step/span), SO-PB-NLC, FO-PB-NLC, and EDC is 6800 km, 3440 km, 3280 km, 2880 km, and 2480 km, respectively. These results indicate that the SO-PB-NLC technique provides an extended transmission reach by 32.2% and 14% when compared to EDC and the FO-PB-NLC techniques, respectively. It can also be inferred from Fig. 6.13 that the nonlinearity threshold of the DBP (16 steps/span), DBP (1 step/span), SO-PB-NLC, and FO-PB-NLC techniques is improved by ~ 11.4, ~ 6.3 dB, ~ 5.3 dB, and ~ 3.6 dB, respectively, when compared to the EDC technique at a transmission distance of 2480 km (i.e., the maximum

reach for the EDC technique). The nonlinearity threshold is defined as the value of the launch power at which the BER performance crosses the FEC limit for a given transmission distance [112]. It is interesting to note that the nonlinearity threshold of the SO-PB-NLC technique is improved by $\sim$ 1.7 dB when compared to the FO-PB-NLC technique.

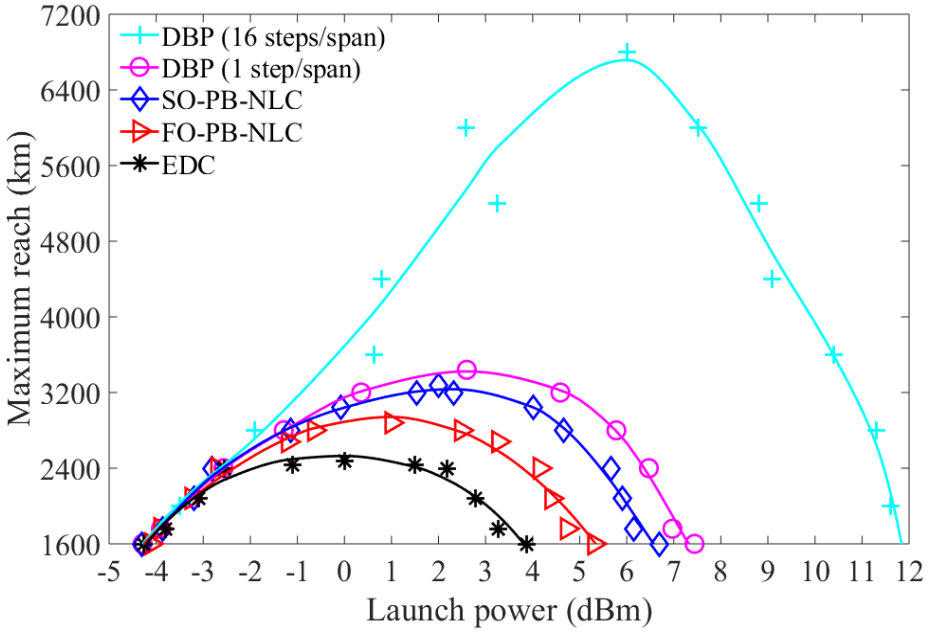


Figure 6.13: Maximum reach as a function of the launch power at 7% OH-HD-FEC limit with a BER value of 4.3×10^{-3} for a single-polarization optical transmission system.

Fig. 6.14 shows the BER as a function of the launch power for the DBP (1 and 16 steps/span), SO-PB-NLC, FO-PB-NLC, and EDC techniques for a dual-polarization and single-channel optical transmission system. It is observed that the BER performance of the SO-PB-NLC technique is significantly better than that of the FO-PB-NLC and EDC techniques. It is also observed that the optimal launch power for the SO-PB-NLC technique is increased by $\sim$ 2 dB and $\sim$ 1 dB when compared to the EDC and FO-PB-NLC techniques, respectively.

Fig. 6.15 shows the plot of the maximum transmission reach as a function of

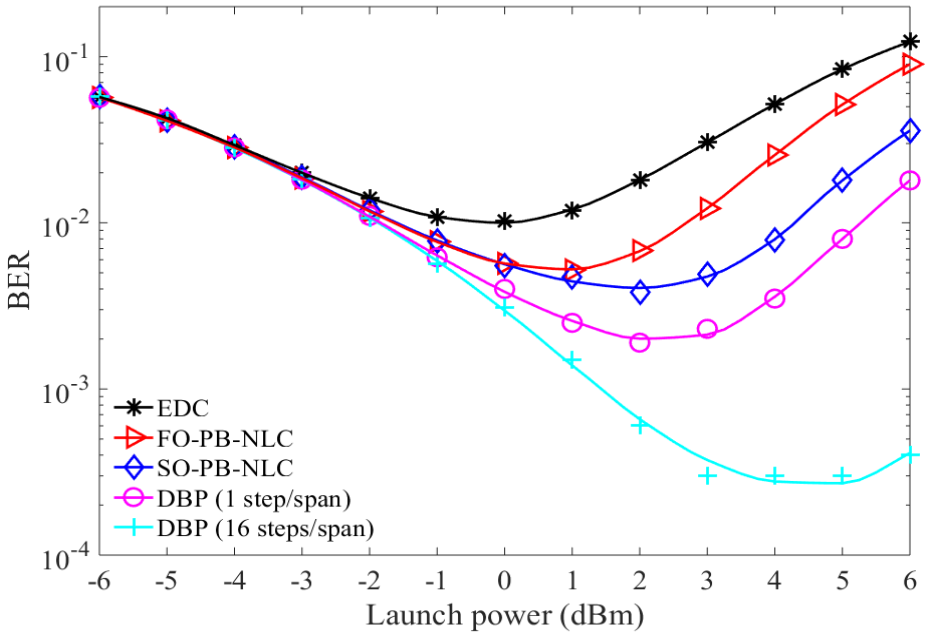


Figure 6.14: BER as a function of the launch power for EDC, FO-PB-NLC, SO-PB-NLC, and DBP (1 and 16 steps/span) techniques at a transmission distance of 2800 km for a dual-polarization optical transmission system.

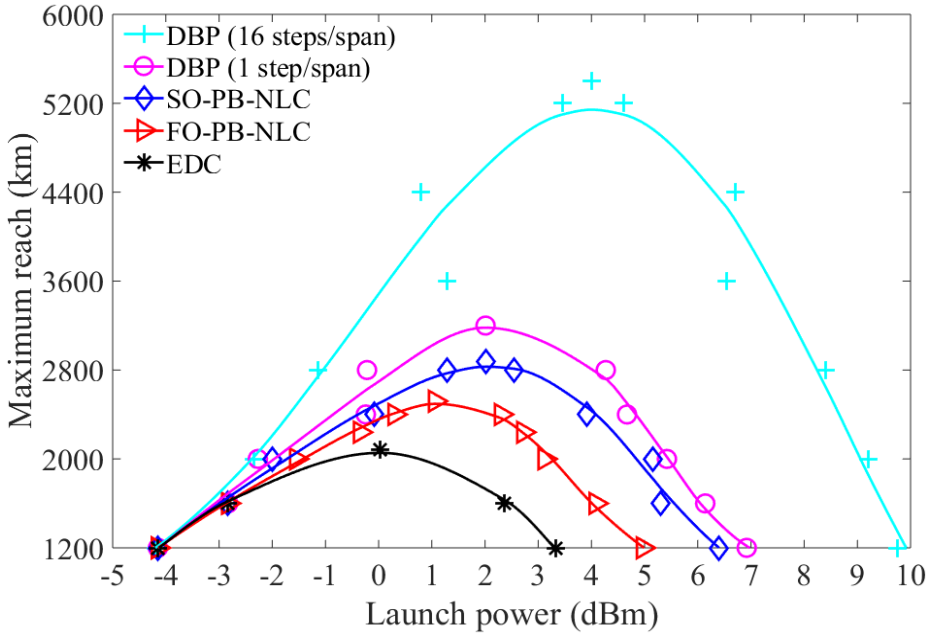


Figure 6.15: Maximum reach as a function of the launch power at 7% OH-HD-FEC limit with a BER value of 4.3×10^{-3} for a dual-polarization optical transmission system.

the launch power for DBP (1 and 16 steps/span), SO-PB-NLC, FO-PB-NLC, and EDC techniques at 7% OH-HD-FEC limit with a BER value of 4.3×10^{-3} for a dual-polarization and single-channel optical transmission system. It is observed that the maximum transmission reach for DBP (16 steps/span), DBP (1 step/span), SO-PB-NLC, FO-PB-NLC, and EDC is 5400 km, 3200 km, 2880 km, 2520 km, and 2080 km, respectively. These results indicate that the SO-PB-NLC technique provides an extended transmission reach by 38.46% and 14% when compared to EDC and the FO-PB-NLC techniques, respectively. Further, it is also observed that the nonlinearity threshold of the SO-PB-NLC technique is improved by $\sim$ 4.7 dB and $\sim$ 1.7 dB when compared to the EDC and FO-PB-NLC techniques, respectively.

6.8 Complexity Evaluation

In this section, the computational complexities of the DBP (1 and 16 steps/span), SO-PB-NLC, FO-PB-NLC, and EDC techniques are evaluated based on the number of real-valued multiplications per symbol for the dual-polarization optical transmission system. It is important to mention that the nonlinearity coefficient matrices of the FO-/SO-PB-NLC techniques are truncated at a threshold of μ=-40 dB [46]. Also, in the implementation, the nonlinearity coefficient matrices are quantized according to the method given in [110]. For DBP with N_{steps} per span, the expression for the number of real-valued multiplications per symbol is given as $8N_{steps}N_{spans}N_{\mathrm{FFT}}(\log_2(N_{\mathrm{FFT}}) + 10.5)/N_s$, where N_{spans} is the number of fiber spans, N_{FFT} is the fast Fourier transform size, and N_s is the number of samples [5]. In case of the FO-/SO-PB-NLC techniques, the triplet/quintuplet symbols in the nonlinear distortion calculation can be stored in LUT; therefore, the number of real-valued multiplications per symbol can be represented as $2(4M + 3)$ [73], where M is the number of significant perturbation

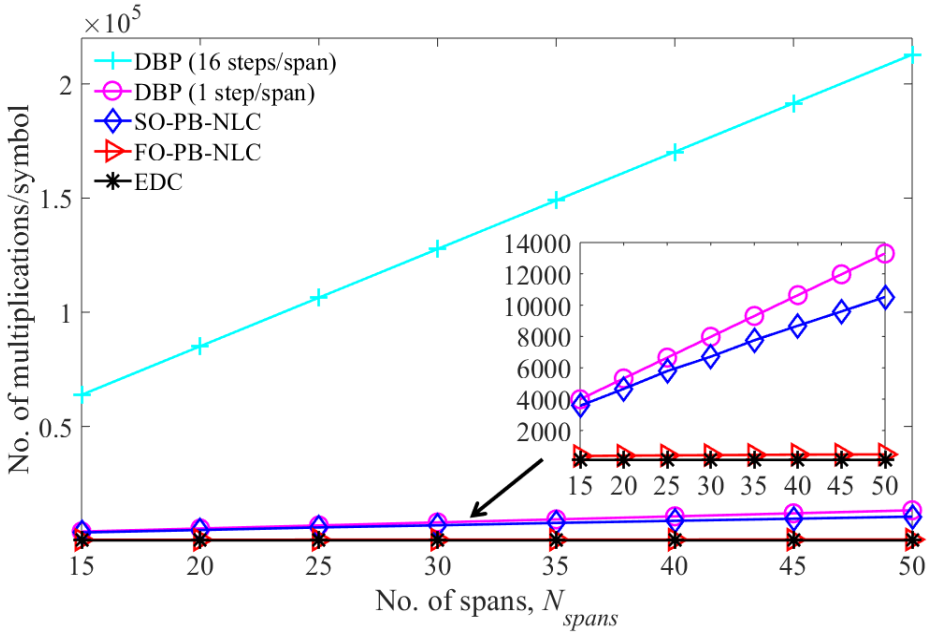


Figure 6.16: The number of real-valued multiplications/symbol for DBP (1 and 16 steps/span), SO-PB-NLC, FO-PB-NLC, and EDC techniques as a function of the number of fiber spans N_{spans}.

coefficients in the nonlinearity coefficient matrix $\mathbf{C}^{FO}_{m,n}/\widetilde{\mathbf{C}}^{SO,Term\,1}_{m,n,k}$. It is worth noting that the value of M increases with increasing the number of fiber spans because of the corresponding increase in the number of coefficients in the nonlinearity coefficient matrix, satisfying the truncation threshold. For the EDC technique, the number of real-valued multiplications per symbol is given as $8N_{\mathrm{FFT}}(\log_2(N_{\mathrm{FFT}})+1)/N_s$ [5].

Fig. 6.16 shows the number of real-valued multiplications per symbol as a function of the number of fiber spans, N_{spans} for DBP (1 and 16 steps/span), SO-PB-NLC, FO-PB-NLC, and EDC techniques. The results indicate that the complexity of the DBP technique increases rapidly as N_{spans} increases, which is attributed to the corresponding increase in the computation steps for the SSFM technique [5]. On the other hand, the complexity increase for the FO-/SO-PB-NLC techniques is due to the increase in the number of quantized nonlinearity coefficients as the number of fiber span increases. It is important to mention that the performance comparison of the

proposed SO-PB-NLC technique with DBP with 1 step/span is more reasonable when compared to the DBP with 16 steps/span. The result in Fig. 6.16 shows that the computational complexity of the proposed SO-PB-NLC technique is less than that of the DBP with 1 step/span. For example, at N_{spans}= 35 (i.e., at 2800 km), the required number of real-valued multiplications for the SO-PB-NLC technique is 1550 fewer than that of the DBP technique with 1 step/span.

6.9 Conclusion

In this chapter, we have proposed to extend the FO-PB-NLC technique to the SO, referred to as the SO-PB-NLC technique. We have presented a detailed mathematical analysis to derive the SO nonlinear distortion field with a Gaussian shape assumption for the input pulse. We have shown through numerical simulations that the NLC performance of the proposed SO-PB-NLC technique is significantly improved when compared to the FO-PB-NLC technique. We have demonstrated that the SO-PB-NLC technique extends the transmission reach by 38.46% and 14% when compared to the EDC and FO-PB-NLC techniques, respectively, for a dual-polarization and single-channel optical transmission system. We have also shown that, for a transmission distance of 2800 km, the number of real-valued multiplications per symbol for the SO-PB-NLC technique is 1550 fewer when compared to the DBP technique with 1 step/span.

Chapter 7

Conclusion and Suggested Future Work

7.1 Conclusion

In this thesis, we studied the impact of Kerr-induced fiber nonlinearity effects on long-haul coherent optical transmission systems. We briefly studied the origin of the nonlinear susceptibility and the Kerr effect in a silica-based optical fiber using the classical electron oscillator model. We briefly discussed the mechanism of the optical pulse propagation in the optical fiber medium using the nonlinear Schrödinger equation derived from Maxwell's equations. Then, we briefly investigated the numerical and analytical methods commonly used to solve the pulse propagation equation, such as the split-step Fourier method, the Volterra series-based analysis, and the regular perturbation series-based analysis. Following that, we briefly discussed various DSP techniques available in the literature to deal with the detrimental effects of fiber nonlinearity. Based on the knowledge gained from the background study, we developed four different DSP techniques for the compensation of fiber nonlinearity in coherent

optical communication systems. We considered both the CO-OFDM-based and the single carrier-based optical transmission systems to evaluate the effectiveness of our developed DSP algorithms. It is important to mention that the first two techniques were designed for the CO-OFDM systems; whereas, the other two techniques were developed for the single-carrier systems.

In Chapter 3, we discussed the spectral efficiency problem associated with the phase-conjugated twin wave (PCTW) technique. On this ground, we developed two linear coding techniques, referred to as linear time/polarization-coded phase-conjugated twin signals (MOD-16-QAM-CDR/MOD-16-QAM-PCTW), to double the spectral efficiency of the PCTW technique. In this scheme, the data symbols on the adjacent subcarriers of the OFDM symbol are linearly combined, one at full amplitude and the other at half amplitude. The linearly coded data is then transmitted as phase conjugate pairs on the orthogonal dimensions (time or polarization). The nonlinear distortions added to these transmitted symbols are essentially anti-correlated since they carry phase conjugate pairs of data. At the receiver, the coherent superposition of the phase conjugate pairs eventually leads to the cancellation of the nonlinear distortions. In Chapter 3, we also investigated the impact of polarization-dependent loss (PDL) on the performance of MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques and demonstrate that the MOD-16-QAM-CDR shows a superior PDL tolerance when compared to the MOD-16-QAM-PCTW, regardless of the PDL model.

In Chapter 4, we discussed the background of the problem associated with the individual implementations of the single-channel digital back-propagation (SC-DBP), multi-channel (MC)-DBP and PCTW techniques to compensate for the intra- and inter-channel nonlinearity effects. In this chapter, we developed a low-complexity joint technique for fiber nonlinearity compensation, which combines the SC-DBP and PCTW techniques. This scheme provides a performance gain higher than applying

the SC-DBP and PCTW techniques individually. The SC-DBP technique compensates for the deterministic intra-channel nonlinearity, whereas, the PCTW technique handles both intra- and inter-channel nonlinearity effects. Thus, the overall performance improvement of the joint SC-DBP-PCTW technique comes from the two-stage compensation for the intra-channel nonlinearity and a first-order cancellation of the inter-channel nonlinear distortion. In addition to that, the developed SC-DBP-PCTW technique shows similar performance as MC-DBP implemented with 16 steps/span.

In Chapter 5, we developed an enhanced regular perturbation (ERP)-based nonlinearity compensation technique, referred to as ERP-NLC, to compensate for the fiber nonlinearity in a polarization-division multiplexed dispersion unmanaged optical communication system. We also developed a modified perturbation-based NLC (PB-NLC) technique by simple phase-rotation (PR) of the nonlinearity coefficient matrix, referred to as the PR-PB-NLC. The PR-PB-NLC can be considered as a by-product of the ERP-NLC technique. We carried out numerical simulations for the single-channel and the wavelength-division multiplexed (WDM) five-channel transmission system to verify the effectiveness of the ERP-NLC technique. Results showed that the ERP-NLC technique provides an improved NLC performance when compared to the electronic dispersion compensation (EDC) and the conventional PB-NLC technique for both single-channel and the WDM transmission systems. Also, the simulation results for the PR-PB-NLC technique for a single or five-channel transmission system showed an improved NLC performance when compared to the EDC and PB-NLC techniques. Finally, we showed that the obtained performance enhancement comes with a negligible increase in the computational complexity for the ERP-NLC and PR-PB-NLC techniques when compared to the PB-NLC technique.

In Chapter 6, we discussed the extension of the first-order (FO) perturbation theory to second-order (SO) and the implementation of SO perturbation-based NLC

(SO-PB-NLC) technique. We first discussed the implementation of the FO-PB-NLC technique to compensate for fiber nonlinearity. Following that, we presented a comprehensive theoretical analysis for the derivation of the SO nonlinear distortion field, which is the foundation for the SO-PB-NLC technique. Then, we investigated a few simplifying assumptions to reduce the implementation complexity of the SO-PB-NLC technique. Through numerical simulations, we showed that the SO-PB-NLC technique significantly enhances the NLC performance and the maximum transmission reach when compared to the FO-PB-NLC technique. In Chapter 6, the performance of the SO-PB-NLC technique is also compared with that of the benchmark DBP. Finally, we carried out the complexity analysis and showed that the performance enhancement of the SO-PB-NLC technique comes with a reduced implementation complexity when compared to the DBP technique with one step per span.

In a nutshell, this thesis contributed several low-complexity DSP solutions to combat the detrimental effects of fiber nonlinearity effects for both CO-OFDM and single-carrier optical long-haul transmission systems. Also, their low-complexity position them for consideration as suitable candidates for real-time implementation to combat the fiber nonlinearity effects.

7.2 Suggested Future Work

7.2.1 Extending SO-PB-NLC Technique to WDM Superchannel Systems

The optical WDM superchannel systems increase the spectral efficiency by closely packing the modulated carriers with minimal wasting of the optical spectrum. These closely packed carriers travel as a single entity from the same origin to the same

destination through the optical fiber link. However, the reduced spectral spacing of the subchannels causes the superchannel system highly vulnerable to the intra- and inter-channel fiber nonlinearity effects. In this context, the SO-PB-NLC technique developed in this thesis to compensate for the intra-channel fiber nonlinearity effect in a single-channel system can be extended by including the inter-channel fiber nonlinearity effects in a WDM superchannel system, to improve the transmission performance.

7.2.2 The Impact of PMD and Its Interplay with PDL on the Performance of Digital Phase Conjugation Techniques

PMD and PDL are two linear impairments that are encountered when dealing with long spans of single-mode fiber. PMD refers to the polarization effects of concatenated birefringent segments of the fiber. Each homogeneous segment produces differential group delay. PMD is generated when two or more differential-group delay segments are placed in cascade. The combination of PMD and PDL creates effects that are quite complicated, which can impair a communication system more than either effect alone. For example, PDL is generally wavelength-independent, while PMD is wavelength-dependent [18]. Addition of some PMD to PDL results in wavelength-dependent PDL. Also, PDL converts the real-valued PMD vector into a complex vector, giving rise to a loss of orthogonality between the principal states of polarization that scales with the link PDL. As both differential-group delays and PDL vary randomly in magnitude and orientation along the transmission link, a study on the impact of combined effects of PMD and PDL on the performance of digital phase conjugation techniques can be done using the statistical-PDL model.

7.2.3 Improve Spectral Efficiency of the SC-DBP-PCTW Technique

As discussed in Chapter 3, the PCTW technique transmits an additional phase-conjugated copy in one of the orthogonal dimensions and occupies the same amount of bandwidth as the original signal. Consequently, the implementation of the PCTW technique halves the spectral efficiency, and thereby the performance improvement of the SC-DBP-PCTW technique comes with a cost of spectral efficiency loss. Recently, we developed MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques to solve the spectral efficiency issue of the PCTW technique. Based on this, the SC-DBP-PCTW technique can be modified by replacing the PCTW scheme with the MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques, to improve the spectral efficiency.

7.2.4 Modifying the ERP-NLC Technique to Compensate for Both Intra- and Inter-Channel Fiber Nonlinearities

The main advantage of the perturbation theory-based nonlinearity compensation technique is the possibility of implementation on a single stage for the entire fiber link, which may significantly reduce the computational complexity in comparison with DBP and Volterra nonlinear equalizer. It can also be implemented with one sample per symbol. As discussed previously, the ERP method solves the energy divergence problem of the RP technique without a significant increase in the computational complexity when compared to the RP-based approximation. The ERP-NLC technique developed in this thesis only considers the intra-channel nonlinearity effects. In this context, the ERP-NLC technique can be modified to compensate for both intra- and inter-channel fiber nonlinearity effects in a WDM optical superchannel transmission system.

7.2.5 Deep Neural Network-Assisted PB-NLC Technique Including SO Quintuplet Pulses

The machine learning-based techniques provide the advantage of directly capturing the effects of fiber nonlinearity [113]- [116]. That is achieved by creating direct input-output relations between the observed and the desired outputs based on training data [117]- [121]. According to the recent works in the literature, deep neural networks (DNN) can be trained with the intra-channel XPM and intra-channel FWM triplets to estimate the nonlinear perturbation coefficients, instead of analytical computation, to implement the PB-NLC technique. However, existing studies are limited to the compensation of the intra-channel nonlinearity effect and only consider FO triplets in the input layer of DNN [122]. Based on this fact, the existing DNN-assisted PB-NLC technique can be modified by including the intra- and inter-channel SO quintuplets along with the FO triplets in the input layer of DNN, to improve the compensation performance.

References

[1] Global Network Performance, "Cisco annual internet report," *Cisco White Paper*, 2020.

[2] M. Chen, W. Saad, C. Yin, and M. Debbah, "Data correlation-aware resource management in wireless virtual reality (VR): an echo state transfer learning approach," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4267-4280, Jun. 2019.

[3] J. Zhang, S. Rajendran, Z. Sun, R. Woods, and L. Hanzo, "Physical layer security for the internet of things: authentication and key generation," *IEEE Wirel. Commun.*, vol. 26, no. 5, pp. 92-98, Oct. 2019.

[4] F. Restuccia, S. D'Oro, and T. Melodia, "Securing the internet of things in the age of machine learning and software-defined networking," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 4829-4842, Dec. 2018.

[5] A. Amari, O. A. Dobre, R. Venkatesan, O. S. Sunish Kumar, P. Ciblat, and Y. Jaouën, "A survey on fiber nonlinearity compensation for 400 Gbps and beyond optical communication systems," *IEEE Commun. Surv. Tutorials*, vol. 19, no. 4, pp. 3097-3113, 4th Quart. 2017.

[6] C. Xu, X. Liu, and X. Wei, "Differential phase-shift keying for high spectral efficiency optical transmissions," *IEEE J. Sel. Topics Quantum Electron.*, vol. 10, no. 2, pp. 281-293, Mar. 2004.

[7] C. Laperle, B. Villeneuve, Z. Zhang, and D. McGhan, "WDM performance and PMD tolerance of a coherent 40-Gbit/s dual-polarization QPSK transceiver," *J. Lightw. Technol.*, vol. 26, no. 1, pp. 168-175, Jan. 2008.

[8] D. Soma, Y. Wakayama, S. Beppu, S. Sumita, T. Tsuritani, T. Hayashi, T. Nagashima, M. Suzuki, M. Yoshida, K. Kasai, M. Nakazawa, H. Takahashi, K. Igarashi, I. Morita, and M. Suzuki, "10.16-peta-B/s dense SDM/WDM Transmission over 6-mode 19-core fiber across the C+L band," *J. Lightw. Technol.*, vol. 36, no. 6, pp. 1362-1368, Jan. 2018.

[9] R. Rios-Müller, J. Renaudier, P. Brindel, A. Ghazisaeidi, I. Fernandez, P. Tran, C. Simonneau, L. Schmalen, and G. Charlet, "Spectrally-efficient 400Gb/s single carrier transport over 7200 km," *J. Lightw. Technol.*, vol. 33, no. 7, pp. 1402–1407, Apr. 2015.

[10] R. Rios-Müller, J. Renaudier, P. Brindel, P. Jennevé, H. Mardoyan, L. Schmalen, and G. Charlet, "Experimental comparison between superchannel and sub-band single-carrier for 400 Gb/s and 800 Gb/s transport," in *Proc. ECOC*, Valencia, Spain, Sep. 2015, pp. 1-3.

[11] S. Chandrasekhar and X. Liu, "OFDM based superchannel transmission technology," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3816-3823, Dec. 2012.

[12] Z. Xia, S. Cui, C. Ke, S. Fu, M. Tang, and D. Liu, "Fast and robust chromatic dispersion estimation for digital optical coherent receivers," in *Proc. OFC*, Anaheim, CA, USA, Mar. 2016, paper Th2A.52.

[13] R. I. Killey, P. M. Watts, V. Mikhailov, M. Glick, and P. Bayvel, "Electronic dispersion compensation by signal predistortion using digital processing and a

dual-drive Mach-Zehnder modulator," *IEEE Photon. Technol. Lett.*, vol. 17, no. 3, pp. 714-716, Mar. 2005.

[14] M. Kuschnerov, F. N. Hauske, K. Piyawanno, B. Spinnler, M. S. Alfiad, A. Napoli, and B. Lankl, "DSP for coherent single-carrier receivers," *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3614-3622, Aug. 2009.

[15] J. Renaudier, O. Bertran-Pardo, G. Charlet, M. Salsi, H. Mardoyan, P. Tran, and S. Bigo, "8 Tb-s long haul transmission over low dispersion fibers using 100 Gbps PDM-QPSK channels paired with coherent detection," *Bell Lab. Tech. J.*, vol. 14, no. 4, pp. 27-45, Feb. 2010.

[16] S. J. Savory, "Optimum electronic dispersion compensation strategies for nonlinear transmission," *Electron. Lett.*, vol. 42, no. 7, pp. 407-408, Apr. 2006.

[17] G. P. Agrawal, *Nonlinear Fiber Optics*, Springer, 2000.

[18] A. O. D. Forno, A. Paradisi, R. Passy, and J. P. von der Weid, "Experimental and theoretical modeling of polarization-mode dispersion in single-mode fibers," *IEEE Photon. Technol. Lett.*, vol. 12, no. 3, pp. 296-298, Mar. 2000.

[19] Z. Yu, X. Yi, J. Zhang, D. Zhao, and K. Qiu, "Experimental demonstration of polarization-dependent loss monitoring and compensation in stokes space for coherent optical PDM-OFDM," *J. Lightw. Technol.*, vol. 32, no. 23, pp. 4528-4533, Dec. 2014.

[20] A. Mecozzi and M. Shtaif, "The statistics of polarization-dependent loss in optical communication systems," *IEEE Photon. Technol. Lett.*, vol. 14, no. 3, pp. 313-315, Mar. 2002.

[21] S. Ten and M. Edwards, "An introduction to the fundamentals of PMD in fibers," *Corning White Paper*, 2006.

[22] B. Huttner, C. Geiser, and N. Gisin, "Polarization-induced distortions in optical fiber networks with polarization-mode dispersion and polarization-dependent losses," *IEEE J. Sel. Topics Quantum Electron.*, vol. 6, no. 2, pp. 317-329, Mar. 2000.

[23] I. Fatadin, D. Ives, and S. J. Savory, "Blind equalization and carrier phase recovery in a 16-QAM optical coherent system," *J. Lightw. Technol.*, vol. 27, no. 15, pp. 3042-3049, Aug. 2009.

[24] H. Louchet, K. Kuzmin, and A. Richter, "Improved DSP algorithms for coherent 16-QAM transmission," *Proc. ECOC*, Brussels, Belgium, Sep. 2008, pp. 1-3.

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

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

[27] D. G. Foursa, O. V. Sinkin, A. Lucero, J. X. Cai, G. Mohs, and A. Pilipetskii, "Nonlinear interaction between signal and amplified spontaneous emission in coherent systems," in *Proc. OFC*, Anaheim, CA, USA, Mar. 2013, paper JTh2A.35.

[28] A. Ghazisaeidi, "A theory of nonlinear interactions between signal and amplified spontaneous emission noise in coherent wavelength division multiplexed systems," *J. Lightw. Technol.*, vol. 35, no. 23, pp. 5150 - 5175, Dec. 2017.

[29] E. Ip, A. P. T. Lau, D. J. F. Barros, and J. M. Kahn, "Coherent detection in optical fiber systems," *Opt. Express*, vol. 16, no. 2, pp. 753-791, Jan. 2008.

[30] L. B. Du, D. Rafique, A. Napoli, B. Spinnler, A. D. Ellis, M. Kuschnerov, and A. J. Lowery, "Digital fiber nonlinearity compensation," *IEEE Signal Process. Mag.*, vol. 31, no. 2, pp. 46-56, Mar. 2014.

[31] D. Rafique, "Fiber nonlinearity compensation: commercial applications and complexity analysis," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 544-553, Jan. 2016.

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

[33] M. Ding, D. Fan, W. Wang, Y. Luo, and G. D. Peng, *Basics of Optical Fiber Measurements.* Springer, Singapore, 2018.

[34] *Understanding and measuring chromatic dispersion*, Fiber Optics Online: The Information Source for the Fiber Optic Communications Industry, 2000. [Online]. Available: https://www.fiberopticsonline.com/doc/understanding-and-measuring-chromatic-dispers-0002

[35] S. Kumar, and M. Jamal, *Fiber Optic Communications: Fundamentals and Applications.* Wiley, 2014.

[36] F. P. Guiomar, J. D. Reis, A. L. Teixeira, and A. N. Pinto, "Digital postcompensation using Volterra series transfer function," *IEEE Photon. Technol. Lett.*, vol. 23, no. 19, pp. 1412-1414, Oct. 2011.

[37] A. Amari, O. A. Dobre, and R. Venkatesan, "Fifth-order Volterra series based nonlinear equalizer for long-haul high data rate optical fiber communications," in *Proc. IEEE ICTON*, Girona, Spain, Jul. 2017, paper We.A1.2.

[38] V. Vgenopoulou, A. Amari, M. Song, E. Pincemin, I. Roudas, and Y. Jaouën, "Volterra-based nonlinear compensation in 400 Gb/s WDM multiband coherent optical OFDM systems," in *Proc. ACPC*, Shanghai, China, Nov. 2014, paper AF1E.4.

[39] A. Bakhshali, W. Y. Chan, J. C. Cartledge, M. OSullivan, C. Laperle, A. Borowiec, and K. Roberts, "Frequency-domain Volterra-based equalization structures for efficient mitigation of intrachannel Kerr nonlinearities," *J. Lightw. Technol.*, vol. 34, no. 8, pp. 1770-1777, Apr. 2016.

[40] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems.* Wiley, New Jersey, 1980.

[41] K. V. Peddanarappagari and M. Brandt-Pearce, "Volterra series transfer function of single-mode fibers," *J. Lightw. Technol.*, vol. 15, no. 12, pp. 2232-2241, Dec. 1997.

[42] N. P. Diamantopoulos, H. Nishi, W. Kobayashi, K. Takeda, T. Kakitsuka, and S. Matsuo, "On the complexity reduction of the second-order Volterra nonlinear equalizer for IM/DD systems," *J. Lightw. Technol.*, vol. 37, no. 4, pp. 1214-1224, Feb. 2019.

[43] A. Mecozzi, C. B. Clausen, and M. Shtaif, "Analysis of intrachannel nonlinear effects in highly dispersed optical pulse transmission," *IEEE Photon. Technol. Lett.*, vol. 12, no. 4, pp. 392-394, Apr. 2000.

[44] 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.

[45] Y. Gao, J. C. Cartledge, A. S. Karar, S. S. H. Yam, M. O'Sullivan, C. Laperle, A. Borowiec, and K. Roberts, "Reducing the complexity of perturbation based nonlinearity pre-compensation using symmetric EDC and pulse shaping," *Opt. Exp.*, vol. 22, no. 2, pp. 1209-1219, Jan. 2014.

[46] 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.

[47] O. S. Sunish Kumar, A. Amari, O. A. Dobre, and R. Venkatesan, "Enhanced regular perturbation-based nonlinearity compensation technique for optical transmission systems," *IEEE Photon. J.*, vol. 11, no. 4, pp. 1-13, Aug. 2019.

[48] S. L. Jansen, D. van den Borne, B. Spinnler, S. Calabrò, H. Suche, P. M. Krummrich, W. Sohler, G. D. Khoe, and H. de Waardt, "Optical phase conjugation for ultra long-haul phase-shift-keyed transmission," *J. Lightw. Technol.*, vol. 24, no. 1, pp. 54-64, Jan. 2006.

[49] 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," *Nat. Photon.*, vol. 7, pp. 560-568, May 2013.

[50] X. Liu, S. Chandrasekhar, and P. J. Winzer, "Phase-conjugated twin waves and fiber nonlinearity compensation," in *Proc. IEEE OECC*, Melbourne, VIC, Australia, Jul. 2014, pp. 938-940.

[51] H. Eliasson, P. Johannisson, M. Karlsson, and P. A. Andrekson, "Mitigation of nonlinearities using conjugate data repetition," *Opt. Exp.*, vol. 23, no. 3, pp. 2392-2402, Jan. 2015.

[52] X. Yi, X. Chen, D. Sharma, C. Li, M. Luo, Q. Yang, Z. Li, and K. Qiu, "Digital coherent superposition of optical OFDM subcarrier pairs with Hermitian symmetry for phase noise mitigation," *Opt. Exp.*, vol. 22, no. 11, pp. 13454-13459, May 2015.

[53] S. T. Le, M. E. McCarthy, N. M. Suibhne, A. D. Ellis, and S. K. Turitsyn, "Phase-conjugated pilots for fiber nonlinearity compensation in CO-OFDM transmission," *J. Lightw. Technol.*, vol. 33, no. 7, pp. 1308-1314, Apr. 2015.

[54] T. Yoshida, T. Sugihara, K. Ishida, and T. Mizuochi, "Spectrally efficient dual phase-conjugate twin waves with orthogonally multiplexed quadrature pulse-shaped signals," in *Proc. OFC*, San Francisco, CA, USA, Mar. 2014, paper M3C.6.

[55] S. T. Le, M. E. McCarthy, N. M. Suibhne, M. A. Z. Al-Khateeb, E. Giacoumidis, N. Doran, A. D. Ellis, and S. K. Turitsyn, "Demonstration of phase-conjugated subcarrier coding for fiber nonlinearity compensation in CO-OFDM transmission," *J. Lightw. Technol.*, vol. 33, no. 11, pp. 2206-2212, Jun. 2015.

[56] A. Ellis and M. Sorokina, *Optical Communication Systems: Limits and Possibilities*, Jenny Stanford Publishing, Singapore, 2019.

[57] X. Liu, S. Chandrasekhar, P. J. Winzer, R. W. Tkach, and A. R. Chraplyvy, "406.6-Gb/s PDM-BPSK superchannel transmission over 12,800-km TWRS fiber via nonlinear noise squeezing," in *Proc. OFC*, Anaheim, CA, USA, Mar. 2013, paper PDP5B.10.

[58] O. S. Sunish Kumar, O. A. Dobre, R. Venkatesan, S. K. Wilson, O. Omomukuyo, A. Amari, and D. Chang, "A spectrally-efficient linear polarization coding scheme

for fiber nonlinearity compensation in CO-OFDM systems," in *Proc. SPIE Opt.*, San Francisco, CA, USA, Jan. 2017, pp. 152-161.

[59] O. V. Sinkin, R. Holzlohner, J. Zweck, and C. R. Menyuk, "Optimization of the split-step Fourier method in modeling optical fiber communications systems," *J. Lightw. Technol.*, vol. 21, no. 1, pp. 61-68, Jan. 2003.

[60] D. Rafique, M. Mussolin, M. Forzati, J. Martensson, M. N. Chugtai, and A. D. Ellis, "Compensation of intra-channel nonlinear fibre impairments using simplified digital back-propagation algorithm," *Opt. Exp.*, vol. 19, no. 10, pp. 9453-9460, May 2011.

[61] 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," *Opt. Exp.*, vol. 16, no. 2, pp. 880-888, Jan. 2008.

[62] D. Rafique, M. Mussolin, M. Forzati, J. Martensson, M. N. Chugtai, and A. D. Ellis, "Modified split-step Fourier method for compensation of nonlinear fibre impairments," in *Proc. IEEE ICTON*, Stockholm, Sweden, Jun. 2011, paper Tu.P.5.

[63] L. Li, Z. Tao, L. Dou, W. Yan, S. Oda, T. Tanimura, T. Hoshida, and J. C. Rasmussen, "Implementation efficient nonlinear equalizer based on correlated digital backpropagation," in *Proc. OFC*, Los Angeles, CA, USA, Mar. 2011, paper OWW3.

[64] 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.

[65] G. Liga, T. Xu, A. Alvarado, R. I. Killey, and P. Bayvel, "On the performance of multichannel digital backpropagation in high-capacity long-haul optical transmission," *Opt. Exp.*, vol. 22, no. 24, pp. 30053-30062, Dec. 2014.

[66] N. K. Fontaine, "Spectrally-sliced coherent receivers for THz bandwidth optical communications," in *Proc. ECOC*, ICC London ExCeL, UK, Sep. 2013, pp. 1-3.

[67] R. Maher, D. Lavery, D. Millar, A. Alvarado, K. Parsons, R. Killey, and P. Bayvel, "Reach enhancement of 100% for a DP-64QAM super-channel using MC-DBP," in *Proc. OFC*, Los Angeles, CA, USA, Mar. 2015, paper Th4D.5.

[68] 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," *J. Lightw. Technol.*, vol. 34, no. 8, pp. 1896-1907, Apr. 2016.

[69] F. Zhang, Q. Zhuge, M. Qiu, W. Wang, M. Chagnon, and D. V. Plant, "XPM model-based digital backpropagation for subcarrier-multiplexing systems," *J. Lightw. Technol.*, vol. 33, no. 24, pp. 5140-5150, Dec. 2015.

[70] X. Liu, S. Chandrasekhar, P. J. Winzer, R. W. Tkach, and A. R. Chraplyvy, "Fiber-nonlinearity-tolerant superchannel transmission via nonlinear noise squeezing and generalized phase-conjugated twin waves," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 766-775, Feb. 2014.

[71] I. Fernandez, A. Ghazisaeidi, E. Awwad, P. Tran, and G. Charlet, "Polarization effects in nonlinearity compensated links," in *Proc. ECOC*, Dusseldorf, Germany, Sep. 2016, pp. 1-3.

[72] C. Xie, "Impact of nonlinear and polarization effects in coherent systems," *Opt. Exp.*, vol. 19, no. 26, pp. 915-930, Dec. 2011.

[73] X. Liang and S. Kumar, "Multi-stage perturbation theory for compensating intra-channel nonlinear impairments in fiber-optic links," *Opt. Express*, vol. 22, no. 24, pp. 29733-29745, Nov. 2014.

[74] E. Collett, *Polarized Light: Fundamentals and Applications*, Marcel Dekker, New York, 1992.

[75] D. Wang and C. R. Menyuk, "Calculation of penalties due to polarization effects in a long-haul WDM system using a stokes parameter model," *J. Lightw. Technol.*, vol. 19, no. 4, pp. 487-494, Apr. 2001.

[76] O. Vassilieva, I. Kim, Y. Akasaka, M. Bouda, and M. Sekiya, "Interplay between PDL and nonlinear effects in coherent polarization multiplexed systems," *Opt. Exp.*, vol. 19, no. 26, pp. 357-362, Dec. 2011.

[77] O. S. Sunish Kumar, A. Amari, O. A. Dobre, and R. Venkatesan, "PDL impact on linearly coded digital phase conjugation techniques in CO-OFDM systems," *IEEE Photon. Technol. Lett.*, vol. 30, no. 9, pp. 769-772, May 2018.

[78] N. Rossi, P. Serena, and A. Bononi, "Polarization-dependent loss impact on coherent optical systems in presence of fiber nonlinearity," *IEEE Photon. Technol. Lett.*, vol. 26, no. 4, pp. 334-337, Feb. 2014.

[79] X. Li, A. Alphones, W. D. Zhong, and C. Yu, "Investigation of PMD in direct-detection optical OFDM with zero padding," *Opt. Exp.*, vol. 21, no. 18, pp. 20851-20856, Aug. 2013.

[80] S. T. Le, T. Kanesan, E. Giacoumidis, N. J. Doran, and A. D. Ellis, "Quasi-pilot aided phase noise estimation for coherent optical OFDM systems," *IEEE Photon. Technol. Lett.*, vol. 26, no. 5, pp. 504-507, Mar. 2014.

[81] T. Umeki, T. Kazama, A. Sano, K. Shibahara, K. Suzuki, M. Abe, H. Takenouchi, and Y. Miyamoto, "Simultaneous nonlinearity mitigation in 92×180-Gbit/s PDM-16QAM transmission over 3840 km using PPLN-based guard-bandless optical phase conjugation," *Opt. Exp.*, vol. 24, no. 15, pp. 16945-16951, Jul. 2016.

[82] X. Liang and S. Kumar, "Correlated digital back propagation based on perturbation theory," *Opt. Express*, vol. 23, no. 11, pp. 14655-14665, Jun. 2015.

[83] O. Vassilieva, I. Kim, and T. Ikeuchi, "Enabling technologies for fiber nonlinearity mitigation in high capacity transmission systems," *J. Lightw. Technol.*, vol. 37, no. 1, pp. 50-60, Jan. 2019.

[84] T. T. Nguyen, P. Harper, O. S. Sunish Kumar, and A. Ellis, "Nonlinear tolerance enhancement based on perturbation theory for optical phase conjugation systems," in *Proc. OFC*, San Diego, CA, USA, Mar. 2020, paper Th2A.52.

[85] E. P. Silva, M. P. Yankov, F. D. Ros, T. Morioka, and L. K. Oxenløwe, "Perturbation-based FEC-assisted iterative nonlinearity compensation for WDM systems," *J. Lightw. Technol.*, vol. 37, no. 3, pp. 875-881, Feb. 2019.

[86] A. Ghazisaeidi, J. Renaudier, P. Tran, and G. Charlet, "Perturbative nonlinear pre-compensation in presence of optical filtering," in *Proc. OFC*, Los Angeles, CA, USA, Mar. 2015, paper ThD3.3.

[87] T. Oyama, H. Nakashima, S. Oda, T. Yamauchi, Z. Tao, T. Hoshida, and J. C. Rasmussen, "Robust and efficient receiver-side compensation method for intra-channel nonlinear effects," in *Proc. OFC*, San Francisco, CA, USA, Mar. 2014, paper Tu3A.3.

[88] Y. Gao, A. S. Karar, J. C. Cartledge, S. S. H. Yam, M. O'Sullivan, C. Laperle, A. Borowiec, and K. Roberts, "Joint pre-compensation and selective post-compensation for fiber nonlinearities," *IEEE Photon. Technol. Lett.*, vol. 26, no. 17, pp. 1746-1749, Jun. 2014.

[89] P. Serena and A. Bononi, "An alternative approach to the Gaussian noise model and its system implications," *J. Lightw. Technol.*, vol. 31, no. 22, pp. 3489-3499, Nov. 2013.

[90] B. R. Washburn, "Dispersion and nonlinearities associated with supercontinuum generation in microstructure fibers," Ph.D thesis, Georgia Institute of Technology, Atlanta, GA, 2005.

[91] E. Mateo, L. Zhu, and G. Li, "Impact of XPM and FWM on the digital implementation of impairment compensation for WDM transmission using backward propagation," *Opt. Express*, vol. 16, no. 20, pp. 16124-16137, Sep. 2008.

[92] O. S. Sunish Kumar, A. Amari, O. A. Dobre, R. Venkatesan, and S. K. Wilson, "A joint technique for nonlinearity compensation in CO-OFDM superchannel systems," in *Proc. ACPC*, Guangzhou, China, Nov. 2017, paper Su4B.4.

[93] M. Malekiha, I. Tselniker, and D. V. Plant, "Efficient nonlinear equalizer for intra-channel nonlinearity compensation for next generation agile and dynamically reconfigurable optical networks," *Opt. Express*, vol. 24, no. 4, pp. 4097-4108, Feb. 2016.

[94] J. C. Cartledge, F. P. Guiomar, F. R. Kschischang, G. Liga, and M. P. Yankov, "Digital signal processing for fiber nonlinearities," *Opt. Express*, vol. 25, no. 3, pp. 1916-1936, Feb. 2017.

[95] A. Amari, S. Goossens, Y. C. Gültekin, O. Vassilieva, I. Kim, T. Ikeuchi, C. M. Okonkwo, F. M. J. Willems, and A. Alvarado, "Introducing enumerative sphere shaping for optical communication systems with short blocklengths," *J. Lightw. Technol.*, vol. 37, no. 23, pp. 5926-5936, Dec. 2019.

[96] 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.

[97] C. B. Czegledi, G. Liga, D. Lavery, M. Karlsson, E. Agrell, S. J. Savory, and P. Bayvel, "Digital backpropagation accounting for polarization-mode dispersion," *Opt. Express*, vol. 25, no. 3, pp. 1903-1915, Feb. 2017.

[98] C. A. Hall and T. A. Porsching, *Numerical Analysis of Partial Differential Equations.* Prentice Hall, New Jersey, 1990.

[99] S. Okamoto, K. Horikoshi, M. Nakamura, A. Matsushita, F. Hamaoka, and Y. Kisaka, "Low complexity sub-band digital back propagation," in *Proc. OFC*, San Diego, CA, USA, Mar. 2019, paper W2A.49.

[100] C. S. Martins, L. Bertignono, A. Nespola, A. Carena, F. P. Guiomar, and A. N. Pinto, "Low-complexity time-domain DBP based on random step-size and partitioned quantization," *J. Lightw. Technol.*, vol. 36, no. 14, pp. 2888-2895, Jul. 2018.

[101] Z. Xiao, Q. Zhuge, S. Fu, F. Zhang, M. Qiu, M. Tang, D. Liu, and D. V. Plant, "Low complexity split digital backpropagation for digital subcarrier-multiplexing optical transmissions," *Opt. Express*, vol. 25, no. 22, pp. 27824-27833, Oct. 2017.

[102] F. Zhang, Q. Zhuge, M. Qiu, and D. V. Plant, "Low complexity digital back-propagation for high baud subcarrier-multiplexing systems," *Opt. Express*, vol. 24, no. 15, pp. 17027-17040, Jul. 2016.

[103] L. Dou, Z. Tao, W. Yan, L. Li, T. Hoshida, and J. C. Rasmussen, "Pre-distortion method for intra-channel nonlinearity compensation with phase-rotated perturbation term," in *Proc. OFC*, Los Angeles, CA, USA, Mar. 2012, paper OTh3C.2.

[104] A. Redyuk, E. Averyanov, O. Sidelnikov, M. Fedoruk, and S. Turitsyn, "Compensation of nonlinear impairments using inverse perturbation theory with reduced complexity," *J. Lightw. Technol.*, vol. 38, no. 6, pp. 1250-1257, Mar. 2020.

[105] O. S. Sunish Kumar, A. Amari, O. A. Dobre, and R. Venkatesan, "Intra-channel nonlinearity compensation based on second-order perturbation theory," arXiv:2005.01191v1 [cs.IT], May 2020.

[106] Y. Li, H. Minn, and R. M. A. P. Rajatheva, "Synchronization, channel estimation, and equalization in MB-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4341-4352, Nov. 2008.

[107] A. N. Pinto and G. P. Agrawal, "Nonlinear interaction between signal and noise in optical fibers," *J. Lightw. Technol.*, vol. 26, no. 13, pp. 1847-1853, Jul. 2008.

[108] A. Ghazisaeidi and R. J. Essiambre, "Calculation of coefficients of perturbative nonlinear pre-compensation for Nyquist pulses," in *Proc. ECOC*, Cannes, France, Sep. 2014, paper We.1.3.3.

[109] T. Oyama, H. Nakashima, T. Hoshida, Z. Tao, C. Ohshima, and J. C. Rasmussen, "Efficient transmitter-side nonlinear equalizer for 16QAM," in *Proc. ECOC*, London, UK, Sep. 2013, paper We.3.C.1.

[110] Z. Tao, L. Dou, W. Yan, Y. Fan, L. Li, S. Oda, Y. Akiyama, H. Nakashima, T. Hoshida, and J. C. Rasmussen, "Complexity-reduced digital nonlinear compensation for coherent optical systems," in *Proc. SPIE Opt.*, San Francisco, CA, USA, Jan. 2013, pp. 115-125.

[111] L. Bomin, "Forward error correcting codes for 100 Gbit/s optical communication systems," Ph.D thesis, Technical University of Denmark, Denmark, 2014.

[112] L. B. Du, M. M. Morshed, and A. J. Lowery, "Fiber nonlinearity compensation for OFDM super-channels using optical phase conjugation," *Opt. Express*, vol. 20, no. 18, pp. 19921-19927, Aug. 2012.

[113] F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini, and M. Tornatore, "An overview on application of machine learning techniques in optical networks," arXiv: 1803.07976v4 [cs. NI], Dec. 2018.

[114] D. Zibar, J. Thrane, J. Wass, R. Jones, M. Piels, and C. Schaeffer, "Machine learning techniques applied to system characterization and equalization," in *Proc. OFC*, Anaheim, CA, USA, Mar. 2016, paper Tu3K.1.

[115] T. S. R. Shen and A. P. T. Lau, "Fiber nonlinearity compensation using extreme learning machine for DSP-based coherent communication systems," in *Proc. OECC*, Kaohsiung, Taiwan, Jul. 2011, pp. 816-817.

[116] D. Wang, M. Zhang, Z. Li, Y. Cui, J. Liu, Y. Yang, and H. Wang, "Nonlinear decision boundary created by a machine learning-based classifier to mitigate nonlinear phase noise," in *Proc. ECOC*, Valencia, Spain, Sep. 2015, pp. 1-3.

[117] R. Borkowski, A. Doberstein, H. Haisch, and D. Zibar, "Single-step emulation of nonlinear fiber-optic link with Gaussian mixture model," in *Proc. SPPCom*, Boston, Massachusetts, USA, Jun. 2015, paper SpS2C.6.

[118] C. Häger and H. D. Pfister, "Deep learning of the nonlinear Schrödinger equation in fiber-optic communications," arXiv: 1804.0299v1 [cs. IT], Apr. 2018.

[119] C. Y. Chuang, L. C. Liu, C. C. Wei, J. J. Liu, L. Henrickson, W. J. Huang, C. L. Wang, Y. K. Chen, and J. Chen, "Convolutional neural network based nonlinear classifier for 112-Gbps high speed optical link," in *Proc. OFC*, San Diego, CA, USA, Mar. 2018, paper W3A.4.

[120] V. Kamalov, L. Jovanovski, V. Vusirikala, S. Zhang, F. Yaman, K. Nakamura, T. Inoue, E. Mateo, and Y. Inada, "Evolution from 8QAM live traffic to PS 64-QAM with neural-network based nonlinearity compensation on 11000 km open subsea cable," in *Proc. ECOC*, Roma, Italy, Sep. 2018, pp. 1-3.

[121] S. Zhang, F. Yaman, E. Mateo, and Y. Inada, "Neuron-network-based nonlinearity compensation algorithm," in *Proc. ECOC*, Roma, Italy, Sep. 2018, pp. 1-3.

[122] Y. Gao, Z. A. El-Sahn, A. Awadalla, D. Yao, H. Sun, P. Mertz, and K. T. Wu, "Reduced complexity nonlinearity compensation via principal component analysis and deep neural networks," in *Proc. OFC*, San Diego, CA, USA, Mar. 2019, paper Th2A.49.

Appendix A

Calculation of the First-Order Perturbative Nonlinear Distortion Field Added to the OFDM Signal

The exact solution of (3.7) is not known in analytical form. The RP method provides an analytical approximate solution to (3.7), starting from the linear solution for $\gamma = 0$ and asymptotically converges to the exact solution for $\gamma \neq 0$. According to the RP method, the optical field complex envelop $\hat{u}(z,t)$ is expanded as a power series in γ as:

$$\hat{u}(z,t) = \sum_{r=0}^{\infty} \gamma^r \hat{u_r}(z,t). \tag{A.1}$$

Substituting (A.1) in (3.7), we obtain:

$$\sum_{r=0}^{\infty} \gamma^r \frac{\partial \hat{u_r}(z,t)}{\partial z} = \sum_{r=0}^{\infty} \gamma^r j \frac{\beta_2}{2} \frac{\partial^2 \hat{u_r}(z,t)}{\partial t^2} - j\gamma \sum_{l'=0}^{\infty} \sum_{m'=0}^{\infty} \sum_{n'=0}^{\infty} \gamma^{l'+m'+n'} \hat{u}_{l'}(z,t) \hat{u}^*_{m'}(z,t) \hat{u}_{n'}(z,t) \exp(-\alpha z), \tag{A.2}$$

where $*$ represents the complex conjugate.

The zeroth-order or the linear solution is obtained by equating the coefficients of γ with $r = 0$ on either side of (A.2) as:

$$\frac{\partial \hat{u_0}(z,t)}{\partial z} = j\frac{\beta_2}{2}\frac{\partial^2 \hat{u_0}(z,t)}{\partial t^2}. \tag{A.3}$$

The solution of (A.3) for a subcarrier w_k can be represented in frequency-domain as:

$$\hat{U}_0(z, w_k) = \hat{U}(0, w_k)\exp(-j\frac{\beta_2}{2}w_k^2 z), \tag{A.4}$$

where the subscript 0 represents the zeroth-order (linear) solution and $\hat{U}(0, w_k)$ is the Fourier transform of $\hat{u}(0,t) = u(t)$.

After down-conversion to the baseband and sampled at $t = nT_s$, the zeroth-order solution in frequency-domain can be represented as:

$$Y_0(z, w_k) = a_k \exp(-j\frac{\beta_2}{2}w_k^2 z). \tag{A.5}$$

For $r = 1$, the only contribution to the triple summation in (A.2) becomes ($l^{'} = m^{'} = n^{'} = 0$). The corresponding linear differential equation for the first-order correction is given by:

$$\frac{\partial \hat{u_1}(z,t)}{\partial z} = j\frac{\beta_2}{2}\frac{\partial^2 \hat{u_1}(z,t)}{\partial t^2} - j\left|\hat{u_0}(z,t)\right|^2 \hat{u_0}(z,t)\exp(-\alpha z). \tag{A.6}$$

The differential equation in (A.6) can be further simplified as:

$$\frac{\partial \hat{u_1}(z,t)}{\partial z} = j\frac{\beta_2}{2}\frac{\partial^2 \hat{u_1}(z,t)}{\partial t^2} - jf(z,T), \tag{A.7}$$

where

$$f(z,t) = \left|\hat{u_0}(z,t)\right|^2 \hat{u_0}(z,t)\exp(-\alpha z). \tag{A.8}$$

Taking the Fourier transform of (A.7), we get for the kth subcarrier frequency as :

$$\frac{\partial \hat{U}_1(z, w_k)}{\partial z} = -j\frac{\beta_2}{2}k^2\hat{U}_1(z, w_k) - jF(z, w_k), \tag{A.9}$$

where $F(z, w_k)$ is the Fourier transform of $f(z, t)$.

The solution for (A.9) can be represented as:

$$\hat{U}_1(z, w_k) \triangleq \hat{G}(z, w_k)\exp(-j\frac{\beta_2}{2}w_k^2 z), \tag{A.10}$$

where $\hat{G}(z, w_k)$ is the unknown distorted field envelop function, which can be derived from (A.10) as:

$$\hat{G}(z, w_k) = \hat{U}_1(z, w_k)\exp(j\frac{\beta_2}{2}w_k^2 z). \tag{A.11}$$

The evolution of $\hat{G}(z, k)$ along the optical fiber can be obtained by differentiating (A.11) with respect to z as:

$$\frac{\partial \hat{G}(z, w_k)}{\partial z} = -j\exp(j\frac{\beta_2}{2}w_k^2 z)F(z, w_k). \tag{A.12}$$

Integrating (A.12) from 0 to z and substituting in (A.10), yields:

$$\hat{U}_1(z, w_k) = \exp(-j\frac{\beta_2}{2}w_k^2 z)(-j)\int_0^z \exp(j\frac{\beta_2}{2}w_k^2 z')F(z', w_k)dz'. \tag{A.13}$$

By taking the inverse Fourier transform of (A.13) and substituting the value for $f(z, t)$, we get:

$$\hat{u}_1(z, t) = \left[-j\int_0^z \exp(j\frac{\beta_2}{2}w_k^2 z')\left|\hat{u_0}(z', t)\right|^2 \hat{u_0}(z', t)\exp(-\alpha z')dz'\right] \circledast h(z, t), \tag{A.14}$$

where we can represent

$$\hat{g}(z,t) = \int_{-\infty}^{\infty} \left(-j \int_0^z \exp(j\frac{\beta_2}{2} w_k^2 z') \left| \hat{u_0}(z',t) \right|^2 \right.$$
$$\left. \times \hat{u_0}(z',t) \exp(-\alpha z') dz' \right) \exp(jwt)\ dw. \quad \text{(A.15)}$$

The field in (A.13) can be further modified by substituting the value for $F(z, w_k)$ and applying the phase matching condition for the four-wave mixing as:

$$\hat{U_1}(z,w_k) = \exp(-j\frac{\beta_2}{2} w_k^2 z)(-j) \int_0^z \exp(j\frac{\beta_2}{2} w_k^2 z') \left(\sum_{m=-N/2}^{N/2-1} \sum_{n=-N/2}^{N/2-1} \hat{U_0}(z', w_k + w_m) \right.$$
$$\left. \times \hat{U_0^*}(z', w_k + w_n) \hat{U_0}(z', w_k + w_m + w_n) \exp(-\alpha z') \right) dz'. \quad \text{(A.16)}$$

Substituting the linear solution in (A.5) into (A.16) and down-converting to the base-band and sampled at $t = nT_s$, the first-order field in frequency-domain can be represented after some algebra as:

$$Y_1(z,w_k) = \exp(-j\frac{\beta_2}{2} w_k^2 z) G(z,w_k), \quad \text{(A.17)}$$

where

$$G(z,w_k) = (-j) \sum_{m=-N/2}^{N/2-1} \sum_{n=-N/2}^{N/2-1} a_{k+m} a_{k+n}^* a_{k+m+n} \Xi_{m,n}^{FO}, \quad \text{(A.18)}$$

and

$$\Xi_{m,n}^{FO} = \frac{1}{L} \int_0^L \exp(-A(z') - j\frac{1}{2} C(z')(w_m - w_k)(w_m - w_n)) dz', \quad \text{(A.19)}$$

where $A(z') = \int_0^{z'} \alpha(\xi)\, d\xi$ and $C(z') = \int_0^{z'} \beta_2(\xi)\, d\xi$ for a realistic transmission link. By separating the terms for self-phase modulation (SPM), intra-channel cross-phase

modulation (IXPM) and intra-channel four-wave mixing (IFWM)* from $G(L, w_k)$, and substituting the value for β_2, (A.18) can be modified as:

$$\begin{aligned}
\hat{a}_{1,k} &= (-j)\gamma \exp(j\varphi_D(w_k)L) \left|a_k\right|^2 a_k \Xi_{0,0}^{FO} + 2(-j)\gamma \exp(j\varphi_D(w_k)L) \sum_{n\neq 0} \left|a_{k+n}\right|^2 a_k \Xi_{0,n}^{FO} \\
&\qquad + (-j)\gamma \exp(j\varphi_D(w_k)L) \sum_{m\neq 0} \sum_{n\neq 0} a_{k+m} a_{k+n}^* a_{k+m+n} \boldsymbol{\Xi}_{m,n}^{FO} \\
&= (-j)\gamma a_k \exp(j\varphi_D(w_k)L) \left[\underbrace{\left|a_k\right|^2 \Xi_{0,0}^{FO} + 2 \sum_{n\neq 0} \left|a_{k+n}\right|^2 \Xi_{0,n}^{FO}}_{\text{SPM+IXPM}} \right] \\
&\qquad + (-j)\gamma \underbrace{\exp(j\varphi_D(w_k)L) \sum_{m\neq 0} \sum_{n\neq 0} a_{k+m} a_{k+n}^* a_{k+m+n} \boldsymbol{\Xi}_{m,n}^{FO}}_{\text{IFWM}} \\
&= (-j) a_k \exp(j\varphi_D(w_k)L) \varphi_{\text{NL}} + \Delta a_{\text{IFWM}}, \qquad \text{(A.20)}
\end{aligned}$$

where

$$\varphi_{\text{NL}} = \gamma \left[\left|a_k\right|^2 \Xi_{0,0}^{FO} + 2 \sum_{n\neq 0} \left|a_{k+n}\right|^2 \Xi_{0,n}^{FO} \right] \qquad \text{(A.21)}$$

and

$$\Delta a_{\text{IFWM}} = (-j)\gamma \exp(j\varphi_D(w_k)L) \sum_{m\neq 0} \sum_{n\neq 0} a_{k+m} a_{k+n}^* a_{k+m+n} \boldsymbol{\Xi}_{m,n}^{FO}. \qquad \text{(A.22)}$$

*Since the analysis is done for the single-channel, only intra-channel nonlinearities are considered.

Appendix B

Calculation of the First-Order ERP-Based Nonlinear Distortion Field for Dual-Polarization Transmission Systems

B.1 First-order ERP-based Nonlinear Distortion Field

In Section 5.2, we obtained the first-order ERP-based solution by applying a change of variable technique in (5.2) to solve the energy divergence problem of the RP-based method. Equation (5.10) represents the first-order ERP distortion field in time-domain. By taking the Fourier transform of (5.10), we get the distortion field in

frequency-domain at a transmission distance $z = L$ as:

$$U_1^{ERP}(L,w) = j\gamma \int_0^L \tilde{F}(z,w)\exp(-j\frac{w^2\beta_2 z}{2})\exp(-\alpha z)dz, \tag{B.1}$$

where $\tilde{F}(z,w)$ is given as:

$$\tilde{F}(z,w) = \mathcal{F}\{[|\tilde{u}_0(z,t)|^2\,\tilde{u}_0(z,t) - P_0\tilde{u}_0(z,t)]\exp(-j\gamma P_0 L_{\text{eff}})\}. \tag{B.2}$$

The input field to the optical fiber can be represented as:

$$\tilde{u}(z=0,t) = \sqrt{P}\sum_k a_k\tilde{g}(z=0,t-kT) = \sqrt{P}\sum_k a_k\tilde{g}(0,k), \tag{B.3}$$

where P is the launch power, a_k is the symbol complex amplitude imposed by data modulation on the k^{th} pulse, $\tilde{g}(z,t)$ is the pulse temporal waveform, and T is the symbol interval. By substituting (B.3) in (B.2) and calculating the FT, we obtain:

$$\tilde{F}(z,w) = P^{3/2}\sum_m\sum_n\sum_l a_m a_n a_l^*[\tilde{G_m}(z,w)\otimes\tilde{G}_l^*(z,-w)\otimes\tilde{G}_n(z,w) - P_0 a_k\tilde{G}(z,w)]\exp(-j\gamma P_0 L_{\text{eff}}), \tag{B.4}$$

where $*$ is the complex conjugation operation, m, n, l are the symbol indices, and $\tilde{G}(z,w) = \mathcal{F}\{\tilde{g}(z,t)\}$. Calculating the convolution operation in (B.4) and substituting

the result in (B.1), we get the ERP kernel term in frequency-domain as:

$$U_1^{ERP}(L,w)$$
$$= \underbrace{\left(\begin{array}{c} j\gamma a_m a_l^* a_n \int_0^L \iint \exp(-\alpha z)\tilde{G}(0,w_1+w)\tilde{G}^*(0,w_1+w-w_2)\tilde{G}(0,w-w_2) \\ \times \exp(-j[w_1(T_m-T_l)+w_2(T_l-T_n)-\beta_2 z w_1 w_2])dw_1 dw_2 dz \\ \times \exp(-jw(T_m+T_n-T_l)\exp(-j\gamma P_0 L_{\text{eff}}) \end{array}\right)}_{\text{Term 1}}$$
$$\underbrace{- j\gamma P_0 a_k \int_0^L \exp(-\alpha z)\tilde{G}(0,w)dz \exp(-j\gamma P_0 L_{\text{eff}})}_{\text{Term 2}}. \quad \text{(B.5)}$$

The ERP kernel in time-domain can be obtained by calculating the IFT of (B.5). First, we consider Term 1 and calculate the IFT. Assuming the Gaussian shape for input pulses, i.e., $\tilde{G}(0,w) = \sqrt{2\pi\tau^2}\exp(-\frac{w^2\tau^2}{2})$, with τ as the pulse width, the product of the triplet pulses in Term 1 can be represented as [46]:

$$\tilde{G}(0,w_1+w)\tilde{G}^*(0,w_1+w-w_2)\tilde{G}(0,w-w_2) = \left(\sqrt{2\pi\tau^2}\right)^3 \exp(-\frac{3\tau^2 w^2}{2})$$
$$\times \exp(-\tau^2[w_1^2+w_2^2+2(w_1-w_2)w-w_1 w_2]). \quad \text{(B.6)}$$

Substituting (B.6) in Term 1 of (B.5) and following a similar procedure with the phase matching condition $m+n-l=k$, as in [46], the IFT of Term 1 is obtained as:

$$\tilde{u}_{1,\text{Term 1}}^{ERP}(L,t+kT) = j\gamma P^{3/2}\exp(-j\gamma P_0 L_{\text{eff}})\sum_m\sum_n\sum_l a_m a_n a_l^*$$
$$\times \exp(-\frac{t^2}{6\tau^2})\int_0^L \frac{\exp(-\alpha z)}{\sqrt{1+2j\beta_2 z/\tau^2+3(\beta_2 z/\tau^2)^2}}$$
$$\times \exp\left\{\begin{array}{c} -\frac{3\left[\frac{2}{3}t+(m-l)T\right]\left[\frac{2}{3}t+(n-l)T\right]}{\tau^2(1+3j\beta_2 z/\tau^2)} \\ -\frac{(n-m)^2T^2}{\tau^2[1+2j\beta_2 z/\tau^2+3(\beta_2 z/\tau^2)^2]} \end{array}\right\} dz. \quad \text{(B.7)}$$

Next, consider Term 2 and calculate the IFT as:

$$\begin{aligned}\tilde{u}^{ERP}_{1,\text{Term 2}}(L,t+kT) &= j\gamma\sqrt{2\pi\tau^2}P_0 a_k \int_0^L \int \exp(-\alpha z)\exp(-\frac{w^2\tau^2}{2})dz \\ &\quad \times \exp(-j\gamma P_0 L_{\text{eff}})\exp(jwt)dwdz \\ &= j\gamma P_0 a_k \exp(-\frac{t^2}{2\tau^2}) \int_0^L \exp(-\alpha z)\exp(-j\gamma P_0 L_{\text{eff}})dz. \end{aligned} \tag{B.8}$$

Combining (B.7) and (B.8), the first-order ERP kernel term in time-domain can be represented as:

$$u^{ERP}_1(L,t+kT) = \tilde{u}^{ERP}_{1,\text{Term 1}}(L,t+kT) - \tilde{u}^{ERP}_{1,\text{Term 2}}(L,t+kT). \tag{B.9}$$

B.2 Extension to Dual-polarization

In the dual-polarization case, the electric field input to the optical fiber is a column vector $\mathbf{u}(z,t) = [u_x(z,t)\ \ u_y(z,t)]^\dagger$, with x, y representing the horizontal and vertical polarization, respectively, and the superscript $\dagger$ as the transpose. The propagation of the vector field $\mathbf{u}(z,t)$ through the optical fiber can be represented using the Manakov equation, where the nonlinear effective length is much longer than the fiber birefringent beating length, as [46]:

$$\frac{\partial}{\partial z}\mathbf{u} + j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}\mathbf{u} = j\frac{8}{9}\gamma(\mathbf{u}^{*\dagger}\mathbf{u}\mathbf{I})\mathbf{u}, \tag{B.10}$$

where $\mathbf{I}$ is the identity matrix. Note that in (B.10), we omitted the space and time variables z, t for the sake of simplicity. After applying the change of variable technique, as in the single polarization case, the modified Manakov equation can be represented

as:

$$\frac{\partial}{\partial z}\tilde{\mathbf{u}} + j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}\tilde{\mathbf{u}} = j\frac{8}{9}\gamma(\tilde{\mathbf{u}}^{*\dagger}\tilde{\mathbf{u}}\mathbf{I} - P_0\mathbf{I})\tilde{\mathbf{u}}, \tag{B.11}$$

where $\tilde{\mathbf{u}}(z,t) = [\tilde{u_x}(z,t) \;\; \tilde{u_y}(z,t)]^\dagger$. After solving (B.11) with the ERP technique, the zeroth- and first-order solutions for the output field can be represented as:

$$u_{0,x/y}^{ERP}(L,t) = \tilde{u}_{0,x/y}(L,t)\exp(-j\gamma P_0 L_{\text{eff}}) \tag{B.12}$$

and

$$u_{1,x/y}^{ERP}(L,t) = j\frac{8}{9}\gamma\int_0^L \exp(-\alpha z)\left(h_L(t)\otimes\left[\left(\left|\tilde{u}_{0,x/y}(z,t)\right|^2 - P_0\right)\tilde{u}_{0,x/y}(z,t)\right]\right)dz$$
$$\times\exp(-j\gamma P_0 L_{\text{eff}}). \tag{B.13}$$

Appendix C

Mathematical Proofs for Lemmas, Theorems, and Corollaries in Chapter 6

C.1 Proof of Lemma 1

Here, we consider Term 1 of the nonlinear part in (6.6). The corresponding propagation equation can be represented as:

$$\frac{\partial}{\partial z}u_2(z,t) + j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}u_2(z,t) = j2\left|u_0(z,t)\right|^2 \tilde{u}_1(z,t)\exp(-\alpha z). \tag{C.1}$$

By taking the Fourier transform of (C.1) and integrating with respect to z from 0 to L with the assumption of an ideal dispersion compensation at $z = L$, we get the solution in frequency-domain as:

$$U_2^{Term\,1}(L,w) = 2\gamma^2 \int_0^L F^{Term\,1}(z,w)\exp\left(-j\frac{w^2\beta_2 z}{2}\right)\exp(-\alpha z)dz. \tag{C.2}$$

Note that (C.1) is obtained by equating the terms that multiply the SO nonlinearity coefficient γ^2 on both sides of the equal sign in (6.7). Therefore, the actual SO distortion field is obtained by scaling the solution of (C.1) with γ^2, as shown in (C.2). The $F^{Term\,1}(z,w)$ in (C.2) is given as:

$$F^{Term\,1}(z,w) = \int_{-\infty}^{\infty} (\tilde{u}_1(z,t)u_0(z,t)u_0^*(z,t)) \exp(-jwt)\, dt, \tag{C.3}$$

where $u_0(z,t)$ is the zeroth-order (linearly dispersed) pulse and $\tilde{u}_1(z,t)$ is the dispersed FO ghost pulse.

The input pulse sequence can be represented as:

$$u_0(z=0,t) = \sqrt{P_0} \sum_{\bar{k}} a_{\bar{k}} \hat{g}(z=0, t-\bar{k}T), \tag{C.4}$$

where $a_{\bar{k}}$ is the data information of the $\bar{k}^{th}$ pulse and $g(z,t)$ is the pulse temporal waveform at z. In our analysis, we consider Gaussian shape assumption for the input pulses, i.e., $\hat{g}(z=0,t) = \exp\left(-\frac{t^2}{2\tau^2}\right)$, where τ represents the pulse width. By substituting (C.4) in (C.3), the equation for $F^{Term\,1}(z,w)$ can be represented as:

$$F^{Term\,1}(z,w) = P_0^{5/2} \sum_m \sum_n \sum_l \sum_k \sum_p a_m a_l^* a_n a_k a_p^* \int_{-\infty}^{\infty} \tilde{g}_{1,m+n-l}(z, t-(m+n-l)T) \times \hat{g}_k(z, t-kT)\hat{g}_p^*(z, t-pT) \exp(-jwt)\, dt. \tag{C.5}$$

Next, substituting (C.5) in (C.2), we obtain:

$$U_2^{Term\,1}(L,w) = \gamma^2 P_0^{5/2} \sum_m \sum_n \sum_l \sum_k \sum_p 2a_m a_l^* a_n a_k a_p^* G_{m,n,l,k,p}^{Term\,1}(z,w), \tag{C.6}$$

where

$$G^{Term\,1}_{m,n,l,k,p}(z,w) = \int_0^L \left(\int_{-\infty}^{\infty} \tilde{g}_{1,m+n-l}(z, t-(m+n-l)T) \right.$$
$$\left. \times\, \hat{g}_k(z, t-kT)\hat{g}^*_p(z, t-pT)\exp\left(-jwt\right)\, dt \right) \exp\left(-j\frac{w^2\beta_2 z}{2}\right)\exp(-\alpha z)dz. \quad \text{(C.7)}$$

It is clear from (C.7) that the function $G^{Term\,1}_{m,n,l,k,p}(z,w)$ calculates the coefficient of nonlinear interaction between five data symbols to generate the SO nonlinear distortion field in frequency-domain. For simplicity of implementation, we carry out the SO predistortion in time-domain. Accordingly, by calculating the inverse Fourier transform of $G^{Term\,1}_{m,n,l,k,p}(z,w)$, we get the corresponding function in time-domain as:

$$g^{Term\,1}_{m,n,l,k,p}(z,t) = \int_0^L \left(\tilde{g}_{1,m+n-l}(z, t-(m+n-l)T)\ \hat{g}_k(z, t-kT)\hat{g}^*_p(z, t-pT) \right)$$
$$\otimes \left(\frac{1}{\sqrt{-2\pi j\beta_2 z}} \exp\left(\frac{-jt^2}{2\beta_2 z}\right) \right) \exp(-\alpha z)dz. \quad \text{(C.8)}$$

In (C.8), $\tilde{g}_{1,m+n-l}(z, t-(m+n-l)T)$ is the dispersed FO ghost pulse obtained by the convolution between the temporal dispersion term $\frac{1}{\sqrt{-2\pi j\beta_2 z}} \exp\left(\frac{-jt^2}{2\beta_2 z}\right)$ and the FO nonlinear coefficient expression in (2.103), which can be represented as:

$$\tilde{g}_{1,m+n-l}(z,t-(m+n-l)T)=\frac{1}{\sqrt{-2\pi j\beta_2 z}}\exp\left(\frac{-jt^2}{2\beta_2 z}\right)$$
$$\otimes\left(\exp(\frac{-t^2}{6\tau^2})\int_0^z\frac{\exp(-\alpha s)}{\sqrt{1+2j\beta_2 s/\tau^2+3(\beta_2 s/\tau^2)^2}}\exp\left\{\begin{array}{c}-\frac{3[2t/3+(m-l)T][2t/3+(n-l)T]}{\tau^2(1+3j\beta_2 s/\tau^2)}\\-\frac{(n-m)^2T^2}{\tau^2[1+2j\beta_2 s/\tau^2+3(\beta_2 s/\tau^2)^2]}\end{array}\right\}ds\right)$$
$$=j\tau^3\int_0^z\frac{\exp(-\alpha s)}{\sqrt{-j\overline{B}(s)\widehat{D}(z,s)}}\exp\left\{\frac{-j}{\overline{B}(s)\widehat{D}(z,s)}\right.$$
$$\left[(\dot{A}_{m,n,l}T^2-t\dot{B}_{m,n,l}T+\frac{3}{2}t^2)\tau^4-j\beta_2((\dot{C}_{m,n,l}s+\dot{D}_{m,n,l}z)T^2-t\dot{E}_{m,n,l}sT+st^2)\tau^2-\beta_2^2 s\right.$$
$$\left.\left.\times(\dot{F}_{m,n,l}zT^2-\frac{1}{2}st^2)\right]\right\}ds, \quad \text{(C.9)}$$

where

$$\dot{A}_{m,n,l}=\left(3l^2-3\left(m+n\right)l+m^2+nm+n^2\right), \quad \text{(C.10)}$$
$$\dot{B}_{m,n,l}=(4l-2(m+n)), \quad \text{(C.11)}$$
$$\dot{C}_{m,n,l}=\left(3\left(l-n\right)\left(l-m\right)\right), \quad \text{(C.12)}$$
$$\dot{D}_{m,n,l}=\left(l^2-\left(m+n\right)l+m^2-nm+n^2\right), \quad \text{(C.13)}$$
$$\dot{E}_{m,n,l}=\left(4-2(m+n)\right), \quad \text{(C.14)}$$
$$\dot{F}_{m,n,l}=\left(\left(l-n\right)\left(l-m\right)\right). \quad \text{(C.15)}$$

Expressions for the zeroth-order (linearly dispersed) pulses $\hat{g}_k(z,t-kT)$ and $\hat{g}_p^*(z,t-pT)$ can be represented as $\frac{\tau}{\sqrt{\tau^2-j\beta_2 z}}\exp\left(\frac{(kT-t)^2}{2(j\beta_2 z-\tau^2)}\right)$ and $\frac{\tau}{\sqrt{\tau^2+j\beta_2 z}}\exp\left(-\frac{(pT-t)^2}{2(\tau^2+j\beta_2 z)}\right)$, respectively. Next, substituting (C.9) and the expressions for the linearly dispersed pulses $\hat{g}_k(z,t-kT)$ and $\hat{g}_p^*(z,t-pT)$ in (C.8), we obtain $g_{m,n,l,k,p}^{Term\,1}(z,s,t)$ as in (C.16). The 4-D nonlinear coefficient matrix $\mathbf{C}_{m,n,l,k}^{SO,\,Term\,1}$ is obtained by substituting the phase-matching condition $p=m+n-l+k$ and $t=0$ (symbol rate operation) in (C.16), i.e.,

$$\mathbf{C}_{m,n,l,k}^{SO,\,Term\,1}=g_{m,n,l,k,p}^{Term\,1}(z,s,t)\Big|_{t=0,\,p=m+n-l+k}. \quad \text{(C.27)}$$

$$g^{Term\,1}_{m,n,l,k,p}(z,s,t) = -\tau^3 \int_0^L \int_0^z \frac{\exp(-\alpha(z+s))}{\sqrt{\overline{A}(z,s)\overline{B}(s)}} \exp\left\{ \frac{1}{2\tau^2 \overline{A}(z,s)\overline{C}(s)} \left[(\ddot{A}_{m,n,l,k,p} T^2 \right.\right.$$
$$- 2t\ddot{B}_{m,n,l,k,p} T + 5t^2)\tau^8 + 2j\beta_2((\ddot{C}_{m,n,l,k,p} z + s\ddot{D}_{m,n,l,k,p})T^2 - 2(\ddot{E}_{m,n,l,k,p} s - (k+2p)z)tT$$
$$+ t^2(s-3z))\tau^6 + \beta_2^2((\ddot{F}_{m,n,l,k,p} z^2 - 4sz\ddot{G}_{m,n,l,k,p} + 3s^2(k^2+p^2))T^2 - 2(\ddot{H}_{m,n,l,k,p} z^2 - 4ksz$$
$$+ 3s^2(k+p))tT + (7s^2 - 4sz + 3z^2)t^2)\tau^4 + 2j\beta_2^3 sz((\ddot{I}_{m,n,l,k,p} z + ks(k-p))T^2 - 2(\ddot{J}_{m,n,l,k,p} z$$
$$\left.\left. + s(k-2p))tT - t^2(s+z))\tau^2 + s^2 z^2 \beta_2^4((k-p)T - t)^2 \right] \right\} ds\, dz, \quad \text{(C.16)}$$

where

$$\ddot{A}_{m,n,l,k,p} = \left(k^2 + p^2 + 6l^2 - (6(m+n))l + 2(m^2 + mn + n^2)\right), \quad \text{(C.17)}$$
$$\ddot{B}_{m,n,l,k,p} = (k + 4l - 2(m+n) + p), \quad \text{(C.18)}$$
$$\ddot{C}_{m,n,l,k,p} = \left(k^2 + (-4l + 2m + 2n - p)k - (2(2l - m - n))p + 6l^2 \right.$$
$$\left. - (6(m+n))l + 2(m^2 + mn + n^2)\right), \quad \text{(C.19)}$$
$$\ddot{D}_{m,n,l,k,p} = \left(k^2 + p^2 - (3(l-n))(l-m)\right), \quad \text{(C.20)}$$
$$\ddot{E}_{m,n,l,k,p} = (k - 2l + m + n + p), \quad \text{(C.21)}$$
$$\ddot{F}_{m,n,l,k,p} = \left(3k^2 + (8l - 4m - 4n - 6p)k + 3p^2 - 4(2l - m - n)p + 10l^2 \right.$$
$$\left. - (10(m+n))l + 6m^2 + 6n^2 - 2mn\right), \quad \text{(C.22)}$$
$$\ddot{G}_{m,n,l,k,p} = \left(k^2 + (2l - m - n - p)k + (2l - m - n)p - (3(l-n))(l-m)\right), \quad \text{(C.23)}$$
$$\ddot{H}_{m,n,l,k,p} = (3k + 4l - 2m - 2n - 3p), \quad \text{(C.24)}$$
$$\ddot{I}_{m,n,l,k,p} = \left(-k^2 - (2(l - m - n - p))k - p^2 + (2(2l - m - n))p \right.$$
$$\left. - (5(l-n))(l-m)\right), \quad \text{(C.25)}$$
$$\ddot{J}_{m,n,l,k,p} = (-k - 2l + m + n + p). \quad \text{(C.26)}$$

C.2 Proof of Lemma 2

By considering Term 2 of the nonlinear part in (6.6), the propagation equation governing the evolution of the SO distortion field can be represented as:

$$\frac{\partial}{\partial z} u_2(z,t) + j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2} u_2(z,t) = j u_0^2(z,t)\tilde{u}_1^*(z,t)\exp(-\alpha z). \quad \text{(C.28)}$$

The solution of (C.28) in frequency-domain can be obtained as:

$$U_2^{Term\,2}(L,w) = \gamma^2 \int_0^L F^{Term\,2}(z,w) \exp\left(-j\frac{w^2\beta_2 z}{2}\right) \exp(-\alpha z) dz, \tag{C.29}$$

where

$$F^{Term\,2}(z,w) = \int_{-\infty}^{\infty} \left(\tilde{u}_1^*(z,t) u_0(z,t) u_0(z,t)\right) \exp\left(-jwt\right) dt. \tag{C.30}$$

By substituting (C.4) in (C.30), we obtain:

$$\begin{aligned} F^{Term\,2}(z,w) = P_0^{5/2} \sum_m \sum_n \sum_l \sum_k \sum_p a_m^* a_l a_n^* a_k a_p \\ \times \int_{-\infty}^{\infty} \tilde{g}_{1,m+n-l}^*(z, t-(m+n-l)T) \\ \times \hat{g}_k(z, t-kT) \hat{g}_p(z, t-pT) \exp\left(-jwt\right) dt. \end{aligned} \tag{C.31}$$

Next, substituting (C.31) in (C.29), we get the nonlinear distortion term as:

$$U_2^{Term\,2}(L,w) = \gamma^2 P_0^{5/2} \sum_m \sum_n \sum_l \sum_k \sum_p a_m^* a_l a_n^* a_k a_p G_{m,n,l,k,p}^{Term\,2}(z,w), \tag{C.32}$$

where

$$\begin{aligned} G_{m,n,l,k,p}^{Term\,2}(z,w) = \int_0^L \Bigg(\int_{-\infty}^{\infty} \tilde{g}_{1,m+n-l}^*(z, t-(m+n-l)T) \\ \times \hat{g}_k(z, t-kT) \hat{g}_p(z, t-pT) \exp\left(-jwt\right) dt \Bigg) \\ \times \exp\left(-j\frac{w^2\beta_2 z}{2}\right) \exp(-\alpha z) dz. \end{aligned} \tag{C.33}$$

By taking inverse Fourier transform of (C.33), we obtain:

$$g_{m,n,l,k,p}^{Term\,2}(z,t) = \int_0^L \left(\tilde{g}_{1,m+n-l}^{*}(z, t-(m+n-l)T)\ \hat{g}_k(z,t-kT)\hat{g}_p(z,t-pT) \right) \otimes \left(\frac{1}{\sqrt{-2\pi j\beta_2 z}} \exp\left(\frac{-jt^2}{2\beta_2 z} \right) \right) \exp(-\alpha z) dz. \quad \text{(C.34)}$$

Substituting (C.9) and the expressions for the linearly dispersed pulses in (C.34), we obtain (C.35). The 4-D nonlinear coefficient matrix $\mathbf{C}_{m,n,l,k}^{SO,\,Term\,2}$ is obtained by substituting the phase-matching condition $p = m+n-l+k$ and $t = 0$ in (C.35), i.e.,

$$\mathbf{C}_{m,n,l,k}^{SO,\,Term\,2} = g_{m,n,l,k,p}^{Term\,2}(z,s,t)\Big|_{t=0,\,p=m+n-l+k}. \quad \text{(C.49)}$$

C.3 Proof of Theorem 1

The solution of the differential equation in (6.6) which governs the evolution of the SO distortion field in the optical fiber can be represented in frequency-domain as:

$$U_2(L,w) = \gamma^2 \int_0^L F(z,w) \exp\left(-j\frac{w^2\beta_2 z}{2} \right) \exp(-\alpha z) dz, \quad \text{(C.50)}$$

where

$$F(z,w) = 2F^{Term\,1}(z,w) + F^{Term\,2}(z,w). \quad \text{(C.51)}$$

By substituting expressions for $F^{Term\,1}(z,w)$ and $F^{Term\,2}(z,w)$ in (C.51), we obtain:

$$F(z,w) = P_0^{5/2} \sum_m \sum_n \sum_l \sum_k \sum_p \Big[2a_m a_l^* a_n a_k a_p^* \int_{-\infty}^{\infty} \tilde{g}_{1,m+n-l}(z, t-(m+n-l)T) \times \hat{g}_k(z,t-kT)\hat{g}_p^*(z,t-pT) \exp(-jwt)\, dt + a_m^* a_l a_n^* a_k a_p \int_{-\infty}^{\infty} \tilde{g}_{1,m+n-l}^{*}(z, t-(m+n-l)T) \times \hat{g}_k(z,t-kT)\ \hat{g}_p(z,t-pT) \exp(-jwt)\, dt \Big]. \quad \text{(C.52)}$$

$$g^{Term\,2}_{m,n,l,k,p}(z,s,t) = \sqrt{3}\tau^4 \int_0^L \int_0^z \frac{\sqrt{\widehat{A}(z,s)}\exp(-\alpha(z+s))}{\sqrt{\widehat{B}(z,s)\widehat{C}(z)}\left(\sqrt{-\overline{B}(s)\widehat{D}(z,s)}\right)^*}$$
$$\times \exp\left\{\frac{-j}{2\tau^2\widehat{B}(z,s)\widehat{E}(z)\widehat{F}(s)}\left[(\dddot{A}_{m,n,l,k,p}T^2 + 2t\ddot{B}_{m,n,l,k,p}T - 5t^2)\tau^8\right.\right.$$
$$+ 2j\beta_2((\dddot{C}_{m,n,l,k,p}z + \ddot{D}_{m,n,l,k,p}s)T^2 - 2t(\ddot{E}_{m,n,l,k,p}z + \dddot{F}_{m,n,l,k,p}s)T + t^2(s+z))\tau^6$$
$$- \beta_2^2((\ddot{G}_{m,n,l,k,p}z^2 - 4\dddot{H}_{m,n,l,k,p}sz + 3s^2(k^2+p^2))T^2 - 2t(\dddot{I}_{m,n,l,k,p}z^2 - 4\ddot{J}_{m,n,l,k,p}sz$$
$$+ 3s^2(p+k))T + t^2(7s^2 - 4sz + 3z^2))\tau^4 - 2jsz\beta_2^3((\ddot{K}_{m,n,l,k,p}z$$
$$+ \dddot{L}_{m,n,l,k,p}s)T^2 + 2t(\dddot{M}_{m,n,l,k,p}z + s(p+k))T - t^2(3s-z))\tau^2$$
$$\left.\left. -((p+k)T - t)^2s^2z^2\beta_2^4\right]\right\} ds\,dz, \quad \text{(C.35)}$$

where

$$\dddot{A}_{m,n,l,k,p} = \left(-k^2 - 6l^2 - p^2 + 6(m+n)l - 2(m^2+nm+n^2)\right), \quad \text{(C.36)}$$
$$\ddot{B}_{m,n,l,k,p} = (p + k + 2(2l - m - n)), \quad \text{(C.37)}$$
$$\dddot{C}_{m,n,l,k,p} = \left(-3k^2 + (4l - 2m - 2n + p)k - 3p^2 + 2(2l - m - n)p - 2l^2\right.$$
$$\left. +2(m+n)l - 2(m^2 - nm + n^2)\right), \quad \text{(C.38)}$$
$$\ddot{D}_{m,n,l,k,p} = \left(k^2 + p^2 - (3(l-n))(l-m)\right), \quad \text{(C.39)}$$
$$\ddot{E}_{m,n,l,k,p} = (-p - k + 2(2l - m - n)), \quad \text{(C.40)}$$
$$\dddot{F}_{m,n,l,k,p} = (p + k - 2l + m + n), \quad \text{(C.41)}$$
$$\dddot{G}_{m,n,l,k,p} = \left(3k^2 - 2(4l - 2m - 2n - 3p)k + 3p^2 - 4(2l - m - n)p + 10l^2\right.$$
$$\left. -10l(m+n) - 2nm + 6(n^2+m^2)\right), \quad \text{(C.42)}$$
$$\dddot{H}_{m,n,l,k,p} = \left(k^2 + (-2l + m + n + p)k + (l - n - p)(l - m - p)\right), \quad \text{(C.43)}$$
$$\dddot{I}_{m,n,l,k,p} = (3(p+k) - 4l + 2(m+n)), \quad \text{(C.44)}$$
$$\ddot{J}_{m,n,l,k,p} = (p + k - 2l + m + n), \quad \text{(C.45)}$$
$$\dddot{K}_{m,n,l,k,p} = \left(k^2 - 2(2l - m - n - p)k + p^2 - 2(2l - m - n)p\right.$$
$$\left. +(5(l-n))(l-m)\right), \quad \text{(C.46)}$$
$$\dddot{L}_{m,n,l,k,p} = \left(k^2 - 3kp + p^2\right), \quad \text{(C.47)}$$
$$\dddot{M}_{m,n,l,k,p} = (-p - k + 2l - m - n). \quad \text{(C.48)}$$

Next, substituting (C.52) in (C.50) and substituting the expressions for the FO ghost pulse and the linearly dispersed pulses, we obtain the SO distortion field as:

$$U_2(L,w) = \gamma^2 P_0^{5/2} \sum_m \sum_n \sum_l \sum_k \sum_p \left[2a_m a_l^* a_n a_k a_p^* G^{Term\,1}_{m,n,l,k,p}(z,w)\right.$$
$$\left. +a_m^* a_l a_n^* a_k a_p G^{Term\,2}_{m,n,l,k,p}(z,w)\right], \quad \text{(C.53)}$$

where $G^{Term\,1}_{m,n,l,k,p}(z,w)$ and $G^{Term\,2}_{m,n,l,k,p}(z,w)$ are given by (C.7) and (C.33), respectively. By calculating the inverse Fourier transform of (C.53) and substituting the phase-matching condition $p = m+n-l+k$ and $t = 0$ followed by some simplifications, we obtain the SO distortion field in time-domain as in (6.35), i.e.,

$$u_2(L,t) = \Bigg(\gamma^2 P_0^{5/2} \sum_m \sum_n \sum_l \sum_k \sum_p \Big[2a_m a_l^* a_n a_k a_p^* g^{Term\,1}_{m,n,l,k,p}(z,s,t) + a_m^* a_l a_n^* a_k a_p g^{Term\,2}_{m,n,l,k,p}(z,s,t)\Big]\Bigg)\Bigg|_{t=0,\, p=m+n-l+k}. \quad \text{(C.54)}$$

C.4 Proof Outline of Corollary 1

The substitution $l = m+n$ is the phase-matching condition for the triplet pulses involved in the generation of the FO ghost pulse at the zeroth time index. By substituting $l = m+n$ in (C.9), the expression for the FO ghost pulse generated at the zeroth time index can be represented as:

$$\tilde{g}_{1,0}(z,t) = j\tau^3 \int_0^z \frac{\exp(-\alpha s)}{\sqrt{-j\overline{B}(s)\widehat{D}(z,s)}} \exp\Bigg\{\frac{-j}{\overline{B}(s)\widehat{D}(z,s)} \Big[(\dddot{A}_{m,n} T^2 - 2t\dddot{B}_{m,n} T + \frac{3}{2}t^2)\tau^4 - j\beta_2((m^2 z + (3s-z)mn + n^2 z)T^2 + 2st\dddot{C}_{m,n} T + st^2)\tau^2 - (mnzT^2 - \frac{1}{2}st^2)\beta_2^2 s\Big]\Bigg\} ds, \quad \text{(C.55)}$$

where

$$\dddot{A}_{m,n} = \left(m^2 + mn + n^2\right), \quad \text{(C.56)}$$
$$\dddot{B}_{m,n} = (m+n), \quad \text{(C.57)}$$
$$\dddot{C}_{m,n} = (m+n-2). \quad \text{(C.58)}$$

By substituting (C.55) in (C.8) and (C.34) and phase-matching conditions $l = m + n$ and $p = k$ at $t = 0$, we can rewrite (C.54) as:

$$\widetilde{u}_2(L,t) = \Bigg(\gamma^2 P_0^{5/2} \sum_m \sum_n \sum_l \sum_k \sum_p \Big[2a_m a_l^* a_n a_k a_p^* g_{m,n,l,k,p}^{Term\,1}(z,s,t) + a_m^* a_l a_n^* a_k a_p g_{m,n,l,k,p}^{Term\,2}(z,s,t) \Big] \Bigg) \Bigg|_{t=0,\, l=m+n,\, p=k} . \quad \text{(C.59)}$$

After some simplifications, the proof is concluded by deriving (6.46).

C.5 Proof of Theorem 2

The propagation of the polarization multiplexed optical signal through the SSMF can be modeled by using the Manakov equation, which is represented as:

$$\frac{\partial}{\partial z}\mathbf{u} + j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}\mathbf{u} = j\frac{8}{9}\gamma(\mathbf{u}^{*\dagger}\mathbf{u}\mathbf{I})\mathbf{u}\exp\left(-\alpha z\right), \quad \text{(C.60)}$$

where $\mathbf{I}$ is the identity matrix and the input to the optical fiber is a column vector $\mathbf{u}(z,t) = [u_x(z,t)\ \ u_y(z,t)]^\dagger$, with x, y representing the horizontal and vertical polarization, respectively, and the superscript $\dagger$ as the transpose.

The differential equation governing the SO distortion field for the polarization multiplexed signal can be represented as:

$$\frac{\partial}{\partial z}u_{2,x/y}(z,t) = \underbrace{-j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}u_{2,x/y}(z,t)}_{\text{Linear part}} + \underbrace{\left(\begin{array}{c} \underbrace{j2\left(\left|u_{0,x/y}(z,t)\right|^2 + \left|u_{0,y/x}(z,t)\right|^2\right)\tilde{u}_{1,x/y}(z,t)\exp(-\alpha z)}_{\text{Term 1}} \\ \underbrace{+j\Big(u_{0,x/y}^2(z,t) + u_{0,y/x}^2(z,t)\Big)\tilde{u}_{1,x/y}^*(z,t)\exp(-\alpha z)}_{\text{Term 2}} \end{array} \right)}_{\text{Nonlinear part}} . \quad \text{(C.61)}$$

The solution of the differential equation in (C.61) which governs the evolution of the SO distortion field in the optical fiber can be represented in frequency-domain as:

$$U_{2,x/y}(L,w) = \frac{64}{81}\gamma^2 \int_0^L F_{x/y}(z,w) \exp\left(-j\frac{w^2\beta_2 z}{2}\right) \exp(-\alpha z) dz, \tag{C.62}$$

where

$$F_{x/y}(z,w) = 2F_{x/y}^{Term\,1}(z,w) + F_{x/y}^{Term\,2}(z,w). \tag{C.63}$$

Following the similar analysis in Section C.1 and Section C.2, we can represent $F_{x/y}^{Term\,1}(z,w)$ and $F_{x/y}^{Term\,2}(z,w)$ for dual-polarization transmission systems as:

$$\begin{aligned} F_{x/y}^{Term\,1}(z,w) &= P_0^{5/2} \sum_m \sum_n \sum_l \sum_k \sum_p \\ &\quad \times \left(a_{m,x/y} a_{l,x/y}^* + a_{m,y/x} a_{l,y/x}^*\right) a_{n,x/y} \left(a_{k,x/y} a_{p,x/y}^* + a_{k,y/x} a_{p,y/x}^*\right) \\ &\times \int_{-\infty}^{\infty} \tilde{g}_{1,m+n-l}(z, t-(m+n-l)T) \hat{g}_k(z, t-kT) \hat{g}_p^*(z, t-pT) \exp\left(-jwt\right) dt. \end{aligned} \tag{C.64}$$

and

$$\begin{aligned} F_{x/y}^{Term\,2}(z,w) &= P_0^{5/2} \sum_m \sum_n \sum_l \sum_k \sum_p \\ &\quad \times \left(a_{m,x/y}^* a_{l,x/y} + a_{m,y/x}^* a_{l,y/x}\right) a_{n,x/y}^* \left(a_{k,x/y} a_{p,x/y} + a_{k,y/x} a_{p,y/x}\right) \\ &\times \int_{-\infty}^{\infty} \tilde{g}_{1,m+n-l}^*(z, t-(m+n-l)T) \hat{g}_k(z, t-kT) \hat{g}_p(z, t-pT) \exp\left(-jwt\right) dt. \end{aligned} \tag{C.65}$$

By substituting (C.64) and (C.65) in (C.63), we obtain:

$$F_{x/y}(z,w) = P_0^{5/2} \sum_m \sum_n \sum_l \sum_k \sum_p \Bigg[2\left(a_{m,x/y} a^*_{l,x/y} + a_{m,y/x} a^*_{l,y/x}\right) a_{n,x/y} \times \left(a_{k,x/y} a^*_{p,x/y} + a_{k,y/x} a^*_{p,y/x}\right) \int_{-\infty}^{\infty} \tilde{g}_{1,m+n-l}(z, t-(m+n-l)T) \hat{g}_k(z, t-kT) \times \hat{g}^*_p(z, t-pT) \exp(-jwt)\, dt + \left(a^*_{m,x/y} a_{l,x/y} + a^*_{m,y/x} a_{l,y/x}\right) a^*_{n,x/y} \times \left(a_{k,x/y} a_{p,x/y} + a_{k,y/x} a_{p,y/x}\right) \int_{-\infty}^{\infty} \tilde{g}^*_{1,m+n-l}(z, t-(m+n-l)T) \times \hat{g}_k(z, t-kT) \hat{g}_p(z, t-pT) \exp(-jwt)\, dt \Bigg]. \quad \text{(C.66)}$$

Next, substituting (C.66) in (C.62) and substituting the expressions for the FO ghost pulse and the linearly dispersed pulses, we obtain the SO distortion field as:

$$U_{2,x/y}(L,w) = \frac{64}{81}\gamma^2 P_0^{5/2} \sum_m \sum_n \sum_l \sum_k \sum_p \Big[2\left(a_{m,x/y} a^*_{l,x/y} + a_{m,y/x} a^*_{l,y/x}\right) a_{n,x/y} \times \left(a_{k,x/y} a^*_{p,x/y} + a_{k,y/x} a^*_{p,y/x}\right) G^{Term\,1}_{m,n,l,k,p}(z,w) + \left(a^*_{m,x/y} a_{l,x/y} + a^*_{m,y/x} a_{l,y/x}\right) \times a^*_{n,x/y} \left(a_{k,x/y} a_{p,x/y} + a_{k,y/x} a_{p,y/x}\right) G^{Term\,2}_{m,n,l,k,p}(z,w) \Big], \quad \text{(C.67)}$$

where $G^{Term\,1}_{m,n,l,k,p}(z,w)$ and $G^{Term\,2}_{m,n,l,k,p}(z,w)$ are given by (C.7) and (C.33), respectively. It is worth noting that the coefficients $G^{Term\,1}_{m,n,l,k,p}(z,w)$ and $G^{Term\,2}_{m,n,l,k,p}(z,w)$ are same for both single-polarization and dual-polarization transmission systems. By calculating the inverse Fourier transform of (C.67) and substituting the phase-matching condition $p = m+n-l+k$ and $t = 0$ followed by some simplifications, we obtain the SO distortion

field in time-domain as:

$$u_{2,x/y}(L,t)=\left(\frac{64}{81}\gamma^2P_0^{5/2}\sum_m\sum_n\sum_l\sum_k\sum_p\Big[2\left(a_{m,x/y}a^*_{l,x/y}+a_{m,y/x}a^*_{l,y/x}\right)a_{n,x/y}\right.$$
$$\times\left(a_{k,x/y}a^*_{p,x/y}+a_{k,y/x}a^*_{p,y/x}\right)g^{Term\,1}_{m,n,l,k,p}(z,s,t)+\left(a^*_{m,x/y}a_{l,x/y}+a^*_{m,y/x}a_{l,y/x}\right)$$
$$\left.\left.\times a^*_{n,x/y}\left(a_{k,x/y}a_{p,x/y}+a_{k,y/x}a_{p,y/x}\right)g^{Term\,2}_{m,n,l,k,p}(z,s,t)\Big]\right)\right|_{t=0,\,p=m+n-l+k}. \quad \text{(C.68)}$$

C.6 Proof Outline of Corollary 2

By substituting the phase-matching conditions $l=m+n$ and $p=k$ at $t=0$, and substituting (C.55) in (C.8) and (C.34), we can rewrite (C.68) as:

$$\widetilde{u}_{2,x/y}(L,t)=\left(\frac{64}{81}\gamma^2P_0^{5/2}\sum_m\sum_n\sum_l\sum_k\sum_p\Big[2\left(a_{m,x/y}a^*_{l,x/y}+a_{m,y/x}a^*_{l,y/x}\right)a_{n,x/y}\right.$$
$$\times\left(a_{k,x/y}a^*_{p,x/y}+a_{k,y/x}a^*_{p,y/x}\right)g^{Term\,1}_{m,n,l,k,p}(z,s,t)+\left(a^*_{m,x/y}a_{l,x/y}+a^*_{m,y/x}a_{l,y/x}\right)$$
$$\left.\left.\times a^*_{n,x/y}\left(a_{k,x/y}a_{p,x/y}+a_{k,y/x}a_{p,y/x}\right)g^{Term\,2}_{m,n,l,k,p}(z,s,t)\Big]\right)\right|_{t=0,\,l=m+n,\,p=k}. \quad \text{(C.69)}$$

After some simplifications, the proof is concluded by deriving (6.49).