Signal Processing for Impairment Control in Fiber Optic Communication

by

Abbas Abolfathimomtaz

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Communications

Department of Electrical and Computer Engineering

University of Alberta

Abstract

Fiber optics is the only infrastructure capable of handling the large data traffic in the backend of many modern communication systems. However, the higher growth rate of data traffic compared to that of fiber system achievable data rate (ADR) raises significant concerns about meeting future demands. Recognizing this potential shortage, many researchers have recently directed their efforts towards increasing fiber ADR.

Fiber ADR is primarily limited by system impairments, which are rooted either in Kerr nonlinearity or the suboptimality of the transmitter/receiver. In the first half of this thesis, we focus on controlling Kerr nonlinearity, while the second half is dedicated to addressing another impairment occurring due to the suboptimal receiver structure.

Kerr nonlinearity is a power-dependent impairment that introduces nonlinear interference noise (NLIN). Based on existing fiber models, Kerr nonlinearity depends on the statistical properties of the power launched into the system; therefore, by controlling the launched power, one can reduce NLIN. One of the main factors impacting signal power behavior in a communication system is the constellation from which symbols are drawn. In the first completed work, we formulate the impact of power fluctuation on NLIN with respect to the launched symbols. We then demonstrate that by properly pairing the transmitted symbols in the wavelength division multiplexing (WDM) channels, we can minimize power fluctuation and consequently reduce fiber NLIN.

In the next completed work, we design the signal power spectral density (PSD) to minimize NLIN. Fiber models express NLIN power based on the PSD of the signal launched into the fiber. Thus, a rational question is: what is the optimal PSD that minimizes NLIN? By formulating the problem as an optimization problem, we find the optimal PSD that minimizes system NLIN, which differs from that achieved by the commonly used raised cosine pulse.

The suboptimal structure of the existing transmitter/receiver also limits fiber ADR. One of these suboptimalities is the order of the phase noise and dispersion compensators. While an optimal receiver should compensate for receiver laser phase noise before dispersion compensation, the lack of phase information before the dispersion compensator forces the incorrect compensation order. This suboptimal order gives rise to equalization-enhanced phase noise (EEPN), which restricts system performance at high baud rates and over long distances. To address the problem of EEPN, in the third completed work, we introduce a new formulation of EEPN that enables its compensation. We then provide two digital signal processing (DSP)-ready implementations of this compensator applicable to different receiver structures.

In the final completed work, we design a carrier phase estimator (CPE) that can extract receiver laser phase noise information before dispersion compensation. Our design leverages the fact that fiber chromatic dispersion causes the positive and negative excess bandwidths of the pulse to propagate at different velocities. This velocity mismatch, along with the fact that both positive and negative excess bandwidths are modulated with the same information, enables us to observe the receiver laser phase variation over time after proper signal processing. Our proposed CPE prevents the generation of EEPN by allowing the optimal sequence of dispersion and receiver laser phase noise compensations.

Preface

The result of **Chapter 2** was published in the following paper:

> Abolfathimomtaz, A., Ardakani, M., Ebrahimzad, H. and Yazdani, R., 2022. **"Constant power constellation grouping for reducing nonlinearity effects in fiber optics,"** *IEEE Journal of Lightwave Technology*, 40(22), pp.7314-7325.

The result of **Chapter 3** has been submitted to the IEEE Journal on Selected Areas in Communications as:

> Abolfathimomtaz, A., Ardakani, M., Ebrahimzad, H. and Zhang, Z., 2024. **"Minimizing Fiber's Nonlinear Interference Noise by Designing Launched Signal PSD,"** *IEEE Journal on Selected Areas in Communications.*

The result of **Chapter 4** has been accepted for publication in the IEEE Journal of Lightwave Technology as:

> Abolfathimomtaz, A., Ardakani, M., Ebrahimzad, H. and Zhang, Z., 2024. **"Equalization-Enhanced Phase Noise Compensation in Coherent Fiber Receivers,"** *IEEE Journal of Lightwave Technology.*

This paper was also presented at the IEEE Communication Theory Workshop (CTW) 2024.

The result of **Chapter 5** has been submitted to the IEEE Journal of Lightwave Technology as:

> Abolfathimomtaz, A., Ardakani, M., Ebrahimzad, H. and Zhang, Z., Chuandong L., 2024. **"Receiver Laser Phase Noise Estimation**

with Application to EEPN Control," *IEEE Journal of Lightwave Technology.*

Education is not the learning of facts, but the training of the mind to think.

– Albert Einstein, 1879–1955.

Acknowledgements

I would like to express my deepest gratitude to all those who have supported me throughout my PhD journey.

First and foremost, I am immensely grateful to my primary supervisor, Prof. Ardakani. His kindness and patience have been invaluable. Prof. Ardakani not only taught me how to conduct my research but also how to think critically about everything. His thorough discussions on every aspect of my research have sparked numerous new ideas. It has been my pleasure to be one of the few who had the privilege to be his PhD student. This research would not have been possible without his valuable supervision. He has been more than just a great supervisor; he has been a wonderful mentor for life.

I extend my heartfelt thanks to my wife, Forough. She was unwavering in her support throughout my PhD, continually encouraging me to put more effort into my research and staying by my side during challenging times. Her constant support and motivation were crucial in helping me persevere through the toughest moments of this journey.

Additionally, I owe a debt of gratitude to my parents, who raised me as a researcher. They instilled values in me that will shape my life forever. Their unwavering belief in my abilities and their endless sacrifices made this journey possible. Their guidance and encouragement have always been a source of strength and inspiration for me.

I am very thankful to Dr. Ebrahimzad and Dr. Yazdani, our contacts at Huawei, who trained me in the area of fiber communications. Their technical insights were instrumental in defining practical problems and addressing them effectively. Their expertise and perspectives greatly enhanced the quality and practical relevance of my research.

My special thanks go to Dr. Frank R. Kschischang, a distinguished professor of digital communication at the University of Toronto, who read this thesis very carefully and provided accurate comments that significantly improved its quality. Prof. Kschischang was one of my PhD examination committee members, and his insightful suggestions illuminated many paths for future research.

I would also like to express my sincere gratitude to my supervisory committee, Prof. Jing and Prof. Tellambura, for their valuable feedback on my research and their insightful suggestions on how to proceed. Their detailed reviews and constructive critiques were essential in refining my work and ensuring its academic rigor.

Finally, I want to thank my groupmates who have directly or indirectly contributed to this thesis. Your support has been invaluable, and I am deeply grateful.

Contents

List of Tables

List of Figures

Acronyms

achievable data rate (ADR)

amplified spontaneous emission (ASE)

bit error rate (BER)

blind phase search (BPS)

carrier phase estimator (CPE)

constant power constellation grouping (CPCG)

constant composition distribution matching (CCDM)

cross-phase modulation (XPM)

digital backpropagation (DBP)

digital signal processing (DSP)

dual reference subcarrier (DRS)

enhanced Gaussian noise model (EGN)

equalization-enhanced phase noise (EEPN)

erbium-doped fiber amplifier (EDFA)

finite impulse response (FIR)

forward error correction (FEC)

four-wave mixing (FWM)

frequency division multiplexing (FDM)

Gaussian noise model (GN)

group velocity dispersion (GVD)

intersymbol interference (ISI)

linearly polarized (LP)

Mach-Zehnder interferometer (MZI)

mean squared error (MSE)

multiple-input and multiple-output (MIMO)

nonlinear Schrödinger equation (NLSE)

non-linear interference noise (NLIN)

phase shift keying (PSK)

photodetector (PD)

power spectral density (PSD)

quadrature amplitude modulation (QAM)

raised cosine (RC)

self-phase modulation (SPM)

signal-to-noise ratio (SNR)

split-step Fourier method (SSFM)

standard single-mode fiber (SSMF)

wavelength division multiplexing (WDM)

zero-forcing phase estimator (ZFP)

Chapter 1

Introduction and Background

Fiber optics is the primary infrastructure for handling immense data traffic in the backbone of communication systems [1]. As emerging technologies demand substantial communication capacity, concerns about meeting future demands arise [2]. Consequently, enhancing fiber achievable data rate (ADR) and overcoming the limitations of this channel is crucial [3]–[5].

One of the primary factors limiting fiber ADR is fiber nonlinearity, known as the Kerr effect [6]. Increasing the power launched into the fiber to achieve a higher signal-to-noise ratio (SNR) also increases nonlinearity effects, ultimately reducing the effective SNR. This behavior fundamentally restricts the attainable SNR and, consequently, ADR. Naturally, reducing nonlinear impairments for the same launched power can increase ADR in practice, which is the goal of Chapters 2 and 3.

Another cause of fiber ADR degradation is the suboptimality of the transmitter/receiver enforced by practical considerations. In particular, we focus on the interaction between the dispersion and phase noise compensators. Although compensation for laser phase noise and dispersion ideally should occur in the reverse order of their occurrence [7], dispersion is compensated first due to practical challenges facing the carrier phase estimator (CPE). Practical CPEs utilize a sequence of known symbols, referred to as pilots, to estimate and mitigate the phase noise in the received symbol sequence. Extracting the pilots requires compensating for dispersion first, leading to a suboptimal sequence of compensations and the occurrence of equalization-enhanced phase noise (EEPN) [8]. Depending on the bandwidth and system reach of the fiber, EEPN can severely affect system performance [9]. In Chapters 4 and 5, we focus on developing two novel digital signal processing (DSP) techniques to overcome the EEPN challenge.

Before delving into the details of our completed work, we believe that reviewing the fundamentals of a fiber communication system will facilitate understanding the challenges and solutions presented in this thesis. Therefore, in this chapter, building upon electromagnetic and communication theory, we first review the fundamentals of fiber optics as a communication channel. Next, we examine the impairments associated with this channel and describe suitable transmitter and receiver structures to address these challenges. Instead of providing a full derivation, our formulation is used solely to explain the properties of fiber based on well-known communication theories. This chapter serves as a foundation for better understanding the upcoming chapters, where we discuss our solutions to the nonlinearity and EEPN challenges.

1.1 Historical Perspective

From a historical perspective, fiber communication emerged in 1966 through the pioneering work of Kao and Hockham [10], demonstrating a significant reduction in glass attenuation. Ten years later, the first communication system based on fiber optics was implemented [11]. During this era, data transmission relied on on-off keying signaling and was predominantly conducted through multimode fibers [12]. Due to fiber attenuation and the absence of inline amplifiers, the system's reach was limited to tens of kilometers.

In the 1980s, the first inline optical amplifier, known as the erbium-doped fiber amplifier (EDFA), was introduced [13]. This amplifier greatly extended the system's reach. EDFAs allowed us to amplify attenuated optical signals every 50–100 km, thereby increasing the system reach to thousands of kilometers.

Another breakthrough that occurred in the 1980s was the adoption of wavelength division multiplexing (WDM) systems, which were designed to fully utilize the available bandwidth of fiber [14]. In a WDM system, the fiber bandwidth is divided into orthogonal wavelength channels, with each channel dedicated to a user. At the time, WDM system increased the fiber ADR by a factor of 1000.

In the modern era, starting in 2000, the advent of powerful DSP units has enabled the design of intricate transmitter and receiver structures specifically tailored to address fiber channel impairments [15]. Consequently, the primary focus has shifted towards enhancing spectral efficiency through sophisticated signal processing techniques. For instance, powerful DSPs allow for electronic dispersion compensation at either the transmitter or receiver end, thereby substantially reducing the impact of nonlinearity compared to inline dispersion compensation [16].

1.2 Overview of Fiber Optic Communication

A communication system is a combination of hardware and protocols designed to transmit data between two or more entities. This structure comprises three main components: transmitter, channel, and receiver. The transmitter modifies the state of a physical channel that connects it to the distant receiver(s). Based on this channel state, the receiver estimates the transmitted data [17]. In fiber optic communication systems, the channel consists of a glass fiber that acts as a waveguide for optical pulses.

The primary advantage of using fiber as a communication channel is its low attenuation. Generally, attenuation varies with signal wavelength. Fiber optics offer exceptionally low attenuations in certain regions, notably the C-band, ranging from 1530 nm (195.2 THz) to 1560 nm (192.2 THz) [18]. A WDM system, analogous to frequency division multiplexing (FDM), divides this band into multiple frequency channels, allowing different users to transmit their data on each. These channels typically have a bandwidth between 40 GHz to 200 GHz, resulting in 20 to 100 WDM channels available within a single fiber.

WDM systems can be categorized based on different features, including the system reach. System reach is a critical parameter in WDM systems, determining the complexity of transmitters and receivers, modulation formats, critical channel impairments, and achievable data rates. Table 1.1 categorizes WDM systems based on their system reach and provides an overview of the challenges each category faces.

System Class	**Reach**	**Challenges**
Access	up to 60 km	Limited DSP
Short Reach	40-500 km	Dispersion
Long-haul	500-2000 km	Nonlinearity, Dispersion
Ultra-long-haul	2000-6000 km	Nonlinearity, Dispersion
Submarine	more than 5000 km	Severe Nonlinearity

Table 1.1: WDM system classes based on the system reach

Since the advent of fiber communication, various system structures have been utilized, including direct detection [19] and coherent detection [20]. In this dissertation, we focus on the most advanced technology: the coherent WDM system. In this structure, each channel provides at least two dimensions: in-phase and quadrature-phase components. Additionally, fiber polarization and fiber mode can be utilized to generate new dimensions.

A typical coherent WDM system structure is presented in Figure 1.1. As depicted, components in this structure operate in either the electrical or optical domains. While the frequency range of the electrical domain is around hundreds of gigahertz, the optical domain operates at frequencies of hundreds of terahertz.

In the optical domain, narrowband lasers are employed to generate a pure carrier at dedicated wavelengths for each channel. Subsequently, optical modulators, based on Mach-Zehnder interferometer (MZI), modulate the laser output according to the electrical communication signal received from the transmitter [21]. The resulting modulated light pulses are then launched into the fiber.

The fiber channel is divided into spans of 50 km to 100 km, each equipped with an EDFA to amplify the light pulse. These amplifiers compensate for fiber attenuation, enabling long-haul transmission.

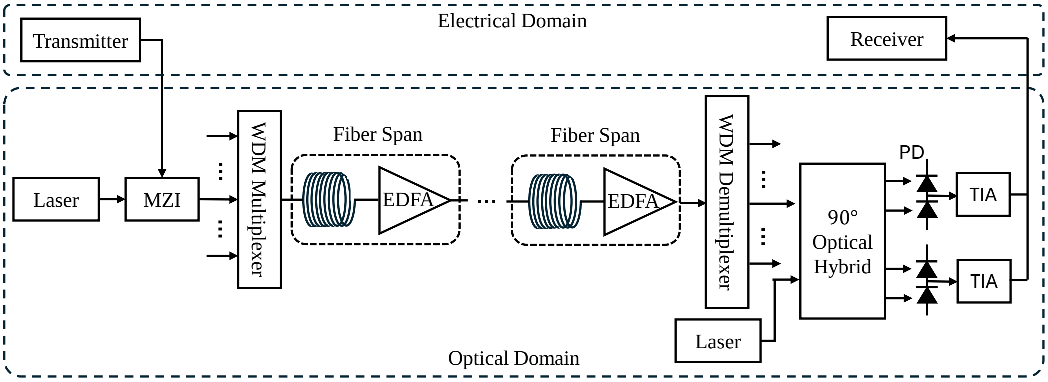


Figure 1.1: WDM structure based on the coherent receiver

Finally, at the receiver side, a combination of the local laser, photodetector (PD), and trans-impedance amplifier(TIA) extract the transmitted signal and transfer it to the electrical domain.

In the following sections, we study the properties of the different components of a WDM system and explain the suitable transmitter and receiver structures in more detail.

1.3 Fiber as a Waveguide

Optical fiber is a silica-based dielectric medium acting as a waveguide for optical pulses. An optical cable consists of three main cylindrical sections: the core, cladding, and jacket, as shown in Figure 1.2. The core has a higher refractive index, denoted as n_0, compared to the cladding, n_c. According to Snell's law, the light entering the fiber core will reflect from the interface between the core and cladding if the incident angle is less than [22]:

$$\theta_A = \arcsin\sqrt{n_0^2 - n_c^2}. \tag{1.1}$$

The total internal reflection enables the light to remain confined within the fiber core. Apart from facilitating light reflection, the cladding plays a crucial role in minimizing waveguide loss, especially when n_c is selected to be close to n_0 [23]. Although the jacket does not contribute to signal transmission, it protects the fiber against physical damage. It is crucial to acknowledge that

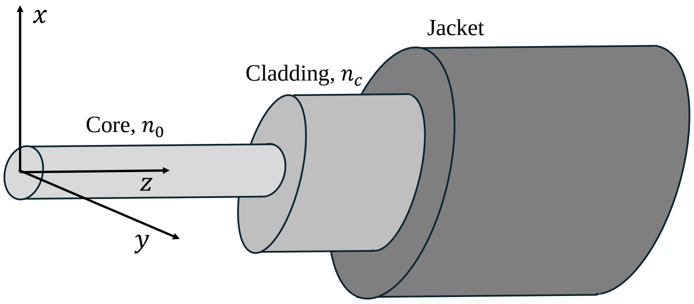


Figure 1.2: Different sections of an optical fiber

achieving a guided pulse within the fiber necessitates requirements beyond (1.1), demanding a more advanced analysis of the fiber, as presented in the following.

Like any electromagnetic wave traveling in a dielectric medium, light propagation inside the fiber is governed by Maxwell's equations, expressed as [24]:

$$\nabla \cdot \mathbf{E} = \rho_f, \tag{1.2}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{1.3}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{1.4}$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}, \tag{1.5}$$

where $\mathbf{E}$ and $\mathbf{H}$ are the electric field and magnetic field, respectively. Current density and charge density are represented by $\mathbf{J}$ and ρ_f, respectively. In addition, $\mathbf{D}$ and $\mathbf{B}$ are the flux densities resulting from the propagation of the electric and magnetic field inside the medium. For a nonmagnetic medium such as fiber, based on the constitutive relation, we have:

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}, \tag{1.6}$$

$$\mathbf{B} = \mu_0 \mathbf{H}, \tag{1.7}$$

where ε_0 represents the vacuum permittivity and μ_0 denotes the vacuum permeability, while $\mathbf{P}$ stands for the polarization density vector.

By substituting (1.6) and (1.7) into Maxwell equations, the wave equation of the electric field inside the fiber can be derived as:

$$\nabla \times \nabla \times \mathbf{E} = -\frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} - \mu_0 \frac{\partial^2 \mathbf{P}}{\partial t^2}. \tag{1.8}$$

In order to fully determine the electric field inside the fiber, a new equation relating **E** and **P** is still required. Due to the directional behavior of the silica crystal, the total polarization density vector in the fiber, **P**, satisfies the tensor relation expressed as:

$$\mathbf{P} = \varepsilon_0 \left(\chi^{(1)} \cdot \mathbf{E} + \chi^{(3)} \vdots \mathbf{EEE} \right), \tag{1.9}$$

where $\chi^{(i)}$ represents the i-th order susceptibility [25]. Due to the crystalline structure of silica, the polarization **P** is influenced by susceptibilities of odd orders exclusively. Moreover, susceptibilities higher than third order are typically deemed negligible. Equations (1.8) and (1.9) are sufficient to determine the behavior of the incident light in the fiber.

Here, we should note that the electric field propagates in $\hat{z}$ direction and can have components in $\hat{x}$, $\hat{y}$ or both directions. To simplify our discussion, let us assume that the electric field has only one component in the $\hat{x}$ direction. With this assumption, and based on (1.9), the frequency-dependent dielectric constant can be defined as:

$$\varepsilon(\omega) = 1 + \chi^{(1)} + \frac{3}{4}\chi^{(3)}_{xxxx}|E_x|^2. \tag{1.10}$$

Based on the definition of the dielectric constant $\varepsilon = (n + j\alpha c/2\omega)^2$, the refractive index, n, and absorption coefficient α of silica can be expressed as:

$$n(\omega) = 1 + \frac{1}{2}\mathrm{Re}\{\chi^{(1)}(\omega)\} + \frac{3}{8n}\mathrm{Re}\{\chi^{(3)}_{xxxx}(\omega)\}|E_x|^2, \tag{1.11}$$

and

$$\alpha(\omega) = \frac{\omega}{nc}\mathrm{Im}\{\chi^{(1)}(\omega)\} + \frac{3\omega}{4nc}\mathrm{Im}\{\chi^{(3)}_{xxxx}(\omega)\}|E_x|^2, \tag{1.12}$$

respectively [25]. As seen in (1.11), the refractive index depends not only on the frequency but also on the electric field intensity, $|E_x|^2$, introducing nonlinearity

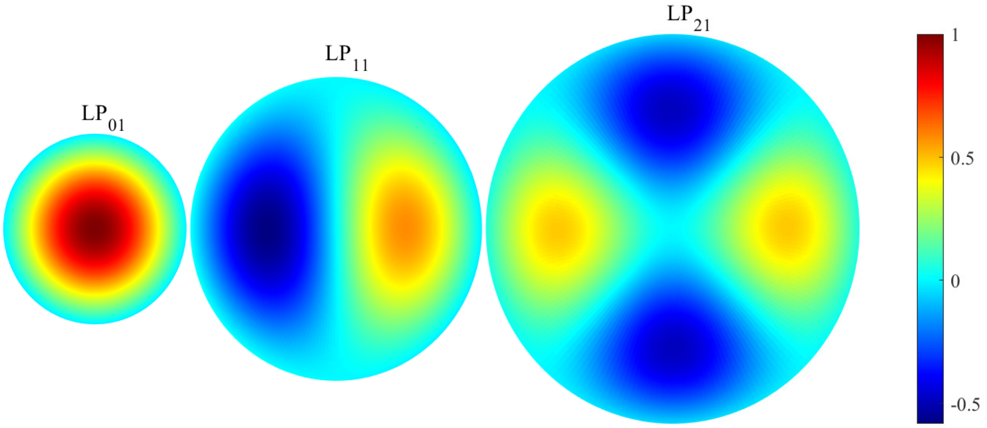


Figure 1.3: Guided modes in fiber optic. SMSF are designed to only allow linearly polarized (LP) mode of LP$_{01}$

into the system. The frequency dependency of the refractive index, $n(\omega)$, leads to chromatic dispersion, which broadens an unchirped signal in the time domain. Let us consider the plane wave equation traveling in a dispersive medium as:

$$\mathbf{E}(t,z) = \hat{x}E_0 \cos{(Bz - \omega_0 t)}, \tag{1.13}$$

where E_0, ω_0 and B are the field amplitude, wave frequency and wave number, respectively. Because of the ω-dependent propagation velocity, the wave number B is a function of frequency. To capture this dependency, the wave number is represented using a Taylor expansion around ω_0 as:

$$B(\omega) = n(\omega)\frac{\omega}{c} = \beta_0 + (\omega - \omega_0)\beta_1 + \frac{1}{2}(\omega - \omega_0)^2\beta_2 + \frac{1}{6}(\omega - \omega_0)^3\beta_3 + \cdots. \tag{1.14}$$

The Taylor coefficients β_i are called dispersion coefficients. Among these coefficients, β_1 and β_2 carry important physical meanings. β_1 determines the group velocity, while β_2 explains the broadening of the pulse along the fiber. β_2 is known as the phase group velocity dispersion (GVD) coefficient.

Even when a light wave enters the fiber, satisfying (1.2), it can either be guided by the fiber structure or radiate out of the fiber. The guided wave must exhibit oscillatory behavior inside the core and exponential decay in the cladding, while satisfying (1.8) and the boundary conditions imposed by the fiber structure. The electric field distribution on the cross-section of the fiber that satisfies these requirements is called a fiber mode and is represented by $F(x, y)$. In other words, fiber modes can be defined as the electric field intensity distribution of the light that can travel inside a cylindrical fiber for long distances. Figure 1.3 demonstrates the most important modes of the fiber. Depending on the core diameter, a fiber can support one or more modes. While fiber modes are orthogonally polarized and can be used to create orthogonal communication channels, they have different propagation velocities that can cause modal dispersion in a multi-mode pulse. In modern fibers, only the first mode is allowed by choosing the fiber core diameter to be less than 10 μm [25].

1.4 Fiber Model for Communication

Although Maxwell's equations and the wave equation sufficiently describe light propagation inside the fiber, they are complex to work with, and the insights into communication systems they provide are limited. In the following, we review the most widely accepted fiber model that serves as a tool for communication purposes.

1.4.1 Nonlinear Schrödinger Equation

In the WDM system structure depicted in Figure 1.1, both the output of the transmitter and the input of the receiver are baseband signals. From a communication perspective, the combination of the laser and MZI serves as an oscillator and mixer, upconverting the baseband signal to the dedicated frequency slot. Since the information is encoded within the baseband signal, our primary interest lies in studying the evolution of this signal along the fiber rather than focusing on the actual electric field, which necessitates the use of Maxwell's equations. This objective can be effectively achieved using the nonlinear Schrödinger equation (NLSE). In other words, the NLSE provides a simplified representation of Maxwell's equations tailored to the transfer of the baseband signal through the fiber structure [26].

Let us denote the baseband signal of the transmitter as A. Given the low-pass signal A and the fiber mode $F(x, y)$, the electric field launched into the fiber can be expressed as:

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{2}\hat{x}\left(F(x, y)A(t)\exp\left(j2\pi f_0 t\right) + c.c.\right), \tag{1.15}$$

where *c.c.* denotes the complex conjugate of the preceding terms. Substituting (1.15) into (1.8) and simplifying the resulting equation with respect to the fiber properties, NLSE can be derived as [25]:

$$\frac{\partial A}{\partial z} + \beta_1 \frac{\partial A}{\partial t} + j\frac{\beta_2}{2}\frac{\partial^2 A}{\partial t^2} + \frac{\beta_3}{6}\frac{\partial^3 A}{\partial t^3} + \frac{\alpha}{2}A = j\gamma|A|^2 A, \tag{1.16}$$

where α represents fiber attenuation, β_is are the dispersion coefficients defined in (1.14), and γ is the nonlinear parameter defined as follows:

$$\gamma(\omega) = \frac{\omega n \iint_{-\infty}^{\infty} |F(x, y)|^4 dxdy}{c \iint_{-\infty}^{\infty} |F(x, y)|^2 dxdy}. \tag{1.17}$$

It is important to note that dispersion, attenuation, and nonlinearity parameters depend on the fiber type and the dedicated wavelength band. Table 1.2 represents these parameters for non-zero dispersion-shifted fiber (NZDSF), standard single-mode fiber (SSMF), and pure-silica-core fiber (PSCF).

Fiber Type	α_{dB} (dB/km)	β_2 (s^2/m)	γ (1/W/m)
SMSF	0.2	-16.7×10^{-27}	0.013
PSCF	0.17	-20.1×10^{-27}	0.008
NZDSF	0.22	-3.8×10^{-27}	0.015

Table 1.2: Parameters for different fiber types

The NLSE consists of three terms that account for different fiber properties. The term $j\frac{\beta_2}{2}\frac{\partial^2 A}{\partial t^2}$ describes fiber dispersion, while $\frac{\alpha}{2}A$ is responsible for fiber attenuation. Finally, the term $j\gamma|A|^2 A$ arises from fiber nonlinearity. In the next section, we will study these terms and their effects on communication signals in detail.

It is important to note that (1.10) is suitable for single polarization fibers. Considering two polarizations, the effect of the other electric field components on the dielectric constant should also be considered. Doing so, (1.10) can be rewritten as:

$$\begin{aligned}\varepsilon_x(\omega) &= 1+\chi^{(1)}+\frac{3}{4}\chi^{(3)}_{xxxx}\left(|E_x|^2+\frac{2}{3}|E_y|^2\right),\\ \varepsilon_y(\omega) &= 1+\chi^{(1)}+\frac{3}{4}\chi^{(3)}_{xxxx}\left(\frac{2}{3}|E_x|^2+|E_y|^2\right).\end{aligned} \tag{1.18}$$

Following the same procedure as for single polarization fibers, the NLSE for a two-polarization fiber becomes a coupled equation as:

$$\begin{aligned}\frac{\partial A_x}{\partial z}+\beta_1\frac{\partial A_x}{\partial t}+j\frac{\beta_2}{2}\frac{\partial^2 A_x}{\partial t^2}+\frac{\beta_3}{6}\frac{\partial^3 A_x}{\partial t^3}+\frac{\alpha}{2}A_x &= j\frac{8}{9}\gamma\left(|A_x|^2+|A_y|^2\right)A_x,\\ \frac{\partial A_y}{\partial z}+\beta_1\frac{\partial A_y}{\partial t}+j\frac{\beta_2}{2}\frac{\partial^2 A_y}{\partial t^2}+\frac{\beta_3}{6}\frac{\partial^3 A_y}{\partial t^3}+\frac{\alpha}{2}A_y &= j\frac{8}{9}\gamma\left(|A_x|^2+|A_y|^2\right)A_y,\end{aligned} \tag{1.19}$$

This version of the NLSE, observing a two-polarization fiber, is referred to as the Manakov equations [27].

The NLSE can easily explain fiber impairments important for communication systems, as we do in Section 1.5. However, in the general case, the closed-form solution of the NLSE is not known. Different approximations or numerical methods are proposed, such as the Volterra series [28] and the split-step Fourier method (SSFM) [29]. The latter method is studied in the following subsection.

1.4.2 Numerical Solver

Since the NLSE does not have a closed-form solution, we need to use a numerical solver to find the output waveform of a fiber. The SSFM is the most widely used numerical solver for NLSE. In this section, we explain the SSFM numerical method, its accuracy, and important parameters [29]–[31]. All the simulation results in the upcoming chapters are based on this method.

To explain the SSFM, we need to define two linear and nonlinear operators as:

$$D=-\beta_1\frac{\partial}{\partial t}-j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}+\frac{\beta_3}{6}\frac{\partial^3}{\partial t^3}-\frac{\alpha}{2}, \tag{1.20}$$

and

$$N=j\gamma|A|^2, \tag{1.21}$$

respectively. Using these operators, the NLSE expressed in (1.16) can be rewritten as follows:

$$\frac{\partial A}{\partial z} = (D + N)A. \tag{1.22}$$

Let us assume D and N are two z-independent operators defined over a step size of h. Then the solution to this equation can be written as:

$$A(t, z + h) = \exp\left((D + N)h\right)A(t, z). \tag{1.23}$$

Based on the fact that D and N are noncommuting operators, the exponential of their sum can be found using the Baker–Hausdorff formula as [32]:

$$\begin{aligned} \exp(Dh)\exp(Nh) = \exp\Bigg(& Dh + Nh + \frac{1}{2}[Dh, Nh] + \\ & \frac{1}{12}[Dh - Nh, [Dh, Nh]] + \cdots \Bigg), \end{aligned} \tag{1.24}$$

where $[a, b] = ab - ba$. Considering h small, we can ignore all the right-hand side terms except for $Dh + Nh$. Therefore, for a small step size, (1.22) can be written as:

$$A(t, z + h) = \exp(Dh)\exp(Nh)A(t, z). \tag{1.25}$$

From this equation, we can infer that over a short step size, the linear and nonlinear operators defined in (1.20) and (1.21) can act separately. Additionally, we should note that the linear operator can operate easily in the frequency domain by replacing $\frac{\partial}{\partial t}$ with $-j\omega$ in (1.20) to obtain:

$$D = -j\beta_1\omega + j\frac{\beta_2}{2}\omega^2 - j\frac{\beta_3}{6}\omega^3 - \frac{\alpha}{2}. \tag{1.26}$$

Therefore, from a complexity point of view, the linear operator needs to operate in the frequency domain, while the nonlinearity operates in the time domain. As a result, for each step of the SSFM, transferring between time and frequency domains is required, resulting in considerable computation. Therefore, efficient implementation of the SSFM and choosing the optimal step size is of great importance.

The step size determines the accuracy of SSFM in solving NLSE. It is crucial to choose a step size smaller than the walk-off length between the frequency channels in the WDM system. Figure 1.4 demonstrates the maximum

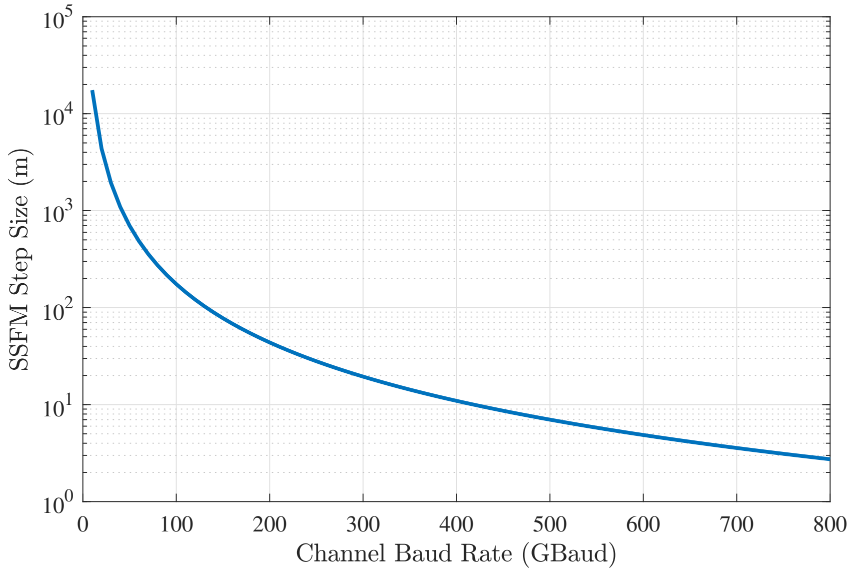


Figure 1.4: Uniform step size for SSFM to accurately simulate a WDM system with 5 channels. These step sizes are determined so that the walk-off length remains larger than the SSFM step size.

step size for different channel numbers and channel bandwidths that guarantee SSFM accuracy.

1.5 Fiber Channel Impairments

From a communication standpoint, important fiber impairments can be divided into two categories: linear and nonlinear. In addition to the fiber impairments, component noises also impact the WDM system. Figure 1.5 illustrates a flowchart depicting these impairments. In the following section, building upon the discussions in the previous sections, we will explain fiber impairments. Understanding these impairments is crucial for designing an effective fiber communication system.

1.5.1 Fiber Linear Impairments

Linear impairments, such as attenuation and dispersion, play a significant role in defining the primary properties of fiber as a communication channel, including channel bandwidth and system reach. However, because of their linearity, these impairments can be easily compensated for without imposing

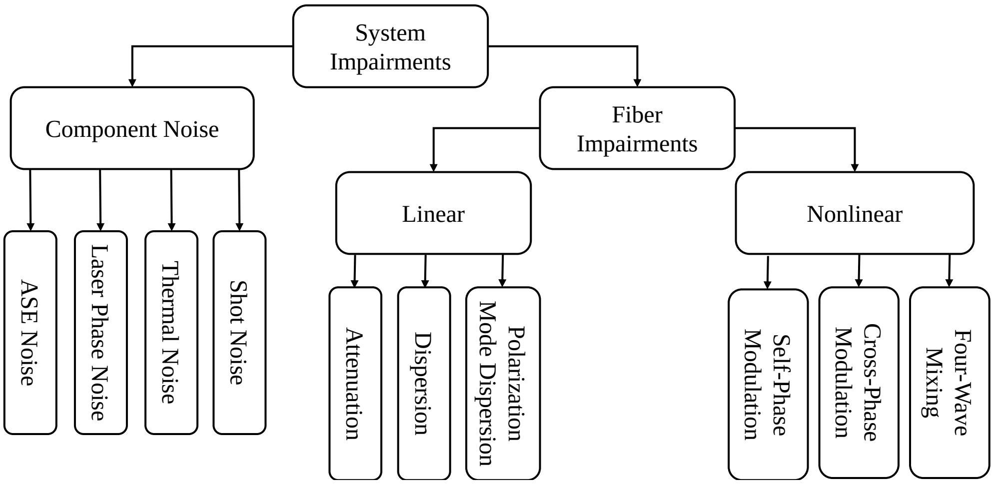


Figure 1.5: Various impairments in a WDM system. This chart is based on the discussions in [25].

critical limitations on the achievable spectral efficiency.

Attenuation

Attenuation causes the signal to lose its power as it propagates inside the fiber. Fiber attenuation stems from intrinsic material absorption, extrinsic impurities absorption, and scattering [23].

As mentioned, the very low attenuation of silica fiber (around 0.2 dB/km) enables long-reach transmission. Figure 1.6 presents the fiber attenuation at different wavelengths. In this figure, the peak at 1390 nm is due to the OH ions as impurities in the fiber [25]. As observed, the fiber provides low attenuation in two regions, making it perfect as a communication channel. The first region is the O-band centered at $\lambda = 1300$ nm. This band was initially preferred for optical systems, as fiber dispersion is small in this region. The other region with low attenuation is centered at $\lambda = 1550$ nm, known as the C-band. Modern systems typically utilize this band. Notably, the C-band offers even lower attenuation compared to the O-band. It spans from $\lambda = 1536$ nm to $\lambda = 1565$ nm, providing a bandwidth of 3.6 THz.

To gain a better understanding of the impact of attenuation on communication signals, let's determine the gain of the EDFA for a typical span of 100 km as an example. To do this, we neglect dispersion and nonlinearity in

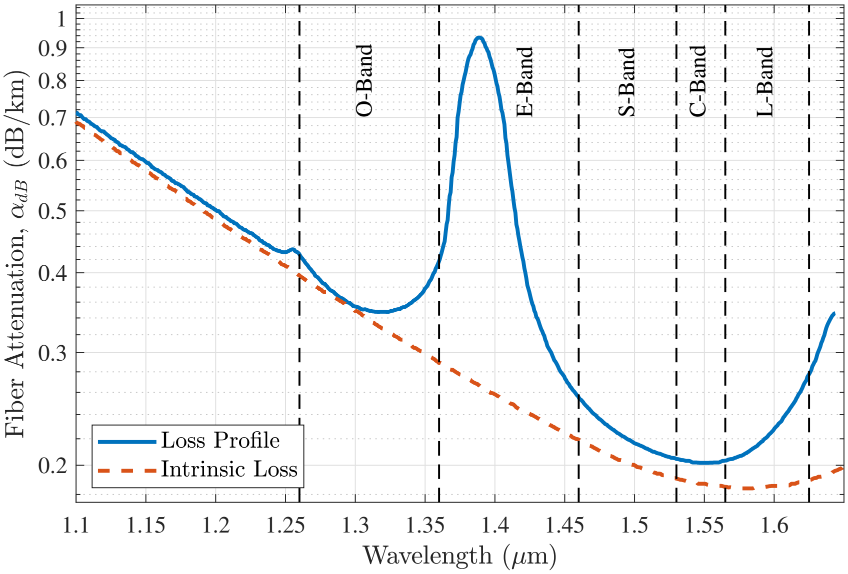


Figure 1.6: Based on the fiber loss, O-band and C-band are chosen for communication applications [25]

(1.16), simplifying the NLSE to:

$$\frac{\partial A}{\partial z} + \frac{\alpha}{2} A = 0. \tag{1.27}$$

For this case, the output of the fiber is:

$$A(t, z) = A(t, 0) \exp\left(-\frac{\alpha}{2} z\right). \tag{1.28}$$

As observed, the fiber exponentially attenuates the signal amplitude. If the fiber attenuation is $\alpha = 0.2$ dB/km, then over every 100 km, the fiber attenuates the signal by a factor of 10. Therefore, amplifiers need to amplify the signal for every span by the same factor.

We should note that in the NLSE, α has the dimension of m^{-1}. On the other hand, fiber attenuation is usually reported as α_{dB} with the dimension of dB/km. To convert α_{dB} to α, we can use the following relation based on (1.28):

$$\alpha_{dB} = -\frac{10}{L} \log_{10}\left(\frac{P_{\mathrm{out}}}{P_{\mathrm{in}}}\right) = -\frac{10}{1000} \log_{10}\left(\exp\left(-\frac{\alpha}{2} \times 1000\right)^2\right) = 4.34\alpha. \tag{1.29}$$

Dispersion

In the following, using the dispersion coefficients β_1 and β_2 defined in (1.14), we explain the main implications of the dispersion, including group velocity, pulse broadening, and walk-off phenomenon.

Group Velocity indicates the velocity at which the envelope of the electric field propagates. Group velocity is determined by β_1. To better explain the impact of β_1, we simplify the NLSE to only include β_1, resulting in:

$$\frac{\partial A}{\partial z} + \beta_1 \frac{\partial A}{\partial t} = 0. \tag{1.30}$$

The solution of this equation can be easily found as:

$$A(t, z) = A(t + \beta_1 z, 0). \tag{1.31}$$

As observed, the envelope of the electric field, A, propagates with the velocity of $v_g = \frac{1}{\beta_1}$.

Pulse Broadening is a direct implication of the GVD parameter, β_2. Considering only β_2 in (1.16), the NLSE can be rewritten as:

$$\frac{\partial A}{\partial z} + j\frac{\beta_2}{2}\frac{\partial^2 A}{\partial t^2} = 0. \tag{1.32}$$

By transforming this equation into the frequency domain, the solution can be easily found as:

$$\widetilde{A}(\omega, z) = \widetilde{A}(\omega, 0) \exp\left(-j\frac{\beta_2}{2}\omega^2 z\right). \tag{1.33}$$

Based on this solution, we can conclude that dispersion does not change the signal power spectral density (PSD), but it does alter the temporal pattern.

As an example, let us consider that the launched pulse has a Gaussian shape and is defined as:

$$A(t, 0) = \exp\left(-\frac{t^2}{2T^2}\right). \tag{1.34}$$

Inserting (1.34) into (1.33), the output of the fiber for a Gaussian pulse is:

$$A(t, z) = \frac{T}{\sqrt{T^2 - j\beta_2 z}} \exp\left(-\frac{t^2}{2(T^2 - j\beta_2 z)}\right). \tag{1.35}$$

As observed, dispersion broadens this pulse by a factor of $\sqrt{1 + z|\beta_2|/T^2}$.

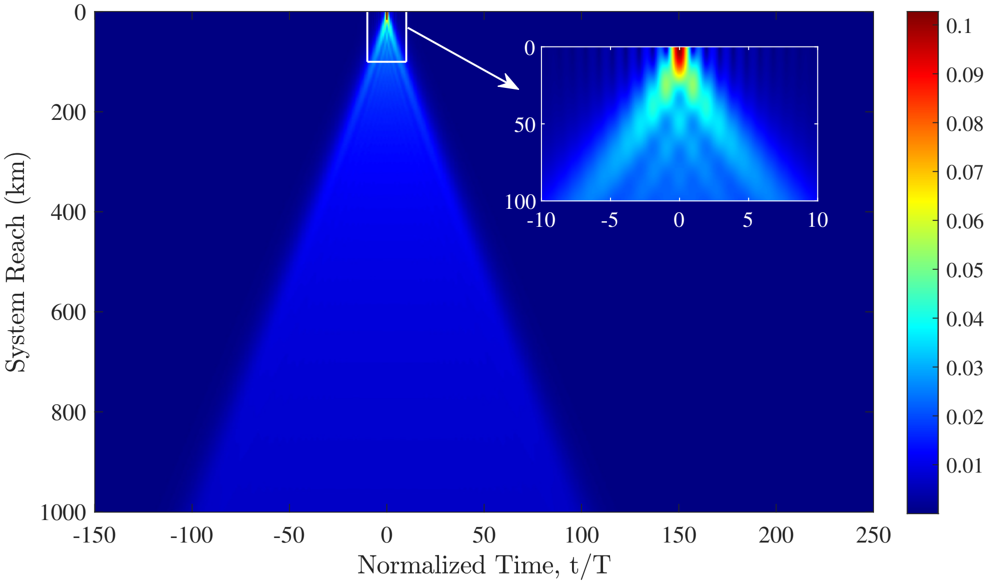


Figure 1.7: Temporal evolution of the raised cosine pulse along a dispersive channel.

As another example, Figure 1.7 represents the temporal evolution of the raised cosine signal along a dispersive channel defined by (1.32). Based on this simulation, the following can be concluded:

- Dispersion reduces the peak power by broadening the pulse energy in the time domain.
- After a few spans, the pulse shape approaches a Gaussian pulse.
- Pulse width broadens linearly with the system reach.

Walk-Off Phenomenon is another implication of the GVD parameter, β_2. GVD causes different frequency channels to have different group velocities proportional to their central frequency. To explain this phenomenon, let us consider two frequency channels as:

$$A(t,0) = A_1(t,0)\exp(\omega_1 t) + A_2(t,0)\exp(\omega_2 t). \tag{1.36}$$

Inserting (1.36) into (1.32) and separating the frequency channels by filtering, we can write the equivalent separated channels as:

$$\frac{\partial A_i}{\partial z} - j\frac{\beta_2}{2}\omega_i^2 A_i - \omega_i\beta_2\frac{\partial A_i}{\partial t} + j\frac{\beta_2}{2}\frac{\partial^2 A_i}{\partial t^2} = 0, \tag{1.37}$$

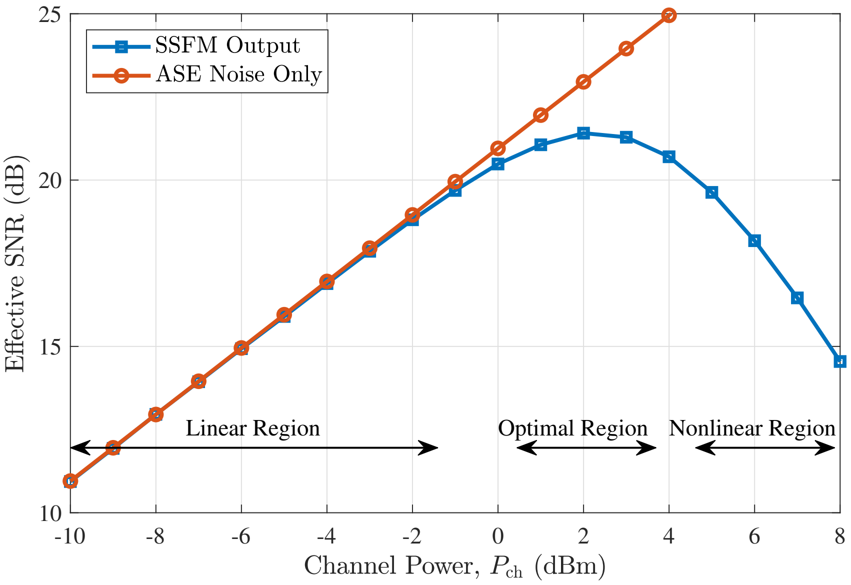


Figure 1.8: Effective SNR of a typical fiber optic communication system.

for $i = 1, 2$. As observed, for frequency channels, there exists a group velocity proportional to the carrier frequency, $-\omega_i\beta_2$, even if $\beta_1 = 0$. This phenomenon causes the WDM channels to have different propagation velocities.

Let us consider two pulses with duration T at different frequencies ω_1 and ω_2. Due to the walk-off phenomenon, these two pulses pass each other over a distance given by:

$$d_{\text{walk-off}} = 2T \left| \frac{1}{\omega_1\beta_2} - \frac{1}{\omega_2\beta_2} \right| \tag{1.38}$$

meters. This distance is known as the walk-off length [25].

1.5.2 Kerr Non-Linear Impairments

As mentioned earlier, the power dependency of the reflective index (1.11), causes Kerr nonlinearity. Due to the Kerr nonlinearity, SNR in the fiber is not a monotonic function of channel power. In fact, increasing the channel power increases the nonlinearity effects and eventually reduces the effective SNR. As an example, Figure 1.8 demonstrates the effective SNR for different channel powers. As observed, based on the channel launch power, we can investigate the effective SNR in three regions. In this figure, for powers less than -2 dBm, the system is in the linear region where the dominant impairment is the amplified spontaneous emission (ASE) noise. It is evident that with increasing channel power, the effective SNR increases accordingly. For channel powers around 2 dBm, the maximum achievable SNR is attained. The optimal channel power refers to the power that achieves the maximum achievable SNR. Finally, for channel powers higher than 4 dBm, the nonlinearity effects dominate and reduce the effective SNR.

The primary fiber impairments arising from Kerr nonlinearities include self-phase modulation (SPM), cross-phase modulation (XPM), and four-wave mixing (FWM). In the following, we will explain the mechanisms and impacts of these impairments [23], [25].

Self-Phase Modulation

The primary intensity-dependent nonlinear effect in fiber systems is SPM. Considering a WDM channel, SPM is an intensity-dependent phase shift that affects the channel proportionally to its intensity [23].

Let us assume a zero dispersion and lossless fiber for which NLSE becomes:

$$\frac{\partial A}{\partial z} = j\gamma |A|^2 A. \tag{1.39}$$

Without loss of generality, we consider the solution to this equation as:

$$A(t, z) = V(t, z) \exp(j\phi(t, z)), \tag{1.40}$$

where V and ϕ are two real-valued functions representing the envelope amplitude and phase, respectively. By inserting this into (1.39) and separating the real and imaginary parts of the result, we obtain:

$$\frac{\partial V}{\partial z} = 0, \tag{1.41}$$

and

$$\frac{\partial \phi}{\partial z} = j\gamma |V|^2. \tag{1.42}$$

As observed, while the amplitude of the time domain representation remains unchanged along the fiber, the spectrum of the pulse changes due to the new phases generated by SPM [25].

The interaction between SPM and dispersion is also of great importance. In a dispersive fiber ($\beta_2 < 0$), dispersion causes the time domain representation of the pulse to broaden. Therefore, it reduces the pulse power, thereby

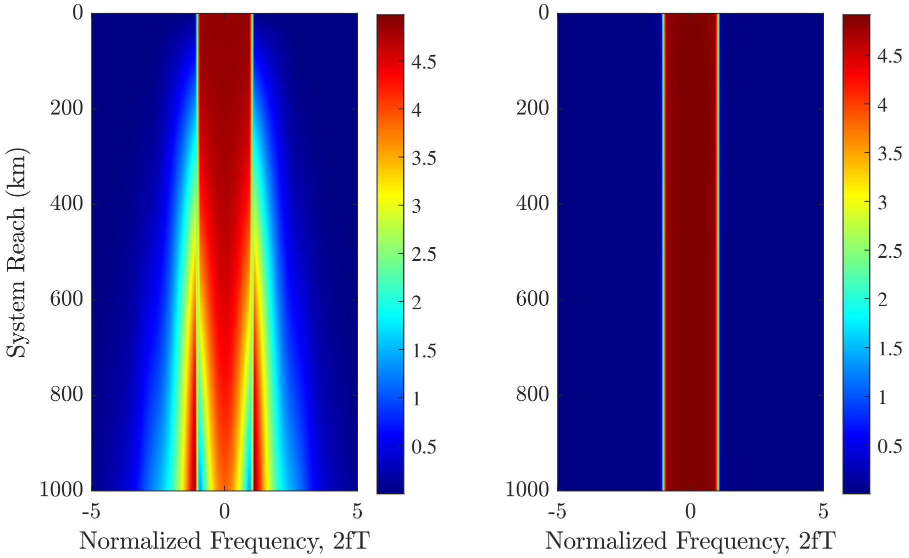


Figure 1.9: Evolution of the raised cosine pulse in dealing with SPM (right) with and (left) without dispersion.

weakening SPM. Figure 1.9 demonstrates the propagation of the same pulse used in Figure 1.7 in the presence of both SPM and dispersion. As observed, spectral broadening due to SPM is reduced considerably in the presence of dispersion. This is why, since 2010, dispersion-uncompensated links have been widely used due to their lower nonlinear effects.

It is important to note that, on average, SPM causes a phase rotation in the received constellation as follows:

$$\Delta\phi_{SPM} = \gamma P_{\mathrm{ch}} L, \tag{1.43}$$

where P is the WDM channel power and L is the fiber length. This phase rotation necessitates the use of the phase estimation in a fiber receiver when nonlinearity presents.

Cross-Phase Modulation

In contrast to SPM, where phase modulation is caused by the intensity of the channel itself, XPM involves phase modulation due to the intensity of neighboring channels [33].

To derive the XPM formulation, we need to consider at least two frequency channels, as follows:

$$A = A_1 \exp(-\omega_0 t) + A_2 \exp(\omega_0 t). \tag{1.44}$$

Substituting (1.44) into (1.16) and filtering the frequency channels, we obtain [34]:

$$\frac{\partial A_1}{\partial z} + \beta_{11}\frac{\partial A_1}{\partial t} + j\frac{\beta_{21}}{2}\frac{\partial^2 A_1}{\partial t^2} - \frac{\beta_{31}}{6}\frac{\partial^3 A_1}{\partial t^3} + \frac{\alpha}{2}A_1 = j\gamma(|A_1|^2 + 2|A_2|^2)A_1, \tag{1.45}$$

$$\frac{\partial A_2}{\partial z} + \beta_{12}\frac{\partial A_1}{\partial t} + j\frac{\beta_{22}}{2}\frac{\partial^2 A_2}{\partial t^2} - \frac{\beta_{32}}{6}\frac{\partial^3 A_2}{\partial t^3} + \frac{\alpha}{2}A_2 = j\gamma(2|A_1|^2 + |A_2|^2)A_2. \tag{1.46}$$

In (1.45), the term $|A_1|^2$ is responsible for SPM, while the term $2|A_2|^2$ represents XPM. As observed, the impact of XPM on phase modulation is twice that of SPM. It is important to note that XPM only changes the phase of neighboring channels and does not transfer energy between these channels.

Four-Wave Mixing

FWM is a scattering process in which power from four frequency channels can interact with each other [35]. To understand the origin of FWM, we need to start from (1.9). As seen in the tensor-based definition of the polarization density vector, three electric field components can affect the polarization density vector, affecting the fourth electric field.

Considering the electric field with four frequency components as [36]:

$$\mathbf{E} = \frac{1}{2}\hat{x}\sum_{k=1}^{4} E_k \exp\big(j(\beta^{(k)}z - \omega_k t)\big) + c.c., \tag{1.47}$$

we can write the polarization density vector affecting the fourth electric field as:

$$\begin{aligned} P_4 = \frac{3\varepsilon_0}{4}\chi^{(3)}_{xxxx}\big[&|E_4|^2E_4 + 2(|E_1|^2 + |E_2|^2 + |E_3|^2)E_4 + 2E_1E_2E_3\exp(i\theta_+) \\ &+ 2E_1E_2E_3^*\exp(i\theta_-) + \cdots\big], \end{aligned} \tag{1.48}$$

where:

$$\theta_+ = (\beta^{(1)} + \beta^{(2)} + \beta^{(3)} - \beta^{(4)}) - (\omega_1 + \omega_2 + \omega_3 - \omega_4), \tag{1.49}$$

$$\theta_- = (\beta^{(1)} + \beta^{(2)} - \beta^{(3)} - \beta^{(4)}) - (\omega_1 + \omega_2 - \omega_3 - \omega_4). \tag{1.50}$$

While in (1.48), $|E_4|^2E_4$ and $2(|E_1|^2+|E_2|^2+|E_3|^2)E_4$ are responsible for SPM and XPM, respectively, $2E_1E_2E_3\exp(i\theta_+)$ and $2E_1E_2E_3^*\exp(i\theta_-)$ are known as FWM. These terms significantly affect the system only when the phase-matching condition is satisfied. The phase matching for θ_+ is hard to achieve, but for θ_-, it is enough to have:

$$\beta^{(1)} + \beta^{(2)} = \beta^{(3)} + \beta^{(4)}, \tag{1.51}$$

$$\omega_1 + \omega_2 = \omega_3 + \omega_4. \tag{1.52}$$

This requirement can be easily met in the case where $\beta_2 = 0$. Therefore, FWM can affect the system performance when zero-dispersion fibers are used. However, the phase-matching condition is typically not satisfied in other practical scenarios, rendering FWM negligible.

1.5.3 Component Noise

In practice, WDM components, including transceivers, receivers, lasers, and amplifiers, are not ideal and introduce noise to the system. In the following, we will explain the important sources of noise in a fiber system.

ASE Noise

EDFAs are used as a main component in WDM systems to amplify the signal after each span. The optical amplifier also adds ASE noise to the light pulse [37]. The amount of this noise is usually characterized by the EDFA noise figure, *NF*.

The ASE noise can be modeled by a circular symmetric Gaussian process with the following PSD [38]:

$$G_{\mathrm{ASE}} = NFh\frac{c}{\lambda}(G-1), \tag{1.53}$$

where h is Planck's constant, λ is the light wavelength, and G is the EDFA gain.

Laser Phase Noise

The lasers used in transmitters and coherent receivers introduce phase noise to the optical system. The laser phase noise follows a Wiener process, and

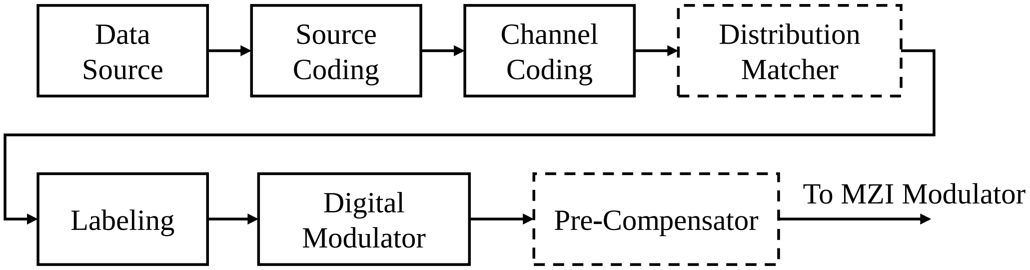


Figure 1.10: Transmitter structure. Dashed blocks are optional with respect to the system properties

consequently, the output of the laser exhibits a Lorentzian spectrum with a linewidth commonly denoted as $\partial\nu$. Therefore, we can model the laser phase output as [39]:

$$\phi = \int_{-\infty}^{t} n(\nu) d\nu \tag{1.54}$$

where n is a Gaussian random process with zero mean and a variance of $\partial\nu/2\pi$.

1.6 Electrical Domain

In any communication system, the transmitter and receiver should be designed to cope with channel impairments efficiently. Fiber transmitters and receivers, in particular, are designed to handle the impairments we discussed earlier in this chapter. In the following, we review the transmitter and receiver structures suitable for fiber communication [18], [23], [25], [40].

1.6.1 Transmitter

The primary goal of the transmitter is to map the input sequence of data into a certain set of waveforms to be sent into the fiber channel. Figure 1.10 demonstrates a typical fiber transmitter structure. Below, we briefly review the components required on the transmitter side.

Similar to any other communication channel, source coding is the first stage in fiber transmitters. Source coding removes unwanted redundancies in the input data and represents the data in binary format [41]. Conversely, channel coding adds redundancy to the bit stream to detect and correct any errors caused by channel impairments [42].

The distribution of the transmitted signal can affect channel quality and capacity. The distribution matcher block changes the source distribution to the desired distribution. In fiber optics, the transmitted distribution can be optimized to achieve different goals, such as:

- Increasing the QAM achievable data rate by shaping the distribution to capture the Gaussian distribution [43]–[46].
- Reducing channel nonlinearity effects [47]–[49].

Changing the distribution of the source is mainly referred to as probabilistic shaping [50], [51].

Next, the mapping block converts the bit sequence into a constellation. A constellation is a set of points in an N-dimensional space that uniquely represents a block of bits. Analogous to probabilistic shaping, constellation points can be chosen so that the desired distribution fits the output signal [52]–[54]. This method is referred to as geometric shaping. The dimension of the constellation depends on the orthogonal subspaces that the channel can provide. In fiber optics, different features, including phase, frequency, polarization, mode, and time, can be used to generate orthogonal bases.

Constellation points are still discrete values that need to be modulated over a band-limited pulse shape. In other words, a pulse shape transforms the discrete constellation points into a continuous domain suitable for transmitting over the channel. Based on classical communication theory, we want the pulse shape to be band-limited and satisfy the Nyquist zero-intersymbol interference (ISI) criterion. The widely used pulse shape in fiber optics is the root-raised cosine pulse, defined as:

$$G(f) = \begin{cases} 1 & \text{if } |f| \leq \frac{1-\beta}{2T} \\ \sqrt{\frac{1}{2}\left[1 + \cos\left(\pi \frac{T}{\beta}\left(|f| - \frac{1-\beta}{2T}\right)\right)\right]} & \text{if } \frac{1-\beta}{2T} < |f| \leq \frac{1+\beta}{2T} \\ 0 & \text{otherwise,} \end{cases} \tag{1.55}$$

where T is the symbol period and β is the pulse roll-off factor [17].

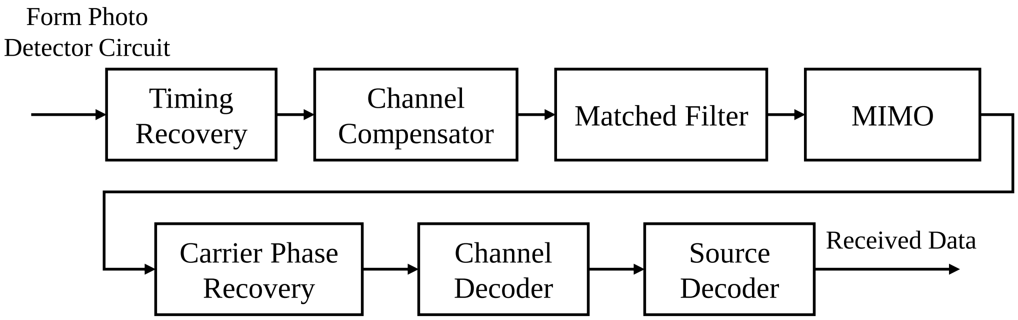


Figure 1.11: A typical coherent receiver structure suitable for fiber optic applications

In some applications, a pre-compensator is implemented in the fiber transmitter. This block can be used to compensate for transmitter laser phase noise or to collaborate with the receiver to compensate for fiber dispersion [55]–[57].

1.6.2 Coherent Receiver

In general, a coherent receiver is a receiver with a local oscillator capable of extracting the in-phase and quadrature components of the signal [20], [58]. In fiber optic systems, in addition to the in-phase and quadrature-phase extraction, a local laser is used to:

- Produce enough power to drive the photodetector,
- Downconvert the received signal, enabling the ADCs to work in the baseband.

A typical fiber coherent receiver is presented in Figure 1.11.

As seen, at the beginning of the receiver chain, timing information is extracted based on the Gardner [59] or Godard Algorithm [60]. This is important because the performance of the rest of the receiver chain depends on the correct timing information.

Receiver algorithms, such as phase recovery, are based on pilots. To enable these algorithms, dispersion needs to be compensated for first. Therefore, a channel compensator is implemented before these blocks. The channel compensator should at least compensate for dispersion [61]. It is important to note that due to the long impulse response caused by dispersion, the dispersion compensator requires the most computation and power resources. Digital backpropagation [1], as the most effective channel compensation method, should also be implemented here [62]–[64]. This compensator can compensate for both dispersion and nonlinearity at the cost of intractable complexity.

In the next step, the matched filter is implemented to extract the received symbols. The matched filter, in addition to the pulse shape used on the receiver side, takes the effect of channel filtering into account [17].

The multiple-input and multiple-output (MIMO) block is responsible for determining the state of polarization and extracting the polarization channels [65], [66]. It can also exploit the diversity in multi-mode fiber systems to increase fiber capacity [67].

Next, the pilot-based carrier phase recovery compensates for the transmitter and receiver laser phase noises. Also, this block compensates for the phase rotation caused by nonlinearity [68]–[70].

Finally, the received symbols are fed to the channel decoder and source decoder to undo the effects of their counterpart blocks and extract the raw data fed to the transmitter.

1.7 Thesis Overview

In this chapter, we reviewed fiber impairments and the structure of existing transmitters and receivers designed to cope with them. While fiber impairments and their mechanisms have been extensively explored, creating a communication system that can handle these impairments simultaneously remains an open challenge. In particular, we focus on the interaction between I) dispersion and nonlinearity and II) dispersion and laser phase noise, aiming to improve signaling techniques.

The first part of my work focuses on nonlinearity control in the presence of dispersion. In this scenario, a practical nonlinearity compensator is not yet

---

[1] Digital backpropagation is a compensation technique where the received signal is launched into a fiber with parameters of the opposite sign. Since a fiber with negative nonlinearity does not exist, this fiber is implemented digitally by numerically solving NLSE.

known. Therefore, reducing the nonlinearity impairment is of great importance. Based on the Gaussian noise model (GN) [71] and enhanced Gaussian noise model (EGN) [72], constellation and pulse shape impact fiber nonlinearity. Consequently, a natural question arises regarding how constellations and pulse shape should be designed to minimize nonlinearity.

In Chapter 2, we demonstrate that the fiber nonlinearity effect is dependent on the accumulated power launched into the WDM channels. We model the accumulated launched power and investigate the impact of constellation design on fiber nonlinearity effects. Through mathematical analysis, we show that controlling the fluctuations in the launched power enables achieving the same noise power at longer distances. Building upon this concept, we introduce a grouping technique for digital subcarrier systems aimed at reducing nonlinearity.

In Chapter 3, we first employ the calculus of variations to identify the optimal band-limited pulse shape that minimizes fiber nonlinearity effect based on the GN model. However, for practical considerations, a pulse shape must also fulfill other communication requirements such as zero-ISI and rapid decay over time. To address this, we develop a general pulse shape model and formulate the problem of finding the minimum non-linear interference noise (NLIN) pulse, subject to the mentioned constraints, as a convex optimization problem. By solving this optimization problem, we introduce the optimal pulse shape aimed at reducing the nonlinearity effect in fiber optical communication systems.

In the second part of my work, I focus on equalization-enhanced phase noise. This impairment occurs due to the interaction between the receiver laser phase noise and the dispersion compensator. Generally, phase noise and dispersion operations do not have an associative property; therefore, compensating for these impairments should be done in the opposite order of their occurrence. However, as we discussed in Section 1.6.2, compensating for dispersion before pilot-based DSP blocks is crucial. Therefore, the dispersion compensator comes after the phase noise estimator and carrier phase recovery. This suboptimal order of compensators causes the phase noise to be amplified in the dispersion compensator, causing EEPN [9].

In Chapter 4, we demonstrate that EEPN can be mitigated through signal processing. We derive the compensation expression and propose two different compensators depending on the availability of the receiver phase noise. Our study shows that by employing a simple time-variant finite impulse response filter, one can effectively compensate for EEPN.

In Chapter 5, we introduce a novel carrier phase recovery that not only estimates receiver phase noise distinguished from the transmitter phase noise but also can operate before the dispersion compensator. These two unique features allow for the optimal order of phase and dispersion compensation, avoiding EEPN.

Finally, Chapter 6 serves as a conclusion to this discussion and offers suggestions for future research directions.

Chapter 2

Constant Power Constellation Grouping

2.1 Introduction

Based on existing fiber channel models such as GN [71] or EGN [72], NLIN depends on the behavior of the power launched into the fiber. Reducing the power or its fluctuation can enhance system immunity against nonlinearity impairments. This is why many studies use phase shift keying (PSK) signaling to increase system reach due to its constant launched power [73]–[75]. PSK constellations, however, suffer from a low ADR. As a solution, using a 4-D PSK to achieve better data rates is proposed [76]. However, generating and detecting 4-dimensional PSKs needs high hardware and computational complexity. This is because high-resolution analog-to-digital converters and sophisticated timing recovery between the x- and y-polarizations are required.

Shaping methods can also be used to control power fluctuation and reduce the fiber nonlinear noise. Shaping gain for nonlinear fiber is studied in [77], where the authors suggest hyper-ball shaped constellations implemented on different time slots. This shaping strategy is explored further in [51] and the authors suggest moving to 4-dimensional symbols instead of only using the time dimension. Many works follow this path and try to propose multidimensional symbol shaping methods to enhance the nonlinear behavior of the fiber [78]–[81]. While these methods are developed based on the memory-less assumption of the fiber channel, in [82], a finite state machine source is proposed to further increase the shaping gain by taking the fiber memory into account. Based on the memory assumption, one can also suggest constant composition distribution matching (CCDM) to reduce the power fluctuation and decrease the fiber nonlinearity effects since the energy of any symbol blocks is constant [83]. Unfortunately, because the fiber memory is limited, the CCDM block size cannot be large enough and therefore, the tradeoff between nonlinear gain and the capacity loss limits the CCDM benefits. Another shaping strategy stems from the EGN model, where it has been observed that the constellation's 4th moment, called kurtosis, has a great impact on the fiber nonlinear noise [72]. To control the nonlinear noise, many works suggest developing temporal shaping methods to limit the constellation kurtosis. In [84], a kurtosis-limited sphere shaping method is proposed. Despite the nonlinear gain that shaping methods achieve, these methods, including multidimensional ball-shaped constellations or kurtosis-limited modulations, restrict themselves to only one WDM channel and neglect the potential of using neighboring channels to reduce the nonlinearity effects.

Spherical codes are another well-known approach to reducing the peak-to-average power ratio and mitigating fiber nonlinearity impairments. Spherical codes are defined as a set of codewords equidistant from the origin in an N-dimensional space [85]. These codewords are selected such that the projection of each codeword onto the others is minimized. While there has been significant work done on spherical codes [86], [87], the direct relationship between the bit sequence in spherical codes and the power fluctuations of the signal launched into the fiber is not well defined.

In this chapter, we propose controlling the power fluctuation by designing a new signaling technique that we refer to as constant power constellation grouping (CPCG). We first use NLSE to show that the fiber nonlinearity effect depends on the accumulated power of all the WDM channels. Therefore, to control the nonlinearity impairments, we can reduce the fluctuation of the accumulated power launched into the WDM channels by correlating the symbols launched into the WDM systems. To this end, we combine the constellations that correspond to these channels in such a way that the accumulated power of the transmitted symbols remains constant for any system use. We analytically show that using the CPCG technique, one can achieve the same noise power at longer distances in comparison with the quadrature amplitude modulation (QAM) family. Also, launching constant power enhances the system ADR for the same distance.

The CPCG technique can be applied to many base constellations. Furthermore, shaping techniques can be used along with CPCG. In this work, we implement the CPCG technique for the 16-QAM constellation. We also conduct a simulation study to investigate our proposed signaling method. Simulation results confirm our theoretical findings. We show that our CPCG technique on a system based on 16-QAM constellation increases the system reach and that for long system reaches, CPCG increases the system ADR. We also show that compared to PSK, for the same system reach, our solution offers a higher system ADR.

The rest of this chapter is organized as follows: In Section 2.2, we describe the system model. Section 2.3 finds the accumulated power model and shows the impact of the constellation design on the power fluctuation in a WDM system. Based on our findings, we proposed our CPCG technique to achieve better power properties in Section 2.4. Simulation results are discussed in Section 2.5. Finally, Section 2.6 concludes this chapter.

2.2 System Model

As for any electromagnetic wave in a dielectric medium, propagation of light pulses inside the fiber is governed by Maxwell's equations. By considering the fiber's physical properties, we can simplify Maxwell's equations for the slowly varying part of a modulated signal as explained in Section 1.4.1. The result is called the nonlinear Schrödinger equation which can be found in (1.16) and repeated here for convenience as:

$$\frac{\delta A(t,z)}{\delta z} + j\frac{\beta_2}{2}\frac{\delta^2 A(t,z)}{\delta t^2} + \frac{\alpha}{2}A(t,z) = j\gamma|A(t,z)|^2 A(t,z), \tag{2.1}$$

where A is the signal envelope propagating in z direction, β_2 represents group velocity dispersion, α is the attenuation parameter, and γ captures the nonlinearity [88].

In this chapter, we focus on the nonlinear effects right after launching the signal into the fiber. While dispersion can spread the signal energy and reduce nonlinear effects after a few spans, nonlinearity can cause severe issues at the beginning of the fiber. Therefore, in this chapter, whenever we refer to A, we mean $A = A(t, 0)$.

Considering a WDM system, the available bandwidth is divided into frequency separated channels. Therefore, the signal envelope A in (2.1) can be expressed based on the WDM scheme as:

$$A = \sum_{i=1}^{N} A_i e^{j2\pi f_i t}, \tag{2.2}$$

where A_i is the i-th channel signal and f_i is the i-th channel central frequency. Also, N is the number of the WDM channels.

Substituting (2.2) into (2.1) yields the pulse propagation equation for the WDM channels, separately. One can derive the channel-wise NLSE by keeping SPM and XPM terms as:

$$\frac{\delta A_i}{\delta z} + \beta_{i1}\frac{\delta A_i}{\delta t} + j\frac{\beta_2}{2}\frac{\delta^2 A_i}{\delta t^2} + \frac{\alpha}{2}A_i = j\gamma(2\sum_{k=1}^{N}|A_k|^2 - |A_i|^2)A_i, \tag{2.3}$$

where β_{i1} is the first order dispersion coefficient for the i-th channel. This formulation is also called XPM based NLSE or coupled NLSE [89], [90]. If β_2 in (2.1) is measured for f_0 then β_{i1} can be expressed as [91]:

$$\beta_{i1} = (f_i - f_0)\beta_2. \tag{2.4}$$

As seen in (2.3), the nonlinear term for a channel depends not only on the launched power into the channel of interest, but also on the sum of the power in the other channels. Throughout this work, we refer to this term as accumulated power, $P(t)$, and define it as follows[1]:

$$P(t) = \sum_{k=1}^{N} |A_k(t)|^2. \tag{2.5}$$

---

[1] This equation assumes the orthogonality of the WDM channels holds true during the propagation inside the fiber. If the nonlinearity effect is too strong it may cause a frequency channel to penetrate into its neighbors. In this case, the orthogonality does not hold true.

In theory, we can compensate the fiber linear and nonlinear impairments by using (2.1) or (2.3) with the knowledge of all the received channels. However, in a typical WDM system, users have access only to their own channels. This forces us to treat the accumulated power term as a destructive unknown in the receiver.

In this work, we investigate the statistical behavior of the accumulated power term, P, and focus on reducing its impact on the system performance.

2.3 Accumulated Power and Its Impact

In this section first, we find the statistics of the accumulated power term. Then, we show the impact of the accumulated power term on the system performance. To this end, we start by modeling the pulse train launched into the channels and gradually consider the impact of different fiber impairments influencing a WDM system.

2.3.1 Accumulated Power Model

Assuming a pulse shape $g(t)$, the signal launched into channel i can be expressed as:

$$A_i(t) = \sum_k m_{ik} g(t - kT), \tag{2.6}$$

where k corresponds to time index and T is the symbol duration. Also, m_{ik} is the symbol sent in time index k and channel i. In a dense WDM system with a large number of channels, channels are not synchronous and hence their initial launch time is random. Besides, the channels experience different velocities according to their β_1 in (2.4). Therefore, at a given time, t_0, channel i contributes to the accumulated power term experiencing a random time delay, τ_i. It is enough to consider τ_i with a uniform distribution in the interval of $[0, T)$. This is because in our study, only the relative position within a symbol duration matters. Considering random time delay and the model in (2.6), accumulated power (2.5) at a given time, t_0, can be expressed as:

$$P(t_0) = \sum_{i=1}^{N} |A_i(t_0 - \tau_i)|^2 = \sum_{i=1}^{N} \left| \sum_k m_{ik} g(t_0 - kT - \tau_i) \right|^2 . \tag{2.7}$$

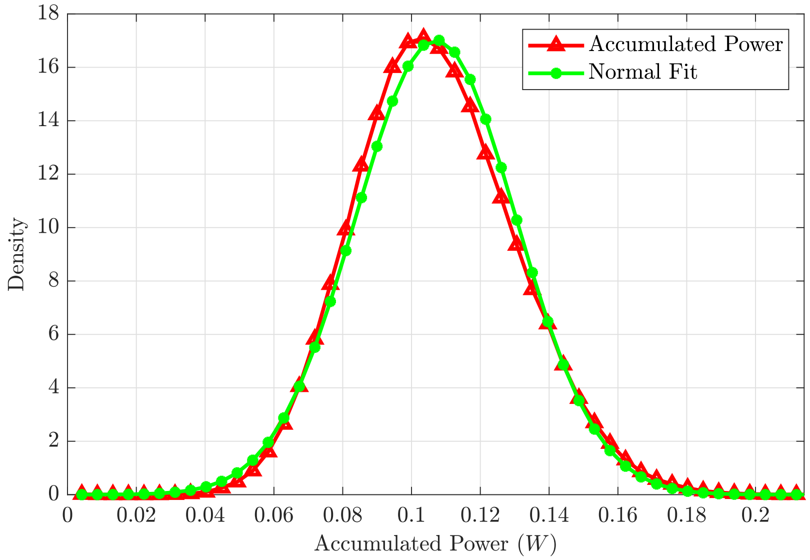


Figure 2.1: Accumulated power distribution for 11 channels carrying 16-QAM symbols on root raised cosine pulse shape. As seen, a Gaussian approximation is quite accurate for the accumulated power. Based on (2.8) and (2.9), we expect the mean and standard deviation to be 0.107 and 0.023, respectively. The simulated values based on a true WDM simulation are 0.107 and 0.023. It is worth mentioning that (2.8) and (2.9) use no approximation while the distribution is only fitted. This is why we observe a slight difference in the fitted curve while the mean and variance value predicted by (2.8) and (2.9) are accurate.

Accumulated power term consists of many random variables such as time delays τ_i and transmitted symbols m_{ik}. Based on the central limit theorem, the distribution for $P(t_0)$ can be approximated using a Gaussian distribution. Simulation results depicted in Figure 2.1 also support this claim. This figure demonstrates the distribution of $P(t_0)$ for a WDM system with 11 channels loaded with root raised cosine pulse shape and its approximated Gaussian distribution.

Assuming a Gaussian distribution, the mean and variance are sufficient to describe accumulated power behavior at a given time. In Appendix A.1, we show that the mean value, μ, of (2.7) can be found as:

$$\mu = N \sum_{k=-q}^{q} \int_0^T \frac{1}{T} g^2(t - kT) dt, \tag{2.8}$$

where we assume the pulse shape, $g(t)$, is non-zero only in $[-qT, qT]$ and the average power of the constellation is one, $E\{|m|^2\} = 1$. As seen in (2.8), the mean value of the accumulated power depends on the pulse shape while the constellation has no effect.

In Appendix A.2, we find the variance of the accumulated power, σ^2, as:

$$\begin{aligned}\sigma^2 = \zeta N \sum_{k=-q}^{q} \int_0^T \frac{1}{T} g^4(t-kT)dt \\ + 4N \sum_{k_0=-q}^{q} \sum_{k_1=k_0+1}^{q} \int_0^T \frac{1}{T} g^2(t-k_0T) g^2(t-k_1T)dt - \frac{\mu^2}{N},\end{aligned} \tag{2.9}$$

where

$$\zeta = E\{\frac{1}{N}\sum_{i=1}^{N}|m_i|^4\}. \tag{2.10}$$

As seen, the variance of the accumulated power depends on the fourth moment of the accumulated symbol power of the used constellations. In other words, we can reduce the fluctuations of the accumulated power by designing and using a group of constellations with smaller ζ. As an example, we evaluate (2.9) for a typical root raised cosine pulse shape (roll-off factor = 0.1, $T = 1.1 \times 10^{-10}$ sec, $P_0 = 0.01$ W) as:

$$Var\{P(t_0)\} = N(6.63\zeta - 3.76) \times 10^{-5}. \tag{2.11}$$

This equation shows the significant impact of ζ on power fluctuation.

To better understand ζ, let us find its value for typical constellations. The minimum value of ζ is 1 for any constellations with constant energy such as PSKs. In case of 2-D M-QAMs, we can substitute the constellation symbols, described as:

$$m_{p,l} = \frac{(2p-k-1)+(2l-k-1)j}{\frac{1}{k}\sqrt{\sum_{p'=1}^{k}\sum_{l'=1}^{k}(2p'-k-1)^2+(2l'-k-1)^2}}, \tag{2.12}$$

into (2.10) and express ζ as:

$$\zeta = \frac{k^2\sum_{p=1}^{k}\sum_{l=1}^{k}\left((2p-k-1)^2+(2l-k-1)^2\right)^2}{\left(\sum_{p=1}^{k}\sum_{l=1}^{k}(2p-k-1)^2+(2l-k-1)^2\right)^2}, \tag{2.13}$$

where $k^2 = M$ and $p, l \in 1, 2, \cdots, q$. For large M, ζ tends to 1.4. Table 2.1 shows ζ's value for commonly used QAM constellations.

	$M=4$	$M=16$	$M=64$	$M=256$	$M=1024$
ζ	1	1.32	1.381	1.3953	1.3988

Table 2.1: ζ value for M-QAM constellation

Accumulated power is a function of time; therefore, to complete the statistical model, we should consider the accumulated power as a random process and find the auto-correlation function. As shown in Appendix A.3, the auto-correlation of this random process can be expressed as:

$$\begin{aligned} R(l) = {} & \zeta N \sum_{k=-q}^{q} \int_0^T \frac{1}{T} g^2(t-kT) g^2(t-l-kT) dt \\ & + N \sum_{k_0=-q}^{q} \sum_{k_1=k_0+1}^{q} \int_0^T \frac{1}{T} \Big(g(t-k_0T) g(t-l-k_1T) \\ & + g(t-k_1T) g(t-l-k_0T) \Big)^2 dt - \frac{\mu^2}{N}. \end{aligned} \tag{2.14}$$

As seen, the auto-correlation function depends on both ζ and the pulse shape, $g(t)$. For a typical pulse shape the concentration of the power is on the main lobe and hence the auto-correlation function is approximately non-zero only for $|l| < 2T$. Figure 2.2 also supports this argument. Symbol duration, T, is negligible in comparison with pulse propagation time. Therefore, we can approximate the accumulated power term by a white Gaussian process.

2.3.2 Impact of the Accumulated Power on Nonlinearity

In the last subsection, we showed that the accumulated power term can be approximated by a white Gaussian process and derived the statistics with respect to the constellation design parameter, ζ. Here, we concentrate on the impact of the accumulated power term on the system performance. For the sake of simplicity, let us assume a nonlinear fiber with the following propagation equation by omitting the dispersion coefficients from (2.3) resulting in:

$$\frac{\delta A_i}{\delta z} = j\gamma \Big(2P(t) - |A_i|^2 \Big) A_i. \tag{2.15}$$

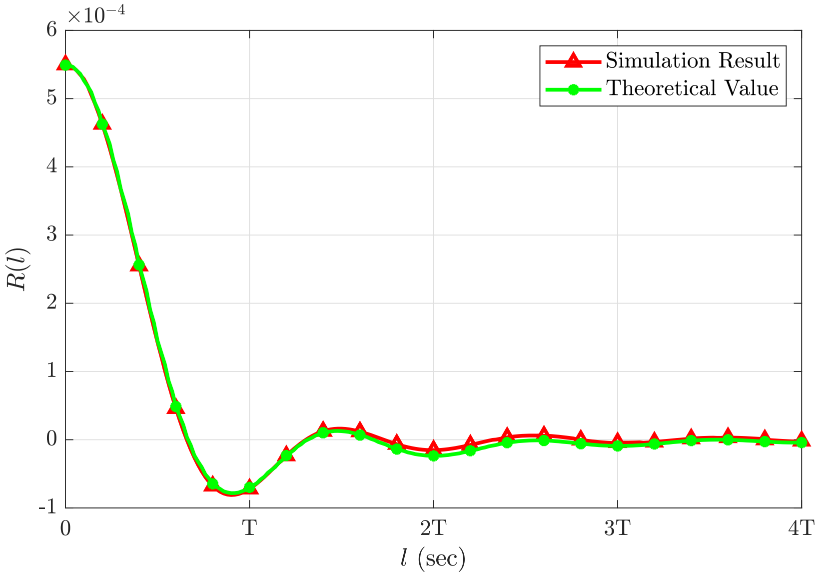


Figure 2.2: Auto-correlation function of the accumulated power term for 11 WDM channels. As seen, auto-correlation function is almost zero for $|l|> 2T$.

In the following discussions and simulations we will continue to consider dispersion effects again.

Based on our discussion in Subsection 2.3.1, accumulated power $P(t)$ can be modeled by a white Gaussian process with mean and variance defined in (2.8) and (2.9), respectively. Let us represent this model using a standard (zero-mean, unit variance) Gaussian random variable, N_n, as:

$$P(t) = \mu + \sigma N_n, \tag{2.16}$$

where μ and σ^2 are defined in (2.8) and (2.9), respectively. Now, we can solve (2.15) analytically as:

$$A_i(t) = A_i(0)e^{j\phi_i}, \tag{2.17}$$

where

$$\phi_i = \gamma z\mu - \gamma|A_i|^2 + \gamma\phi_l + \gamma\sigma\int_0^z N_n dz. \tag{2.18}$$

The term $\gamma\sigma\int_0^z N_n dz$ represents the nonlinear phase noise (NLPN) and can be considered as a Wiener process since N_n is a standard Gaussian random variable [92]. Furthermore, ϕ_l models the phase noise due to the transmitter and amplifiers noise. Following [93], we also consider ϕ_l as a Wiener process. The terms $z\mu$ and $\gamma|A_i|^2$ are constant or known to the receiver. They cause the constellation to rotate. We assume the receiver can compensate for this rotation and hence neglect these terms. Therefore, the power of the remaining phase noise can be expressed as:

$$P_{\phi_i} = \gamma^2(\sigma^2 + \sigma_l^2)z, \tag{2.19}$$

where σ_l^2 represents the power of ϕ_l. Here, we argue that based on (2.9), σ^2 is a linear function of ζ. Therefore, by reducing ζ, we can decrease σ, and hence z can increase while maintaining the same phase noise power. In other words, for the same constellation, we can achieve the same noise power at longer system reach when we reduce ζ using CPCG. For example let us consider two cases where we reduce σ_1^2 to σ_2^2 by reducing ζ by a factor of 1.4 using CPCG. Based on the numerical values in (2.11) we can say:

$$\sigma_2^2 = 0.52\sigma_1^2. \tag{2.20}$$

Also, we assume the linear noise power is twice the nonlinear noise in the optimal launched power [71]:

$$\sigma_l^2 = 2\sigma_1^2. \tag{2.21}$$

For the same phase noise power, we can equate (2.19) for both σ_1^2 and σ_2^2 and find the system reach ratio for these two cases as:

$$\frac{z_2}{z_1} = \frac{\sigma_1^2 + \sigma_l^2}{\sigma_2^2 + \sigma_l^2}. \tag{2.22}$$

Substituting (2.20) and (2.21), the system reach ratio in (2.22) is 1.19 which means CPCG generates the same noise power at a 19% longer distance when we reduce ζ by a factor of 1.4. With the same argument, CPCG based on 16-QAM generates the same noise power at a 16% longer distance in comparison with 16-QAM.

From another point of view, we can reduce nonlinear phase noise by using constant power constellations and hence increase the system's ADR at the same distance.

2.4 Constant Power Constellation Grouping

In the last section, we investigated the effects of the constellation design on power fluctuation, system reach, and fiber nonlinearity impairment. In this section, we answer the question of how to achieve the best ζ from an existing constellation.

PSKs are one of the well-known classes of the constant power constellations. PSKs have $\zeta = 1$ and hence take full advantage of nonlinearity impairment reduction due to lower power fluctuation. However, two limiting factors make PSK not a good choice for today's fiber systems: (i) PSKs in low-dimension spaces offer a low ADR. (ii) higher-dimensional PSKs need high computation for creation and detection. In this section, we introduce CPCG as a technique to achieve constant power constellations with better properties than PSK.

CPCG allows a set of channels to send only combinations of symbols with a predefined constant accumulated power. Let us assume we have N constellations as $X_1, \cdots, X_N$ and call them base constellations. For the sake of simplicity, in this work, we assume all of the base constellations are the same. We transmit the symbols of these constellations, $(x_1, \cdots, x_N)$, over N parallel channels, where x_i represents the symbols of X_i constellation. With CPCG, we are only allowed to transmit combinations of symbols with constant accumulated power. Let us form a set containing all combinations of symbols with constant accumulated power and call it constant power constellation set, S_c. In mathematical form, the constant power constellation set is as follows:

$$S_c = \left\{ (x_1, x_2, ..., x_N) \middle| x_i \in X_i \text{ for } i = 1 \cdots, N, \sum_{i=1}^{N} |x_i|^2 = N \right\}. \quad (2.23)$$

This definition might resemble sphere coding, but we should remember that sphere codes are defined as a set of bits known as codewords, whereas our definition is based on constellation symbols. While the symbol energy directly impacts the behavior of the launched power, the impact of the codewords on the launched power is not well defined.

Transmitting symbols combinations in a constant power constellation set on N parallel channels causes the accumulated power of the symbols to remain

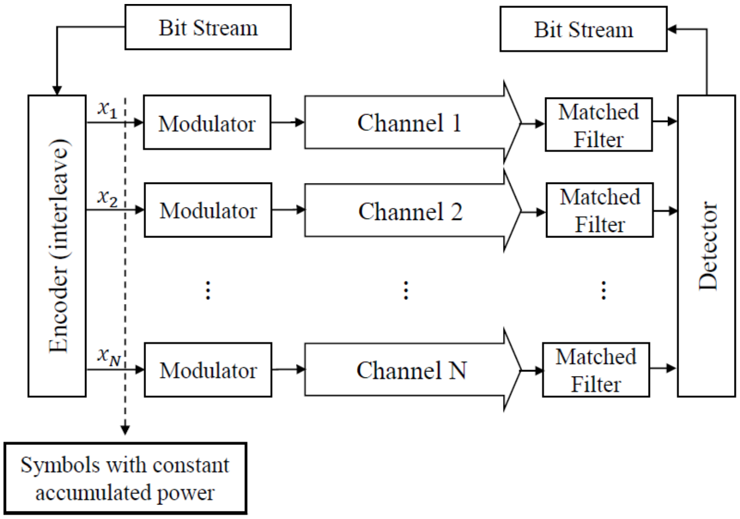


Figure 2.3: Block diagram of a system suitable for proposed CPCG. Based on the input bitstream, encoder chooses one of the combinations, $(x_1, x_2, \cdots, x_N)$, in the constant power constellation set (2.23) and transmits them on N parallel channels. In this scheme the accumulated power launched into the system is constant.

a predefined value, N. To implement these parallel channels, we propose sending data by interleaving N symbols and sending them in parallel as depicted in Figure 2.3. These parallel channels can be implemented using polarization and/or frequency orthogonality. Within a set of parallel channels any channel selection for CPCG is allowed since the accumulated launched power into the main system remains constant.

CPCG introduces constant accumulated power as a constraint and hence reduces the source's entropy. We mathematically represent the reduction in the information content of the source caused by CPCG by defining a grouping efficiency for N joint constellations as:

$$\theta(N) = \frac{H\left(X_1, X_2, \cdots, X_N \,\middle|\, \sum_{i=1}^{N} |X_i|^2 = N\right)}{\sum_{i=1}^{N} H(X_i)}, \tag{2.24}$$

where $H(X)$ is the entropy of the random variable X.

In the following subsections first, we study the entropy of a CPCG system and how to maximize the efficiency. Later, we introduce the 16-QAM constellation as a candidate to implement the CPCG technique in the real world.

2.4.1 Grouping Performance

In this subsection, we investigate the entropy of a CPCG system and find the ultimate ADR in such a system. As a starting point, let us define entropy per channel for N grouped constellations as:

$$H_p(N) = \frac{H\left(X_1, X_2, \cdots, X_N \middle| \sum_{i=1}^{N} |X_i|^2 = N\right)}{N}. \tag{2.25}$$

Based on the chain rule, we can write (2.25) as:

$$H_p(N) = \sum_{i=1}^{N} \frac{H\left(X_i \middle| X_1, \cdots, X_{i-1}, \sum_{i=1}^{N} |X_i|^2 = N\right)}{N}. \tag{2.26}$$

One can simply show this sum is bounded as:

$$0 \leq H_p(N) \leq \max\left\{ H(X_1), \cdots, H(X_N) \right\}. \tag{2.27}$$

Furthermore, the series $H_p(N)$ is a monotonic function of N because:

$$\begin{aligned} &H\left(X_i \middle| X_1, \cdots, X_{i-1}, \sum_{i\,=1}^{N} |X_i|^2 = N\right) \geq \\ &H\left(X_{i+1} \middle| X_1, \cdots, X_i, \sum_{i=1}^{N} |X_i|^2 = N\right). \end{aligned} \tag{2.28}$$

Therefore, the sum in (2.26) is convergent and we can denote the convergence value as η for large N. In mathematical form, we can write, $\lim_{N\to\infty} H_P(N) = \eta$. This allows us to conclude the following for large enough N:

$$H\left(X_1, X_2, \cdots, X_N \middle| \sum_{i=1}^{N} |X_i|^2 = N\right) = N\eta. \tag{2.29}$$

Also creating any other power level which is realizable with countable constellation combinations would result in $H_P(N) = \eta$ for large N. The proof is provided in Appendix A.4. Therefore, we can write:

$$H\left(X_2, \cdots, X_N \middle| \sum_{i=2}^{N} |X_i|^2 = N - |X_1|^2\right) = (N-1)\eta. \tag{2.30}$$

We can find η by applying the chain rule to (2.29) and substituting (2.30) as:

$$\begin{aligned}N\eta &= H\Bigg(X_1, X_2, \cdots, X_N\Bigg|\sum_{i=1}^{N}|X_i|^2{=}N\Bigg)\\ &= H\Bigg(X_1\Bigg|\sum_{i=1}^{N}|X_i|^2{=}N\Bigg) + H\Bigg(X_2, \cdots, X_N\Bigg|\sum_{i=2}^{N}|X_i|^2{=}N-|X_1|^2\Bigg)\\ &= H\Bigg(X_1\Bigg|\sum_{i=1}^{N}|X_i|^2{=}N\Bigg) + (N-1)\eta.\end{aligned} \tag{2.31}$$

Therefore:

$$\eta = H\Bigg(X_1\Bigg|\sum_{i=1}^{N}|X_i|^2{=}N\Bigg). \tag{2.32}$$

The constant power constraint in (2.32) can reduce the entropy of X_1. In other words, $\eta \leq H(X_1)$. However, for a well-chosen base constellation η approaches $H(X_1)$ as N increases because the condition is automatically satisfied. As an example, we will find the probability of the symbols for a 16-QAM based CPCG in the next section. As N grows, the symbols become equiprobable which means constraint has no effect and by relaxing it we can write $\eta = H(X_1)$.

Now, we can conclude from (2.32) that ultimately for a large enough N, the reduction in the entropy due to the grouping vanishes and each channel transmits data at its maximum possible entropy. In other words grouping efficiency tends to one for a large N or $\lim_{N\to\infty}\theta(N) = 1$, where θ is defined in (2.24). In short, for better performance of a CPCG design, we need to increase the number of the channels.

2.4.2 16-QAM Based CPCG

Until now, we supposed there is a base constellation that satisfies the constant accumulated power condition. Here, we work on 16-QAM as one of the candidates for the base constellation. This constellation has been already employed in long-reach fiber systems. Symbols in a 16-QAM constellation can be categorized into three groups based on their power. These groups are shown in Figure 2.4 as A, B and C. An interesting property of these groups is that the accumulated power of any pairs of (a, c), (c, a) or (b, b) is constant where

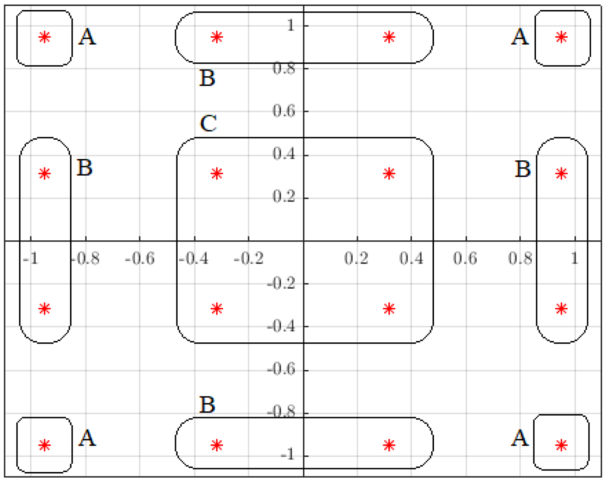


Figure 2.4: 16-QAM constellation with three groups of symbols with the same power level A, B, C. This groups satisfy the constant accumulated power as $|a|^2+|c|^2= |b|^2+|b|^2= 2$, where $a \in A, b \in B, c \in C$.

$a \in A$, $b \in B$ and $c \in C$. Therefore, to implement a CPCG of two 16-QAMs ($N = 2$), one can freely choose a symbol from the first 16-QAM to be sent on one channel and select one of the groups on the second channel such that the accumulated power stays constant.

For a pair of two 16-QAMs, the constant power condition allows us to transmit a total of 96 pairs of symbols. If we consider all of these 96 combinations of symbols with equal probability, we can transmit up to $\frac{\log_2 96}{2} = 3.29$ bits/symbol/channel in this scheme. The total number of available symbols in groups of N 16-QAM constellation with constant accumulated power is given by:

$$|S_{\text{16-QAM}}|= \begin{cases} \sum_{i=0}^{N/2} \binom{N}{2i} \binom{N-2i}{(N-2i)/2} 2^{2i+2N} & , N \in \text{even} \\ \sum_{i=0}^{(N-1)/2} \binom{N}{2i+1} \binom{N-2i-1}{(N-2i-1)/2} 2^{2i+2N+1} & , N \in \text{odd}, \end{cases} \tag{2.33}$$

where, $|\cdot|$ represents the set cardinality. Then for a 16-QAM based CPCG, grouping efficiency would be:

$$\theta_{\text{16-QAM}} = \frac{\log_2(|S_{\text{16-QAM}}|)}{4N}. \tag{2.34}$$

Based on our argument in Subsection 2.4.1, we expect (2.33) to have an asymptote of 2^{4N}.

Furthermore, we can find the symbol probabilities in a 16-QAM constant power constellation set for the symbols from the sets A, B, C in Figure 2.4 as

($N \in$ even):

$$Pb(N) = \frac{\sum_{i=0}^{N/2} \binom{N}{2i} \binom{N-2i}{(N-2i)/2} 2^{2i+2N+1} i}{8N \sum_{i=0}^{N/2} \binom{N}{2i} \binom{N-2i}{(N-2i)/2} 2^{2i+2N}}, \tag{2.35}$$

and

$$Pa(N) = Pc(N) = \frac{1}{8} - Pb(N). \tag{2.36}$$

Numerical results show that for large N we have:

$$\lim_{N\to\infty} Pa(N) = \lim_{N\to\infty} Pb(N) = \lim_{N\to\infty} Pc(N) = 1/16. \tag{2.37}$$

Therefore, in a constant power constellation set, symbols probabilities are equal for large N. We used this result in interpreting (2.32).

The main reason for the uniform distribution of the symbols in this setup is to choose a constant power for the symbols that is equal to their average power. However, choosing a lower value for the constant power will alter the distribution of the symbols, and one might even achieve shaping gains by emphasizing symbols with lower energy.

Figure 2.5 demonstrates the grouping efficiency for 16-QAM based CPCG. Based on this figure, as the number of constellations increases, grouping efficiency tends to one which means the entropy of each CPCG becomes equal to a 16-QAM constellation.

Another interesting aspect of the CPCG is the channel capacity for a paired constellation. Accepting the Gaussian noise model for the fiber impairments [71], [72], we find the capacity of the CPCG for an additive white Gaussian noise channel. In general, the required calculation for more than 4 grouped channels is intractable. We use Ungerboeck method [94] to find the capacity with low complexity based on a Monte Carlo approach. Based on this method, the capacity for the constant power constellation set, S_c (2.23), can be calculated as:

$$C_{S_c} = \log_2(|S_c|_0) - \frac{1}{|S_c|_0} \sum_{k=0}^{|S_c|_0-1} E\Big\{\log_2 \sum_{i=0}^{|S_c|_0-1} \exp(-\frac{|a_k+n-a_i|^2-|n|^2}{2\sigma^2})\Big\}, \tag{2.38}$$

where a_i is the i-th combination in S_c and n is an N-dimensional Gaussian noise with variance of σ_n^2. To further reduce the complexity, we calculate the

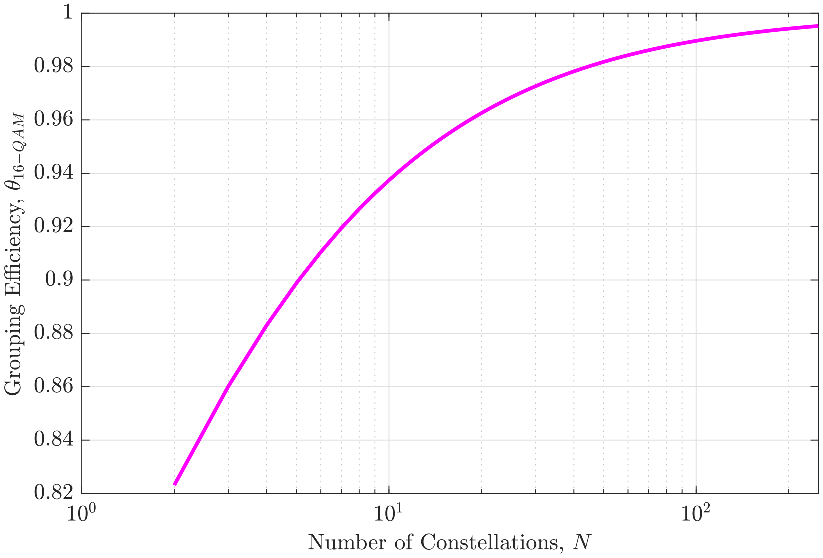


Figure 2.5: Grouping efficiency for 16-QAM. As seen, by increasing the number of the constellations, entropy loss due to grouping is compensated. Increasing the number of constellations from 2 to 6, reduces the grouping loss to less than half.

capacity of the sign bit and symbol amplitude separately. Figure 2.6 represents the normalized capacity of the channels, $\frac{I_{S_c}}{N}$, for constant power constellation set. We use this figure to find the capacity of CPCG in the fiber setting in Section 2.5 by calculating the equivalent SNR. As seen, CPCG offers lower entropy in comparison with 16-QAM but approaches the 16-QAM capacity as N increases. As for 10 and 14-PSK constellations, please note that both CPCG and PSK constellations produce $\zeta = 1$. Hence, when comparing with PSKs, for fairness, we picked PSKs with the same entropy with CPCG $N = 2$ and $N = 12$ which are PSKs with 10 points and 14 points respectively. We called these equivalent PSKs. As seen, CPCGs provide a higher capacity in comparison with their equivalent PSKs.

2.5 Simulations

In this section, we first describe the system configuration we use to validate our proposed signaling scheme. Then, we present and discuss the simulation results.

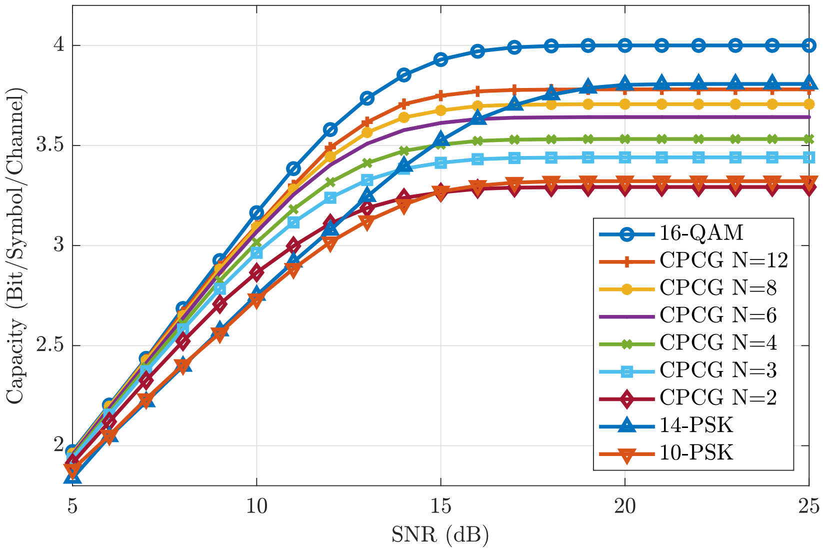


Figure 2.6: Capacity for CPCGs. As seen, the CPCG technique can approach QAM constellations by increasing the dimension. The CPCG technique also offers higher capacity in comparison with the equivalent PSK constellation in lower SNR.

2.5.1 Simulation configuration

The optical system we consider in this study works as follows: on the transmitter side, we first generate a constant power constellation set and for each channel use select one of the possible constant-power combinations randomly. This symbol combination is modulated on the raised cosine pulse with a roll-off factor of 0.1 and time duration of $T_0 = 1.1^{-10}$ sec. Then these pulse trains are launched into a digital subcarrier multiplexing system with 8 channels of 11.95 GHz in both polarization with a baud rate of 144 GBaud as described in [95]. Finally, the resulting signal is launched into a dual-polarization fiber. Table 2.2 demonstrates the specifics of this fiber. Fiber attenuation is compensated using in-line EDFA every 100 km. EDFA's noise figure is 5 dB. This

α	β_1	β_2	γ
$0.2\ dB/km$	$0\ s/m$	$-20 \times 10^{-27} s^2/m$	$0.013\ (mW)^{-1}$

Table 2.2: Fiber parameters for the central channel of a hypothetical terrestrial link

fiber system is simulated based on (2.1) using SSFM [29] to solve the Manakov equation [91]. The SSFM step size for the fiber simulator is set to 500 meters.

On the receiver side, dispersion is compensated first and then frequency channels are extracted. Next, constellation rotation due to the fiber nonlinearity is calculated and canceled. Afterward, we calculate the received symbol points using matched filters. Our results are based on Monte Carlo evaluation method with 2^{16} trials. Results are reported for the $\hat{x}$ polarization of the fifth frequency channel.

2.5.2 Simulation Results

In this section, we present the simulation results for maximum achievable SNR, system reach, and ADR in a CPCG design.

Maximum Achievable SNR

Nonlinearity in fiber causes abnormal behavior of the achievable SNR. Kerr nonlinearity is proportional to the launched power; therefore, we cannot increase SNR unlimitedly by launching more power into the fiber. In this simulation, we find the achievable SNR for the constellations with different ζs over a 6000 km fiber. Results are shown in Figure 2.7. Here, we emphasize that $\zeta = 1$ corresponds to PSK and CPCG. As seen, signaling schemes with smaller ζ achieve higher SNRs and shift the maximum achievable SNR to the higher launched power. This can be interpreted to better immunity against nonlinearity which was predicted in (2.19). Higher SNR allows us to achieve higher ADR. On the other hand, higher launched power empowers us to increase span length between EDFAs for the same minimum acceptable signal power before amplification. This means the total number of EDFAs in the system can be reduced and hence the system experiences lower ASE noise. We can show that for an increase of λ dB in the launched power, span length can be increased by $\frac{\lambda}{\alpha}$ where α is the fiber attenuation parameter. Using a signaling technique with $\zeta = 1$ instead of 16-QAM shifts the maximum SNR by 0.5 dB to a higher power which means a 2.5 km increase in span length for the assumed fiber.

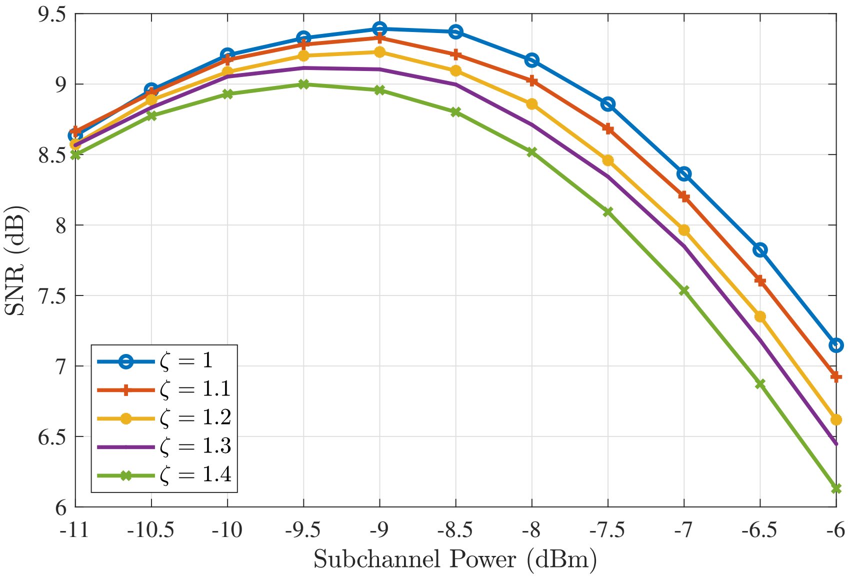


Figure 2.7: Using signaling techniques with smaller ζ results in i. achieving higher maximum SNR and ii. shifting the maximum SNR point to higher launched powers.

System reach

Our findings in (2.19) predict a linear relationship between noise power and system reach when the nonlinearity effect is dominant. To investigate reach-noise relation and compare the impact of the traveled distance on constellations with different ζs, we plot the noise power versus system reach in Figure 2.8. In this simulation, we use the optimal launched power of the base constellation, 16-QAM, at each distance for all the methods. It is worth mentioning that all the CPCG family with different N produces the same $\zeta = 1$ and hence the same noise power. For the comparison, we also plot the noise power for CCDM [83] (block length of 128 and energy of 110) and kurtosis-limited sphere modulation (K-ESS, block length of 128 and maximum kurtosis of 4360) [96]. For a fair comparison, CCDM and K-ESS methods are chosen to provide a source entropy very close to CPCG[2]. These results confirm the linear relation between noise power and system reach. Also, the CPCG technique offers lower noise power as expected based on (2.19) because it results in lower ζ in comparison with CCDM and kurtosis-limited sphere modulation.

---

[2] CPCG $N = 12$ provides a source entropy of 3.78 bits/symbol while CCDM and K-ESS have 3.82 bits/symbol.

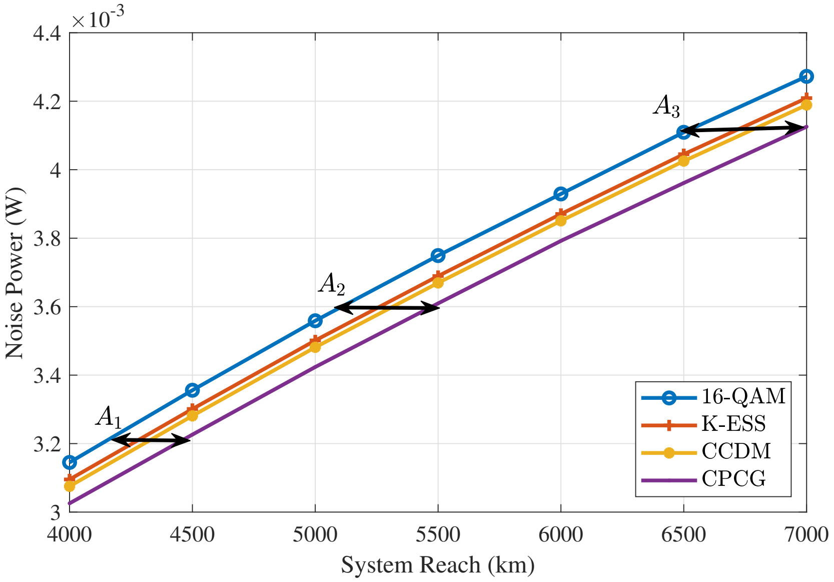


Figure 2.8: Noise power at different distances. This figure confirms the linear relation between noise power and system reach. In this simulation, the CPCG technique experiences the same nonlinear noise at 14% longer distances in comparison with its base constellation, 16-QAM. For example, CPCG at 4000 km causes a noise power the same as 16-QAM at 3500 km (as indicated by arrow A_1). Arrows A_2 and A_3 also show increases from 4830 km and 6250 km to 5500 km and 7000 km, respectively.

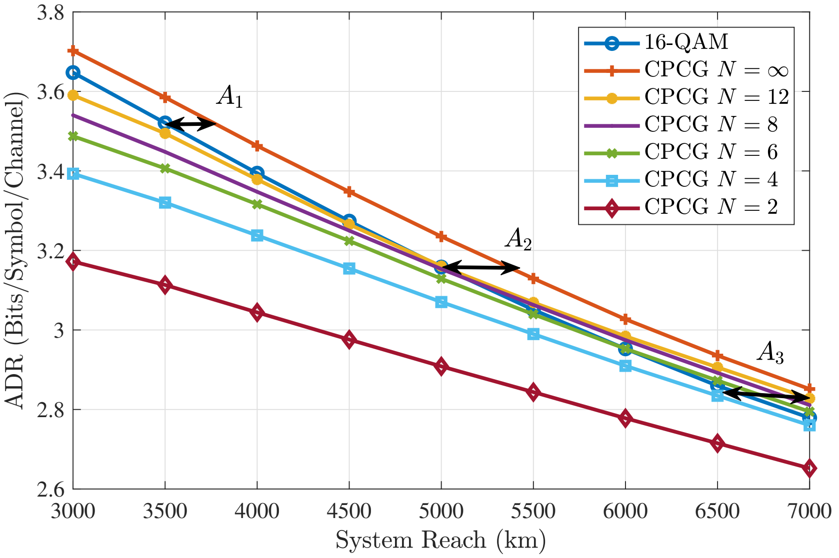


Figure 2.9: ADR at different distances. The CPCG technique offers higher ADR at longer distances in comparison with its base constellation, 16-QAM. For $N = \infty$, CPCG technique increases the system reach by 14% as indicated by A_1 and A_2 where CPCG increase the system reach from 3280 km and 4620 km to 3750 km and 5250 km, respectively for the same ADR as 16-QAM. For $N = 12$, CPCG starts beating 16-QAM at 4200 km and increases the system reach by a factor of 10% at 6000 km as indicated by arrow A_3.

The CPCG technique experiences a noise power the same as 16-QAM at 1.14 longer distances. Based on Section 2.3.2, we expect this increase to be by a factor 1.16. We explain the difference by putting emphasis on the neglected dispersion effects in the theoretical discussion in Section 2.3.2. To gain a better insight into system reach, we find the ADR we get at different distances. Figure 2.9 represents the results for the different numbers of the constellations. Here, we can conclude that 16-QAM ADR drops rapidly in comparison with CPCG. This causes CPCG to offer higher ADR at longer distances despite their lower entropy. As seen, when we completely compensate for the grouping entropy loss by using large enough group, $N = \infty$, system reach increases by 14% in comparison with 16-QAM for the same ADR. When the fiber length is more than 4200 km, CPCG $N = 12$ outperforms 16-QAM, meaning that it achieves the same ADR as 16-QAM at longer distances. For instance, CPCG $N = 12$ at 6600 km has the same ADR as 16-QAM at 6000 km. This can

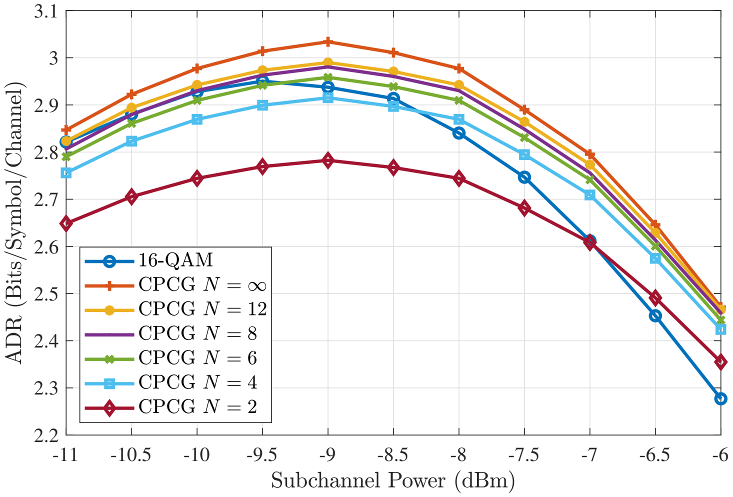


Figure 2.10: ADR for different signalling schemes. CPCG can achieve a higher data rate in comparison with 16-QAM for a large number of channels.

be interpreted as a 10% increase in the system reach. In addition, lower-dimensional CPCG offers higher ADR in comparison with 16-QAM at longer distances.

Achievable Data Rate

As explained in Section 2.4, the entropy reduction due to grouping can be mitigated by increasing the number of the constellations. Figure 2.10 demonstrates the ADR for a fiber length of 6000 km. As seen, for the case of two paired channels, we have a reduction in ADR in comparison with 16-QAM. At this reach, 6-D CPCG technique beats 16-QAM. On the other hand, the ultimate aADR (CPCG $N = \infty$) suggests that by compensating the entropy loss of the CPCG scheme, we can achieve a higher data rate using CPCG in a fiber system in comparison with 16-QAM.

2.6 Conclusion

In this chapter, we concentrated on increasing system tolerance against fiber nonlinearity by reducing the signal power fluctuation. We mathematically tracked the impact of the power fluctuation and showed the contribution of the constellation design on the system nonlinearity. Then we introduced a constant power constellation grouping technique that achieves the minimum possible variance of the launched power. This scheme outperforms its constant-power counterparts in both complexity and efficiency. We showed that this signaling method can achieve the QAM efficiency as the number of the channels increases. We also theoretically showed that using our proposed signaling, one can achieve the same noise power at longer distances using CPCG in comparison with the QAM constellation. We also showed that for a given distance, the proposed scheme can offer a higher ADR than its QAM counterparts. This can be attributed to the higher SNR that constant launched power can provide in the fiber systems.

Chapter 3

Minimizing Fiber's Nonlinear Interference Noise by Designing Launched Signal PSD

3.1 Introduction

According to the GN model, NLIN is dependent on the PSD of the signal launched into the WDM system [72]. In a general communication system, two primary factors affect the signal PSD: (i) the second moment of the constellation from which the transmitted symbols are drawn and (ii) the modulating pulse used to convert these symbols into continuous-time signals. Therefore, the design of the modulating pulse PSD can effectively reduce NLIN in fiber systems. While numerous constellations have been proposed to mitigate fiber nonlinearity [97]–[100], less attention has been given to the design of pulses for NLIN reduction.

In this study, our focus lies in designing the pulse PSD to minimize NLIN. A common concern regarding pulse design for fiber communication arises from the alteration of the pulse's temporal shape by dispersive fiber shortly after launch, potentially rendering pulse design ineffective. However, it is crucial to differentiate between the pulse's temporal shape and its PSD. While the pulse's temporal behavior undergoes drastic changes as it travels through the fiber, its PSD, the primary factor influencing NLIN, remains nearly unchanged throughout the entire channel. Therefore, designing the pulse PSD can effectively minimize NLIN over the entire fiber channel.

In our approach, we first utilize the calculus of variations to determine the band-limited pulse PSD that minimizes NLIN based on the GN model. We demonstrate that the optimal band-limited pulse, aimed at reducing nonlinearity, features a flat PSD spanning the entire available bandwidth. However, despite the effectiveness of this pulse in minimizing NLIN, it is not a practical pulse (e.g., it introduces ISI). Therefore, our next step involves imposing more constraints on the pulse design.

For practical considerations, in addition to (i) band-limitation, we introduce two constraints on the pulse shape [17]: (ii) the pulse should not introduce ISI, and (iii) it should exhibit rapid decay over time. We develop a comprehensive model for the pulse that meets these criteria and evaluate the NLIN power using the GN model. We formulate an optimization problem aimed at identifying the pulse PSD that minimizes NLIN. Our analysis reveals that for very small and very large excess bandwidths, the optimal pulse for minimizing NLIN remains the flat PSD pulse, even while satisfying the practical constraints. However, for the general case with arbitrary excess bandwidth, we express the problem as an optimization problem and numerically solve it.

Our formulation and pulse optimization are performed in the frequency domain. To present a new angle, we study our pulse in the time-domain too. This study provides a deeper understanding of why this pulse has its desirable properties. In particular, we show that our pulse better spreads the signal energy over time.

To investigate the performance of our proposed pulses, we conduct a simulation study based on SSFM [29]. Our simulation focuses on system tolerance against nonlinearity and compares our pulses with some other existing pulses. In all cases, our pulse generates the lowest NLIN in the fiber system. In short, our results confirm that compared to other pulses, our pulse increases (i) the maximum achievable SNR, (ii) the optimal launched power, and (iii) the system reach.

The remainder of this chapter is organized as follows: Section 3.2 provides the necessary background. In Section 3.3, we delve into the requirements for the pulse shape, present the general pulse model, and outline the optimization process to determine the optimal pulse shape for minimizing NLIN. Section 3.4 offers insights into the nonlinearity gains that our pulse achieves based on a time domain approach. Simulation results are presented and discussed in Section 3.5, followed by the conclusion in Section 3.6.

3.2 Background and System Model

3.2.1 Related Works

While there has been less emphasis on pulse design compared to constellation design for addressing fiber nonlinearity, some recent studies have suggested tailored pulses specifically for fiber communication. In the following, we review some of these studies.

In [101], the authors test the performances of different pulse shapes and suggest using the Gaussian monocycle and Scholz's monocycle to achieve a better bit error rate performance. Also, in [102], the authors compare the nonlinear performance of the non-return to zero and return to zero pulse shapes with different duty cycles. Their study interestingly reveals a trade-off between spectral efficiency and nonlinearity tolerance. In addition, an analytical approach to design airy pulses for fiber communication is proposed in [103].

Based on the GN model, the nonlinear noise has small components at higher frequencies [71]. Using this fact, an M-shape pulse is proposed in [104]–[106] that provides higher SNR than raised cosine (RC) by putting emphasis on high frequencies and interfering less with the nonlinear noise.

The polynomial pulse is also claimed to have an improved nonlinearity performance because of its fast decay rate and low peak-to-average ratio[107], [108]. In all these studies, for practical reasons, the pulses are chosen to be ISI-free and band-limited.

While these studies confirm the impact of pulse shapes on fiber nonlinearity, they lack an analytical approach to identify the optimal pulse for this purpose, leaving the problem as an ongoing challenge.

3.2.2 System Model

In this study, we consider a fiber system based on the SSMF with β_2, α, and γ as group velocity dispersion, attenuation, and nonliterary parameters, respectively [26]. The propagation of light pulses in this channel is governed by the nonlinear Schrödinger equation expresed in (1.16) and repeated here as:

$$\frac{\delta a}{\delta z} + j\frac{\beta_2}{2}\frac{\delta^2 a}{\delta t^2} + \frac{\alpha}{2}a = j\gamma|a|^2 a, \tag{3.1}$$

where $a(t, z)$ represents the signal traveling in z direction at time t.

In this system, fiber attenuation is compensated for by EDFAs [109]. We do not use inline dispersion compensation components, as is the case in modern systems, to enhance nonlinearity tolerance.

Considering the pulse shape as $h(t)$ and the transmitted symbols of s_i where i represents the symbol index in time, we can express the transmitted signal a, over the central channel of the WDM system as:

$$a(t, 0) = \sqrt{P_g}\sum_i s_i h(t - iT). \tag{3.2}$$

Here, we assume that $h(t)$ has unit energy and P_g represents the channel power. Given the assumption of independent and identically distributed (i.i.d.) transmitted symbols, we can express the channel PSD before launch as:

$$G(f) = P_g|H(f)|^2, \tag{3.3}$$

where $H(f)$ is the Fourier transform of the pulse shape, $h(t)$. Furthermore, considering other channels in the WDM system, the launched signal PSD is:

$$G_{\text{WDM}}(f) = P_g\sum_{i=-N/2}^{N/2} G(f - if_0), \tag{3.4}$$

where we assume there are $(N + 1)$ WDM channels with central frequency of if_0 for $\frac{-N}{2} \leq i \leq \frac{N}{2}$.

As we show in the next subsection, based on the GN model, the NLIN power is a function of G_{WDM}. According to (3.4), G_{WDM} is itself a function of the pulse PSD, $G(f)$. Therefore, our objective in this study is to design a pulse

PSD, $G(f)$, that minimizes the NLIN power. Furthermore, we design our pulse PSD to satisfy classical communication requirements such as band-limitation and zero-ISI.

Throughout this work, we denote the time domain functions and their Fourier transform using lowercase and uppercase letters, respectively.

3.2.3 Gaussian Noise Model

Because of nonlinearity, the propagation of the signal inside the fiber causes NLIN. GN models the NLIN PSD using the following expression [72]:

$$\begin{aligned} G_{\mathrm{NLIN}}(f) = \frac{16}{27}\gamma^2 L_{\mathrm{eff}}^2 \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} & G_{\mathrm{WDM}}(f_1) G_{\mathrm{WDM}}(f_2) \\ G_{\mathrm{WDM}}(f_1+f_2-f)\rho(f_1,f_2,f)&\chi(f_1,f_2,f)df_1df_2, \end{aligned} \tag{3.5}$$

where $G_{\mathrm{WDM}}(f)$ is the PSD of the WDM signal launched into the fiber defined in (3.4). Also, ρ and χ are the functions that capture the fiber non-degenerate FWM efficiency and NLIN accumulation in multi-span links, respectively. These functions are pulse-shape independent and are defined as:

$$\rho(f_1,f_2,f) = \left| \frac{1-e^{-2\alpha L_s} e^{j4\pi^2(f_1-f)(f_2-f)\beta_2 L_s}}{2\alpha - j4\pi^2(f_1-f)(f_2-f)\beta_2} \right|^2 L_{\mathrm{eff}}^{-2}, \tag{3.6}$$

and

$$\chi(f_1,f_2,f) = \frac{\sin^2\left(2N_s\pi^2(f_1-f)(f_2-f)\beta_2 L_s\right)}{\sin^2\left(2\pi^2(f_1-f)(f_2-f)\beta_2 L_s\right)}, \tag{3.7}$$

where L_s and N_s are the span length and the span number, respectively. Also, L_{eff} is defined as:

$$L_{\mathrm{eff}} = (1-e^{-2\alpha L_s})/2\alpha. \tag{3.8}$$

The NLIN expression in (3.5) clearly demonstrates the dependency of NLIN on the signal PSD, $G_{\mathrm{WDM}}(f)$.

3.3 Pulse Shape Design

This section begins by determining the optimal band-limited pulse PSD based on the GN model using the calculus of variations technique. The resulting pulse, however, may not be practical for communication systems. Therefore, we later add practical constraints for zero-ISI and fast decay on the band-limited pulse.

3.3.1 Optimal Band-limited Pulse for Nonlinearity

In this subsection, we determine the optimal band-limited pulse PSD, $G(f)$, that minimizes nonlinearity based on the GN model in a densely packed WDM system. For now, no other constraints are put on the pulse shape.

Required by the WDM system, we assume the pulse shape bandwidth is limited to $(1+\beta)/2T$, where T is the symbol period and $\beta \in [0,1]$ is the pulse excess bandwidth, which represents the percentage of the bandwidth allocated for a specific design. Therefore, we have:

$$G(f) = 0, \qquad \text{for} \qquad |f| \geq \frac{1+\beta}{2T}. \tag{3.9}$$

To initiate the pulse PSD design process, we calculate the NLIN power interfering with the received symbol energy after the matched-filter as:

$$P_{\text{NLIN}} = \int_{-\infty}^{+\infty} G_{\text{NLIN}}(f) G(f) df. \tag{3.10}$$

Substituting (3.5) and (3.4) into (3.10), we can calculate the NLIN power based on $G(f)$, as:

$$\begin{aligned} P_{\text{NLIN}} = \frac{16}{27}\gamma^2 L_{\text{eff}}^2 P_g^3 & \sum_{i_1=-N/2}^{N/2} \sum_{i_2=-N/2}^{N/2} \sum_{i_3=-N/2}^{N/2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \\ & G(f_1 - i_1 f_0) G(f_2 - i_2 f_0) G(f_1 + f_2 - f - i_3 f_0) G(f) \\ & \rho(f_1, f_2, f) \chi(f_1, f_2, f) df_1 df_2 df. \end{aligned} \tag{3.11}$$

Because we assume a densely packed WDM system in this subsection, $f_0 = (1+\beta)/T$. Equation (3.11) directly represents P_{NLIN} in terms of the pulse PSD, $G(f)$. Additionally, for a normalized symbol energy, we impose the following constraint on the pulse PSD:

$$\int_{-\infty}^{+\infty} G(f) df = 1. \tag{3.12}$$

To minimize (3.11), we apply the calculus of variations technique.

As the starting step, we introduce a transformation to incorporate the constraint in (3.12) automatically. This transformation facilitates the optimization problem significantly. Given that $G(f)$ is a nonnegative function, we represent it as:

$$G(f) = U^2(f) + \zeta, \tag{3.13}$$

where $U(f)$ represents the auxiliary unconstrained functional and $-U^2(f) \leq \zeta$ for $f \in \mathbb{R}$. Additionally, by applying the constraint from (3.12) to (3.13), we can determine ζ as:

$$\zeta = \frac{T}{1+\beta} - \frac{T}{1+\beta} \int_{-\infty}^{+\infty} U^2(f) df. \tag{3.14}$$

Substituting (3.14) into (3.13), we obtain:

$$G(f) = U^2(f) + \frac{T}{1+\beta} - \frac{T}{1+\beta} \int_{-\infty}^{+\infty} U^2(f) df. \tag{3.15}$$

This transformation automatically satisfies the constraint in (3.12). It is important to note that there is not necessarily a one-to-one correspondence between $G(f)$ and $U(f)$. However, for any given $G(f)$, we can identify at least one corresponding $U(f)$ [1]. Therefore, solving the optimization problem based on $U(f)$ and finding its associated $G(f)$, if it exists, is sufficient.

By substituting (3.15) into (3.11), we aim to find the functional $U(f)$ that minimizes P_{NLIN}. Based on the calculus of variations technique, we introduce a scalar, ϵ, and a free function, $\eta(f)$, such that any given $U(f)$ can be expressed as follows:

$$U(f) = U_{\text{opt}}(f) + \epsilon\eta(f). \tag{3.16}$$

Here, $U_{\text{opt}}(f)$ represents the optimal $U(f)$ that minimizes (3.11). By substituting (3.16) into (3.15) and (3.11), P_{NLIN} becomes a function of the scalar ϵ, with its minimum occurring at $\epsilon = 0$. Therefore, we can employ Newton's calculus to locate this minimum with respect to ϵ. Given that this minimum is at $\epsilon = 0$, the following criterion holds true for any η:

$$\frac{\delta}{\delta\epsilon} P_{\text{NLIN}} \Big|_{\epsilon=0} = 0. \tag{3.17}$$

---

[1] For instance, consider $U(f) = \sqrt{G(f)}$.

This criterion serves as a method for determining $U_{\text{opt}}(f)$. In Appendix B.1, we demonstrate that the sole function satisfying this criterion is $U(f) = 0$ which results in the following by substitution in (3.15):

$$G(f) = \begin{cases} \frac{T}{1+\beta} & |f|< \frac{1+\beta}{2T} \\ 0 & \text{else.} \end{cases} \tag{3.18}$$

In addition to minimizing nonlinearity, this pulse achieves zero-ISI for two cases: $\beta = 0$ and $\beta = 1$. Therefore, achieving the minimum possible NLIN and the zero-ISI is possible with this pulse for very low excess bandwidth (for $\beta = 0$) or very high excess bandwidth (for $\beta = 1$). However, for arbitrary values of β, the optimal NLIN pulse, given in (3.18), introduces ISI and is, therefore, impractical for most communication systems. In the next subsection, we will also impose practical constraints on the pulse shape while attempting to minimize NLIN.

3.3.2 Pulse Shape Requirements and Model

A practical pulse needs to satisfy the following fundamental criteria [17]: (i) it must be band-limited, (ii) it should result in zero ISI, and (iii) the pulse shape should decay faster than $\frac{\lambda T}{\pi t}$ to avoid strong ISI in case of timing error. Here, λ controls the decay rate. In this subsection, we establish a pulse model satisfying these criteria. We also relax the dense WDM assumption, i.e., $f_0 \geq (1+\beta)/T$.

To have a band-limited pulse, it is enough that the pulse PSD, $G(f)$, satisfies the condition expressed in (3.9). Also, in [17], it is shown that for zero ISI, we need:

$$g(iT) = \begin{cases} 1 & i = 0 \\ 0 & \text{else.} \end{cases} \tag{3.19}$$

This zero-ISI criterion is based on the assumption of using a filter matched to the pulse shape, $h(t)$, on the receiver side. This criterion guarantees that only one symbol contributes in each pulse sample at $t = 0, \pm T, \cdots$.

The following is a general model for a pulse that satisfies the requirements in (3.9) and (3.19):

$$g(t) = \text{sinc}\left(\frac{t}{T}\right) \cdot \int_0^{\frac{\beta}{2T}} X(f)\cos(2\pi ft)df, \tag{3.20}$$

where $X(f)$ is a free function which can be nonzero only in $[0, \frac{\beta}{2T})$ and with the following constraint:

$$\int_0^{\frac{\beta}{2T}} X(f)df = 1. \tag{3.21}$$

This constraint guarantees a normalized symbol energy as from (20), it is clear that $g(0) = 1$, i.e., (3.12) holds true. Also, in Appendix B.2, we show that to satisfy the third criterion, it is enough to have the following:

$$\int_0^{\frac{\beta}{2T}} |X(f)|df \leq \lambda, \tag{3.22}$$

where $\lambda > 0$ controls the decaying rate.

Considering the band-limited, zero-ISI, and fast decay properties, the pulse model becomes (3.20), and finding the optimal $G(f)$ reduces to selecting $X(f)$ that minimizes (3.11).

3.3.3 Pulse Optimization

In this subsection, we numerically optimize the pulse PSD, $G(f)$, based on the model discussed in the previous subsection, considering an arbitrary excess bandwidth.

As the first step, to convert our problem into a well-defined discrete optimization problem, we utilize the Riemann sum approximation for the integral in (3.20). Consequently, we can define a discrete design parameter X_j as follows:

$$X_j = X\left(\frac{(2j-1)\beta}{4MT}\right), \tag{3.23}$$

where M represents the number of grid points in the Riemann sum. Additionally, using the same technique, (3.20) can be approximated as follows:

$$g(t) = \frac{\beta}{2MT}\text{sinc}\left(\frac{t}{T}\right)\sum_{j=1}^{M} X_j\cos\left(\frac{2\pi(2j-1)\beta t}{4MT}\right). \tag{3.24}$$

In the frequency domain, we can represent (3.24) as follows:

$$G(f) = \frac{\beta}{4M}\sum_{j=1}^{M} X_j \left(\text{rect}\left(fT - \frac{(2j-1)\beta}{4M}\right) + (\text{rect}\left(fT + \frac{(2j-1)\beta}{4M}\right)\right), \tag{3.25}$$

where rect(f) is the rectangular function defined as:

$$\text{rect}(f) = \begin{cases} 1 & |f| < \frac{1}{2} \\ \frac{1}{2} & |f| = \frac{1}{2} \\ 0 & \text{else.} \end{cases} \tag{3.26}$$

Substituting (3.25) into (3.11) results in the NLIN loss function expressed in (3.27) based on the discrete design parameter, X_i. Furthermore, through the application of the same Riemann approximation for the constraints in (3.21) and (3.22), we can formulate the optimization problem for minimizing NLIN as follows:

Minimize

$$P_{\text{NLIN}} = \sum_{j_1=1}^{M}\sum_{j_2=1}^{M}\sum_{j_3=1}^{M}\sum_{j_4=1}^{M} X_{j_1}X_{j_2}X_{j_3}X_{j_4}\theta(j_1, j_2, j_3, j_4), \tag{3.27}$$

Subject to:

$$\sum_{i=1}^{M} X_i = \frac{2MT}{\beta}, \tag{3.28}$$

$$\sum_{i=1}^{M} |X_i| \leq \frac{2\lambda MT}{\beta}, \tag{3.29}$$

where $\theta(j_1, j_2, j_3, j_4)$ is defined as:

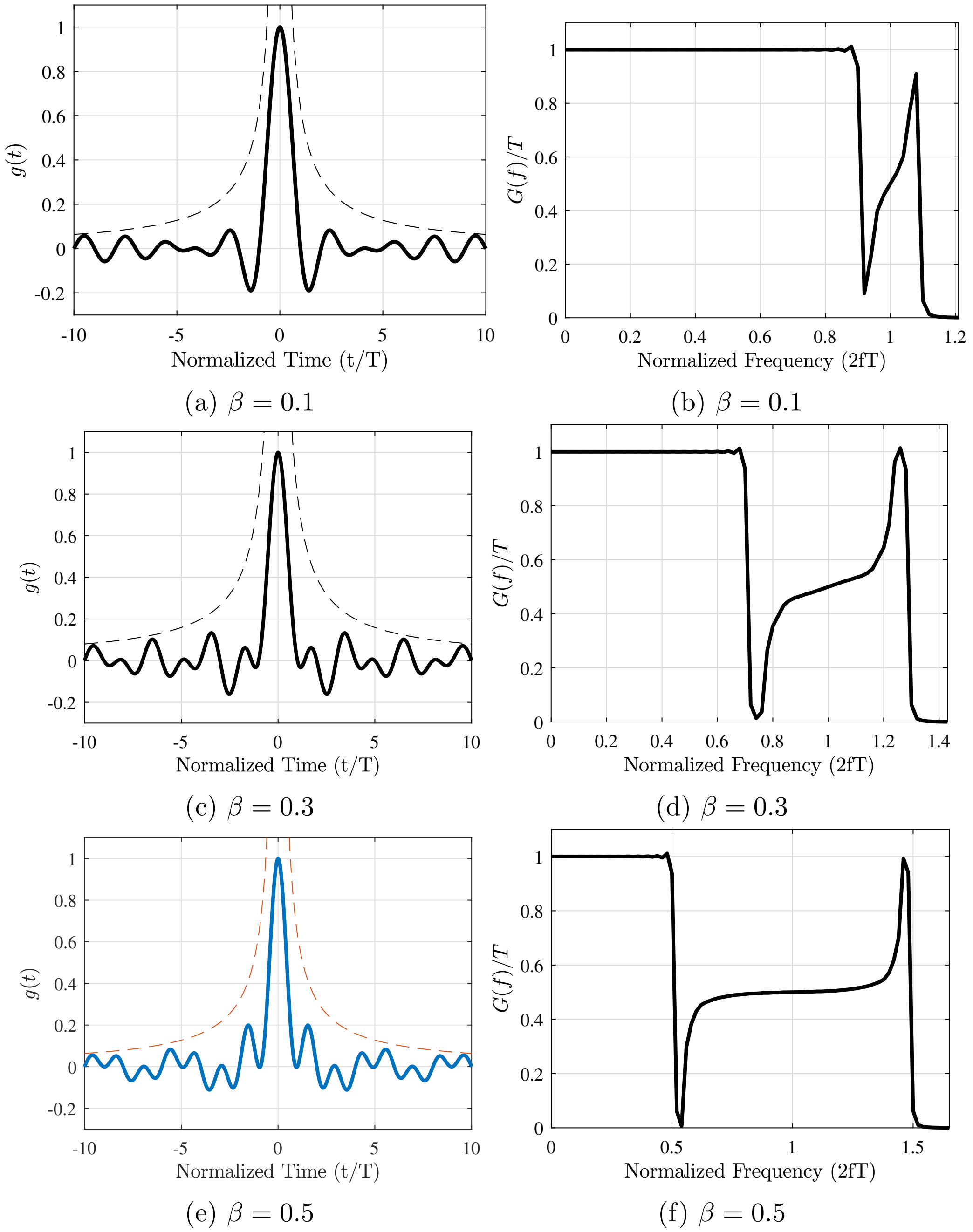
$$\begin{aligned}
&\theta(j_1, j_2, j_3, j_4) = \frac{16}{27}\gamma^2 L_{\text{eff}}^2 P_g^3 \left(\frac{\beta}{4M}\right)^4 \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \\
&\cdot \sum_{i_1=-N/2}^{N/2} \left(\text{rect}\left(f_1T - i_1f_0T - \frac{(2j_1-1)\beta}{4M}\right) + \text{rect}\left(f_1T - i_1f_0T + \frac{(2j_1-1)\beta}{4M}\right)\right) \\
&\cdot \sum_{i_2=-N/2}^{N/2} \left(\text{rect}\left(f_2T - i_2f_0T - \frac{(2j_2-1)\beta}{4M}\right) + \text{rect}\left(f_2T - i_2f_0T + \frac{(2j_2-1)\beta}{4M}\right)\right) \\
&\cdot \sum_{i_3=-N/2}^{N/2} \left(\text{rect}\left((f_1+f_2-f)T - i_3f_0T - \frac{(2j_3-1)\beta}{4M}\right)\right. \\
&\left. + \,\text{rect}\left((f_1+f_2-f)T - i_3f_0T + \frac{(2j_3-1)\beta}{4M}\right)\right) \\
&\cdot \left(\text{rect}\left(fT - \frac{(2j_4-1)\beta}{4M}\right) + \text{rect}\left(fT + \frac{(2j_4-1)\beta}{4M}\right)\right) \\
&\quad \cdot \rho(f_1, f_2, f)\chi(f_1, f_2, f)df_1df_2d_f.
\end{aligned} \tag{3.30}$$

The loss function in (3.27) is a polynomial of degree four. Furthermore, stemming from the physical meaning of P_{NLIN}, we infer that (3.27) is always nonnegative on its domain. In [110]–[112], it is shown that a nonnegative polynomial, such as P_{NLIN}, can be expressed as a sum of squared polynomials:

$$P_{\text{NLIN}} = \sum_i \left(q_i(X_1, \cdots, X_M)\right)^2, \tag{3.31}$$

where q_i is a polynomial to be determined using semidefinite programming. This technique is commonly referred to as the sum of squares (SOS). The result of SOS is in a format suitable for convex optimization. Furthermore, the constraints in (3.28) and (3.29) are convex. Therefore, we have the flexibility to utilize any convex optimization toolbox [113], [114] or a gradient descent algorithm [115] to efficiently solve the optimization problem given in (3.27)-(3.29) after applying the SOS technique.

Considering a typical 3000 km fiber system with parameters expressed in Subsection 3.5.1, the optimal pulse PSDs in the frequency and time domains for three different values of β and $\lambda = 2$ are depicted in Fig. 3.1. Additionally, the numerical values of X for various β values are provided in Appendix B.3.



(a) $\beta = 0.1$ (b) $\beta = 0.1$

(c) $\beta = 0.3$ (d) $\beta = 0.3$

(e) $\beta = 0.5$ (f) $\beta = 0.5$

Figure 3.1: (a) to (c): Time domain representation of the optimal pulse PSD. (d) to (f): Frequency domain representation of the optimal pulse PSD. Assumed fiber system parameters are given in Subsection 3.5.1.

3.4 Discussion on the Designed PSDs

So far, our formulation and pulse optimization were performed in the frequency domain. Thinking in the time-domain, however, one may find our results counter-intuitive. More specifically, since the fiber is a highly dispersive channel, the pulse shape very quickly changes over the fiber length and one may think that any attempt on pulse design is misguided.

In this section, we provide a time domain study of our pulse to ease this tension. Earlier, thinking in the frequency domain, we clarified that this work designs the PSD of the pulse and that dispersion does not change the PSD. Here, we stay focused on time domain. and provide a new understanding of why our proposed pulse has its desirable properties.

The pulse PSD plays an important role in the pulse time domain behavior in a dispersive channel. It is important to note that in a dispersive channel, after traversing a few spans, the temporal shape of the pulse converges towards its spectrum. Consequently, the pulse can be approximated as [16], [116]:

$$h(t,z) \approx \sqrt{\frac{j}{2\pi B_2 z}} \exp\left(-\frac{jt^2}{2B_2 z}\right) H\left(\frac{t}{B_2 z}, 0\right). \tag{3.32}$$

As seen, the pulse in the time domain at distance z, $h(t,z)$, is expressed based on the frequency domain representation that is broadened by a factor of $B_2 z$, $H(\frac{t}{B_2 z}, 0)$. This approximation can be explained by the fact that the dispersive channel separates the frequency components of the pulse over time. Therefore, we anticipate observing the pulse spectrum in the time domain. For example, Fig. 3.2 illustrates the intensity of the optimal pulse with $\beta = 0.5$ propagating in a dispersive fiber channel. As depicted, after traversing a few spans, the temporal shape starts to resemble the spectrum of the pulse and maintains this shape throughout the rest of the fiber except for broadening. Therefore, the pulse PSD has a direct impact on the signal time domain behavior across the fiber length.

The signal power in the time domain determines the nonlinearity power in fiber optics. This fact can be derived by applying the first-order perturbation technique on the NLSE, (3.1). To this end, let us consider the zero term $a^{(0)}(t,L)$ as the solution to the dispersion operator as:

$$a^{(0)}(t,L) = U(L)a(t,0), \tag{3.33}$$

where $U(L)$ is the dispersion operator for a length of L. Then the first-order

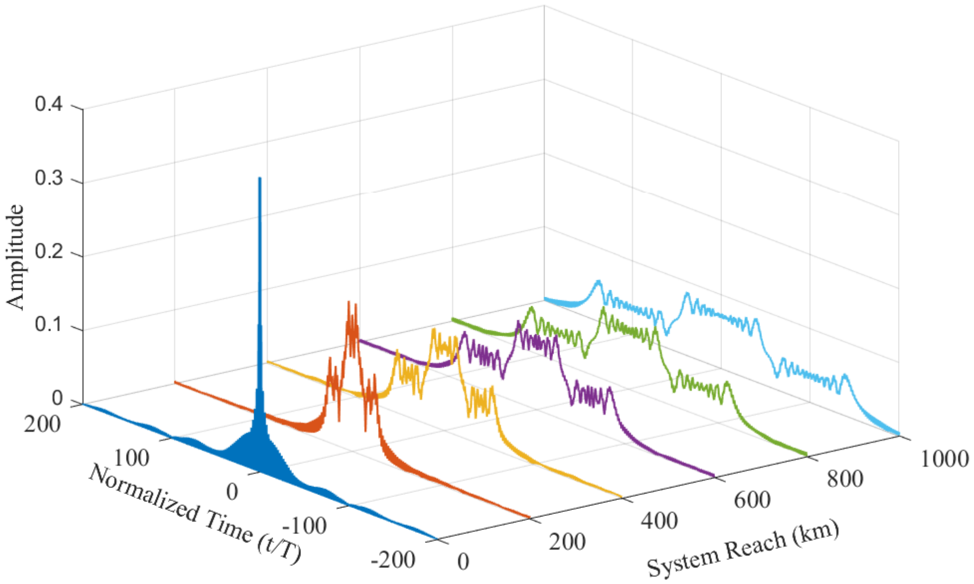


Figure 3.2: The temporal evolution of the optimal pulse intensity with $\beta = 0.5$ in a dispersive fiber channel. After a few spans, the temporal shape starts to resemble the pulse spectrum.

term representing the nonlinearity can be expressed as:

$$a^{(1)}(t, L) = j\gamma \int_0^L U(L - z)|a^{(0)}(t, z)|^2 a^{(0)}(t, z)dz. \tag{3.34}$$

As seen, the nonlinearity term depends on the dispersed signal intensity, $|a^{(0)}(t, z)|^2$. Therefore, a pulse with high power can evidently generate strong NLIN. Based on the observation that the time-domain pulse in a dispersive channel resembles the PSD, we argue that our designed pulse has a smaller region with high power compared to the RC pulse. To illustrate this point, Fig. 3.3 depicts the RC pulse PSD and the designed PSD for $\beta = 0.5$. As shown, our pulse has a small portion with high power in the excess bandwidth. Therefore, we anticipate a lower amount of nonlinearity excited by this pulse compared to the RC pulse.

Finally, to justify the high-power edges of our pulses, we should note that NLIN PSD in a WDM system with reasonable channel spacing has lower power for the high frequencies. Therefore, by emphasizing these regions, our pulse can increase the effective SNR.

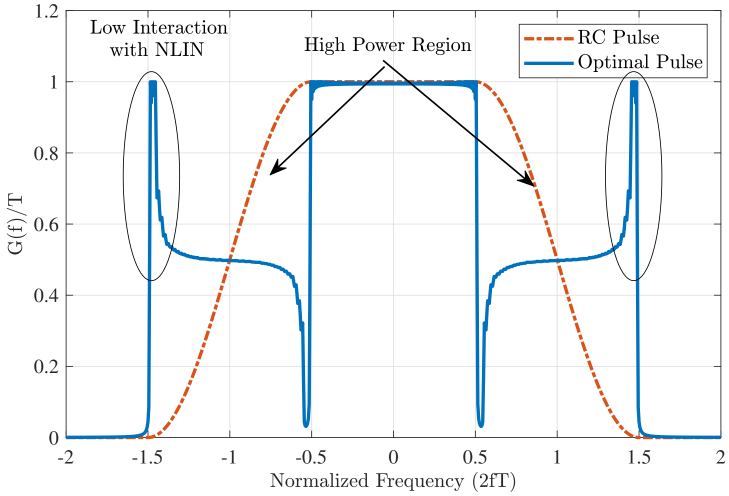


Figure 3.3: Comparison between RC and optimal pulse PSDs with $\beta = 0.5$.

In short, the nonlinear benefits of our pulse stem from the following:

- It reduces the pulse's high-power regions by spreading the symbol energy over time.
- It puts emphasis on the frequencies where NLIN has less effect.

In the end, it is worth mentioning that the sharp edges of the pulse are intended to be loyal to the theoretical findings. In practice, these sharp edges could increase the required filter size for pulse shape implementation. If this is not possible or desirable in a specific design, one can remove the sharp edges to achieve better frequency behavior and a shorter filter time response. Even in this case, with the flat spectrum in the excess bandwidth, a reduction in nonlinearity effects is expected based on our simulation results.

3.5 Simulations

3.5.1 Simulation Setup

We consider a transmitter that chooses a symbol from the 16-QAM constellation and modulates it on a pulse shape with a baud rate of 200 Gbaud. Along with our proposed pulse, we consider a few other pulses with the same bandwidth, symbol rate and roll-off factor of $\beta = 0.3$, including root M-shape [104] and root polynomial [107] pulses. For the simulations, we consider 5 WDM channels with a bandwidth of 260 GHz each and a 10 GHz channel spacing. Our pulse is optimized for the decay control factor of $\lambda = 2$. The results are reported for the central channel on x-polarization.

For this study, we consider a fiber system with a length of 3000 km. The fiber type is SSMF, and its parameters are $\beta_2 = -20 \times 10^{-27}$ s^2/m, $\gamma = 0.013$ (mW)$^{-1}$, and $\alpha = 0.2$ dB/km. The fiber is divided into 100 km spans with an EDFA at the end of each span to compensate for the fiber attenuation. The EDFAs noise figure is 5 dB. In order to find the fiber output, we use SSFM with a step size of 50 m to solve NLSE, [26], accurately.

On the receiver side, we first extract the baseband version of the frequency channels. Symbols are extracted by a filter matched to the transmitted pulse, and then any rotation in the received constellation is calculated and canceled.

3.5.2 Simulation Results

Effective SNR

A pulse shape resulting in lower nonlinearity not only increases the maximum achievable SNR but also shifts the optimal launched power to higher levels. To investigate the optimal launched power of the pulse shapes and the maximum achievable SNR, we launch different power levels into the system and calculate the effective SNR. Fig. 3.4 demonstrates the results. As observed, all four pulses exhibit the same performance in the linear region. This similarity can be attributed to the classical pulse requirements, including band limitation and zero-ISI criterion that our design is constrained to. However, as the channel power increases, resulting in higher NLIN, the pulses demonstrate varying performance.

As observed, our pulse increases the maximum achievable effective SNR by 0.3 dB compared to the RC pulse. Additionally, compared to the M-shape pulse, this increase is 0.2 dB.

On the other hand, our pulse shifts the optimal launch power to higher levels by 0.5 dBm compared to the RC pulse. It is worth mentioning that the

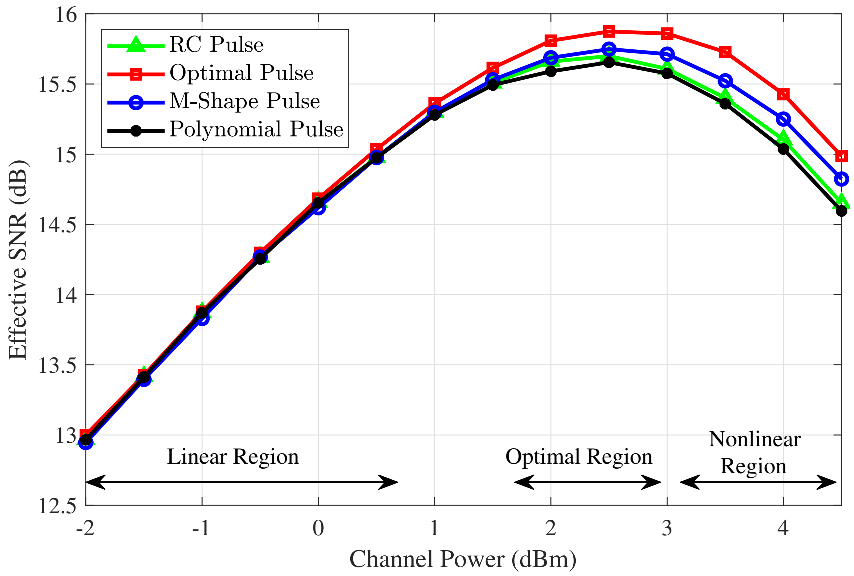


Figure 3.4: Effective SNR for different pulses with $\beta = 0.3$ versus the channel power.

higher optimal launched power can also be interpreted as a longer span length for the same minimum acceptable channel power before amplification, which can reduce the number of required EDFAs and, consequently, the linear noise as well.

In short, in this simulation, we observe that our proposed pulse can better tolerate fiber nonlinearity in terms of maximum achievable SNR and optimal launch power, while experiencing no loss in the linear region compared to existing pulses.

System Reach

A better nonlinearity profile not only increases the effective SNR but can also extend the system's reach. In other words, we can achieve the same signal quality at longer distances when using our proposed pulse. To demonstrate this benefit, we calculate the effective SNR at different distances for the optimal channel power at each distance. Fig. 3.5 shows the results. First, let us consider the line L_2, which indicates the system's reach difference between different pulses around 3000 km. As observed, at this reach, our pulse increases the system's reach by 190 km compared to the RC pulse and by 100 km

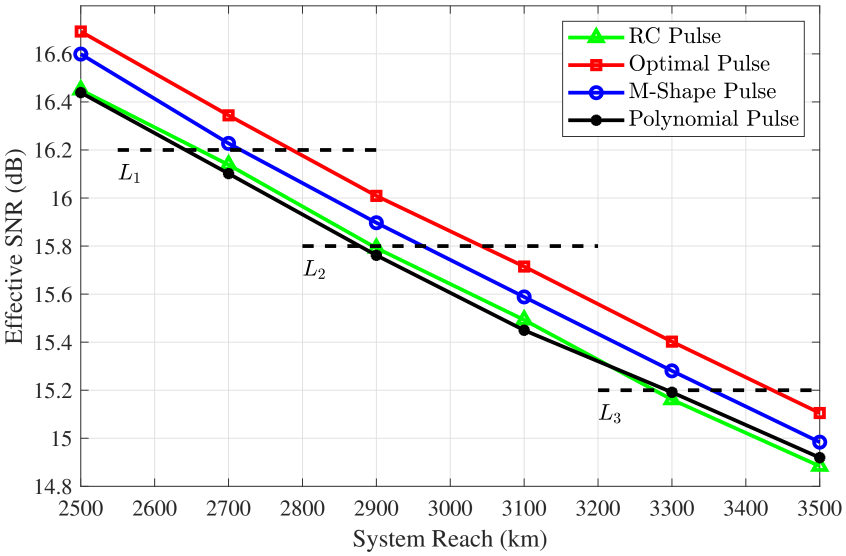


Figure 3.5: Effective SNR for different pulses with $\beta = 0.3$ at different distances.

compared to the M-shape pulse.

On the other hand, a comparison between the lines L_1 and L_2 reveals that for low system reach, the M-shape pulse tends to have performance closer to our pulse compared to higher system reaches, where it approaches the RC pulse. To explain this behavior, we should note that the M-shape pulse only emphasizes low interaction with the NLIN PSD, which is more effective at low system reaches. However, for high system reaches, low pulse power becomes more important, causing it to comparatively lose its performance compared to our pulse.

In summary, in this simulation, we observed a system reach increase of 6% compared to the RC pulse at 3000 km, which can be attributed to the nonlinear benefits of our pulse.

Channel Baud Rate

Increasing the channel baud rate increases the dispersion and nonlinearity effects, limiting the effective SNR. In this simulation, we determine the effective SNR for different channel baud rates while keeping the channel spacing the same at 10 GHz. Fig. 3.6 demonstrates the results.

As indicated by Line L_1, through the use of our pulse, we can increase the

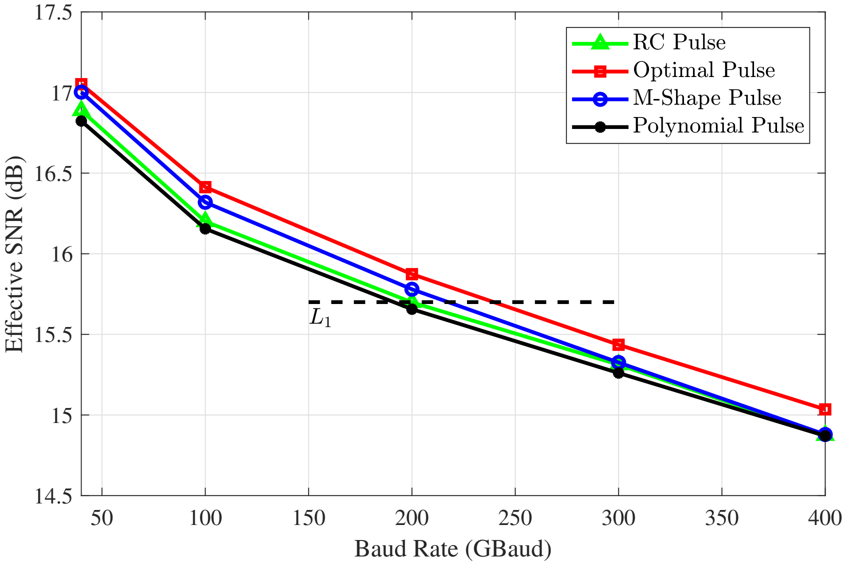


Figure 3.6: Effective SNR for different pulses with $\beta = 0.3$ versus the channel baud rate.

channel baud rate from 190 GBaud for the RC pulse to 240 GBaud for our pulse.

Additionally, when comparing the effective SNR at 40 GBaud and 400 GBaud, we can infer that when the channel spacing is negligible compared to the channel bandwidth, the M-shape pulse loses its performance gain. This is expected because in highly packed WDM systems, the NLIN PSD approaches a flat PSD, and thus emphasizing on the high frequency cannot increase system performance. On the other hand, our pulse can increase the effective SNR even for cases with negligible channel spacing since it reduces the pulse regions with high powers.

In summary, this simulation showed that using our pulse, we can increase the baud rate by 26%. Furthermore, our pulse is effective even for cases where the channel spacing is negligible.

Eye Diagram

An eye diagram generally provides valuable information on ISI and immunity against timing jitter. In this simulation, we plot the eye diagram for the RC and the optimal pulse that we proposed in this work. This eye diagram is extracted on the receiver side after the matched filter and rotation canceler.

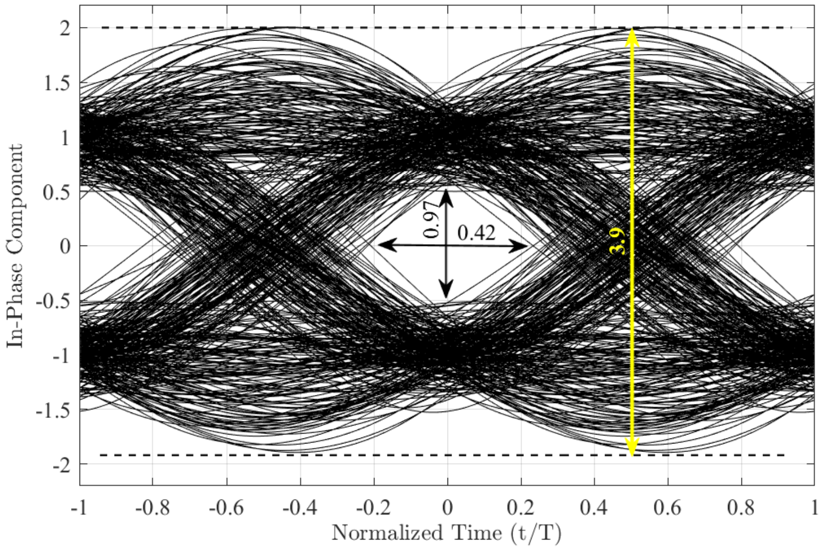


Figure 3.7: Effective SNR for RC Pulse with $\beta = 0.3$ versus the channel baud rate.

The transmitter sends QPSK symbols on a 3000 km fiber with a channel power of 2.5 dBm. Fig. 3.7 and Fig. 3.8 demonstrate the eye diagram for RC and the optimal pulse, both with $\beta = 0.3$, respectively.

As seen, the eye width for the optimal pulse is 0.47, while the RC pulse has an eye width of 0.42. This 12% increase in the eye width suggests better immunity of the optimal pulse against timing jitter compared to RC. Furthermore, the eye height for the optimal pulse and RC are 0.97 and 1.14, respectively. The larger eye height guarantees the higher SNR our pulse shape can offer compared to RC. The main reason for the larger eye height is the lower NLIN in the fiber. As we can observe from the eye diagram, the peak-to-peak fluctuations of RC is 3.9, while the optimal pulse shape has fluctuations of 3.67. Therefore, our pulse shape results in a smaller nonlinearity and suggests a better immunity against fiber nonlinearity and timing jitter.

3.6 Conclusion

Based on the Gaussian noise model, NLIN in optical fibers depends on the pulse PSD. In this work, we focused on finding the optimal pulse PSD that

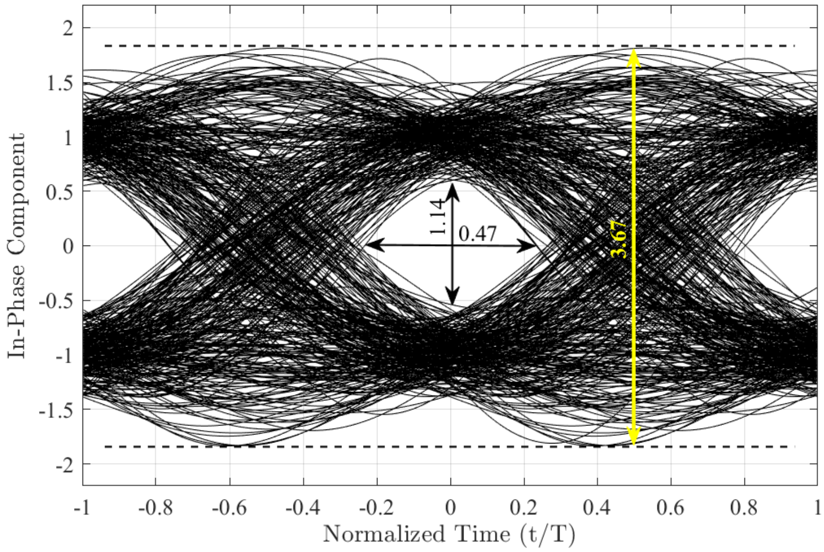


Figure 3.8: Effective SNR for the optimal pulse with $\beta = 0.3$ versus the channel baud rate.

results in the minimum NLIN. We analytically proved that the optimal band-limited pulse with respect to nonlinearity has a flat PSD spanning the entire available bandwidth. This pulse can meet conventional communication requirements, including zero ISI, for very small and very large excess bandwidths. For arbitrary excess bandwidth, we developed a pulse design framework that minimizes NLIN and results in a band-limited zero-ISI pulse with a controlled speed of decay. We also conducted a time-domain analysis to further elucidate the nonlinearity gain achieved by our pulse. Specifically, the nonlinearity gain stems from the following features of our designed pulse PSD. First, our pulse emphasizes frequencies where NLIN has small components. Second, our pulse spreads the symbol energy over time to reduce the portions of the signal with high power. The resulting pulse shape improves various aspects of the system's performance. The simulation study confirmed that our proposed pulse shape reduces NLIN and, in turn, leads to higher optimal launch power, achievable SNR, and system reach compared to other candidate pulses.

Chapter 4

Equalization Enhanced Phase Noise Compensation in Coherent Fiber Receivers

4.1 Introduction

EEPN emerges as a critical impairment, limiting fiber capacity in high baud rate and long reach applications [117]. EEPN results from the interplay between laser phase noise and the dispersion compensator in a suboptimal receiver [7], [118]. In an ideal receiver, the compensation of receiver phase noise should precede dispersion compensation. However, practically estimating the phase noise is infeasible unless dispersion is compensated first [119]. In practical systems, as a result of compensating for the receiver phase noise after the dispersion compensator, the phase noise power increases due to the dispersion compensator.

Studies on EEPN date back to 2008, when dispersion-uncompensated links were proven to be beneficial for fiber nonlinearity. One of the pioneering works defining and studying EEPN [7] accurately predicted a stricter constraint on laser linewidth for systems with high data rates compared to fiber systems with low data rates. In [120], the authors suggest using a dedicated photodetector for estimating the receiver laser phase noise before dispersion compensation to avoid EEPN. However, high-accuracy compensation of the phase noise is not achieved due to practical challenges in phase estimation. The statistical properties of EEPN have been analytically derived in [8]. Additionally, the authors noted that the impact of laser phase noise on the transmitter side is negligible, as it can be compensated by the carrier phase recovery block in the correct order of compensation. Despite these works defining and predicting EEPN properties, EEPN was not considered a serious impairment due to the lower bandwidth of fiber systems in those days.

As the bandwidth of fiber channels has increased in recent years, subsequent works have increasingly focused on the impact of EEPN on systems with high data rates. A comprehensive analysis of EEPN is provided in [9], where the nature of EEPN and its impact on different constellations is studied. The authors propose measures in the system design to limit EEPN, concluding that it can significantly affect system performance, even for metro links utilizing higher-order constellations.

Another recent study on EEPN [3] investigates the impact of EEPN on the carrier phase recovery algorithm and derives characteristics of EEPN for modern receivers. The authors conclude that using a blind carrier phase recovery algorithm can partially compensate for EEPN.

The considerable impact of EEPN on system performance, particularly in long-haul systems, is explored in [121], where the nonlinear behavior of the fiber is also considered. This work integrates the EEPN effect with the Kerr nonlinearity effect based on the Gaussian noise model [71]. In [117], the researchers investigate the impact of the constellation on EEPN impairments. The authors demonstrate that EEPN impairments can be categorized into phase and amplitude noise. Furthermore, they derive the probability density function of these induced noises. In [122], a dual-reference phase estimator is proposed, capable of estimating and compensating for the phase noise for multicarrier systems, showing better tolerance against EEPN. Lastly, [123] presents a comprehensive study on phase recovery algorithms, closely investigating the effect of EEPN.

Traditionally, existing works have regarded EEPN as a random phenomenon, aiming to extract its statistical properties or design the receiver to tolerate the EEPN noise. Although these studies provide valuable insights, there exists a gap in the design of EEPN compensators aimed at removing the EEPN effect.

This work aims to address and fill this void.

In this chapter, we introduce a unique formulation of EEPN that allows for compensation at the receiver's output. To accomplish this, we initially derive the expression of the output in an ideal receiver, where the compensation for receiver phase noise precedes dispersion compensation. Subsequently, we illustrate that this output can be represented in terms of the dispersion compensator output within the existing suboptimal receivers, where phase noise is compensated after dispersion. Consequently, this formulation serves as the basis for our EEPN compensation.

With this novel formulation, we develop two distinct compensators primarily contingent on the availability of the receiver laser phase noise. In cases where receiver laser phase noise is estimated, we show that the EEPN compensator can be implemented based on a basic time-variant finite impulse response (FIR) filter. We refer to it as the phase-dependent compensator, as it relies on the availability of the receiver phase noise. An example of a carrier phase estimator capable of extracting the receiver phase noise can be found in [122], where transmitter and receiver phase noise are separately extracted based on the walk-off between two frequency channels. Additionally, the phase-dependent compensator can be used in scenarios where transmitter phase noise is compensated on the transmitter side [124]. Despite the simplicity of our phase-dependent compensator, extracting the receiver laser phase noise might not be possible for all receiver structures. This is why we proceed to discuss the second compensator.

The second compensator is designed for more practical scenarios in which the receiver phase noise is not known. In this case, we demonstrate that EEPN FIR coefficients can be estimated based on a pilot. For this pilot-based compensator, a simple carrier phase recovery suffices as proposed by [68], [125], [126].

For both compensators, we elaborate on the design parameters and their suitable values for different fiber systems. Finally, we conduct a complexity analysis, revealing that the complexity of the phase-dependent compensator is linearly proportional to the fiber length and data rate, whereas for the pilot-based compensator, this complexity grows cubically with the fiber length and quadratically with data rate.

To validate our theoretical findings, we conduct a simulation study using a typical long-haul fiber system. Our simulations confirm the significant impact of EEPN on high baud rate systems. We showcase the effectiveness of our compensators in mitigating EEPN at the output of a suboptimal receiver, outperforming the existing methods. This improvement in performance is evident across various metrics, including bit error rate (BER), system reach, and effective SNR. Lastly, we illustrate that the runtime of our EEPN compensator is comparable to existing carrier phase recovery methods, ensuring the practical implementability of the design.

This chapter is organized as follows: Section 4.2 presents the system model. In Section 4.3, we derive the compensator expression. Section 4.4 proposes two compensation methods based on the availability of the phase noise, explains the appropriate design parameters for different fiber systems, and conducts a complexity analysis. Simulation results are discussed in Section 4.5, and Section 4.6 concludes this work.

4.2 System Model

In this work, we consider a fiber optic system based on an SMSF. The propagation of light in this channel is governed by NLSE, expressed in (1.16) [88]. In this work, we assume complete fiber dispersion compensation at the receiver. However, it is also possible to derive the compensator for cases where dispersion compensation is divided between the transmitter and receiver.

For WDM systems, the signal needs to be up-converted and down-converted using local lasers at the transmitter and receiver sides, respectively. These lasers introduce phase noise to the signal. Figure 4.1 demonstrates a typical fiber system structure with respect to the transmitter and receiver phase noises, ϕ_t and ϕ_r, respectively. The laser phase noise follows a Wiener process, and consequently, the output of the laser exhibits a Lorentzian spectrum with a linewidth commonly denoted as $\Delta\nu$ [127], [128].

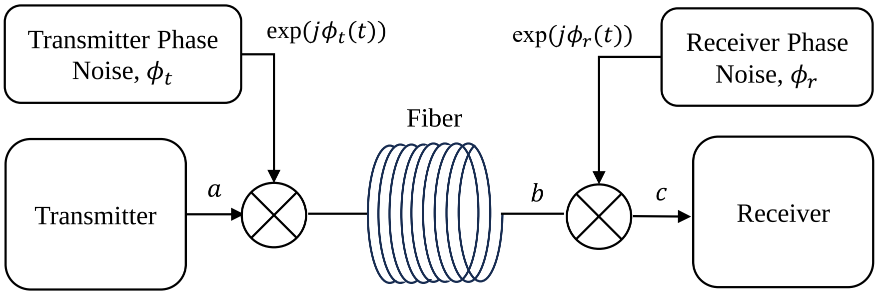


Figure 4.1: A typical fiber system structure with respect to transmitter and receiver laser phase noise.

In an optimal receiver, compensation for transmitter laser phase noise, fiber dispersion, and receiver laser phase noise should occur in the opposite order of their occurrence. Considering the practical challenges, the receiver phase noise is compensated after dispersion compensation, giving rise to EEPN. In [9], the noise caused by EEPN in such a suboptimal receiver is derived as:

$$\sigma^2_{\mathrm{EEPN}} = \frac{\pi^2 \beta_2 L \Delta\nu}{T}, \tag{4.1}$$

where T is the symbol period and L is the fiber length. Therefore, the influential factors on the EEPN are fiber dispersion parameter, fiber length, data rate, and laser linewidth.

In this work, our goal is to design a compensator capable of removing EEPN at the output of such a suboptimal receiver. This compensation relies on either the estimated receiver phase noise or a known pilot. Throughout our formulation, we assume ideal carrier phase recovery, which accurately extracts the laser phase noise, while employing realistic phase estimators in our simulation study.

In the remainder of this chapter, time-domain signals are denoted with regular lowercase letters, and their Fourier transforms are indicated by regular uppercase letters. Matrices and vectors are represented by bold uppercase and bold lowercase letters, respectively. Furthermore, we represent function-function products and matrix products by the symbols $\cdot$ and $\times$, respectively. Continuous signal convolution and discrete signal convolutions are represented by $*$ and $\otimes$, respectively. In this chapter, when a function such as $\exp(\cdot)$ or

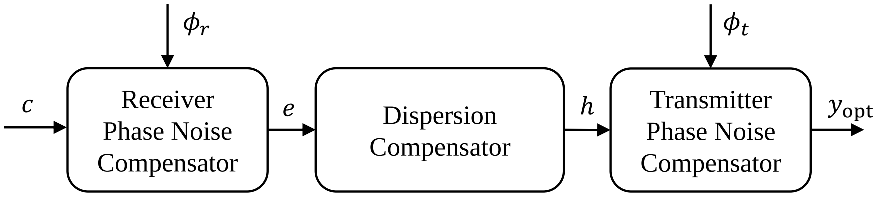


Figure 4.2: Optimal receiver structure, where the receiver's phase noise is compensated for before the dispersion compensator. Despite the simplicity of this receiver, extracting the phase information before dispersion compensation is not possible based on the existing techniques.

sinc(·) is applied to a vector or matrix, it acts on each element individually, and the result is a vector or matrix with the same dimension as the input, respectively. Finally, the superscript ^{T} applied to a matrix represents the transpose operation, which involves reading the rows of the original matrix and writing them as the columns of the resulting matrix.

4.3 EEPN Formulation

In this section, we build upon an optimal receiver structure to derive the EEPN-free output expression. Next, we demonstrate that this expression can be written based on the output of the suboptimal receiver, indicating that it can serve as the EEPN compensator expression.

Figure 4.2 illustrates the optimal receiver structure. The input to the receiver, c, is affected by the transmitter laser phase noise, dispersion, and receiver laser phase noise, as shown in Figure 4.1. To obtain an EEPN-free output in the optimal receiver, the laser phase noises and dispersion are compensated in the reverse order of their occurrence. Despite having the simplest structure and not causing EEPN, this receiver cannot be implemented with existing techniques because pilot-based carrier phase estimators require dispersion to be compensated first. We analyze this structure solely to derive the expression of the EEPN-free output, y_{opt}.

In the optimal receiver, the input to the dispersion compensator, $e(t)$, is the receiver phase noise-compensated version of the received signal, $c(t)$.

Therefore, $e(t)$ can be modeled as follows:

$$e(t) = c(t) \cdot \exp\left(-j\phi_r(t)\right). \tag{4.2}$$

Due to the small linewidth of the laser[1], we know that it can be considered as a slowly varying signal compared to the received signal, $c(t)$. In other words, $\exp\left(-j\phi(t)\right)$ can be accurately approximated with few frequency components. Therefore, within a time window of T_ϕ, we can represent this term based on its frequency components as:

$$\exp\left(-j\phi_r(t)\right) = \sum_{i=-N/2}^{N/2} \zeta_i \exp\left(\frac{j2\pi it}{T_\phi}\right), \tag{4.3}$$

where $N+1$ is the number of nonzero frequency elements, and ζ_i represents the Fourier coefficients, defined as:

$$\zeta_i = \frac{1}{T_\phi} \int_{-\frac{1}{2T_\phi}}^{+\frac{1}{2T_\phi}} \exp\left(-j\phi_r(t) - \frac{j2\pi it}{T_\phi}\right) dt. \tag{4.4}$$

While Fourier series are typically used to represent periodic signals, we are considering only a single period corresponding to the observation window. In the rest of this chapter, we refer to ζ as the phase parameter. Expressing the receiver phase function based on its frequency components enables us to easily track its impact on the dispersion compensator output. By substituting (4.3) into (4.2), we can rewrite the input to the dispersion compensator, $E(f)$, as:

$$\begin{aligned} E(f) &= C(f) * \sum_{i=-N/2}^{N/2} \zeta_i \delta\left(f - \frac{i}{T_\phi}\right) \\ &= \sum_{i=-N/2}^{N/2} \zeta_i C\left(f - \frac{i}{T_\phi}\right), \end{aligned} \tag{4.5}$$

where $*$ denotes the convolution operator. Based on (4.5), we can infer that, in the presence of receiver phase noise, the input to the dispersion compensator is a linear combination of frequency-shifted versions of the received signal $C(f)$

---

[1] Usually $\Delta\nu$, the bandwidth of $\exp\left(j\phi\right)$, is less than 1 MHz.

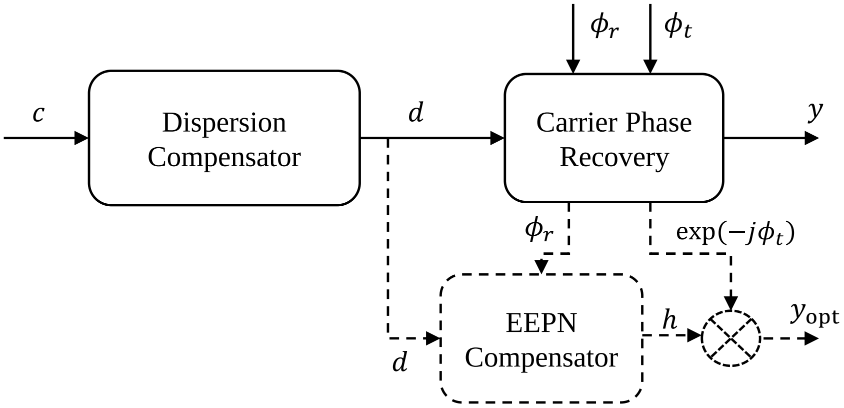


Figure 4.3: Suboptimal receiver structure shown by the solid lines. Due to practical considerations, transmitter and receiver phase noises are estimated and compensated after the dispersion compensator, resulting in EEPN in output y. Our objective is to design an EEPN compensator, depicted by the dashed line, that generates the EEPN-free output y_{opt}.

with a frequency shift of i/T_ϕ. Applying the dispersion operator to $E(f)$, we can express the output of the dispersion compensator as:

$$\begin{aligned} H(f) &= E(f) \cdot \exp\left(-j2\pi^2\beta_2 L f^2\right) \\ &= \sum_{i=-N/2}^{N/2} \zeta_i C\left(f - \frac{i}{T_\phi}\right) \cdot \exp\left(-2j\pi^2\beta_2 L f^2\right). \end{aligned} \tag{4.6}$$

And finally, the optimal output, $y_{\text{opt}}(t)$, is:

$$y_{\text{opt}}(t) = h(t) \cdot \exp\left(-j\phi_t(t)\right). \tag{4.7}$$

In the next step, our objective is to rewrite the EEPN-free output expression, as given in (4.7), relying on the dispersion output within a suboptimal receiver structure. The resulting expression can serve as the EEPN compensator. The solid lines in Figure 4.3 demonstrate a suboptimal structure in which the receiver phase noise is not compensated before the dispersion compensator. This suboptimal structure results in the EEPN in the output y. Our objective is to design the EEPN compensator, as shown by the dashed lines in Figure 4.3, such that it generates the EEPN-free output y_{opt}, as defined in (4.7).

As depicted in Figure 4.3, $D(f)$ represents the output of the dispersion compensator with the input $C(f)$. Therefore, we can express $D(f)$ as:

$$D(f) = C(f) \cdot \exp\left(-j2\pi^2\beta_2 L f^2\right). \tag{4.8}$$

In other to rewrite (4.7) based on $D(f)$, we equate $C(f)$ from (4.8) and substitute it into (4.6), to have:

$$H(f) = \sum_{i=-N/2}^{N/2} \zeta_i D(f - \frac{i}{T_\phi}) \cdot \exp\left(-\frac{j4\pi^2\beta_2 L i f}{T_\phi} + \frac{j2\pi^2\beta_2 L i^2}{T_\phi^2}\right). \tag{4.9}$$

Also, in the time domain, we have:

$$h(t) = \sum_{i=-N/2}^{N/2} \zeta_i d\left(t - \frac{2\pi\beta_2 L i}{T_\phi}\right) \cdot \exp\left(-\frac{j2\pi^2\beta_2 L i^2}{T_\phi^2}\right) \cdot \exp\left(\frac{j2\pi i t}{T_\phi}\right). \tag{4.10}$$

Finally, substituting $h(t)$ into (4.7), we can find the EEPN-free output as:

$$\begin{aligned} y_{\mathrm{opt}}(t) = \exp\left(-j\phi_t(t)\right) &\sum_{i=-N/2}^{N/2} \zeta_i d\left(t - \frac{2\pi\beta_2 L i}{T_\phi}\right) \\ &\cdot \exp\left(-\frac{j2\pi^2\beta_2 L i^2}{T_\phi^2}\right) \cdot \exp\left(\frac{j2\pi i t}{T_\phi}\right). \end{aligned} \tag{4.11}$$

As seen, this expression is based on the dispersion compensator output in a suboptimal receiver, d, and therefore, it can compensate for EEPN in this receiver.

In summary, the expressions derived in (4.10) formulate the EEPN compensator, removing the EEPN effects based on the dispersion output in a suboptimal receiver. This formulation depends on $\boldsymbol{\zeta}$, which is a function of the receiver phase noise. In practice, this phase noise may or may not be available. For example, using special carrier phase recovery techniques such as the one proposed in [122], one can extract the receiver phase noise. Additionally, compensating for the transmitter laser phase noise on the transmitter side [124], [129], [130] enables a simple carrier phase recovery method to extract the receiver phase noise. However, in other cases, the carrier phase recovery extracts the sum of the transmitter and receiver phase noises, making the distinct receiver phase noise unavailable. In the following sections, we design suitable compensators for both scenarios.

4.4 EEPN Compensator

In this section, building upon the proposed expression for EEPN compensation, we introduce two implementations of the EEPN compensator with varying levels of complexity, requirements, and accuracy. More specifically, depending on the availability of the receiver laser phase information, ϕ_r, we propose the following compensators:

- phase-dependent compensator when the receiver phase noise is available,
- pilot-based compensator when total phase noise, $\phi_{tot} = \phi_t + \phi_r$, is compensated, and a pilot sequence is available.

In the following subsections, we consider a DSP-based receiver in which the carrier phase recovery and EEPN compensator are implemented in the time domain. We assume that the EEPN compensator block has access to a block of M_d symbols with a symbol rate of $1/T$. Therefore, $\boldsymbol{d}$, $\boldsymbol{\phi}_t$, $\boldsymbol{\phi}_r$, $\boldsymbol{h}$, and $\boldsymbol{y}$ are all vectors of size M_d.

4.4.1 Phase-dependent compensator for known ϕ_r

In this subsection, we assume that $\boldsymbol{\phi}_r$ is known, and thus $\boldsymbol{\zeta}$ can be found using (4.4). Therefore, the phase carrier recovery block should either be capable of separately estimating the transmitter and receiver laser phase noise [122] or the transmitter phase should be compensated for at the transmitter [124], [129], [130]. In the latter case, any carrier recovery method can estimate $\boldsymbol{\phi}_r$.

Given $\boldsymbol{\phi}_r$, to generate the EEPN-free output, $\boldsymbol{y}$, it is sufficient to implement $\boldsymbol{h}$ according to discrete-time versions of (4.10) and (4.7). Later, we demonstrate that this structure can be simplified to a basic time-variant FIR filter.

Figure 4.4 illustrates the receiver structure suitable for this EEPN implementation. As seen, the carrier phase recovery block determines the receiver phase noise, $\boldsymbol{\phi}_r$, and forwards it to the EEPN compensator. Additionally, the dispersion-compensated signal is obtained from the output of the down-

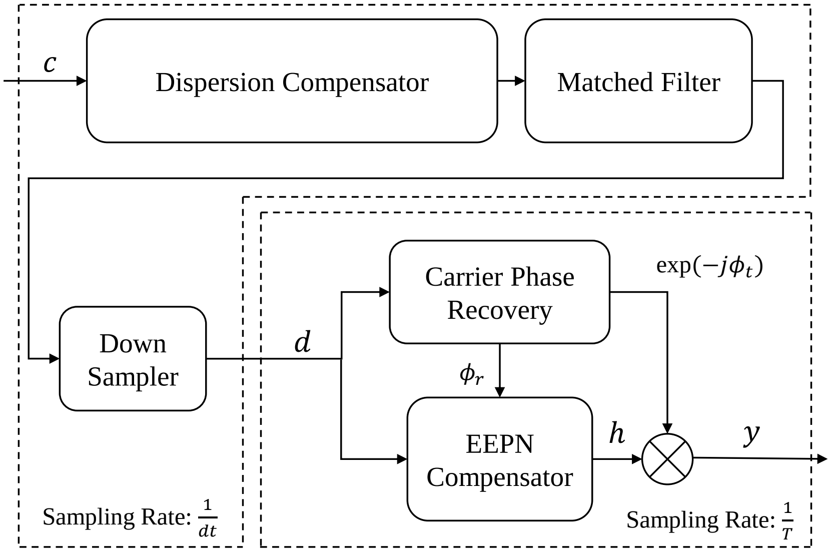


Figure 4.4: Receiver structure suitable for EEPN compensator implementation when ϕ_r is known.

sampler block. The output of the EEPN compensator implements $\boldsymbol{h}$, which is transformed into the optimal output, $\boldsymbol{y}$, by multiplying it with $\exp(-j\boldsymbol{\phi}_t)$.

With the inputs in place, our focus shifts to the design of the EEPN compensator. The initial step involves determining the phase parameter ζ_i as defined in (4.4), based on the estimated phase noise, $\boldsymbol{\phi}_r$. We assume the phase noise has $N+1$ low-frequency components. Therefore, we first down-sample the receiver phase vector $\boldsymbol{\phi}_r$ such that $N+1$ samples remain as:

$$\boldsymbol{\phi}_{\text{red}} = \left[\boldsymbol{\phi}_r\left[k_{\text{init}}\right], \boldsymbol{\phi}_r\left[k_{\text{init}} + \frac{M_d}{N+1}\right], \ldots, \boldsymbol{\phi}_r\left[k_{\text{init}} + \frac{NM_d}{N+1}\right]\right]^{\intercal}, \quad (4.12)$$

where k_{init} represents the index of the first picked sample and can be any value less than $M_d/(N+1)$. Next, we utilize the discrete Fourier transform to find $\boldsymbol{\zeta} = [\zeta_{-N/2}, \zeta_{1-N/2}, \cdots, \zeta_{N/2}]^{\intercal}$, as:

$$\boldsymbol{\zeta} = W \times \exp(-j\boldsymbol{\phi}_{\text{red}}), \quad (4.13)$$

where $W_{(N+1)\times(N+1)}$ entries are defined as:

$$W[i,k] = \frac{1}{N+1}\exp\left(-j2\pi\frac{(i-N/2-1)(k-1)}{N+1}\right). \quad (4.14)$$

As mentioned, ζ_i is the Fourier transform coefficient of the function $\exp\left(-j\phi_r(t)\right)$ over a time window of $T_\phi = M_d T$. It is important to note that (4.13) is the discrete equivalent of (4.4), which can be used to determine the phase parameter $\boldsymbol{\zeta}$ when phase noise $\boldsymbol{\phi}_r$ is known.

In the next step, we implement the discrete version of (4.10) as:

$$\boldsymbol{h} = \sum_{i=-N/2}^{N/2} \zeta_i (\boldsymbol{d} \otimes \boldsymbol{s}^{(i)}) \cdot \exp\left(-\frac{j2\pi^2\beta_2 L i^2}{M_d^2 T^2}\mathbf{1} + \frac{j2\pi i}{M_d T}\boldsymbol{t}\right), \tag{4.15}$$

where

$$\boldsymbol{t} = [0, 1, \ldots, M_d - 1]T, \tag{4.16}$$

and "$\otimes$" denotes the discrete convolution operator with the output length the same as the length of the first input. In this expression, $\mathbf{1}$ is a vector with a size similar to $\boldsymbol{t}$ with all elements equal to 1. The function $\exp(\cdot)$ acts on the input vector elements. In (4.15), the time shift of $2\pi\beta_2 i L/T_\phi$ is implemented using the interpolation filter with the following taps:

$$\boldsymbol{s}^{(i)} = \operatorname{sinc}\left(\frac{\boldsymbol{\tau}}{T} - \frac{2\pi\beta_2 L i}{M_d T^2}\right), \tag{4.17}$$

where $\boldsymbol{\tau} = [-M_s/2, 1 - M_s/2, \ldots, M_s/2]T$ and $M_s + 1$ is the number of the taps for the interpolation filter.

Assuming a block of data is available with size M_d, (4.15) is suitable for EEPN compensator implementation. However, in the following, we demonstrate that (4.15) can also be implemented using a simple time-variant FIR filter. Finally, we derive the update rule for the FIR coefficients based on ϕ_r.

The design principle is to determine the output solely for the symbol at the center of the block, $\boldsymbol{h}$. We demonstrate that this central symbol can be compensated using an FIR filter. Evidently, by shifting the block, we can compensate for all the received symbols. To derive the FIR filter, we need to express (4.15) for the center symbol, $\boldsymbol{h}[M_d/2]$, as:

$$\boldsymbol{h}\left[\frac{M_d}{2}\right] = \left(\boldsymbol{d} \otimes \sum_{i=-N/2}^{N/2} \zeta_i \boldsymbol{s}^{(i)} \cdot \exp\left(-\frac{j2\pi^2\beta_2 i^2 L}{M_d^2 T^2} + j\pi i - \frac{j2\pi i}{M_d}\right)\right)\left[\frac{M_d}{2}\right]. \tag{4.18}$$

In the derivation of (4.18), the distributive property of the convolution operator was also utilized. We can express (4.18) in matrix format as:

$$\begin{aligned} \boldsymbol{h}\left[\frac{M_d}{2}\right] &= (\boldsymbol{d}\otimes\boldsymbol{\eta})\left[\frac{M_d}{2}\right] \\ &= \boldsymbol{d}\left[\frac{M_d+M_s}{2},\ldots,\frac{M_d-M_s}{2}\right]\times\boldsymbol{\eta}, \end{aligned} \tag{4.19}$$

where $\boldsymbol{\eta}_{(M_s+1)\times 1}$ can be expressed as:

$$\begin{aligned} \boldsymbol{\eta} &= \boldsymbol{\Delta}\times\boldsymbol{\zeta} \\ &= \boldsymbol{\Delta}\times\boldsymbol{W}\times\exp\left(-j\boldsymbol{\phi}_{\text{red}}\right). \end{aligned} \tag{4.20}$$

Also, $\boldsymbol{\zeta}$ is defined in (4.13) and the entries of $\boldsymbol{\Delta}_{(M_s+1)\times(N+1)}$ are defined as:

$$\begin{aligned} &\boldsymbol{\Delta}_{(M_s+1)\times(N+1)}[k,i] \\ &\quad= \operatorname{sinc}\left(1-k+M_s/2-\frac{2\pi\beta_2 L(i-N/2-1)}{M_d T^2}\right) \\ &\qquad\cdot\exp\left(-\frac{j2\pi^2\beta_2 L(i-N/2-1)^2}{M_d^2T^2}+j\pi(i-N/2-1)-\frac{j2\pi(i-N/2-1)}{Md}\right). \end{aligned} \tag{4.21}$$

It is important to note that (4.19) defines a time-variant FIR filter with coefficients $\boldsymbol{\eta}$ capable of compensating for EEPN, and the update rule for this EEPN FIR filter is defined in (4.20). Additionally, we should note that $\boldsymbol{\Delta}\times\boldsymbol{W}$ represents a fixed matrix, requiring calculation and storage only once.

Figure 4.5 illustrates the block diagram depicting the FIR implementation of the EEPN compensator. The delay units synchronize the update rule with respect to the EEPN FIR filter. According to (4.12), the down-sampler samples the receiver laser phase noise every $M_d/(N+1)$ symbols. Consequently, the EEPN FIR filter should be updated every $M_d/(N+1)$ symbols because, during this period, the phase parameter $\boldsymbol{\zeta}$ remains constant. Overall, this implementation introduces a symbol delay of $M_d/2$ symbols in its output.

It is worth noting that there is no need for receiver phase noise cancellation after the EEPN compensator because the output of the EEPN compensator is equivalent to the optimal receiver, where the receiver phase noise is compensated before the dispersion compensator.

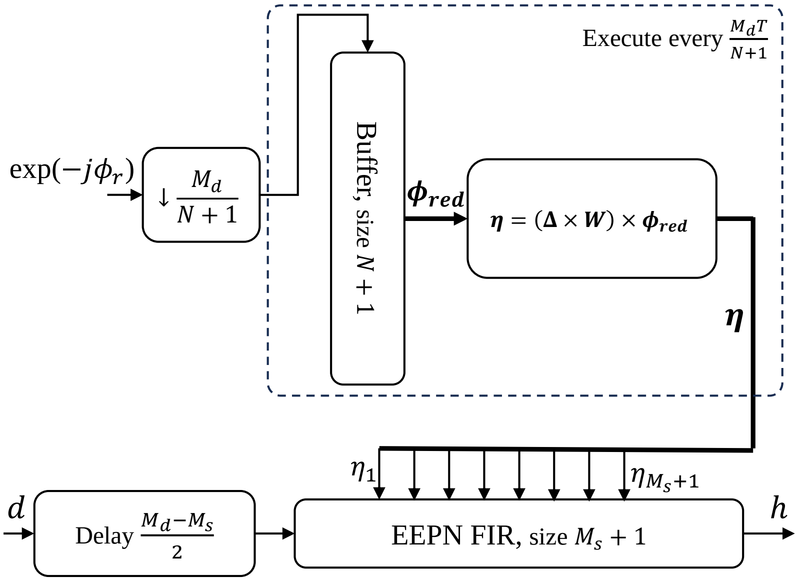


Figure 4.5: Phase-dependent EEPN compensator implementation based on the time-variant FIR defined in (4.20).

In short, in this subsection, assuming $\boldsymbol{\phi}_r$ is known, we derive a simple time-varying FIR filter to compensate for EEPN. The biggest advantage of this implementation is its simplicity and accuracy, as our simulation results reveal in Section 4.5.

In some systems, although distinguishing the phase noises $\boldsymbol{\phi}_t$ and $\boldsymbol{\phi}_r$ might be challenging. Thus, in the following, we propose another compensator relaxing the need for receiver laser phase noise, $\boldsymbol{\phi}_r$.

4.4.2 Pilot-based compensator

In this subsection, we consider a more practical scenario in which we determine the EEPN FIR filter coefficients based on a pilot signal and the sum of the receiver and transmitter laser phase noises, $\phi_{\text{tot}} = \phi_t + \phi_r$. This approach alleviates the need to know the explicit receiver laser phase noise, $\boldsymbol{\phi}_r$.

To determine the EEPN FIR coefficients independently of the receiver phase noise, we initially establish a system of equations to find the phase parameter $\boldsymbol{\zeta}$ based on the pilot. Subsequently, we employ the extracted $\boldsymbol{\zeta}$ to find the FIR coefficients and compensate for EEPN.

In Figure 4.6, the structure of a symbol block, including the pilot sequence

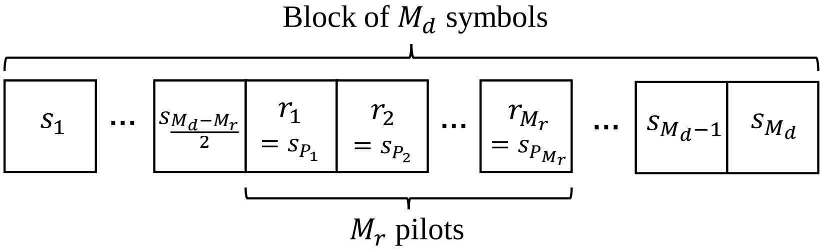


Figure 4.6: A sequence of M_r symbols is considered as the pilot in a block of M_d symbols. These symbols are denoted by $r_i = s_{P_i}$, where P_i is the pilot symbol index and defined as $P_i = i + (M_d - M_r)/2$. In our design, pilots are assumed to be a random but known sequence of symbols drawn from the same constellation as the data.

$\boldsymbol{r}$, is illustrated. Within a block of M_d symbols, there are M_r pilot symbols. The decision to use a centralized pilot section is to allow us to assume constant phase noises for all the pilots. As seen, we denote the i-th pilot index as $P_i = i + (M_d - M_r)/2$ for $i = 1, 2, \ldots, M_r$. It is essential to note that while we consider the pilots to be known, in practice, they can be the detected received symbols.

Furthermore, in Figure 4.7, the suitable receiver structure for the pilot-based compensator is depicted. As observed, for this compensator, we relocate the phase noise compensator before the EEPN FIR filter resulting in the total phase noises, $\phi_{\text{tot}} = \phi_t + \phi_r$, to be compensated before the EEPN compensator. This adjustment is possible because the phase noises within each EEPN FIR update can be treated as constant. Consequently, the input of the EEPN FIR filter is the output of the suboptimal receiver, $\boldsymbol{y}$.

In order to find $\boldsymbol{\zeta}$ based on the pilot sequence $\boldsymbol{r}$, we begin by deriving the pilot expression based on the unknown $\boldsymbol{\zeta}$. Assuming $M_s + 1$ FIR taps as $\boldsymbol{\eta}$, we write the FIR output for pilot $\boldsymbol{r}[i]$ based on (4.19) as:

$$\boldsymbol{r}[i] = \boldsymbol{u}_i \times \boldsymbol{\eta}, \tag{4.22}$$

where, $\boldsymbol{u}_i = \boldsymbol{y}[P_i + M_s/2, P_i + M_s/2 - 1, \ldots, P_i - M_s/2]$. Substituting the filter coefficients based on (4.20), we have:

$$\boldsymbol{r}[i] = \boldsymbol{u}_i \times \boldsymbol{\Delta} \times \boldsymbol{\zeta}. \tag{4.23}$$

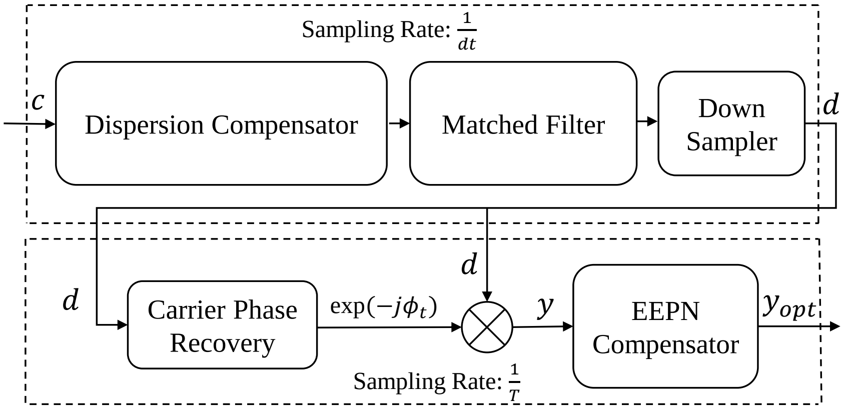


Figure 4.7: Receiver structure for pilot-based EEPN compensator.

This expression establishes one equation based on the known pilot $\boldsymbol{r}[i]$ and the unknown phase parameter $\boldsymbol{\zeta}$. Considering M_r pilots, we can formulate the following system of equations:

$$\boldsymbol{r} = \boldsymbol{U} \times \boldsymbol{\Delta} \times \boldsymbol{\zeta}, \tag{4.24}$$

where $\boldsymbol{U}_{M_r \times (M_s+1)}$ is a matrix with the i-th row as $\boldsymbol{u}_i$. We can find the solution of (4.24) for phase parameter, $\boldsymbol{\zeta}$, as:

$$\boldsymbol{\zeta} = ((\boldsymbol{U} \times \boldsymbol{\Delta})^T \times \boldsymbol{U} \times \boldsymbol{\Delta})^{-1} \times (\boldsymbol{U} \times \boldsymbol{\Delta})^T \times \boldsymbol{r}. \tag{4.25}$$

Then, we determine the EEPN FIR coefficients to compensate for EEPN by substituting (4.25) into (4.20) to obtain:

$$\boldsymbol{\eta} = \boldsymbol{\Delta} \times ((\boldsymbol{U} \times \boldsymbol{\Delta})^T \times \boldsymbol{U} \times \boldsymbol{\Delta})^{-1} \times (\boldsymbol{U} \times \boldsymbol{\Delta})^T \times \boldsymbol{r}. \tag{4.26}$$

It is worth mentioning that in calculating $\boldsymbol{\zeta}$, we use a pseudo-inverse to solve the overdetermined system of equations in (4.24). This approach is optimal when assuming Gaussian noise impacts the received pilots.

As seen, (4.26) defines the coefficients of the EEPN FIR filters that can be adaptively updated based on the known pilot $\boldsymbol{r}$ and the received symbols matrix $\boldsymbol{U}$. It is crucial to note that the matrices $\boldsymbol{\Delta}$ and $\boldsymbol{U}$ are not square, and therefore, the equation in (4.25) cannot be simplified for the general case.

Based on our formulation in (4.26), we propose the block diagram presented in Figure 4.8 for the implementation of the pilot-based EEPN compensator.

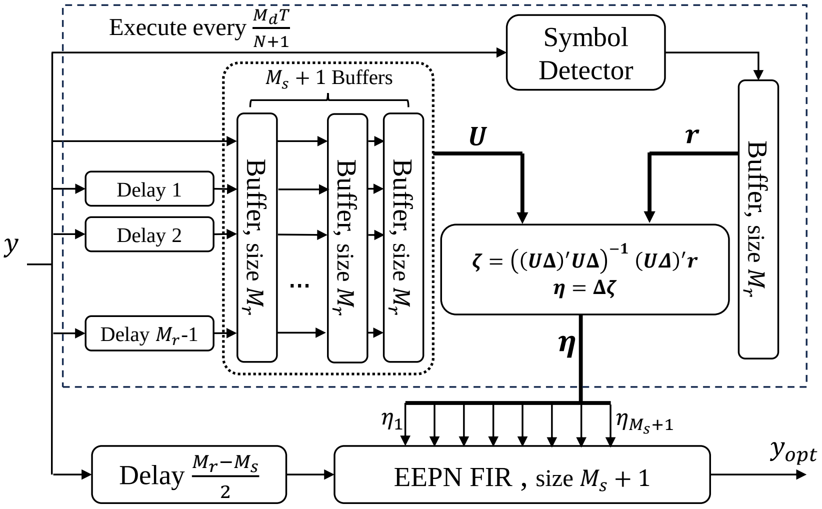


Figure 4.8: Pilot-based EEPN compensator implementation based on the pilot-driven phase parameter defined in (4.25).

In this diagram, we leverage a symbol detector to generate the pilot sequence, eliminating the need for transmitting preset symbols. As mentioned in the previous subsection, the EEPN FIR coefficients need to be updated for every $M_d/(N+1)$ symbols.

While EEPN FIR coefficients can be directly determined based on the phase parameter $\boldsymbol{\zeta}$ as derived in (4.26), it is worth mentioning that one can use (4.25) to separately estimate the receiver laser phase noise in a suboptimal receiver. Based on (4.3), $\boldsymbol{\zeta}$ represents the Fourier transform coefficients of $\exp(-j\phi_r)$, allowing one to extract ϕ_r using:

$$\phi_r(t) = j \ln \left(\sum_{i=-N/2}^{N/2} \zeta_i \exp\left(\frac{j2\pi i t}{M_d T}\right) \right), \tag{4.27}$$

where $\ln(\cdot)$ denotes the natural logarithm.

4.4.3 Discussions on the Design Parameters

The compensators designed in Subsections 4.4.1 and 4.4.2 include design parameters such as block size M_d, phase parameter number N, the number of pilots M_r, and FIR length M_s. In this subsection, we will elaborate on each parameter, its meaning, and the appropriate value based on the system properties. In the end, we present the numerical values of the parameters for commonly used systems.

Block length, M_d

This parameter is defined as the symbol interval during which the compensator observes the laser phase noise to compensate for EEPN.

Let us assume our aim is to compensate for the symbol at time t_0. Clearly, the minimum interval required to observe the phase noise is the interval during which the receiver laser phase noise affects the symbol at time t_0. To determine this minimum required interval, we derive the impulse response of the receiver laser phase noise. To this end, in Figure 4.1, we assume that the output of the laser is an impulse, $\delta(t)$, resulting in $c = b(0)\delta(t)$. Based on (4.8), the output of the dispersion compensator is given by:

$$d(t) = b(0)\mathrm{iFT}\left\{\exp\left(-j2\pi^2\beta_2 L f^2\right)\right\}, \tag{4.28}$$

where iFT is the inverse Fourier transform. Hence, the suboptimal receiver broadens the impact of the phase noise by the length of the dispersion impulse response. We can infer that the symbol at time t_0 is affected by the receiver laser phase noise in the interval $[t_0 - t_d/2, t_0 + t_d/2]$, where t_d represents the impulse response length for the dispersion operator. Therefore, for EEPN compensation, it is sufficient to consider $M_d = t_d/T$. In [131], the dispersion impulse response length, t_d, is derived. Utilizing their findings, we set the EEPN block size to match the dispersion impulse length as:

$$M_d = \frac{2\pi|\beta_2|L}{T^2} + 1. \tag{4.29}$$

With this equation, we anticipate the block size to increase at a rate proportional to the square of the baud rate and linearly with respect to the fiber length.

Phase Parameters Number, N

As indicated in (4.4), $N + 1$ is the number of nonzero Fourier coefficients of the laser phase noise function, $\exp(-j\phi_r(t))$. This parameter also determines the EEPN FIR update rate given by $M_d/(N + 1)$ symbols.

To determine N, we recognize that the frequency resolution of the discrete Fourier transform in (4.13) is $1/(M_dT)$. Therefore, with a bandwidth of ω for $\exp(-j\phi_r(t))$, we deduce that the number of nonzero ζ_is is:

$$N = M_dT\omega. \tag{4.30}$$

Here, we should mention that ω is different from the laser linewidth in this manner: ω observes the nonzero frequency components of the laser phase noise, while the laser linewidth represents the 3-dB bandwidth. Although the laser linewidth has a Lorentzian spectrum with tails stretching to infinity, an approximate absolute bandwidth can be used. In our study, based on numerical trials, we found that $\omega = 1000\Delta\nu$ is sufficient.

FIR length, M_s

FIR filters in (4.15) are used to implement the maximum delay of $\pm\pi\beta_2NL/M_dT$. The EEPN FIR filter in (4.18) is also a sum of these FIR filters. Therefore, the EEPN FIR filter also has the same number of taps.

As seen in (4.17), the coefficients of $s^{(i)}$ are a sinc function. Based on our extensive simulations, for good accuracy, it is enough to keep five cycles of the sinc function as the FIR taps. Keeping this number of cycles maintains 98% of the ideal filter energy while facilitating a low-complexity implementation. Therefore, to implement the maximum delay of $\pm\pi\beta_2NL/M_dT$, we need:

$$M_s = 10 + \frac{2\pi\beta_2LN}{M_dT^2} \tag{4.31}$$

taps for the FIR filters.

Number of Pilots, M_r

Pilots are employed in (4.25) to determine the phase parameter, $\boldsymbol{\zeta}$. Since we have $N+1$ nonzero ζ, we require at least $N+1$ pilots ($M_r > N$). Based on the numerical values presented in Table 4.1 for different fiber systems, the pilot overhead is less than 1.3%. However, increasing the number of pilots in noisy scenarios can be advantageous.

The derived expressions in (4.29), (4.30), and (4.31) can be utilized to determine the optimal design parameters for EEPN compensators. Table 4.1 also presents these parameters for various channel bandwidths and laser linewidths. In this table, we assume the fiber length is $L = 1000$ km. However, considering the linear relation impact of the fiber length on M_d and N, we can find suitable parameters for any fiber length with the proper scaling.

	Linewidth = 50 kHz			Linewidth = 100 kHz		
BW GHz	M_d	N	M_s	M_d	N	M_s
50	315.2	0.3	10.3	315.2	0.6	10.6
100	1257.6	0.6	10.6	1257.6	1.3	11.3
150	2828.4	0.9	10.9	2828.4	1.9	11.9
200	5027.5	1.3	11.3	5027.5	2.5	12.5
250	7855.0	1.6	11.6	7855.0	3.1	13.1
300	11310.7	1.9	11.9	11310.7	3.8	13.8
350	15394.8	2.2	12.2	15394.8	4.4	14.4
400	20107.2	2.5	12.5	20107.2	5.0	15.0
450	25447.9	2.8	12.8	25447.9	5.7	15.7
500	31416.9	3.1	13.1	31416.9	6.3	16.3

	Linewidth = 200 kHz			Linewidth = 500 kHz		
BW GHz	M_d	N	M_s	M_d	N	M_s
50	315.2	1.3	11.3	315.2	3.2	13.1
100	1257.6	2.5	12.5	1257.6	6.3	16.3
150	2828.4	3.8	13.8	2828.4	9.4	19.4
200	5027.5	5.0	15.0	5027.5	12.6	22.6
250	7855.0	6.3	16.3	7855.0	15.7	25.7
300	11310.7	7.5	17.5	11310.7	18.9	28.8
350	15394.8	8.8	18.8	15394.8	22.0	32.0
400	20107.2	10.1	20.1	20107.2	25.1	35.1
450	25447.9	11.3	21.3	25447.9	28.3	38.3
500	31416.9	12.6	22.6	31416.9	31.4	41.4

Table 4.1: Compensator parameters for different fiber systems. In preparing this table, we adhered to the exact numbers provided by (4.29) to (4.30). In practice, however, the ceiling of these numbers should be used to determine the integer values for the filter taps or pilot numbers.

4.4.4 Complexity Analysis

In this subsection, we analyze the complexity of the proposed compensators. Specifically, we derive the number of complex multiplications and summations required per one symbol. We assume the compensators have the parameters M_d as the block size, N as the number of phase parameters, M_s as the filter length, and M_r as the pilot count.

We begin by counting the number of operations required for the phase-dependent compensator proposed in Subsection 4.4.1. For every output symbol, the EEPN FIR filter should be computed, which results in M_s+1 complex products and M_s complex sums. Furthermore, the update rule necessitates a matrix-vector multiplication where the matrix is of size (M_s+1) by $(N+1)$. Therefore, the number of complex products is $(M_s+1)\times(N+1)$, and the number of complex sums is $(M_s+1)\times N$. Because updating the EEPN FIR filter is required every $M_d/(N+1)$ symbols, the overall number of complex products and sums normalized for one symbol for this compensator is:

$$N_{p1} = \frac{(N+1)^2(M_s+1)}{M_d} + M_s + 1, \tag{4.32}$$

and

$$N_{s1} = \frac{N(N+1)(M_s+1)}{M_d} + M_s, \tag{4.33}$$

respectively. According to the numerical values presented in Table 4.1 for different fiber systems, we realize that the complexity of the update procedure is negligible compared to the EEPN FIR filter. By substituting (4.29), (4.30), and (4.31), this complexity is of order $\mathcal{O}(L\omega/T)$.

For the pilot-based compensator, similar to the first compensator, we require M_s+1 complex products and M_s complex sums for the implementation of the EEPN FIR filter. On the other hand, for updating the filter coefficients based on (4.25), we first compute $\boldsymbol{U}_{M_r\times(M_s+1)}\times\boldsymbol{\Delta}_{(M_s+1)\times(N+1)}$, which requires $M_r(M_s+1)(N+1)$ complex products and $M_rM_s(N+1)$ complex sums. We then calculate the term $(\boldsymbol{U}\times\boldsymbol{\Delta})^T\times\boldsymbol{U}\times\boldsymbol{\Delta}$ with $(N+1)^2M_r$ complex products and $(N+1)^2(M_r-1)$ complex sums. Next, we determine the inverse of the result with a maximum of $(N+1)^3/3$ complex products and $(N+1)^3/3$ complex sums based on the Gaussian elimination method. Calculating the rest of the expression results in:

$$(M_s+1)\left((N+1)^2+(N+1)M_r+M_r\right) \tag{4.34}$$

complex products and

$$(M_s+1)\left(N(N+1)+NM_r+M_r-1\right) \tag{4.35}$$

complex sums. Considering the update rule for $M_d/(N+1)$ symbols, the normalized complex products and sums for the pilot-based EEPN compensator are:

$$\begin{aligned} N_{p2} &\approx N_{s2} \\ &\approx \frac{M_r N(N^2+2M_r N+M_r)+(N+1)^4/3}{M_d}+M_s. \end{aligned} \tag{4.36}$$

Thus, substituting (4.29), (4.30), and (4.31), we determine that the complexity of the pilot-based compensator is governed by the FIR coefficient update procedure, and it is of the order $\mathcal{O}(L^3\omega^4/T^2)$.

Based on this analysis, we observe that the complexity of the phase-dependent compensator is mainly determined by the EEPN FIR filter, which is linearly proportional to the fiber length and data rate, while the complexity of the pilot-based compensator is determined by the update rule and is proportional to the cube of the fiber length and the square of the data rate.

4.5 Results and Discussions

In this section, we first present our simulation setup, followed by the results on the performance of the proposed compensators. As a benchmark, we compare our compensators with the blind phase search (BPS) method [132]. BPS has been shown to tolerate EEPN in various works, including [3], [117], [133], [134]. In addition, we compare our compensator's performance to the recently proposed dual reference subcarrier (DRS) approach [122]. As the optimal phase compensator in the absence of dispersion, we present the results of zero-forcing phase estimator (ZFP) recovery with ideal knowledge of phase noises.

We refer to our compensator in Subsection 4.4.1 and Subsection 4.4.2 as Phase-Dependent Comp and Pilot-Based Comp, respectively. In light of (4.1), we have opted to investigate the performance of our compensator concerning the impact of data rate, laser linewidth, and fiber length.

4.5.1 Simulation Setup

We consider our simulation setup to be aligned with recent advancements in carrier phase recovery and the EEPN challenge, as demonstrated in works such as [121]–[123]. Accordingly, we model a 2000 km fiber link with 5 WDM dual polarization channels and a combined laser linewidth of 400 kHz, evenly distributed between the transmitter and receiver. In the following, we describe our setup in more detail.

On the transmitter side, we generate a stream of 2^{23} random bits and, using gray labeling, map them to 16-QAM symbols. These symbols are then modulated to a root-raised cosine pulse with a roll-off factor of 10%. We consider a channel bandwidth of 200 GHz with a channel power of 2 dBm unless otherwise specified. The resulting signal is multiplied by the laser phase noise and then launched into the fiber.

We consider a typical fiber with the following parameters: $\beta_2 = -20 \times 10^{-27}$ s^2/m, $\gamma = 0.013$ $(\mathrm{mW})^{-1}$, and $\alpha = 0.2$ dB/km. This fiber is simulated based on NLSE and is implemented using the SSFM [29]. The step size we consider for this method is 50 m to ensure high accuracy. To compensate for fiber attenuation, we employ EDFA every 100 km. The EDFA noise figure is 5 dB. No inline dispersion compensation is implemented, necessitating complete dispersion compensation on the receiver side.

On the receiver side, we first apply receiver laser phase noise, and then compensate for dispersion and apply the matched filter. For laser phase noise estimation, we assume a pilot sequence is transmitted every 40 symbols, causing an overhead of 2.5%. Finally, the EEPN compensator is implemented. Considering the mentioned system properties and the discussions in Subsection 4.4.3, we select a suitable set of compensator parameters for each simulation. For the phase-dependent compensator, we assume effective laser linewidth on

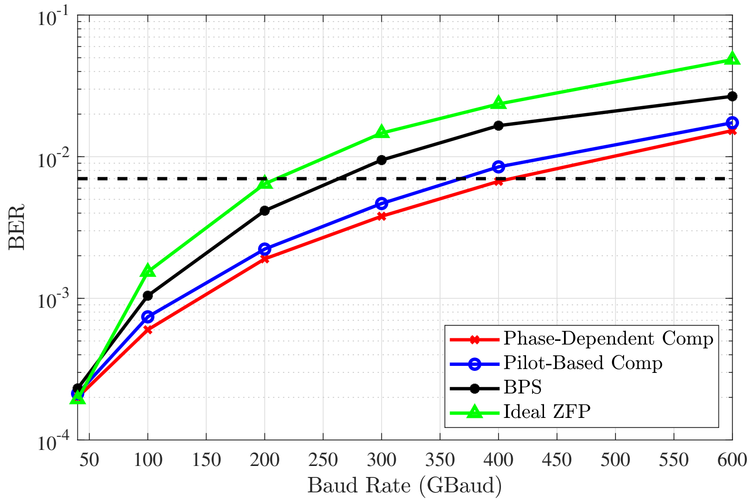


Figure 4.9: BER performance of the compensators for different channel baud rates.

the transmitter side is reduced by 90% using the feedback compensation technique. This value is chosen based on the results reported in [124], Figure 4 for our system frequency range. To ensure a fair comparison, the remaining portion of the combined laser linewidth is implemented on the receiver side. This approach allows a straightforward pilot-based phase noise estimator to extract the necessary receiver laser phase noise for the phase-dependent compensator.

Following [121], we consider a practical raw forward error correction (FEC) BER as an indication of the system performance threshold. In our simulations, we consider a raw BER of 0.007. As shown in [135], a FEC with 9% overhead can effectively work with this raw BER. The following results are based on the central channel on the x-polarization.

4.5.2 Simulation Results

Impact of Channel Baud Rate

Based on (4.1), EEPN noise power is proportional to the baud rate, $1/T$. Therefore, EEPN is a primary factor limiting the baud rate in fiber systems. In this simulation, we evaluate the performance of the fiber system under various channel baud rates, both with and without the EEPN compensator. The simulation results are illustrated in Figure 4.9. The impact of EEPN with respect to the channel baud rate should be studied in three regions: low baud rates where EEPN is negligible, intermediate baud rates where EEPN is one of the main system limiting factors, and higher baud rates where nonlinearity dominates.

For low baud rates (less than 200 GBaud), all methods, including uncompensated ideal ZFP, can provide acceptable BER performance. This is because at low channel baud rates, nonlinearity and EEPN are negligible.

At practical baud rates (in this setup, 200 GBaud to 400 GBaud), EEPN is one of the main system limiting factors. Therefore, EEPN uncompensated methods or partially compensated methods like BPS lose their performance. As observed, our proposed methods can significantly increase system performance in this region. Furthermore, based on this simulation, we observe slightly lower BER performance of the pilot-based EEPN compensator compared to the phase-based compensator. This can be explained by the fact that estimating the EEPN FIR filter coefficients in a noisy situation is also a source of error.

For high baud rates (400 GBaud and beyond), nonlinearity dominates. This is because higher channel bandwidth necessitates launching higher power, which in turn increases the nonlinearity effects. Even though our method still outperforms existing methods, the considered raw BER is not satisfied in this region.

The dashed line demonstrates the considered raw BER threshold. As observed, BPS can work with the data rate up to 270 GBaud, while our compensators can operate up to 400 GBaud. In other words, using our proposed compensators, one can increase the system baud rate by 48% for this setup.

In short, this simulation confirms that using the proposed EEPN compensators, the fiber system can support higher data rates compared to EEPN uncompensated receivers.

Impact of Laser Linewidth

Laser phase noise not only causes EEPN in a suboptimal receiver but also can severely impact the phase estimation performance. In this simulation, we in-

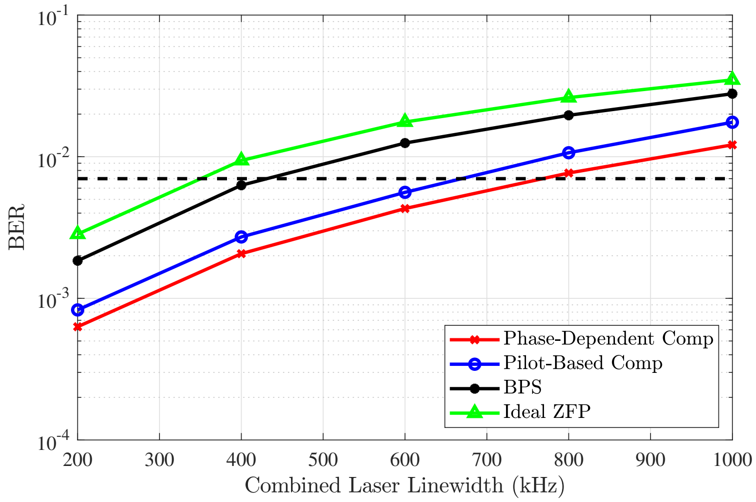


Figure 4.10: BER performance of the compensators for different combined laser linewidths.

vestigate the performance of different methods under various laser linewidths. Figure 4.10 demonstrates the results. In this figure, the BER performance of the fiber system for different laser linewidths is presented. As observed, the uncompensated systems, including ideal ZFP and BPS, lose their BER performance immediately with increasing laser linewidth. This is because, based on (4.1), the EEPN power is proportional to the laser linewidth, and hence increasing the laser linewidth causes the EEPN to degrade the system performance.

On the other hand, EEPN compensators can tolerate significantly higher laser linewidths by canceling the EEPN effects. However, considering the low performance of the carrier phase estimation for high laser linewidths, eventually, our compensators surpass the raw BER requirement that we considered. This is evident at a laser linewidth of $\Delta\nu = 750$ kHz.

In short, using our compensators, we can tolerate a much higher laser linewidth compared to the EEPN uncompensated systems. In this simulation, we observe a 85% increase in acceptable laser linewidth.

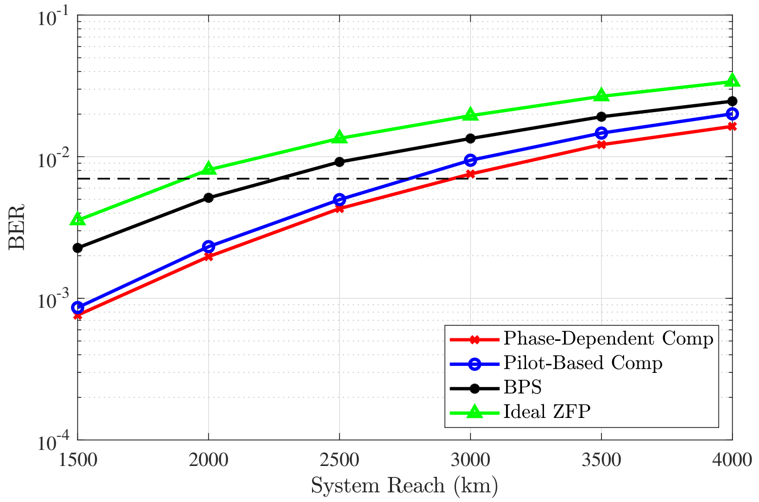


Figure 4.11: BER performance for different system reaches

System Reach

As predicted by (4.1), increasing the fiber length increases the EEPN noise power. Therefore, the minimum achievable BER increases for long fibers compared to short fibers. In this simulation, we investigate the impact of fiber length on the fiber system. Figure 4.11 depicts the results. As observed, all studied methods provide acceptable BER performance for fiber lengths up to 2000 km. We can conclude that for these lengths and the considered baud rate of 200 GBaud, the EEPN effect is not significant. However, for system reaches higher than 2000 km, EEPN limits the minimum achievable BER. The partially EEPN compensated method, BPS, increases the system reach to 2300 km, while our compensator can tolerate a system reach of 2900 km. In other words, in this system setup, using our compensators, one can increase the system reach by 26%.

It is important to note that the effects of nonlinearity also grow with increasing system reach. As observed, for high system reach, the performance gain achieved by our EEPN compensators decreases because fiber nonlinearity dominates at these reaches.

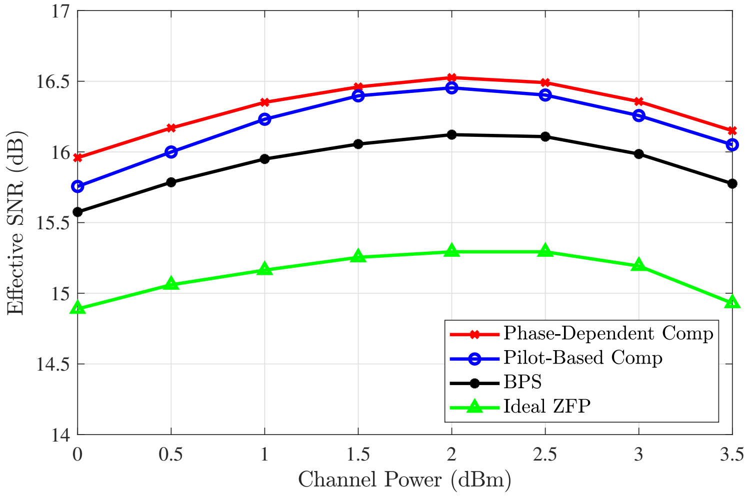


Figure 4.12: Effective SNR for different launch power.

Launch Power

Channel power is the primary factor in exciting fiber nonlinearity. While channel power does not directly affect dispersion and EEPN, compensating for EEPN can decrease the optimal launch power, enabling us to reduce nonlinearity. This is supported by [71], which indicates that the optimal launch power occurs at:

$$P_{\text{opt}} = \sqrt[3]{\frac{P_{\text{lin}}}{2\eta}}, \tag{4.37}$$

where η is a constant independent of power, and P_{lin} is the power of the linear noises (power-independent). Therefore, by reducing the EEPN noise power, one of the sources of power-independent noise, we can lower the optimal launch power and increase the effective SNR. To assess the impact of EEPN compensation on the optimal launch power, we conduct this simulation in which the effective SNR is determined for different launch powers. Figure 4.12 presents the results. We explain this simulation result with respect to the maximum achievable effective SNR and the optimal channel power.

As observed, using our compensators, we can achieve a maximum SNR of 16.5 dB. This is while using the BPS and Ideal ZFP methods, the maximum achievable SNRs are 16.5 dB and 15.4 dB, respectively. The 0.4 dB gain in effective SNR comparing our compensators and BPS method can be attributed to the higher ability of our method to compensate for EEPN. Comparing the results of ideal ZFP and our methods, we can conclude that EEPN causes a 1.1 dB reduction in maximum effective SNR in this system setup.

On the other hand, observing the optimal launch power, we realize that by compensating for the EEPN, our method achieves the maximum SNR at a channel power of 2 dBm, while the ideal ZFP has an optimal launch power of 2.5 dBm. This means compensating for EEPN causes the optimal launch power to occur at lower channel powers compared to uncompensated systems.

It is also worth mentioning that the limited effective SNR reported in this simulation is due to the presence of nonlinearity. In an ideal case where fiber nonlinearity is compensated for and perfect knowledge of the laser phase noises is available, we expect the effective SNR to grow linearly with the channel power when our proposed compensator is used. This is in contrast to other existing methods, which provide a limited SNR due to the EEPN effect.

In short, we can conclude that using EEPN compensators not only increases the maximum achievable SNR but also enables us to achieve this SNR at lower launch powers compared to uncompensated receivers. This, in turn, allows us to work with lower nonlinearity in the system.

Comparison to DRS

One of the recent CPE techniques proposed for subcarrier systems is DRS [122]. DRS is designed to improve compensation for receiver phase noise through the use of the walk-off phenomenon among different frequency subchannels. Specifically, DRS receives pilots at two distinct frequency channels that experience a time delay due to the walk-off effect. This time delay causes the pilots to be mixed with receiver phase noise at different time instances. By comparing these pilots, DRS can extract the receiver phase noise. Although DRS can only extract the receiver phase noise after dispersion compensation, it has been shown to better tolerate EEPN.

The main objective of this simulation is to compare our method with the DRS method in handling EEPN. As noted by the authors, the DRS method necessitates a minimum SNR of 18 dB for a subcarrier system with 8 subchannels

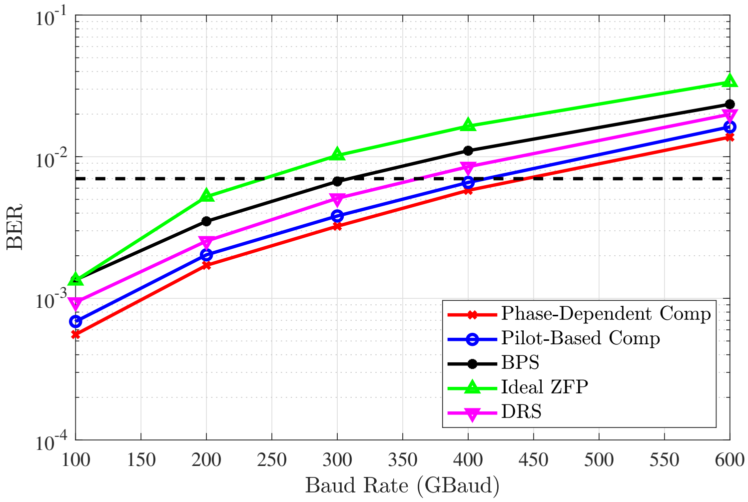


Figure 4.13: BER performance of the system with 500 km system reach for different baud rates.

[122]. However, as depicted in Figure 4.12, the considered long-haul system fails to meet the minimum effective SNR across a broad range. To tackle this issue, we reduce the system reach to 500 km and employ a 64-QAM constellation. The remaining system properties remain consistent with those discussed in Subsection 4.5.1. The DRS method is implemented for a subcarrier system with 8 subchannels. In this simulation, we examine the BER performance for different baud rates, as illustrated in Figure 4.13. As observed, for data rates lower than 300 GBaud, ZFP and BPS fail to achieve the assumed raw BER. On the other hand, DRS increases the baud rate with acceptable BER to 350 GBaud. This is while our compensators satisfy the raw BER criteria up to 420 GBaud. Although increasing the number of subchannels can enhance DRS accuracy, it results in a higher computational burden. In summary, DRS and our compensators can reduce the EEPN effect, while our method can work for a wide range of SNRs and provide lower BER. As illustrated by the dashed line in Figure 4.13, our compensators can enhance the data rate for the considered raw BER by 20%.

Runtime

Due to the very high data rates in fiber systems, the complexity of DSP methods is to be considered. In this simulation, we compare the complexity of different compensators. For maximum fairness in our comparison, the complexity of the prerequisite DSP blocks for each technique is also considered. In particular, the complexity of the carrier phase recovery algorithm needed along with each compensator is also taken into account. Table 4.2 provides the results. These results demonstrate that our EEPN compensators offer complexity comparable to existing methods, ensuring the practical implementability of these compensators. Additionally, based on this simulation, we observe that the phase-based compensator has lower time complexity compared to the pilot-based EEPN compensator. This was expected based on our analysis in Subsection 4.4.4, where we concluded that the pilot-based complexity is determined by the update rule with more complexity required.

Method	M_d	N	Mr	Time (ns)
ZFP	-	-	-	36
BPS	-	-	-	337
DRS	-	-	-	433
Phase-Dependent Comp	5000	10	-	520
	10000	14	-	528
	15000	18	-	534
Pilot-Based Comp	5000	10	125	588
	10000	14	250	626
	15000	18	375	682

Table 4.2: The time complexity of various methods on a Core i7 CPU operating at 2.26 GHz.

4.6 Conclusion

In this chapter, we first derived the output expression of an optimal EEPN-free receiver. Subsequently, we demonstrated that this EEPN-free output could be expressed based on the dispersion-compensated signal in a suboptimal receiver, serving as a basis for EEPN compensation. Building upon this new formulation, we proposed an EEPN compensator under the assumption of the availability of the receiver phase noise. We illustrated that this compensator can be effectively implemented using a time-varying FIR filter. Following this, we introduced a more practical compensator that determines the FIR filter coefficients based on a pre-known pilot, thus relaxing the requirement for the receiver phase noise. We emphasized the practical aspects of the proposed compensators by explaining the suitable design parameters for different fiber systems and conducting a complexity analysis. Our simulations corroborated the superior performance of our compensator compared to other existing EEPN-tolerant approaches. Notably, our compensators exhibited the ability to enhance system reach, maximum achievable SNR, and system data rate in comparison with existing methods for addressing EEPN.

Chapter 5

Receiver Laser Phase Noise Estimation with Application to EEPN Control

5.1 Introduction

As discussed in the previous chapter, compensating for receiver laser phase noise after dispersion compensation gives rise to EEPN, which limits system performance [118]. Although compensation for laser phase noise and dispersion ideally should occur in the reverse order of their occurrence [7], dispersion is compensated for first due to practical challenges facing CPE. Practical CPEs utilize a sequence of known symbols, referred to as pilots, to estimate and mitigate the phase noise in the received symbol sequence. Extracting the pilots requires compensating for dispersion first, leading to a suboptimal sequence of compensations and the occurrence of EEPN [8]. Depending on the bandwidth and system reach of the fiber, EEPN can severely affect system performance [9].

CPE, as the main bottleneck causing EEPN, can be designed to control the EEPN effects and hence enhance the system capacity. In the following, we have a closer look into the existing CPE techniques developed for fiber optic systems.

The fundamentals of pilot-based carrier phase estimation have been explored in various studies [68], [126], [136]. These methodologies, commonly referred to as ZFP, aim to identify the phase mismatch between the transmitted pilot and the corresponding received symbol. Despite its straightforward implementation, the susceptibility to noise and the limited accuracy of this approach restrict its applications.

Incorporating a BPS after the ZFP has been shown to enhance phase estimation accuracy and performance [137]. BPS assumes the presence of a slowly-varying phase noise in the received symbols; thus, it employs a symbol detector to estimate the transmitted symbols and extracts the phase noise by finding the phase mismatch. In a recent study [117], the authors illustrate that using BPS can partially mitigate EEPN. Although this approach does not require a known pilot, it necessitates the symbols to be estimated by the receiver before the CPE. Symbol estimation, in turn, requires compensation for dispersion beforehand leading to EEPN.

As fiber technologies evolve, CPE needs to evolve as well. Digital subcarrier fiber systems are among the recent developments in fiber technology, hence CPE for digital subcarrier is needed [138], [139]. Several studies have delved into the problem of carrier phase recovery in digital subcarrier systems [140]–[142]. However, the walk-off phenomenon among the subchannels poses a challenge for pilot-based estimation of phase noise. In [123], it has been demonstrated that the pilot position in the subchannels can significantly affect CPE performance. The authors propose a pilot structure and a specialized CPE tailored to the unique properties of digital subchannel systems. Additionally, in [122], a specific CPE for digital subcarrier systems is proposed by considering the time delay between the two subchannels. This approach referred to as DRS, is shown to have EEPN control capabilities. However, to fully leverage this method, a large number of subchannels are required, which can substantially increase the complexity of both the receiver and transmitter.

Despite advancements in CPE technology, current techniques still suffer from the following shortcomings, which prevent the optimal sequence of compensators. Existing CPEs:

I. are implemented after the dispersion compensator as they need the extracted symbols;

II. estimate the sum of the lasers' phase noise, making separate compensation of receiver and transmitter phase noise —needed for an optimal receiver— impossible.

In this work, we propose a new CPE technique to overcome these challenges, allowing the implementation of a receiver with the optimal order of compensating blocks and, therefore, preventing the generation of EEPN.

We design our CPE to operate with the dispersed signal without the need for dispersion compensation. Dispersion has two important impacts on the signal: group velocity and GVD. Group velocity causes signals with different carrier frequencies to experience different delays, while GVD broadens the pulse, preventing symbol extraction. In our design, we exploit the fact that the impact of GVD on both negative and positive excess bandwidths[1] is the same. Therefore, in a differential manner, we can use these excess bandwidths to remove the GVD impact and extract the receiver phase noise without the need for first performing dispersion compensation.

To extract the receiver phase noise differentially, the frequency components of the data-modulated signal at negative and positive frequencies should be the same. While this may not hold true for all signal frequency components, we demonstrate that the excess bandwidths at negative and positive frequencies satisfy this condition. Furthermore, leveraging the group velocity property, these bandwidths experience different but known delays. Consequently, after appropriate time adjustment and filtering, the excess bandwidths at negative and positive frequencies contain synchronized data and the same transmitter phase noise, while the receiver laser phase noise affecting them originates from two distinct time slots. Therefore, by extracting the phase difference between these two signal portions, we can discern variations in receiver laser phase noise over time. Finally, we extract the receiver laser phase noise through the cumulative sum of the extracted time variations, separately from the transmitter

---

[1] By excess bandwidth, we mean the additional bandwidth beyond the symbol rate that is allocated to a pulse shape to aid in its implementation. This excess bandwidth is governed by a parameter known as the roll-off factor, denoted by β. For a symbol period of T, the negative and positive excess bandwidths for a pulse are expressed as $[-\frac{1+\beta}{2T}, -\frac{1-\beta}{2T}]$ and $[\frac{1-\beta}{2T}, \frac{1+\beta}{2T}]$, respectively.

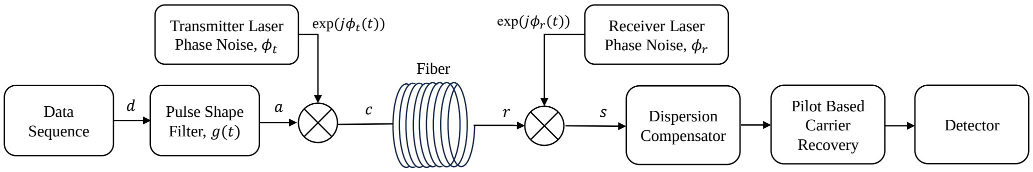


Figure 5.1: A typical fiber system structure with respect to transmitter and receiver laser phase noise. This structure is suboptimal regarding the order of receiver and dispersion compensation causing EEPN.

phase noise.

Expanding on this concept, we present a DSP- ready implementation of our CPE technique. Furthermore, we illustrate the configuration of our proposed receiver with the optimal order of compensation blocks, hence eliminating EEPN.

Using a simulation study, we investigate the performance of our CPE under different scenarios. Our main objectives in the simulation study are twofold: firstly, to examine the factors influencing our CPE's performance, including channel bandwidth and channel quality; and secondly, to evaluate the performance gains that our proposed receiver can offer due to the absence of EEPN. Our results confirm performance improvements compared to existing methods in various aspects, including system reach, effective SNR, and BER. We further demonstrate that our CPE can accurately estimate phase even within a much wider SNR range compared to existing methods, thus enabling its application in a broader range of scenarios.

This chapter is structured as follows: In Section 5.2, we elucidate the system model. Section 5.3 introduces the main idea underlying our proposed CPE, followed by the presentation of a practical implementation. Simulation outcomes are detailed in Section 5.4, and lastly, Section 5.5 offers concluding remarks for this study.

5.2 System Model

In this chapter, for the ease of notation and discussions, we consider the minimal system required for estimating receiver laser phase noise based on the pulse excess bandwidth as demonstrated by Figure 5.1. However, extending the same idea to more advanced systems is straightforward.

On the transmitter side, the data sequence is modulated with a bandlimited and zero-phase pulse shape, $g(t)$. This pulse shape satisfies the Nyquist criterion to avoid introducing ISI [17]. For simplicity, we consider the pulse shape filter to be implemented entirely on the transmitter side only in our formulation. However, splitting the pulse shape between the transmitter and a matched-filter receiver, as we did in the implementation and simulation, does not interfere with our results.

The resulting baseband signal is frequency up-converted using the transmitter laser. The lasers introduce phase noise to the system. We refer to the transmitter and receiver phase noises as $\phi_t(t)$ and $\phi_r(t)$, respectively. The laser phase noises, $\phi_t(t)$ and $\phi_r(t)$, can be modeled by a Wiener process as [143]:

$$\phi(t) = \int_{-\infty}^{t} n(\nu) d\nu, \tag{5.1}$$

where n is a Gaussian random process with zero mean and a variance of $\Delta\nu/2\pi$, and $\Delta\nu$ is the laser linewidth.

In the next step, the signal is launched into the fiber. We consider a SSMF in which the signal propagation is governed by the nonlinear Schrödinger equation described in (1.16) and repeated here as [88]:

$$\frac{\delta a}{\delta z} + j\frac{\beta_2}{2}\frac{\delta^2 a}{\delta t^2} + \frac{\alpha}{2} a = j\gamma |a|^2 a. \tag{5.2}$$

In this equation, a symbolizes the signal envelope traveling in the z direction. Additionally, β_2, α, and γ represent GVD, attenuation, and nonlinearity parameters, respectively.

At the receiver, the receiver laser down-converts the signal, imposing the receiver laser phase noise, $\phi_r(t)$. It is important to note that phase noise and fiber dispersion do not have a commutative property; hence, they should be compensated for in the opposite order of their occurrence. However, pilot-based CPEs require the dispersion compensator to be applied first, causing EEPN. In [7], it is demonstrated that the power of EEPN noise can be expressed as:

$$\sigma^2_{\mathrm{EEPN}} = \frac{\pi^2 \beta_2 L \Delta\nu}{T}, \tag{5.3}$$

where T denotes the symbol period and L represents the fiber length. Based on (5.3), we can deduce that for high data rates and long system reaches, EEPN grows significantly, thereby degrading system performance.

In this chapter, our goal is to propose a novel pilot-independent CPE capable of extracting receiver phase noise before dispersion compensation. This CPE enables us to compensate for phase noises and dispersion in the optimal order, avoiding EEPN.

In the subsequent sections, lowercase letters are employed to denote time-domain signals, and their corresponding Fourier transforms are specified using uppercase letters. Bold lowercase letters are utilized for vectors, while bold uppercase letters indicate matrices.

5.3 Carrier Phase Recovery

In this section, we first present the main idea behind our CPE. Subsequently, we utilize the derived expression for phase estimation to propose a DSP-ready implementation of our CPE. Finally, we present the optimal structure based on our CPE that should be used instead of the one depicted in Figure 5.1 to avoid EEPN.

5.3.1 Design Principle

In this subsection, our main objective is to extract the temporal variation of receiver laser phase noise using the pulse excess bandwidths in negative and positive frequencies. The cumulative sum of this temporal variation will then estimate the receiver phase noise.

As a starting step, we derive the expression of the modulated signal at the transmitter, emphasizing the redundancy in the excess bandwidth in negative and positive frequencies. Let us consider the discrete-time symbol sequence $\boldsymbol{q}$ over complex numbers to be transmitted at a rate of $1/T$. We can represent this symbol sequence in a continuous time domain using the Dirac delta function as:

$$d(t) = \sum_{n=-\infty}^{+\infty} q_n \delta(t - nT). \tag{5.4}$$

Using the Fourier transform, the continuous frequency representation of $d(t)$ can be expressed as follows [144]:

$$D(f) = \sum_{n=-\infty}^{+\infty} q_n \exp\left(-j2\pi f nT\right). \tag{5.5}$$

It is crucial to note that $D(f)$ is a periodic function with a period of $f_r = 1/T$. An example of the frequency representation of the data sequence is depicted in Figure 5.2a. As observed, due to periodicity, the data contains the same information over the pulse's excess bandwidth on both negative and positive frequencies.

To transform the data sequence into a continuous bandlimited signal, we utilize a pulse shape represented by $G(f)$. To meet practical considerations and prevent introducing ISI, the pulse shape must adhere to the Nyquist criterion [17]. Therefore, it should conform to the following model:

$$G(f) = \begin{cases} 1, & |f| < \frac{1-\beta}{2T} \\ 1 - X\left(\frac{1}{2T} - f\right), & \frac{1-\beta}{2T} \leq |f| < \frac{1}{2T} \\ X\left(f - \frac{1}{2T}\right), & \frac{1}{2T} \leq |f| < \frac{1+\beta}{2T} \\ 0, & \text{otherwise,} \end{cases} \tag{5.6}$$

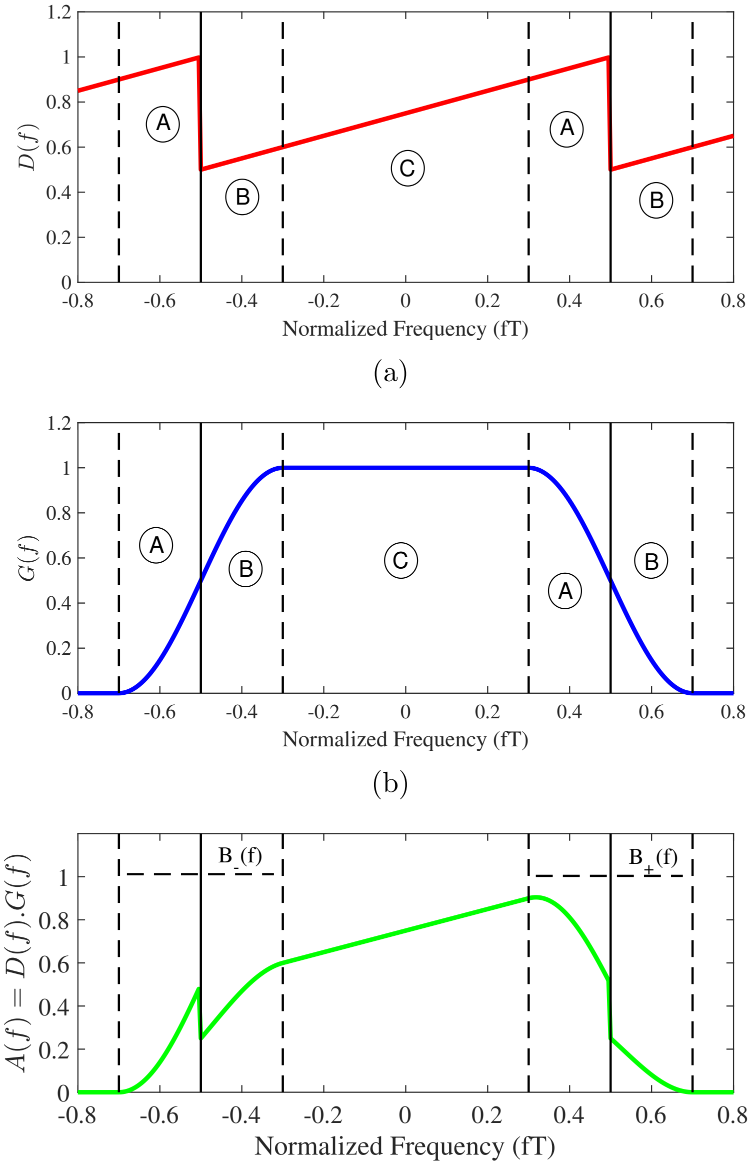
where β is the roll-off factor representing the portion of the bandwidth dedicated to the pulse edge, and $0 \leq X(f) \leq 1$ is a real-valued arbitrary function. For example, in the case of the well-known raised cosine pulse, $X(f) = \frac{1}{2}\cos\left(\frac{\pi f T}{\beta}\right)$. Figure 5.2b represents a raised cosine pulse with $\beta = 0.4$.

Having the data and pulse shape, the modulated signal on the transmitter side can be obtained as:

$$A(f) = G(f) \cdot D(f), \tag{5.7}$$

as depicted in Figure 5.2c for our example. As seen in this figure, we define the following two signals representing the portions of the modulated signal in the negative and positive excess bandwidths as:

$$B_{-}(f) = \begin{cases} G\left(f - \frac{1}{2T}\right) \cdot D\left(f - \frac{1}{2T}\right), & |f| < \frac{\beta}{2T} \\ 0, & \text{else,} \end{cases} \tag{5.8}$$



(a)

(b)

(c)

Figure 5.2: Frequency representation of (a) the data sequence, (b) the pulse shape, and (c) the modulated signal on the pulse shape. In Subfigure (a), circles A and B represent the data spectrum sections with the same information. These sections are then multiplied by their corresponding pulse shape sections in Subfigure (b).

and,

$$B_{+}(f) = \begin{cases} G\left(f + \frac{1}{2T}\right) \cdot D\left(f + \frac{1}{2T}\right), & |f| < \frac{\beta}{2T} \\ 0, & \text{else,} \end{cases} \tag{5.9}$$

respectively. It is important to note that B_{-} and B_{+} have identical phase information. This is because, based on their definitions, B_{-} and B_{+} result from the zero-phase filter $G(f)$ acting on the same data $D(f - \frac{1}{2T}) = D(f + \frac{1}{2T})$. The equality stems from the periodicity of the data in the frequency domain with a period of $\frac{1}{T}$, as expressed in (5.5).

Furthermore, based on $B_{+}(f)$ and $B_{-}(f)$, we can rewrite the modulated signal, expressed in (5.7), as:

$$A(f) = B_{-}\left(f + \frac{1}{2T}\right) + D(f) \cdot \operatorname{rect}\left(\frac{fT}{1-\beta}\right) + B_{+}\left(f - \frac{1}{2T}\right), \tag{5.10}$$

where $\operatorname{rect}(f)$ is the rectangular function defined as:

$$\operatorname{rect}(f) = \begin{cases} 1, & |f| < 0.5 \\ 0.5, & |f| = 0.5 \\ 0, & \text{else.} \end{cases} \tag{5.11}$$

Writing (5.10) in the time domain, we get:

$$a(t) = b_{-}(t) \cdot \exp\left(\frac{-j\pi t}{T}\right) + d(t) * \operatorname{sinc}\left(\frac{t(1-\beta)}{T}\right) + b_{+}(t) \cdot \exp\left(\frac{j\pi t}{T}\right). \tag{5.12}$$

To this point, we represented the modulated signal based on the baseband versions of the excess bandwidth and redundancy in the transmitted signal, as expressed by (5.12). In the next step, our goal is to investigate the effects of the laser phase noise and dispersive fiber channel on the transmitted signal. We then derive the expression of the received signal before dispersion compensation.

Based on Figure 5.1, the modulated signal experiences the transmitter laser phase noise before launching into the fiber. Considering this phase noise, we can write the transmitted signal as:

$$c(t) = a(t) \cdot \exp\left(j\phi_t(t)\right). \tag{5.13}$$

Afterward, $c(t)$ is launched into the fiber, experiencing fiber dispersion. Therefore, the received signal can be expressed as:

$$r(t) = c(t) * u(t), \tag{5.14}$$

where $u(t)$ is the time impulse response of the dispersion, defined as:

$$u(t) = \int_{-\frac{1+\beta}{2T}}^{\frac{1+\beta}{2T}} \exp(2j\pi^2\beta_2 f^2 L + 2j\pi f t)df. \tag{5.15}$$

As depicted in Figure 5.1, at the receiver, the received signal encounters the receiver laser phase noise. Therefore, we can represent the signal before dispersion compensation as:

$$s(t) = r(t) \cdot \exp\left(j\phi_r(t)\right). \tag{5.16}$$

To incorporate the data redundancy and account for the time delay due to dispersion, we substitute (5.12), (5.13), and (5.14) into (5.16), resulting in:

$$s(t) = s_-(t) + s_0(t) + s_+(t), \tag{5.17}$$

where

$$\begin{aligned} s_-(t) = &\left(\left(b_-(t) \cdot \exp(j\phi_t(t)) \cdot \exp\left(-\frac{j\pi t}{T}\right)\right) * u(t)\right) \\ &\cdot \exp(j\phi_r(t)), \end{aligned} \tag{5.18}$$

and,

$$\begin{aligned} s_0(t) = &\left(\left(\left(d(t) * \operatorname{sinc}\left(\frac{t(1-\beta)}{T}\right)\right) \cdot \exp\left(j\phi_t(t)\right)\right) * u(t)\right) \\ &\cdot \exp\left(j\phi_r(t)\right), \end{aligned} \tag{5.19}$$

and,

$$\begin{aligned} s_+(t) = &\left(\left(b_+(t) \cdot \exp(j\phi_t(t)) \cdot \exp\left(\frac{j\pi t}{T}\right)\right) * u(t)\right) \\ &\cdot \exp(j\phi_r(t)). \end{aligned} \tag{5.20}$$

Here, we claim that by using proper bandpass filtering, we can extract $s_-(t)$ and $s_+(t)$. This is because the terms $\exp\left(j\phi_t(t)\right)$ and $\exp\left(j\phi_r(t)\right)$ have negligible bandwidth compared to that of the channel. Consequently, $s_-(t)$, $s_0(t)$, and $s_+(t)$ do not overlap in the frequency domain, thus allowing for their separation.

The signal in the excess bandwidth, $s_-(t)$ and $s_+(t)$, can be considered as two narrow-band signals. To show the impact of the dispersion operator on a narrow-band signal, we introduce the following lemma:

Lemma 1. *Consider the input of the dispersion operator with the impulse response $u(t)$ as $x(t) \cdot \exp(j2\pi f_0 t)$. The output $y(t)$ can be expressed as:*

$$y(t) = (x(t) * u(t) * \delta(t - 2\pi\beta_2 f_0 L)) \cdot \exp(j2\pi f_0(t - \pi\beta_2 f_0 L)). \quad (5.21)$$

Proof. The output of the dispersion operator is $(x(t) \cdot \exp(j2\pi f_0 t)) * u(t)$. It is easy to show that the Fourier transform of (5.21) and the dispersion operator output are the same. □

Based on this lemma, we can infer that the narrow-band signals in the excess bandwidths experience a time delay corresponding to their central frequency. This can be considered as a special case of the well-known walk-off phenomenon.

To formulate the impact of the dispersive fiber on the narrow-band signals in the excess bandwidth, we apply Lemma 1 to (5.18) and (5.20), to obtain:

$$\begin{aligned} s_-(t) = &\left(\left(b_-(t) \cdot \exp(j\phi_t(t))\right) * u(t) * \delta\left(t - \frac{\pi\beta_2 L}{T}\right)\right) \\ &\cdot \exp(j\phi_r(t)) \cdot \exp\left(-j\frac{\pi}{T}\left(t + \frac{\pi\beta_2 L}{2T}\right)\right), \end{aligned} \quad (5.22)$$

and

$$\begin{aligned} s_+(t) = &\left(\left(b_+(t) \cdot \exp(j\phi_t(t))\right) * u(t) * \delta\left(t + \frac{\pi\beta_2 L}{T}\right)\right) \\ &\cdot \exp(j\phi_r(t)) \cdot \exp\left(j\frac{\pi}{T}\left(t - \frac{\pi\beta_2 L}{2T}\right)\right). \end{aligned} \quad (5.23)$$

As seen, there is a time delay and phase mismatch between (5.22) and (5.23). In order to align the data embedded in the negative and positive excess bandwidth, we compensate for this time and phase difference on (5.22) and (5.23), using the following filters:

$$U_-^i(f) = \begin{cases} \exp\left(\frac{2j\pi^2\beta_2 L}{T}f - \frac{j\pi^2\beta_2 L}{2T^2}\right), & \frac{1-\beta}{2T} < f < \frac{1+\beta}{2T} \\ 0, & \text{else} \end{cases} \quad (5.24)$$

and

$$U_+^i(f) = \begin{cases} \exp\left(-\frac{2j\pi^2\beta_2 L}{T}f + \frac{j\pi^2\beta_2 L}{2T^2}\right), & -\frac{1+\beta}{2T} < f < -\frac{1-\beta}{2T} \\ 0, & \text{else,} \end{cases} \quad (5.25)$$

respectively. Applying these filters and down-converting the results to the baseband, $s_-(t)$ and $s_+(t)$ become:

$$v_-(t) = \left[\Big(b_-(t) \cdot \exp(j\phi_t(t))\Big) * u(t)\right] \cdot \exp\left(j\phi_r\left(t + \frac{\pi\beta_2 L}{T}\right)\right), \quad (5.26)$$

and

$$v_+(t) = \left[\Big(b_+(t) \cdot \exp(j\phi_t(t))\Big) * u(t)\right] \cdot \exp\left(j\phi_r\left(t - \frac{\pi\beta_2 L}{T}\right)\right), \quad (5.27)$$

respectively.

In the final step, we demonstrate that based on the phase difference between (5.26) and (5.27), we can extract the receiver phase noise. To this end, we utilize the fact that $b_-(t)$ and $b_+(t)$ have the same phase information. Additionally, noting that the identical dispersion operator is acting on both $b_-(t) \cdot \exp(j\phi_t(t))$ and $b_+(t) \cdot \exp(j\phi_t(t))$, we can conclude that the expressions in brackets in (5.26) and (5.27) share the same phase. Therefore, the phase difference between $v_-(t)$ and $v_+(t)$ is:

$$\begin{aligned} \angle v_+(t) - \angle v_-(t) &= \angle v_+(t) \cdot v_-^*(t) \\ &= \phi_r\left(t + \frac{\pi\beta_2 L}{T}\right) - \phi_r\left(t - \frac{\pi\beta_2 L}{T}\right) \end{aligned} \quad (5.28)$$

As seen on the right-hand side of (5.28), this expression represents the variation of the receiver phase noise over the time interval:

$$T_d = \frac{2\pi\beta_2 L}{T}. \quad (5.29)$$

It is important to note that (5.29) defines the time resolution over which the receiver phase noise should be extracted. According to this equation, the time resolution of the estimated receiver laser phase noise decreases for high data rates.

To enhance the estimation robustness against system noise, we can exploit the characteristic of laser phase noise, which behaves as a low-pass signal with a bandwidth proportional to the laser linewidth, $\Delta\nu$. Therefore, to estimate the temporal variation of the receiver phase noise, $\Delta\phi_r'$, from (5.28), we can employ a low-pass filter as follows:

$$\Delta\phi_r'[i] = \arctan\Big(\mathrm{LPF}\{v_+[i] \cdot v_-^*[i]\}\Big), \quad (5.30)$$

where $v[i] = v\left(iT_d\right)$, LPF represents a low-pass filter with a bandwidth equal to or greater than the receiver laser phase noise. The $\arctan(\cdot)$ function is used to extract the angle of the term $v_+[i] \cdot v_-^*[i]$. Finally, by employing the following cumulative sum, we can extract the receiver phase noise:

$$\phi_r'[i] = \phi_r'[i-1] + \Delta\phi_r'[i]. \tag{5.31}$$

For the initial receiver phase noise, we can consider any value for $\phi_r'[0]$ since it only causes a constant rotation in the received symbol. This phase offset does not cause EEPN and can be easily compensated for.

In summary, this subsection demonstrates that the excess bandwidths in the negative and positive frequencies can be treated as two narrowband signals, each modulated with the same data as expressed in (5.18) and (5.20). Due to their frequency difference, they undergo different time delays because of the dispersion as shown by (5.22) and (5.23). Compensating for this time difference results in two identical and synchronized copies of the transmitted data corrupted with receiver phase noise at two different time slots as described in (5.26) and (5.27). By extracting the phase difference between these synchronized signals and utilizing a cumulative sum, expressed in (5.31), we can extract the receiver laser phase noise, albeit subject to an ambiguity in the initial value of the phase noise.

5.3.2 Implementation of Carrier Phase Recovery

In this subsection, building upon our derivations in the previous subsection, we present a practical implementation of CPE applicable to DSP-based receivers.

To begin with, we explain the steps needed for receiver phase noise extraction based on the signal's excess bandwidth, as presented in Algorithm 1.

The primary input to this algorithm is the received signal before dispersion compensation, represented as s in Figure 5.1. Subsequently, the excess bandwidth on both negative and positive frequencies is extracted through bandpass filtering, as outlined in steps 1 and 2. Following this, the dispersion-induced time delay is compensated using filters u_-^i and u_+^i, with the results downcon-

Algorithm 1 Receiver Laser Phase Noise Estimator

```
Input: signal $s$, Dispersion coefficient $\beta_2$, Fiber length $L$, Symbol period $T$.
Output: Receiver Phase $\phi_r$.
1: extract $s_-(t)$ by bandpass filter of $s$.
2: extract $s_+(t)$ by bandpass filter of $s$.
3: find $v_-$ by $(s_- * u_-^i) \cdot \exp(+j\pi t/T)$.
4: find $v_+$ by $(s_+ * u_+^i) \cdot \exp(-j\pi t/T)$.
5: find $\Delta\phi'_r[i]$ based on (5.30).
6: find $\phi_r[i] = \phi'_r[i-1] + \Delta\phi'_r[i]$ as expressed in (5.31).
7: return $\phi_r$.
```

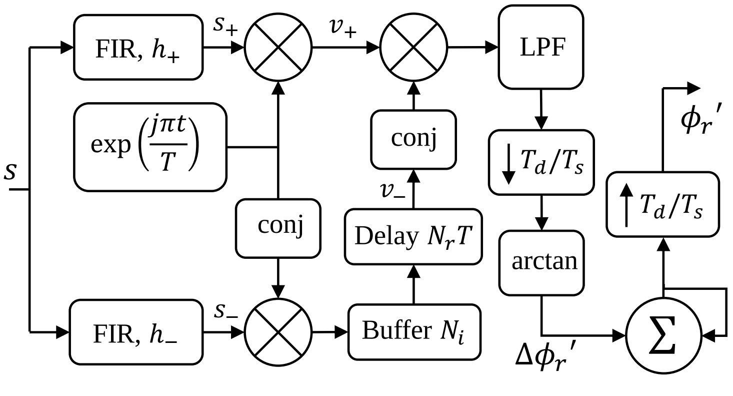


Figure 5.3: FIR-based implementation of the proposed CPE.

verted to the baseband to generate signals v_- and v_+. This synchronization process occurs in steps 3 and 4. Then, the temporal receiver phase noise variation is extracted based on (5.30) in step 5. Finally, the receiver phase noise can be extracted through a cumulative sum, as executed in step 6.

Algorithm 1 can be implemented using DSP techniques as demonstrated by Figure 5.3. In this block diagram, FIR filters are responsible for extracting $s_-(t)$ and $s_+(t)$. These filters are bandpass with the pass bands of $[-\frac{1+\beta}{2T}, -\frac{1-\beta}{2T}]$ and $[\frac{1-\beta}{2T}, \frac{1+\beta}{2T}]$, respectively. To avoid any phase distortion affecting the estimated ϕ_r, we propose using FIR filters with linear phase response [145]. The tap coefficients for these FIR implementations can be:

$$h_-[k] = \frac{\beta T_s}{T}\text{sinc}\left(\frac{\beta k T_s}{T}\right) \cdot \exp\left(\frac{j2\pi k T_s}{T}\right), \tag{5.32}$$

and

$$h_+[k] = \frac{\beta T_s}{T}\text{sinc}\left(\frac{\beta k T_s}{T}\right) \cdot \exp\left(-\frac{j2\pi k T_s}{T}\right), \tag{5.33}$$

respectively. Here, $k = -M_s/2, 1 - M_s/2, \ldots, M_s/2$ and $M_s + 1$ is the number of the filter tabs.While we use ideal filtering with $\mathrm{sinc}(\cdot)$, any low-pass filtering that is more practical for implementation can be used instead.

After down-converting the FIR filter outputs to generate v_- and v_+, the time delay of T_d is applied. To preserve causality within the system, the entire delay is allocated to the v_- path. This adjustment of time delay also results in a $T_d/2$ delay in the estimated phase.

The FIR implementation of T_d can require a large FIR size depending on the accumulated dispersion value. To reduce complexity, we suggest dividing the required delay, T_d, into an integer multiple of the sampling period, T_s, and a residual non-integer delay. Consequently, we obtain:

$$T_d = N_i T_s + N_r T_s, \tag{5.34}$$

where $N_i = \lfloor \frac{T_d}{T_s} \rfloor$ represents the integer part and $N_r = \frac{T_d}{T_s} - N_i$ represents the fractional part of the delay. The time delay equivalent to $N_i T_s$ can be efficiently implemented using a straightforward buffer of size N_i. Conversely, the fractional delay of $N_r T_s$ can be implemented using an interpolation filter, with the coefficients being:

$$h_d = \frac{\beta T_s}{T} \mathrm{sinc}\left(\frac{\beta k T_s}{T} - \frac{\beta N_r T_s}{T} \right). \tag{5.35}$$

After the time delay block, the data and the phase noise transmitted through the negative and positive excess bandwidths are synchronized. Subsequently, we apply a conjugation operator and find the product of the two paths. By low pass filtering of the result and using the arctan function, we can extract the temporal variation of the receiver phase noise expressed in (5.30). Finally, a cumulative sum is implemented in accordance with (5.31). It is important to note that the downsampler by a factor of T_d/T_s is employed to implement (5.30) with a time resolution of T_d. Subsequently, we upsample the result of the cumulative sum to match the sampling rate of the estimated receiver phase noise with the signal sampling rate.

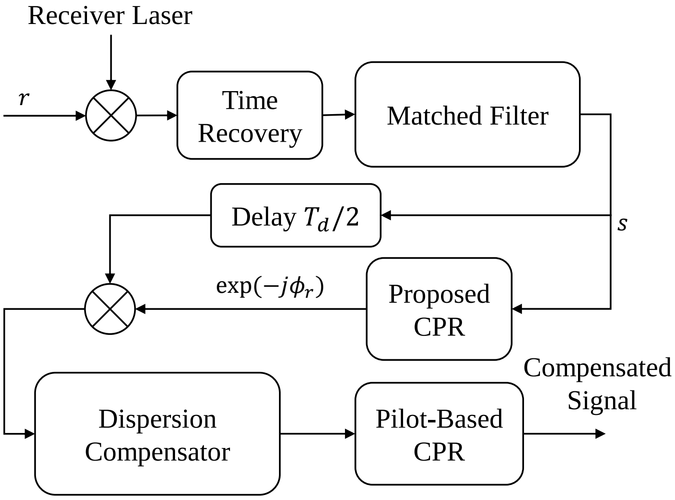


Figure 5.4: Receiver structure proposed based on the receiver laser phase noise estimation avoiding EEPN.

5.3.3 Optimal Receiver Structure

Our proposed CPE enables the implementation of the optimal receiver structure, as illustrated in Figure 5.4. The optimal arrangement of receiver laser phase noise, dispersion, and transmitter laser phase noise in this structure effectively prevents the generation of EEPN. Our CPE is implemented after the time recovery block and before the dispersion compensator. Compensating for timing errors before our proposed CPE is essential since they can cause a drift in the cumulative sum result. A suitable timing recovery block for this implementation could be the Gardner [60] or Godard [59] techniques. It is important to note that these techniques are also pilot-independent, making them readily implementable before dispersion compensation.

Before entering the dispersion compensator, receiver phase noise is estimated and compensated by our proposed CPE. As seen in Figure 5.4, a delay of $T_d/2$ is required to adjust the signal with the estimated receiver laser phase noise.

As mentioned earlier, our CPE can estimate the receiver phase noise up to a constant phase offset, which is due to the initial value for the cumulative sum. A constant receiver phase offset cannot cause EEPN since it has a commutative property with the dispersion convolution operator. In other words, this offset can be compensated jointly with the transmitter phase noise after dispersion compensation.

After the dispersion compensator, a pilot-based CPE is employed to mitigate transmitter phase noise and the offset of the receiver phase noise. Furthermore, this pilot CPE can correct any unwanted drift or error in receiver phase noise estimation.

5.4 Results and Discussion

To assess the performance of our proposed CPE, we conduct an extensive simulation study. For comparison, we included well-known CPE techniques such as ZFP [68], [126], [136] compensator, BPS [117], [137], and DRS[122], [123] phase estimator. While ZFP is optimal for phase recovery in nondispersive channels when perfect knowledge of laser phases is available, it causes EEPN in dispersive channels. Conversely, BPS and DRS are tailored to handle dispersive channels and mitigate the EEPN effect. Although DRS outperforms BPS, it requires higher signal quality for convergence.

In the subsequent simulation results, we denote our CPE technique as RCPE, short for receiver carrier phase estimator. Our simulation focuses on two primary objectives: determining the minimum requirements for acceptable accuracy of our CPE and evaluating the performance gain resulting from avoiding EEPN generation based on our CPE.

5.4.1 Simulation Setup

As our primary configuration, we consider a dual polarization fiber system of length 2000 km and with 5 WDM channels. This specific setup has been used in numerous recent studies to investigate the effects of EEPN on fiber systems [121]–[123].

On the transmitter side, a random bit sequence of size 2^{24} is generated and mapped to the 16-QAM constellation using Gray labeling. These symbols are then modulated onto a root-raised cosine pulse shape with a roll-off factor of $\beta = 0.1$ and a symbol duration of $T = 5 \times 10^{-12}$ seconds, unless otherwise specified. Therefore, the system baud rate is set to 200 GBaud. We consider a channel spacing of 5 GHz between adjacent channels. Subsequently, the different WDM channels are upconverted to their respective frequency slots and launched into the fiber by a laser with a linewidth of $\Delta\nu = 200$ kHz, which can be modeled as a Wiener process expressed in (5.1). For all simulations, the channel power is set to maximize the effective SNR.

In this study, we consider a typical single mode fiber with the following parameters: $\beta_2 = -20 \times 10^{-27}$ s^2/m, $\gamma = 0.013$ (mW)$^{-1}$, and $\alpha = 0.2$ dB/km, operating at a wavelength of 1550 nm. The fiber length is divided into 100 km spans, each equipped with an EDFA with a noise figure of 5 dB to compensate for the fiber loss. To simulate this fiber, we employ the SSFM to solve NLSE expressed in (5.2). The step size for SSFM is set to 50 m to ensure accurate results.

On the receiver side, we begin by downconverting the received light pulses using the receiver laser, which introduces receiver laser phase noise mirroring that of the transmitter side. For ZFP, BPS, and DRS, our process involves initial dispersion compensation, symbol extraction through matched filtering, and subsequent estimation and mitigation of combined phase noises. This receiver structure is depicted in Figure 5.1. In contrast, our proposed CPE initiates by estimating and compensating for receiver phase noise, followed by dispersion compensation. Eventually, the transmitter phase noise is addressed through pilot-based ZFP. All the reported results pertain to the central channel on the x-polarization.

While the aforementioned long-haul fiber system is suitable for EEPN studies, it fails to meet the minimum required SNR for the DRS technique. As outlined in [122], an SNR of at least 18 dB is necessary for DRS to yield accurate estimations. Therefore, we introduce a second short-reach system with high bandwidth to surpass this threshold. In this system, we maintain the same parameters as those mentioned for the long-haul system, except for the fiber length, reduced to 500 km, and the channel bandwidth increased to 400 GBaud. Additionally, we employ 64-QAM to elevate the system data rate for this channel.

5.4.2 Simulation Results

In this subsection, we begin by examining the minimum signal requirements necessary for an accurate receiver phase noise estimation using our CPE technique. We then proceed to study the performance improvement achieved by mitigating EEPN. Our investigation focuses on evaluating the system performance concerning key factors influencing EEPN, including baud rate, system reach, and laser linewidth, as suggested by (5.3). To establish a baseline for BER comparison, we assume that a raw BER of 0.007 is necessary for the forward error correction code. A practical implementation of such a code is outlined in [135] based on staircase coding with 9% overhead.

Minimum SNR for Accurate Estimation

Similar to other phase estimation techniques, our CPE also requires a minimum SNR to ensure acceptable phase estimation results. We utilize the mean squared error (MSE) between the estimated receiver phase, ϕ_r', and the actual phase, ϕ_r, as a metric for assessing the accuracy of phase estimation. The MSE is defined as follows:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}\Big(\phi_r'[i] - \phi_r[i]\Big)^2, \tag{5.36}$$

where N is the number of the estimated phase noise samples. Based on our observations, we have chosen an MSE of 0.1 rad^2 as the minimum accuracy required for receiver phase noise estimation to control EEPN.

In this simulation, we determine the minimum SNR required for our proposed CPE to achieve phase estimation with MSE = 0.1 rad^2 across different pulse excess bandwidths. Table 5.1 presents the results. This simulation demonstrates that our CPE can achieve the specified minimum MSE with very

β	0.05	0.1	0.15	0.2	0.25	0.3
Minimum SNR (dB)	0.5	-0.5	-1	-1.5	-2.5	-4

Table 5.1: Minimum SNR required for our CPE to result in a phase estimation with MSE = 0.1 rad^2.

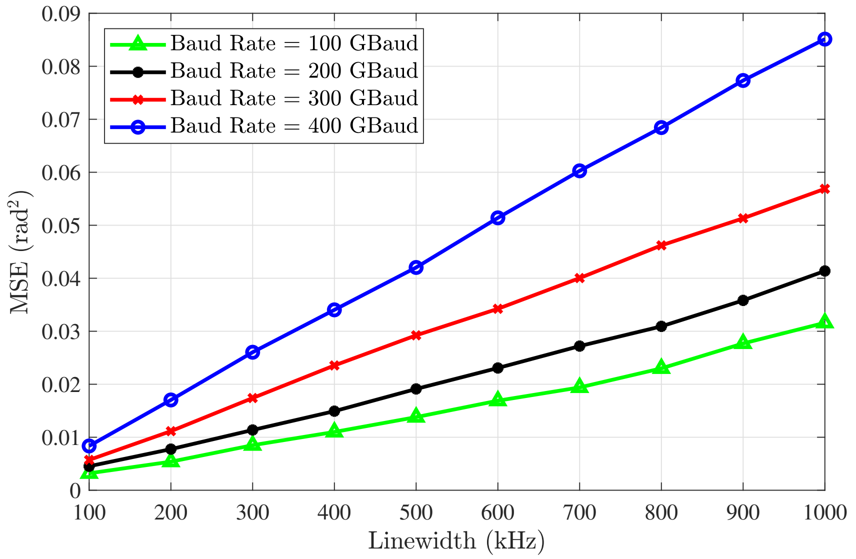


Figure 5.5: MSE performance of the phase noise estimation based on our proposed CPE versus laser linewidth.

low SNR values. Additionally, our CPE can estimate receiver phase noise even for cases with very small roll-off factors, such as $\beta = 0.05$. However, increasing the roll-off factor can decrease the minimum required SNR. This is because, as the excess bandwidth increases, more information becomes available to our CPE, resulting in expected performance improvements.

Impact of Laser Linewidth on Estimation Accuracy

Large laser linewidth causes rapid variations in phase noise in the time domain, making estimation challenging. The time resolution of our CPE is determined by the channel baud rate, as indicated in (5.29). Therefore, for high laser linewidths, we anticipate a reduction in phase estimation accuracy due to the maximum achievable time resolution. In this simulation, we examine the MSE of the estimated phase noise for different laser linewidths. Figure 5.5 illustrates the results. As observed, our CPE suffers from a higher MSE for larger laser linewidths. Furthermore, the MSE for channels with higher baud rates is larger than for channels with smaller baud rates, as expected. Therefore, in cases with high laser linewidths, it is desirable to use a smaller channel baud rate to provide more time resolution for phase estimation using our CPE.

In summary, in this simulation, we found that MSE increases with laser

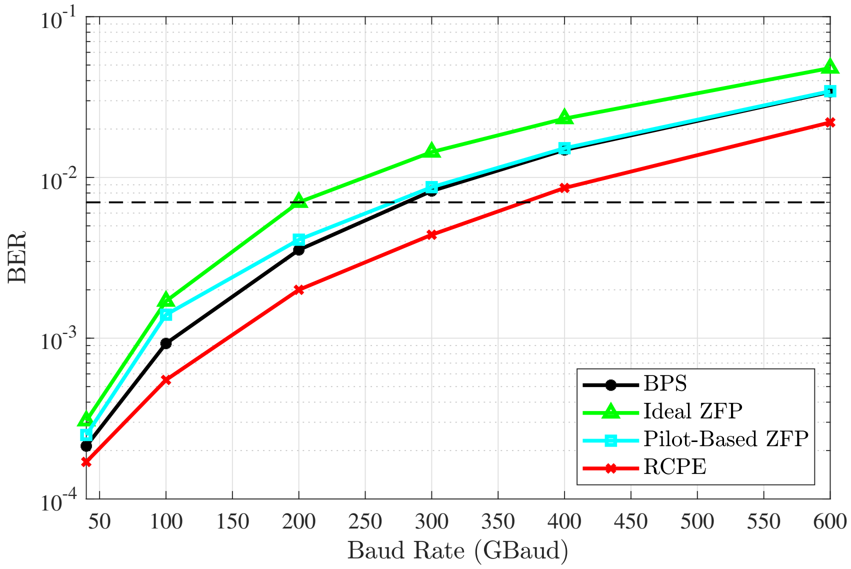


Figure 5.6: BER performance of the receivers based on different CPEs versus the channel baud rates.

linewidth due to the limited time resolution of the proposed CPE. Reducing the channel baud rate can enhance CPE estimation accuracy by increasing the time resolution for higher laser linewidths.

Impact of Channel Baud Rate

According to (5.3), EEPN noise power increases proportional to the baud rate, making EEPN one of the main limiting factors for increasing the system baud rate. Additionally, as per (5.29), the high baud rate reduces the time resolution of our proposed CPE. To assess the impact of the channel baud rate on system performance and our CPE's ability to mitigate EEPN generation, we calculate the BER performance for different baud rates, as shown in Figure 5.6. As observed, for small baud rates where EEPN is not a limiting factor, very low BER can be achieved. However, increasing the channel baud rate increases the BER not only due to EEPN but also because of nonlinear interference noise. In this simulation, ideal ZFP and BPS can achieve the target BER criteria up to 200 GBaud and 275 GBaud, respectively, while our proposed CPE extends the acceptable baud rate to 375 GBaud. Furthermore, we confirm that throughout the bandwidth range studied in this simulation, our CPE meets the MSE requirements, even for 600 GBaud, indicating that the time

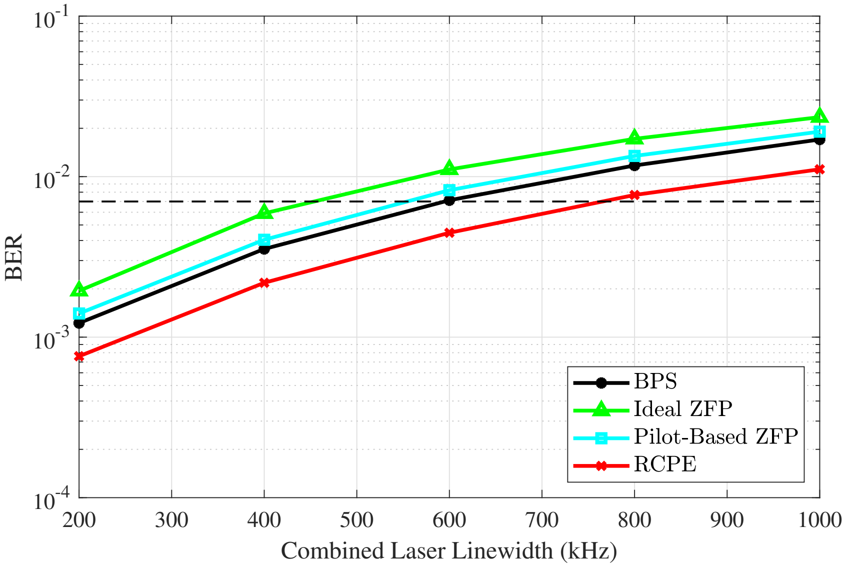


Figure 5.7: BER performance of the receivers based on different CPEs versus the combined laser linewidths.

resolution is sufficient.

Impact of Laser Linewidth

According to (5.3), EEPN noise power scales linearly with laser linewidth. Therefore, systems with higher laser linewidths experience more performance degradation due to EEPN. Here, we investigate the impact of laser linewidth on the BER performance of systems employing different CPEs. Figure 5.7 illustrates the results. As seen, the ideal ZFP is the most sensitive technique to laser linewidth, yet it can withstand combined laser linewidths of up to 450 kHz. In contrast, BPS method extends the acceptable laser linewidth to 600 kHz by partially mitigating the effects of EEPN. Notably, our proposed CPE significantly enhances system tolerance against laser linewidth, allowing for operation with combined linewidths of up to 780 kHz while meeting the target BER. It's crucial to acknowledge that laser phase noise acts as a detrimental source of noise within the system, and even with correct compensation, it can degrade system performance.

In summary, our CPE enables an increase in the acceptable combined laser linewidth by 30% compared to other proposed methods for EEPN control.

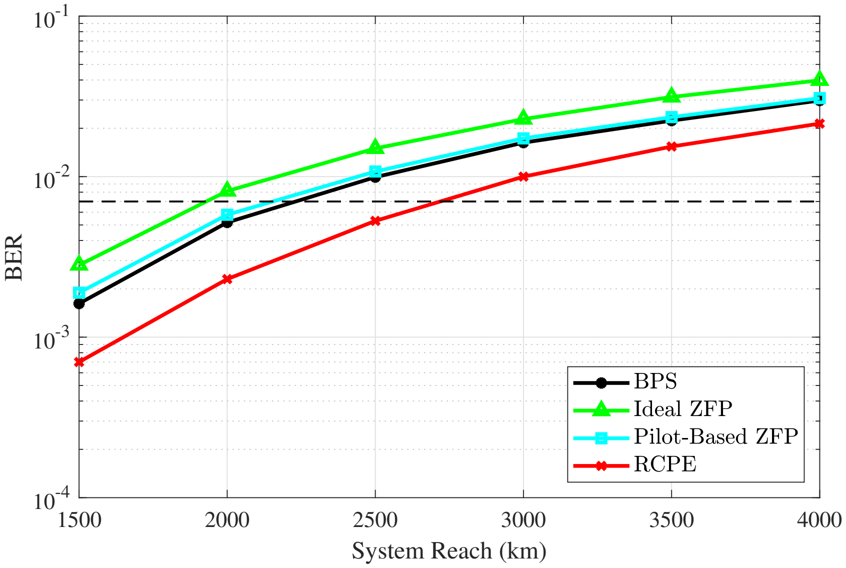


Figure 5.8: BER performance for different system reaches

System Reach

According to (5.3), EEPN power increases linearly with the system reach. Therefore, as we extend the system reach, we anticipate degradation in system performance due to EEPN. In this simulation, we examine the performance improvements achieved by employing our CPE to mitigate EEPN across various system reaches. Figure 5.8 illustrates the corresponding BER performance for different system reaches. Observations indicate that the ideal ZFP method can maintain the assumed raw BER up to 1900 km, while the BPS technique extends the system reach to 2250 km by partially mitigating the effects of EEPN. Conversely, our CPE technique enhances the system reach to 2750 km by avoiding EEPN generation.

In summary, utilizing our CPE technique, we can increase the system reach by 22% in this system setup while guaranteeing the target BER.

Impact of Launch Power

The launch power into the fiber directly determines the operating region of the fiber. In this simulation, we calculate the effective SNR of the fiber system for various launch powers to investigate the effect of EEPN on different operating regions. Figure 5.9 illustrates the results. Observations indicate that

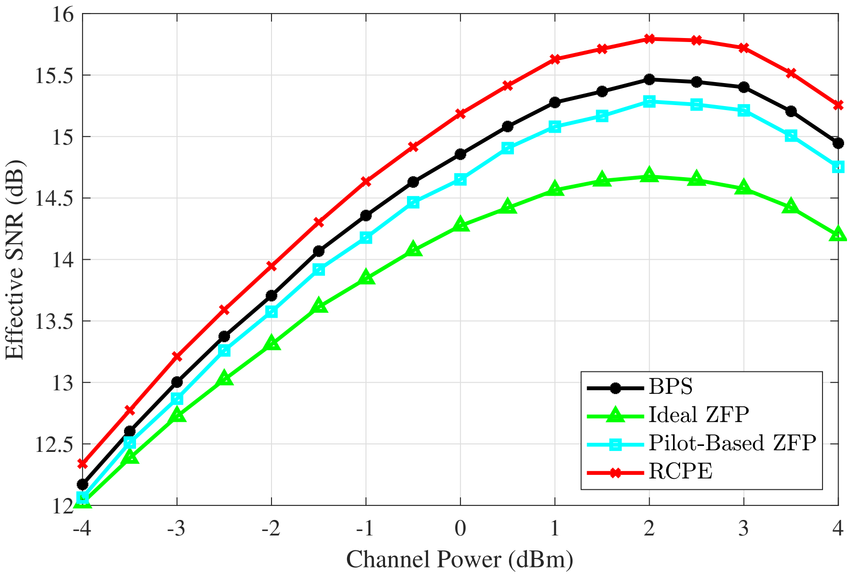


Figure 5.9: Effective SNR for different launch power.

for the considered system setup, the maximum achievable SNR is lower than the minimum threshold for DRS, which is 18 dB. Our proposed CPE method enhances the effective SNR in both linear and optimal regions by avoiding EEPN generation.

The performance gain is particularly pronounced in the optimal region. Compared to the ideal ZFP method, our CPE technique can enhance the maximum effective SNR by 1.1 dB. Furthermore, we observe that the optimal launch power is shifted to a lower value for our CPE compared to other methods by 0.5 dB. This effect can be explained based on the derived expression for the optimal launch power in [71] as:

$$P_{\text{opt}} = \sqrt[3]{\frac{P_{\text{lin}}}{2\eta}}, \tag{5.37}$$

where P_{lin} represents the linear noise power, ASE noise and EEPN, and η is a constant. Consequently, decreasing the linear noise power will shift the optimal launch power towards lower values. This adjustment effectively diminishes nonlinearity within the system.

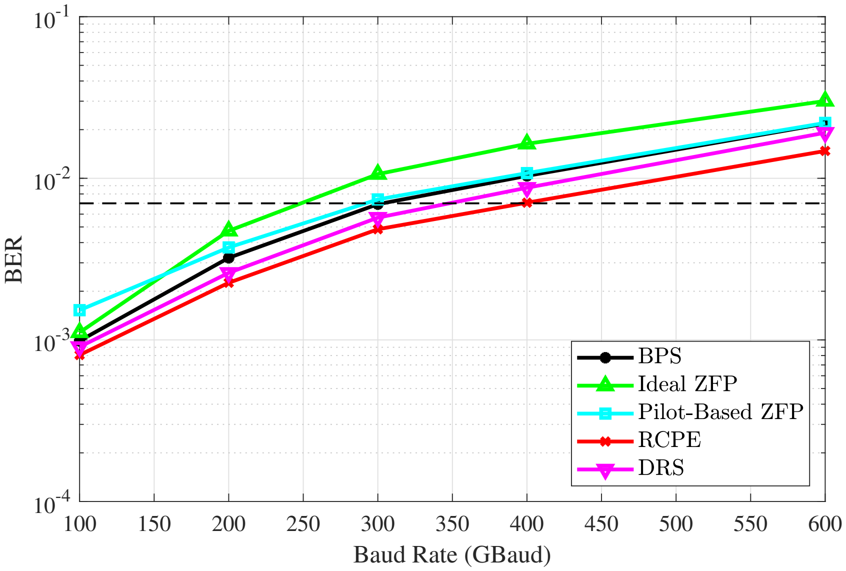


Figure 5.10: BER performance of the receivers based on different CPEs versus the channel baud rates in a short-reach system.

Comparison to DRS on Short-Reach Systems

As discussed in Simulation 5.4.2, the long-haul system considered does not meet the minimum required SNR for the DRS technique. DRS shows promise for phase estimation in digital subcarrier systems, leveraging two known pilot sequences in different subchannels. This technique extracts receiver phase noise based on the time difference between the arrival of pilots on different subchannels. However, since DRS requires symbols to be extracted, receiver phase noise is not available before the dispersion compensator, making optimal compensation order impossible.

Nevertheless, DRS can enhance phase estimation in digital subcarrier systems. In [123], it is shown that DRS offers benefits in terms of EEPN when subchannel bandwidth is small. In this simulation, we investigate the performance of different CPEs in a short-reach system where the minimum SNR of 18 dB is met. For DRS, we implement 8 subchannels with pilots transmitted on the side subchannels. Figure 5.10 illustrates the results. Observations show that while DRS effectively mitigates the EEPN effect for low baud rates, its performance diminishes with an increase in channel baud rate. Conversely, our CPE technique outperforms existing methods across both low and high baud rates. It is worth noting that increasing the number of subchannels in DRS can enhance EEPN mitigation for high baud rate channels; however, this also increases complexity, which may not be feasible in all applications.

In summary, compared to DRS, our CPE requires significantly lower SNRs for accurate receiver phase estimation. Furthermore, unlike DRS, our CPE can extract receiver phase noise before dispersion compensation, allowing for the optimal compensation order. As observed in this simulation, our CPE can increase the system baud rate by 15% compared to DRS for the same raw BER in a short-reach system.

5.5 Conclusion

In a fiber system, compensation for receiver laser phase noise, dispersion, and transmitter laser phase noise should ideally occur in the inverse order of their occurrence. However, practical limitations in phase estimation have forced existing receivers to estimate and mitigate receiver phase noise combined with transmitter phase after dispersion compensation. This suboptimal order introduces EEPN, degrading system performance, especially for high baud rates over long reaches. In this chapter, we addressed the problem of receiver phase noise estimation before dispersion compensation, allowing for the optimal order of compensators. Leveraging the fact that the signal excess bandwidth is modulated with the same data but experiences different time delays due to the dispersive fiber channel, we extracted receiver phase noise variation in time after proper signal processing. Based on this approach, we proposed a DSP-ready implementation applicable to modern fiber receivers. We then conducted a simulation study to validate our designed CPE. The results demonstrated that our CPE requires very low SNR for accurate estimation of receiver phase noise compared to existing methods. Furthermore, our CPE enhanced system performance in various aspects, such as system reach, optimal channel power, and baud rate, by avoiding the generation of EEPN.

Chapter 6

Summary and Future Work

6.1 Summary

The growing popularity of data-hungry applications, such as video streaming, virtual reality, and generative artificial intelligence, raises concerns about the future capacity of communication infrastructure [146]. Therefore, fiber optics, as the backbone of high-capacity communication systems, is subjected to extensive research investigation to meet future data demands [2], [147]. This research is mainly directed towards designing sophisticated signal processing algorithms that enhance receiver/transmitter compatibility with fiber channel impairments.

Fiber channel capacity is mainly limited due to Kerr nonlinearity effects [23], [25]. As reviewed in Chapter 1, Kerr nonlinearity occurs because of the power dependency of the silica refractive index. This phenomenon results in a series of impairments, including SPM, XPM, and FWM. Although the mechanisms of these impairments are known, their interaction with fiber dispersion causes unknown effects. Therefore, reducing NLIN – as a result of Kerr nonlinearity – is of great importance.

In Chapter 2, we showed that NLIN power is a function of power variations. Therefore, one could minimize the NLIN power by reducing power fluctuations. We then developed a framework that pairs the constellations transmitted in different subchannels of a digital subcarrier system to reduce the nonlinearity impact. We showed an example of this framework based on the 16-QAM constellation. Although pairing the constellations reduces the entropy, we analytically showed that by increasing the number of paired constellations, one can easily approach the unpaired system while gaining nonlinearity benefits.

In Chapter 3, we focused on the pulse PSD as another factor impacting NLIN. Based on the GN model, the PSD of the pulse launched into the system determines the NLIN power. Therefore, pulse PSD can be designed to minimize system nonlinearity. To this end, we formulated the NLIN as a function of a general pulse model subject to classical communication requirements such as zero-ISI and band limitation. We then established an optimization problem to find the optimal pulse shape. The resulting pulse improved different aspects of system performance, including system reach and maximum achievable SNR.

Another degradation in fiber systems is due to the suboptimality of transmitters and receivers imposed by practical considerations. One of the main challenges in existing receivers is EEPN. EEPN is caused by the suboptimal order of the receiver phase noise compensator and dispersion compensator. While in an ideal receiver, the dispersion compensator should be implemented after receiver laser phase noise compensation, phase noise estimation is only possible after dispersion compensation due to the required pilots. EEPN power grows with data rate and system reach, making it hard to achieve high data rates over long distances.

To overcome the challenge of EEPN, in Chapter 4, we proposed a new formulation of EEPN that enabled us to compensate for it in the output of a suboptimal receiver. This compensator requires the receiver laser phase noise to be distinguished from the transmitter laser phase noise. Based on the availability of the receiver laser phase noise, we proposed two DSP-ready implementations of our compensator. We showed that this compensator can be implemented based on a simple time-variant FIR filter. We performed a complexity analysis and showed that our compensator's complexity is comparable to that of the existing CPEs.

As another solution to EEPN, in Chapter 5, we designed a new CPE capable of (I) extracting the receiver laser phase noise before dispersion and (II) distinguishing it from the transmitter laser phase noise. These two properties allowed the optimal sequence of compensators to avoid EEPN. Our proposed

CPE was based on the fact that pulse excess bandwidth on the positive and negative frequencies is modulated with the same information while undergoing different delays because of fiber dispersion. Therefore, after proper signal processing, we could extract the impacting receiver laser phase noise temporal variations by observing the positive and negative excess bandwidths. Based on this idea, we developed a DSP-ready implementation of our CPE. Our CPE exhibits very robust performance compared to existing methods for phase estimation. Our techniques for EEPN control enhance the overall system performance, including system reach and baud rate.

6.2 Future Work

6.2.1 Operator-based Parallel Digital Backpropagation in Fiber Optics

In the first part of this thesis, we focused on controlling the fiber nonlinearity. As discussed, reducing fiber nonlinearity can effectively increase fiber capacity. Whenever the neighboring channels in a WDM system are available, one of the most effective methods to compensate for fiber nonlinearity is the wide-band digital backpropagation (DBP). Unfortunately, the computational complexity of DBP makes its practical application challenging [92]. Therefore, reducing the complexity of DBP can be considered another practical solution to mitigate fiber nonlinearity.

In DBP, received signals propagate in an imaginary fiber with parameters of opposite sign to reverse the fiber effects on the propagated pulses [92]. An optical fiber with a negative nonlinear parameter is not physically realizable. Therefore, DBP simulates such an optical fiber by considering the NLSE [26]. There are two main types of NLSE known as the regular NLSE and the coupled NLSE [89]. The regular NLSE can deal with all intra and inter-channel impairments if all the WDM channels are available as an ultra-wideband signal. Unfortunately, working with such a wideband signal in real-time exceeds today's processing capabilities. On the other hand, the coupled NLSE works on the baseband channels separately but only compensates for XPM [90]. Despite its lower accuracy compared to the regular NLSE, the coupled NLSE allows us to distribute the complexity of the compensation method among parallel processors. Although parallelization of DBP is a practical solution to its complexity challenges, existing parallel DBP implementations are based on the coupled NLSE, which cannot consider FWM.

As a future work regarding fiber nonlinearity, we propose a new parallel DBP scheme based on the regular NLSE, capable of compensating for any fiber impairments, including FWM. In this approach, one can still use the regular NLSE and focus on parallelizing the NLSE solver instead of using the coupled NLSE. It can be shown that the linear operator is equivalent to an ordinary differential equation in which we can easily treat the WDM channels separately. However, in the nonlinear operator, a WDM channel can affect its neighbor channels. Performing this operator in parallel requires computing many new terms, which may question the idea of reducing DBP complexity. Therefore, deriving these terms mathematically and effectively handling them is of great importance.

6.2.2 Constellation Design to Combat Equalization Enhanced Phase Noise

Constellation shaping is a promising approach in fiber optics to enhance system performance. For instance, in [148], authors develop a constellation to enhance fiber nonlinear behavior based on the EGN model. Additionally, kurtosis-limited constellations [84], constant composition distribution matching [149], and constant power constellation grouping [100] methods are proposed to mitigate nonlinearity in the system. While the existing focus of constellation design has primarily been on mitigating nonlinearity, EEPN control has received less attention.

As a future work regarding EEPN, we propose designing constellations that mitigate the impact of EEPN on the system. Existing models reveal the non-circular symmetry of EEPN-induced noise, with the phase component dominant for high-power symbols [117]. One can build upon existing models for EEPN-induced noise to develop a metric for constellation symbol error rate.

By utilizing the proposed metric, we can design EEPN-aware constellations tailored for fiber systems. Essentially, these constellations should offer an extended phase direction distance for points with higher energy compared to lower energy symbols to reduce the symbol error rate.

References

[1] C. De Alwis, A. Kalla, Q.-V. Pham, *et al.*, "Survey on 6G frontiers: Trends, applications, requirements, technologies and future research," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 836–886, 2021.

[2] P. J. Winzer, D. T. Neilson, and A. R. Chraplyvy, "Fiber-optic transmission and networking: The previous 20 and the next 20 years," *Optics express*, vol. 26, no. 18, pp. 24 190–24 239, 2018.

[3] A. Arnould and A. Ghazisaeidi, "Equalization enhanced phase noise in coherent receivers: DSP-aware analysis and shaped constellations," *Journal of Lightwave Technology*, vol. 37, no. 20, pp. 5282–5290, 2019.

[4] S. Popov, G. Jacobsen, and S. Sergeyev, "High capacity coherent optical systems: Advanced modulation formats and margins for transmission impairments," in *2015 17th International Conference on Transparent Optical Networks (ICTON)*, IEEE, 2015, pp. 1–2.

[5] M. Qiu, Q. Zhuge, M. Y. Sowailem, *et al.*, "Equalization-enhanced phase noise in stokes-vector direct detection systems," *IEEE Photonics Journal*, vol. 8, no. 6, pp. 1–7, 2016.

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

[7] W. Shieh and K.-P. Ho, "Equalization-enhanced phase noise for coherent-detection systems using electronic digital signal processing," *Optics Express*, vol. 16, no. 20, pp. 15 718–15 727, 2008.

[8] A. P. T. Lau, T. S. R. Shen, W. Shieh, and K.-P. Ho, "Equalization-enhanced phase noise for 100Gb/s transmission and beyond with coherent detection," *Optics Express*, vol. 18, no. 16, pp. 17 239–17 251, 2010.

[9] A. Kakkar, J. R. Navarro, R. Schatz, *et al.*, "Comprehensive study of equalization-enhanced phase noise in coherent optical systems," *Journal of lightwave technology*, vol. 33, no. 23, pp. 4834–4841, 2015.

[10] K. C. Kao and G. A. Hockham, "Dielectric-fibre surface waveguides for optical frequencies," in *Proceedings of the Institution of Electrical Engineers*, IET, vol. 113, 1966, pp. 1151–1158.

[11] R. D. Maurer and P. C. Schultz, *Fused silica optical waveguide*, US Patent 3,659,915, May 1972.

[12] D. Gloge, "Optical power flow in multimode fibers," *Bell System Technical Journal*, vol. 51, no. 8, pp. 1767–1783, 1972.

[13] E. Desurvire, J. R. Simpson, and P. Becker, "High-gain erbium-doped traveling-wave fiber amplifier," *Optics letters*, vol. 12, no. 11, pp. 888–890, 1987.

[14] O. E. DeLange, "Wide-band optical communication systems: Part II frequency division multiplexing," *Proceedings of the IEEE*, vol. 58, no. 10, pp. 1683–1690, 1970.

[15] P. J. Winzer and R.-J. Essiambre, "Advanced optical modulation formats," in *Optical Fiber Telecommunications VB*, Elsevier, 2008, pp. 23–93.

[16] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Optics Express*, vol. 21, no. 22, pp. 25 685–25 699, 2013.

[17] P. Massoud Salehi and J. Proakis, *Digital Communications*. McGraw-Hill Education, 2007, ISBN: 9780072957167.

[18] G. P. Agrawal, *Fiber-optic communication systems*. John Wiley & Sons, 2012.

[19] M. Chagnon, "Optical communications for short reach," *Journal of Lightwave Technology*, vol. 37, no. 8, pp. 1779–1797, 2019.

[20] K. Kikuchi, "Fundamentals of coherent optical fiber communications," *Journal of lightwave technology*, vol. 34, no. 1, pp. 157–179, 2015.

[21] A. Tervonen, P. Poyhonen, S. Honkanen, and M. Tahkokorpi, "A guided-wave Mach-Zehnder interferometer structure for wavelength multiplexing," *IEEE photonics technology letters*, vol. 3, no. 6, pp. 516–518, 1991.

[22] D. Halliday, R. Resnick, and J. Walker, *Fundamentals of physics*. John Wiley & Sons, 2013.

[23] G. Keiser, *Optical fiber communications*. McGraw-Hill New York, 2000, vol. 2.

[24] D. K. Cheng *et al.*, *Field and wave electromagnetics*. Pearson Education India, 1989.

[25] G. P. Agrawal, "Nonlinear fiber optics," in *Nonlinear Science at the Dawn of the 21st Century*, Springer, 2000, pp. 195–211.

[26] G. Fibich, *The nonlinear Schrödinger equation*. Springer, 2015.

[27] S. V. Manakov, "On the theory of two-dimensional stationary self-focusing of electromagnetic waves," *Soviet Physics-JETP*, vol. 38, no. 2, pp. 248–253, 1974.

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

[29] J. Shao, X. Liang, and S. Kumar, "Comparison of split-step Fourier schemes for simulating fiber optic communication systems," *IEEE Photonics Journal*, vol. 6, no. 4, pp. 1–15, 2014.

[30] R. Deiterding, R. Glowinski, H. Oliver, and S. Poole, "A reliable split-step fourier method for the propagation equation of ultra-fast pulses in single-mode optical fibers," *Journal of Lightwave Technology*, vol. 31, no. 12, pp. 2008–2017, 2013.

[31] A. Yunakovsky *et al.*, "Split-step fourier method for nonlinear schrodinger equation," in *DAYS on DIFFRACTION 2006*, IEEE, 2006, pp. 34–42.

[32] R. Achilles and A. Bonfiglioli, "The early proofs of the theorem of Campbell, Baker, Hausdorff, and Dynkin," *Archive for history of exact sciences*, vol. 66, no. 3, pp. 295–358, 2012.

[33] X. Liang and S. Kumar, "Analytical modeling of XPM in dispersion-managed coherent fiber-optic systems," *Optics Express*, vol. 22, no. 9, pp. 10 579–10 592, 2014.

[34] Z. Tao, W. Yan, L. Liu, *et al.*, "Simple fiber model for determination of XPM effects," *Journal of Lightwave Technology*, vol. 29, no. 7, pp. 974–986, 2011.

[35] M. Ajmani and P. Singh, "FWM in WDM system, effects and techniques to minimize: A review," in *2015 Fifth International Conference on Advanced Computing & Communication Technologies*, IEEE, 2015, pp. 385–389.

[36] C. Wehmann, L. Fernandes, C. Sobrinho, *et al.*, "Analysis of the four wave mixing effect (FWM) in a dispersion decreasing fiber (DDF) for a WDM system," *Optical Fiber Technology*, vol. 11, no. 3, pp. 306–318, 2005.

[37] J. Huang, "Impact of ASE noise in WDM systems," *Optik*, vol. 122, no. 15, pp. 1376–1380, 2011.

[38] W. Jiang, J. Chen, T. Shang, and G. Wu, "Analysis of ASE noise in optical fiber raman amplifiers," *Journal of optical communications*, vol. 26, no. 5, pp. 201–203, 2005.

[39] Y. Bai, F. Yan, T. Feng, *et al.*, "Demonstration of linewidth measurement based on phase noise analysis for a single frequency fiber laser in the 2 μm band," *Laser Physics*, vol. 29, no. 7, p. 075 102, 2019.

[40] L. N. Binh, *Optical fiber communication systems with Matlab and Simulink models*. CRC press, 2015.

[41] T. M. Cover, *Elements of information theory*. John Wiley & Sons, 1999.

[42] A. Graell i Amat and L. Schmalen, “Forward error correction for optical transponders,” *Springer Handbook of Optical Networks*, pp. 177–257, 2020.

[43] M. P. Yankov, D. Zibar, K. J. Larsen, L. P. Christensen, and S. Forchhammer, “Constellation shaping for fiber-optic channels with QAM and high spectral efficiency,” *IEEE Photonics Technology Letters*, vol. 26, no. 23, pp. 2407–2410, 2014.

[44] J. Barrueco, J. Montalban, E. Iradier, and P. Angueira, “Constellation design for future communication systems: A comprehensive survey,” *IEEE Access*, vol. 9, pp. 89 778–89 797, 2021.

[45] M. P. Yankov, F. Da Ros, E. P. da Silva, *et al.*, “Constellation shaping for WDM systems using 256qam/1024qam with probabilistic optimization,” *Journal of Lightwave Technology*, vol. 34, no. 22, pp. 5146–5156, 2016.

[46] J. Cho and P. J. Winzer, “Probabilistic constellation shaping for optical fiber communications,” *Journal of Lightwave Technology*, vol. 37, no. 6, pp. 1590–1607, 2019.

[47] A. Soleimanzade, M. Ardakani, and H. Ebrahimzad, “Optimization of the non-linearity tolerant 4D geometric shaped constellations for optical fiber communication systems using neural networks,” *Journal of Lightwave Technology*, 2023.

[48] D. Pilori, A. Nespola, F. Forghieri, and G. Bosco, “Non-linear phase noise mitigation over systems using constellation shaping,” *Journal of Lightwave Technology*, vol. 37, no. 14, pp. 3475–3482, 2019.

[49] E. Sillekens, G. Liga, D. Lavery, P. Bayvel, and R. I. Killey, “High-cardinality geometrical constellation shaping for the nonlinear fibre channel,” *Journal of Lightwave Technology*, vol. 40, no. 19, pp. 6374–6387, 2022.

[50] G. Böcherer, P. Schulte, and F. Steiner, “Probabilistic shaping and forward error correction for fiber-optic communication systems,” *Journal of Lightwave Technology*, vol. 37, no. 2, pp. 230–244, 2019.

[51] T. Fehenberger, A. Alvarado, G. Böcherer, and N. Hanik, “On probabilistic shaping of quadrature amplitude modulation for the nonlinear fiber channel,” *Journal of Lightwave Technology*, vol. 34, no. 21, pp. 5063–5073, 2016.

[52] A. Mirani, E. Agrell, and M. Karlsson, “Low-complexity geometric shaping,” *Journal of Lightwave Technology*, vol. 39, no. 2, pp. 363–371, 2020.

[53] Z. Qu and I. B. Djordjevic, “On the probabilistic shaping and geometric shaping in optical communication systems,” *IEEE Access*, vol. 7, pp. 21 454–21 464, 2019.

[54] H. Dzieciol, G. Liga, E. Sillekens, P. Bayvel, and D. Lavery, "Geometric shaping of 2-D constellations in the presence of laser phase noise," *Journal of Lightwave Technology*, vol. 39, no. 2, pp. 481–490, 2020.

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

[56] N. Sultana, H. Tada, H. Imai, and T. Shioda, "Dispersion pre compensation of 25.6 Tbps waveforms using an optical frequency comb synthesizer/analyzer," *Optics Communications*, vol. 475, p. 126 196, 2020.

[57] V. Düendorfs, S. Spolitis, and V. Bobrovs, "Comparison of dispersion compensation methods for 40Gbit/s WDM-PON transmission systems," in *2017 Progress In Electromagnetics Research Symposium-Spring (PIERS)*, IEEE, 2017, pp. 2431–2436.

[58] T. Kimura, "Coherent optical fiber transmission," *Journal of lightwave technology*, vol. 5, no. 4, pp. 414–428, 1987.

[59] F. Gardner, "A BPSK/QPSK timing-error detector for sampled receivers," *IEEE Transactions on communications*, vol. 34, no. 5, pp. 423–429, 1986.

[60] D. Godard, "Passband timing recovery in an all-digital modem receiver," *IEEE Transactions on Communications*, vol. 26, no. 5, pp. 517–523, 1978.

[61] L. Grüner-Nielsen, M. Wandel, P. Kristensen, *et al.*, "Dispersion compensating fibers," *Journal of Lightwave Technology*, vol. 23, no. 11, p. 3566, 2005.

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

[63] 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," *Optics Express*, vol. 22, no. 24, pp. 30 053–30 062, 2014.

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

[65] R. Puerta, J. Yu, X. Li, Y. Xu, J. J. V. Olmos, and I. T. Monroy, "Single-carrier dual-polarization 328-Gb/s wireless transmission in a D-band millimeter wave 2× 2 MU-MIMO radio-over-fiber system," *Journal of Lightwave Technology*, vol. 36, no. 2, pp. 587–593, 2018.

[66] S. L. Jansen, I. Morita, T. C. Schenk, and H. Tanaka, “Long-haul transmission of16× 52.5 Gbits/s polarization-division-multiplexed OFDM enabled by MIMO processing,” *Journal of Optical Networking*, vol. 7, no. 2, pp. 173–182, 2008.

[67] S. O. Arik, J. M. Kahn, and K.-P. Ho, “MIMO signal processing for mode-division multiplexing: An overview of channel models and signal processing architectures,” *IEEE Signal Processing Magazine*, vol. 31, no. 2, pp. 25–34, 2014.

[68] M. Morsy-Osman, Q. Zhuge, L. R. Chen, and D. V. Plant, “Feedforward carrier recovery via pilot-aided transmission for single-carrier systems with arbitrary M-QAM constellations,” *Optics express*, vol. 19, no. 24, pp. 24 331–24 343, 2011.

[69] M. Magarini, L. Barletta, A. Spalvieri, *et al.*, “Pilot-symbols-aided carrier-phase recovery for 100-G PM-QPSK digital coherent receivers,” *IEEE Photonics Technology Letters*, vol. 24, no. 9, pp. 739–741, 2012.

[70] W.-R. Peng, T. Tsuritani, and I. Morita, “Simple carrier recovery approach for RF-pilot-assisted PDM-CO-OFDM systems,” *Journal of lightwave technology*, vol. 31, no. 15, pp. 2555–2564, 2013.

[71] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, “The GN-model of fiber non-linear propagation and its applications,” *Journal of lightwave technology*, vol. 32, no. 4, pp. 694–721, 2013.

[72] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, “EGN model of non-linear fiber propagation,” *Optics express*, vol. 22, no. 13, pp. 16 335–16 362, 2014.

[73] M. Sjodin, E. Agrell, and M. Karlsson, “Subset-optimized polarization-multiplexed PSK for fiber-optic communications,” *IEEE communications letters*, vol. 17, no. 5, pp. 838–840, 2013.

[74] L. Beygi, E. Agrell, J. M. Kahn, and M. Karlsson, “Coded modulation for fiber-optic networks: Toward better tradeoff between signal processing complexity and optical transparent reach,” *IEEE Signal Processing Magazine*, vol. 31, no. 2, pp. 93–103, 2014.

[75] Y. Ozeki, M. Kishi, and A. Tsuchiya, “Fiber-optic transmission of 60 GHz DBPSK signal employing the dual-mode PSK modulation (DMPM) method,” in *2001 International Topical Meeting on Microwave Photonics. Technical Digest. MWP'01 (Cat. No. 01EX476)*, IEEE, 2002, pp. 65–68.

[76] H. Khodakarami and W. Shieh, “On the energy efficiency of modulation formats for optical communications,” *IEEE photonics technology letters*, vol. 25, no. 3, pp. 275–278, 2012.

[77] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "On shaping gain in the nonlinear fiber-optic channel," in *2014 IEEE International Symposium on Information Theory*, IEEE, 2014, pp. 2794–2798.

[78] B. Chen, G. Liga, Y. Lei, *et al.*, "Shaped four-dimensional modulation formats for optical fiber communication systems," in *Optical Fiber Communication Conference*, Optica Publishing Group, 2022.

[79] P. Skvortcov, I. Phillips, W. Forysiak, *et al.*, "Huffman-coded sphere shaping for extended-reach single-span links," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 27, no. 3, pp. 1–15, 2021.

[80] A. Amari, S. Goossens, Y. C. Gültekin, *et al.*, "Introducing enumerative sphere shaping for optical communication systems with short blocklengths," *Journal of Lightwave Technology*, vol. 37, no. 23, pp. 5926–5936, 2019.

[81] W.-R. Peng, A. Li, Q. Guo, Y. Cui, and Y. Bai, "Transmission method of improved fiber nonlinearity tolerance for probabilistic amplitude shaping," *Optics Express*, vol. 28, no. 20, pp. 29 430–29 441, 2020.

[82] M. P. Yankov, K. J. Larsen, and S. Forchhammer, "Temporal probabilistic shaping for mitigation of nonlinearities in optical fiber systems," *Journal of Lightwave Technology*, vol. 35, no. 10, pp. 1803–1810, 2017.

[83] P. Schulte and G. Böcherer, "Constant composition distribution matching," *IEEE Transactions on Information Theory*, vol. 62, no. 1, pp. 430–434, 2015.

[84] Y. C. Gültekin, A. Alvarado, O. Vassilieva, *et al.*, "Kurtosis-limited sphere shaping for nonlinear interference noise reduction in optical channels," *Journal of Lightwave Technology*, vol. 40, no. 1, pp. 101–112, 2022.

[85] P. Delsarte, J.-M. Goethals, and J. J. Seidel, "Spherical codes and designs," in *Geometry and Combinatorics*, Elsevier, 1991, pp. 68–93.

[86] J. Bao, Z. Ma, M. A. Mahamadu, Z. Zhu, and D. Chen, "Spherical codes for SCMA codebook," in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, IEEE, 2016, pp. 1–5.

[87] R. R. Borujeny, S. E. Rumsey, S. C. Draper, and F. R. Kschischang, "Soft-in soft-out decoding of spherical codes from cartesian powers of PAM constellations," *arXiv preprint arXiv:2404.04776*, 2024.

[88] G. Fibich, *The nonlinear Schrödinger equation.* Springer, 2015, vol. 192.

[89] M. Secondini and E. Forestieri, "Analytical fiber-optic channel model in the presence of cross-phase modulation," *IEEE Photonics Technology Letters*, vol. 24, no. 22, pp. 2016–2019, 2012.

[90] C. Li, F. Zhang, L. Zhang, *et al.*, "Inter-channel fiber nonlinearity mitigation in high baud-rate optical communication systems," *Journal of Lightwave Technology*, vol. 39, no. 6, pp. 1653–1661, 2020.

[91] G. P. Agrawal, *Nonlinear fiber optics*. Elsevier, 2012.

[92] A. Amari, O. A. Dobre, R. Venkatesan, O. S. Kumar, P. Ciblat, and Y. Jaouën, "A survey on fiber nonlinearity compensation for 400 Gb/s and beyond optical communication systems," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 3097–3113, 2017.

[93] M. P. Yankov, T. Fehenberger, L. Barletta, and N. Hanik, "Low complexity tracking of laser and nonlinear phase noise in wdm optical fiber systems," *Journal of Lightwave Technology*, vol. 33, no. 23, pp. 4975–4984, 2015.

[94] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE transactions on Information Theory*, vol. 28, no. 1, pp. 55–67, 1982.

[95] H. Sun, M. Torbatian, M. Karimi, *et al.*, "800g DSP asic design using probabilistic shaping and digital sub-carrier multiplexing," *Journal of lightwave technology*, vol. 38, no. 17, pp. 4744–4756, Sep. 2020.

[96] J. Cho and R. Tkach, "On the kurtosis of modulation formats for characterizing the nonlinear fiber propagation," *Journal of Lightwave Technology*, 2022.

[97] M. P. Yankov, O. Jovanovic, D. Zibar, and F. Da Ros, "Recent advances in constellation optimization for fiber-optic channels," in *European Conference and Exhibition on Optical Communication*, Optica Publishing Group, 2022, Mo3D–4.

[98] A. Amari, S. Goossens, Y. C. Gültekin, *et al.*, "Introducing enumerative sphere shaping for optical communication systems with short blocklengths," *Journal of Lightwave Technology*, vol. 37, no. 23, pp. 5926–5936, 2019.

[99] A. Soleimanzade and M. Ardakani, "EGN-based optimization of the apsk constellations for the non-linear fiber channel based on the symbolwise mutual information," *Journal of Lightwave Technology*, vol. 40, no. 7, pp. 1937–1952, 2021.

[100] A. Abolfathimomtaz, M. Ardakani, H. Ebrahimzad, and R. Yazdani, "Constant power constellation grouping for reducing nonlinearity effects in fiber optics," *Journal of Lightwave Technology*, 2022.

[101] Y. Liu and C. Wang, "The research on transmission effect of different pulse shapes in MPPM," in *2022 14th International Conference on Computer Research and Development (ICCRD)*, IEEE, 2022, pp. 373–377.

[102] Y.-H. Wang and I. Lyubomirsky, "Impact of DP-QPSK pulse shape in nonlinear 100 G transmission," *Journal of lightwave technology*, vol. 28, no. 18, pp. 2750–2756, 2010.

[103] M. A. Preciado and M. A. Muriel, “Bandlimited airy pulses for invariant propagation in single-mode fibers,” *Journal of lightwave technology*, vol. 30, no. 23, pp. 3660–3666, 2012.

[104] X. Xu, B. Châtelain, Q. Zhuge, *et al.*, “Frequency domain M-shaped pulse for SPM nonlinearity mitigation in coherent optical communications,” in *National Fiber Optic Engineers Conference*, Optica Publishing Group, 2013, JTh2A–38.

[105] X. Xu, Q. Zhuge, B. Châtelain, *et al.*, “A nonlinearity-tolerant frequency domain root M-shaped pulse for coherent optical communication systems,” *Optics Express*, vol. 21, no. 26, pp. 31 966–31 982, 2013.

[106] X. Xu, Q. Zhuge, B. Châtelain, *et al.*, “Experimental investigation on the nonlinear tolerance of root M-shaped pulse in spectrally efficient coherent transmissions,” *Optics express*, vol. 23, no. 2, pp. 882–894, 2015.

[107] A. Karar, S. Gazor, Y. Gao, *et al.*, “Polynomial pulses for mitigating fiber nonlinearity in coherent optical fiber communications,” *IEEE Photonics Technology Letters*, vol. 27, no. 15, pp. 1653–1655, 2015.

[108] A. Karar, Y. Gao, J. C. Cartledge, *et al.*, “Mitigating intra-channel nonlinearity in coherent optical communications using ISI-free polynomial pulses,” in *OFC 2014*, IEEE, 2014, pp. 1–3.

[109] A. Naji, B. A. Hamida, X. Cheng, *et al.*, “Review of erbium-doped fiber amplifier,” *International Journal of the Physical Sciences*, vol. 6, no. 20, pp. 4674–4689, 2011.

[110] P. A. Parrilo and R. R. Thomas, *Sum of Squares: Theory and Applications*. American Mathematical Soc., 2020, vol. 77.

[111] J. B. Lasserre, “A sum of squares approximation of nonnegative polynomials,” *IFAC Proceedings Volumes*, vol. 38, no. 1, pp. 441–444, 2005.

[112] T. Roh, B. Dumitrescu, and L. Vandenberghe, “Interior-point algorithms for sum-of-squares optimization of multidimensional trigonometric polynomials,” in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP’07*, IEEE, vol. 3, 2007, pp. III–905.

[113] M. Grant and S. Boyd, *CVX: Matlab software for disciplined convex programming, version 2.1*, `http://cvxr.com/cvx`, Mar. 2014.

[114] M. Grant and S. Boyd, “Graph implementations for nonsmooth convex programs,” in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds., `http://stanford.edu/~boyd/graph_dcp.html`, Springer-Verlag Limited, 2008, pp. 95–110.

[115] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004, ISBN: 978-0521833783.

[116] A. Papoulis, "Pulse compression, fiber communications, and diffraction: A unified approach," *JOSA A*, vol. 11, no. 1, pp. 3–13, 1994.

[117] H. Wang, X. Yi, J. Zhang, and F. Li, "Extended study on equalization-enhanced phase noise for high-order modulation formats," *Journal of Lightwave Technology*, vol. 40, no. 24, pp. 7808–7813, 2022.

[118] I. Fatadin and S. J. Savory, "Impact of phase to amplitude noise conversion in coherent optical systems with digital dispersion compensation," *Optics express*, vol. 18, no. 15, pp. 16 273–16 278, 2010.

[119] A. F. Alfredsson, E. Agrell, H. Wymeersch, *et al.*, "Pilot-aided joint-channel carrier-phase estimation in space-division multiplexed multicore fiber transmission," *Journal of Lightwave Technology*, vol. 37, no. 4, pp. 1133–1142, 2018.

[120] G. Colavolpe, T. Foggi, E. Forestieri, and M. Secondini, "Impact of phase noise and compensation techniques in coherent optical systems," *Journal of Lightwave Technology*, vol. 29, no. 18, pp. 2790–2800, 2011.

[121] C. Jin, N. A. Shevchenko, Z. Li, S. Popov, Y. Chen, and T. Xu, "Nonlinear coherent optical systems in the presence of equalization enhanced phase noise," *Journal of Lightwave Technology*, vol. 39, no. 14, pp. 4646–4653, 2021.

[122] M. S. Neves, P. P. Monteiro, and F. P. Guiomar, "Enhanced phase estimation for long-haul multi-carrier systems using a dual-reference subcarrier approach," *Journal of Lightwave Technology*, vol. 39, no. 9, pp. 2714–2724, 2021.

[123] M. S. Neves, A. Lorences-Riesgo, C. S. Martins, *et al.*, "Carrier-phase recovery for coherent optical systems: Algorithms, challenges and solutions," *Journal of Lightwave Technology*, 2023.

[124] T. Verolet, G. Aubin, Y. Lin, *et al.*, "Mode locked laser phase noise reduction under optical feedback for coherent DWDM communication," *Journal of Lightwave Technology*, vol. 38, no. 20, pp. 5708–5715, 2020.

[125] X. Zhou, "An improved feed-forward carrier recovery algorithm for coherent receivers with M-qam modulation format," *IEEE Photonics Technology Letters*, vol. 22, no. 14, pp. 1051–1053, 2010.

[126] A. Spalvieri and L. Barletta, "Pilot-aided carrier recovery in the presence of phase noise," *IEEE Transactions on Communications*, vol. 59, no. 7, pp. 1966–1974, 2011.

[127] M. Magarini, A. Spalvieri, F. Vacondio, M. Bertolini, M. Pepe, and G. Gavioli, "Empirical modeling and simulation of phase noise in long-haul coherent optical transmission systems," *Optics Express*, vol. 19, no. 23, pp. 22 455–22 461, 2011.

[128] M. Al-Qadi, M. O'Sullivan, C. Xie, and R. Hui, "Phase noise measurements and performance of lasers with non-white FM noise for use in digital coherent optical systems," *Journal of Lightwave Technology*, vol. 38, no. 6, pp. 1157–1167, 2019.

[129] J. S. Suelzer, T. B. Simpson, P. Devgan, and N. G. Usechak, "Tunable, low-phase-noise microwave signals from an optically injected semiconductor laser with opto-electronic feedback," *Optics letters*, vol. 42, no. 16, pp. 3181–3184, 2017.

[130] F. Aflatouni, M. Bagheri, and H. Hashemi, "Design methodology and architectures to reduce the semiconductor laser phase noise using electrical feedforward schemes," *IEEE transactions on microwave theory and techniques*, vol. 58, no. 11, pp. 3290–3303, 2010.

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

[132] J. H. Ke, K. P. Zhong, Y. Gao, J. C. Cartledge, A. S. Karar, and M. A. Rezania, "Linewidth-tolerant and low complexity two-stage carrier phase estimation for dual-polarization 16-QAM coherent optical fiber communications," *Journal of lightwave technology*, vol. 30, no. 24, pp. 3987–3992, 2012.

[133] R. Farhoudi, A. Ghazisaeidi, and L. A. Rusch, "Performance of carrier phase recovery for electronically dispersion compensated coherent systems," *Optics express*, vol. 20, no. 24, pp. 26 568–26 582, 2012.

[134] X. Zhou, K. Zhong, Y. Gao, C. Lu, A. P. T. Lau, and K. Long, "Modulation format independent blind phase search algorithm for coherent optical square M-QAM systems," *Optics Express*, vol. 22, no. 20, pp. 24 044–24 054, 2014.

[135] L. M. Zhang and F. R. Kschischang, "Staircase codes with 6% to 33% overhead," *Journal of Lightwave Technology*, vol. 32, no. 10, pp. 1999–2002, 2014.

[136] V. Simon, A. Senst, M. Speth, and H. Meyr, "Phase noise estimation via adapted interpolation," in *GLOBECOM'01. IEEE Global Telecommunications Conference (Cat. No. 01CH37270)*, IEEE, vol. 6, 2001, pp. 3297–3301.

[137] T. Pfau, S. Hoffmann, and R. Noé, "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for M-qam constellations," *Journal of Lightwave Technology*, vol. 27, no. 8, pp. 989–999, 2009.

[138] M. Qiu, Q. Zhuge, M. Chagnon, *et al.*, "Digital subcarrier multiplexing for fiber nonlinearity mitigation in coherent optical communication systems," *Optics Express*, vol. 22, no. 15, pp. 18 770–18 777, 2014.

[139] D. Welch, A. Napoli, J. Bäck, *et al.*, "Digital subcarrier multiplexing: Enabling software-configurable optical networks," *Journal of Lightwave Technology*, vol. 41, no. 4, pp. 1175–1191, 2023.

[140] S. Bilal, C. Fludger, and G. Bosco, "Carrier phase estimation in multi-subcarrier coherent optical systems," *IEEE Photonics Technology Letters*, vol. 28, no. 19, pp. 2090–2093, 2016.

[141] M. P. Yankov, L. Barletta, and D. Zibar, "Low-complexity joint subcarrier phase noise compensation for digital multi-carrier systems," in *2017 European Conference on Optical Communication (ECOC)*, IEEE, 2017, pp. 1–3.

[142] D. Martinez, Y. Mori, H. Hasegawa, and K.-i. Sato, "Novel subcarrier multiplexing and subcarrier-synchronous phase estimation tolerant to laser phase noise," in *Signal Processing in Photonic Communications*, Optica Publishing Group, 2017, SpW1F–5.

[143] X. Yi, W. Shieh, and Y. Ma, "Phase noise effects on high spectral efficiency coherent optical OFDM transmission," *Journal of Lightwave Technology*, vol. 26, no. 10, pp. 1309–1316, 2008.

[144] A. V. Oppenheim, *Discrete-time signal processing.* Pearson Education India, 1999.

[145] E. C. Ifeachor and B. W. Jervis, *Digital signal processing: a practical approach.* Pearson Education, 2002.

[146] N. Chen and M. Okada, "Toward 6G internet of things and the convergence with RoF system," *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 8719–8733, 2020.

[147] G. Forecast, "Cisco visual networking index: Global mobile data traffic forecast update, 2017–2022," *Update*, vol. 2017, p. 2022, 2019.

[148] A. Soleimanzade, M. Ardakani, and H. Ebrahimzad, "Optimization of the non-linearity tolerant 4d geometric shaped constellations for optical fiber communication systems using neural networks," *Journal of Lightwave Technology*, pp. 1–11, 2023. DOI: `10.1109/JLT.2023.3317149`.

[149] P. Schulte and G. Böcherer, "Constant composition distribution matching," *IEEE Transactions on Information Theory*, vol. 62, no. 1, pp. 430–434, 2015.

Appendix A

Constant Power Constellation Grouping

A.1 Mean of Accumulated Power

Accepting the accumulated power model in (2.7), we can find the mean of the accumulated power as:

$$\begin{aligned}
\mu &= E\{P(t_0)\} \\
&= \sum_{i=1}^{N} E_{m_i,\tau_i}\{|A_i(t_0-\tau_i)|^2\} \\
&= \sum_{i=1}^{N} \Bigg(\sum_{k=-q}^{q} E_{m_i,\tau_i}\Big\{|m_{ik}|^2 g^2(t-\tau_i-kT\Big\} \\
&+ 2\sum_{k_0=-q}^{q}\sum_{k_1=k_0+1}^{q} E_{m_i,\tau_i}\Big\{Re\{m_{ik_0}m_{ik_1}^*\}\times g(t-\tau_i-k_0T)g(t-\tau_i-k_1T)\Big\}\Bigg) \\
&= N\left(\sum_{k=-q}^{q} E_\tau\Big\{g^2(t-\tau-kT)\Big\}\right),
\end{aligned} \tag{A.1}$$

where τ and τ_i are uniform random variables in the interval of $[0,T)$, representing the initial launched time delay. Furthermore, we assume the pulse shape, g(t), is non-zero in $[-qT,qT]$. Therefore:

$$\mu = N\left(\sum_{k=-q}^{q}\int_0^T \frac{1}{T} g^2(t-kT)dt\right). \tag{A.2}$$

In the derivation of the mean, we assume the symbols, m_{ik}, are symmetrically distributed and have the following properties:

$$\begin{aligned} E\{|m_{ik}|^2\} &= 1, \\ E\{m_{ik}\} &= 0. \end{aligned} \tag{A.3}$$

A.2 Variance of Accumulated Power

Based on the signal model in (2.6), variance of the power in channel i can be found as:

$$\begin{aligned} &Var\{P_i(t_0)\} \\ &\quad = E\{P_i(t_0)^2\} - \Big(E\{P_i(t_0)\}\Big)^2 \\ &\quad = E\bigg\{\Big(\sum_{k=-q}^{q} |m_{ik}|^2 g^2(t-\tau_i-kT) \\ &\qquad + 2\sum_{k_0=-q}^{q}\sum_{k_1=k_0+1}^{q} \mathrm{Re}\{m_{ik_0}m_{ik_1}^*\} \times g(t-\tau_i-k_0T)g(t-\tau_i-k_1T)\Big)^2\bigg\} - \mu_i^2 \\ &\quad = \zeta_i \sum_{k=-q}^{q} E\{g^4(t-\tau_i-kT)\} + 2\sum_{k_0=-q}^{q}\sum_{k_1=k_0+1}^{q} E\{g^2(t-\tau_i-k_0T)g^2(t-\tau_i-k_1T)\} \\ &\qquad + 4\sum_{k_0=-q}^{q}\sum_{k_1=k_0+1}^{q} E\{\mathrm{Re}^2\{m_{ik_0}m_{ik_1}^*\} \times g^2(t-\tau_i-k_0T)g^2(t-\tau_i-k_1T)\} - \mu_i^2, \end{aligned} \tag{A.4}$$

where $\zeta_i = E\{|m_i|^4\}$ and $\mu_i = E\{P_i(t_0)\}$ and the pulse shape, g(t), is non-zero in $[-qT, qT]$. Given the fact that used, designed constellations are symmetric with power 1, we can show that:

$$\mathrm{Re}^2\{m_{ik_0}m_{ik_1}^*\} = \frac{1}{2}. \tag{A.5}$$

Therefore, accumulated power variance, σ^2, for N channels can be expressed as:

$$\begin{aligned} \sigma^2 = \zeta N &\sum_{k=-q}^{q} \int_0^T \frac{1}{T} g^4(t-kT)dt \\ &+ 4N \sum_{k_0=-q}^{q}\sum_{k_1=k_0+1}^{q} \int_0^T \frac{1}{T} g^2(t-k_0T)g^2(t-k_1T)dt - \frac{\mu^2}{N}, \end{aligned} \tag{A.6}$$

where $\zeta = E\{\frac{1}{N}\sum_{i=1}^{N}|m_i|^4\}$ and μ is the accumulated power mean defined in (A.2).

A.3 Auto-correlation of Accumulated Power

Assuming accumulated power satisfies wide-sense stationary (WSS) condition, we can find the auto-correlation as:

$$\begin{aligned} R(l) &= \sum_{i=1}^{N}\Bigg(E\bigg\{\Big(P_i(t-\tau_i)-\mu_i\Big)\Big(P_i(t-\tau_i-l)-\mu_i\Big)^*\bigg\}\Bigg) \\ &= \zeta N\sum_{k=-q}^{q}E_{\tau_i}\Big\{g^2(t-\tau-kT)g^2(t-\tau-l-kT)\Big\} \\ &\quad+\sum_{k_0=-q}^{q}\sum_{k_1=k_0+1}^{q}E_{\tau}\bigg\{\Big(g(t-\tau-k_0T)g(t-\tau-l-k_1T) \\ &\quad+g(t-\tau-k_1T)g(t-\tau-l-k_0T)\Big)^2\bigg\}-\frac{\mu^2}{N}. \end{aligned} \tag{A.7}$$

Since τ and τ_i are uniform random variables in the interval of $[0,T)$ modeling channels' initial launched time, we can simplify (A.7) as:

$$\begin{aligned} R(l) &= \zeta N\sum_{k=-q}^{q}\int_0^T\frac{1}{T}g^2(t-kT)g^2(t-l-kT)dt \\ &\quad+\sum_{k_0=-q}^{q}\sum_{k_1=k_0+1}^{q}\int_0^T\frac{1}{T}\Big(g(t-k_0T)g(t-l-k_1T) \\ &\quad+g(t-k_1T)g(t-l-k_0T)\Big)^2dt-\frac{\mu^2}{N}, \end{aligned} \tag{A.8}$$

where μ is the accumulated power mean (A.1) and the pulse shape, g(t), is non-zero in $[-qT,qT]$.

A.4 Entropy for a Different Power Level

Let us say we are going to generate the power level of $N-M+\sum_{i=1}^{M}|x_i|^2$ using N paired constellations, where $\sum_{i=1}^{M}|x_i|^2$ is a realizable power level using M constellations. We further assume the entropy of generating the power level of $\sum_{i=1}^{M}|x_i|^2$ using M constellations is h. Then, for generating the desired power level it is enough to generate the power level of $\sum_{i=1}^{M}|x_i|^2$ using the first M constellations and generate $N-M$ power level using the $N-M$ remained constellations. In this case for total entropy, we have:

$$\begin{aligned} & H(X_1, X_2, \cdots, X_N \Big| \sum_{i=1}^{N}|X_i|^2 = N - M + \sum_{i=1}^{M}|x_i|^2) = \\ & \quad H(X_1, X_2, \cdots, X_M \Big| \sum_{i=1}^{N}|X_i|^2 = \sum_{i=1}^{M}|x_i|^2) + \\ & \quad H(X_{M+1}, X_{M+2}, \cdots, X_N \Big| \sum_{i=1}^{N}|X_i|^2 = N - M) = \\ & \quad h + (N-M)\eta, \end{aligned} \tag{A.9}$$

where we use (2.29) to find the entropy of the constellations with the second power level. Calculating the entropy per channel,(2.25), yields the following for this system:

$$H_p(N) = \frac{h + (N-M)\eta}{N}. \tag{A.10}$$

For large N, we have $\lim_{N\to\infty} H_P = \eta$. Therefore, creating any realizable power level using countable constellations results in entropy of $N\eta$.

Appendix B

Optimal Power Spectral Density

B.1 Band-limited Optimal Pulse

In this appendix, we demonstrate that (3.18) minimizes (3.11) subject to (3.12). To start, substituding (3.15) into (3.11), we have:

$$\begin{aligned} P_{\text{NLIN}} = \frac{16}{27}\gamma^2 L_{\text{eff}}^2 P_g^3 & \sum_{i_1=-N/2}^{N/2} \sum_{i_2=-N/2}^{N/2} \sum_{i_3=-N/2}^{N/2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \\ & \left(U^2(f_1 - i_1 f_0) + \frac{T}{1+\beta} - \frac{T}{1+\beta} \int_{-\infty}^{+\infty} U^2(f_3) df_3 \right) \\ & \left(U^2(f_2 - i_2 f_0) + \frac{T}{1+\beta} - \frac{T}{1+\beta} \int_{-\infty}^{+\infty} U^2(f_4) df_4 \right) \\ & \left(U^2(f_1 + f_2 - f - i_3 f_0) + \frac{T}{1+\beta} - \frac{T}{1+\beta} \int_{-\infty}^{+\infty} U^2(f_5) df_5 \right) \\ & \left(U^2(f) + \frac{T}{1+\beta} - \frac{T}{1+\beta} \int_{-\infty}^{+\infty} U^2(f_6) df_6 \right) \\ & \rho(f_1, f_2, f)\chi(f_1, f_2, f) df_1 df_2 df. \end{aligned} \tag{B.1}$$

Upon substituting (3.16) into (B.1) and applying the criterion described in (3.17), we can write:

$$
\begin{aligned}
&\frac{\delta}{\delta\epsilon} P_{\text{NLIN}}\Big|_{\epsilon\,=0} = \\
&\qquad \frac{16}{27}\gamma^2 L_{\text{eff}}^2 P_g^3 \sum_{i_1=-N/2}^{N/2}\sum_{i_2=-N/2}^{N/2}\sum_{i_3=-N/2}^{N/2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \\
&\qquad \left(\frac{U_{\text{opt}}(f_1-i_1f_0)\eta(f_1-i_1f_0)-\frac{T}{1+\beta}\int_{-\infty}^{+\infty}U_{\text{opt}}(f_3)\eta(f_3)df_3}{U_{\text{opt}}(f_1-i_1f_0)+\frac{T}{1+\beta}-\frac{T}{1+\beta}\int_{-\infty}^{+\infty}U_{\text{opt}}(f_3)df_3}+\right. \\
&\qquad \frac{U_{\text{opt}}(f_2-i_2f_0)\eta(f_2-i_2f_0)-\frac{T}{1+\beta}\int_{-\infty}^{+\infty}U_{\text{opt}}(f_4)\eta(f_4)df_4}{U_{\text{opt}}(f_2-i_2f_0)+\frac{T}{1+\beta}-\frac{T}{1+\beta}\int_{-\infty}^{+\infty}U_{\text{opt}}(f_4)df_4}+ \\
&\qquad \frac{U_{\text{opt}}(f')\eta(f')-\frac{T}{1+\beta}\int_{-\infty}^{+\infty}U_{\text{opt}}(f_5)\eta(f_5)df_5}{U_{\text{opt}}(f')+\frac{T}{1+\beta}-\frac{T}{1+\beta}\int_{-\infty}^{+\infty}U_{\text{opt}}(f_5)df_5}+ \\
&\qquad \left.\frac{U_{\text{opt}}(f)\eta(f)-\frac{T}{1+\beta}\int_{-\infty}^{+\infty}U_{\text{opt}}\eta(f_6)df_6}{(U_{\text{opt}}(f)+\frac{T}{1+\beta}-\frac{T}{1+\beta}\int_{-\infty}^{+\infty}U_{\text{opt}}(f_6)df_6}\right) \\
&\qquad \left(U_{\text{opt}}(f_1-i_1f_0)+\frac{T}{1+\beta}-\frac{T}{1+\beta}\int_{-\infty}^{+\infty}U_{\text{opt}}(f_3)df_3\right) \\
&\qquad \left(U_{\text{opt}}(f_2-i_2f_0)+\frac{T}{1+\beta}-\frac{T}{1+\beta}\int_{-\infty}^{+\infty}U_{\text{opt}}(f_4)df_4\right) \\
&\qquad \left(U_{\text{opt}}(f')+\frac{T}{1+\beta}-\frac{T}{1+\beta}\int_{-\infty}^{+\infty}U_{\text{opt}}(f_5)df_5\right) \\
&\qquad \left(U_{\text{opt}}(f)+\frac{T}{1+\beta}-\frac{T}{1+\beta}\int_{-\infty}^{+\infty}U_{\text{opt}}(f_6)df_6\right) \\
&\qquad \rho(f_1,f_2,f)\chi(f_1,f_2,f)df_1df_2df,
\end{aligned}
\tag{B.2}
$$

where $f' = f_1 + f_2 - f - i_3 f_0$ is defined to save room. Clearly to make (B.2) equal to 0 for any $\eta(f)$, it is enough to have:

$$U_{\text{opt}}(f) = 0. \tag{B.3}$$

Please note that we utilize the fact that due to the dense WDM assumption, $f_0 = (1+\beta)/T$, shifted versions of $U_{\text{opt}}(f)$ by integer multiples of f_0 span the entire interval over which the integral is defined. By substituting $U_{\text{opt}}(f)$ into (3.15), we have:

$$G_{\text{opt}}(f) = \frac{T}{1+\beta} \qquad \text{for} \qquad |f| < \frac{1+\beta}{2T}, \tag{B.4}$$

and thus (3.18) is the functional that minimizes (3.11).

B.2 Band-limited Zero-ISI Pulse with Fast Decay Model

In this appendix, we show that to avoid unbounded ISI error the condition in (3.22) is enough. To guarantee a bounded ISI error in case of timing jitter, we should have:

$$|g(t)| \leq \frac{\lambda T}{\pi t}. \tag{B.5}$$

Applying this condition on (3.20) we have:

$$\left| \sin\left(\frac{\pi t}{T}\right) \cdot \int_0^{\frac{\beta}{2T}} X(f)\cos(2\pi f t) df \right| \leq \lambda. \tag{B.6}$$

Starting from the left side of (B.6), we have:

$$\begin{aligned} &\left| \sin\left(\frac{\pi t}{T}\right) \cdot \int_0^{\frac{\beta}{2T}} X(f)\cos(2\pi f t) df \right| \\ &\leq \left| \int_0^{\frac{\beta}{2T}} X(f)\cos(2\pi f t) df \right| \\ &\leq \int_0^{\frac{\beta}{2T}} \left| X(f)\cos(2\pi f t) \right| df \\ &\leq \int_0^{\frac{\beta}{2T}} \left| X(f) \right| df \\ &\leq \lambda, \end{aligned} \tag{B.7}$$

where the last inequality holds true because of the condition (3.22). Therefore, condition (3.22) is enough to guarantee the bounded ISI error in case of timing jitter.

B.3 Numerical Values for the Optimal Pulse Shape

To produce the optimal pulse shape, one can solve the optimization problem expressed in (3.27). Table B.1 and Table B.2 provide the results for $X_{\text{opt}}(f)$ for some values of β considering a typical 3000 km fiber with parameters expressed in Subsection 3.5.1. In all cases $\lambda = 2$. By inserting these values in (3.20), we can calculate $G_{\text{opt}}(f)$.

j	$\beta = 0.1$	$\beta = 0.2$	$\beta = 0.3$	$\beta = 0.4$	$\beta = 0.5$
1	-13.916	-1.754	-1.201	-5.436	-0.043
2	-4.878	-2.771	-1.687	-6.820	-0.114
3	-9.890	-2.322	-0.956	-2.437	-0.235
4	-6.455	-0.386	-1.886	1.189	-0.092
5	-10.746	-2.458	-1.209	3.288	-0.090
6	-9.623	-3.235	-1.605	1.952	-0.231
7	-9.699	-1.244	-1.647	0.510	-0.221
8	-19.952	-1.593	-0.984	-0.680	-0.174
9	-16.603	-2.903	-1.971	-1.516	-0.401
10	-24.246	-3.075	-1.508	-1.765	-0.467
11	-22.904	-4.143	-4.014	-1.648	-0.663
12	-27.822	-10.427	-4.843	-1.358	-0.697
13	-30.003	-11.456	-7.426	-2.676	-0.823
14	-29.000	-21.877	-10.945	-5.249	-1.411
15	-19.674	-40.433	-19.315	-7.842	-2.108
16	-9.040	-43.607	-55.380	-12.459	-7.082
17	3.304	-46.047	-16.757	-40.406	-7.926
18	36.981	-0.270	-0.000	-16.648	-42.026
19	182.313	-0.000	-0.000	-0.000	-15.196
20	441.852	400.000	266.667	200.000	160.000

Table B.1: Numerical values for the optima $X_j(f)$ and $\beta = 0.1$ to 0.5

j	$\beta = 0.6$	$\beta = 0.7$	$\beta = 0.8$	$\beta = 0.9$	$\beta = 1$
1	2.032	−1.052	0.525	−0.516	−0.000
2	3.034	−0.000	0.000	−0.000	−0.001
3	1.157	−0.000	0.000	0.621	0.001
4	0.000	−0.000	0.000	0.473	−0.001
5	0.000	0.000	−0.000	−0.001	−0.000
6	0.000	0.000	−1.381	−0.706	0.002
7	0.000	2.333	−1.094	−0.798	−0.004
8	0.000	1.328	−0.000	0.000	0.002
9	0.000	0.000	2.704	1.905	0.001
10	−0.000	0.000	0.000	0.183	−0.001
11	−0.000	0.000	0.000	−0.607	−0.002
12	−0.000	0.000	0.790	−1.180	0.003
13	−3.530	−0.000	0.000	−2.686	−0.000
14	−7.059	−0.000	0.000	3.211	0.001
15	−6.077	−0.000	−0.000	4.461	−0.011
16	−5.261	−5.788	−0.000	−2.759	0.036
17	−19.584	−6.065	−4.321	−1.974	−0.156
18	−8.488	−22.203	−9.172	−3.426	0.776
19	−0.000	−7.748	−21.532	−18.679	−4.276
20	110.444	96.338	83.479	66.920	43.629

Table B.2: Numerical values for the optima $X_j(f)$ and $\beta = 0.6$ to 1