2016

Efficient Modeling and Simulation of Wavelength Division Multiplexing Dual Polarization QPSK Optical Fiber Transmission

Siva Kumar Raju Nadimpalli
Minnesota State University Mankato

Efficient Modeling and Simulation of Wavelength Division Multiplexing Dual Polarization QPSK Optical Fiber Transmission

By

Siva Kumar Raju Nadimpalli

A Thesis Submitted in Partial Fulfillment of the Requirements for

Master of Science

In

Electrical Engineering

Minnesota State University

Mankato, Minnesota

MAY 2016

Date: 04/13/2016

Title: Efficient Modeling and Simulation of Wavelength Division Multiplexing Dual Polarization QPSK Optical Fiber Transmission

Student's Name: Siva Kumar Raju Nadimpalli

This thesis has been examined and approved by the following members of the student's committee.

Dr. Qun Zhang (Chairperson)

Dr. Xuanhui Wu

Dr. Hyekyung Min

ACKNOWLEDGEMENT

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study. First and foremost, my utmost gratitude to my advisor Dr. Qun Zhang whose seriousness and encouragement I will never forget. Dr. Zhang has been my inspiration as I hurdle all the obstacles related from code development to result collection in the completion of optical communication related research work.

Besides my advisor, I would like to thank Dr. Xuanhui Wu, Dr. Hyekyung Min for serving as members in my thesis committee and for their insightful comments and encouragement.

Last but not the least, I would like to thank my grandfather Raghupathi Raju Uddharaju, my parents, SubbaRaju Nadimpalli and Vijaya Lakshmi Nadimpalli, and my brother Prasada Raju Nadimpalli, and my cousin Rama Krishna Raju Uddharaju for all their love and support.

I would like to dedicate my thesis to my Family

ABSTRACT

Due to enormous growth in communications, wavelength division multiplexing (WDM) systems are popular because these systems allow us to expand the capacity of the networks without laying more optical fiber cables. In this thesis, we have systematically derived the coupled nonlinear Schrödinger (CNLS) equations, including a consistent definition of the complex envelope, Fourier transform, the state of polarization, and derivation under the engineering notation. After a discussion of coarse step based second order symmetrized split-step Fourier (SSSF) simulation method, which is applicable to the numerical solution of the CNLS equations, an analytical step-size selection based local error method is applied to the WDM optical fiber communication systems. With systematical simulation study of both standard single mode fiber (SSMF) fiber links and true-wave reduced slope (TWRS) fiber links. It is found that similar to the single channel systems, the global simulation accuracy for the vector propagation can be satisfied using the local error bound (LEB) obtained from a scalar propagation model for the same global error over a large range of simulation accuracy and differential group delay (DGD). Furthermore, carefully designed numerical simulations are used to show that the proposed local error method leads to higher computational efficiency compared to other prevalent step-size selection schemes in vector WDM simulations. The scaling of the global simulation error with respect to the number of optical fiber spans is demonstrated, and global error control for multi-span WDM simulations is proposed.

TABLE OF CONTENTS

Chapter 1

Introduction

1.1 Introduction to coherent optical fiber communication system with dual polarization

Optical communications systems use light to transmit information from source to destination from a few kilometers to several hundreds of kilometers. The research and development in optical fiber communication systems were started in the beginning of 1970s. These early systems used the intensity modulation of lasers and the modulated optical signal transmitted through an optical fiber was converted into photocurrent by a photodiode, in which the photocurrent is proportional to the optical power received. This combination of transmitter and receiver is called intensity modulation and direct detection (IMDD). The advantage of IMDD lies in that the receiver sensitivity is independent of the carrier phase and the state of polarization of the incoming signal, which is arbitrary in real systems. The state of polarization (SOP) will be explained in detail in the next chapter.

The above mentioned direct detection is formally termed as noncoherent detection in the context of communication system engineering, where the receiver estimates the input data based on measured signal energy. In intensity modulated on-off keying or binary orthogonal signaling, a bit "1" is represented by a pulse and bit "0" is represented by the

absence of a pulse. In the simplest communication channel model, the channel is assumed to add the additive white Gaussian noise (AWGN) to the signal field thus the signal is corrupted. Under this assumption, the bit error probability for binary orthogonal signaling scheme with noncoherent detection is $P_b = \frac{1}{2}e^{-\varepsilon_b/2N_0}$ where ε_b/N_0 is called signal-to-noise ratio per bit, or SNR of the communication system [1].

With signal phase information utilized, we can enhance the receiver sensitivity. This is via the means of coherent detection where the received signal is interfered with a local oscillator (LO) in order to extract the phase information of the signal. Compared to the intensity modulated on-off keying with incoherent direct detection, the phase modulated signal with coherent detection can have similar bit error probability performance even with larger noise power. Coherent detection increases receiver sensitivity which leads to power efficiency, since, the receiver detects even the weak power signals. Eventually, this leads us to use less number of repeaters in the optical fiber communication (OFC) system.

We start with an introduction to the simplest phase modulated signaling i.e., the binary phase shift keying (BPSK), also called as binary antipodal signaling (or 2-PSK). In this modulation scheme, we transmit signal pulse train including signals modulated with two different phases, 0 (bit 1) and 180 (bit 0) degrees. For this signaling scheme, the bit error probability is $P_b = Q\left(\sqrt{2\,\varepsilon_b/N_0}\right)$, assume the channel is AWGN [1]. The bit error probability for binary orthogonal signaling scheme is $P_b = Q\left(\sqrt{\varepsilon_b/N_0}\right)$. So for coherent detection in terms of power efficiency, binary orthogonal signaling is inferior to binary antipodal signaling by a factor of 2, or equivalently by 3dB [2].

The above introduction has been focusing on the receiver sensitivity or energy efficiency. Now we explain that the spectrum efficiency can be increased by using high order modulation. Quadrature phase shift keying (QPSK) or 4-PSK is one of the most common and widely used modulation formats in practice because it can be viewed as the superposition of two independent BPSKs thus they can be separated into in-phase (I-channel) and quadrature phase (Q-channel) BPSK components. The energy efficiency of QPSK is the same as the BPSK, but the spectrum efficiency is doubled for QPSK when compared to BPSK, i.e., a single QPSK signal pulse carries 2 information bits compared to only 1 bit carried by the BPSK. This is why we don't see BPSK utilized in practical optical fiber communication transceivers.

For large values of M, the arrangement of M constellation points on a circle becomes progressively less energy efficient, in the case of M-PSK modulation format. Specifically, with the doubling of M, the distance between the neighboring constellations points is nearly halved, so they are more susceptible to noise. Therefore, for *M-PSK* the required increase of energy to noise density ratio with the increase of M, and doubling of M requires the increase of *SNR* by 6 dB to achieve similar error performance thus results in almost 6 dB loss in energy efficiency. If we need data rates beyond the data rates offered by 8-psk it is common to move to Quadrature amplitude modulation(QAM) since the spacing between adjacent points are larger so they are less prone to noise compared to PSK. Note that QAM is a form of modulation which is a combination of both phase modulation and amplitude modulation. In M-QAM for large M, doubling M results in a 3 dB power penalty in signal power, compared to 6 dB power penalty in M-PSK [3].

So far we discussed about different modulation schemes in terms of their energy efficiency. In currently deployed optical fiber communication systems, the most favorable detection technique for achieving high spectral efficiency while maximizing power (or SNR) efficiency, is coherent detection with polarization multiplexing. Here the symbol decisions can be made based on in-phase and quadrature phase channels in the two field polarizations, which allows information to be encoded in phase, amplitude, and polarization degrees of freedom. In exchange, for such advantages, coherent receivers are sensitive to phase and SOP of the incoming signal. A digital signal processing (DSP) based receiver allows us to compensate all kinds of transmission impairments and recover full information of the electric field.

The traditional communication channel capacity given by Shannon [17] is

$$C = W \log_2(1 + SNR) \tag{1.1}$$

where W is the optical filter bandwidth (in Hertz), SNR is the signal to noise power ratio, and C is the channel capacity (in bits/second). According to Shannon precise communication can be achieved if we send information at a speed lower than or equal to C bits/second. But by using polarization division multiplexing (PDM), channel capacity becomes doubled $(2C)$, since we are using both x- and y- optical fields to carry information. In PDM, both polarizations carries independent multilevel modulated streams, which improves overall spectral efficiency. So PDM is a very efficient method to double the spectral efficiency of a transmission system. Polarization of light has defined with two kinds of conventions, which along with corresponding mathematical representations will be explained briefly in the next chapter (Chapter 2). The high

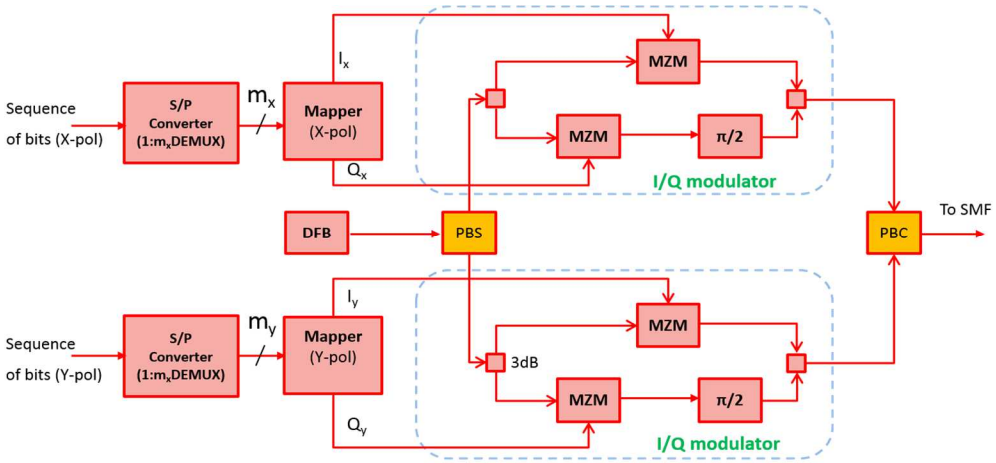
complexity of the current communication systems is a driving force behind the extensive use of simulation [5]. So it is crucial to enhance the computational efficiency of waveform level simulation of optical signal propagation through the dispersive and nonlinear single mode fiber [6].The electric field of plane linearly polarized CW waves propagating in the z-direction is represented as [7]

$$E(z,t) = E_x(z,t) + E_y(z,t) = e_x E_{0x}\cos(\omega t - kz) + e_y E_{0y}\cos(\omega t - kz + \delta) \quad (1.2)$$

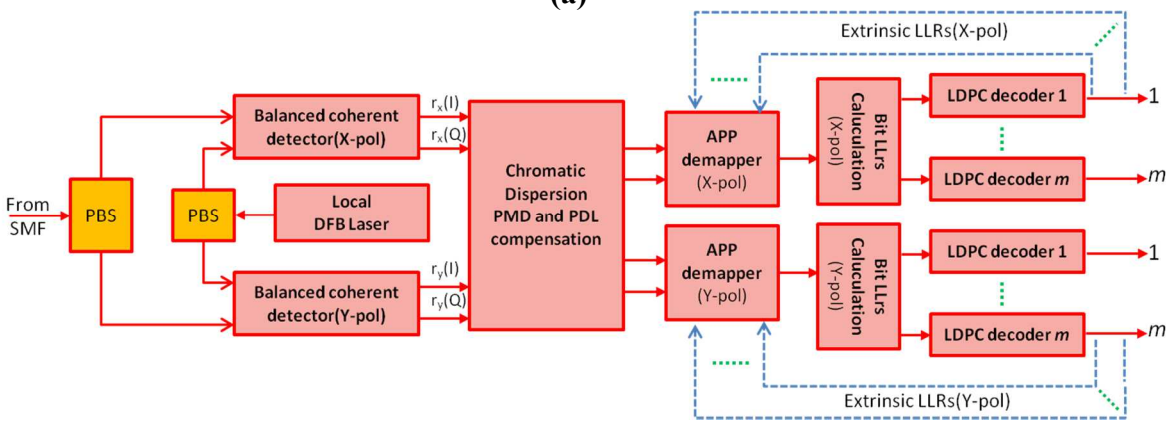
where $k = 2\pi/\lambda$ is the magnitude of the wave propagation vector (λ-wavelength), and δ is the relative phase difference between two orthogonal polarizations. Here we denoted the x-polarization state as E_x and y-polarization state as E_y, and we assumed that the initial phase in E_x is 0.

We use a QAM transceiver to illustrate the polarization multiplexed transceivers with coherent detection, as in Figure 1.1(a). The CW distributed feedback (DFB) semiconductor laser is separated into two orthogonal polarizations by a polarization beam splitter (PBS). The independent 2-D data streams are multiplexed together by a PBC and transmitted through an optical fiber cable. Each polarization branch contains a single I/Q modulator. In PDM applications, the QAM constellation coordinates in x- and y-polarizations are used, after the pulse shaping, as in-phase (I) component and quadrature (Q) inputs of the corresponding I/Q modulator. In I/Q modulator we have two Mach-Zehnder modulators (converts CW light into an optical bit stream), each allowing independent modulation of I and Q components of the optical electrical fields for both x- and y-polarizations [8]. The independent QAM streams are multiplexed together by a polarization beam combiner (PBC). In PDM both polarizations carry independent data

streams. This increases the overall spectral efficiency. In the transmitter configuration shown in Figure 1.1(a), we are using two independent mappers to drive two independent 2-D data streams through I/Q modulators. The corresponding coherent detector receiver architecture is shown in Figure 1.1(b). In this advanced receiver design we have a dispersion compensation block which compensates the effects of chromatic dispersion,



(a)



(b)

Figure 1.1: Block diagram of the coherent optical communication system based on polarization division multiplexing scheme (a) Transmitter configuration and (b) Receiver configuration. DFB: distributed feedback laser, PBS/C: polarization beam splitter/combiner, MZM: Mach-Zehnder modulator, LDPC: low density parity-check.

polarization mode dispersion (PMD), polarization dependent loss (PDL) and other channel impairments. We don't explain this receiver block in detail because it is beyond the scope of this thesis.

1.2 Mathematical Equations used for modeling the optical fiber propagation

In IMDD systems, the propagation of the scalar optical signal is governed by non-linear Schrödinger equation (NLSE), which can be given by [6, 9],

$$\frac{\partial A(z,t)}{\partial z} = \left(\widehat{D} + \widehat{N}(A(z,t))\right) A(z,t) \tag{1.3}$$

where $\widehat{D} = \frac{i}{2}\beta_2 \frac{\partial^2}{\partial t^2} - \frac{\alpha}{2}$ and $\widehat{N}(A(z,t)) = -i\gamma|A(z,t)|^2$ denotes linear and nonlinear operators respectively, and $A(z,t)$ is the complex envelope of optical pulse, t is the retarded time, z is the propagation distance and at the beginning of simulation $z = 0$, so $A(0,t) = \sqrt{|A_x(0,t)|^2 + |A_y(0,t)|^2}$ where signal phase is assumed to be zero. Parameters β_2, α, and γ represents fiber group velocity dispersion, attenuation, and nonlinearity respectively. The linear operator $\widehat{D}$ involves attenuation and chromatic dispersion. In a fiber optic system we have two different fibers, one, is the transmission fiber and the other is the dispersion compensation fiber (DCF). The value of β_2 is significantly different for transmission fiber and DCF. Fiber nonlinearity coefficient γ represents the fiber nonlinearity coefficient when the polarization effects of optical field is averaged during the propagation over the fiber, and it is often expressed as $\gamma = 8\,a_0/9$ where a_0 represents nonlinear coefficient that is measured from the propagation of an optical field aligned to either the fast or slow axis of a polarization maintaining fiber (PMF) with the same waveguide and doping profile [6]. We used engineering notation in

this thesis, but Max Born et.al used traditional notation while deriving the Schrodinger's equation [16, 18, 35], the difference between traditional physics and engineering notations along with Fourier transforms is elaborated in appendix A and the mathematical derivation of coupled nonlinear Schrödinger (CNLS) equations using engineering notation is given briefly in the next chapter (Chapter 2).

Coupled vector equations are needed for modeling dual polarization systems with coherent detection. The vector propagation model is governed by the coupled nonlinear Schrödinger (CNLS) equations, and the coarse step method is often used in the simulation of the vector propagation for practical fiber optic links [7]. With a coordinate system moving at the average group velocity of the two polarized signals, the following equations can model the vector signal propagation over any coarse step after omitting the nonlinear terms involving significant amount of phase mismatch due to the large fiber birefringence [6, 14, 25],

$$\frac{\partial A_x}{\partial z} + b' \frac{\partial A_x}{\partial t} - \frac{i}{2}\beta_2 \frac{\partial^2 A_x}{\partial t^2} + i a_0 (|A_x|^2 + \frac{2}{3}|A_y|^2) A_x = 0 \qquad (1.4\text{ a})$$

$$\frac{\partial A_y}{\partial z} - b' \frac{\partial A_y}{\partial t} - \frac{i}{2}\beta_2 \frac{\partial^2 A_y}{\partial t^2} + i a_0 (|A_y|^2 + \frac{2}{3}|A_x|^2) A_y = 0 \qquad (1.4\text{ b})$$

In the above equations (1.4 a) and (1.4 b), $A_x(z,t)$ and $A_y(z,t)$ models the complex envelope of the two polarized signals at distance z and retarded time t. Only considering the x-polarized field, for single wavelength channel systems we have $A_x(z,t) = \sum_{n=1}^{N} d_n\, a(t - nT_{sym})$ in our simulation model, where N is the number of symbols used in the model, d_n is the n^{th} data symbol, $a(t)$ is the pulse shaping function, and T_{sym} is

the symbol time interval. System bandwidth is approximately inversely proportional to T_{sym} in single channel systems [37].

1.3 The local error method: introduction and motivation

The NLSE governs the optical pulse propagation through the optical fiber. The accurate solution to (1.3) after one simulation step h can be given by the following equation [10],

$$A(z + h, t) = \exp[h(\hat{D} + \hat{N})]A(z, t) \tag{1.5}$$

For practical cases for currently deployed and studied optical fiber communication system, NLSE cannot be solved analytically because of the complex interplay between fiber nonlinearity and chromatic dispersion, though there is an exception under some special conditions in which inverse scattering method can be employed [20]. Therefore, a numerical method is needed for studying the nonlinear effects in the optical fiber. A large number of numerical methods are created for this purpose and has been divided into two broad categories known as finite difference and pseudospectral methods. In general, pseudospectral methods are faster compared to finite difference methods to achieve similar accuracy. The split-step Fourier (SSF) method is a kind of pseudospectral methods and in fact the most widely used numerical scheme for solving the NLSE due to ease of implementation and high computational efficiency [31, 32]. The SSF method has been used widely to solve the pulse propagation in the nonlinear dispersive medium. In general, by using the split-step method the long fiber is divided into many segments or steps with individual steps being usually small.

It is important to devise step-size selection rules i.e., ways of dividing the fiber into small steps because the step-size selection can significantly affect the simulation

efficiency. Many methods are available to select the step-size. Constant step-size method is one of the methods to implement SSF method. In this method, the whole length of the fiber link i.e., both the transmission fiber and DCF is divided into equal size steps in the transmission path (recall that we have discussed that there may be two different kinds of fibers i.e. transmission fiber and DCF used in an optical fiber system). The global accuracy can be improved only by increasing the number of steps. Another commonly used step-size selection method to implement SSF method is called walk-off method, where the step-size will be selected such that $|\beta_2| \times h(z) = C$ with C being a constant, where step-size $h(z)$ is a function of z when fibers with different $|\beta_2|$ are used. With walk-off method the step-size in the DCF is much less compared to that in transmission fiber [11]. The global accuracy can be improved by reducing C.

The order of SSF method is defined on the relation between global error and step-size. Note that for an n^{th} order SSF method the one step simulation error or local error (LE) is proportional to h^{n+1}, where h is the simulation step-size, and the global error (GE) or system simulation error is,

$$\text{Global error} \propto \quad h^n \tag{1.6}$$

New numerical methods in terms of step-size selection have been proposed to improve the computation efficiency or the simulation accuracy. Sinkin et al. have used an iterative method to compute each step-size for bounding **the one step simulation error or local error (LE).** In this method, the step-size is selected by bounding the local error in each step using a technique of doubling or halving the step-size and estimating the local error. In this paper, he also mentioned that his method is computationally efficient about the

order of two compared to walk-off, constant step methods at higher accuracies [21]. But there are several shortcomings associated with this method, in the range of global simulation error or system simulation error $10^{-3} - 10^{-1}$ which is the region of interest in optical fiber links where this method is performing poor compared to other methods. This is because the method used is a third order method. Furthermore, bounding the local error alone will not guarantee a comparable global simulation error or global error (GE) since the global error is not only controlled by LE but also depends on the total number of steps. Comparable global simulation accuracy for a system with varying parameter values can only be achieved by consecutively scaling LE and the step-size in determining the total number of steps, which was not studied in this method [11].

In a step-size optimization method proposed by Zhang et.al in [11], the above discussed problems are solved by using the second order symmetrized SSF (SSSF) method and also by deriving and using an analytical expression involving both system parameters and step-size to calculate one step simulation error. Based on this analytical expression, simple step-size scaling rules were derived to achieve comparable global error accuracy for different system parameter values. This step-size optimization method can achieve similar LE for all simulation steps so it is called local error method [11, 21]. The SSSF, a second order method, is the frequently used numerical method for the simulation of optical signal transmission through an optical fiber, again due to the fact that it is the most efficient SSF method in the region of interest in terms of simulation accuracy for optical fiber links [34].

With step-size optimized to achieve similar one-step simulation error or local error, the SSSF is more efficient the SSSF is a more efficient method of simulating scalar optical fiber propagation. That's why the study here is limited to the second order SSSF method. We emphasize that similar LE means that there is no computational waste from any simulation step, so for SSSF simulations, the local error method has higher computational efficiency than other step-size selection methods including the constant step-size method, walk-off method, and nonlinear phase shift method [11]. This means that for a certain GE level, the local error method leads to the least number of required simulation steps, compared to those other step-size selection methods. The study in [11] is limited to the simulation of signal propagation in scalar optical fiber channel.

Here we explain the analytical formula based local error method more in detail. We used a step splitting method in which the dispersion operation is calculated first, followed by nonlinearity and then followed by dispersion operation (denoted as the DND method).We use the following analytical formula to select the simulation step-size, approximately constant one step or local simulation error ($\Delta\xi$) can be achieved [10, 11],

$$\gamma P_{max}(z) D\, \Delta\lambda\, \Delta f h^2(z) = \Delta\xi \tag{1.7}$$

where D denotes the fiber dispersion parameter, $P_{max}(z)$ denotes the peak power at certain propagation distance z, and the step-size h is written as $h(z)$ since h is calculated on fly so that it will depend on z. Local error bound $\Delta\xi$ should be understood as a bound on the pulse-width error. Since in modern optical communication systems where the nonlinearity is moderate, signal bandwidth $\Delta\lambda$ and Δf do not change significantly during simulation. In order to estimate the step-size $h(z)$, the only parameter that need to be

calculated during the simulation is $P_{max}(z)$. In this method as we are going from the transmitter towards the receiver end the power decreases exponentially. As a result, per (1.7) the step-size goes on increasing along the transmission path in Local error method. We will compute the approximate solution by dividing a fiber into coarse steps, later we will compute the true solution by dividing the coarse step into very fine steps of equal length. The calculation of GE will be explained in Chapter 3.

In [32], the analytical step-size selection formula derived for the scalar optical fiber channel [11] was modified and proposed for single polarized optical signal propagation through the vector optical fiber channel. By using a scalar model based local error bound (LEB) finding software package [24], the LEB found for a certain global simulation error level can be used as the LEB for the vector model with single polarized input, and the same global simulation error can be satisfied for the vector case. In [9] we extended the above results to the dual polarization input or polarization multiplexed (PM) input, and the same conclusion is obtained. In [12] we verified for the PM signal that similar local error (LE) is indeed achieved for all simulation steps when using the analytical formula based step-size optimization method. Thus, we adopt the local error method as the name of the studied step-size optimization method for vector SSSF simulation in this work. This is well described in the vector simulation in Chapter 3. The computational efficiency of two other methods (constant step and walk-off) was compared with local error method in single-channel systems and verified that local error method requires 1.25 to 2 times less number of steps compared to other methods, Note that the number of FFT/IFFTs used per simulation step can be made same for all the three studied step-size selection

methods [36].

The results summarized in the above paragraph are limited to standard single mode fiber (SSMF) based systems. So we are mainly focusing on wavelength division multiplexing (WDM) systems in this thesis. Please note that WDM is a scheme in which multiple optical carriers (which are not in phase with each other) at different wavelength are modulated by using independent electrical bit streams and are then transmitted over the same fiber. For simulating real-world WDM systems, the fiber length is quite long, varying from few hundreds of kilometers for metropolitan systems to few thousands of kilometers for transoceanic systems. Furthermore, optical communication system design is a multidimensional optimization problem, requiring long fiber simulations to be repeated in multiple times for the interested design parameters to explore the design space. As a result, the computation burden during the system design is very high. So, there has been a strong interest in studying efficient numerical techniques for solving the NLSE.

For current WDM systems, multiple wavelength channels need to be used in the simulation, resulting in large simulation bandwidth. Long waveform, high sampling rate, and small simulation steps combined leads to high computational complexity and often results in time-consuming simulations. So it is significant to enhance the simulation efficiency for the purpose of efficient WDM system optimization and design. There are different notations in modern engineering and physics/optics and traditional physics/optics which may lead to confusion, so we added a unified notation for deriving NLSE in the next chapter (Chapter 2).

In this work, we extend the research from single channel system to WDM systems for dual polarization or polarization multiplexed (PM) vector input signals. We systematically investigate if the local error method is still applicable and also if the local error method is still more efficient than the walk-off and the constant step-size methods, with the significantly increased simulation bandwidth in WDM systems. Though the details are not included, we studied the pulse propagation in TWRS and SSMF fibers while average power is considered instead of peak power during the calculation of step-size in single channel single-span and multi-span systems with dispersion compensated and uncompensated links.

1.4 Outline

In this thesis we apply the local error method to the simulation of the polarization multiplexed signal propagation through dispersion compensated standard single mode fiber links and true-wave reduced slope fiber links. The goal is to validate the proposed local error method and its high computational efficiency as compared to other prevalent step-size selection methods. In Chapter 2 we explain different conventions regarding polarization and Poincaré sphere. Besides that we derived the coupled nonlinear Schrödinger equations using unified notation presented in appendix A. In appendix A, we briefly explain about various notations used for the complex envelope and Fourier transforms by modern engineering and traditional physics/optics. To facilitate derivations and theoretical explanations, we include the derivation of a linear model for the pulse propagation in the fiber is presented in appendix B. Also, in appendix C, we explain the calculation method for differential group delay. In Chapter 3, we explain basics of the

WDM fiber optic system, the scalar fiber propagation, and the vector propagation. In addition to that, we describe the simulation system setup and summarize results for single-span and multi-span dispersion compensated WDM systems. We also analyzed the computational efficiency of WDM single-span and multi-span systems via comparisons to results with constant step method and walk-off method. In Chapter 4, we conclude the thesis with proposed future work.

Chapter 2

A Formal Derivation of the Coupled Nonlinear Schrodinger Equations

2.1 Conventions regarding Polarization and Poincaré Sphere

Polarization of lightwave signal refers to the orientation of the electric field vector in the plane that is perpendicular to the direction of propagation [38]. Consider a lightwave propagating along the z-axis with its propagation direction modeled by a unit vector $\vec{k}$, then electric field intensity can be expressed by using x- and y- components as

$$\vec{E}(z,t) = \hat{x}E_x(z,t) + \hat{y}E_y(z,t) \tag{2.1}$$

with

$$E_x(z,t) = a_x \cos(\omega t - kz + \delta_x)\ ;\ E_y(z,t) = a_y \cos\big(\omega t - kz + \delta_y\big) \tag{2.2}$$

For communication purpose, the transmitter and receiver are placed at fixed points in space. Therefore, we often focus on a constant z value, and study signal dynamics with time t at that particular z. Now without the loss of generality, we assume $z = 0$ and $\delta = \delta_y - \delta_x$. We explain three different types of polarizations, linear polarization, circular polarization and elliptical polarization. For linear polarization the electric field of light is oscillating with time on a straight line in the transversal plane. For circular polarization the electric field rotates on a circle in the transversal plane, and can be viewed as being formed by two field orthogonal components which have equal amplitude and $\pi/2$ phase

difference, just as shown in (2.2). Figure 2.1 is used to illustrate the definition of circular lights. For general elliptical polarization the two field components will not have the same amplitude and furthermore, the phase difference can be arbitrary.

There are two different kinds of notations for circular polarization, one is right hand circular (RHC) polarization which will be used by modern engineering/physics and the other one is left hand circular (LHC) polarization used in traditional terminology. The engineering convention is adopted by Ulaby and Balanis which is used in our work [16, 35]. The other convention is adopted by Derickson, and Born & Wolf (traditional), which is opposite to the engineering notation [15, 18]. So the left hand circular polarization in one terminology becomes right hand circular polarization and vice versa, so there exists a lot of confusion because of two opposing conventions.

Polarization handedness is defined in terms of the rotation of electric field vector $\vec{E}(z_0, t)$ as a function of time in a fixed plane ($z = z_0$) orthogonal to the direction of propagation, which is opposite of the direction of rotation of $\vec{E}(z, t_0)$ as a function of distance at a fixed point in time ($t = t_0$) [16]. Per our engineering notation, the EM wave as shown in Figure 2.1(a), is called left-handed circularly polarized because, when the thumb of the left hand is pointed towards the direction of propagation (in the direction of $+z$) the other curled four fingers point in the direction of rotation of electric field $\vec{E}(z, t)$. In another words, a left handed rotation is defined when the thumb is pointing towards the wave propagation direction, or when the wave is viewed from the perspective of the transmitter. So for our engineering notation, the wave is from the view of the transmitter. Similarly, it is easy to see that with the same notation, the EM wave as shown in Figure

2.1(b) is right-handed circularly polarized. Note that in optical fiber communication measurement literature, traditional physics books, and classical optics books, the opposite notation is adopted, and there the thumb is pointing to a direction that is consistent to the view from the receiver.

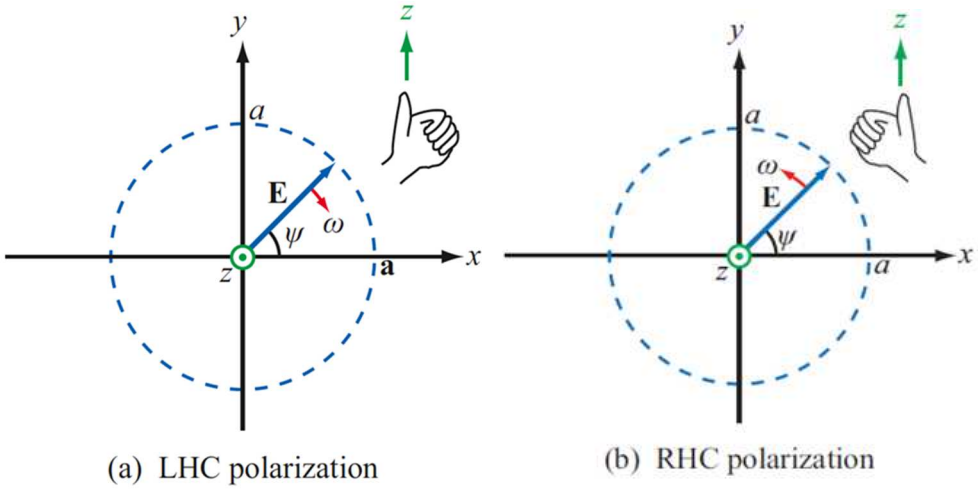


Figure 2.1: Circularly polarized plane waves propagating in the $+z$ direction.

With the notations explained, we elaborate the formulation for circularly polarized light following our engineering notation. For left hand circular polarization, we will have $\delta = \pi/2$ and $a_x = a_y = a$ so (2.1) becomes the following (2.3) assuming $\delta_x = 0$.

$$\vec{E}(z,t) = \Re e\left[\vec{E}(z)e^{i\omega t}\right] = \ = \frac{1}{2}(\hat{x}a\cos(\omega t - kz) - \hat{y}a\sin(\omega t - kz)) \qquad (2.3)$$

Note that $\vec{E}(z)$ is complex envelope (Appendix A), and $\vec{E}(z) = \hat{x}E_x(z) + \hat{y}E_y(z)$. Because we are concerned with time harmonic waves, the complex envelope $\vec{E}(z)$ does not contain t, so for a fixed location the complex envelope is simply a complex vector here.

Similarly, for right hand circular polarization ($\delta = -\pi/2$) and $a_x = a_y = a$, we have

$$\vec{E}(z,t) = \Re e\left[\vec{E}(z)e^{i\omega t}\right] = \frac{1}{2}(\hat{x}a\cos(\omega t - kz) + \hat{y}a\sin(\omega t - kz)) \quad (2.4)$$

The state of polarization of an optical field can be represented by using Jones vector and the Poincaré ellipse. The Jones vector illustrates the polarization of light in free space or another homogeneous isotropic medium, where light can be well described as transverse waves [18]. A monochromatic plane wave of frequency ω can be completely described by its scalar complex envelopes $E_x = A_x e^{i\delta_x}$ and $E_y = A_y e^{i\delta_y}$ at a fixed location e.g. $z = 0$ without the loss of generality. So polarized light can be represented by a complex vector with two elements specifying respectively the x- and y- components of the electric field for a particular point in space. The Jones vector has the form [15],

$$\underline{E} = \begin{bmatrix} A_x e^{i\delta_x} \\ A_y e^{i\delta_y} \end{bmatrix} \quad (2.5)$$

If we are interested in the state of polarization of a wave, then we can use the normalized Jones vector,

$$\underline{E}^n = \frac{1}{\sqrt{|A_x|^2 + |A_y|^2}} \begin{bmatrix} A_x e^{i\delta_x} \\ A_y e^{i\delta_y} \end{bmatrix} 1 \quad (2.6)$$

which satisfies the condition $\underline{E}^{n*} . \underline{E}^n = 1$.

We need three independent parameters to describe elliptical polarization. For example, we need the two magnitudes of the electric field components in both axes (x and y) and phase difference δ. In another representation, we will need 1. major axis, 2. minor axis, and 3. angle ψ which represents the orientation of the ellipse. Due to the wide usage in measurement equipment, for practical purposes it is better to characterize the SOP of a signal by parameters having same physical dimensions, i.e., the Stokes

parameters. These parameters were introduced by G. G. Stokes in 1852. The Stokes parameters of a plane monochromatic light are given by [18],

$$S_0 = |A_x|^2 + |A_y|^2$$

$$S_1 = |A_x|^2 - |A_y|^2$$

$$S_2 = 2|A_x||A_y|\cos\delta$$

$$S_3 = 2|A_x||A_y|\sin\delta \tag{2.7}$$

Note that in (2.7) there are only three independent variables among the four Stokes parameters, due to the fact that they are related by the following equation,

$$S_0{}^2 = S_1{}^2 + S_2{}^2 + S_3{}^2 \tag{2.8}$$

The state of polarization can be visualized on a sphere with the use of the representation of Stokes parameters. This sphere is called the Poincaré sphere, illustrated by Figure 2.2. The figure clearly shows that parameter S_0, which is proportional to intensity of light, represents the radius of the sphere, and parameters S_1, S_2 and S_3 are the Cartesian coordinates. Note that in the Poincaré sphere representation, S_0 is normalized to 1. A point on the Poincaré sphere represents a state of polarization. For point P in Figure 2.2 that represents a general elliptical polarization state, 2χ and 2ψ are the spherical angular coordinates. The angle ψ $(0 \leq \psi < \pi)$ represents the orientation of the ellipse, and the angle χ $(-\pi / 4 \leq \chi \leq \pi / 4)$ characterizes the ellipticity. The factor of two before ψ represents that any polarization is identical with an ellipse rotated by 180^o, whereas the two before χ represents that the ellipse is identical from one with semi-axis lengths swapped followed by 90^o rotation. We list the following useful equations,

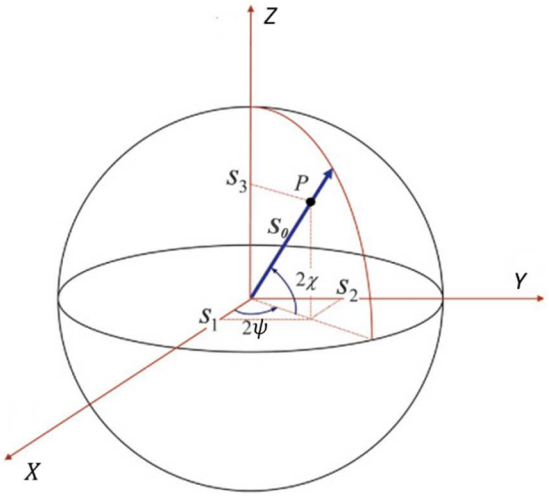


Figure 2.2: Stokes parameters represented on a Poincaré sphere.

$$S_1 = S_0 \cos 2\chi \cos 2\psi$$

$$S_2 = S_0 \cos 2\chi \sin 2\psi$$

$$S_3 = S_0 \sin 2\chi \tag{2.9}$$

For every possible state of polarization of a plane monochromatic wave with a specific intensity ($S_0 = constant$), then there exists a corresponding point on the Poincaré sphere shown in Figure 2.3. The value of each Stokes parameter will be in the range from -1 to 1 due to the normalization of S_0. For a linear polarized light the value of δ is zero or an integral multiple of π, so the Stokes parameter S_3 becomes zero. Therefore, linear polarization is represented by points on equatorial plane. For circular polarization we have $|A_x| = |A_y|$ and $\delta = \pi/2$ or $-\pi/2$. In traditional physics and optics, the right handed circular polarization is represented by North Pole ($S_1 = S_2 = 0$ and $S_3 = S_0$) and left handed circular polarization on the South Pole ($S_1 = S_2 = 0$ and $S_3 = -S_0$) on the Poincaré sphere. We note that for engineering notation, we have opposite results.

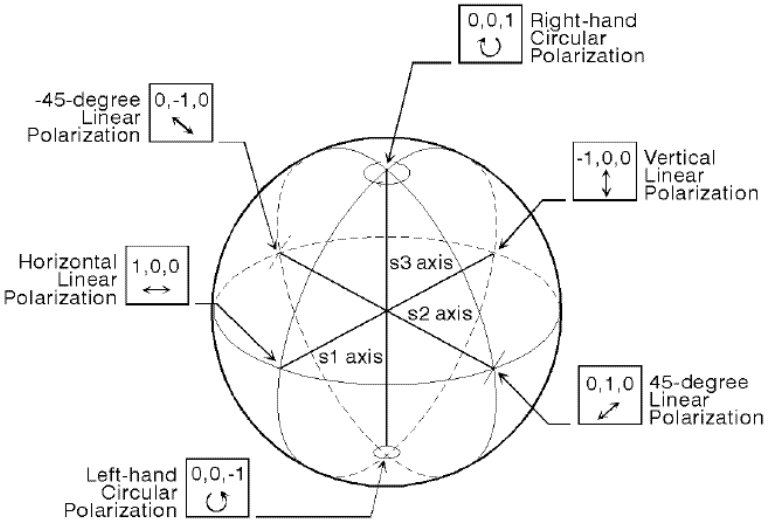


Figure 2.3: Poincaré sphere representing Polarization states.

2.2 Nonlinear Schrödinger equation derivation using perturbation method and Engineering notation

In order to derive non-linear Schrödinger equation (NLSE) many people used the traditional physics notation but here we are using the engineering notations for the derivation [14, 18]. We discussed about the difference between these two notations before and it is further explained in appendix A. The derivation of a linear model for pulse propagation is explained in Appendix B. The study of all nonlinear effects in optical fibers comprises the use of short pulses with widths ranging from ~ 10 ns to 10 fs. When such small width optical pulses propagate through a fiber, both dispersive and nonlinear effects modify their shape and spectra. Our work here in this section mainly refers to [6, 22]. It is necessary to make several simplifying assumptions to derive the NLSE using perturbation method. First, we assume that $\vec{P}_{NL}$ is a small perturbation to $\vec{P}_L$. This is acceptable because nonlinear changes in the refractive

index are $< 10^{-6}$ in practice. Second, the optical field is supposed to maintain its polarization throughout the fiber (linear polarization) so a scalar approach is valid. Third, optical fiber is assumed to be ideal, meaning that the fiber is in perfect cylindrical shape. We assume electric field being only in the x-direction. Then the following expressions are obtained,

$$\vec{E}(\vec{r},t) = \frac{1}{2}\hat{x}\left[E(\vec{r},t)e^{i\omega_0 t} + c.c.\right] \tag{2.10}$$

$$\vec{P}_L(\vec{r},t) = \frac{1}{2}\hat{x}\left[P_L(\vec{r},t)e^{i\omega_0 t} + c.c.\right] \tag{2.11}$$

$$\vec{P}_{NL}(\vec{r},t) = \frac{1}{2}\hat{x}[P_{NL}(\vec{r},t)e^{i\omega_0 t} + c.c.] \tag{2.12}$$

Note that in the above equations, the complex envelope contains both $\vec{r}$ and t. So at a location of interest ($\vec{r}$ being constant), we are consider a general time domain pulse, not a monochromatic light anymore.

In general, the induced polarization is not an immediate response of the electric field, so it in general leads to chromatic dispersion. The polarization component can be expressed as

$$\vec{P}_L(\vec{r},t) = \varepsilon_0 \hat{x} \int_{-\infty}^{\infty} \chi^{(1)}(t-t')\,\vec{E}(\vec{r},t)dt'$$

$$\vec{P}_L(\vec{r},t) = \frac{\varepsilon_0}{2\pi}\hat{x} \int_{-\infty}^{\infty} \chi^{(1)^F}(\omega)E^F(\vec{r},\omega-\omega_0)e^{-j(\omega-\omega_0)t}\,d\omega$$

$$\vec{P}_L(\vec{r},\omega-\omega_0) = \varepsilon_0\,\chi^{(1)^F}(\omega)\,E^F(\vec{r},\omega-\omega_0) \tag{2.13}$$

By assuming instantaneous nonlinear polarization response, the equation turns into

$$\vec{P}_{NL}(\vec{r},t) = \varepsilon_0\,\chi^{(3)}\vec{E}(\vec{r},t)\vec{E}(\vec{r},t)\vec{E}(\vec{r},t) \tag{2.14}$$

$$\vec{P}_{NL}(\vec{r},t) = \varepsilon_0\,\varepsilon_{NL}E(\vec{r},t) \tag{2.15}$$

where $\varepsilon_{NL} = \frac{3}{4}\chi^{(3)}|E(\vec{r},t)|^2$ is assumed to be a constant.

By including nonlinear polarization term, following the derivation process of (B.8), the following equation will be obtained

$$\nabla^2\vec{E} - \frac{1}{c^2}\frac{\partial^2\vec{E}}{\partial t^2} = \mu_0\frac{\partial^2\vec{P}_L}{\partial t^2} + \mu_0\frac{\partial^2\vec{P}_{NL}}{\partial t^2} \tag{2.16}$$

By using a linear polarization model as in (2.13), from (2.16), we obtain

$$\nabla^2\vec{E}^F + \varepsilon(\omega)k_0^2\vec{E}^F = 0 \tag{2.17}$$

where $\varepsilon(\omega) = 1 + \chi^{(1)^F}(\omega) + \varepsilon_{NL}$ (2.18)

By generalizing (B.7), we will get the following expression

Denoting $\varepsilon(\omega) = \left(\bar{n} + i\frac{\bar{\alpha}}{2k_0}\right)^2$ with $\bar{n} = n + \bar{n}_2|E|^2$ and $\bar{\alpha} = \alpha + \alpha_2|E|^2$ due to $\alpha_2 \ll \alpha$ and $\bar{n}_2 \ll n$, omitting $\bar{n}_2\alpha_2|E|^4$, $\bar{n}_2{}^2|E|^4$, $\bar{n}_2\alpha|E|^2$, we obtain $\bar{n}_2$and α_2 as follows

$$\bar{n}_2 = \frac{3}{8n}\mathrm{Re}\left(\chi^{(3)}_{xxxx}\right),\ \alpha_2 = \frac{3\omega_0}{4nc}\mathrm{Im}\left[\chi^{(3)}_{xxxx}\right] \tag{2.19}$$

By assuming $E^F(\vec{r},\omega-\omega_0) = F(x,y)A^F(z,\omega-\omega_0)e^{i\beta_0 z}$, where A^F is a slow varying function of the transmission distance z, we have $\frac{\partial^2 A^F}{\partial t^2} \approx 0$ (slow varying envelope approximation). Then (2.18) becomes,

$$\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} + \left[\varepsilon(\omega)k_0^2 - \overline{\beta^2}\right]F = 0 \tag{2.20a}$$

$$2i\beta_0\frac{\partial A^F}{\partial z} + \left(\overline{\beta^2} - \beta_0^2\right)A^F = 0 \tag{2.20b}$$

The dielectric constant $\varepsilon(\omega)$ in (2.20a) can be approximated by

$$\varepsilon(\omega) = (n+\Delta n)^2 \approx n^2 + 2n\Delta n \tag{2.20c}$$

where Δn is a small perturbation given by

$$\Delta n = \bar{n}_2|E|^2 + \frac{i\bar{\alpha}}{2k_0} \tag{2.20d}$$

Now we include the effect of Δn in (2.20a). In the first-order perturbation theory, Δn does not affect the modal distribution $F(x,y)$. However, the eigenvalue $\bar{\beta}$ becomes $\bar{\beta} = \beta(\omega) + \Delta\beta$ where

$$\Delta\beta(\omega) = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k_0 \Delta n|F|^2\,dx\,dy}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|F|^2\,dx\,dy} = \frac{n_2|A|^2 k_0}{S_{eff}} + \frac{i\alpha}{2} \tag{2.20e}$$

In (2.20e), the effective mode area of the fiber is given by $S_{eff} = \frac{\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|F(x,y)|^2\,dx\,dy\right]^2}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|F(x,y)|^4\,dx\,dy}$.

Note that here A is normalized so that $|A|^2$ has unit power. For fundamental mode HE_{11}, we have with terms up to the group velocity dispersion (GVD) included

$$\frac{\partial A^F}{\partial z} = i[\beta + \Delta\beta - \beta_0]A^F$$

$$= i\left[\beta_0 + (\omega - \omega_0)\beta_1 + \frac{1}{2}(\omega - \omega_0)^2\beta_2 + \frac{n_2|A|^2 k_0}{S_{eff}} + \frac{i\alpha}{2} - \beta_0\right]A^F \tag{2.21a}$$

which leads to

$$\frac{\partial A}{\partial z} + \beta_1\frac{\partial A}{\partial t} + \frac{i}{2}\beta_2\frac{\partial^2 A}{\partial t^2} + \frac{\alpha}{2}A = ia_0|A|^2 A \text{ with } a_0 = \frac{n_2\omega_0}{cS_{eff}} \tag{2.21b}$$

where a_0 is the nonlinearity coefficient. Note that to obtain (2.21b) we assumed the use of a polarization maintaining fiber that leads to enhanced nonlinear Kerr effect. For the scalar model applicable to the single mode optical fiber, where polarization state gets scrambled due to the imperfect core shape and imperfection in the fiber core materials, we need to modify the nonlinearity coefficient to $\gamma = 8a_0/9$. This is a key result for the Manakov equation [14] for which we will have more discussion in the next Section.

When the pulse width is less than 0.1 ps, the higher order dispersion and higher order nonlinearities need to be included [6].

2.3 Derivation of the Coupled Nonlinear Schrödinger Equations

In this section, we derive the coupled NLSE using a unified notation based on appendix A. This section mainly refers to [6]. As I explained in Chapter 1, optical signal experiences chromatic dispersion during propagation through an optical fiber. In a fiber under ideal conditions (perfect cylindrical symmetry and stress-free), a mode excited with polarization in the x-direcion will not couple to the mode with orthogonal y-polarization state. But in real fibers, they are not stress free so small departure from cylindrical symmetry results in mixing of two polarization states due to the breaking of mode degeneracy. The mode-propagation constant β is slightly different for modes polarized in x- and y-direcions. If a material or material system has a difference between β_x and β_y, then it is birefringent, and in the case of optical fiber propagation, signal will experience modal dispersion and it is termed as polarization mode dispersion (PMD). The strength of modal birefringence is given by a dimensionless parameter [6].

$$B_m = \frac{|\beta_x - \beta_y|}{k_0} = |n_x - n_y| \tag{2.22}$$

where n_x and n_y are refractive indices. For a fixed value of B_m, the two modes exchange their power periodically while propagating inside the fiber with period given by $L_B = \frac{2\pi}{|\beta_x - \beta_y|} = \frac{\lambda}{B_m}$. The length L_B is called beat length. The axis for which the mode refractive index is smaller is called the fast axis because the group velocity is larger for light polarized in this axis. The axis with large mode refractive index is called as slow axis. In

standard optical fibers, B_m is not constant along the fiber because of core shape and departure of cylindrical symmetry due to anisotropic stress. As a result, the beat length is can be viewed as being random inside a long optical fiber.

If we assume that the axial component E_z of the electromagnetic field is very small, then it can be ignored compared with the transverse components. Another assumption for the following derivations is that the optical fiber is again assumed to be a polarization maintaining fiber (PMF). The real long fiber can be viewed as may PMF segments concatenated together. The electrical field associated with randomly polarized optical wave can be given as

$$\vec{E}(\vec{r},t) = \frac{1}{2}\left(\hat{x}E_x(\vec{r},t) + \hat{y}E_y(\vec{r},t)\right)e^{i\omega_0 t} + c.c. \tag{2.23}$$

where $\hat{x}$ and $\hat{y}$ are unit vectors, $E_x(\vec{r},t)$ and $E_y(\vec{r},t)$ are the complex amplitudes of x- and y-polarization components. The carrier frequency is ω_0, and $c.c.$ stands for complex conjugate. In the following, we use a shortened notation. For example, for $E_x(\vec{r},t)$ and $E_y(\vec{r},t)$ we use E_x and E_y. In an isotropic medium like silica glass, only three elements are independent to one another, the third order susceptibility can be given as [4]

$$\chi^{(3)}_{ijkl} = \chi^{(3)}_{xxyy}\delta_{ij}\delta_{kl} + \chi^{(3)}_{xyxy}\delta_{ik}\delta_{jl} + \chi^{(3)}_{xyyx}\delta_{il}\delta_{jk} \tag{2.24}$$

where δ_{ij} is the Kronecker delta function defined such that $\delta_{ij} = 1$ when $i = j$ and zero otherwise. Using this result in (2.23), the nonlinear polarization $\vec{P}_{NL}$ can be written as

$$\vec{P}_{NL}(\vec{r},t) = \frac{1}{2}\left(\hat{x}P_x^{NL} + \hat{y}P_y^{NL}\right)e^{i\omega_0 t} + c.c \tag{2.25}$$

where $P_x^{NL}$and P_y^{NL} are given by

$$P_i^{NL} = \frac{3\varepsilon_0}{4}\sum_j\left(\chi^{(3)}_{iijj}E_iE_jE_j^* + \chi^{(3)}_{ijij}E_jE_iE_j^* + \chi^{(3)}_{ijji}E_jE_jE_i^*\right) \tag{2.26}$$

where $i, j = x$ or y. From the rotational symmetry of an isotropic medium, we have the following relation [4]

$$\chi_{xxxx}^{(3)} = \chi_{xxyy}^{(3)} + \chi_{xyxy}^{(3)} + \chi_{xyyx}^{(3)} \tag{2.27}$$

The relative magnitudes of all the three components in (2.27) depend on the physical mechanism that contributes to $\chi^{(3)}$. For silica fibers all the three components have equal magnitude. If they are assumed to be the same then P_x^{NL} and P_y^{NL} in (2.26) can be written as

$$P_x^{NL} = \frac{3\varepsilon_0}{4}\chi_{xxxx}^{(3)}\left[\left(|E_x|^2 + \frac{2}{3}\left|E_y\right|^2\right)E_x + \frac{1}{3}(E_x^* E_y)E_y\right], \tag{2.28a}$$

$$P_y^{NL} = \frac{3\varepsilon_0}{4}\chi_{xxxx}^{(3)}\left[\left(\left|E_y\right|^2 + \frac{2}{3}|E_x|^2\right)E_y + \frac{1}{3}(E_y^* E_x)E_x\right], \tag{2.28b}$$

The nonlinear contribution Δn_x to the refractive index is controlled by the term proportional to E_x in (2.28a). Writing $P_j^{NL} = \varepsilon_0\, \varepsilon_j^{NL} E_j$ and relating to

$$\varepsilon_j = \varepsilon_j^L + \varepsilon_j^{NL} = \left(n_j^L + \Delta n_j\right)^2 \tag{2.29}$$

where n_j^L is linear part of the refractive index ($j = x, y$), the nonlinear contributions Δn_x and Δn_y are given by

$$\Delta n_x = \bar{n}_2\left(|E_x|^2 + \frac{2}{3}\left|E_y\right|^2\right) \text{ and } \Delta n_y = \bar{n}_2\left(\left|E_y\right|^2 + \frac{2}{3}|E_x|^2\right) \tag{2.30}$$

where $\bar{n}_2$ is the nonlinearity index coefficient defined in (2.19). In the above equation the first term is responsible for the self-phase modulation (SPM). The second term is responsible for cross-phase modulation (XPM) because the nonlinear phase shift on one polarization component depends on the intensity of the other polarization component. The presence of this term generates nonlinear coupling between the field components E_x and E_y. The nonlinear contributions Δn_x and Δn_y are generally not equal so it creates

nonlinear birefringence whose magnitude will depend on the intensity and SOP of the incident light.

Assuming that there is no significant affect to the fiber mode due to the nonlinearities, the transverse dependence of E_x and E_y can be factored out using

$$E_j\,(\,\vec{r},\mathrm{t}) = F(x,y)A_j(z,\mathrm{t})e^{i\beta_{0j}z} \tag{2.31}$$

where $F(x,y)$ is the spatial distribution of the single mode supported by the fiber, $A_j(z,\mathrm{t})$ is the slowly varying amplitude and β_{0j} is the corresponding propagation constant for $j = x$ and y. The slowly changing amplitudes A_x and A_y has to follow the set of two coupled mode equations [6],

$$\frac{\partial A_x}{\partial z} + \beta_{1x}\frac{\partial A_x}{\partial t} + \frac{i\beta_2}{2}\frac{\partial^2 A_x}{\partial t^2} + \frac{\alpha}{2}A_x = i\gamma'\left(|A_x|^2 + \frac{2}{3}|A_y|^2\right)A_x + \frac{i\gamma'}{3}A_x^* A_y^2 e^{-2i\Delta\beta z} \tag{2.32}$$

$$\frac{\partial A_y}{\partial z} + \beta_{1y}\frac{\partial A_y}{\partial t} + \frac{i\beta_2}{2}\frac{\partial^2 A_y}{\partial t^2} + \frac{\alpha}{2}A_y = i\gamma'\left(|A_y|^2 + \frac{2}{3}|A_x|^2\right)A_y + \frac{i\gamma'}{3}A_y^* A_x^2 e^{2i\Delta\beta z} \tag{2.33}$$

where $\Delta\beta = \beta_{0x} - \beta_{0y} = \left(\frac{2\pi}{\lambda}\right)B_\mathrm{m} = 2\pi/L_\mathrm{B}$ and $\gamma' = \frac{9\gamma}{8} = a_0$ (2.34)

In (2.34), $\Delta\beta$ is related to linear birefringence of the fiber. In (2.32) and (2.33), β_{1x} and β_{1y} represent the group delay per unit length. The linear or modal birefringence gives us two different group velocities for the two polarization components because β_{1x} is different from β_{1y} in general. On the other hand, the parameters β_2 and γ are assumed to be same for both polarization components.

Menyuk [14] showed that the coupled nonlinear Schrödinger (CNLS) equations

apply to practical optical fibers with random birefringence, where the birefringence is large so the beat lengths will be typically around 10 to 100 m, and the length scale on which the birefringence orientation varies from 0.3 to 300 m. When we ignore the effect of nonlinearity, the large but quickly changing birefringence in fiber leads to PMD effect. But nonlinearity cannot be ignored, because the length of one segment transmission optical fiber can be up to several hundreds of kilometers. In order to deal with both nonlinearity and randomly varying birefringence Menyuk et.al averaged the polarization state of the light over the Poincaré sphere, and they showed that the CNLS reduces to the Manakov equation.

The key idea for Manakov-PMD equation from CNLS equation is to restrict the linear motion of the signal's center frequency on the Poincaré sphere. Since the movement of other frequencies in the signal and variations due to nonlinearity are slow compared to this point's motion on the sphere, we can take long computational steps. This actually justifies the coarse step method [25]. By considering the combined physical effects, the propagation of light pulses can be described by the coupled nonlinear Schrödinger equations [14],

$$i\frac{\partial \underline{A}(z,t)}{\partial z} + b\Sigma\underline{A} + ib'\Sigma\frac{\partial \underline{A}}{\partial t} - \frac{1}{2}\beta''\frac{\partial^2 \underline{A}}{\partial t^2} + n_2 k_0\left[\frac{5}{6}\left|\underline{A}\right|^2\underline{A} + \frac{1}{6}\left(\underline{A}^\dagger\sigma_3\,\underline{A}\right)\sigma_3\underline{A} + \frac{1}{3}\vec{B}\right] = 0 \quad (2.35)$$

or equivalently

$$i\frac{\partial \underline{A}(z,t)}{\partial z} + b\Sigma\underline{A} + ib'\Sigma\frac{\partial \underline{A}}{\partial t} - \frac{1}{2}\beta''\frac{\partial^2 \underline{A}}{\partial t^2} + n_2 k_0\left[\left|\underline{A}\right|^2\underline{A} - \frac{1}{3}\left(\underline{A}^\dagger\sigma_2\underline{A}\right)\sigma_2\underline{A}\right] = 0 \quad (2.36)$$

which are written in a form that is applicable for a fiber with arbitrarily varying birefringence orientation. Note that (2.35) and (2.36) are coupled equations and they each involves two scalar equations. Here, $\underline{A} = \left[A_x\ A_y\right]^T$ or $[A_1\ A_2]^T$ is a column vector with

the complex envelopes of the two polarization components as its elements and † represents the complex conjugate transpose. Vector $\vec{B}$ in (2.35) follows $\vec{B} = [A_1^* A_2^2 \quad A_1^2 A_2^*]^T$, where $*$ represents complex conjugation. The matrix

$$\Sigma = \sigma_3 \cos(2\theta) + \sigma_1 \sin(2\theta) \tag{2.37}$$

is defined in terms of Pauli's matrices

$$\mathrm{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.38}$$

Because the orientation of axes of birefringence is a weak function of frequency in the concerned bandwidth, the same Σ is used in the second and the third terms of (2.36). The parameter β'' in equation (2.36) doesn't have any subscript, since we are assuming $\beta_1'' \approx \beta_2'' = \beta''$.

Substituting all the parameters of (2.38) in (2.37) then the following result is obtained.

$$\Sigma = \sigma_3 \cos 2\theta + \sigma_1 \sin 2\theta = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} \tag{2.39}$$

Insert (2.39) to (2.36), the scalar form of equation (2.36) becomes,

$$i\frac{\partial A_1}{\partial z} + b[(\cos 2\theta)\, A_1 + (\sin 2\theta)\, A_2] + ib'\left[(\cos 2\theta)\, \frac{\partial A_1}{\partial t} + (\sin 2\theta)\, \frac{\partial A_2}{\partial t}\right]$$

$$-\frac{1}{2}\beta''\frac{\partial^2 A_1}{\partial t^2} + n_2 k_0 \begin{bmatrix} \frac{5}{6}(|A_1|^2 + |A_2|^2)A_1 \\ + \frac{1}{6}(|A_1|^2 - |A_2|^2)A_1 + \frac{1}{3}A_1^* A_2^2 \end{bmatrix} = 0$$

$$i\frac{\partial A_2}{\partial z} + b[(\sin 2\theta)\, A_1 - (\cos 2\theta)\, A_2] + ib'\left[(\sin 2\theta)\,\frac{\partial A_1}{\partial t} - (\cos 2\theta)\,\frac{\partial A_2}{\partial t}\right]$$

$$-\frac{1}{2}\beta''\frac{\partial^2 A_2}{\partial t^2} + n_2 k_0 \begin{bmatrix} \frac{5}{6}(|A_1|^2 + |A_2|^2)A_2 \\ + \frac{1}{6}(|A_2|^2 - |A_1|^2)A_2 + \frac{1}{3}A_2^* A_1^2 \end{bmatrix} = 0 \tag{2.40}$$

After simplification, we obtain the following equations

$$i\frac{\partial A_1}{\partial z} + b[(\cos 2\theta)\, A_1 + (\sin 2\theta)\, A_2] + ib'\left[(\cos 2\theta)\,\frac{\partial A_1}{\partial t} + (\sin 2\theta)\,\frac{\partial A_2}{\partial t}\right]$$

$$-\frac{1}{2}\beta''\frac{\partial^2 A_1}{\partial t^2} + n_2 k_0\left[|A_1|^2 A_1 + \frac{2}{3}|A_2|^2 A_1 + \frac{1}{3}A_1^* A_2^2\right] = 0$$

$$i\frac{\partial A_2}{\partial z} + b[(\sin 2\theta)\, A_1 - (\cos 2\theta)\, A_2] + ib'\left[(\sin 2\theta)\,\frac{\partial A_1}{\partial t} - (\cos 2\theta)\,\frac{\partial A_2}{\partial t}\right]$$

$$-\frac{1}{2}\beta''\frac{\partial^2 A_2}{\partial t^2} + n_2 k_0\left[|A_2|^2 A_2 + \frac{2}{3}|A_1|^2 A_2 + \frac{1}{3}A_2^* A_1^2\right] = 0 \tag{2.41}$$

When the rotation angle $\theta = 0$, then equations in 2.41 can be written as,

$$-i\frac{\partial A_1}{\partial z} = bA_1 + ib'\frac{\partial A_1}{\partial t} - \frac{1}{2}\beta''\frac{\partial^2 A_1}{\partial t^2} + n_2 k_0\left[|A_1|^2 A_1 + \frac{2}{3}|A_2|^2 A_1 + \frac{1}{3}A_1^* A_2^2\right]$$

$$-i\frac{\partial A_2}{\partial z} = -bA_2 - ib'\frac{\partial A_2}{\partial t} - \frac{1}{2}\beta''\frac{\partial^2 A_2}{\partial t^2} + n_2 k_0\left[|A_2|^2 A_2 + \frac{2}{3}|A_1|^2 A_2 + \frac{1}{3}A_2^* A_1^2\right] \tag{2.42}$$

Now we are ready to compare the derived CNLS equations from Menyuk (equations in (2.42) to that of Agarwal (equations (2.32) and (2.33)). Note in equations (2.32) and (2.33), Δβ is related to linear birefringence of the fiber, and β_{1x} and β_{1y} represent the group delay per unit length. The linear birefringence gives us two different group velocities for the two polarization components because β_{1x} is different from β_{1y} in general. On the other hand, the parameters β_2 and γ are assumed to be same for both polarization components.

Let us take $\Delta\beta = \beta_{0x} - \beta_{0y} = \beta_{1x} - \beta_{1y}$ and replacing this Δβ in (2.32) and (2.33).

Also, we use the following transformation,

$$\underline{A}(r,t,z) = e^{-\frac{\alpha}{2}z}\begin{bmatrix} A_x(r,t,z)e^{i\beta_{1x}z} \\ A_y(r,t,z)e^{i\beta_{1y}z} \end{bmatrix}, \tag{2.43}$$

$$A_1 = e^{-\frac{\alpha}{2}z}A_x(r,t,z)e^{i\beta_{1x}z} \tag{2.44}$$

$$A_2 = e^{-\frac{\alpha}{2}z}A_y(r,t,z)e^{i\beta_{1y}z} \tag{2.45}$$

Applying (2.44) to (2.32) we obtain the following equation after some manipulation such as dividing both sides of the equation by $e^{\frac{\alpha}{2}z}e^{-i\beta_{1x}z}$,

$$-\frac{i\partial A_1}{\partial z} = \beta_{1x}A_1 + i{\beta_x}'\frac{\partial A_1}{\partial t} - \frac{\beta_2}{2}\frac{\partial^2 A_1}{\partial t^2}$$

$$+\gamma'\left(|A_1|^2 + \frac{2}{3}|A_2|^2\right)A_1 + \frac{\gamma'}{3}A_1^*A_2^2 \tag{2.46}$$

Similarly, applying (2.45) to (2.33) we obtain

$$-\frac{i\partial A_2}{\partial z} = \beta_{1y}A_2 + i{\beta_y}'\frac{\partial A_2}{\partial t} - \frac{\beta_2}{2}\frac{\partial^2 A_2}{\partial t^2}$$

$$+\gamma'\left(|A_2|^2 + \frac{2}{3}|A_1|^2\right)A_2 + \frac{\gamma'}{3}A_2^*A_1^2 = 0 \tag{2.47}$$

we use the following transformation,

$$A_1 = A_1'e^{\frac{i(\beta_{1x}+\beta_{1y})z}{2}} \tag{2.48a}$$

$$A_2 = A_2'e^{\frac{i(\beta_{1x}+\beta_{1y})z}{2}} \tag{2.48b}$$

Applying (2.48) to (2.46) after some manipulation such as dividing both sides of the equation by $e^{\frac{i(\beta_{1x}+\beta_{1y})z}{2}}$ we obtain

$$-\frac{i\partial A_1'}{\partial z} = \frac{(\beta_{1x}-\beta_{1y})}{2}A_1' + i{\beta_x}'\frac{\partial A_1'}{\partial t} - \frac{\beta_2}{2}\frac{\partial^2 A_1'}{\partial t^2}$$

$$+\gamma'\left(|A_1'|^2 + \frac{2}{3}|A_2'|^2\right)A_1 + \frac{\gamma'}{3}{A_1'}^*{A_2'}^2 \tag{2.49a}$$

Similarly, applying (2.48) to (2.47) we obtain

$$-\frac{i\partial A_2'}{\partial z} = \frac{(\beta_{1y}-\beta_{1x})}{2}A_1' + i{\beta_y}'\frac{\partial A_2'}{\partial t} - \frac{\beta_2}{2}\frac{\partial^2 A_2'}{\partial t^2}$$

$$+\gamma'\left(|A_2'|^2 + \frac{2}{3}|A_1'|^2\right)A_2 + \frac{\gamma'}{3}{A_2'}^* {A_1'}^2 \qquad (2.49b)$$

Compare (2.49) with the CNLS equations used by Menyuk (equations in (2.42)), we can say that (2.42) is consistence with Agarwal CNLs and the relation between different parameters are tabulated below.

Menyuk	**Agarwal**
b	$\frac{\beta_{1x}-\beta_{1y}}{2}$
b'	${\beta_x}' = -{\beta_y}'$
β''	β_2
$n_2 k_0$	γ'

Table 1: Comparison of parameters

The coupled nonlinear Schrödinger equations used by Evangelides, describing propagation delay in a linear birefringent lossless fiber are [25]

$$-i\frac{\partial u}{\partial z_1} = \beta u + i\delta\frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial^2 u}{\partial t^2} + |u|^2 u + \frac{2}{3}|v|^2 u + \frac{1}{3}v^2 u^* \qquad (2.50a)$$

$$-i\frac{\partial v}{\partial z_1} = -\beta v - i\delta\frac{\partial v}{\partial t} + \frac{1}{2}\frac{\partial^2 v}{\partial t^2} + |v|^2 v + \frac{2}{3}|u|^2 v + \frac{1}{3}u^2 v^* \qquad (2.50b)$$

where u and v are the components of a normalized field corresponding to slow and fast axis respectively. This results in $\delta > 0$. In addition, 2β is the wavenumber difference, π/β is the beat length, and 2δ is the corresponding inverse group velocity difference, which is a derivative of 2β with respect to frequency ω.

Compare (2.50) with the CNLS equations used by Menyuk (equations in (2.42)), and the relation between different parameters are tabulated below.

Menyuk	**Evangelides**
b	β
b'	δ
β''	-1
$n_2 k_0$	1
A_1	u
A_2	v
z	z_1

Table 2: Comparison of parameters

Finally we conclude that, coupled NLS used by Menyuk and Evangelides are same, after scaling u, v and z_1 parameters in Evangelides equation (2.50). The scaling of these parameters are done as follows: $u = \sqrt{n_2 k_0} A_1$, $v = \sqrt{n_2 k_0} A_2$, and $z_1 = z|\beta''|$. In Evangelides's work, $\beta'' < 0$ is assumed, so we are taking the absolute value of β'' while scaling z_1, since we considered that the optical pulse propagates in the $+z$ direction.

Chapter 3

Study of a new simulation scheme for WDM vector fiber propagation

3.1 WDM Fiber Optic Communication system

We introduced the wavelength division multiplexing (WDM) systems and the concept of PDM in Chapter 1. The mathematical equations used for modeling the WDM system i.e., scalar and coupled vector Schrödinger equations, was explained in Section 1.2.

For clarity, we list the following coupled nonlinear Schrödinger (CNLS) equations again and under the engineering notation,

$$\frac{\partial A_x}{\partial z} + b' \frac{\partial A_x}{\partial t} - \frac{i}{2}\beta_2 \frac{\partial^2 A_x}{\partial t^2} + i a_0(|A_x|^2 + \frac{2}{3}|A_y|^2)A_x = 0 \qquad (3.0\text{ a})$$

$$\frac{\partial A_y}{\partial z} - b' \frac{\partial A_y}{\partial t} - \frac{i}{2}\beta_2 \frac{\partial^2 A_y}{\partial t^2} + i a_0(|A_y|^2 + \frac{2}{3}|A_x|^2)A_y = 0 \qquad (3.0\text{ b})$$

In the above equations (3.0 a) and (3.0 b), $A_x(z,t)$ and $A_y(z,t)$ represents the complex envelopes of the x- and y- polarized signals at distance z and retarded time t. The coarse step is assumed to be polarization maintaining, the group delay per unit length is represented by β_{1x} and β_{1y}, and b' is $(\beta_{1x} - \beta_{1y})/2$ rescaled to obtain the correct statistics of polarization mode dispersion (PMD) [14]. The fiber chromatic dispersion (CD) is modeled by $\beta_{2x} = \beta_{2y} = \beta_2$. The fiber attenuation is absorbed into A_x and A_y in (1.4). For each simulation step the second order symmetrized split-step Fourier (SSSF)

method is used in vector simulations [11]. After each coarse step, the simulated vector optical fields are scattered over the Poincaré sphere.

For the *x*-polarized field, in single channel systems we, when omitting its special direction, have $A_x(z,t) = \sum_{n=1}^{N} d_n\, a\left(t - nT_{sym}\right)$, where N is the number of symbols used in the simulation, d_n is the n^{th} data symbol, $a(t)$ is the pulse shaping function, and T_{sym} is the symbol time interval. System bandwidth is approximately inversely proportional to T_{sym} in single channel systems.

For WDM systems, $A_x(z,t)$ represents a total signal field that is superimposed by multiple WDM channels. Assuming that five channels are included in the simulation with the center channel as our channel of interest, we have

$$A_x(z,t) = \sum_{m=-2}^{2} \sum_{n=1}^{N} d_n^m\, a\left(t - nT_{sym}\right) \exp(j2\pi m F_{wdm} t) \tag{3.1}$$

where F_{wdm} is the channel spacing, and d_n^m is the n^{th} data symbol for the m^{th} wavelength channel. System bandwidth is approximately $(M-1)F_{wdm}$. Where M is the total number of channels. Since F_{wdm} is often much greater than $1/T_{sym}$, the total field of WDM signals has a much larger bandwidth compared to single channel signals. Correspondingly, the interference among these individual wavelength channels produces narrow spikes in time domain, and WDM systems is subject to more pronounced dispersion or walk-off effects. Note that in (3.1), the five wavelength channel numbers are denoted by $m = -2, -1, 0, 1,$ and 2 .

Here we use the following analytical formula to select the simulation step-size $h(z)$ for the coarse step,

$$\gamma P_{max}(z) h(z) (D \Delta\lambda \Delta f h(z))^2 = \Delta\xi \tag{3.2}$$

In (3.2), D is the fiber dispersion parameter, $\Delta\lambda$ or Δf is signal bandwidth, $P_{max}(z)$ denotes the peak optical power of the simulated waveform at propagation distance z. The step-size h is written as $h(z)$ since h depends on z. Parameter $\Delta\xi$ is called the local error bound and it represents a pulse-width error due to a finite h [11]. Note that for WDM systems, the bandwidth of the total field needs to be used for $\Delta\lambda$ or Δf in (3.2), so Δf will be approximately equal to the number of channels multiplying the channel spacing for simulations involving a large number of channels. In (3.2) the only parameter that need to be calculated during the simulation run is $P_{max}(z)$, which must be modified as follows for PM signals,

$$P_{max}(z) = \max_{t}\left(|A_x(z,t)|^2 + \left|A_y(z,t)\right|^2\right) \tag{3.3}$$

When a specific channel e.g., a center channel or an edge channel, is studied, the total field needs to be filtered to obtain the signal of interest for the GE computation. For vector fields, the nsd is modified for PM-QPSK system with coherent detection,

$$nsd(L) = \sqrt{\frac{\int(|A_x(L,t)-A_{xt}(L,t)|^2+\left|A_y(L,t)-A_{yt}(L,t)\right|^2)dt}{\int(|A_{xin}(0,t)|^2+\left|A_{yin}(0,t)\right|^2)dt}} \tag{3.4}$$

where $A_x(L,t)$ and $A_y(L,t)$ represent the simulated approximate waveforms at distance L, $A_{xt}(L,t)$ and $A_{yt}(L,t)$ represent "true" solutions using a finer step-size for each saved step, and $A_{xin}(0,t)$ and $A_{yin}(0,t)$ are the inputs. Equation (3.4) needs to be modified to represent the LE of a step $(z, z+h(z))$: L changes to $z+h(z)$, A_x and A_y becomes the saved approximate solution at $z+h(z)$, and A_{xt} and A_{yt} are obtained using the saved

approximate solution at z as the input and using finer step-size for simulation over $(z, z + h(z))$.

3.2 Simulation model and simulation package for the study of Local Error method

In scalar fiber propagation we have a simulation software developed by Zhang et.al, the simulation software consists of two portions. In the first portion, we can find the local error bound (LEB) for the scalar simulation to achieve a prescribed global error accuracy. In the second portion, we will use the LEB values obtained from scalar simulation software and validate them with the vector simulation software.

We note that for single channel systems the square root of the combined dual polarized signal power is used as the input to the scalar model in LEB finding simulations. In contrast, for WDM systems, we used the scaled total x-input field as input to the LEB finding package.

3.2.1 Simulation of scalar fiber propagation

A trial $\Delta\xi$ value is used as a starting point and more specifically is used for the first run of the waveform level SSSF simulation using our step-size selection local error method. In our simulation, we selected an LEB of 10^{-11} to simulate the "true" solution where the step-size is very small, so that the obtained optical field satisfies the required accuracy of the true field. The obtained output optical signal is compared with the "true" solution, and the normalized standard deviation (nsd) is used as a criterion for the global simulation error or GE.

$$nsd(z) = \|A_0(z,t) - A_t(z,t)\| / \|A_{in}(0,t)\| \tag{3.5}$$

In the above equation (3.5), $A_0(z,t)$ is the simulated scalar optical field for a specific $\Delta\xi$ and $A_t(z,t)$ is the "true" optical solution obtained by taking very small $\Delta\xi$ satisfies the required accuracy of the true field. Please note that $A_{in}(0,t)$ is the input scalar optical field (recall it is the scaled total x-input field) and $\|A(z,t)\|$ denotes the operation of $\sqrt{\int |A(z,t)|^2\, dt}$. The GE is defined by $nsd(L)$, where L is the total propagation distance.

If the calculated nsd is within range $\pm 10\%$ of the prescribed GE, then the trail $\Delta\xi$ corresponding to this global error will be saved, and this $\Delta\xi$ is used for rest of the system level simulations. Otherwise the new $\Delta\xi$ is calculated by considering the prescribed GE, obtained GE and the current trail $\Delta\xi$.

$$\text{new } \Delta\xi = \frac{\text{prescribed GE}}{\text{obtained GE}} \times \text{current or trail } \Delta\xi \tag{3.6}$$

We will calculate the nsd again with this newly generated $\Delta\xi$. This process will continue until we achieve the prescribed GE. Usually only a few iterations are needed during the LEB finding process. Once we got the prescribed GE the corresponding $\Delta\xi$ will be taken as the final GE, and it will be used for all simulation runs to achieve a consistent GE.

3.2.2 Simulation of vector fiber propagation

In the main simulation, we use the obtained $\Delta\xi$ from the scalar simulations for the vector simulations to achieve same GE (recall it is characterized by the nsd defined in (3.4)). Here we assume that the optical fields $A_{xin}(0,t)$ and $A_{yin}(0,t)$ are aligned with the fast and slow axis respectively of the optical fiber at the input. We assume random polarization coupling in the fiber with a PMD coefficient of $\varsigma = 0.1\ \text{ps}/\sqrt{km}$ for small PMD case (CASE 1) and $\varsigma = 0.1\ \text{ps}/\sqrt{km}$ for large PMD case (CASE 2). For each $\Delta\xi$

we perform 5 simulations each with a unique PMD realization and record all polarization scattering angles, step-sizes and other polarization related parameters. In order to obtain the true solution $A_{xt}(L,t)$ and $A_{yt}(L,t)$, we load the saved parameters, and equally divide each step into 10 finer steps of equal length and perform SSSF simulation for each finer step. By using the true solution, approximate solution and input optic field in (3.4), we calculate the GE for the optical fiber communication system.

3.3 Simulation System Setup

We simulate a 120 Gbps PM Quadrature phase shift keying (PM-QPSK) system as shown in Figure 3.1. The transmitter DSP helps to generate raised cosine pulses with a roll-off factor of 0.1 [23]. The 3dB bandwidth for the modulator is 26 GHz, and it is 40 GHz for the arrayed waveguide grating (AWG). The system includes repeated transparent spans each with 100 km SSMF for CD compensated links with 100% inline CD compensation. The 17.7 km dispersion compensating fiber (DCF) is placed in-between the two-stage EDFAs in every span, and the signal power into DCF is kept 7 dB lower than that into the SSMF. We study the cases for two PMD coefficients: $\varsigma = 0.1$ ps/$\sqrt{km}$ (CASE 1) vs. 1.0 ps/$\sqrt{km}$ (CASE 2). The two dramatically different PMD coefficients lead to significantly different differential group delays (DGDs) in our simulations.

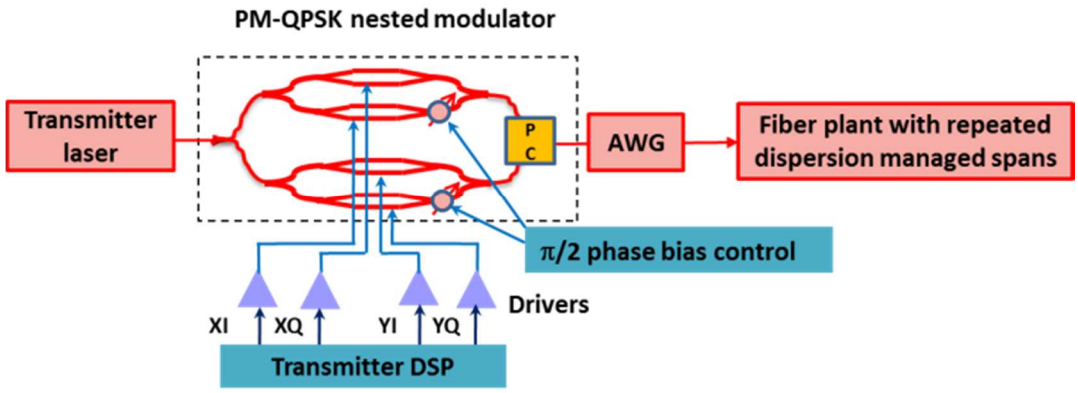


Figure 3.1: System setup: note that PC stands for polarization combiner.

In Figure 3.1, only the center wavelength channel is shown, and its carrier frequency is 194 THz (or the wavelength is 1545.32 nm). The studied WDM system contains a total of five wavelength channels. The channel of interest is the center channel, so on each side of the channel of interest, there are two other channels. The WDM channel spacing is 50 GHz. A 1024-bit pseudorandom bit sequence is used, and the simulation bandwidth is 960 GHz. Figure 3.2a shows the simulated eye diagram of the x-input field for the channel of interest, and Figure 3.2b shows the energy spectrum density (ESD) of the total x-input field including all the five wavelength channels. Note that only the baseband simulation signals are shown in Figure 3.2. Also, in Figure 3.2a, the channel of interest i.e., the center channel is filtered out from the WDM total field using an ideal rectangular shaped filter with a full-width bandwidth of 50 GHz.

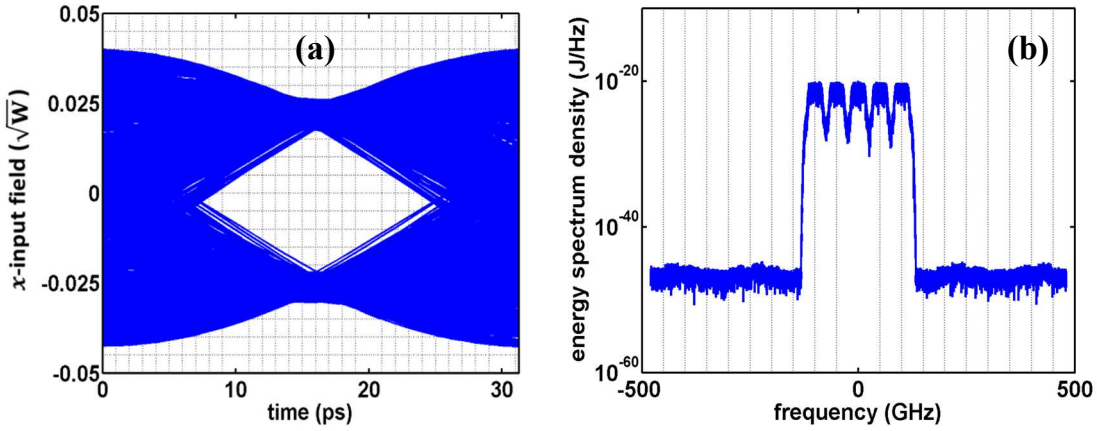


FIGURE 3.2: (a) and (b) shows the eye diagram of the filtered center channel in the total x-input field, and the energy spectrum density of the x-input total field respectively.

3.4 Simulation Results

As explained earlier, for GE control, we use a recently developed scalar LEB finding software package, to obtain the LEB for a prescribed GE in single span systems [24]. The original total x-input field is normalized to have a launch power of 10 dBm, and then it is used as the input to the scalar software. Recall that the software is configured to stop the simulation once the obtained GE is within $90\% - 110\%$ of the prescribed GE. The prescribed GEs are $nsd = 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$, and 10^{-5}, for which we obtain the corresponding LEBs via LEB finding simulations. Now for each obtained LEB or $\Delta\xi$ we use the proposed local error method to perform vector simulations each with five PMD realizations, and save all the rotation angles, phase scattering angles, and step-sizes [12, 13]. For vector simulations the input is composed of both x and y total fields, and the total launch power is 10 dBm. This constitutes the first simulation run. In the second simulation run, we reload the saved parameters, and divide each saved step into 10 equal steps to simulate the accurate or the "true" solutions. The GEs can then be computed

using the nsd defined in [12] for vector signals. The study is carried out for both CASE 1 (small PMD) and CASE 2 (large PMD) in a standard single mode fiber (SSMF). The results are summarized in Figure 3.3. The GE vs. number of steps during the scalar LEB study is also plotted as a reference.

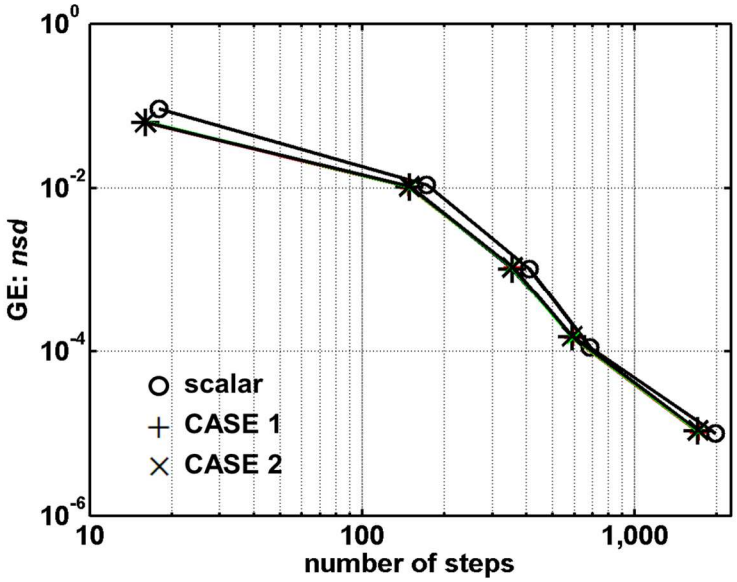


FIGURE 3.3: Single span GE vs. total number of steps for scalar, CASE 1, and CASE 2. Note for each of the two studied vector simulation cases, 2 curves are plotted with each curve for a particular PMD realization.

There are several observations for our WDM results as shown in Figure 3.3. Although the curves follow the correct trend from the second order SSSF method i.e., the four orders of magnitude change in GE (from 10^{-1} to 10^{-5}) corresponds to approximately two orders of magnitude change in the number of steps, the curves deviates from the straight line shape which is present in single channel cases [12, 13]. This is due to the fact that for WDM systems, dispersion effect is significantly more pronounced when compared to single channel systems due to the much increased bandwidth. Nonetheless, by using the LEBs from the scalar simulation, the same target GEs are satisfied for WDM systems in

vector simulations. Furthermore, the computational efficiency of the vector simulation is similar to that of the scalar simulation in terms of the number of steps needed to achieve the same GE. This means that when we use the LEBs obtained from the WDM simulation for vector WDM simulations, there is almost no waste of computation.

Now we investigate for WDM systems whether the proposed local error method leads to higher computational efficiency when compared to two often-used step-size selection methods: the walk-off method and the constant step-size method [12, 13]. Because the step-sizes behave differently for the three studied step-size schemes, and also because random polarization scattering is applied for each step in coarse step method, it takes carefully designed simulations to obtain GE results for the three compared methods with **the same PMD realization.** Here we extend the method proposed for single channel study in [12, 13]. One PMD realization is selected from vector WDM simulations using the local error method with a large $\Delta\xi$. This results in large step-sizes and less than 20 simulation steps. With the saved polarization scattering angles and the saved large step-sizes, we redo simulations but now we treat each saved large step as a polarization maintaining fiber. Within each polarization maintaining fiber, we apply the three step-size schemes with much smaller $\Delta\xi$ values. To obtain accurate or "true" solutions we use extremely small step-sizes in polarization maintaining fiber segments.

In Figure 3.4 we compare the computational efficiency of the three studied methods for the same PMD realization. For both CASE 1 and CASE 2, we found that over a large GE range e.g., from 10^{-6} to 10^{-4}, the proposed local error method requires from 40% to 75% of the number of simulation steps required by the walk-off method. Over an even

larger GE range i.e., $nsd = 10^{-6}$ to 10^{-3}, the proposed local error method needs less than 50% of the number of simulation steps required by the constant step-size method. Though the observed computational savings fall within the region of low global accuracy (or very high GE), our results has practical significance for multi-span WDM simulations. Assume the target GE is 10^{-3} for a 100-span simulation in a transoceanic optical fiber communication system, a conservative GE target will be 10^{-5} in single span simulation. So for long distance multi-span WDM simulations the results obtained from local error method represent a significant computational efficiency enhancement over the other two often-used methods. So for WDM vector simulations, the local error method maintains its optimal step-size nature, and it still requires the least number of simulation steps to achieve a certain GE level.

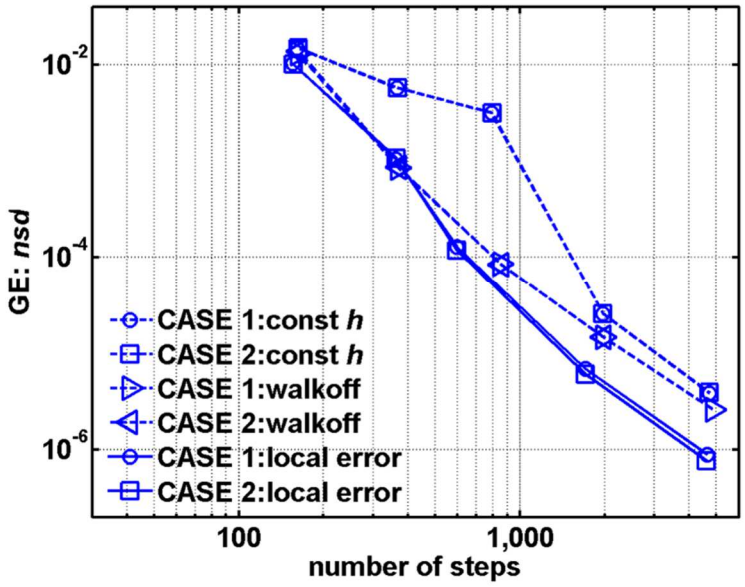


FIGURE 3.4: Single span GE vs. number of steps for three step-size methods. In the legend, "const h" means the constant step-size method, "walkoff" means the walk-off method, and "local error" means the local error method.

We simulate a 12-span system with the proposed local error method. In Figure 3.5,

we did the simulation for both cases with one PMD realization. Due to some constraints in the simulation software, we cannot run more than one PMD realization during the same simulation run. Both CASEs are studied, and for each of them we have five PMD realizations, we run one PMD realization each time and represented them in a single plot as shown in Figure 3.6. For both cases CASE 1 and CASE 2, $\Delta\xi = 5 \times 10^{-3}$ is used.

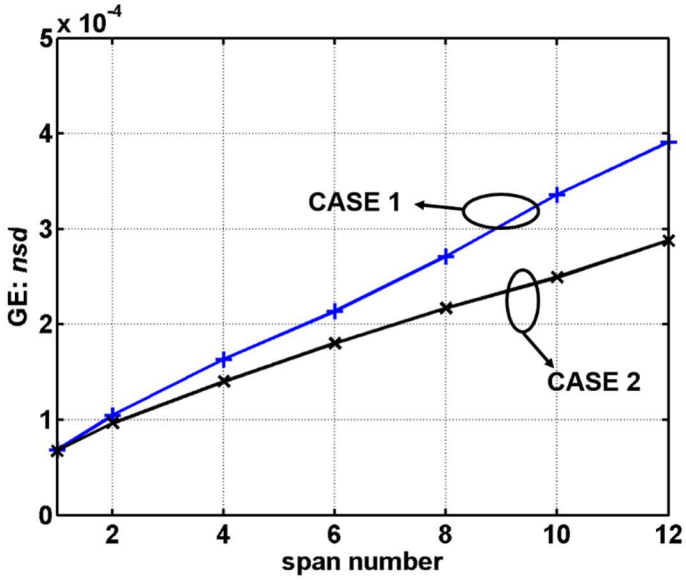


FIGURE 3.5: Multi-span GE vs. span number for the two studied CASEs.

We found that for all our simulations, the number of simulation steps for each span is similar. Key results are summarized in Figure 3.6. Comparing to the high accuracy single channel result in [12], we do observe two key differences. One, compared to [12], there exists a bigger gap between slopes of the curves for simulations with small PMD (CASE 1) and these with large PMD (CASE 2). Furthermore, with large PMD, the accumulation of the GE is slower with the increase of the number of spans, as compared to the small PMD case. Two, here in WDM study, with large PMD there is less spread in the curves

for different PMD realizations while in single channel study the opposite happens: with large PMD there is more spread in the curves for different PMD realizations. Nonetheless, for all studied cases the spread is not significant. The above multi-span simulation results extend GE control to multi-span WDM systems: to achieve a prescribed GE for a N-span system, we can simply scale the GE by a factor of $1/N$, and use this target GE to obtain the desirable $\Delta\xi$ with the single span scalar LEB finding software. Figure 3.6 show that for CASE 1, the GE after 12 spans is about seven times of the GE value after one single span, and for CASE 2, it is about five times. As a result, the computation overhead in our proposed multi-span GE control it is not significant.

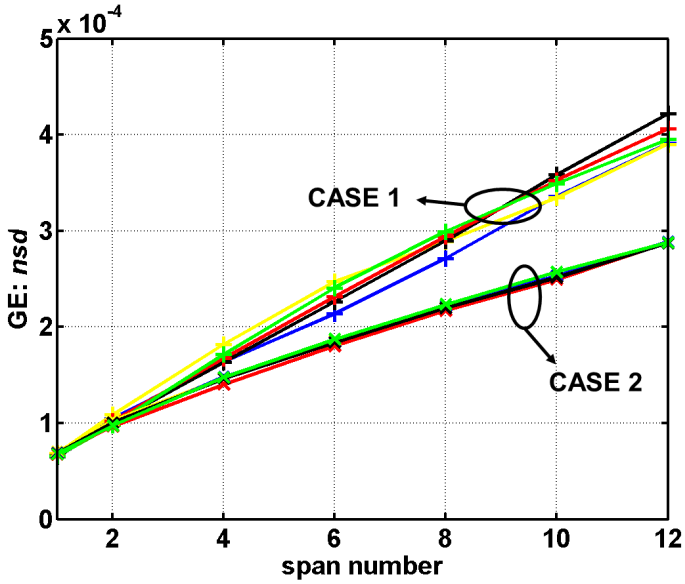


FIGURE 3.6: Multi-span GE vs. span number for the two studied CASEs.

So far we studied single span and multi span results for WDM systems using SSMF fiber links. Now we will go through the single span and multi span results for WDM systems using TWRS fiber links. In Figure 3.7 we observed that the vector curves are almost similar to the scalar curves with only a slight deviation and four orders of

magnitude change in GE (from 10^{-1} to 10^{-5}) corresponds to nearly two orders of magnitude in the number of steps as per the second order SSSF method. The scalar and vector curves are almost similar to a straight line just like the single channel case [12, 13]. By comparing Figure 3.3 with 3.7 we can clearly see that the number of steps needed to achieve a similar accuracy is reduced significantly when we used a TWRS compared to an SSMF fiber since the dispersion effect is low in a TWRS fiber. Finally, we can see that the computational efficiency of the vector simulation is similar to that of the scalar simulation in terms of number of steps needed to achieve same prescribed GE.

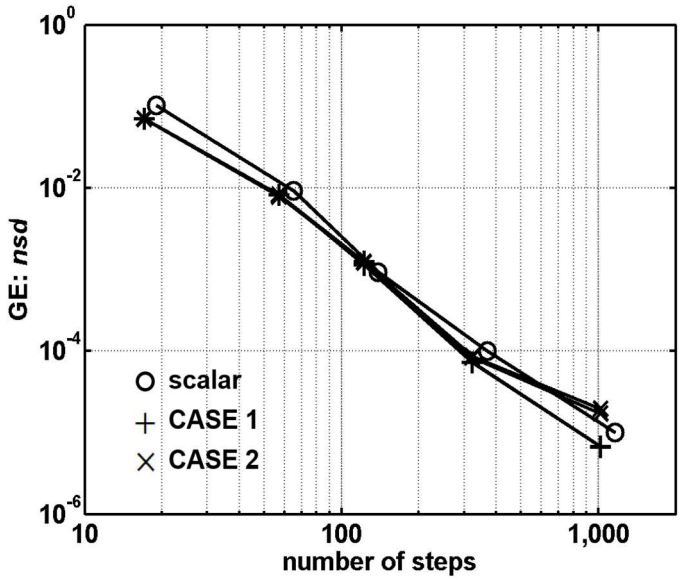


FIGURE 3.7: Single span GE vs. total number of steps for CASE 1 and CASE 2 in a TWRS fiber link. Note for each of the two studied vector simulation cases, 2 curves are plotted with each curve for a particular PMD realization.

The computational efficiency of the three studied methods for a single PMD realization is represented in Figure 3.8. For both CASE 1 and CASE 2, we found that over a large GE range i.e., from 10^{-6} to 10^{-2}, the proposed local error method needs only 40% to 70% of the number of steps needed by the walk-off method. Over a large GE

range i.e.., $nsd = 10^{-6}$ to 10^{-4}, the local error method needs even less than 50% of the number of steps required by constant-step method. Note that the number of FFT/IFFTs per simulation step is same for all the three step-size selection methods.

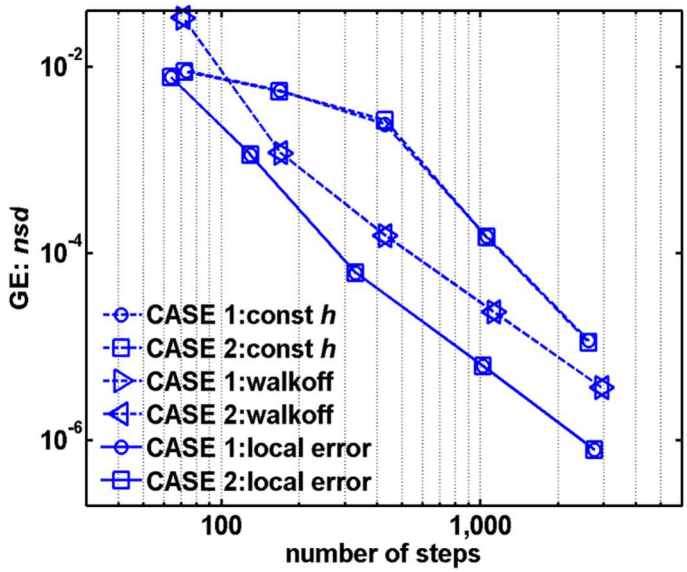


FIGURE 3.8: Single span GE vs. number of steps for three step-size methods in a TWRS fiber link. In the legend, "const h" means the constant step-size method, "walkoff" means the walk-off method, and "local error" means the local error method.

We simulate a 12-span system with the proposed local error method. In Figure 3.9, we did the simulation for both CASESs with one PMD realization. Due to some constraints in the simulation software, we cannot run with more than one PMD realization at the same time. Both CASEs are studied with five different PMD realizations, we run one PMD realization each time and its result is shown by a single curve as shown in Figure 3.10. There are altogether five curves for each case. The LEB with $\Delta\xi = 4 \times 10^{-3}$ is used. We find that the number of simulation steps for each span is similar. Comparing with the single channel results in [12], there exists a bigger gap between slopes of the curves for simulations with small PMD (CASE 1) and these with

large PMD (CASE 2). Furthermore, with large PMD, the increase of the GE is slower with the increase of the number of spans, as compared with the small PMD case. Similar to the SSMF case, here in WDM study, with large PMD there is less spread in the curves for different PMD realizations, which is quite opposite in single channel case where we have more spread in the large PMD case. Figure 3.10 show that for CASE 1, the GE after 12 spans is about 10 times of the GE value after one single span, and for CASE 2, it is about 7 times. This is due to the smaller chromatic dispersion effect for TWRS systems.

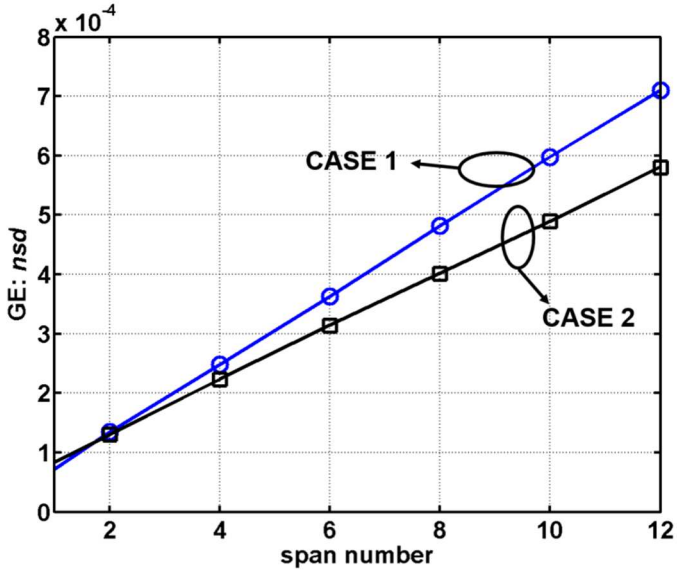


FIGURE 3.9: Multi-span GE vs. span number for the two studied CASEs with only one PMD realization in a TWRS fiber link.

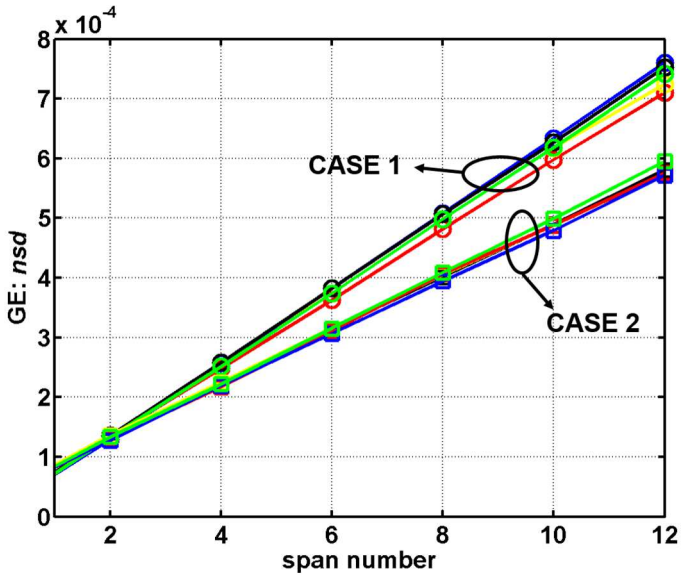


FIGURE 3.10: Multi-span GE vs. span number for the two studied CASEs with 5 PMD realizations in a TWRS fiber link.

Chapter 4

Summary and Future Work

4.1 Summary of this thesis

In Chapter 1, the concept of coherent optical communication with the dual polarization and related mathematical equations for modeling the wavelength division multiplexing (WDM) system is introduced. The local error method to achieve similar local simulation error in the simulation of signal propagation through the optical fiber is introduced. For local error method, the relation between the one step simulation error and global simulation error with the step-size is explained.

In Chapter 2, different conventions used for the polarization and Poincaré sphere is explained. The polarization handedness is different for different notations, e.g., the right hand circular polarization in engineering notation was described as left hand circular polarization in the traditional physics notation and vice versa. Ways used to represent the state of polarization is explained. To assist our discussions, different notations used for the complex envelope and Fourier transforms are described in Appendix A. The detailed derivation of coupled nonlinear Schrödinger (CNLS) equations is presented using an engineering notation. We showed that the coupled nonlinear Schrödinger equations used by Agarwal, Menyuk and Evangelides are the same.

In Chapter 3, we applied a local error method employing an analytical step-size selection rule to the second order SSSF simulation of WDM systems with dual polarization QPSK signal transmission through the vector optical fiber channel. Five wavelength channels are used.

It is found that similar to the single channel systems, the global simulation accuracy for the vector propagation can be satisfied using the local error bound (LEB) obtained from a scalar propagation model for the same global error over a large range of simulation accuracy and differential group delay (DGD).

With extensive simulations, we show that for both single span and multi-span 100% dispersion compensated SSMF and TWRS fiber links and for a wide range of different group delay, the proposed method results in higher computational efficiency than the walk-off and constant step-size methods for a wide range of global error (GE) levels. We show that the computational efficiency for one span simulation can be maintained for multi-span WDM simulations, due to the fact that the number of steps is similar for all the spans. The scaling of the global simulation error with respect to the number of optical fiber spans is also proposed, and we show that global error control for multi-span simulation can be extended to WDM simulations.

4.2 Future work

The presented work represents a systematic study of strategies of most efficient simulation of optical signal propagation through the optical fiber for WDM systems, in terms of step-size selection. Due to time constraint we have not directly verify the similar local error achieved at different simulation steps, while we have done that for single

channel systems. This may constitute a future work. Furthermore, step-size pre-calculation also represents a possible future work for WDM systems. Finally, future work may include the study of new methods that further enhances computational efficiency within one simulation step, the study of other GE scaling rules, and the validation of the local error method for dispersion uncompensated systems.

Appendix A

Notations for complex envelope and Fourier transform

A uniform plane wave is characterized by electric and magnetic fields that have uniform properties at all points across an infinite plane. The properties of an electromagnetic wave, such as its phase velocity v_p and wavelength depends on the angular frequency and the three parameters electrical permittivity (ε) , magnetic permeability (μ) and conductivity (σ). If the medium is lossless ($\sigma = 0$) then the wave does not suffer any attenuation during the propagation. For lossless media, there will not be any dispersion and wavenumber (k) as:

$$k = \omega\sqrt{\mu\varepsilon}$$

The wave impedance of the electromagnetic wave is the ratio of the transverse components of electric and magnetic fields. For a transverse electromagnetic (TEM) wave propagating through a homogeneous medium, the wave impedance is always equal to the intrinsic impedance of the medium. The intrinsic impedance of a lossless medium is defined as:

$$\eta = \omega\mu/k = \omega\mu/\omega\sqrt{\mu\varepsilon} = \sqrt{\mu/\varepsilon}$$

The phase velocity of the wave is

$$v_p = \omega/k = 1/\sqrt{\mu\varepsilon}$$

and its wavelength is

$$\lambda = 2\pi/k = v_p/f$$

There are two kinds of notations for the complex envelope of pass band signal. In the physics and traditional optics, we have

$$\begin{aligned} E(t,z) &= Re[\, B(t,z)e^{-i\omega_c t}] \\ &= \tfrac{1}{2}[\, B(t,z)e^{-i\omega_c t} + c.c.] \\ &= \tfrac{1}{2}[\, B(t,z)e^{-i\omega_c t} + B^*(t,z)e^{i\omega_c t}] \end{aligned} \tag{1.1}$$

where ω_c is the carrier frequency, $c.c.$ stands for complex conjugate, and $B(t,z)$ is the complex envelope of the real passband signal $E(t,z)$.For the same signal $E(t,z)$, the complex envelope is denoted as $A(t,z)$ in engineering, modern physics, and modern optics notations, such that

$$\begin{aligned} E(t,z) &= Re[\, A(t,z)e^{i\omega_c t}] \\ &= \tfrac{1}{2}[\, A(t,z)e^{i\omega_c t} + c.c.] \\ &= \tfrac{1}{2}[\, A(t,z)e^{i\omega_c t} + A^*(t,z)e^{-i\omega_c t}] \end{aligned} \tag{1.2}$$

From (1.1) and (1.2) we derive the relationship between the two complex notations as follows:

$$A(t,z) = B^*(t,z) \tag{1.3}$$

Now we will discuss about phase velocity under two different notations. We simply take a look at plane waves with the following constant phase planes i.e..,$kz - \omega t = constant$ for the traditional physics notation, and $\omega t - kz = constant$ for the engineering notation. The phase velocity is same, $v_p = \Delta z/\Delta t = \omega/k$

We illustrate the two different notations of Fourier transform in the continuous-time domain. According to engineering and modern physics/optics notations, we have the following Fourier transform of the complex envelope signal $A(t,z)$,

$$\tilde{A}(\omega,z) = \int_{-\infty}^{\infty} A(t,z)e^{-i\omega t}\,dt \tag{2.1}$$

On the other hand, we have the following counterpart with the traditional physics and traditional optics notations,

$$\tilde{B}(\omega,z) = \int_{-\infty}^{\infty} B(t,z)e^{i\omega t}\,dt \tag{2.2}$$

Let's assume that $A(t,z)$ and $B(t,z)$ are complex envelope signals defined in (1.1) and (1.2), then we have

$$\begin{aligned}\tilde{A}^*(\omega,z) &= \int_{-\infty}^{\infty} A^*(t,z)e^{i\omega t}\,dt \\ &= \int_{-\infty}^{\infty} B(t,z)e^{i\omega t}\,dt = \tilde{B}(\omega,z)\end{aligned} \tag{2.3}$$

As a result, we can summarize that if $A(t,z)$ and $B(t,z)$ are the complex envelopes for the same passband signal $E(t,z)$ in respectively engineering and traditional physics/optics notations, then $A(t,z) = B^*(t,z)$. Furthermore, $\tilde{A}(\omega,z)$ and $\tilde{B}(\omega,z)$ obtained from the Fourier transforms with respectively engineering and traditional physics/optics notations, will also forms complex conjugate pairs i.e., $\tilde{A}(\omega,z) = \tilde{B}^*(\omega,z)$.

In engineering notation, we have

$$a\cos((\omega_c + \Delta\omega)\,t - kz) = \frac{1}{2}[\,ae^{i\Delta\omega t}e^{i\omega_c t - ikz} + \,ae^{-i\Delta\omega t}e^{-i\omega_c t + ikz}]$$

$$a\cos((\omega_c + \Delta\omega)\,t - kz) = \frac{1}{2}[\,A(t,z)e^{i\omega_c t} + A^*(t,z)e^{-i\omega_c t}]$$

Where $A(t,z) = a\,e^{i(\Delta\omega t - kz)}$.

In physics notation, we have

$$a\cos(kz-(\omega_c+\Delta\omega)\,t)=\frac{1}{2}[\,ae^{-i\Delta\omega t}e^{-i\omega_c t+kz}+\,ae^{i\Delta\omega t}e^{i\omega_c t-kz}e^{-ikz}]$$

$$=\frac{1}{2}[\,B(t,z)e^{-i\omega_c t}+B^*(t,z)e^{i\omega_c t}]$$

where $B(t,z)=a\,e^{-i\Delta\omega t}e^{ikz}=ae^{i(kz-\Delta\omega t)}$.

Now let's study an example. We assume to have a plane wave as the carrier signal for an information signal. We assume that the information signal is a simple constant with a carrier offset $\Delta\omega$, we have the complex envelope as $ae^{i(\Delta\omega t-kz)}$ and the Fourier transform as $a2\pi\delta(\omega-\Delta\omega)e^{-ikz}$ with engineering notation. Now with physics notation we have the complex envelope as $ae^{i(kz-\Delta\omega t)}$ and the Fourier transform as

$$\int_{-\infty}^{\infty}a\,e^{i(kz-\Delta\omega t)}e^{i\omega t}\,dt\;=-\int_{-\infty}^{\infty}a\,e^{i(kz-\Delta\omega t)}e^{-i\omega(-t)}\,d(-t)$$

$$=-\int_{\infty}^{-\infty}a\,e^{i(kz+\Delta\omega t)}e^{-i\omega t}\,dt$$

$$=\;\int_{-\infty}^{\infty}a\,e^{ikz}e^{-i(\omega-\Delta\omega)t}\,dt=a\,2\pi\delta(\omega-\Delta\omega)e^{ikz}\quad(2.4)$$

From the above example, we can clearly see the relations in (1.3) and (2.3). One key remark is that Fourier transform or the power spectrum of the complex envelope identifies the same frequency location for the two systems of notations, with a key requirement that both the complex envelope and the Fourier transform follow the same system of notation. Another way to show (2.4) is as follows

$$\int_{-\infty}^{\infty}a\,e^{i(kz-\Delta\omega t)}e^{i\omega t}\,dt\;=ae^{ikz}\int_{-\infty}^{\infty}\,e^{i(\omega-\Delta\omega)t}\,dt$$

$$=ae^{ikz}\int_{-\infty}^{\infty}\,e^{-i(\Delta\omega-\omega)t}\,dt$$

$$=ae^{ikz}2\pi\delta(\Delta\omega-\omega)$$

$$= ae^{ikz}2\pi\delta(\omega - \Delta\omega) \quad \text{since } \delta(-x) = \delta(x). \qquad (2.4)$$

Another important remark is that some properties of Fourier transform vary with different systems of notations. For example the operator $\frac{\partial}{\partial t}$ has $i\omega X$ in the engineering notation, and it has $-i\omega X$ in the traditional physics and optics notation because

$$B(t,z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{B}(\omega,z)\, e^{-i\omega t}\, d\omega, \text{ and}$$

$$\frac{dB(t,z)}{dt} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{B}(\omega,z)\, e^{-i\omega t}(-i\omega)\, d\omega$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} [-i\omega\tilde{B}(\omega,z)]\, e^{-i\omega t}\, d\omega \qquad (2.5)$$

We conclude this section with the following discussion. If someone uses complex envelope defined with physics notation and uses Fourier transform with engineering notation then for a single frequency ω_0, the following will be obtained,

$$\tilde{X}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega_0 t} e^{-i\omega t} dt = 2\pi\delta(\omega + \omega_0) \qquad (2.6)$$

Figure A.1 shows the Fourier transform, and the signal is located at $-\omega_0$ that is not a good representation of the signal.

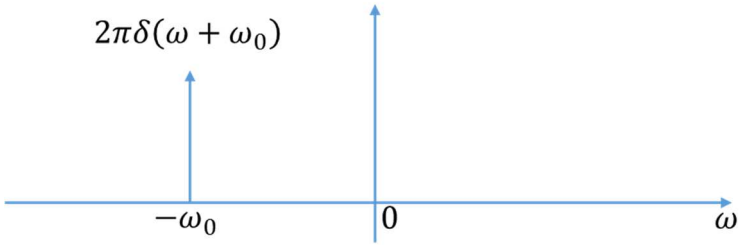


Figure A.1: Fourier Transform of the signal.

Even with the above discussed scenario, when the inverse Fourier transform is applied to $\tilde{X}(\omega)$ again define with the engineering notation, the single frequency signal at $\omega = \omega_0$ is recovered. Furthermore, when the above action i.e., physics notation for complex envelope and engineering notation for Fourier transform is consistently applied to both signals and linear time-invariant system response, we obtain the correct system output. This is why for many equations which take physics notation for the complex envelope representation of signals, tools such as Matlab can be used to conduct Fourier and inverse Fourier transforms, and still correct time domain results can be obtained.

In summary, we recommend to use one system of notation, and in this regard, either physics or engineering notation. The mixing of the notations, for example, the use of physics notation for the complex envelope and engineering notation for Fourier transform may lead to confusing results. The notations are different since the complex envelope they are using is different in engineering and traditional physics and left hand circular (LHC) polarization in one notation is used as right hand circular (RHC) polarization in another notation and vice versa.

The differences between engineering and modern physics/optics compared to traditional physics/optics notation is summarized below.

	Engineering and modern Physics/optics	**Traditional Physics/optics**
Real plane wave:	$E(t) = a\cos((\omega_c + \Delta\omega)\,t - kz)$	$E(t) = a\cos(kz - (\omega_c + \Delta\omega)\,t)$
Complex Envelope	$E(t,z) = \ Re[\ A(t,z)e^{i\omega_c t}]$	$E(t,z) = \ Re[\ B(t,z)e^{-i\omega_c t}]$
Fourier Transform	$\tilde{A}(\omega,z) = \int_{-\infty}^{\infty} A(t,z)e^{-i\omega t}\,dt$	$\tilde{B}(\omega,z) = \int_{-\infty}^{\infty} B(t,z)e^{i\omega t}\,dt$

Appendix B

Derivation of Linear Model for Pulse propagation in fibers

Here we formally derive the equation governing pulse propagation in optical fibers referring to a derivation in [6, 26-30], but with an engineering notation on Fourier transform and the complex envelope was discussed in appendix A.

Like all electromagnetic phenomena, the pulse propagation in optical fibers is governed by Maxwell's equations.

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad , \quad \nabla \times \vec{H} = \vec{J}_f + \frac{\partial \vec{D}}{\partial t}$$

$$\nabla \cdot \vec{D} = \rho_f \quad , \quad \nabla \cdot \vec{B} = 0 \tag{B.1}$$

where ρ_f is the density of the free charge, $\vec{J}_f$ is the current density, $\vec{E} = \vec{E}(\vec{r}, t)$ is the macroscopic electric field, $\vec{B} = \vec{B}(\vec{r}, t)$ is the macroscopic magnetic field, where t is time and $\vec{r}$ is the unit vector from the source charge to the point of interest, $\vec{D} = \vec{D}(\vec{r}, t)$ is called the electrical flux density, and it is related to $\vec{E}$ by $\vec{D} = \varepsilon_0 \vec{E} + \vec{P}$ and $\vec{B} = \mu_0 \vec{H} + \vec{M}$, where ε_0 is the electric permittivity of free space, μ_0 is the free space magnetic permeability and $\vec{P}$ and $\vec{M}$ are the induced electric and magnetic polarizations. Note that the polarization $\vec{P}$ mentioned here is different from the polarization of light, this polarization represents the polarization of a material.

For a nonmagnetic medium like optical fiber, $\vec{M} = 0$, so $\vec{B} = \mu_0 \vec{H}$. Because of the absence of free charges in a medium like optical fibers, $\rho_f = 0$ and $\vec{J}_f = 0$. By using the above assumptions and (B.1) the following equation can be obtained,

We know that $\nabla \times \nabla \times \vec{E} = \nabla(\nabla \cdot \vec{E}) - \nabla^2 \vec{E}$, so the following result can be obtained,

$$\nabla \times \nabla \times \vec{E} = -\nabla \times \frac{\partial \vec{B}}{\partial t} = -\frac{\partial}{\partial t}(\nabla \times \vec{B}) = -\mu_0 \frac{\partial}{\partial t}(\nabla \times \vec{H})$$
$$= -\mu_0 \frac{\partial}{\partial t}(\frac{\partial \vec{D}}{\partial t}) = -\mu_0 \frac{\partial^2}{\partial t^2}(\varepsilon_0 \vec{E} + \vec{P})$$

$$\Rightarrow \nabla \times \nabla \times \vec{E} = -\mu_0 \varepsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} - \mu_0 \frac{\partial^2 \vec{P}}{\partial t^2} = -\frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} - \mu_0 \frac{\partial^2 \vec{P}}{\partial t^2} \quad \text{(B.2)}$$

where c is the velocity of light in vacuum. By assuming the locality of response as well as the instantaneous response of nonlinear polarization, the induced polarization consists of two parts such that

$$\vec{P}(\vec{r}, t) = \vec{P}_L(\vec{r}, t) + \vec{P}_{NL}(\vec{r}, t) = \varepsilon_0 \left(\chi^{(1)} \cdot \vec{E} + \chi^{(3)} \vdots \vec{E}\vec{E}\vec{E} \right) \quad \text{(B.3)}$$

where $\vec{P}_L$ is linear part, $\vec{P}_{NL}$ is the nonlinear part, $\chi^{(1)}$ and $\chi^{(3)}$ are first and third order electric susceptibility, respectively. In this context they are simply called as susceptibility. Because SiO_2 has inversion symmetry so the second order susceptibility is zero. In order to simplify (B.2), let us treat the nonlinear polarization $\vec{P}_{NL}$ as a small perturbation to the total induced polarization, since, the nonlinear effects are relatively weak in silica fiber. Then (B.2) turns into

$$\nabla \times \nabla \times \vec{E} = -\frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} - \mu_0 \frac{\partial^2 \vec{P}_L}{\partial t^2} \quad \text{(B.4)}$$

Following engineering convention, the Fourier transform and its inverse Fourier transform are defined as

$$\vec{E}^F(\vec{r},\omega) = \int_{-\infty}^{\infty} \vec{E}(\vec{r},t) e^{-i\omega t}\, dt$$

$$\vec{E}(\vec{r},t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \vec{E}^F(\vec{r},\omega) e^{i\omega t}\, dt \qquad \text{(B.5)}$$

Assuming the material is isotropic, only the diagonal elements of $\chi^{(1)^F}(\omega)$ are nonzero.

$$\vec{P}_L(t) = \varepsilon_0\ \chi^{(1)}(t) * \vec{E}(t) = \varepsilon_0 \int \chi^{(1)}(\tau)\, \vec{E}(t-\tau) d\tau$$

The transfer function will be $H(f) = \varepsilon_0\ \chi^{(1)}(f)$. Also all the elements of $\chi^{(1)^F}(\omega)$ have same value. From (B.3), (B.4), and (B.3) and with $\varepsilon(\omega) = 1 + \chi^{(1)^F}(\omega)$ which is frequency dependent, the resulting equation is

$$\nabla \times \nabla \times \vec{E}^F(\vec{r},\omega) - \varepsilon(\omega)\frac{\omega^2}{c^2}\vec{E}^F(\vec{r},\omega) = 0 \qquad \text{(B.6)}$$

As $\chi^{(1)^F}(\omega)$ is in general complex, so is $\varepsilon(\omega)$. Its real and imaginary parts can be linked to the refractive index $n(\omega)$ and the absorption coefficient $\alpha(\omega)$ by using the definition $\varepsilon(\omega) = \left(n + i\frac{\alpha c}{2\omega}\right)^2$, assuming α is small, the following approximate result can be obtained,

$$n(\omega) = \left[1 + \mathrm{Re}\left[\chi^{(1)^F}(\omega)\right]\right]^{\frac{1}{2}}$$

$$\alpha(\omega) = \frac{\omega}{nc}\mathrm{Im}\left[\chi^{(1)^F}(\omega)\right] \qquad \text{(B.7)}$$

We can simplify the equation (B.6) before solving it. One, because of very few losses in fibers in the wavelength region of interest, the imaginary part of $\varepsilon(\omega)$ is small in comparison with the real one. So we can replace $\varepsilon(\omega)$ with $n(\omega)$. Two, $n(\omega)$ is

independent of spatial coordinates in both core and cladding of step-index fibers then we can use the equality

$$\nabla \times \nabla \times \vec{E} = \nabla\left(\nabla \cdot \vec{E}\right) - \nabla^2\vec{E} = -\nabla^2\vec{E}$$

where the relation $\nabla \cdot \vec{D} = \varepsilon\nabla \cdot \vec{E} = 0$ was used from $\varepsilon(\omega) = 1 + \chi^{(1)^F}(\omega)$. By doing these simplifications, equation (B.6) takes form of the Helmholtz equation:

$$\nabla^2\vec{E}^F + n^2(\omega)\frac{\omega^2}{c^2}\vec{E}^F = 0 \tag{B.8}$$

Assume that the fiber has a perfect cylindrical geometry and the loss is negligible, the unguided radiation mode is not an important problem. For any frequency fiber can support a finite number of guided modes, whose spatial components satisfy the Helmholtz equation and boundary conditions. It is useful to express the wave equation in cylindrical coordinates and denoting $k_0 = \frac{\omega}{c}$, (B.8) becomes

$$\frac{\partial^2\vec{E}^F}{\partial r^2} + \frac{1}{r}\frac{\partial\vec{E}^F}{\partial r} + \frac{1}{r^2}\frac{\partial^2\vec{E}^F}{\partial\phi^2} + \frac{\partial^2\vec{E}^F}{\partial z^2} + n^2\,{k_0}^2\vec{E}^F = 0 \tag{B.9}$$

Similar relations exists for the magnetic field $\vec{H}$. As $\vec{E}$ and $\vec{H}$ satisfy the Maxwell's equations, only two out of six components are independent variables because we have four equations as constraints in the Maxwell's equations. Without loss of generality, choose ${E_z}^F$ and ${H_z}^F$ as independent components. ${E_z}^F$ can be denoted as ${E_z}^F(\vec{r},\omega) = U(\omega)F(r)e^{iq\phi}e^{i\beta z}$, where $U(\omega)$ is a normalization constant, q is an integer. Then by using the designation $\boldsymbol{\mathcal{K}}^2 = n(\omega)^2\,{k_0}^2 - \beta(\omega)^2$, where $k_0 = \omega/c$ i.e., In vacuum (non-dispersive medium), n is independent of ω, where the phase velocity and group velocity are equal. Also, effective propagation constant $\beta(\omega)$ must be lied in between

$n_{clad}{}^2\, k_0{}^2$ and $n_{core}{}^2\, k_0{}^2$ (or equivalently the effective refractive index must be lied in between n_{core} and n_{clad}). The following equation is obtained from (B.9),

$$\frac{d^2F}{dr^2} + \frac{1}{r}\frac{dF}{dr} + \left[\boldsymbol{\mathcal{K}}^2 - \frac{q^2}{r^2}\right] F = 0 \qquad \text{(B.10)}$$

Equation (B.10) is the well-known differential equation for Bessel functions. Its general solution inside the core can be written as

$$\boldsymbol{\mathcal{K}}^2 > 0 \Rightarrow F(r) = \mathrm{c}_1 J_q(\mathcal{K}r) + \mathrm{c}_2 Y_q(\mathcal{K}r) \qquad (\mathcal{K} = \boldsymbol{\mathcal{K}})$$

$$\boldsymbol{\mathcal{K}}^2 < 0 \Rightarrow F(r) = \mathrm{d}_1 I_q(\gamma r) + \mathrm{d}_2 K_q(\gamma r) \qquad (\gamma = -\boldsymbol{\mathcal{K}}) \quad \text{(Cladding)} \quad \text{(B.11)}$$

where J_q and Y_q are the Bessel functions of the first and second kind respectively, while I_q and K_q are modified Bessel functions of the first and the second kind, respectively. In the above equation, c_1, c_2, d_1 and d_2 are constants.

We used n_{core} to represent n for $r < r_0$ and n_{clad} to represent n for $r > r_0$. Here r_0 is the radius of the core of the fiber. For confined modes, $n_{clad}{}^2\, k_0{}^2 < \beta^2 < n_{core}{}^2\, k_0{}^2$. With a boundary condition that tangential components $\vec{E}^F$ and $\vec{H}^F$ are continuous across the core-cladding interface requires that $E_z{}^F$, $H_z{}^F$, $E_\phi{}^F$ and $H_\phi{}^F$ must be same when $r = r_0$, is approached either from inside or from oust side of the core, the physical solution is

$$F(r) \propto J_q(\mathcal{K}r) \;\; (r < r_0); \qquad F(r) \propto K_q(\gamma r) \;\; (r > r_0) \qquad \text{(B.12)}$$

Due to the equality of the field components when $r = r_0$ which yields an eigenvalue equation,

$$\left[\frac{q\beta k_0 (n_{core}{}^2 - n_{clad}{}^2)}{r_0 \mathcal{K}^2 \gamma^2 n_{in}}\right]^2 = \left[\frac{J_q{}'(\mathcal{K}r_0)}{\mathcal{K} J_q(\mathcal{K}r_0)} + \frac{K_q{}^1(\gamma r_0)}{\gamma K_q(\gamma r_0)}\right] + \left[\frac{J_q{}'(\mathcal{K}r_0)}{\mathcal{K} J_q(\mathcal{K}r_0)} + \frac{n_{core}{}^2}{n_{clad}{}^2}\frac{K_q{}^1(\gamma r_0)}{\gamma K_q(\gamma r_0)}\right] \qquad \text{(B.13)}$$

where both n_{core} and n_{clad} are functions of ω, "prime" denotes differentiation. The above (B.13) is a transcendental function of β for each q. For each q and ω, only a finite

set of ω can be found that satisfy the equality within the range $({n_{clad}}^2\,{k_0}^2,\ {n_{core}}^2\,{k_0}^2)$. After β is found, both $\vec{E}^F$ and $\vec{H}^F$ can be obtained.

By solving this eigenvalue equation we will get different solutions for β for each integer value of q. In general we will express this solutions by β_{ql}, where both q and l are integers. Each eigenvalue β_{ql} corresponds to one specific mode supported by fiber. There are two types of fiber modes [6], designated as HE_{ql} and EH_{ql}. When $q = 0$, these modes are analogous to transverse-electric (TE) and transverse-magnetic (TM) modes of a planar wave guide since, the axial components of electric and magnetic fields disappears. In TE mode there will be no electric field in the direction of propagation i.e., $E_z = 0$ and TM mode no magnetic field in the direction of propagation i.e.., $H_z = 0$. When q IS nonzero the solution has two fold degeneracy since the fiber itself does not possess any preferred sense of rotation. For $q > 0$, fiber modes becomes hybrid, i.e., all six components of the electromagnetic fields are nonzero.

The study of this weekly guided step index fiber has been done before but for completeness, we added about the weekly guided step index fiber here by referring a paper by Gloge [33]. In an optical fiber when the difference between the refractive index of core and cladding is small, i.e., $n_{core} - n_{clad} \ll 1,\ \mathcal{K}, \gamma \ll \beta$. Then E_z and H_z are not dominant, but E_x , E_y , H_x and H_y are significant. In general $E_\phi = -E_x \sin\phi + E_y \cos\phi$ exists. There will be $E_\phi \propto E_x$ or E_y respectively if E_y or E_x vanishes respectively. Also due to the continuity of E_ϕ on boundaries, E_x and E_y continuous on boundaries. If the electric field is $y-$ polarized, E_y and H_x dominate. If $x-$ polarized, E_x and H_y dominate. By inspecting the solutions from the above two situations, there

always $H_\phi \propto E_\phi$. So H_ϕ's continuity is also ensured. Both the requirements of E_z's continuity and that of H_z's leads to two equivalent mode conditions. The resulting mode conditions are $\frac{\mathcal{K}J_{q+1}(\mathcal{K}r)}{J_q(\mathcal{K}r_0)} = \frac{\gamma K_{q+1}(\gamma r_0)}{K_q(\gamma r_0)}$ and $\frac{\mathcal{K}J_{q-1}(\mathcal{K}r_0)}{J_q(\mathcal{K}r_0)} = \frac{\gamma K_{q-1}(\gamma r_0)}{K_q(\gamma r_0)}$. The above two identities are equivalent according to Bessel function properties. This means that the x-polarized and y-polarized solutions give the same eigenvalue equation. So these two transversely orthogonal modes are degenerate in the propagation constant β. Note that this approximately degenerates $HE_{q+1,l}$ and $EH_{q-1,l}$. When n_{core} goes to n_{clad} they are strictly degenerate. Here $HE_{q+1,l}$ and $EH_{q-1,l}$ can be denoted as LP_{ql}. Now the lowest order mode HE_{11} turns into LP_{01}. The weakly guided fiber model will be used in the following context for good degree of approximation, the fundamental fiber mode is linearly polarized in either x or y direction depending on which electric field dominates among E_x and E_y.

Appendix C

Computation of Differential Group Delay

The difference in the propagation delay between the light traveling on the fast and slow principal states of polarization (PSP) plays a role in determining the output polarization state of the light. PMD measurements based in the frequency domain measures same differential group delay (DGD) just like the time domain measurement but in a different point of view. The most often used approach for DGD measurement in the frequency domain involves a differential method.

Let us take four points on a sphere with center at origin O, as shown in Figure C.1.

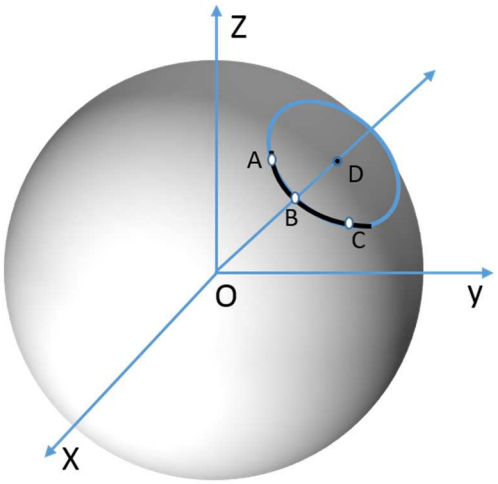


Figure C.1: Poincaré Sphere

We list the four points and their coordinates as

$A: (A_x, A_{y,} A_z)$, $B: (B_x, B_{y,} B_z)$, $C: (C_x, C_{y,} C_z)$, and $D: (D_x, D_{y,} D_z)$

Assume that points A, B and C are on the circumference of a circle with center D. then we can say that

$$\overrightarrow{OD} \perp \overrightarrow{AB},\ \overrightarrow{OD} \perp \overrightarrow{AC} \text{ and } \overline{AD} = \overline{BD} \tag{C.1}$$

By using relations in (C.1) we write the following equations

$$D_x(B_x - A_x) + D_y\left(B_y - A_y\right) + D_z(B_z - A_z) = 0 \tag{C.2}$$

$$D_x(C_x - A_x) + D_y\left(C_y - A_y\right) + D_z(C_z - A_z) = 0 \tag{C.3}$$

$$(D_x - A_x)^2 + \left(D_y - A_y\right)^2 + (D_z - A_z)^2 = (D_x - B_x)^2 + \left(D_y - B_y\right)^2 + (D_z - B_z)^2 \tag{C.4}$$

By solving (C.4) we get

$$D_x(2B_x - 2A_x) + D_y\left(2B_y - 2A_y\right) + D_z(2B_z - 2A_z) = (B_x^2 - A_x^2) + \left(B_y^2 - A_y^2\right) + (B_z^2 - A_z^2) \tag{C.5}$$

$$\begin{bmatrix} (B_x - A_x) & \left(B_y - A_y\right) & (B_z - A_z) \\ (C_x - A_x) & \left(C_y - A_y\right) & (C_z - A_z) \\ 2(B_x - A_x) & 2\left(B_y - A_y\right) & 2(B_z - A_z) \end{bmatrix} \begin{bmatrix} D_x \\ D_y \\ D_z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \overline{OB}^2 - \overline{OA}^2 \end{bmatrix}$$

But we know that $\overline{OB}^2 - \overline{OA}^2 = 0$ then the equation corresponding to the third row in the 3×3 matrix is redundant with the equation corresponding to the first row. Since it is mandatory that $\overrightarrow{AD} \perp \overrightarrow{OD}$.

$$D_x(D_x - A_x) + D_y\left(D_y - A_y\right) + D_z(D_z - A_z) = 0 \tag{C.6}$$

or, $$D_x^2 - A_xD_x + D_y^2 - A_yD_y + D_z^2 - A_zD_z = 0 \tag{C.7}$$

By using the two equations (C.5) and (C.6) or (C.7) we can calculate the DGD.

Appendix D

List of Acronyms

APP – A Posteriori Probability

AWG – Arrayed Waveguide Grating

AWGN – Additive White Gaussian Noise

BPSK – Binary Phase Shift Keying

CD – Chromatic Dispersion

CNLS – Coupled Nonlinear Schrödinger

CW – Continuous Wave

DCF – Dispersion Compensating Fiber

DFB – Distributed Feedback

DND – Dispersion Nonlinearity Dispersion

DSP – Digital Signal Processing

FFT/IFFT – Fast Fourier Transform/Inverse Fast Fourier Transform

GE – Global Error

GVD – Group Velocity Dispersion

IMDD – Intensity Modulation and Direct Detection

LDPC – Low Density Parity Check

LE – Local Error

LEB – Local Error Bound

LHC – Left Hand Circular

LLR – Log-likelihood Ratio

LO – Local Oscillator

MZM – Mach-Zehnder modulators

NLSE – Nonlinear Schrödinger equation

PBS – Polarization Beam Splitter

PM – Polarization Multiplexed

PDM – Polarization Division Multiplexing

PMD – Polarization Mode Dispersion

PSK –Phase Shift Keying

QPSK – Quadrature Phase Shift Keying

QAM – Quadrature amplitude modulation

RHC – Right hand circular

SNR – Signal to Noise Ratio

SOP – State of Polarization

SSF – Split-Step Fourier

SSMF – Standard Single-mode Fiber

SSSF – Symmetrized Split-Step Fourier

TWRS – True Wave Reduced slope

WDM – Wavelength Division Multiplexing

References

[1] John G Proakis and Masoud Salehi, “Digital Communications,” 5th edition, McGraw-Hill, 2008.

[2] Stephen G Wilson, “Digital modulation and coding,” Prentice Hall, 1996.

[3] Fuqin Xiong, “Digital Modulation Techniques,” 2nd edition, Artech House, 2006.

[4] R. W. Boyd, “Nonlinear Optics,” 3rd edition, Academic Press, 2008.

[5] M. C. Jeruchim, P. Balaban, and K. S. Shanmuga, “Simulation of Communication Systems: Modeling, Methodology, and Techniques,” 2nd edition, New York: Springer, 2000.

[6] Govind P Agrawal, “Nonlinear Fiber Optics,” 5th edition, Oxford: Academic Press, 2013.

[7] Milorad Cvijetic and Ivan B.Djordjevic, “Advanced optical communication systems and networks,” Artech House, 2013.

[8] L.E.Nelson, S.L woodward, S.foo, and X.zhou, “Performance of a 46-Gbps Dual polarization QPSK Transceiver with Real-time coherent equalization over High PMD fiber,” Journal of Light wave Technology, Vol.27, No.3, February 1, 2009.

[9] Q. Zhang, Santosh karri, Muhammad Khaliq, Liudong Xing and M. I. Hayee, “Global Simulation Accuracy Control in the Split-Step Fourier Simulation of Vector Optical Fiber Communication Channel,” Journal of Communications Vol. 10, No. 1, January 2015

[10] Q. Zhang and M. I. Hayee, “An SSF scheme to achieve comparable global simulation accuracy for WDM systems,” IEEE Photon. Technol. Lett., 17, 1869-1871 (2005).

[11] Q. Zhang and M. I. Hayee, "Symmetrized SSF scheme to control global simulation accuracy in fiber optic communication systems," IEEE Journal of Light wave Technology, Vol 26, 302-316 (2008).

[12] Q. Zhang, L. Xing, H. Min, and M. I Hayee, “Analytical step-size selection rule for simulation of signal propagation in vector optical fiber channel,” CLEO: 2015, OSA Technol. Digest, paper JTh2A.51, 2015.

[13] Q. Zhang, M. I. Hayee, V. Winstead, X. Wu, D. Huang, J. Lian, S. Phadke, and M. Khaliq, “A local error method for sssf simulation of signal propagation in dispersion compensated optical links," Proceedings, The 14th International Conference on Optical Communications and Networks (ICOCN 2015), paper Sa1D.6, Nanjing, China, July 2015.

[14] D. Marcuse, C. R. Menyuk, P. K. A. Wai, “Application of the Manakov-PMD Equation to Studies of Signal Propagation in Optical Fibers with Randomly Varying Birefringence," Journal of Light wave Technology, Vol.25, No.9, September 1997.

[15] D. Derickson, “Fiber Optic Test and Measurements,” Prentice Hall, New Jersey 1998.

[16] Fawwaz T.Ulaby and Umberto Ravaioli, “Fundamentals of applied electromagnetics,” seventh edition. Pearson Education Inc.

[17] C. E. Shannon, “A mathematical theory of communications,” Bell Syst. Techn. J., Vol. 27, pp 379-423 and 623-656, Jul and OCT. 1948

[18] M. Born and E.Wolf, Principles of Optics, 6th ed. (Cambridge University Press, Cambridge, 1997).

[19] P. K. A. Wai and C. R. Menyuk, “Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence,” J. Lightwave Technol., vol. 14, pp. 148–157, 1996.

[20] V. E. Zakarov and A. B. Shabat, Sov, Phys.JETP 34, 62(1972).

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

[22] R.H Stolen and J.E. Bjorkholm, “Parametric amplification and frequency conversion in optical fiber,” J.Quantum Electron., vol. 18, pp. 1062-1072, July 1982.

[23] Q. Zhang, G. Namburu, S. Karri, and P. V. Mamyshev, "Optimization of transceiver signal clipping in polarization multiplexing QPSK (PM QPSK) systems with Nyquist signals," 2013 IEEE Photon. Conf. (IPC), pp. 517– 518, 2013.

[24] Q. Zhang, S. Shrestha, R. Rashid, S. Karri, M. Khaliq, "An efficient split-step optical fiber simulation package with global simulation accuracy control," Proceedings 2013 IEEE/CIC International Conference on Communications in China (ICCC), pp. 158–164, 2013.

[25] S. G. Evangelides, L. F. Mollenauer, J. P. Gordon and N. **S.** Bergano, "Polarization Multiplexing with Solitons," Journal of Light wave Technology, Vol. 10, No. 1, January 1992.

[26] A. Yariv, "Optical Electronics in Modern Communications," Oxford University Press, Inc., New York, NY, 1996.

[27] Y.R. Shen, "The principles of Nonlinear Optics," John Wiley and Sons, New York, 1984.

[28] A. Ghatak and K. Thayagarajan, "Introduction to Fiber Optics," Cambridge University Press, Cambridge, UK, 1998.

[29] J. D. Jackson, "Classical Electrodynamics," John Wiley and Sons, New York, 1999.

[30] A. Yariv, "Quantum Electronics," John Wiley and Sons, New York, 1997.

[31] T. R. Taha and M. J. Ablowitz, "Analytical and numerical aspects of certain nonlinear evolution equations part I: Analytical," J. Comp. Phys. 55, 192-202 (1984).

[32] Q. Zhang, V. Winstead, J. Lian and L. Xing, "Simulation step-size optimization in the split-step Fourier simulation of polarized optical signal propagation through single mode optical fiber," 9th International Conference on Communications and Networking in China (CHINACOM), 289-292 (2014).

[33] D. Gloge, "weekly guiding fibers," Appl. Opt., Vol. 10, pp. 2252, 1971.

[34] Zhang, Q., Menyuk, C. R., Hayee, M. I., Tavva, C., Huang, H.-W., Nair, R., and Khaliq, M.: Computation step-size study in digital post-compensation of fiber

optic communication systems. 43rd Conference on Information Science and Systems (CISS), FA4, (2009).

[35] C. A. Balanis, “Advanced Engineering Electromagnetics," Wiley, 1989.

[36] Q. Zhang, M. I. Hayee, S. Nadimpalli, V. Winstead, X. Wu, D. Huang, J. Lian, and M. Khaliq, “Application of local error method to SSSF simulation of vector propagation in dispersion compensated optical links,” Springer Photon. Network. Comm. Journal, march 2016.

[37] Q. Zhang, S. Nadimpalli, V. Winstead, S. Shrestha, X. Wu, P. S. Megat Hamari, H. Min, and J. Lian, “WDM simulation of vector propagation in dispersion compensated SSMF links,” accepted conference paper.

[38] R. Hui and M. O. Sullivan, “Fiber Optic Measurement Techniques,” Elsevier Academic Press, 2009.