Study of Fiber Nonlinear Effects on Fiber Optic Communication Systems

---

A Ph. D. Dissertation

Presented to

the Faculty of School of Engineering and Applied Sciences

University of Virginia

In Partial Fulfillment

of the Requirements for the Degree

Doctorate of Philosophy

(Electrical Engineering)

---

by

Bo Xu

August, 2003

APPROVAL SHEET

This dissertation is submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Electrical Engineering

Bo Xu, Author

This Dissertation has been read and approved by the examining committee:

Prof. Maïté Brandt-Pearce, Advisor

Prof. Yibin Zheng, Committee Chairman

Prof. Ira Jacobs (Virginia Tech)

Prof. Stephen G. Wilson

Prof. Tommy R. Guess

Accepted for the School of Engineering and Applied Science:

Dean, School of Engineering and Applied Science

August, 2003

Abstract

This dissertation studies different effects from the fiber Kerr nonlinearity on optical fiber communication systems with the focus on efficient methods to compute the system degradations from these nonlinear effects.

We first review and develop several analysis tools which are essential to the study of optical fiber systems. These analysis tools focus on solving the nonlinear Schrödinger (NLS) equation that describes the propagation of optical pulses through optical fiber channels. Even though the NLS equation can be solved numerically by the split-step Fourier (SSF) method, analytical solutions like the Volterra series transfer function (VSTF) method are still preferable for system performance evaluation when stochastic parameters of the signal and fiber channel are considered. The modified VSTF methods is introduced to solve the energy divergence problem associated with the truncated third-order VSTF method, offering a more accurate result.

We then study the two most important nonlinear effects, cross-phase modulation (XPM) and four-wave mixing (FWM), in the wavelength-division multiplexing (WDM) systems based on the VSTF method. These two nonlinear effects both cause intensity fluctuation of a channel from its neighboring channels. The same system model is used for the study of both effects and consequently direct comparisons between the two effects is possible in our study. As an indication of the system performance, we compute the variance of the intensity fluctuation for both synchronous and asynchronous WDM systems in this dissertation. Different methodologies to reduce the nonlinear effects based on optimizing the system parameters is also discussed.

Another approach to improve the system performance of WDM systems based on mul-

tiuser detection is introduced in this dissertation too. The key idea is to make use of the information carried by the channel correlations due to fiber nonlinearity by the use of multiuser detection. The analysis is significantly complicated by the presence of the photodetector, a square-law device, in the system. Using the Gaussian approximation after its validation, both error performance of the multiuser detector and its asymptotic behavior in such systems are thoroughly studied by upper- and lower-bounds and compared with the conventional single-user detector.

The noise process in a communication system must be known in order to compute the system performance. However, due to nonlinear interaction between the amplified spontaneous emission (ASE) noise from optical amplifiers and the signal, the ASE noise is being amplified by the signal during propagation. The noise nonlinear amplification with any arbitrarily modulated signal is studied based on the perturbation theory up to the second-order in nonlinearity for both real and imaginary parts of the noise where we fully characterize the output noise process by the correlation function in the frequency domain. Significant noise amplification and energy transfer between the real and imaginary parts are identified. The validation range of our result is also discussed with some simple application given for system performance study.

Presently, optical fiber communication systems are evolving from 10Gb/s to 40Gb/s per channel. At 40Gb/s, since the signals have wider bandwidth and higher signal power, the dispersion compensation must be carefully optimized to control the interaction between the dispersion and nonlinearity. However, different device parameters have been used in previous system performance studies. We focus here on the effect of different device parameters on the dispersion compensation optimization and we found that both system performance

and optimal dispersion compensation vary significantly depending on the parameters used for transmitters, receivers, fibers and dispersion compensators in the study.

With all our studies, we hope to better understand the optical fiber communication systems and to achieve better design for such systems.

Acknowledgements

Obtaining a Ph. D. degree is always a challenging task as if frustrations, unexpected, and depression can happen at any time since it is supposed to be a hard task as we are studying a part of the unknown world, which is always a mystery. At this point, when I take a look at the past four years, I find so many people to whom I am grateful.

I would like to thank my dissertation advisor, Prof. Maite Brandt-Pearce, for her guidance and support over the past four years. Her encouragement during my frustration and depression time made the difficult and long-time Ph. D. program easier, tolerable, and possible to finish.

I would then like to thank my Ph. D. program committee members, Prof. Wilson, Prof. Zheng, Prof. Guess and Prof. Jacobs (Virginia Tech), for their generous help and numerous advices during every stage of my program. I have learned from them not only the knowledge but also the invaluable methodologies to solve problems.

I would also like to thank the other students from our group with whom I can chat and from whom I can learn. Their friendship and kindness made my four years stay a enjoyable one.

Lastly, this hard work is impossible to finish without the continuous love and support from my family. To them, I dedicate this dissertation.

Contents

List of Figures

Chapter 1

Introduction

Since the development of the first-generation optical fiber communication systems in the early 80's, the optical fiber communication technology has developed rapidly to achieve larger transmission capacity and longer transmission distance, partly due to the increasing demand of computer network. The demand on the increasing system and network capacity is expected to remain as more bandwidth-needing technologies like video conference and real-time image transmission emerge. To keep in step with the capacity increasing requirement, new devices and technologies are in great need. On the other hand, the systems must be carefully studied and designed.

The most important transmission impairments associated with long-distance high-bit-rate optical fiber communication systems include fiber dispersion, fiber Kerr nonlinearities, fiber polarization mode dispersion (PMD), noise accumulation from amplifiers and the interaction between them. How to achieve optimal systems is being studied extensively around the world. However, the complicated interaction between the impairments make the study on the optical fiber communication systems a challendging task. We focus on the study of fiber nonlinear effect on the fiber optic communication systems in this dissertation and PMD is then neglected in the following.

The key objective of this dissertation is to study the fiber Kerr nonlinear effects on optical fiber communication systems. One traditional approach to improve the system performance of WDM systems is to optimize the system parameters by reducing the nonlinear interference from other channels. Two dominating nonlinear effects in present WDM systems are cross-phase modulation (XPM) and four-wave mixing (FWM). Their effects on WDM systems are thoroughly studied and methods to reduce their effects are also discussed in this dissertation.

Another approach to improve the system performance of a WDM system is to make use of the nonlinearity-induced correlation between channels (users) by using multiuser detection. The error performance of such multiuser detector and its asymptotic behavior are included in this dissertation.

The interaction between the amplified spontaneous emission (ASE) noise and the signal is also studied in this dissertation. Due to the signal-dependent nature of the output noise after its interaction with the signal, the noise is characterized by its cross-correlation.

In this chapter, a simplified optical fiber communication system is first given in Section 1.1. Then the nonlinear Schrödinger (NLS) equation which describes the slowly varying complex envelope of the optical field in the fiber is introduced in Section 1.2. Finally, Section 1.3 gives the organization of this dissertation.

1.1 A Simplified System Model

As most optical fiber communication system nowadays use intensity modulation-direct detection (IM/DD), we focus on this kind of systems in this dissertation. A simple multi-span fiber system is shown in Figure 1.1.

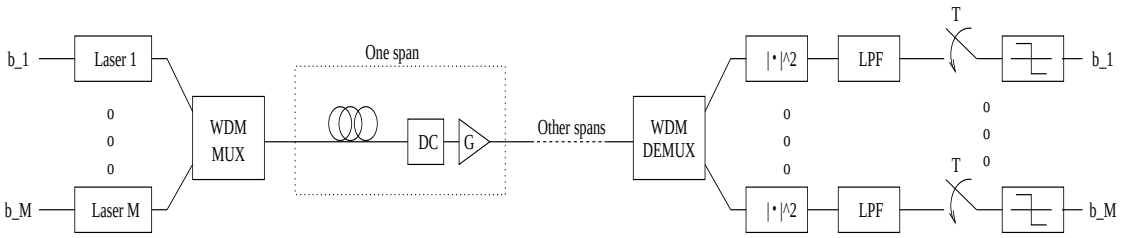


Figure 1.1: A schematic of a fiber optical communication system.

At the transmitter side, single or multiple lasers are used as the transmitters which send out on-off modulated pulses streams over M frequencies. Directly modulating laser diodes or external modulators can be used. A power booster based on Erbium-doped fiber amplifier (EDFA) or semiconductor optical amplifier (SOA) is also used right after the transmitter in some systems. If WDM is used, a multiplexer (MUX) is needed before transmitting the signals through fiber spans.

Each fiber span consists of a piece of single-mode fiber (SMF) or dispersion-shifted fiber (DSF) followed by an optical amplifier (G) and (possibly) a dispersion compensator (DC). The optical amplifier is used to compensate for the fiber attenuation with the cost of introducing a considerable amount of ASE noise. Since fiber dispersion broadens the signal pulses, a dispersion compensator is usually used to reduce this effect. Each fiber span could be the same or could be different. An especially important case is systems using dispersion management where fiber spans with dispersion of opposite signs are both used to achieve optimal operation exploiting the interaction between dispersion and nonlinearities. However, as was said earlier, the optimal system design is not known yet and it also strongly depends on the system parameters. We will further address this issue in this dissertation.

At the receiver side, the signal is passed through a WDM demultiplexer (DEMUX), which is effectively an optical filter bank, to separate the signals, if WDM is used. In

each branch the signal is photodetected and sampled at the bit rate, then passed through a threshold device to restore the information bits. The photodetector can usually be modeled as a square-law receiver whose output is proportional to the intensity of the optical signal. The optimal receiver using a photodetector is to put an optical matched-filter (MF) before the photodetector. However, optical filters with such narrow bandwidth are usually not available with present techniques. What is usually done is to put a narrow-band electrical filter before the sampler to further reduce the noise from the much wider optical filter. As always, synchronization is needed for the receiver to work correctly.

In this dissertation, we focus on the nonlinear effects from the fiber. To do this, the transmitter is usually modeled to send out Gaussian or super-Gaussian pulses with controlled chirping. The receiver assumes perfect synchronization together with optical and electrical filters of commonly used bandwidth.

1.2 Nonlinear Schrödinger Equation

As always, a mathematical model is needed to describe the physical transmission channel. The nonlinear Schrödinger (NLS) equation is found to serve this role and it is used to describe the slowly-varying envelope of the optical field. Neglecting the polarization effect, the NLS equation written in a reference frame moving at the group velocity of the pulse is given by [1]

$$\frac{\partial A}{\partial z} = -\frac{\alpha}{2}A + \frac{j}{2}\beta_2\frac{\partial^2 A}{\partial t^2} + \frac{1}{6}\beta_3\frac{\partial^3 A}{\partial t^3} - j\gamma\left[|A|^2A + \frac{j}{\omega_c}\frac{\partial}{\partial t}|A|^2A - T_RA\frac{\partial(|A|^2)}{\partial t}\right], \quad (1.1)$$

where $A = A(t, z)$ is the slowly varying complex envelope of the optical field at time t and position z along the fiber. α is the fiber loss, β_2 and β_3 are the second- and third-order dispersion coefficients respectively. These three effects are usually combined and referred

to as the linear term of the NLS equation. γ is the nonlinearity coefficient of the fiber with γ/ω_c and γT_R representing the strength of the self-steepening effect and stimulated Raman scattering (SRS) effect respectively. This equation approximates experimental results very well for optical pulses of duration over 50fs.

A simplified version of the NLS equation is given as [1]

$$\frac{\partial A}{\partial z} = -\frac{\alpha}{2}A + \frac{j}{2}\beta_2\frac{\partial^2 A}{\partial t^2} - j\gamma|A|^2A, \tag{1.2}$$

where self-steepening and SRS effects are neglected, which is valid whenever the pulses have duration longer than 1ps. This simplified version of the NLS equation is used throughout this dissertation. For systems transmitting at 10Gb/s or lower, as long as the channel is not at the wavelength where the second-order dispersion is zero (the zero-dispersion wavelength), the third-order dispersion term can also be neglected. For higher bit-rate than 10Gb/s, the effects from the third-order dispersion term will be discussed in the following chapters whenever applicable.

To give an idea of how the linear and nonlinear terms affect IM/DD systems, we study two simple cases first in this section. For the first one, we assume that $\gamma = 0$. Then (1.2) can be solved directly in the frequency domain where we get

$$A(\omega, z) = \exp(-\frac{\alpha z}{2} - j\frac{\beta_2\omega^2 z}{2})A(\omega, 0) \tag{1.3}$$

where $A(\omega, z)$ is the Fourier transform of $A(t, z)$. In this case, the fiber acts as a linear filter with attenuation and phase shift. Both the attenuation and the phase shift depend on the transmission length. These two effects can be compensated by an optical amplifier and a dispersion compensator respectively.

For the second case we assume that $\beta_2 = 0$, then (1.2) can be solved to give

$$A(t,z) = A(t,0)\exp(-\frac{\alpha z}{2})\exp\left(-j\gamma|A(t,0)|^2\frac{1-\exp(-\alpha z)}{\alpha}\right). \tag{1.4}$$

In this case, if a single channel is transmitted, the nonlinearity introduces a time-varying phase called self-phase modulation (SPM), $\gamma|A(t,0)|^2\frac{1-\exp(-\alpha z)}{\alpha}$, to the input pulses. However, for an IM/DD systems, a phase does not introduce any penalty to the system.

However, if both dispersion and nonlinearity are present, we will have interaction between these two effects. As a result, phase modulation/intensity modulation (PM/IM) conversion occurs and the SPM is converted into intensity fluctuation which affects the system performance of a direct detection system. Unfortunately, a simple analytical solution to the NLS equation does not exist for this case and we will thoroughly study this case in the next chapter.

1.3 Outline of the Dissertation

As said in the title of this dissertation, the key objective of this dissertation is to study the fiber nonlinear effects on optical communication systems. The dissertation is organized as follows.

In Chapter 2, we develop several analysis tools which are essential to the study of the optical fiber systems. These analysis tools focus on how to solve the NLS equation. They are the split-step Fourier (SSF) method, the Volterra series transfer function (VSTF) method, and the modified VSTF (MVSTF) method. The SSF method provides an efficient numerical approach to solve the NLS equation. The VSTF method gives an analytical solution to the NLS equation based on the Volterra series expansion. The MVSTF is a modified version of the VSTF which is more accurate than the truncated third-order VSTF

method.

Chapter 3 is dedicated to the study of nonlinear effects in WDM systems based on the VSTF method. Two most important nonlinear effects in WDM systems are XPM and FWM which both introduce intensity fluctuation from the interfering channels. The variance of the intensity fluctuation is found for both synchronous and asynchronous WDM systems.

In Chapter 4, another approach based on multiuser detection to improve the performance of WDM systems is thoroughly studied by the means of the error performance and the asymptotic behavior as signal-to-noise (SNR) increases under the Gaussian approximation. The key idea behind this is that since the nonlinear processes are deterministic given the information bits sent across channels, the optical WDM systems should benefit from multiuser detectors.

Due to the nonlinear interaction between the ASE noise and the signal, the noise is amplified by the signal. We study the noise nonlinear amplification from any arbitrarily modulated signal in Chapter 5. At the output, the noise is signal-dependent and non-stationary. We describe the output noise process by its cross-correlation function in the frequency domain.

In Chapter 6, we study the optimization of dispersion compensation for 40Gb/s systems. We focus on the effects of using different device parameters on the dispersion compensation optimization.

The last chapter focuses on conclusions and future work. Due to the fast-developing technologies, one can never finish the exploration. Together these seven chapters describe the work for my Ph.D. dissertation.

Chapter 2

Modified Volterra Series Transfer Function Method

We develop several tools for analyzing optical fiber communication systems in this chapter. We give a brief discussion on the split-step Fourier (SSF) method and the Volterra series transfer function (VSTF) method first, then we focus on the Modified VSTF method which gives more accurate analytical results than the truncated third-order VSTF method.

2.1 Split-Step Fourier Method

As mentioned in the introduction, the NLS equation is used to describe the slowly varying complex envelope of the optical field in the fiber and it can be written as

$$\frac{\partial A}{\partial z} = -\frac{\alpha}{2}A + \frac{j}{2}\beta_2\frac{\partial^2 A}{\partial t^2} - j\gamma|A|^2A. \tag{2.1}$$

This equation is valid for optical pulses with duration no shorter than 1ps. When shorter pulses (or wider bandwidth) are used, extra terms from self-steepening and stimulated Raman scattering (SRS) should be included. (2.1) is valid for most of the work in this dissertation unless otherwise indicated.

The NLS equation is generally hard to solve analytically except for some special cases and numerical computation is widely used instead. The most widely used numerical method

to solve the NLS equation is the SSF method. Detailed discussion on the SSF method can be found in [1]. The basic idea of SSF method is to divide the fiber span into many steps. Within each step, the linear and nonlinear effects are applied consecutively. To see how the SSF method works, we rewrite (2.1) as

$$\frac{\partial A(t,z)}{\partial z} = (\hat{D} + \hat{N})A(t,z) \tag{2.2}$$

where $\hat{D}$ and $\hat{N}$ are operators corresponding to linear and nonlinear effects respectively. In particular,

$$\hat{D} = -\frac{\alpha}{2} + \frac{j}{2}\beta_2\frac{\partial^2}{\partial t^2}, \tag{2.3}$$

$$\hat{N} = -j\gamma|A|^2. \tag{2.4}$$

Then within each step, we have

$$\begin{aligned} A(t,z+h) &= \exp(h(\hat{D}+\hat{N}))A(t,z) \\ &\approx \exp(h\hat{D})\exp(h\hat{N})A(t,z) \\ &= \mathcal{F}^{-1}\left\{\exp(-\frac{\alpha h}{2})\exp(-j\frac{\beta_2\omega^2 h}{2})\mathcal{F}\{A(t,z)\exp(-j\gamma|A(t,z)|^2 h)\}\right\} \end{aligned} \tag{2.5}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ stand for the Fourier and inverse Fourier transform. It is clear that the solution from the SSF method converges to the exact solution of (2.1) as the step size decreases to 0. It is also known that the error introduced in each step is in the order of h^2 in (2.5). In practice, in order for the SSF to be accurate enough, the step size h should satisfy $\gamma \max|A(t,z)|^2 h < 0.005$ rad. Throughout this dissertation, the results from the SSF simulation are assumed to be the exact solutions to the NLS equation and are used as a basis for comparison with the results from our theoretical analysis.

Several variants of the SSF method exist to reduce the computation complexity or the computation error. To reduce the computation error, a symmetric SSF method can be used. To reduce the computation time, an adaptive step size can be used as the signal power is reduced along the transmission due to fiber attenuation.

2.2 Volterra Series Transfer Function Method

Even though numerical methods to solve the NLS equation exist, an analytical solution to the NLS equation is still preferred. In [2], Peddanarappagari and Brandt-Pearce derived the VSTF as an approximate analytical solution to the NLS equation in the frequency domain. To do this, we first write (2.1) in the frequency domain as

$$\frac{\partial A(\omega,z)}{\partial z} = (-\frac{\alpha}{2}-\frac{j}{2}\beta_2\omega^2)A(\omega,z)-\frac{j\gamma}{4\pi^2}\iint A(\omega_1,z)A^*(\omega_2,z)A(\omega-\omega_1+\omega_2,z)\,d\omega_1\,d\omega_2 \tag{2.6}$$

where $A(\omega,z)$ is the Fourier transform of $A(t,z)$. After applying the Volterra series method and substituting

$$\begin{aligned} A(\omega,z) &= H_1(\omega,z)A(\omega,0) \\ &+ \iint H_3(\omega_1,\omega_2,\omega,z)A(\omega_1,0)A^*(\omega_2,0)A(\omega-\omega_1+\omega_2,0)\,d\omega_1\,d\omega_2 \\ &+ \ldots \end{aligned} \tag{2.7}$$

into (2.6), we finally obtain

$$H_1(\omega,z) = \exp(-\frac{\alpha z}{2})\exp(-j\frac{\beta_2\omega^2 z}{2}) \tag{2.8}$$

and

$$H_3(\omega_1,\omega_2,\omega,z) = -j\frac{\gamma}{4\pi^2}H_1(\omega,z)\frac{1-\exp(-\alpha z - j\beta_2(\omega_1-\omega)(\omega_1-\omega_2)z)}{\alpha+j\beta_2(\omega_1-\omega)(\omega_1-\omega_2)} \tag{2.9}$$

as the first- and third-order Volterra series kernel coefficients. (2.8) represents the linear effect from fiber loss and dispersion. (2.9) represents the nonlinear effect. Due to high computational complexity associated with integration, we truncate the VSTF to third order.

The first- and third-order VSTF series kernel coefficients can also be derived by using the same approach as the SSF method under the assumption that the nonlinear effect is small compared with the linear effects. We assume that a fixed step size is used, and later we let the step size go to zero. For the first step, we have

$$A(t,h) = \mathcal{F}^{-1}\left\{\exp(-\frac{\alpha h}{2})\exp(-j\frac{\beta_2\omega^2 h}{2})\mathcal{F}\{A(t,0)\exp(-j\gamma|A(t,0)|^2 h)\}\right\} \tag{2.10}$$

$$\approx \mathcal{F}^{-1}\left\{\exp(-\frac{\alpha h}{2})\exp(-j\frac{\beta_2\omega^2 h}{2})\mathcal{F}\{A(t,0)(1-j\gamma|A(t,0)|^2 h)\}\right\} \tag{2.11}$$

or in the frequency domain

$$\begin{aligned} A(\omega,h) \approx\ & \exp(-\frac{\alpha h}{2})\exp(-j\frac{\beta_2\omega^2 h}{2})A(\omega,0) - j\gamma h\exp(-\frac{\alpha h}{2})\exp(-j\frac{\beta_2\omega^2 h}{2}) \\ & \iint A(\omega_1,0)A^*(\omega_2,0)A(\omega-\omega_1+\omega_2,0)\,d\omega_1\,d\omega_2 \end{aligned} \tag{2.12}$$

Note that we have used the approximation that $\exp(-j\gamma|A(t,0)|^2 h) \approx 1 - j\gamma|A(t,0)|^2 h$ above. Following the same steps and keeping up to the third-order term only, we obtain

$$\begin{aligned} A(\omega,nh) \approx\ & \exp(-\frac{\alpha nh}{2})\exp(-j\frac{\beta_2\omega^2 nh}{2})A(\omega,0) - j\gamma h\exp(-\frac{\alpha nh}{2})\exp(-j\frac{\beta_2\omega^2 nh}{2}) \\ & \iint \sum_{k=0}^{n-1}\exp(-\alpha kh - j\beta_2(\omega_1-\omega_2)(\omega_1-\omega)kh) \\ & A(\omega_1,0)A^*(\omega_2,0)A(\omega-\omega_1+\omega_2,0)\,d\omega_1\,d\omega_2 \\ =\ & \exp(-\frac{\alpha nh}{2})\exp(-j\frac{\beta_2\omega^2 nh}{2})A(\omega,0) - j\gamma h\exp(-\frac{\alpha nh}{2})\exp(-j\frac{\beta_2\omega^2 nh}{2}) \\ & \iint \frac{1-\exp(-\alpha nh - j\beta_2(\omega_1-\omega_2)(\omega_1-\omega)nh)}{1-\exp(-\alpha h - j\beta_2(\omega_1-\omega_2)(\omega_1-\omega)h)} \end{aligned}$$

$$A(\omega_1, 0)A^*(\omega_2, 0)A(\omega - \omega_1 + \omega_2, 0)\, d\omega_1\, d\omega_2.$$

Now letting $h \to 0$ and $z = nh$ as the total transmission distance, we get

$$\begin{aligned} A(\omega, z) \approx\ & \exp(-\frac{\alpha z}{2}) \exp(-j\frac{\beta_2\omega^2 z}{2}) A(\omega, 0) - j\gamma \exp(-\frac{\alpha z}{2}) \exp(-j\frac{\beta_2\omega^2 z}{2}) \\ & \iint \frac{1 - \exp(-\alpha z - j\beta_2(\omega_1 - \omega_2)(\omega_1 - \omega)z)}{\alpha + j\beta_2(\omega_1 - \omega_2)(\omega_1 - \omega)} \\ & A(\omega_1, 0)A^*(\omega_2, 0)A(\omega - \omega_1 + \omega_2, 0)\, d\omega_1\, d\omega_2 \end{aligned} \tag{2.13}$$

which gives the first- and third-order Volterra kernel coefficients of the VSTF method.

It has been shown that the VSTF method is equivalent to the perturbation method [12, 13]. As an important application of the analytical solution to the NLS equation, Narimanov and Mitra have calculated the channel capacity of a fiber optics communication system based on perturbation theory [31].

Unfortunately, the method has only limited application because the predicted or modeled optical pulse energy diverges quickly when the input power is large. Thus a method to solve the energy divergence problem is required to extend the application range of the truncated VSTF method. In the next section, we present a way to modify the truncated third-order VSTF method. The modified VSTF (MVSTF) method can successfully remove the energy divergence and thus gives more accurate results than the truncated third-order VSTF method.

2.3 Modified Volterra Series Transfer Function Method

To get the third-order Volterra kernel, we have used the assumption that $\exp(-j\gamma|A(t, z)|^2 h) \approx 1 - j\gamma|A(t, z)|^2 h$. This assumption fails if the signal power is large enough, then higher-order terms should be included in this case. What the MVSTF method does is to

restore the truncated nonlinear terms back into the phase. To do this, rewrite (2.13) as

$$A(\omega, z) \approx A_L(\omega, z) + A_{NL}(\omega, z) \tag{2.14}$$

where $A_L(\omega, z)$ and $A_{NL}(\omega, z)$ correspond to the first- and third-order terms in (2.13) respectively, Then the modification is made as follows

$$A(t,z) = \begin{cases} A_L(t,z)\exp\left(\frac{A_{NL}(t,z)}{A_L(t,z)}\right) & \text{if } |A_{NL}(t,z)| < |A_L(t,z)| \\ A_L(t,z) + A_{NL}(t,z) & \text{if } |A_{NL}(t,z)| > |A_L(t,z)| \end{cases} \tag{2.15}$$

where $A_L(t, z)$ and $A_{NL}(t, z)$ are the inverse Fourier transformation of $A_L(\omega, z)$ and A_{NL} (ω, z) respectively. Note that at the rising and falling edge of the pulses where the input signal is very small, we keep the original VSTF simply to avoid dividing $A_{NL}(t, z)$ by zero. For single-channel systems, $A_{NL}(t, z)$ is simply the self-phase modulation (SPM) term. We show later that the MVSTF method gives a much more accurate result than the truncated third-order VSTF method.

It can also be shown that the MVSTF method agrees with theoretical results for the important case of a small β_2. Due to fiber attenuation, the nonlinear effect decreases as the signal power decreases along the transmission. Thus, the nonlinear effect is only important at the front-end of a fiber. The rest of the fiber acts as a dispersive-only fiber and we assume that the dispersion is compensated by a dispersion compensator at the output of the fiber. During the part where the nonlinear effect is important, we write

$$A(t,z) = A(t,0)\exp(\zeta(t,z))\exp(j\theta(t,z))\exp(-\frac{\alpha z}{2}). \tag{2.16}$$

For simplicity, we assume that the input $A(t, 0)$ is real in the following derivation.

Substituting (2.16) into (2.1), we get

$$\begin{aligned} A(t,0)\frac{\partial\theta(t,z)}{\partial z} = & - \gamma A^3(t,0)\exp(2\zeta(t,z))\exp(-\alpha z) + \frac{\beta_2}{2}\{A''(t,0) + 2A'(t,0)\zeta'(t,z) \\ & + A(t,0)\zeta''(t,z) + A(t,0)(\zeta'(t,z))^2 - A(t,0)(\theta'(t,z))^2\}, \end{aligned} \tag{2.17}$$

where $A'(t,0) = \frac{\partial}{\partial t}A(t,0)$, $\zeta'(t,z) = \frac{\partial}{\partial t}\zeta(t,z)$ and $\theta'(t,z) = \frac{\partial}{\partial t}\theta(t,z)$. Since β_2 is small, the dominating term in (2.17) is the first term, and the solution for the phase change becomes

$$\theta(t,z) = -\gamma A^2(t,0)\frac{1-\exp(-\alpha z)}{\alpha}. \tag{2.18}$$

This phase term is exactly the same as (1.4) when $\beta_2 = 0$.

Another equation we obtained by substituting (2.16) into (2.1) is

$$A(t,0)\frac{\partial\zeta(t,z)}{\partial z} = -\frac{\beta_2}{2}\left(2A'(t,0)\theta'(t,z) + A(t,0)\theta''(t,z) + 2A(t,0)\zeta'(t,z)\theta'(t,z)\right) \tag{2.19}$$

$$\approx -\frac{\beta_2}{2}\left(2A'(t,0)\theta'(t,z) + A(t,0)\theta''(t,z)\right). \tag{2.20}$$

When β_2 is small, the PM/IM conversion effect $\zeta(t,z)$ is small and is proportional to β_2. Thus, the last term on the right side of (2.19) can be dropped because it is smaller than the other terms. Substituting (2.18) into (2.20), we obtain

$$\zeta(t,z) \approx \frac{\gamma\beta_2}{\alpha}\left(3(A'(t,0))^2 + A(t,0)A''(t,0)\right)\frac{1-\exp(-\alpha z)}{\alpha} \tag{2.21}$$

since $z \ll \frac{1}{\alpha^2}$ for practical situations and $\zeta'(t,z)$ is small compared with the other terms.

Using the assumption that β_2 is small so that $\beta_2\omega_B^2 \ll \alpha$ where ω_B is the bandwidth of the signal, we can expand the denominator in the double integral of (2.13) as

$$\frac{1}{\alpha + j\beta_2(\omega_1-\omega)(\omega_1-\omega_2)} \approx \frac{1}{\alpha} - j\frac{\beta_2}{\alpha^2}(\omega_1-\omega)(\omega_1-\omega_2), \tag{2.22}$$

where we have used the simplification that $\exp(-\alpha z) \ll 1$. Note also that $H_1(\omega) = 1$ due to the dispersion compensator used. The first term on the right side of (2.22) determines the nonlinear phase, and the second term determines the amplitude change. It is easy to show that the two terms of (2.22), after they are converted into the time-domain and divided by

$A(t,0)$, are the same as (2.18) and (2.21) respectively, thereby showing the validity of the third-order MVSTF for the case of a small β_2. It is also very easy to show that the MVSTF gives the correct analytical solution to the NLS equation when $\beta_2 = 0$.

We now compare the results from the SSF method, the VSTF method and the MVSTF method for different β_2. We first show simulation results of the MVSTF method for a single channel, single pulse case for both a single span and 10 spans of 100km each. The system model we simulate is included in Figure 2.1. The amplifier compensates for the fiber loss and the dispersion compensator compensates for the linear dispersion from the optical fiber. The input is a chirpless single Gaussian pulse. The output pulses from the SSF simulation, VSTF and MVSTF methods are compared in Figure 2.2. The energy of the output pulses for the SSF, original VSTF and the MVSTF method are given for the first span in Table 2.1. The improvement by the MVSTF is clearly seen in Figure 2.2 and from the above energy comparison. The improvement is even more dramatic for 10 spans where the result from VSTF method diverges greatly (beyond the limit of our numerical simulator) and is therefore not included in the plot.

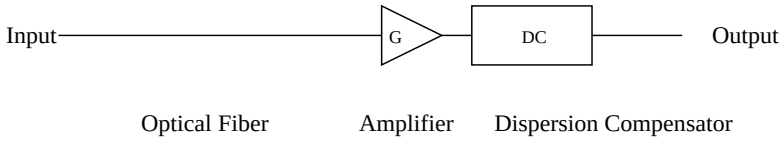


Figure 2.1: System model of one span. $G = \exp(\frac{\alpha z}{2})$.

The situation gets more complicated in WDM system. As always, one can take the multiple channels as a big single-channel and use the above modification. However, this scheme hides the details on the cross-phase modulation (XPM) and four-wave-mixing (FWM) from the interactions between channels. Figure 2.3 shows the output of a two-

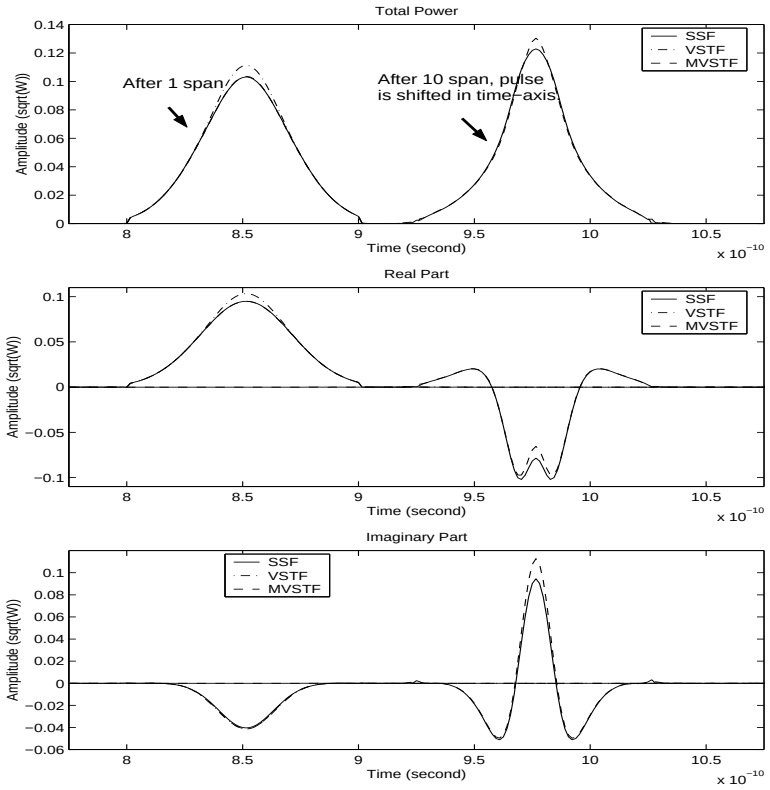


Figure 2.2: Comparison between the fiber outputs computed using SSF, VSTF and MVSTF for single channel, single pulse case. Simulation parameters are $\beta_2 = 2\text{ps}^2/\text{km}$, $\alpha = 0.2\text{dB/km}$, input peak power $P_0 = 10\text{mW}$, $\gamma = 2\text{W}^{-1}\text{km}^{-1}$, span length $= 100\text{km}$.

channel system taken as a big single-channel. The total energies of the output pulses are also given in Table 2.1.

For any multi-channel system, suppose that the Fourier transform of the input can be written as $A(\omega, 0) = \sum_{k=0}^{M-1} A_k(\omega, 0)$, where $A_k(\omega, 0)$ is the input at channel k of M channels; then $A_{NL}(\omega, z)$ consists of $\binom{M}{3}$ terms from the double integral of (2.13). For example,

$$A_{k,SPM}(\omega, z) = \int\int H_3(\omega_1, \omega_2, \omega, z) A_k(\omega_1, 0) A_k^*(\omega_2, 0) A_k(\omega - \omega_1 + \omega_2, 0)\, d\omega_1\, d\omega_2$$

is the SPM term for channel k. Now instead of lumping all signals together as a single channel, we can do channel by channel by modifying the third-order VSTF expansion for

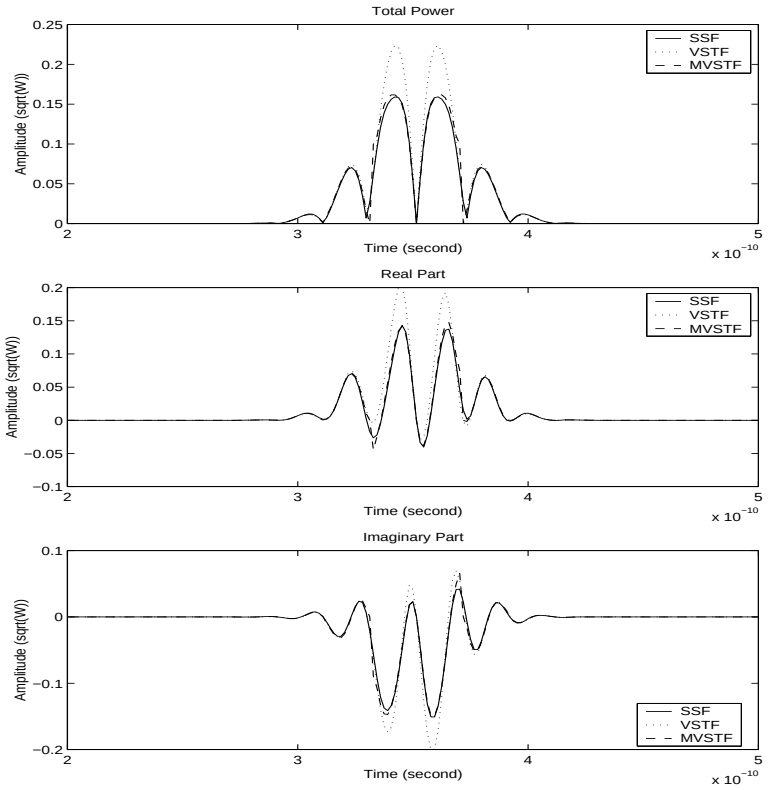


Figure 2.3: Comparison between fiber output computed using SSF, VSTF and MVSTF for two-channel, single pulse case. Simulation parameters are $\beta_2 = -2\text{ps}^2/\text{km}$, $\alpha = 0.2\text{dB/km}$, $P_0 = 10\text{mW}$, $\gamma = 2\text{W}^{-1}\text{km}^{-1}$, single span of length $= 50\text{km}$.

the output at channel k as

$$A_k(t,z) = \begin{cases} A_{k,L}(t,z)\exp\left(\frac{A_{k,NL}(t,z)}{A_{k,L}(t,z)}\right) & \text{if } |A_{k,NL}(t,z)| < |A_{k,L}(t,z)| \\ A_{k,L}(t,z) + A_{k,NL}(t,z) & \text{if } |A_{k,NL}(t,z)| > |A_{k,L}(t,z)| \end{cases} \tag{2.24}$$

where $A_{k,NL}(t,z)$ is the sum of the nonlinear terms for channel k including its SPM term $A_{k,SPM}(t,z)$, the XPM terms from other channels and any FWM terms on it. The results for the central channel of a three-channel WDM system with equal channel spacing and random input bits for each of the pump channel are shown in Figure 2.4 for a single span. It is clearly seen that the MVSTF method performs equally well as the SSF method and successfully solves the energy divergence problem of the original truncated third-order VSTF method.

	1 Channel	2 Channels
SSF	3.5449 $\times 10^{-13}$J	7.0894 $\times 10^{-13}$J
VSTF	3.9194 $\times 10^{-13}$J	11.8150 $\times 10^{-13}$J
MVSTF	3.5511 $\times 10^{-13}$J	7.6820 $\times 10^{-13}$J

Table 2.1: Comparison between fiber output pulse energy computed using SSF, VSTF and MVSTF with single pulse input for both single-channel and two-channel cases.

2.4 Multiple Spans

Both the VSTF and MVSTF can be applied span by span to multiple-span systems, as the SSF method does. An example is shown in Figure 2.2 with a single pulse for a single channel system. On the other hand, the transfer function up to the third-order for the whole system can be found, then (2.13) is calculated. The latter approach requires much less computation, and its accuracy is checked in this section comparing with simulation results and the results from a span-by-span calculation.

The multiple-span system studied in this dissertation consists of a concatenation of spans shown in Figure 2.1. A post-dispersion compensator is used at the fiber output to remove the accumulated dispersion if a partial dispersion compensator is used in each span to compensate partially the fiber dispersion. The VSTF kept up to the third-order is found with $H_1(\omega, z = N_T L) = 1$ due to post-dispersion compensation and

$$\begin{aligned} H_3(\omega_1, \omega_2, \omega, z = N_T L) &= -\frac{j\gamma}{4\pi^2} \frac{1 - \exp(-\alpha L - j\beta_2(\omega_1 - \omega)(\omega_1 - \omega_2)L)}{\alpha + j\beta_2(\omega_1 - \omega)(\omega_1 - \omega_2)} \\ &\times \sum_{k=0}^{N_T - 1} \exp(-jkD(\omega_1 - \omega)(\omega_1 - \omega_2)) \end{aligned} \tag{2.25}$$

where N_T is the number of total spans, L is the span length and D is the accumulated dispersion for each span. The MVSTF is again obtained by modifying VSTF using (2.15) and (2.24) respectively.

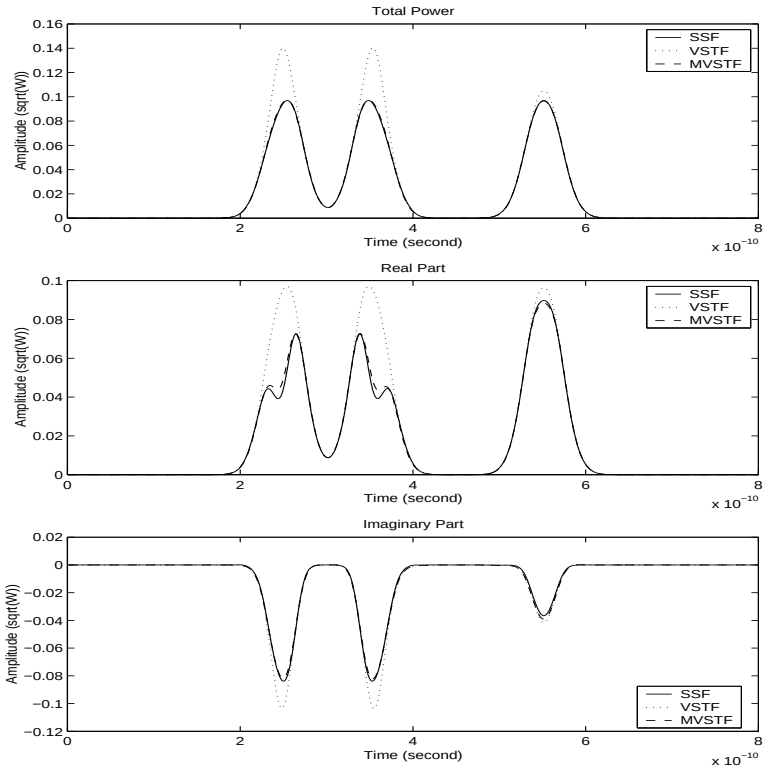


Figure 2.4: Comparison between fiber output computed using SSF, VSTF and MVSTF for the central channel of a three-channel system. Simulation parameters are $\beta_2 = -8\text{ps}^2/\text{km}$, $\alpha = 0.2$dB/km, $P_0 = 10$mW, $\gamma = 2$ W^{-1}km^{-1}, single span of length $= 50$km. Note that random input bits of 0 and 1 are used for the three channels.

We compare the VSTF and MVSTF results with the SSF simulation results for two different systems. The first system consists of a concatenation of fiber spans without dispersion compensation for each span while the second system consists a concatenation of fiber spans with 100% dispersion compensation for each span. The peak power is set to be 2mW. For a fair comparison, we define the normalized energy deviation (NED) as

$$NED = \frac{\int \mid A_{ana}(t,z) - A_{sim}(t,z) \mid^2 dt}{\int \mid A_{sim}(t,z) \mid^2 dt} \tag{2.26}$$

where $A_{ana}(t,z)$ is the analytical results obtained from the VSTF or MVSTF method with span-by-span calculation or with total transfer function and $A_{sim}(t,z)$ is obtained from the SSF simulation. The NEDs are shown in Figures 2.5 and 2.6 for the two systems separately

as span number increases.

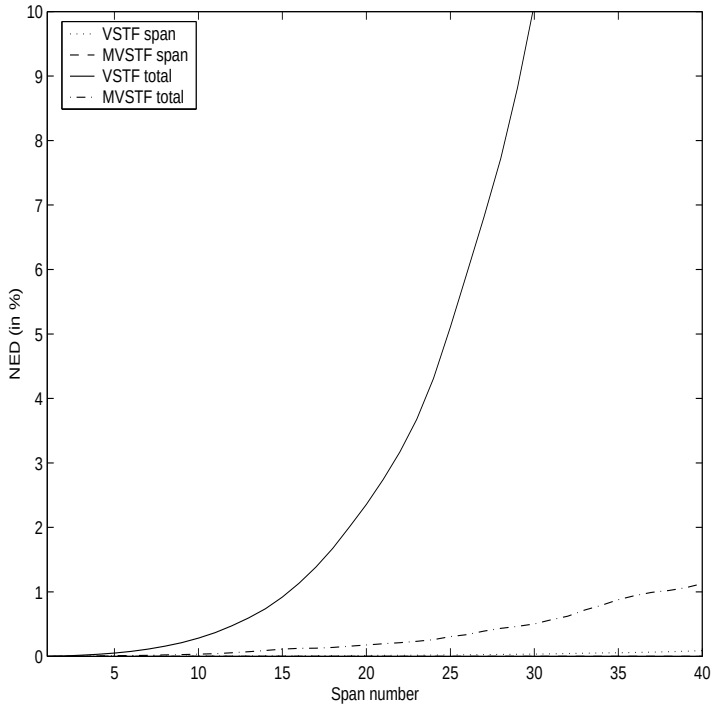


Figure 2.5: Normalized energy deviation (NED) for multiple-span system without dispersion compensation for each span.

In either system, the MVSTF with span-by-span calculation gives negligible error and the MVSTF with total transfer function is much more accurate than the truncated third-order VSTF with total transfer function. As channel increases, the errors of all the four approaches also increase as expected, but at different rates. The four approaches are found to work much better for the system without dispersion compensation for each span than for the system with 100% dispersion compensation for each span. The results for the system with 100% dispersion compensation for each span is believed to have more error because the nonlinear effects from each span for this system add constructively with each other and thus are much stronger than for the other system.

In conclusion, we have shown that a simple modification to the truncaed third-order

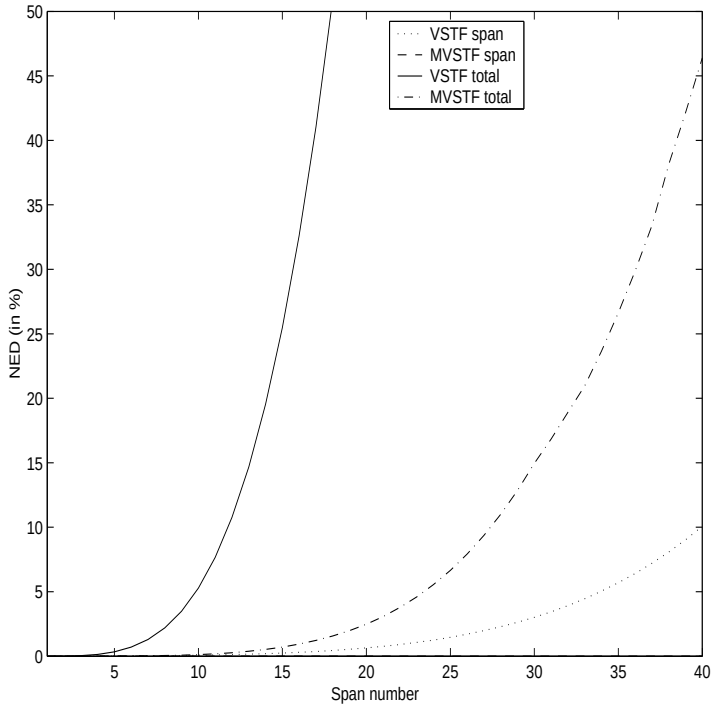


Figure 2.6: Normalized energy deviation (NED) for multiple-span system with 100% dispersion compensation for each span.

VSTF method can successfully solve the energy divergence problem of the VSTF method and can improve the model accuracy by orders of magnitude. The MVSTF method can extend the application range of the original VSTF method significantly while inheriting all the advantage of the VSTF method. It offers a better method for finding an analytical approximation to the overall transfer function of the cascade of all the elements of an optical link, which can be used to determine optimal system parameters. In the following chapters, we study the nonlinear effects on the fiber optical communication systems based on the analytical tools developed in this chapter.

Chapter 3

Cross-Phase Modulation and Four-Wave Mixing

As the capacity requirement of optical fiber communication systems increase, channels with smaller channel spacing will be used in wavelength-division multiplexing (WDM) systems. As a result, the two dominant nonlinear effects, cross-phase modulation (XPM) and four-wave mixing (FWM), become more and more pronounced. Both nonlinear effects introduce intensity fluctuations which are dependent on the neighboring channels, thus causing inter-channel interference (ICI) at the receiver. In this chapter, new variance analyses for both XPM- and FWM-induced intensity distortion based on the Volterra series transfer function (VSTF) method are studied and compared.

The split-step Fourier (SSF) method, which is the mostly commonly used method for fiber optic communication system design and analysis, is not well suited to variance analysis from nonlinearity-induced intensity fluctuation. It can be used to study XPM- and FWM-induced crosstalk in WDM systems, but the computation time increases in the order of at least M^3, where M is the total number of channels in the system for each simulation run. Moreover, many simulation trials should be run in order to obtain a good estimate for the statistics of the receiver observations when taking into account the possible random

natures of the channel phases and the channel delays. Suppose the random channel phases and channel delays are modeled as uniformly distributed with N_θ different phases and N_D different channel delays, then the total computation time with the SSF method grows exponentially with respect to the number of users as $(N_\theta N_D)^M M^3$. On the other hand, as an analytical approach, the VSTF method does not need extra computation with random channel phases and delays. Thus, the computation time with the VSTF method only increases polynomially with the number of channels.

Great efforts have been spent on the theoretical analysis of XPM- and FWM-induced crosstalk. The intensity fluctuations caused by XPM have been studied previously for systems with a continuous wave (CW) probe channel under the interference of a modulated pump channel [3, 4, 5]. These models use a small signal approximation and assume that the pump channel is undistorted by dispersion in the fiber. We call the previous approach the power approximation in this dissertation since in their results the XPM-induced crosstalk depends only on the power of the pump channels. The validity of this analysis based on a CW probe has been verified experimentally [5, 6, 7]. However, the power approximation predicts the same XPM-induced intensity distortion for pumps with different input frequency chirping. In this chapter we show that the distortion depends on the frequency chirping when the frequency chirping is large enough.

FWM has been studied with channels of CW format, with emphasis on predicting the FWM efficiency [8, 9]. [10] extends the analysis on FWM by forming a statistical model including random phases and random bits of each channel. The analysis based on CW signals is a worst-case analysis where each channel is assumed to remain at the pulse peak. The performance evaluation from this worst-case analysis is usually too conservative com-

pared with real systems. A new model which takes into account the modulated nature of the channels is offered in this chapter for a better performance evaluation.

The intensity fluctuations caused by XPM and FWM are compared in [11] for a WDM system, where different models are used for the two distortions. A CW probe plus modulated pumps are used for the XPM effect while a CW probe plus CW pumps are used for the FWM effect. It is not clear whether the comparison is fair: the result may exaggerate the effect from FWM.

In this chapter, we give new analysis for both XPM- and FWM-induced intensity distortion based on the VSTF method [2]. The XPM analysis does not rely on the assumption of undistorted pump channels, and the FWM analysis can be used with modulated pumps; the models for the two effects are identical.

This chapter is organized as follows. We use the VSTF method to derive both the XPM- and FWM-induced intensity distortion in Section 3.1. In Section 3.2, we first verify our methods by comparing them with the SSF simulation for a 3-channel single-span WDM system. We then study and compare the two nonlinear effects as the number of channels increases for both single-span synchronous WDM systems and single-span asynchronous WDM systems where the variance of the nonlinearity-induced intensity fluctuation is found for both systems. Section 3.3 extends the results into multispan systems based on the VSTF method using the approximate total transfer function of the system. We also briefly discuss the accuracy of this method for multispan WDM systems in this section. We summarize our work in Section 3.4.

3.1 Analysis

In WDM systems, the input consists of multiple channels and we let $A(\omega,0)=\sum_{k=-M}^{M} A_k(\omega-k\Delta\omega,0)$, where $A_k(\omega,0)$ is the baseband input of the k-th channel and $\Delta\omega$ is the channel spacing. Here we assume WDM systems of equal channel spacing for simplicity; a similar analysis holds for WDM systems with unequal channel spacing. After substituting $\sum_{k=-M}^{M} A_k(\omega-k\Delta\omega,0)$ into (2.13), we obtain $(2M+1)^3$ terms containing the factors of the form $A_{k_1}(\omega_1,0)A^*_{k_2}(\omega_2,0)A_{k_3}(\omega-\omega_1+\omega_2,0)$ in the double integral of (2.13) where k_1, k_2, and k_3 correspond to the index of any channel between $-M$ and M. The SPM, XPM, and FWM terms are found to be [13]

$$A_{k_1}A^*_{k_2}A_{k_3} \text{ is } \begin{cases} \text{SPM on Channel } k_1 & \text{if } k_1=k_2=k_3, \\ \text{XPM on Channel } k_3 & \text{if } k_1=k_2\neq k_3, \\ \text{XPM on Channel } k_1 & \text{if } k_1\neq k_2=k_3, \\ \text{FWM on Channel } k_1-k_2+k_3 & \text{otherwise.} \end{cases} \tag{3.1}$$

Our analysis assumes that channel 0 is a CW probe channel with input $A_0(t,0)=\sqrt{P_0}$ while the other channels are equally spaced on both sides of the central channel, i.e. the total input to the fiber can be written as $A(t,0)=\sum_{k=-M,k\neq 0}^{M} A_k(t,0)\exp(jk\Delta\omega t)+\sqrt{P_0}$ where $A_k(t,0)$ is an amplitude modulated baseband signal for the k-th channel.

The SPM effect must be studied using a modulated signal as was done using the VSTF method by Peddanarappagari and Brandt-Pearce in [2, 14]. The SPM effect is therefore not considered here; the SPM effect on a CW probe channel is simply a constant nonlinear phase which has no effect on the output of a IM/DD system. The SPM effects of the modulated pumps are also neglected here because their effects on the XPM and FWM of the probe channel are small and due to a higher order nonlinear term.

3.1.1 Analysis of XPM

The model we study in this section is included in Figure 3.1 for a single-span fiber with optional post-dispersion compensator. With the CW probe, the XPM term from the k-th channel to the probe channel at point A in Figure 3.1 is found to be from $(k_1, k_2, k_3) = (0, k, k)$ or $(k, k, 0)$ in (3.1), giving a term

$$A_{X,k}(\omega, L) = -2j\gamma \exp(-\frac{\alpha}{2}L) \exp(-\frac{j}{2}D\omega^2)\sqrt{P_0}$$
$$\frac{1}{2\pi}\int \frac{1-\exp(-(\alpha + j\beta_2(k\Delta\omega)\omega + j\beta_2\omega\omega_1 - j\beta_2\omega^2)L)}{\alpha + j\beta_2(k\Delta\omega)\omega + j\beta_2\omega\omega_1 - j\beta_2\omega^2} A_k(\omega_1, 0)A_k^*(\omega_1 - \omega, 0)\, d\omega_1 \tag{3.2}$$

where L is the fiber span length and D is the accumulated dispersion for the span. When no dispersion compensator is used after the fiber span, $D = \beta_2 L$; when a perfect dispersion compensator is used, $D = 0$. Note that a factor of 2 is included for the XPM term in (3.2) to account for both term $(0, k, k)$ and $(k, k, 0)$. Under a small signal approximation, the intensity distortion of the probe channel induced by the XPM from the k-th channel after the square-law detector can then be approximated by

$$\begin{aligned} \Delta P_{X,k}(t, L) &= \left|\sqrt{P_0}\exp\left(-\frac{\alpha}{2}L\right) + A_{X,k}(t, L)\right|^2 - P_0\exp(-\alpha L) \\ &\approx \sqrt{P_0}\exp\left(-\frac{\alpha}{2}L\right)(A_{X,k}(t, L) + A_{X,k}^*(t, L)), \end{aligned} \tag{3.3}$$

neglecting the much smaller term $|A_{X,k}(t, L)|^2$ and the beat terms among the XPM and FWM terms. Note that in the previous chapter, we have proposed a modified VSTF (MVSTF) method which gives a more accurate solution to the NLS equation than the VSTF method. If MVSTF method is used, the output with the XPM effect will be written as $\sqrt{P_0}\exp(-\frac{\alpha}{2}L)\exp(\sum A_{X,k}(t, L)/(\sqrt{P_0}\exp(-\frac{\alpha}{2}L)))$. Then, the intensity distortion of

the probe channel induced by the XPM is

$$\begin{aligned}\Delta P_{X,k}(t,L) &= \left|\sqrt{P_0}\exp(-\frac{\alpha}{2}L)\exp\left(\frac{\sum_k A_{X,k}(t,L)}{\sqrt{P_0}\exp(-\alpha L/2)}\right)\right|^2 - P_0\exp(-\alpha L)\\ &\approx P_0\exp(-\alpha L)\left(\exp\left(\frac{\sum_k A_{X,k}(t,L)+\sum_k A^*_{X,k}(t,L)}{\sqrt{P_0}\exp(-\alpha L/2)}\right)-1\right)\\ &\approx \sqrt{P_0}\exp\left(-\frac{\alpha}{2}L\right)\sum_k (A_{X,k}(t,L)+A^*_{X,k}(t,L)) \qquad (3.4)\end{aligned}$$

which is the same as (3.3) once linearization is used.

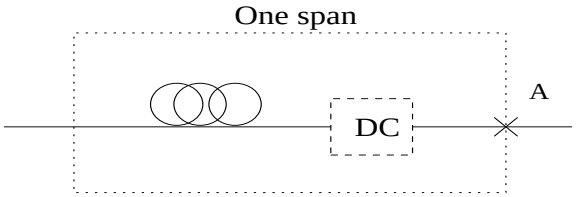


Figure 3.1: Simplified system model for single-span WDM system. DC is the optional dispersion compensator.

Using the same assumption as in [5] that $\exp(-\alpha L) \ll 1$ and that the modulation bandwidths of the pump signals are much smaller than the channel spacing, we obtain

$$A_{X,k}(\omega,L) \approx -2j\gamma\frac{\exp(-\frac{\alpha}{2}L)\exp(-\frac{j}{2}D\omega^2)\sqrt{P_0}}{\alpha+j\beta_2(k\Delta\omega)\omega}\frac{1}{2\pi}\int A_k(\omega_1,0)A_k^*(\omega_1-\omega,0)\,d\omega_1 \qquad (3.5)$$

$$= -2j\gamma\frac{\exp(-\frac{\alpha}{2}L)\exp(-\frac{j}{2}D\omega^2)\sqrt{P_0}}{\alpha+j\beta_2(k\Delta\omega)\omega}P_k(\omega,0). \qquad (3.6)$$

When substituting (3.6) into (3.3), we obtain

$$\Delta P_{X,k}(\omega,L) \approx 4\gamma P_0\exp(-\alpha L)P_k(\omega,0)\frac{\sin(D\omega^2/2)}{\alpha+j\beta_2(k\Delta\omega)\omega}. \qquad (3.7)$$

This result is exactly the same as (6) in [5] with $D=\beta_2 L$ when dispersion compensation is not used and $\beta_2(k\Delta\omega)$ is the walk-off defined in [5]. The different sign in the denominator of (3.7) can be traced back to the different definition of the NLS equation in (2.1); another

difference is that we have used a frame of reference moving with the pulse at the group velocity, thus the propagation delay is also absent in (3.7). Note that our analysis does not assume an undistorted pump (as done in [5]), thus it expands the previous result to cases where the dispersion effect on the pulse shape can not be neglected, i.e., when β_2 is not small or when the input pulses have significant pre-chirping [15].

When a perfect dispersion compensator is used, $D = 0$, (3.5) and (3.6) are oversimplified and give zero XPM-induced intensity distortion. Either (3.2) or the result from the power approximation in [5]

$$A_{X,k}(\omega, L) \approx -2j \frac{\gamma \exp(-\frac{\alpha}{2}L) \exp(-\frac{j}{2}D\omega^2) P_k(\omega, 0)\sqrt{P_0}}{\alpha + j\beta_2(k\Delta\omega)\omega - \frac{j}{2}\beta_2\omega^2} \tag{3.8}$$

can be used in this case. (3.8) gives good results as long as the input pump channels have negligible frequency chirping. One source of input frequency chirping comes from the laser chirping associated with the transmitters. The SSF method of simulation is used to check the accuracy of the VSTF-based XPM analysis, (3.2), and the power approximation, (3.8), for a simple 2-channel system with chirped Gaussian input pulses for the pump channel. From the results shown in Figure 3.2, when the pump channel input has significant frequency chirping, (3.2) should be used instead of (3.8). It is also clearly seen from the plot that prechirping can be used to reduce the XPM effects as studied in [16].

3.1.2 Analysis of FWM

The VSTF method can also be applied to study FWM. From (3.1), for any FWM term to fall on the probe channel at channel 0, $k_2 = k_1 + k_3$ must hold. As a result, the FWM term from the channel combination (k_1, k_2, k_3) with $k_2 = k_1 + k_3$ is found to be

$$A_{F,k_1,k_3}(\omega, L) \quad = \quad -j\gamma \exp\left(-\frac{\alpha}{2}L\right) \exp\left(-\frac{j}{2}D\omega^2\right)$$

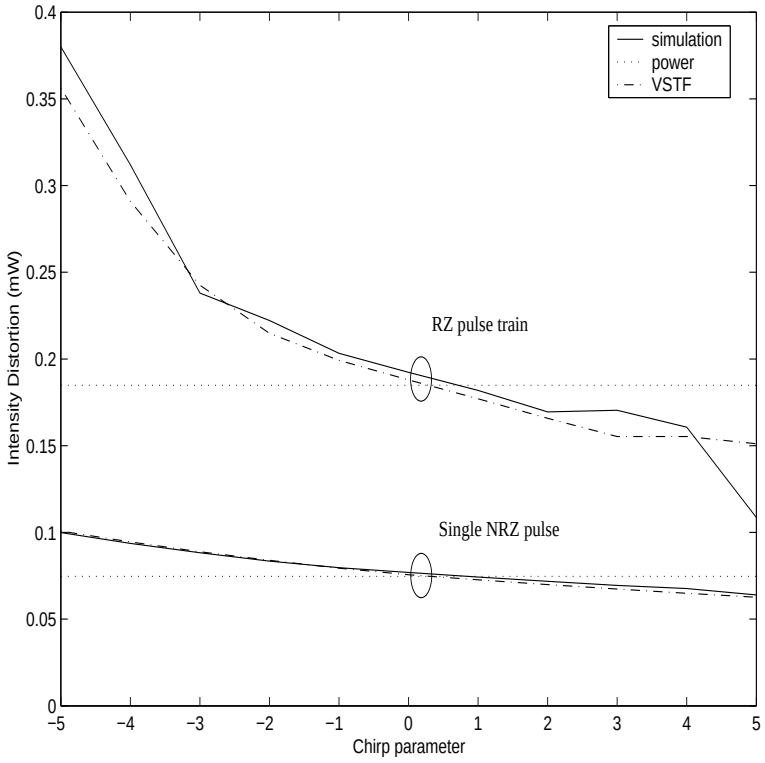


Figure 3.2: Comparison of XPM-induced intensity distortion using simulation, power approximation and VSTF methods with chirped pump input.

$$\times \quad \frac{1}{4\pi^2}\iint \frac{A_{k_1}(\omega_1,0)A^*_{k_1+k_3}(\omega_2,0)A_{k_3}(\omega-\omega_1+\omega_2,0)}{\alpha+j\beta_2(\omega_1+k_1\Delta\omega-\omega)(\omega_1-k_3\Delta\omega-\omega_2)}\,d\omega_1\,d\omega_2 \tag{3.9}$$

where we have used the assumption that $\exp(-\alpha L) << 1$ as before. When the bandwidths of the pump signals are much smaller than the channel spacing, since $k_1 \neq 0$ and $k_3 \neq 0$ for the FWM term to fall on the central channel, we can simplify (3.9) to obtain

$$\begin{aligned} A_{F,k_1,k_3}(\omega,L) \quad &\approx \quad -j\gamma\frac{\exp(-\frac{\alpha}{2}L)\exp(-\frac{j}{2}D\omega^2)}{\alpha-jk_1k_3\beta_2\Delta\omega^2} \\ &\qquad \frac{1}{4\pi^2}\iint A_{k_1}(\omega_1,0)A^*_{k_1+k_3}(\omega_2,0)A_{k_3}(\omega-\omega_1+\omega_2,0)\,d\omega_1\,d\omega_2, \end{aligned} \tag{3.10}$$

or, equivalently,

$$\begin{aligned} A_{F,k_1,k_3}(t,L) \quad &\approx \quad -j\frac{\gamma\exp(-\frac{\alpha}{2}L)}{\alpha-jk_1k_3\beta_2\Delta\omega^2} \\ &\qquad \mathcal{F}^{-1}\left\{\exp(-\frac{j}{2}D\omega^2)\mathcal{F}\{A_{k_1}(t,0)A^*_{k_1+k_3}(t,0)A_{k_3}(t,0)\}\right\} \end{aligned} \tag{3.11}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ stand for the Fourier and the inverse Fourier transform respectively. The intensity distortion of the probe channel from the $(k_1, k_1 + k_3, k_3)$-th FWM term can then be approximated by

$$\Delta P_{F,k_1,k_3}(t, L) \approx \sqrt{P_0}\exp(-\frac{\alpha}{2}L)(A_{F,k_1,k_3}(t, L) + A^*_{F,k_1,k_3}(t, L)). \tag{3.12}$$

After substituting $A_{k_1}(t, 0)$, $A_{k_1+k_3}(t, 0)$ and $A_{k_3}(t, 0)$ with CW signals, one can obtain the well-known CW FWM phase matching efficiency as (4) in [9]. (The only difference is that the dispersion slope coefficient is neglected here; it can be easily incorporated into the analysis.) However, our analysis can be applied to any general modulated pump signals. When CW pumps are used, the dispersion effect introduces only a constant phase, whilst, when modulated pumps are used, the dispersion can broaden the FWM pulses generated and lower the interference power. One can therefore not accurately predict the power of the interference from FWM by using CW pumps.

One interesting result coming directly from (3.11) is that the FWM vanishes whenever the input pumps do not overlap in time no matter how large the walk-off is within the span. This result holds as long as the channel spacing dominates over the signal bandwidth. The effectiveness of reducing FWM effects by reducing pulse overlapping has been studied in [17] for the zero-dispersion wavelength region and has been used in [18].

3.2 Single-Span WDM Systems

3.2.1 Model Verification

As a first step, we check the accuracy of our analytical results on XPM- and FWM-induced intensity fluctuations by comparing with simulation results using the SSF method for a very simple 3-channel single-span WDM system, as shown in Figure 3.1, followed by an optical

amplifier to compensate for the fiber transmission loss. The system consists of a CW probe channel at channel 0 and two independent intensity modulated channels at frequencies $-\Delta\omega$ and $\Delta\omega$ respectively. The two pump channels are aligned in time at the input, giving a worst-case XPM- and FWM-induced intensity fluctuations. For this system configuration, the FWM terms possible on the central channel are from terms $(k_1, k_2, k_3) = (-1, 0, 1)$ and $(1, 0, -1)$.

Figure 3.3 shows the intensity fluctuation of the probe channel at the output of the fiber span for systems both with and without dispersion compensation. In both cases, the pump channels consist of independent random bit sequences of 10Gb/s chirpless return-to-zero (RZ) pulses of Gaussian shape with full width at half maximum (FWHM) of 35ps. Other parameters used are $\gamma = 2.2\text{W}^{-1}\text{km}^{-1}$, $\beta_2 = -2.6\text{ps}^2/\text{km}$, $\alpha = 0.25\text{dB/km}$, pump channel pulse peak power $P = 2\text{mW}$, probe channel power $P_0 = 2\text{mW}$, span length $L = 80\text{km}$ and channel spacing $\Delta f = 50\text{GHz}$. It is clearly seen that our analytical results are very close to the result obtained by the SSF method for both cases. For the parameters used in this example, the XPM-induced distortion adds to the FWM-induced distortion, but the two effects can cancel each other for other situations.

One interesting result that can be derived from (3.11) is that when the phases of the two chirpless pump channels add to $(k + \frac{1}{2})\pi$ for some integer k, the FWM term generated should have a very small real part. As a result, the FWM-induced intensity distortion should also be very small; the XPM-induced intensity distortion is independent of the phases of the pump channels. This hypothesis is confirmed in Figure 3.4 where the phases of the two pump channels are 0 and $\frac{\pi}{2}$ respectively. Comparing Figure 3.4 with Figure 3.3, the FWM-induced intensity distortion is significantly reduced by this input phase offset while

the XPM-induced intensity distortion remains the same. Another effective way to reduce the FWM-induced intensity distortion with RZ pulses is to intentionally introduce delays between channels, thus making the overlapping time between pump channels small at the input.

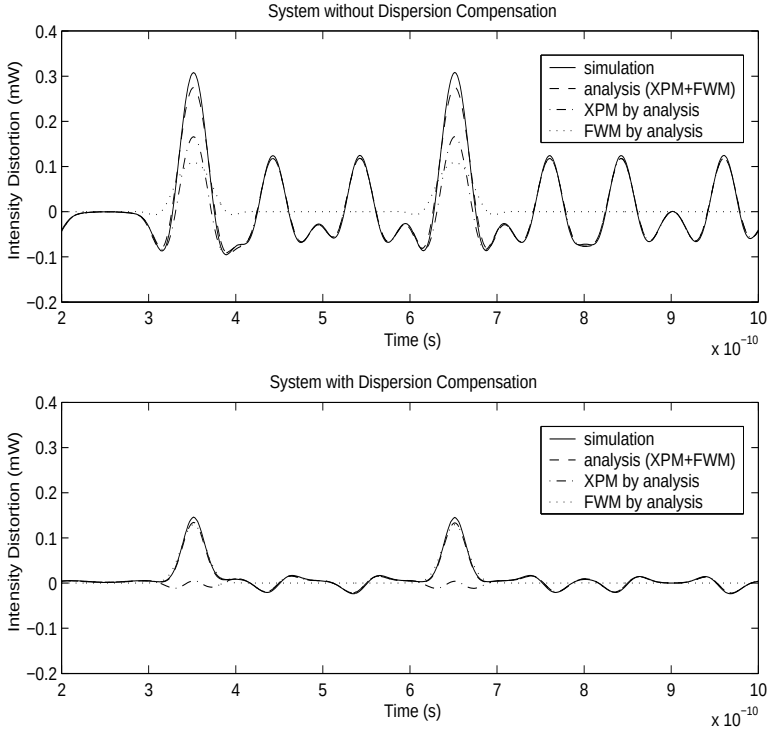


Figure 3.3: Comparison of intensity distortion between SSF simulation results and VSTF results for a 3-channel single-span WDM system. Both pump channels have zero input phases.

The most compelling advantage of our analytical approach is that since the system conditions are identical we can compare the FWM-induced intensity distortion with the XPM-induced intensity distortion fairly. For the 3-channel single-span WDM system, we first plot the normalized intensity distortion defined as $(\max[\Delta P] - \min[\Delta P])/(2P_0)$ [5] versus the percentage of dispersion compensation in Figure 3.5. The FWM-induced intensity distortion is almost constant over the percentage of dispersion compensation used while the

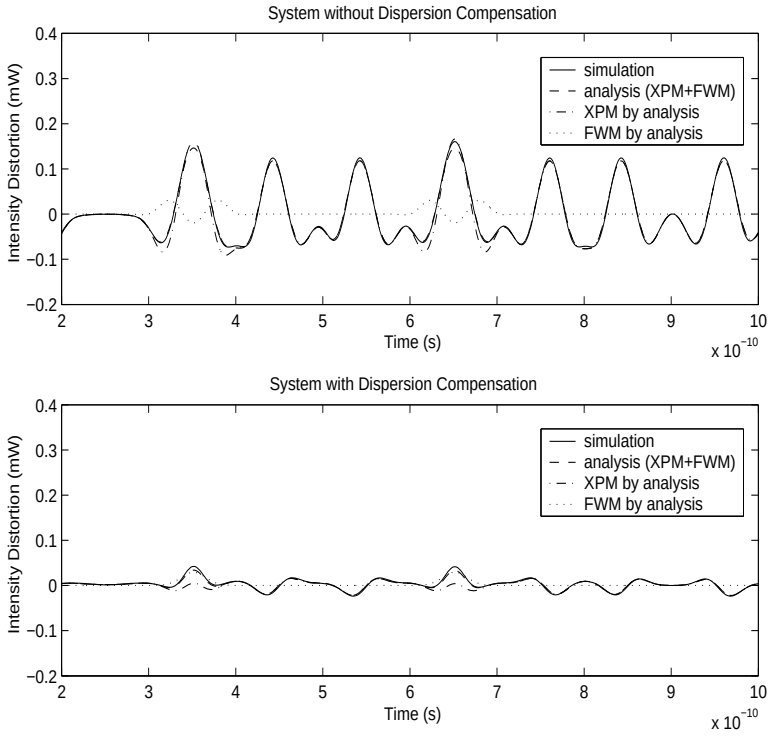


Figure 3.4: Comparison of intensity distortion between SSF simulation results and VSTF results for a 3-channel single-span WDM system. The input phases for the two pump channels are 0 and $\frac{\pi}{2}$ respectively.

XPM-induced intensity distortion shows a strong dependence on dispersion compensation. This behavior of FWM can be explained from (3.11): since the dispersion of our system, $\beta_2 = -2.6\text{ps}^2/\text{km}$, is small, it does not affect the overlapped pulse $A_{-1}(t,0)A_1(t,0)$ much even without dispersion compensation.

In Figure 3.6, we plot the normalized intensity distortion versus dispersion for systems both without and with dispersion compensation. As expected from the theoretical analysis, the FWM-induced intensity distortion falls much faster than the XPM-induced intensity distortion as $|\beta_2|$ increases when dispersion compensation is not used. From (3.11), due to the fact that $k_1 \neq 0$ and $k_3 \neq 0$, the dominating term in the denominator is $jk_1k_3\beta_2\Delta\omega^2$, thus the FWM distortion decreases as $1/\beta_2$. However, for XPM, the dominating term in the

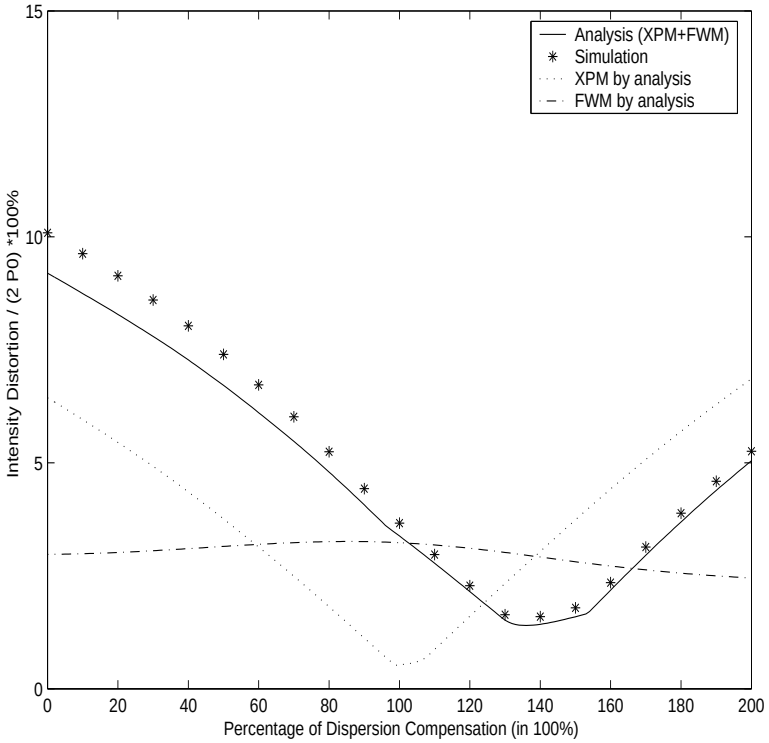


Figure 3.5: Intensity distortion vs. percentage of dispersion compensation for a 3-channel single-span WDM system.

denominator of (3.8) is α when ω is small, thus we expect the XPM effect to decrease much slower than $1/\beta_2$. When dispersion compensation is used, the FWM-induced intensity distortion dominates over the XPM-induced intensity distortion effect.

Figure 3.7 shows the normalized intensity distortion versus channel spacing for systems both without and with dispersion compensation. The plot confirms that the FWM-induced intensity distortion decreases as $1/\Delta\omega^2$ as predicted by (3.11) while the XPM-induced intensity distortion decreases much slower.

From the above analysis and comparison, we conclude that the XPM-induced intensity distortion dominates when dispersion compensation is not used while FWM-induced intensity distortion dominates when dispersion compensation is used for this 3-channel

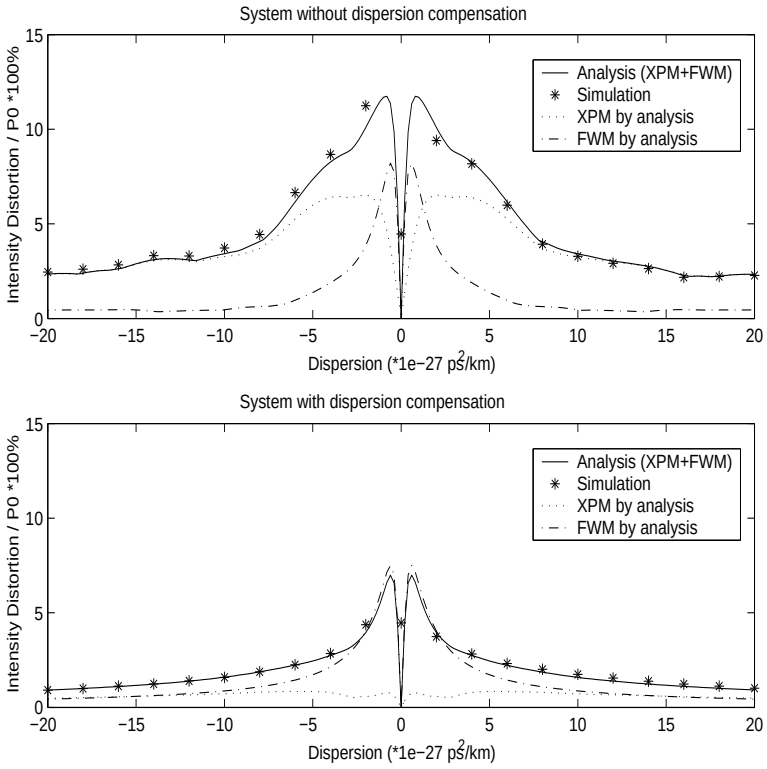


Figure 3.6: Intensity distortion vs. dispersion parameter β_2 for a 3-channel single-span WDM system.

single-span WDM case. The above analysis and comparison also show ways to reduce the XPM and FWM effects when designing WDM systems. It is better to use dispersion compensation to reduce XPM-induced intensity distortion and use optimal channel phases and channel time misalignment to lower the FWM-induced intensity distortion. As an example, in Figure 3.8 we show the eye-diagrams for the central channel of a 3-channel WDM system for four different cases with and without dispersion compensation and with and without input phase offset. For this example, the input of the central channel is also intensity-modulated instead of CW, as found in real situations. The parameters used in the SSF simulation are the same as for Figure 3.3 except that 100GHz channel spacing is used. A 5th-order Butterworth optical filter is used to filter out the central channel. The 100GHz

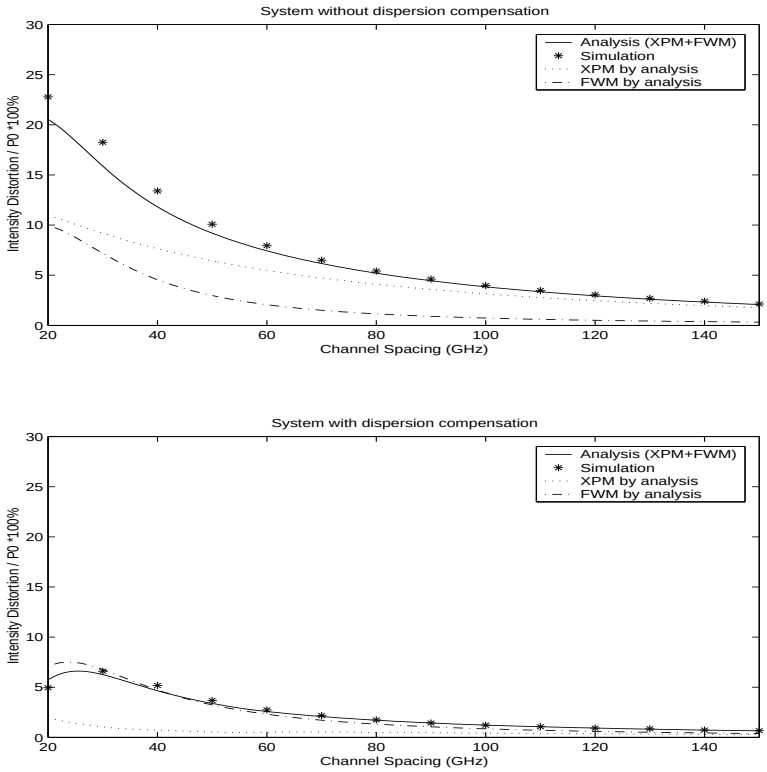


Figure 3.7: Intensity distortion vs. channel spacing for a 3-channel single-span WDM system.

channel spacing is used to avoid ICI introduced by optical filtering which dominates over the nonlinearity induced intensity distortion if 50GHz channel spacing is used. No narrow-bandwidth electrical filter is used to avoid introducing intersymbol interference (ISI). The effectiveness of our design on reducing the XPM and FWM effects is clearly seen.

3.2.2 Variance for Time Synchronized System

When the number of channels M increases, both XPM and FWM effects increase. The number of XPM terms generated on the central channel is M, and the number of FWM terms generated on the central channel is on the order of M^2. However, the magnitude of FWM terms decreases faster than the magnitude of XPM terms as the channel spacing

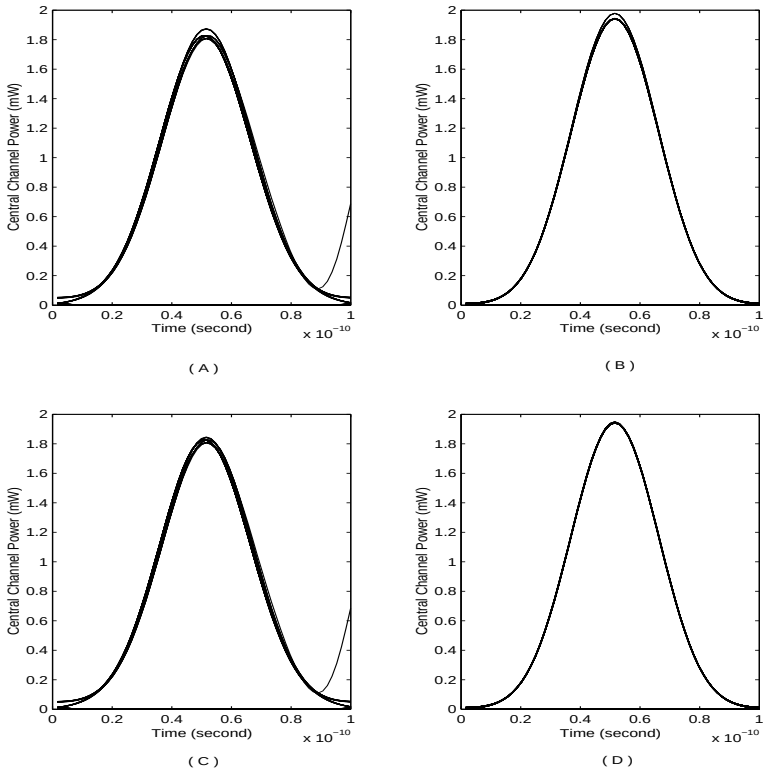


Figure 3.8: Comparison of eye-diagrams with modulated probe and two pump channels for single span. (A) without dispersion compensation with same channel input phases; (B) with dispersion compensation and same channel input phases; (C) without dispersion compensation with channel input phases arrangement $(0, 0, \frac{\pi}{2})$; (D) with dispersion compensation with channel input phases arrangement $(0, 0, \frac{\pi}{2})$.

increases. As a result, whether FWM- or XPM-induced intensity distortion dominates remains a question. In [11], Ten et. al. made a comparison of FWM and XPM penalties for a 40×10Gb/s system with channel spacing of 100GHz where they found that XPM penalties dominate. As discussed above, different system models for XPM- and FWM-induced intensity distortion which favor XPM are used. To avoid this bias, in this section, we use the same system model to study XPM- and FWM-induced intensity distortion and compare the two effects as the number of channels increases. We assume in the following that each channel is transmitting independent random bit sequence from each other and

that each channel has a random input phase that is uniformly distributed over $[-\pi, \pi]$. The central channel is still assumed to be a CW probe channel.

We first consider WDM systems where the channels are time synchronized at the input (or the channels have fixed delays at the input). We consider the asynchronous WDM system later in this section. One example of a synchronous WDM systems is one which uses external modulators driven by the same clock. WDM networks with add/drop operations are examples of asynchronous WDM systems. The system we model here and the next subsection is the same as the system used in the 3-channel model verification above. The only difference is that more channels are present now.

Suppose the input for channel k is $A_k(t,0) = \sum_{l_k} b_{k,l_k} g_k(t - l_k T_b - T_k) \exp(j\theta_k)$ where b_{k,l_k} is the l_k-th bit for channel k; $b_{k,l_k} = 0$ or 1 with probability $\frac{1}{2}$ and is independent of all other bits as appropriate for an uncoded system; $g_k(t)$ represents the pulse shape; T_b is the bit period, T_k is the fixed channel delay and θ_k is the channel input random phase which is assumed to be uniformly distributed over $[-\pi, \pi]$. Then, from (3.8), define $g_{X,k}(t, L) = -2j\gamma\sqrt{P_0}\mathcal{F}^{-1}\left\{\frac{\exp(-\frac{j}{2}D\omega^2)}{\alpha + j\beta_2 k\Delta\omega\omega - \frac{1}{2}\beta_2\omega^2}\mathcal{F}\{|g_k(t)|^2\}\right\}$ as the XPM pulse shape generated at the fiber output. We obtain the XPM on channel 0 due to channel k as

$$A_{X,k}(t, L) = \sum_{l_k} b_{k,l_k} g_{X,k}(t - l_k T_b - T_k, L). \tag{3.13}$$

Note that in obtaining (3.13), the fiber loss is assumed to be compensated by a perfect optical amplifier and the XPM pulses generated by overlapping of different pulses are very small and thus neglected. Then,

$$\begin{aligned} \Delta P_X(t, L) &= \sum_{k=-M, k\neq 0}^{M} \sqrt{P_0}(A_{X,k}(t, L) + A^*_{X,k}(t, L)) \\ &= \sum_{k=-M, k\neq 0}^{M} \sum_{l_k} b_{k,l_k} X_k(t - l_k T_b - T_k, L) \end{aligned} \tag{3.14}$$

where $X_k(t,L)$ is a deterministic function defined by $X_k(t,L) = \sqrt{P_0}(g_{X,k}(t,L)+g^*_{X,k}(t,L))$. The variance of the XPM-induced distortion, defined as $E[\Delta P_X^2(t,L)] - E^2[\Delta P_X(t,\,L)]$ over random phases and bits, [1] is easily found to be

$$\sigma_X^2(t,L) = \frac{1}{4}\sum_{k=-M,k\neq 0}^{M}\sum_{l_k} X_k^2(t-l_kT_b-T_k,L). \tag{3.15}$$

We define the normalized worst-case variance from XPM effects as

$$\hat{\sigma}_X^2 = \frac{1}{4}\max_t\left\{\sum_{k=-M,k\neq 0}^{M}\sum_{l_k} X_k^2(t-l_kT_b-T_k,L)/P_0^2\right\}, \tag{3.16}$$

which corresponds to sampling at the worst time.

For FWM-induced intensity distortion, we substitute $A_k(t,0) = \sum_{l_k} b_{k,l_k} g_k(t-l_kT_b - T_k)\exp(j\theta_k)$ into (3.11). For each channel combination (k_1,k_2,k_3) with $k_2 = k_1+k_3$, we obtain

$$\begin{aligned} A_{F,k_1,k_3}(t,L) &= \frac{-j\gamma}{\alpha - j\beta_2 k_1 k_3 \Delta\omega^2}\sum_{l_{k_1}}\sum_{l_{k_1+k_3}}\sum_{l_{k_3}} b_{k_1,l_{k_1}} b_{k_1+k_3,l_{k_1+k_3}} b_{k_3,l_{k_3}} \\ &\quad \exp(j(\theta_{k_1}+\theta_{k_3}-\theta_{k_1+k_3}))\mathcal{F}^{-1}\{\exp(-\frac{j}{2}D\omega^2)\mathcal{F}\{g_{k_1}(t-l_{k_1}T_b-T_{k_1}) \\ &\quad g^*_{k_1+k_3}(t-l_{k_1+k_3}T_b-T_{k_1+k_3})g_{k_3}(t-l_{k_3}T_b-T_{k_3})\}\} \end{aligned} \tag{3.17}$$

and

$$\begin{aligned} \Delta P_F(t,L) &= \sum_{k_1,k_3=-M}^{M}\sum_{l_{k_1}}\sum_{l_{k_1+k_3}}\sum_{l_{k_3}} b_{k_1,l_{k_1}} b_{k_1+k_3,l_{k_1+k_3}} b_{k_3,l_{k_3}} \\ &\quad \left\{F_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t,L)\exp(j(\theta_{k_1}+\theta_{k_3}-\theta_{k_1+k_3}))+c.c.\right\} \end{aligned} \tag{3.18}$$

where $F_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t,L)$ is a deterministic function defined by

$$\begin{aligned} F_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t,L) &= \frac{-j\gamma\sqrt{P_0}}{\alpha - j\beta_2 k_1 k_3\Delta\omega^2}\mathcal{F}^{-1}\{\exp(-\frac{j}{2}D\omega^2) \\ &\quad \mathcal{F}\{g_{k_1}(t-l_{k_1}T_b-T_{k_1})g^*_{k_1+k_3}(t-l_{k_1+k_3}T_b-T_{k_1+k_3})g_{k_3}(t-l_{k_3}T_b-T_{k_3})\}\}, \end{aligned} \tag{3.19}$$

[1] We use the notation $E[X]$ to mean the expected value of the random variable X.

and c.c. stands for complex conjugate. Note that most of the $F_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t,L)$ terms are zero because the corresponding pulses generated from the three channels have very little overlapping at the input. To find the variance, we use the fact that the contribution to the variance from each $F_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t,L)$ term is uncorrelated to obtain

$$\sigma_F^2(t,L) = \sum_{\substack{k_1,k_3=-M \\ k_1,k_3\neq 0 \\ |k_1+k_3|\leq M}}^{M} \sum_{l_{k_1}} \sum_{l_{k_1+k_3}} \sum_{l_{k_3}} C_{k_1,k_3} |F_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t,L)|^2 \tag{3.20}$$

with

$$C_{k_1,k_3} = \begin{cases} \frac{1}{2} & k_1 = k_3, \\ 1 & k_1 = -k_3, \\ \frac{1}{2} & \text{otherwise}. \end{cases} \tag{3.21}$$

We also define the normalized worst-case variance from the FWM effects as

$$\hat{\sigma}_F^2 = \max_t \left\{ \sigma_{\Delta P_F}^2(t,L)/P_0^2 \right\}. \tag{3.22}$$

In our discussion on 3-channel single-span WDM systems, we found that optimized channel phases can help to reduce the FWM effect. The same property can be extended to more channels. As the number of channels increases, the dominating FWM terms on the central probe channel are from channel combinations $(-1,0,1)$, $(1,2,1)$ and $(-1,-2,-1)$. To reduce the FWM effect from these three combinations, we need optimal channel phases arranged as $..., -\frac{\pi}{2}, -\frac{\pi}{2}, 0, 0, \frac{\pi}{2}, \frac{\pi}{2}, \pi, \pi, ...$

In Figure 3.9 and 3.10, we compare the normalized worst-case variance for XPM-induced and FWM-induced intensity distortion for a single-span system with 25GHz and 50GHz channel spacing respectively. Same parameters are used as for the 3-channel WDM system except that $P = 1$mW and $P_0 = 1$mW here. Both systems with and without dispersion compensation are studied. The channels at the input are aligned for worst-case analysis, that is, $T_k = 0$. The input pump channels we used are chirp-free, thus (3.2)

from our model and (3.8) from the power approximation gives very close results on the XPM-induced intensity distortion, therefore we only included the results calculated using our model. We consider a random phase system and two cases of fixed channel phases: in one case each channel has the same phase; in the second case the channel input phases are arranged as above to reduce the FWM effects. When the channel spacing is 25GHz, the FWM-induced intensity distortion dominates no matter whether dispersion compensation is used or not. The total intensity distortion from the sum of FWM and XPM effects for the system without dispersion compensation is slightly less than that for the system with dispersion compensation. When the channel spacing is 50GHz, whether the FWM or the XPM effect dominates depends strongly on whether dispersion compensation is used or not. The system with dispersion compensation is preferred considering the total effect. Note the effectiveness of using phase arrangement to reduce the FWM-induced intensity distortion, especially for the system with 50GHz channel spacing.

In these figures, we also include results for FWM if CW pumps are used in the model instead of modulated pumps. The magnitude of the CW pumps is set to be equal to the peak power of the modulated pumps. When dispersion compensation is used, the same variance is found for both CW pumps and modulated pumps with random phase which suggests that the worst case for the modulated pumps occurs around the peak of the pump pulses. However, when dispersion compensation is not used, the FWM effect is reduced, yet the analysis using a CW pump assumption does not show this reduction. With modulated pumps, the residual dispersion broadens the generated FWM pulses, thus reduces the intensity distortion. However, with CW pumps, the FWM term is exaggerated because the residual dispersion does not change the magnitude of the generated CW FWM.

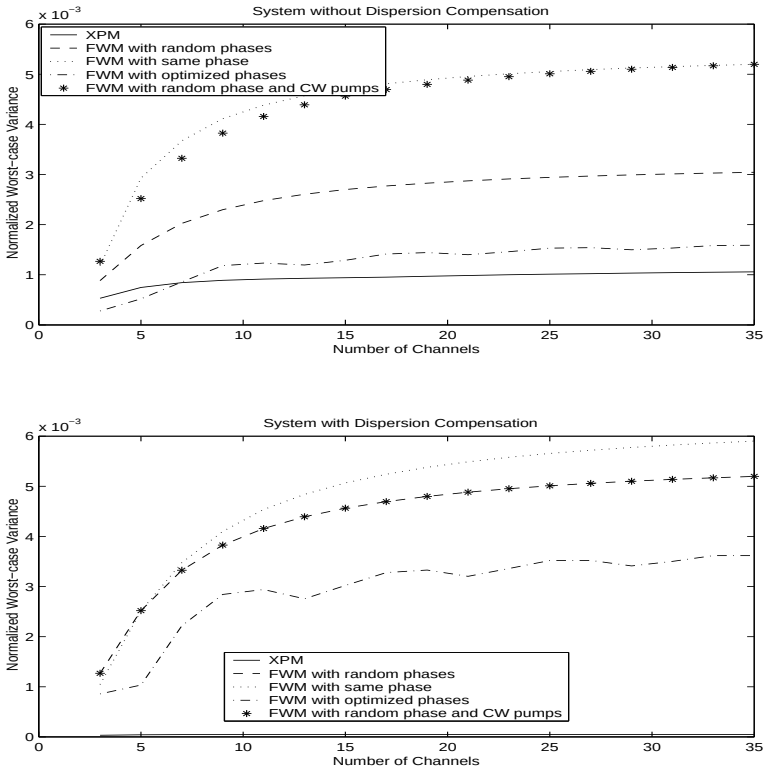


Figure 3.9: Normalized worst-case variance for synchronous WDM system with 25GHz channel spacing.

3.2.3 Variance for Time Asynchronous System

Asynchronous WDM systems are modeled by introducing a random delay uniformly distributed over $[0, T_b]$ for each pump channel instead of a fixed channel delay as in the synchronous systems. The use of a CW probe makes the analysis independent of the sampling time of the receiver. As a result, we focus on calculating the variance of the receiver observations sampled at T_b only.

The analysis of the XPM effect is easily extended to the asynchronous WDM systems. The random delay of the pump channel simply passes over to the XPM generated pulses to give $X_k(t - T_k, L)$ where T_k is uniformly distributed over $[0, T_b]$ now. By sampling at T_b and taking an average with respect to the random phases, random bits and random delays,

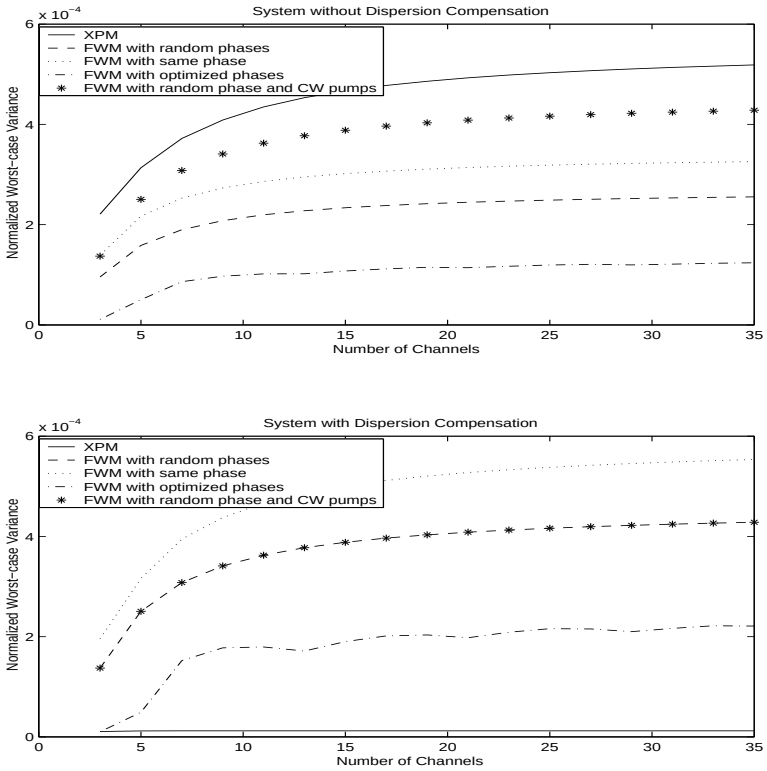


Figure 3.10: Normalized worst-case variance for synchronous WDM system with 50GHz channel spacing.

we obtain

$$\begin{aligned}\sigma_X^2 &= \sum_{k=-M,k\neq 0}^{M} \sigma_{X,k}^2 \\ &= \sum_{k=-M,k\neq 0}^{M} \sum_{l_k} \frac{1}{2} E[X_k^2(T_b - l_k T_b - T_k)] - \frac{1}{4}(E[X_k(T_b - l_k T_b - T_k)])^2,\end{aligned} \tag{3.23}$$

where the remaining expectation is over time delays.

The analysis for the FWM effect becomes quite complicated due to the interactions between channels with random delays. The mean of $\Delta P_F(t, L)$ with random delays is easily shown to be zero. For the variance, the random delays in $F_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t, L)$ of (3.19), T_{k_1}, $T_{k_1+k_3}$, and T_{k_3}, should be taken into account. Sampling $F_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t, L)$

at T_b and taking the expectation with respect to the random delays, we obtain

$$\sigma_F^2 = \sum_{\substack{k_1,k_3=-M \\ k_1,k_3\neq 0 \\ |k_1+k_3|\leq M}}^{M} \sum_{l_{k_1}} \sum_{l_{k_1+k_3}} \sum_{l_{k_3}} C_{k_1,k_3} E[|F_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t,L)|^2] \mid_{t=T_b} \tag{3.24}$$

To get $E[|F_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t,L)|^2]$, we first exchange the order of taking the Fourier transform and taking the expectation, (which is valid for energy-limited pulse as in the case of our analysis), and define $h_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t,t') = E[g_{k_1}(t - l_{k_1}T_b - T_{k_1}) g^*_{k_1+k_3}(t - l_{k_1+k_3}T_b - T_{k_1+k_3}) g_{k_3}(t - l_{k_3}T_b - T_{k_3}) g^*_{k_1}(t' - l_{k_1}T_b - T_{k_1}) g_{k_1+k_3}(t' - l_{k_1+k_3}T_b - T_{k_1+k_3}) g^*_{k_3}(t' - l_{k_3}T_b - T_{k_3})]$. $h_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t,t')$ is a deterministic function depending on the pulse shapes and can be evaluated numerically for each t and t'. Also define $H_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(\omega,\omega')$ as the two-dimensional Fourier transform of $h_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t,t')$ and define

$$\begin{aligned} \tilde{h}_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t,t') &= \frac{1}{4\pi^2} \iint \exp(-\frac{j}{2}D\omega^2) \exp(\frac{j}{2}D\omega'^2) \\ &\quad H_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(\omega,\omega') \exp(j\omega t) \exp(j\omega' t')\, d\omega\, d\omega'. \end{aligned} \tag{3.25}$$

We thus obtain

$$E[|F_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t,L)|^2] = \frac{\gamma^2 P_0}{\alpha^2 + (\beta_2 k_1 k_3 \Delta\omega^2)^2} \tilde{h}_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t,-t). \tag{3.26}$$

One needs to be careful when computing $h_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t,t')$. There are two special cases which need separate computation. One case is when $k_1 = k_3$; a different result is obtained when $k_1 = -k_3$ because $g_{k_1+k_3}(t) = \sqrt{P_0}$ is the probe channel. Note that as above, most of the $F_{k_1,k_3,l_{k_1},l_{k_1+k_3},l_{k_3}}(t,L)$ terms are very small either because the generating pulses have very little overlapping or because the generated FWM pulse is away from T_b.

Figure 3.11 and 3.12 show the variance of the XPM- and FWM-induced distortion for asynchronous WDM systems as the number of channels increases for both 25GHz and

50GHz channel spacing respectively. The same parameters as in the last sub-section are used here. Both system with and without dispersion compensation are included. Similar results are found for both systems with 25GHz and 50GHz channel spacing. When dispersion compensation is not used, the XPM effects dominate while the FWM effects dominate when dispersion compensation is used. Systems with dispersion compensation show less total variance than systems without dispersion compensation, but the difference between systems without and with dispersion compensation becomes smaller as the channel spacing is reduced. When asynchronous WDM systems are compared with synchronous WDM systems, the asynchronous WDM systems are found to have one order less variance than the corresponding worst-case synchronous WDM systems considered. This huge difference is partially due to our use of RZ pulses for the pump channels. This confirms that large system performance improvement can be achieved by correctly aligning the WDM channels with suitable delays between each other.

Simulation results from the SSF method have been included in Figure 3.12 for several small number of channels from 3 to 9 for the 50GHz channel spacing case. An optical 9th-order Butterworth filter with 25GHz bandwidth is used to filter out the probe channel in the simulation. The results for the case without dispersion compensation match our analytical results very well. The variance for the system with dispersion compensation is one order less than that without dispersion compensation. As a result, some discrepancy between our theoretical results and simulation results exists and part of this discrepancy can be tracked down to channel leakage from filtering effects as also shown in the plot. The rest of the discrepancy may come from the inaccuracy of simulation results or the simplifying assumptions we made in our analytical model.

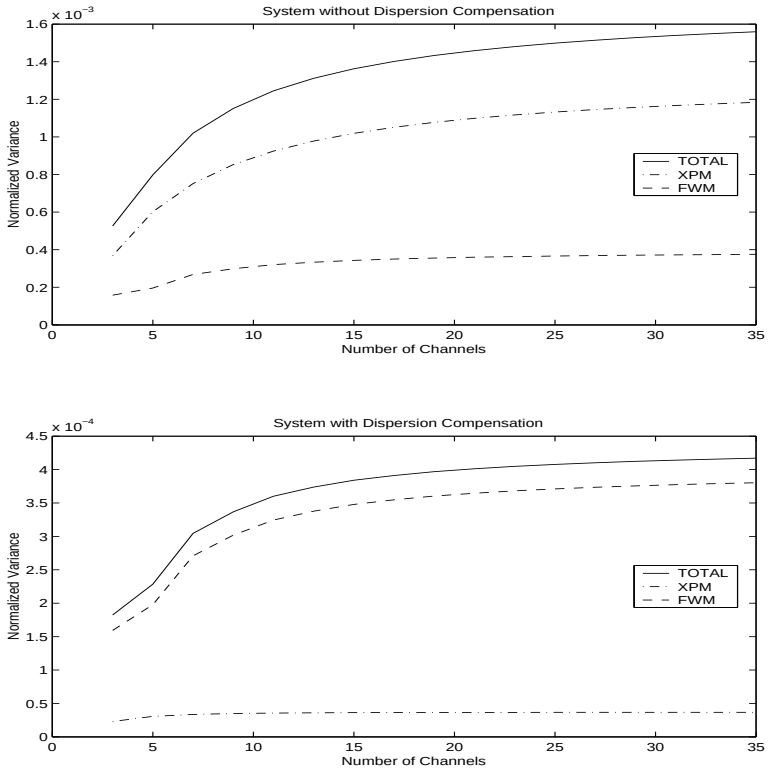


Figure 3.11: Normalized variance for asynchronous WDM system with 25GHz channel spacing.

3.3 Multispan Optically Amplified Systems

Under a small signal approximation, the total intensity fluctuation at the receiver is the accumulation of the intensity fluctuations created by each fiber span. Under this condition, (3.8) can be generalized to analyze the multispan systems as was done in [5]. The same approach can be adopted for the VSTF-based XPM and FWM analysis. However, in order to find the total intensity fluctuations at the receiver, the input of each channel to each span must be known. The SSF method can be used followed by filtering out each channel, but the simulation based on the SSF method is time-consuming and the filtered inputs are not accurate if the channels are tightly packed.

Another way to find the intensity distortion for multispan systems is to apply the VSTF

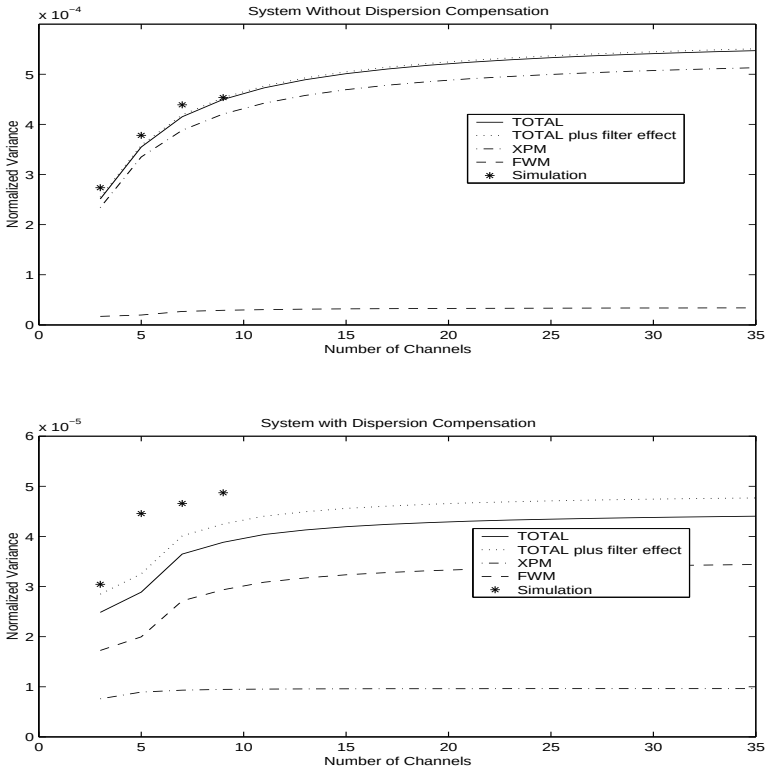


Figure 3.12: Normalized variance for asynchronous WDM system with 50GHz channel spacing.

method by approximating the total transfer function of the whole system. The system we model in this section is shown in Figure 3.13. The system consists of multiple copies of the same segment. Within each segment, there are N_A spans of fibers with both pre- and post-dispersion compensation (DC) for each fiber span and a residual dispersion compensation after N_A fiber spans. Thus, the total number of spans is $N_S N_A$ with N_S as the number of segments. Letting $D_{pre}^{(i)}$ and $D_{post}^{(i)}$ be the amounts of dispersion compensation provided by the pre- and post-DC of the i-th span respectively, then $\exp(\frac{j}{2}\omega^2 \sum_{i=1}^{N_A}(\beta_2^{(i)} L^{(i)} - D_{pre}^{(i)} - D_{post}^{(i)}))$ is the transfer function for the residual dispersion compensation where $\beta_2^{(i)}$ and $L^{(i)}$ are the dispersion and span length of the i-th fiber span.

At point A in Figure 3.13, with the total system transfer function truncated to third-order

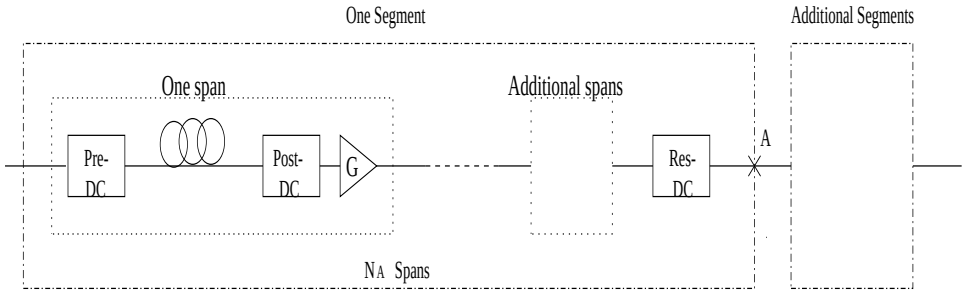


Figure 3.13: System model for multispan WDM systems.

as in [33], the distance propagated is z_A and the Volterra kernels are given by $H_1(\omega, z_A) = 1$ and

$$H_3(\omega_1, \omega_2, \omega, z_A) = -j\frac{1}{4\pi^2}\sum_{i=1}^{N_A}\frac{\gamma^{(i)}\exp(-jD^{(i-1)}(\omega_1-\omega)(\omega_1-\omega_2))}{\alpha^{(i)}+j\beta_2^{(i)}(\omega_1-\omega)(\omega_1-\omega_2)} \tag{3.27}$$

where $D^{(i)}$ is the accumulated dispersion up to the i-th fiber span, i.e., $D^{(i)} = \sum_{k=1}^{i-1}(\beta_2^{(k)} L^{(k)} - D_{pre}^{(k)} - D_{post}^{(k)}) - D_{pre}^{(i)}$, and $\alpha^{(i)}$ and $\gamma^{(i)}$ are the fiber loss and nonlinearity coefficient for the i-th fiber span, respectively. (We assume in the following that the same parameters are used for each span for simplicity; the result can be extended to general cases with distributed parameters.) Then after N_S segments, the distance the signal has propagated is $z_T = N_S z_A$, and the total transfer function is approximated as $H_1(\omega, z_T) = 1$ and $H_3(\omega_1, \omega_2, \omega, z_T) = N_S H_3(\omega_1, \omega_2, \omega, z_A)$.

To find the variance of the nonlinearity-induced intensity fluctuation for a multispan system as was done in Section 3.2 for a single span, we first notice that since $H_1(\omega, z_T) = 1$ and $H_3(\omega_1, \omega_2, \omega, z_T) = N_S H_3(\omega_1, \omega_2, \omega, z_A)$, we have $\sigma^2(z_T) = N_S^2\sigma^2(z_A)$.

Now we can concentrate on finding $\sigma^2(z_A)$ within each segment. Due to the dispersion-induced walk-off effect, we assume that the intensity fluctuation from each span is uncorrelated from each other, that is $\sigma^2(z_A) = \sum_{i=1}^{N_A}\sigma_i^2$ where σ_i^2 is the interference variance

contribution from the i-th fiber span within each segment. This assumption is valid as long as the pump channels walk-off by at least one bit period with respect to each other.

To find σ_i^2, we need the input of each pump channel to the i-th fiber span. With the approximate total transfer function up to the third-order, we have the input signal into the i-th fiber span from channel k as

$$A_{in,k}^{(i)}(\omega) = \exp(-jD^{(i)}k\Delta\omega\omega)\exp(-\frac{j}{2}D^{(i)}\omega^2)A_k(\omega,0). \tag{3.28}$$

The first term $\exp(-jD^{(i)}k\Delta\omega\omega)$ is simply a constant time delay due to the walk-off effect, and the second term $\exp(-\frac{j}{2}D^{(i)}\omega^2)$ is the pulse broadening due to the residual dispersion. Then, the XPM term generated by the k-th channel from the i-th fiber span is very similar to (3.2) and is found to be

$$\begin{aligned} A_{X,k}^{(i)}(\omega,z_A) &= -2j\gamma\exp(-\frac{j}{2}D^{(i)}\omega^2)\sqrt{P_0} \\ &\quad \frac{1}{2\pi}\int \frac{A_{in,k}^{(i)}(\omega_1)(A_{in,k}^{(i)}(\omega_1-\omega))^*}{\alpha+j\beta_2(k\Delta\omega)\omega+j\beta_2\omega\omega_1-j\beta_2\omega^2}\,d\omega_1. \end{aligned} \tag{3.29}$$

The FWM term is found to be

$$\begin{aligned} A_{F,k_1,k_3}^{(i)}(\omega,z_A) &= -j\frac{\gamma}{\alpha-jk_1k_3\beta_2\Delta\omega^2}\exp(-\frac{j}{2}D^{(i)}\omega^2) \\ &\quad \mathcal{F}\{A_{in,k_1}^{(i)}(t)(A_{in,k_1+k_3}^{(i)}(t))^*A_{in,k_3}^{(i)}(t)\} \end{aligned} \tag{3.30}$$

which is very similar to the previous analysis on FWM. Then σ_i^2 can easily be obtained following the discussion in Section 3.2.2 and 3.2.3 for both synchronous and asynchronous WDM systems for each fiber span with the broadened input pulse shape of (3.28) and a new channel delay of $T_k + D^{(i)}k\Delta\omega$.

In Figure 3.14, we show the standard deviation of the XPM- and FWM-induced distortion for a multispan WDM systems with different number of channels. Since perfect

post-DC is used for each span, we have $N_A = 1$. (This is also often used in practice.) For this system, we expect the variance to increase as N_T^2 where N_T is the number of total spans. To emphasize the linear relationship between the standard deviation of the intensity distortion and the number of spans, we plot here the normalized standard deviation instead of variance of the intensity fluctuation. It is clear that our model is reasonably accurate for a small number of spans. As the number of spans increases, discrepancy due to higher-order nonlinear terms emerges.

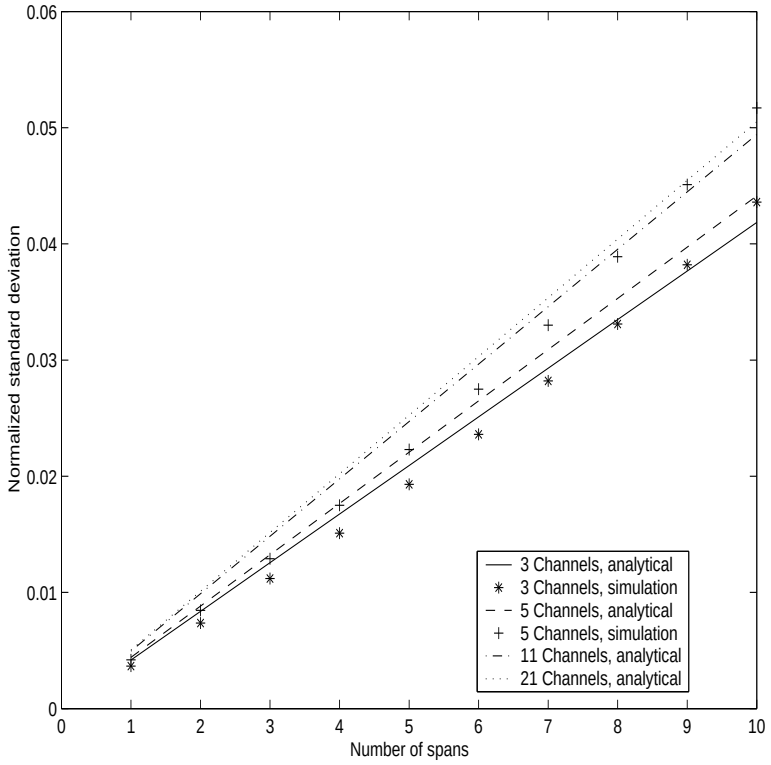


Figure 3.14: Normalized standard deviation for asynchronous multispan WDM system with 50GHz channel spacing. Perfect post-DC is used for each span.

Perfect dispersion compensation for each span has also been used to study the XPM effect in [4] for three different WDM systems. It is interesting to note that their approach, based on an undistorted pump assumption, also predicts a linear relationship between the

total XPM effect and the number of spans. This fact is obscured in Figure 3 in [4] because of the use of a semi-log scale. Our model, which is based on a different mathematical model, provides similar results to [4] for their restricted case, yet is powerful enough to model FWM and dispersion-distorted pumps as well. [2]

3.4 Chapter Summary

We have derived new analytical tools to study the XPM- and FWM-induced intensity distortion by using the VSTF method. The analysis shows good accuracy for practical system parameters. With the new results, the XPM and FWM effects can be compared fairly under identical system configurations. We also present methods to calculate the variance of the intensity fluctuation for both synchronous WDM systems and asynchronous WDM systems. This is the first study which includes the effect of random channel delays on the nonlinearity induced intensity distortion, to the authors' knowledge. This requires a FWM analysis using modulated pump channels as is done here; all previous works were done using CW pump channels. After the variances of the nonlinearity induced intensity distortion, σ_X^2 and σ_F^2, are known, the variance of the receiver samples are found to be $\sigma^2 = \sigma_X^2 + \sigma_F^2 + \sigma_N^2$ where the last term σ_N^2 is the contribution from the noise. The SNR can then be approximated as P_0/σ. The analysis here gives an easy and accurate way to predict how tight one can pack the WDM channels while maintaining a reasonable SNR based on the variance of the intensity fluctuation.

Our analysis also suggests ways to reduce the nonlinearity induced distortion. Results from Figure 3.5 show that dispersion compensation is an effective way to reduce XPM

[2] [4] focuses on the XPM-induced intensity fluctuation, thus the FWM effect is not included. In their calculation, a channel spacing of 100 GHz is used which leaves a very small FWM effect.

while FWM is robust to dispersion compensation. On the other hand, due to the fact that FWM decreases quickly as fiber dispersion increases, dispersion management can be used to reduce FWM effects. Both dispersion compensation and dispersion management have been extensively studied by both simulation and experiment, for example in [19, 20, 21, 22, 23]. Channel misalignment and optimal channel phases can also be used to reduce the FWM-induced intensity fluctuations. Other known ways to reduce the nonlinearity-induced distortion are polarization-division multiplexing [24] and unequal channel spacing [25, 26, 27].

The optimization of dispersion compensation or dispersion management depends strongly on the system parameters, thus an accurate and easy-to-use theoretical model is highly preferred over exhaustive computer simulation. The results presented in this chapter give better tools to find the optimal system parameters when designing a WDM system in the nonlinear effects limited regime. The optimization could be taken with respect to, (but is not limited to), channel input power, dispersion, dispersion compensation and pulse shape. The analytical tools are shown to be more convenient than exhaustive computer simulation.

In this chapter, the FWM-induced intensity distortion is studied with a CW probe and modulated pumps so that it can be compared to the XPM effect. Similar analysis on FWM (only) could be done with both probe and pumps modulated as in real situations; this may give a better understanding for FWM limited systems than the analysis based on a CW probe, and is the subject of future work.

Chapter 4

Multiuser Detection for WDM systems

In the previous chapter, we thoroughly studied the XPM and FWM effects on optical fiber WDM systems. Methodology on how to reduce the nonlinear effects is also given by adjusting the system parameters like dispersion, dispersion compensation and channel input phases. However, the nonlinear degradation is entirely deterministic given the data bits across the channels. Nonlinear crosstalk terms appearing in WDM, including FWM and XPM, result in a correlated statistic across wavelengths. This indicates possible better system performance through the use of a multiuser (multichannel) detector over a conventional single-user detector for optical fiber WDM systems that are nonlinearity limited.

Multi-user detection has been successfully used for processing code-division multiple access (CDMA) signals, both in radio and optical channels ([28] and references therein). To date no attempt has been made to use this technique in WDM systems. The primary method of addressing nonlinearity based degradation in fiber is to reduce their source by manipulating the channel itself, i.e., the fiber link. Several techniques successfully suppress the main crosstalk components by strategic channel placement [42, 43] and dispersion management [44, 45, 46]. These techniques always leave residual terms which, as throughput demand of these links increases, dominate the error rate.

What makes multi-user detection of WDM signals unique and what makes them mathematically interesting is that both the interference and the noise are signal dependent. This is due to the nonlinearity of the fiber and the presence of the photodetector which acts as a square-law device on the signal plus interference plus noise.

Noncoherent multiuser detection has been studied in [47, 48, 49, 50, 51], but none of these results can be applied here. Noncoherent multiuser detection was introduced in the context of differential phase-shifted-keying (DPSK) modulation for the synchronous Gaussian channel in [47] and for the asynchronous Gaussian channel in [48]. [49] focuses on the performance of noncoherent decorrelative detection for nonorthogonal multipulse modulation under multiuser interference. [50] applies decision-feedback to noncoherent decorrelative detection studied in [49] and [51] studies noncoherent multiuser detection over the Rayleigh-fading channels. Our work here considers the performance of the optimal noncoherent multiuser detector with square-law receivers.

This chapter is organized as follows. In Section 4.1 we introduce the system model studied with a square-law receiver and the Gaussian approximation for error performance evaluation. We then propose an optimum multiuser detector for this system in Section 4.2 and study its performance. The asymptotic behavior of this optimum multiuser detector is then studied in Section 4.3. We apply the multiuser detector to several different systems including a WDM system in Section 4.4. We discuss our conclusions and future work in Section 4.5.

4.1 System Description

Square-law receivers have been widely used in optical communication systems where we can write the decision statistic after the square-law receiver and an ideal sampler as

$$Y_i = |b_i A_i + \rho_i(\underline{b}) + N_i|^2, \ i = 1, \ldots, M \tag{4.1}$$

for the $i-$th user of the total M users, where $b_i \in \{0, 1\}$ is the bit sent by user i, A_i is the signal amplitude of user i at the sampling instant, $\rho_i(\underline{b})$ is the correlation among the users due to the fiber nonlinearity and filtering leakage when the vector of information bits $\underline{b}$ is sent . In this chapter, the size (dimensionality) of the observation vector $\underline{Y}$ is used exchangeably with the number of users. N_i is the additive complex Gaussian noise experienced by user i due to fiber amplifiers before the square-law receiver and assumed to dominate over the noise introduced by the square-law receiver. We also assume independent noise from different users here. Other assumptions used in (4.1) are: 1) inter-symbol interference is absent for every user; 2) the users transmit at the same rate and the system is synchronized; 3) all the signals A_i and correlations $\rho_i(\underline{b})$ are real for simplicity. For later use, we define two different cases based on the correlation between users. We call the correlation linear if

$$\rho_i(\underline{b}) = \sum_{k \neq i} \rho_{i,k} b_k A_k, \tag{4.2}$$

i.e. the correlation is nonzero whenever other users send a “one”, and nonlinear if

$$\rho_i(\underline{b}) = \sum_{k \neq i} \rho_{i,k} b_i b_k A_k, \tag{4.3}$$

i.e. the correlation is zero unless users k and i both send out pulses. Linear correlation is very common in wireless channels. Nonlinear correlation shows up in optical WDM communications where XPM introduces crosstalk between two different channels (users).

Due to the square-law receiver, the decision statistic Y_i is distributed as central or non-central χ^2 distribution of 2 degrees of freedom. In optical communication system analysis, the Gaussian distribution has been widely used to approximate the receiver statistics ([52] and references therein). However, both the central or non-central χ^2 distribution have quite different tail behavior from the Gaussian distribution, plus the impossibility of negative data.

We first study the accuracy of this approximation for the single-user case, where we have $Y = |bA + N|^2$. Suppose the noise is complex Gaussian distributed with zero mean and variance σ^2. Then Y is distributed as central or non-central χ^2 distribution of 2 degrees of freedom. The error probability of the optimal detector is derived using the optimal threshold from the two χ^2 distributions. It is easy to show that the performance of the optimal threshold detector with a χ^2 distribution depends on A and σ only through the ratio and we define the signal-to-noise ratio (SNR) before the square-law receiver as A/σ. The error probability of the χ^2 optimal detector can be approximated by deriving the error probability of an optimal detector designed for a Gaussian statistic assuming the same means and variances. It can also be shown that the approximated error probability from the Gaussian approximation depends solely on the signal-to-noise ratio (SNR) A/σ before the square-law detector. As a result, we only need to compare the error probabilities from the χ^2 and Gaussian distribution with a fixed signal level A. This must be done numerically because the optimal threshold between central and non-central χ^2 distribution can not be written in closed-form. The result is shown in Figure 4.1. It is clearly seen that the Gaussian approximation works quite well within error probability range of practical interest for the single-user case. The asymptotic behavior of the optimal detector as SNR increases with

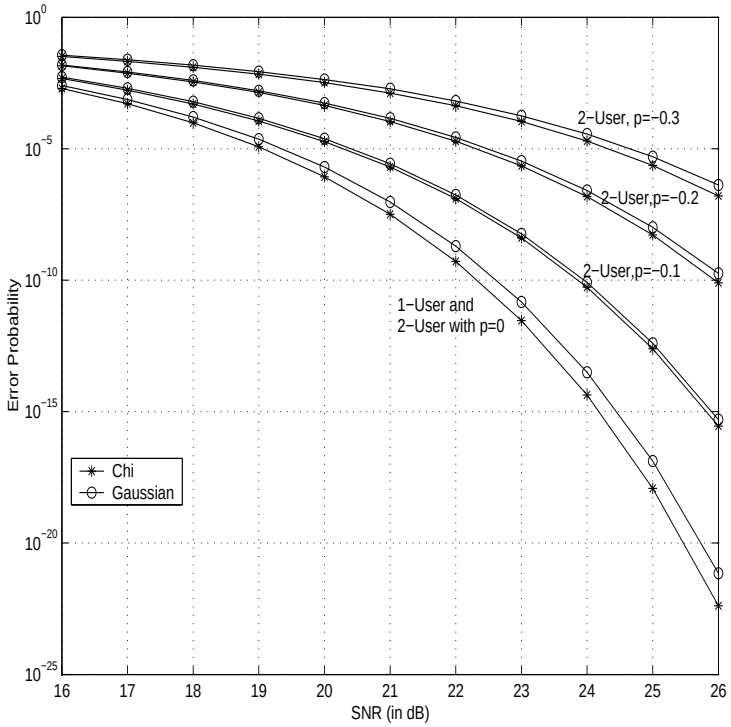


Figure 4.1: Comparison of error probability of optimal single-user and 2-user detector under χ^2 and Gaussian distributions.

the χ^2 distribution and its Gaussian approximation in the single-user case will be studied in Section 4.3 using the asymptotic efficiency. Here, Y is of central or non-central χ^2 distribution of 2 degrees of freedom. If χ^2 distributions of degrees of freedom higher than 2 are involved, the performance from the Gaussian approximation is again very close to the exact performance [53].

Unfortunately, the optimal thresholds for these two distributions are quite different as shown in Figure 4.2. This means that we can only use the Gaussian approximation to estimate the error performance of the system, not to estimate the threshold. A similar result is obtained in [53] for χ^2 distribution of higher degrees of freedom.

Our next step is to study whether this fortuitous coincidence extrapolates to the 2-user case. The optimal multiuser detector is defined as one which maximize the likelihood of

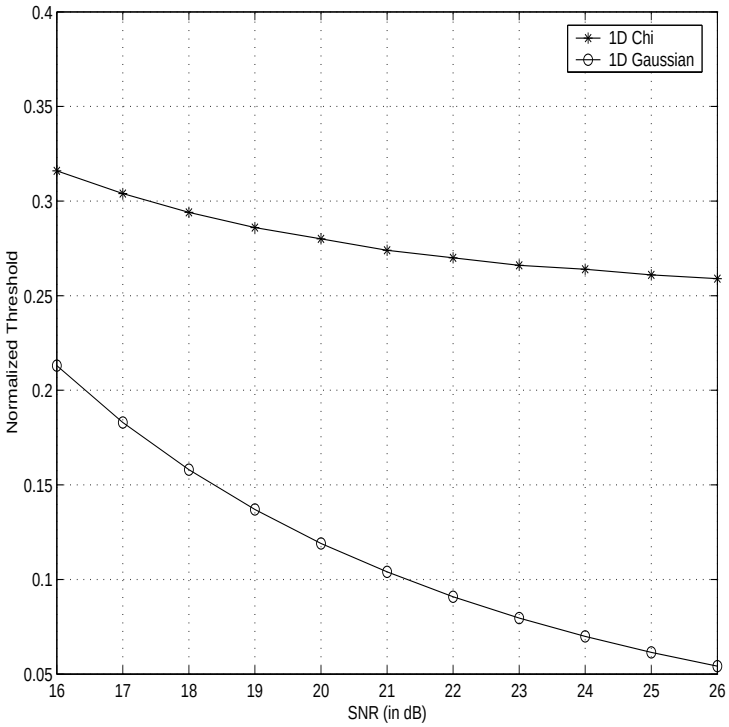


Figure 4.2: Comparison of normalized threshold of optimal single-user detector under χ^2 and Gaussian distribution.

the data of all users under the assumed statistics. In the Gaussian case the detector reduces to a quadratic form, i.e., the decision boundaries are quadratic. In the χ^2 case, the detector can not be simplified so, leading to more complex implicitly-defined decision boundaries. With a non-zero correlation between the two users, the error probability can not be written in closed-form even for the Gaussian approximation. We include a numerical integration of the error probabilities for the linear correlation case in Figure 4.1 where we observe that the Gaussian approximation again works well. However, the decision boundaries are found to be quite different, as for the 1-D case. Similar results are found for the case of nonlinear correlation.

Based on the above observations, we conclude that a Gaussian approximation can be used to approximate error performance for the square-law multiuser detector. We use it ex-

clusively in the following discussion. We also assume without validation that the Gaussian approximation works reasonably well on the error probability analysis for more than two users. However, one needs to bear in mind that the Gaussian approximation can not be used to derive decision boundaries on detection statistics.

4.2 Performance Analysis

In this section, we study the performance of the optimal square-law multiuser detector in the presence of user interference under the Gaussian approximation. The optimal square-law multiuser detector is simply a maximum-likelihood (ML) detector. However, due to the fact that the noise variance is signal-dependent after the square-law operation, the decision boundaries are no longer linear but quadratic and no closed-form expression exists for the error probability. Since numerical integration of tails of probability densities is time-consuming in higher dimensions, we prefer to bound the error probability.

The bit error probability for user i of an M-user system can be upper bounded using the union bound as

$$P_{e,i} \leq \frac{1}{2^M} \sum_{\underline{\tilde{b}}'_i, \underline{\tilde{b}}_i} P(b'_i = 1, \underline{\tilde{b}}'_i | b_i = 0, \underline{\tilde{b}}_i) + \frac{1}{2^M} \sum_{\underline{\tilde{b}}'_i, \underline{\tilde{b}}_i} P(b'_i = 0, \underline{\tilde{b}}'_i | b_i = 1, \underline{\tilde{b}}_i) \tag{4.4}$$

where $\underline{\tilde{b}}_i \in \{0,1\}^{M-1}$ is the vector of transmitted bits excluding element b_i and b'_i is the decision on bit b_i. With the Gaussian approximation, we first define $\mu_{i,\underline{b}}$ and $s^2_{i,\underline{b}}$ as the mean and variance of Y_i under hypothesis $\underline{b}$, we then have that $\mu_{i,\underline{b}} = m^2_{i,\underline{b}} + 2\sigma^2_{i,\underline{b}}$ and $s^2_{i,\underline{b}} = 2\sigma_{i,\underline{b}}\sqrt{m^2_{i,\underline{b}} + \sigma^2_{i,\underline{b}}}$ where $m_{i,\underline{b}}$ and $\sigma^2_{i,\underline{b}}$ are the mean and variance of user i's signal before the square-law receiver. Second, we define $r^2_{\underline{b}}(\underline{y}) = \sum_{k=1}^{M} (y_k - \mu_{k,\underline{b}})^2 / s^2_{k,\underline{b}}$ as the square of the normalized distance between an observation vector and the mean of the signal constellation point corresponding to $\underline{b}$. The normalization refers to the fact that

the distance measure is taken with respect to the noise covariance. Third, the normalized minimum distance (NMD) between the hypothesis pair $\underline{b}$ and $\underline{b}'$ if $\underline{b}$ is sent is defined as $D(\underline{b}'|\underline{b}) = \min_{\underline{y} \in B(\underline{b},\underline{b}')} r^2_{\underline{b}}(\underline{y})$ where $B(\underline{b},\underline{b}')$ is the decision boundary between the hypothesis pair $\underline{b}$ and $\underline{b}'$. Fourth, we define $D_i(\underline{b}) = \min_{\underline{\tilde{b}}'_i} D_i(b'_i \neq b_i, \underline{\tilde{b}}'_i|\underline{b})$ as the NMD over all of the constellation points corresponding to an error on bit i under hypothesis $\underline{b}$. With these definitions, we can further upper-bound (4.4) with

$$\begin{aligned}\sum_{\underline{\tilde{b}}'_i \in \{0,1\}^{M-1}} & P(b'_i \neq b_i, \underline{\tilde{b}}'_i|\underline{b}) \leq P(r^2_{\underline{b}}(\underline{y}) > D_i(\underline{b})) \\ &= \int_{r^2_{\underline{b}}(\underline{y}) > D_i(\underline{b})} \left(\frac{1}{\sqrt{2\pi}}\right)^M \prod_{k=1}^{M} \left[\frac{1}{s_{k,\underline{b}}} \exp\left(-\frac{(y_k - \mu_{k,\underline{b}})^2}{2s^2_{k,\underline{b}}}\right) dy_k\right]. \end{aligned} \tag{4.5}$$

This last integral can be simplified to a single integral of the form

$$\frac{M}{\sqrt{2}^M \Gamma(\frac{M}{2}+1)} \int_{\sqrt{D_i(\underline{b})}}^{\infty} r^{M-1} \exp(-r^2/2)\, dr. \tag{4.6}$$

Similar ideas based on NMD have been used to calculate error probability with multidimensional (M-D) Gaussian distributions where M-D hyperplanes have been used instead of the M-D ellipsoids used here [29, 30]. The upper-bound is expected to be tight at high SNR, thus it has been used as an approximation to the true error probability directly by several authors [29, 30].

The lower bound we use is

$$\begin{aligned} P_{e,i} \;&\geq\; \frac{1}{2^M} \sum_{\underline{\tilde{b}}_i} \max_{\underline{\tilde{b}}'_i} \left\{ P(b'_i = 1, \underline{\tilde{b}}'_i | b_i = 0, \underline{\tilde{b}}_i) \right\} \\ &+\; \frac{1}{2^M} \sum_{\underline{\tilde{b}}_i} \max_{\underline{\tilde{b}}'_i} \left\{ P(b'_i = 0, \underline{\tilde{b}}'_i | b_i = 1, \underline{\tilde{b}}_i) \right\}. \end{aligned} \tag{4.7}$$

We then further lower-bound each term using the assumption of independent channel noises,

giving

$$\begin{aligned} P(b_i', \underline{\tilde{b}}_i' | b_i, \underline{\tilde{b}}_i) &= P(\Lambda_{b_i', \underline{\tilde{b}}_i'} > \Lambda_{b_i, \underline{\tilde{b}}_i}) \\ &= P\left(\prod_{k=1}^{M} \Lambda_{b_k'}^1 > \prod_{k=1}^{M} \Lambda_{b_k}^1 | \underline{b}, \underline{b}' \right) \\ &\geq \prod_{k=1}^{M} P(\Lambda_{b_k'}^1 > \Lambda_{b_k}^1 | \underline{b}, \underline{b}'). \end{aligned} \tag{4.8}$$

where $\Lambda_{\underline{b}}$ is the likelihood function given the data vector $\underline{b}$, and $\Lambda_{\underline{b}}^1$ is the corresponding likelihood function in the 1-D case. Now, $P(\Lambda_{b_k'}^1 > \Lambda_{b_k}^1 | \underline{b}, \underline{b}')$ can be expressed as the sums of standard Q-functions. Note that this lower-bound may be quite loose when the SNR or the number of users is large.

4.3 Asymptotic Efficiency

In this section we study the asymptotic behavior of the multiuser square-law detector as the SNR increases. Throughout this section, we assume that the complex noise before the square-law receiver has the same variance for each user, and we denote it as σ^2. We define the square-law detector asymptotic multiuser-detector efficiency (AMDE), as

$$\eta_i = \sup \left\{ 0 \leq t : \lim_{\sigma \to 0} P_{e,i}(\sigma) \Big/ Q\left(\frac{\sqrt{t} A_i}{2\sigma} \right) < \infty \right\} \tag{4.9}$$

where $Q\left(\frac{A_i}{2\sigma}\right)$ is the error probability of the single-user matched filter detector in the absence of user interference, i.e., the optimal detector with a linear receiver for the same constellation (OOK). $P_{e,i}(\sigma)$ is the error probability of the square-law multiuser detector for user i. For a fair comparison, we also define the square-law *single-user* detector asymptotic efficiency (ASDE) as

$$\eta_{s,i} = \sup \left\{ 0 \leq t : \lim_{\sigma \to 0} P_{s,i}(\sigma) \Big/ Q\left(\frac{\sqrt{t} A_i}{2\sigma} \right) < \infty \right\} \tag{4.10}$$

where $P_{s,i}(\sigma)$ is the error probability of the square-law single-user detector.

4.3.1 AMDE: Upper-Bound

Since the lower- and upper-bounds on $P_{e,i}(\sigma)$ do not coincide at high SNR for the square-law multiuser detector, we focus on obtaining lower- and upper-bounds to the AMDE. We use the lower-bound on the error probability to find an upper-bound to the AMDE, and the result is the following theorem.

Theorem 4.1: Suppose that before the square-law receiver, the signal for user i has mean of $m_{i,\underline{b}}$ and $m_{i,\underline{b}'}$ for hypothesis $\underline{b}$ and $\underline{b}'$ respectively and $b_i' \neq b_i$, then

$$\eta_i \leq \min_{\{\underline{b},\underline{b}':b_i'\neq b_i\}} \sum_{k=1}^{M} t_{i,k,\underline{b},\underline{b}'} \tag{4.11}$$

with

$$t_{i,k,\underline{b},\underline{b}'} = (m_{k,\underline{b}} - m_{k,\underline{b}'})^2/A_i^2 \tag{4.12}$$

as the contribution to the AMDE from the k-th user.

To prove the above theorem, we need the following lemma.

Lemma 4.1: If $0 \leq \lim_{\sigma\to 0} \frac{X_k(\sigma)}{A/2\sigma} = \sqrt{t_k} < \infty, k = 1, ..., M$, and $t > \sum_{k=1}^{M} t_k$, then

$$\lim_{\sigma\to 0} \frac{\prod_{k=1}^{M} Q(X_k(\sigma))}{Q\left(\frac{\sqrt{t}A}{2\sigma}\right)} = \infty. \tag{4.13}$$

Proof of Lemma 4.1: For any $\epsilon > 0$, there exists σ_0 such that if $\sigma < \sigma_0$, we have $|\frac{X_k(\sigma)}{A/2\sigma} - \sqrt{t_k}| \leq \epsilon$ for any k, thus $Q((\sqrt{t_k}+\epsilon)\frac{A}{2\sigma}) \leq Q(X_k(\sigma))$. Using the following inequality from Page 98 of [28]

$$(1 - \frac{1}{x^2})\frac{\exp(-x^2/2)}{\sqrt{2\pi}x} < Q(x) < \frac{\exp(-x^2/2)}{\sqrt{2\pi}x}, x > 1 \tag{4.14}$$

we have

$$\lim_{\sigma\to 0} \frac{\prod_{k=1}^{M} Q(X_k(\sigma))}{Q(\frac{\sqrt{t}A}{2\sigma})} \geq \lim_{\sigma\to 0} \frac{\prod_{k=1}^{M} Q((\sqrt{t_k}+\epsilon)\frac{A}{2\sigma})}{Q(\frac{\sqrt{t}A}{2\sigma})}$$

$$> \lim_{\sigma \to 0} \frac{\prod_{k=1}^{M} \left(1 - \frac{1}{(\sqrt{t_k}+\epsilon)^2 A^2/4\sigma^2}\right) \exp(-\frac{(\sqrt{t_k}+\epsilon)^2}{2}\frac{A^2}{4\sigma^2})/\sqrt{2\pi}(\sqrt{t_k}+\epsilon)\frac{A}{2\sigma}}{\exp(-\frac{t}{2}\frac{A^2}{4\sigma^2})/\sqrt{2\pi}\sqrt{t}\frac{A}{2\sigma}}, \quad (4.15)$$

and the dominating term as $\sigma \to 0$ is $\exp(-(\sum_{k=1}^{M}(\sqrt{t_k}+\epsilon)^2 - t)A^2/8\sigma^2)$. Thus, if $t > \sum_{k=1}^{M}(\sqrt{t_k}+\epsilon)^2$, then

$$\lim_{\sigma \to 0} \frac{\prod_{k=1}^{M} Q(X_k(\sigma))}{Q(\frac{\sqrt{t}A}{2\sigma})} = \infty. \quad (4.16)$$

Since ϵ can be arbitrarily small, we obtain lemma 1. $\diamondsuit$

Proof of Theorem 4.1: For each factor $P(\Lambda^1_{b'_k} > \Lambda^1_{b_k}|\underline{b}, \underline{b}')$ in (4.8), the likelihood functions $\Lambda^1_{b_k}$ and $\Lambda^1_{b'_k}$ are Gaussian probability density functions. Suppose that $\Lambda^1_{b_k}(y_k) = \Lambda^1_{b'_k}(y_k)$ at $y_k = T_1$ and T_2 with $T_1 < T_2$, then T_1 and T_2 are the two thresholds. We also define $X_1 = \frac{\mu_{k,\underline{b}} - T_2}{s_{k,\underline{b}}}$, $X_2 = \frac{\mu_{k,\underline{b}} - T_1}{s_{k,\underline{b}}}$, $X_3 = \frac{T_2 - \mu_{k,\underline{b}}}{s_{k,\underline{b}}}$ and $X_4 = \frac{\mu_{k,\underline{b}} - T_1}{s_{k,\underline{b}}}$. Then we have $P(\Lambda^1_{b'_k} > \Lambda^1_{b_k}|\underline{b}, \underline{b}') = Q(X_1) - Q(X_2)$ if $m_{k,\underline{b}} > m_{k,\underline{b}'}$ or $P(\Lambda^1_{b'_k} > \Lambda^1_{b_k}|\underline{b}, \underline{b}') = Q(X_3) + Q(X_4)$ if $m_{k,\underline{b}} < m_{k,\underline{b}'}$. Through algebraic substitution, we obtain

$$\lim_{\sigma \to 0} \frac{X_1}{A_i/2\sigma} = \lim_{\sigma \to 0} \frac{X_3}{A_i/2\sigma} = (m_{k,\underline{b}} - m_{k,\underline{b}'})^2/A_i^2, \quad (4.17)$$

and

$$\lim_{\sigma \to 0} \frac{X_2}{A_i/2\sigma} = \lim_{\sigma \to 0} \frac{X_4}{A_i/2\sigma} = (m_{k,\underline{b}} + m_{k,\underline{b}'})^2/A_i^2. \quad (4.18)$$

Thus, as SNR increases, $Q(X_1)$ and $Q(X_3)$ always dominate over $Q(X_2)$ and $Q(X_4)$ respectively unless $m_{k,\underline{b}} = m_{k,\underline{b}'}$. (If $m_{k,\underline{b}} = m_{k,\underline{b}'}$, then $P(\Lambda^1_{b'_k} > \Lambda^1_{b_k}|\underline{b}, \underline{b}') = \frac{1}{2}$.) Thus, the lower-bound of $P_{e,i}(\sigma)$ is dominated by a term of the form $\prod_{k=1}^{M} Q(X_k)$ and this term has the slowest decay rate, i.e.,

$$\min_{\{\underline{b}, \underline{b}' : b'_i \neq b_i\}} \sum_{k=1}^{M} t_{i,k,\underline{b},\underline{b}'}. \quad (4.19)$$

If t is larger than (4.19), then at least one of the terms in the lower bound of $P_{e,i}(\sigma)$ is

diverging by Lemma 4.1, as a result,

$$\lim_{\sigma\to 0} P_{e,i}(\sigma) \Big/ Q\left(\frac{\sqrt{t}A}{2\sigma}\right) = \infty, \tag{4.20}$$

and we obtained an upper-bound to the AMDE as stated in Theorem 4.1. ♢

4.3.2 AMDE: Lower-Bound

To find a lower-bound for the AMDE, we use the upper-bound of the error probability. The result is the following theorem.

Theorem 4.2: Define $D_i(\underline{b},\sigma)$ as the NMD over all of the constellation points corresponding to an error on user i under hypothesis $\underline{b}$. Then, if $\lim_{\sigma\to 0} \frac{D_i(\underline{b},\sigma)}{A_i^2/4\sigma^2} \geq \kappa_{i,\underline{b}}$,

$$\eta_i \geq \min_{\underline{b}} \kappa_{i,\underline{b}}. \tag{4.21}$$

We summarize here how to prove Theorem 4.2 first and show in the following how to lower-bound $\lim_{\sigma\to 0} \frac{D_i(\underline{b},\sigma)}{A_i^2/4\sigma^2}$.

To prove Theorem 4.2, notice first from (4.4) and (4.6) that the upper-bound of the error probability consists of sums of integral of the form

$$P_{e,i}^U(\sigma|\underline{b}) = \frac{M}{\sqrt{2}^M \Gamma(\frac{M}{2}+1)} \int_{r>\sqrt{D_i(\underline{b},\sigma)}} r^{M-1} \exp(-r^2/2)\, dr. \tag{4.22}$$

As $\sigma \to 0$, we have $D_i(\underline{b},\sigma) \to \infty$, and the above integral is dominated by a function defined as

$$F(D_i(\underline{b},\sigma)) = \zeta D_i(\underline{b},\sigma)^{\frac{M-2}{2}} \exp\left(-\frac{D_i(\underline{b},\sigma)}{2}\right) \tag{4.23}$$

with some finite constant ζ. Then, it is easy to show (4.21) with the following lemma.

Lemma 4.2: If $\lim_{\sigma\to 0} \frac{D(\sigma)}{A^2/4\sigma^2} \geq \kappa$ with some finite κ, and $0 < t < \kappa$, then

$$\lim_{\sigma\to 0} \frac{F(D(\sigma))}{Q(\frac{\sqrt{t}A}{2\sigma})} < \infty. \tag{4.24}$$

The proof of the above lemma is quite straight-forward by using (4.14). First we replace $Q(\frac{\sqrt{t}A}{2\sigma})$ with its lower bound and then compare the exponential of the numerator with the exponential of the denominator.

As SNR increases, the upper-bound of $P_{e,i}$ is dominated by the one of the $P_{e,i}^{U}(\sigma|\underline{b})$ with smallest $\kappa_{i,\underline{b}}$. As a result, if $t < \min_{\underline{b}} \kappa_{i,\underline{b}}$, we know that every $P_{e,i}^{U}(\sigma|\underline{b})/Q(\frac{\sqrt{t}A_i}{2\sigma})$ is converging as $\sigma \to 0$. Thus, $P_{e,i}(\sigma)/Q(\frac{\sqrt{t}A_i}{2\sigma})$ is also converging and $\min_{\underline{b}} \kappa_{i,\underline{b}}$ is a lower-bound to the asymptotic efficiency.

We now show how to calculate a lower-bound of $\lim_{\sigma\to 0} \frac{D_i(\underline{b},\sigma)}{A_i^2/4\sigma^2}$. Since $D_i(\underline{b},\sigma)$ is the NMD over all of the constellation points corresponding to b_i to any decision surface that changes the decision on bit i under hypothesis $\underline{b}$, we have

$$D_i(\underline{b},\sigma) = \min_{\underline{\tilde{b}}_i'} D_i(b_i' \neq b_i, \underline{\tilde{b}}_i' | \underline{b}, \sigma) \tag{4.25}$$

with $D_i(b_i' \neq b_i, \underline{\tilde{b}}_i' | \underline{b}, \sigma)$ the NMD between the hypothesis pair $\underline{b}$ and $\underline{b}'$ for user i if $\underline{b}$ is sent and $b_i' \neq b_i$. Here we explicitly include σ to show that both $D_i(\underline{b},,\sigma)$ and $D_i(b_i' \neq b_i, \underline{\tilde{b}}_i' | \underline{b}, \sigma)$ are functions of σ. To find a lower-bound of $\lim_{\sigma\to 0} \frac{D_i(b_i'\neq b_i, \underline{\tilde{b}}_i' | \underline{b},\sigma)}{A_i^2/4\sigma^2}$, we study two different cases.

As above, we define $m_{k,\underline{b}}$ and $m_{k,\underline{b}'}$ as the means of the signal of the k-user ($k = 1 \ldots M$) under the two hypothesis $\underline{b}$ or $\underline{b}'$ before the square-law receiver.

Case I: All $m_{k,\underline{b}}, m_{k,\underline{b}'} \neq 0$.

For this case, with the Gaussian approximation, the decision boundary between the two hypotheses is

$$B_\sigma(\underline{y}) : -\sum_{k=1}^{M} \frac{(y_k - (m_{k,\underline{b}}^2 + 2\sigma^2))^2}{(m_{k,\underline{b}}^2 + \sigma^2)} + \sum_{k=1}^{M} \frac{(y_k - (m_{k,\underline{b}'}^2 + 2\sigma^2))^2}{(m_{k,\underline{b}'}^2 + \sigma^2)}$$

$$= 4\sigma^2 \ln \left(\frac{\prod_{k=1}^{M} (m_{k,\underline{b}}^2 + \sigma^2)}{\prod_{k=1}^{M} (m_{k,\underline{b}'}^2 + \sigma^2)} \right). \tag{4.26}$$

As $\sigma \to 0$, the decision boundary becomes

$$B_\sigma(\underline{y}) \to B_0(\underline{y}) : -\sum_{k=1}^{M} \frac{(y_k - m_{k,\underline{b}}^2)^2}{m_{k,\underline{b}}^2} + \sum_{k=1}^{M} \frac{(y_k - m_{k,\underline{b}'}^2)^2}{m_{k,\underline{b}'}^2} = 0. \tag{4.27}$$

Define C as

$$C = \frac{1}{A_i^2} \times \min_{\underline{y} \in B_0(\underline{y})} \sum_{k=1}^{M} \frac{(y_k - m_{k,\underline{b}}^2)^2}{m_{k,\underline{b}}^2} = \frac{1}{A_i^2} \times \min_{\underline{y} \in B_0(\underline{y})} \sum_{k=1}^{M} \frac{(y_k - m_{k,\underline{b}'}^2)^2}{m_{k,\underline{b}'}^2}, \tag{4.28}$$

then it can be shown that

$$\lim_{\sigma \to 0} \frac{D_i(\underline{b}'|\underline{b}, \sigma)}{A_i^2/4\sigma^2} = \lim_{\sigma \to 0} \frac{D_i(\underline{b}|\underline{b}', \sigma)}{A_i^2/4\sigma^2} = C. \tag{4.29}$$

To prove the above result, we need the following lemma.

Lemma 4.3: Suppose that $g_\sigma(\underline{y})$ and $g_0(\underline{y}) = \lim_{\sigma \to 0} g_\sigma(\underline{y})$ are uniformly continuous over sets $B_\sigma(\underline{y})$ and $B_0(\underline{y})$ respectively, and $\lim_{\sigma \to 0} B_\sigma(\underline{y}) = B_0(\underline{y})$, then

$$\min_{\underline{y} \in B_0(\underline{y})} g_0(\underline{y}) = \lim_{\sigma \to 0} \min_{\underline{y} \in B_\sigma(\underline{y})} g_\sigma(\underline{y}). \tag{4.30}$$

To prove this lemma, notice that for any point $\underline{y}$ in $B_\sigma(\underline{y})$, we can find point $\underline{y}_0$ in $B_0(\underline{y})$ with $|g_\sigma(\underline{y}) - g_0(\underline{y}_0)| < \epsilon$ for sufficiently small σ due to the uniform continuity. Thus,

$$\min_{\underline{y} \in B_\sigma(\underline{y})} g_\sigma(\underline{y}) \geq \min_{\underline{y} \in B_0(\underline{y})} g_0(\underline{y}) - \epsilon. \tag{4.31}$$

Use the same method but for any point in $B_0(\underline{y})$, we have

$$\min_{\underline{y} \in B_0(\underline{y})} g_0(\underline{y}) \geq \min_{\underline{y} \in B_\sigma(\underline{y})} g_\sigma(\underline{y}) - \epsilon. \tag{4.32}$$

Combine the results from (4.31) and (4.32), we obtain (4.30).

To apply the lemma, we first define

$$g_\sigma(\underline{y}) = \sum_{k=1}^{M} \frac{(y_k - (m_{k,\underline{b}}^2 + 2\sigma^2))^2}{(m_{k,\underline{b}}^2 + \sigma^2)}. \tag{4.33}$$

For the boundaries defined in (4.26) and (4.27), y_k can extend to ∞. Thus, $g_\sigma(\underline{y})$ defined in (4.33) and the corresponding $g_0(\underline{y})$ are not uniformly continuous over the sets defined by the boundaries. However, we know that the normalized minimum distance must occur at some point between the two mean points, that is,

$$\underline{y} \in B_\sigma^*(\underline{y}) = \{\underline{y} : \min\{m_{k,\underline{b}}^2, m_{k,\underline{b}'}^2\} + 2\sigma^2 \leq y_k \leq \max\{m_{k,\underline{b}}^2, m_{k,\underline{b}'}^2\} + 2\sigma^2\} \tag{4.34}$$

or a relaxed version of above

$$\underline{y} \in B^*(\underline{y}) = \{\underline{y} : \min\{m_{k,\underline{b}}^2, m_{k,\underline{b}'}^2\} \leq y_k \leq 2 \times \max\{m_{k,\underline{b}}^2, m_{k,\underline{b}'}^2\}\}. \tag{4.35}$$

We then find the minimization of $g_\sigma(\underline{y})$ over $B_\sigma(\underline{y}) \cap B^*(\underline{y})$. Now $g_\sigma(\underline{y})$ is uniformly continuous over $B_\sigma(\underline{y}) \cap B^*(\underline{y})$ and (4.29) follows.

Case II: Some $m_{k,\underline{b}}$ or $m_{k,\underline{b}'}$ are zero.

We first study the simplest case where $m_{1,\underline{b}} = 0$, and all other $m_{k,\underline{b}}, m_{k,\underline{b}'} \neq 0$. Similar to (4.33), define

$$g_\sigma(\underline{y}, \underline{b}) = \frac{(y_1 - 2\sigma^2)^2}{\sigma^2} + \sum_{k=2}^{M} \frac{(y_k - (m_{k,\underline{b}}^2 + 2\sigma^2))^2}{(m_{k,\underline{b}}^2 + \sigma^2)}, \tag{4.36}$$

$$g_\sigma(\underline{y}, \underline{b}') = \frac{(y_1 - (m_{1,\underline{b}'}^2 + 2\sigma^2))^2}{(m_{1,\underline{b}'}^2 + \sigma^2)} + \sum_{k=2}^{M} \frac{(y_k - (m_{k,\underline{b}'}^2 + 2\sigma^2))^2}{(m_{k,\underline{b}'}^2 + \sigma^2)}, \tag{4.37}$$

then the boundary condition can be written as

$$B_\sigma(\underline{y}) : -g_\sigma(\underline{y}, \underline{b}) + g_\sigma(\underline{y}, \underline{b}') = 4\sigma^2 \ln\left(\frac{\sigma^2 \prod_{k=2}^{M}(m_{k,\underline{b}}^2 + \sigma^2)}{(m_{1,\underline{b}'}^2 + \sigma^2)\prod_{k=2}^{M}(m_{k,\underline{b}'}^2 + \sigma^2)}\right). \tag{4.38}$$

Clearly, $D_i(\underline{b}'|\underline{b},\sigma) = \min_{\underline{y}\in B_\sigma(\underline{y})} g_\sigma(\underline{y},\underline{b})/4\sigma^2$ and $D_i(\underline{b}|\underline{b}',\sigma) = \min_{\underline{y}\in B_\sigma(\underline{y})} g_\sigma(\underline{y},\underline{b}')/4\sigma^2$ are the NMD for user i between hypothesis pair $\underline{b}$ and $\underline{b}'$ if $\underline{b}$ or $\underline{b}'$ is sent respectively.

To find $\lim_{\sigma\to 0}\frac{D_i(\underline{b}'|\underline{b},\sigma)}{A_i^2/4\sigma^2}$, first notice that $\lim_{\sigma\to 0} g_\sigma(\underline{y},\underline{b}) = \lim_{\sigma\to 0} g_\sigma(\underline{y},\underline{b}')$, thus $\lim_{\sigma\to 0}\frac{D_i(\underline{b}'|\underline{b},\sigma)}{A_i^2/4\sigma^2} = \lim_{\sigma\to 0}\frac{D_i(\underline{b}|\underline{b}',\sigma)}{A_i^2/4\sigma^2}$. Now suppose the NMD occurs at $y_1 = 2\sigma^2+\delta$, then it can be shown that

$$\lim_{\sigma\to 0}\frac{\delta^2}{\sigma^2} < \infty; \tag{4.39}$$

otherwise, we have $\lim_{\sigma\to 0}\frac{D_i(\underline{b}'|\underline{b},\sigma)}{A_i^2/4\sigma^2} = \lim_{\sigma\to 0}\frac{D_i(\underline{b}|\underline{b}',\sigma)}{A_i^2/4\sigma^2} = \infty$.

Using above result, we obtain that for any point on the boundary with $y_1 = 2\sigma^2+\delta$, we have $g_\sigma(\underline{y},\underline{b}') \geq \frac{(\delta - m_{1,\underline{b}'}^2)^2}{m_{1,\underline{b}'}^2+\sigma^2}$. Thus,

$$\lim_{\sigma\to 0}\frac{D_i(\underline{b}'|\underline{b},\sigma)}{A_i^2/4\sigma^2} = \lim_{\sigma\to 0}\frac{D_i(\underline{b}|\underline{b}',\sigma)}{A_i^2/4\sigma^2} \geq \lim_{\sigma\to 0}\frac{(\delta - m_{1,\underline{b}'}^2)^2}{A_i^2(m_{1,\underline{b}'}^2+\sigma^2)} = m_{1,\underline{b}'}^2/A_i^2. \tag{4.40}$$

If more than one $m_{k,\underline{b}}$ or $m_{k,\underline{b}'}$ are zero, define

$$\hat{m}_{k,\underline{b}} = \begin{cases} m_{k,\underline{b}}, & \text{if } m_{k,\underline{b}'} = 0, \\ 0, & \text{otherwise} \end{cases} \tag{4.41}$$

and similarly $\hat{m}_{k,\underline{b}'}$, then we obtain that

$$\lim_{\sigma\to 0}\frac{D_i(\underline{b}'|\underline{b},\sigma)}{A_i^2/4\sigma^2} = \lim_{\sigma\to 0}\frac{D_i(\underline{b}|\underline{b}',\sigma)}{A_i^2/4\sigma^2} \geq \frac{1}{A_i^2}\max\left\{\sum_{k=1}^{M}\hat{m}_{k,\underline{b}}^2, \sum_{k=1}^{M}\hat{m}_{k,\underline{b}'}^2\right\}. \tag{4.42}$$

4.3.3 ASDE: without User Interference

In the absence of user interference, it can be shown that the single-user detector has a ASDE of $\eta_{s,i} = 1$. Moreover, the square-law single-user detector with a χ^2 distribution in the absence of interference also has ASDE of 1. Thus, our result based on the Gaussian approximation is asymptotically equivalent to the result from the true distribution under this condition. We now show how to prove the above two results. We first prove that $\eta_{s,i} = 1$ under the Gaussian approximation and we need the following lemma for this purpose.

Lemma 4.4: If $0 < \lim_{\sigma\to 0} \frac{X}{A/\sigma} = \sqrt{t_0} < \infty$ with $X > 0$, then

$$\lim_{\sigma\to 0} Q(X) \Big/ Q\left(\frac{\sqrt{t}A}{\sigma}\right) \quad < \quad \infty, \text{if } t < t_0. \tag{4.43}$$

Proof: For any $\epsilon > 0$, there exists σ_0 such that if $\sigma < \sigma_0$, then $|\frac{X}{A/\sigma} - \sqrt{t_0}| \leq \epsilon$, or equivalently, $\sqrt{t_0} - \epsilon \leq \frac{X}{A/\sigma} \leq \sqrt{t_0} + \epsilon$. Thus,

$$Q\left((\sqrt{t_0} - \epsilon)\frac{A}{\sigma}\right) \geq Q(X). \tag{4.44}$$

If $\sqrt{t} < \sqrt{t_0} - \epsilon$, then

$$\lim_{\sigma\to 0} Q(X) \Big/ Q\left(\frac{\sqrt{t}A}{\sigma}\right) \leq \lim_{\sigma\to 0} Q\left((\sqrt{t_0} - \epsilon)\frac{A}{\sigma}\right) \Big/ Q\left(\frac{\sqrt{t}A}{\sigma}\right) < \infty. \tag{4.45}$$

Since ϵ can be arbitrarily small, we get the result we want. ♢

Now we are ready to show that $\eta_{s,i} = 1$ without any user interference using Lemma 4.1 with $M = 1$ from Section 4.3.1. and above Lemma 4.4. Letting the user energy be A_i and using the Gaussian approximation, we can write the distributions of the observation Y_i under hypothesis 0 and 1 as

$$f_{Y_i}(y|b_i = 0) = \mathcal{N}(2\sigma^2, 4\sigma^4) = \frac{1}{\sqrt{2\pi}2\sigma^2} \exp\left(-\frac{(y - 2\sigma^2)^2}{2(4\sigma^4)}\right), \tag{4.46}$$

$$\begin{aligned} f_{Y_i}(y|b_i = 1) &= \mathcal{N}(A_i^2 + 2\sigma^2, 4\sigma^2(A_i^2 + \sigma^2)) \\ &= \frac{1}{\sqrt{2\pi}2\sigma\sqrt{A_i^2 + \sigma^2}} \exp\left(-\frac{(y - (A_i^2 + 2\sigma^2))^2}{2(4\sigma^2(A_i^2 + \sigma^2))}\right). \end{aligned} \tag{4.47}$$

The two thresholds are found from $f_{Y_i}(y|b_i = 0) = f_{Y_i}(y|b_i = 1)$ which gives an equation of quadratic form. Denote the two thresholds as T_1 and T_2 with $T_1 < T_2$ as before, then the error probability is found to be

$$P_{s,i}(\sigma) = \frac{1}{2}P_{s,i}(\sigma|b_i = 1) + \frac{1}{2}P_{s,i}(\sigma|b_i = 0) \tag{4.48}$$

$$= \frac{1}{2}\left[Q(X_1(\sigma)) - Q(X_2(\sigma))\right] + \frac{1}{2}\left[Q(X_3(\sigma)) + Q(X_4(\sigma))\right] \tag{4.49}$$

with $X_1(\sigma)$, $X_2(\sigma)$, $X_3(\sigma)$ and $X_4(\sigma)$ defined in Section 4.3.1. It is easy to show that

$$\lim_{\sigma\to 0}\frac{X_1(\sigma)}{A_i/2\sigma}=\lim_{\sigma\to 0}\frac{X_2(\sigma)}{A_i/2\sigma}=\lim_{\sigma\to 0}\frac{X_3(\sigma)}{A_i/2\sigma}=\lim_{\sigma\to 0}\frac{X_4(\sigma)}{A_i/2\sigma}=1 \tag{4.50}$$

in this case.

Using Lemma 4.1 with $M=1$ and Lemma 4.4 together with

$$\frac{1}{2}\left[Q(X_3(\sigma))+Q(X_4(\sigma))\right]\le P_{s,i}(\sigma)\le\frac{1}{2}\left[Q(X_1(\sigma))+Q(X_3(\sigma))+Q(X_4(\sigma))\right], \tag{4.51}$$

it is then easy to show that $\eta_{s,i}=1$ when user interference is absent.

We then prove that $\eta_{s,i}=1$ too with χ^2-distribution in the absence of user interference. In this case, we have

$$f_{Y_i}(y|b_i=0)=\frac{1}{2\sigma^2}\exp\left(-\frac{y}{2\sigma^2}\right), y\ge 0 \tag{4.52}$$

$$f_{Y_i}(y|b_i=1)=\frac{1}{2\sigma^2}\exp\left(-\frac{y+A_i^2}{2\sigma^2}\right)I_0\left(\frac{\sqrt{y}A_i}{\sigma^2}\right), y\ge 0 \tag{4.53}$$

where $I_0(x)$ is the modified Bessel function of 0-th order. At high SNR, using $I_0(x)\approx\exp(x)/\sqrt{2\pi x}$ and following the argument in [49], we neglect $\sqrt{2\pi x}$ in the denominator as it is much smaller than the exponential part at high SNR. As a result, the threshold of the optimal detector is found to converge to $A_i^2/4$ as SNR increases. This is confirmed in Figure 4.2.

To find the ASDE, we find the asymptotic efficiency of a fixed-threshold detector whose threshold is fixed at $A_i^2/4$ as SNR increases. Since the optimal detector's threshold converges to $A_i^2/4$, the two detectors have the same asymptotic efficiency.

For this fixed-threshold detector, we have

$$P_{s,i}(\sigma)=\frac{1}{2}\exp\left(-\frac{A_i^2}{8\sigma^2}\right)+\frac{1}{2}\left[1-Q\left(\frac{A_i}{\sigma},\frac{A_i}{2\sigma}\right)\right] \tag{4.54}$$

where $Q(a,b)$ is the Marcus Q-function defined in [28] and $1-Q\left(\frac{A_i}{\sigma},\frac{A_i}{2\sigma}\right)=Q\left(\frac{A_i}{2\sigma},\frac{A_i}{\sigma}\right)-\exp\left(-\frac{5A_i^2}{8\sigma^2}\right)I_0\left(\frac{A_i^2}{2\sigma^2}\right)$ from (A-3-1) in [54]. It is easy to show that

$$\lim_{\sigma\to 0}\frac{1-Q(A_i/\sigma,A_i/2\sigma)}{Q(A_i/2\sigma)}=\frac{1}{\sqrt{2}} \tag{4.55}$$

using (3.45) and (3.46) in [28], while

$$\lim_{\sigma\to 0}\frac{\exp(-A_i^2/8\sigma^2)}{Q(A_i/2\sigma)}=\infty \tag{4.56}$$

by using (4.14) as done in the proof of Lemma 4.2. Thus, the first term of (4.54) dominates over the second term as $\sigma\to 0$ and $\eta_{s,i}\le 1$ because $\lim_{\sigma\to 0}P_{s,i}(\sigma)/Q(A_i/2\sigma)=\infty$. On the other hand, we can show that $\eta_{s,i}\ge 1$ by applying Lemma 4.2 with $F(y)=\exp(-y/2)$, $y=A_i^2/4\sigma^2$ and $M=1$. Then, if $0<t<1$,

$$\lim_{\sigma\to 0}\frac{\exp(-A_i^2/8\sigma^2)}{Q(\sqrt{t}A_i/2\sigma)}<\infty \tag{4.57}$$

and $\eta_{s,i}\ge 1$ follows. Thus, $\eta_{s,i}=1$ for the optimal single-user detector in the absence of user interference under a χ^2-distribution.

4.3.4 ASDE: with User Interference

For a fair comparison, we need to find the asymptotic efficiency of the square-law single-user detector under user interference and the result is

$$\eta_{s,i}=\left(\min_{\underline{\tilde{b}}_i',\underline{\tilde{b}}_i}\left|m_{i,b_i=0,\underline{\tilde{b}}_i'}-m_{i,b_i=1,\underline{\tilde{b}}}\right|/A_i\right)^2. \tag{4.58}$$

Without user interference, we have $m_{i,\underline{b}}=0$ or A_i and $\eta_{s,i}=1$.

Proof: In this case, the likelihood function is

$$f_{Y_i}(y|b_i)=\frac{1}{2^{M-1}}\sum_{\underline{\tilde{b}}_i}\mathcal{N}(m_{i,b_i,\underline{\tilde{b}}_i}^2+2\sigma^2,4\sigma^2(m_{i,b_i,\underline{\tilde{b}}_i}^2+\sigma^2)) \tag{4.59}$$

for $b_i = 0$ or 1 where $\mathcal{N}(\mu, s^2)$ is the Gaussian probability density function of mean μ and variance s^2.

Let $S_0 = \{m_{i,b_i=0,\underline{\tilde{b}}_i}\}$ and $S_1 = \{m_{i,b_i=1,\underline{\tilde{b}}_i}\}$ be the sets composed of means for user i under each hypothesis $\underline{b}$ before the square-law receiver if $b_i = 0$ or 1 is sent respectively.

(i). If $S_0 \bigcap S_1 \neq \emptyset$, that is, some $m_{i,b_i=0,\underline{\tilde{b}}_i}$ is equal to some $m_{i,b_i=1,\underline{\tilde{b}}_i}$, then the single-user detector has zero asymptotic efficiency because the corresponding hypotheses are not distinguishable no matter how large the SNR is. For this case, (4.58) also gives $\eta_{s,i} = 0$.

(ii). If S_0 and S_1 are disjoint, then we have nonzero asymptotic efficiency for the single-user detector. Let $S = S_0 \bigcup S_1$ and rearrange the components in S in the order from small to large as $m_1, m_2, \ldots, m_{2^M}$. We then discard all the components m_k in S whose adjacent neighbors (m_{k-1}, m_{k+1}) are from the same sub-sets (S_0 or S_1) as m_k and denote the remaining components as $\hat{S} = \{\hat{m}_1, \hat{m}_2, \ldots, \hat{m}_{\bar{M}}\}$ from small to large with $\bar{M}$ to be the number of remaining components. As $\sigma \to 0$, it is easy to show that the thresholds between $\hat{m}_k^2$ and $\hat{m}_{k+1}^2$ depend solely on term $\mathcal{N}(\hat{m}_k^2 + 2\sigma^2, 4\sigma^2(\hat{m}_k^2 + \sigma^2))$ and $\mathcal{N}(\hat{m}_{k+1}^2 + 2\sigma^2, 4\sigma^2(\hat{m}_{k+1}^2 + \sigma^2))$ in (4.59) due to the exponential decay of the Gaussian distribution. As a result, the thresholds can be found from a quadratic equation, that is, $\mathcal{N}(\hat{m}_k^2 + 2\sigma^2, 4\sigma^2(\hat{m}_k^2 + \sigma^2)) = \mathcal{N}(\hat{m}_{k+1}^2 + 2\sigma^2, 4\sigma^2(\hat{m}_{k+1}^2 + \sigma^2))$.

We give an example of 3 users in Figure 4.3 to clarify the above discussion. Suppose there are eight hypothesis in total and they are shown in Figure 4.3(a) as black dots with $S_0 = \{m_{0,000}, m_{0,001}, m_{0,010}, m_{0,011}\}$ and $S_1 = \{m_{1,100}, m_{1,101}, m_{1,110}, m_{1,111}\}$. Note that we include the case that $m_{1,111} < m_{0,011}$ which may occur for sufficiently large correlation. In this case, multiple thresholds may exist for the optimal single-user detector and the crosses in Figure 4.3(a) represent the possible thresholds for this signal constellation. For

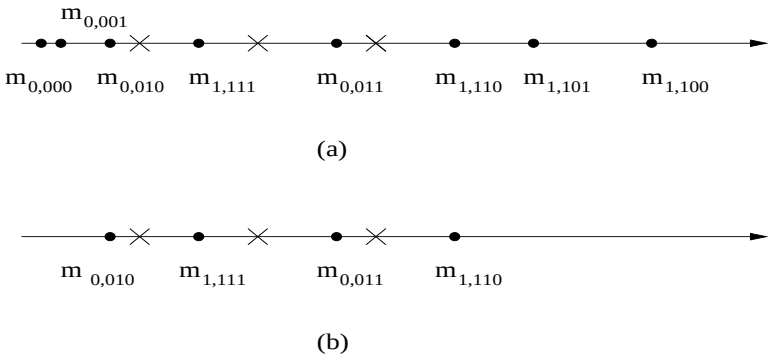


Figure 4.3: An example on how to calculate the asymptotic efficiency of single-user detector. (a). Signal constellation points in S_0 and S_1; (b). Signal constellation points in $\hat{S}$.

this example, $\hat{S} = \{m_{0,010}, m_{1,111}, m_{0,011}, m_{1,110}\}$ are shown in Figure 4.3(b) because all other points have neighbors from the same subset as themselves. The thresholds as $\sigma \to 0$ are also shown in Figure 4.3(b) by crosses.

With the thresholds determined, $P_{s,i}(\sigma)$ can be written as sums of Q-functions as is done in the proof of Theorem 4.1 in Section 4.3.1. As $\sigma \to 0$, the dominating term is the one with the slowest decay rate and the asymptotic efficiency of the optimal single-user detector is given by (4.58) for this case also. $\diamondsuit$

4.4 Applications

In this section, we first study the square-law multiuser detector for a simple case with 2 users of equal user energy. Then the square-law multiuser detector is applied to more practical optical WDM systems.

4.4.1 2-User Case

Using the results in Section 4.2 on the error probability bounds under Gaussain approximation, we compare the square-law multiuser detector with the square-law single-user

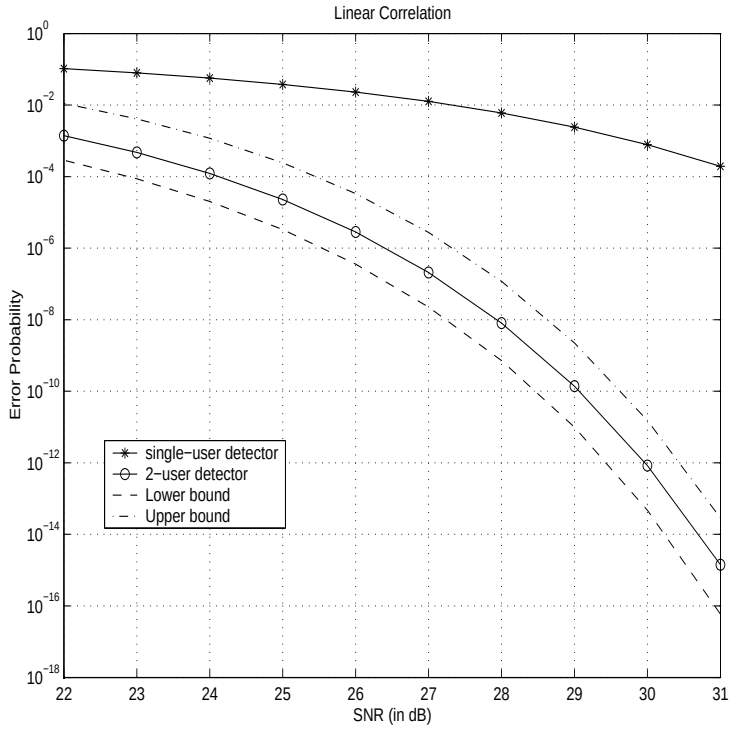


Figure 4.4: Error performance comparison for single-user and multiuser detector for 2-user case with linear correlation. User energies are equal and $\rho = -0.4$.

detector under interference with both linear and nonlinear correlations in Figure 4.4 and 4.5, respectively. The correlation is set to be symmetric and equal to -0.4. The users have equal energies. The advantage of the multiuser detector is clearly seen for both linear and nonlinear correlation cases. We also confirm that the nonlinear correlated system suffer less from interference than the linear correlated case. Note that a correlation of -0.4 is used which gives a signal-to-interference ratio (SIR) of 8 dB for linear correlation and 14 dB for nonlinear correlation. Consequently, the single-user detector has a very poor performance even at the asymptotically high SNR we used in the plot because here the detector error performance is dominated by the interference.

From Figures 4.4 and 4.5, we see that the lower- and upper-bounds do not coincide even at the high SNR that we use. Moreover, we can not prove whether the lower- and

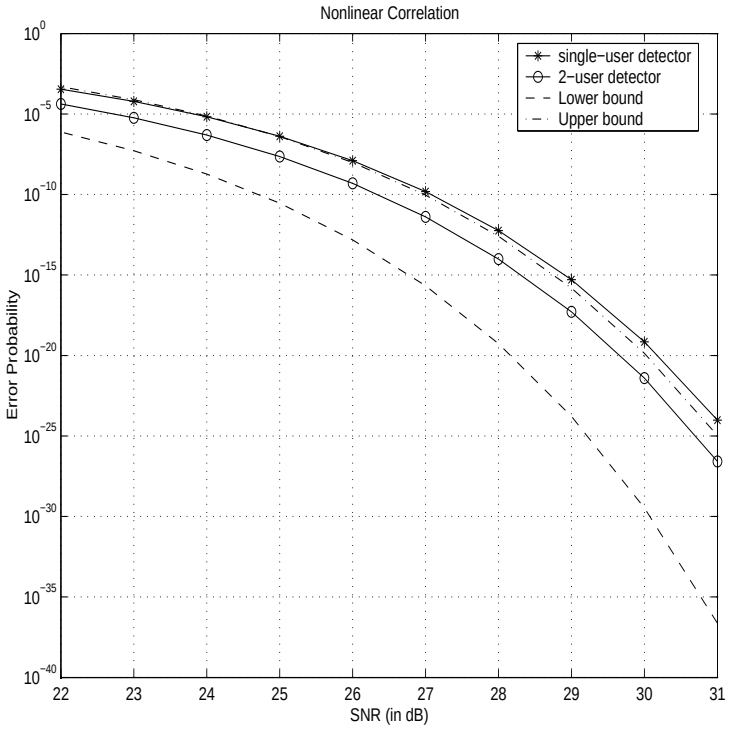


Figure 4.5: Error performance comparison for single-user and multiuser detector for 2-user case with nonlinear correlation. User energies are equal and $\rho = -0.4$.

upper-bounds are exponentially equivalent or not. To give an idea of which bound is closer to the true performance, we include the error probability from numerical integration for the 2-D case in Figures 4.4 and 4.5. The lower-bound is closer to the true error probability for low SNR and the upper-bound is closer to the true error probability for high SNR, as expected.

Another significant difference between the multiuser detector with a linear receiver and a square-law receiver is that with the square-law receiver a correlation with positive value always helps to reduce the error probability while a correlation with negative value always worsens the error performance. However, positive and negative correlations are equivalent for the linear receiver due to symmetry.

Following the discussion in the previous section on the asymptotic efficiency, the upper-

and lower-bounds on the AMDE of either user for the 2-user multiuser detector is found to be

$$\eta \leq \begin{cases} \min\{1+\rho^2, (1+2\rho)^2+\rho^2, 2(1+\rho)^2\} & \text{if } \rho < 0, \\ \min\{1+\rho^2, 2(1-\rho)^2\} & \text{if } \rho > 0 \end{cases} \tag{4.60}$$

for linear correlation and

$$\min\{1, (1+\rho)^2\} \leq \eta \leq \min\{1, (1+\rho)^2+\rho^2, 2(1+\rho)^2\} \tag{4.61}$$

for nonlinear correlation. Numerical computation is needed for the lower bound in the linear correlation case. Meanwhile, the ASDE is found to be:

$$\eta_s = \min\{1, (1+\rho)^2\} \tag{4.62}$$

for nonlinear correlation and

$$\eta_s = \begin{cases} \min\{(1+\rho)^2, (1+2\rho)^2\} & \text{if } \rho < 0, \\ (1-\rho)^2 & \text{if } \rho > 0 \end{cases} \tag{4.63}$$

for linear correlation.

Figures 4.6 and 4.7 compare the upper- and lower-bounds to the AMDE together with the ASDE for linear and nonlinear correlations respectively. For systems with linear correlation, the lower bound to the AMDE is larger than the ASDE. An interesting observation is made around $\rho = -0.5$. When $\rho = -0.5$, the square-law single-user detector has zero asymptotic efficiency because the hypotheses (0,1) and (1,1) both have mean of $A/2$ for user 1 for equal user energy which makes the detection error persist no matter how large the SNR is. This is an example of case (i) in the proof of (4.58). However, for the square-law 2-user detector, correct detection can be made via the observation statistic from user 2. This also explains why the lower- and upper-bound coincide at $\rho = -0.5$.

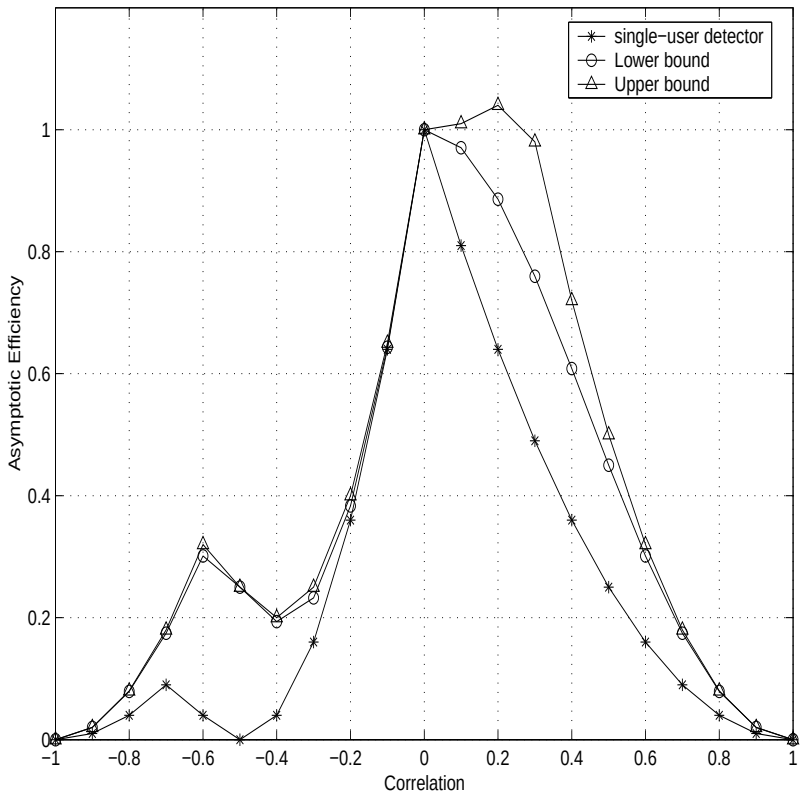


Figure 4.6: Asymptotic efficiency of square-law multiuser and single-user detector under linear interference for 2-user case.

Note that for nonlinear correlation, the lower bound of the AMDE is the same as the ASDE. As a result, we only know that the actual AMDE lies somewhere between the upper-bound to the AMDE and the ASDE, and it is likely to be very close to the ASDE because the actual error probability of the square-law multiuser detector is closer to its upper-bound while SNR increases.

We also include the asymptotic efficiency of the multiuser and single-user detector with a linear receiver in Fig. 4.8 from equations (4.65) and (3.122) of [28]. Note that the detector for a square-law receiver behaves quite differently from the detector with a linear receiver under linear correlation. The detectors with a linear receiver have symmetric behavior with respect to ρ.

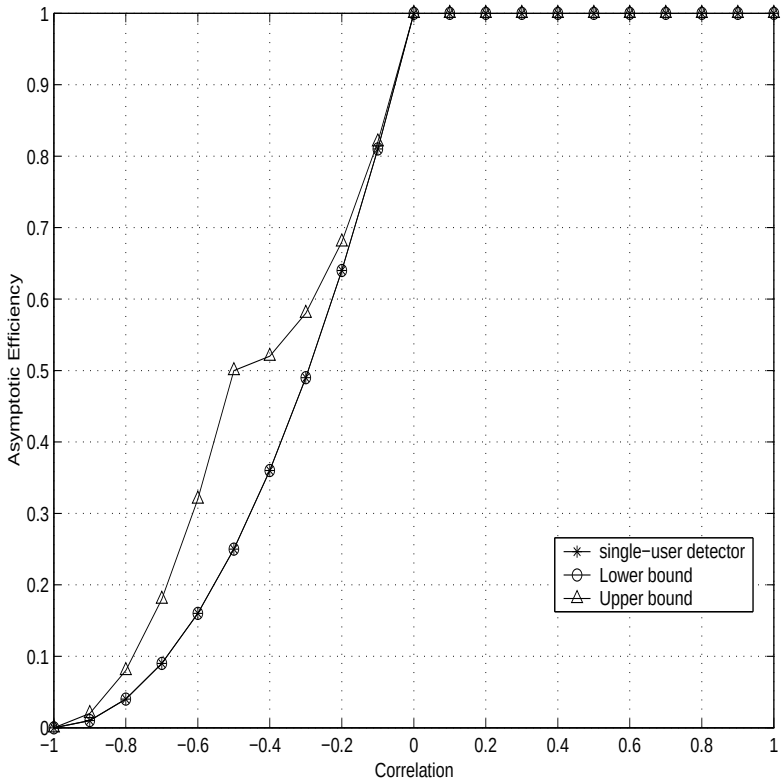


Figure 4.7: Asymptotic efficiency of square-law multiuser and single-user detector under nonlinear interference for 2-user case.

4.4.2 Optical WDM Systems

In this section, we apply the square-law multiuser detector to optical WDM systems. A simple multiple span fiber system is illustrated in Figure 4.9. M channels transmit on-off modulated pulse streams over M frequencies multiplexed (MUX) onto a single fiber. The pulse propagation through the fiber obeys the nonlinear Schroedinger (NLS) equation. Following each span of fiber is a dispersion compensator (DC) which removes all phase shift due to material dispersion, and a fiber amplifier (G) to compensate for all power loss. The optical amplifiers also add amplified spontaneous emission (ASE) noise which can be modeled as wide-band Gaussian noise. The resulting signal plus noise is passed through a WDM demultiplexer (DEMUX), which is effectively an optical filterbank, to

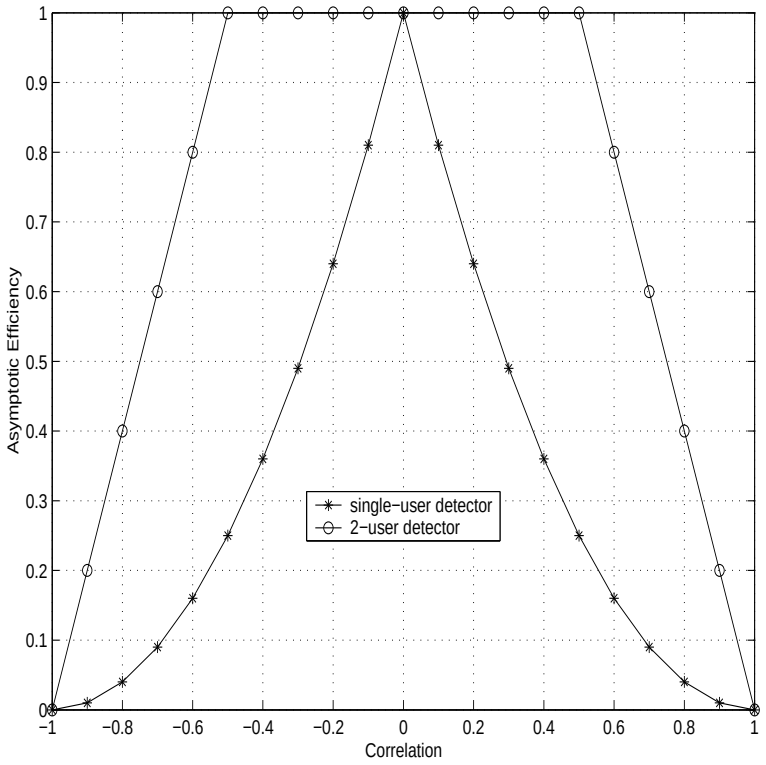


Figure 4.8: Asymptotic efficiency of multiuser and single-user detector with conventional linear receiver under linear interference and equal user energy for 2-user case.

separate the signals. In each branch the signal is photodetected and sampled at the bit rate T. Then, multiuser detectors are used to decide the transmitted bits of multiple channels simultaneously instead of using multiple single-user detectors as in Figure 1.1.

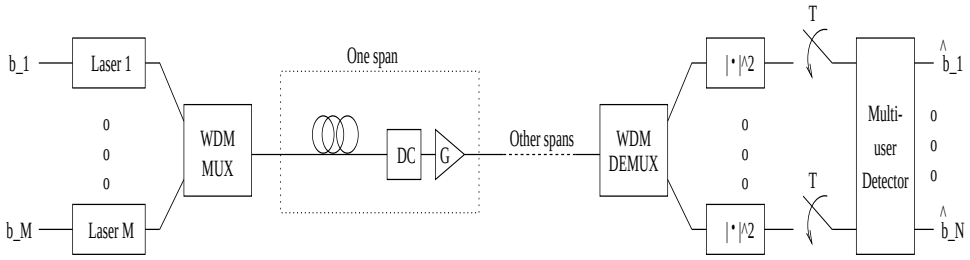


Figure 4.9: WDM fiber communication system.

Modern fiber communication systems operate using tens or hundreds of WDM channels. Given that optimal multiuser detectors have complexity that is NP complete in the

number of users, we restrict our detection scheme to N out of a total of M channels. Other channels are modeled as providing additional correlated noise. If four-wave mixing (FWM) is considered as well as cross-phase modulation (XPM), the correlation between users for channel j must be written as

$$\rho_i(\underline{b}) = \sum_{k,\ell,m=1}^{M} \rho_{k,\ell,m,i} b_k b_\ell b_m A_k A_\ell^* A_m \tag{4.64}$$

because of three-channel interaction. The channel leakage due to imperfect filtering introduces linear correlation while XPM results in nonlinear correlation between channels. FWM can introduce both linear and nonlinear correlations. The optimal detector is easily modified to this type of correlation.

Figure 4.10 shows upper and lower bounds for the error probability for $M = 3$ and 5 users using a Gaussian approximation. The mean and variance are estimated from a WDM channel simulation, then the upper- and lower-bounds are applied to study the error performance. The figure shows that for $M = 3$ the advantage of using multiuser detection over conventional detection ($M = 3$, $N = 1$) is on the order of 4dB at an error rate of 10^{-10}. For $M = 5$, the result is even more dramatic since the eye-opening is entirely closed if multiuser detection is not employed. In the simulation, the WDM signals are transmitted through 10 spans of 80km non-zero dispersion-shifted fibers (NZ-DSF) with dispersion compensation used after each fiber span. Each channel is modulated independently by return-to-zero (RZ) pulses at 10Gb/s. The channel spacing is only 25GHz. Multiuser detection is the only known way to achieve WDM transmissions with such long distances and small channel spacing.

Note that the single-user detector used in this case is a simple single threshold detector, which is what is done in practice. The threshold and the error probability are both

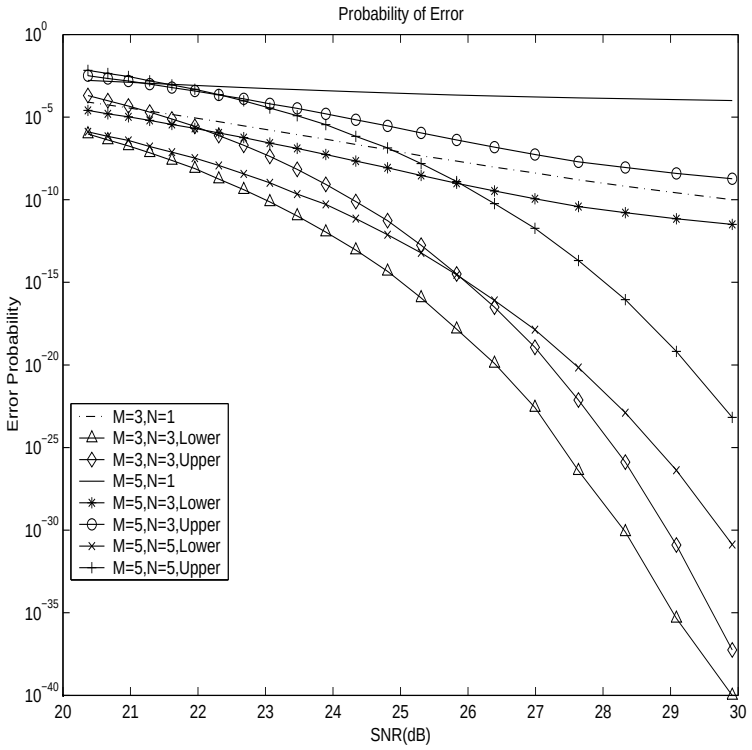


Figure 4.10: Probability of error bounds for total M users, N users modeled in the multiuser detector.

calculated using a Gaussian assumption.

4.5 Chapter Summary

In this chapter, the multiuser detector for a square-law receiver has been studied based on the upper- and lower-bounds to the error probability and asymptotic efficiency using the Gaussian approximation. The square-law multiuser detector is then applied to optical communication systems where the nonlinearity introduces correlation between channels (users). The advantage of the multiuser detector over the single-user detector is clearly shown for optical communication systems where nonlinear effects dominate.

A fiber-optic multiuser system with dominant ASE noise can be modeled to have χ^2-distributed observations because of the square-law detector. We approximate the χ^2 dis-

tribution with a Gaussian distribution which gives a close error performance estimate, but gives thresholds or decision boundaries for 1-D and 2-D cases which are vastly different from optimal under χ^2 statistics. We study the performance using a Gaussian approximation by deriving lower- and upper-bounds to the error probability. We find that, unlike the linear receiver for a symmetric Gaussian channel, the lower- and upper-bounds do not coincide at high SNR for the square-law multiuser detector.

The square-law multiuser detector is studied thoroughly for the 2-user case and compared with the conventional square-law single-user detector for both linear and nonlinear correlation. Great advantage from the multiuser detector is observed when the correlation is large enough. The asymptotic behavior of both the square-law multiuser and single-user detectors is studied based on the asymptotic efficiency compared with the conventional single-user detector for the linear receiver in the absence of interference. The error performances are quite different for the square-law multiuser detectors with linear and nonlinear correlation.

Significant differences between the detectors based on a square-law receiver and a conventional linear receiver are found for the linear correlation case. For the detector with a linear receiver, positive and negative correlations are equivalent due to symmetry; however, for the square-law receiver, positive correlation always helps to reduce the error probability while negative correlation increases the error probability of the multiuser detector.

Under a Gaussian approximation, the ML detector can be simplified to a quadratic detector which simplifies the detector design and saves computation. This is very important for optical communication systems operating at 10Gb/s or higher where a simple detector with little computation is required. The optimal χ^2-statistic detector does not have

such a simple form. Since the Gaussian approximation cannot be used to derive the form of a detector for the χ^2-statistic, deriving implementable suboptimal detectors is still an open question. A suboptimal quadratic detector with coefficients optimized to the actual statistics is feasible, but we do not know how it would perform under χ^2-distributed observations. The coefficients should be quite different from those calculated with the Gaussian approximation. Another possibility is finding an optimal linear detector. This problem is nontrivial due to the presence of signal-dependent noise.

Throughout this chapter, synchronous systems are assumed and studied. However, WDM networks with add/drop operations are examples of asynchronous systems. How to apply the square-law multiuser detectors to an asynchronous system and how well the multiuser detectors perform are still open questions.

The above analysis of the square-law detector based on the Gaussian approximation can be directly applied to detectors with other types of signal-dependent noise, such as optical systems using avalanche photodiodes (APD).

Chapter 5

Noise Nonlinear Amplification by Signal

In the previous chapters, we have focused on the system performance degradation due to the signal waveform degradation and ICI in WDM systems. In this chapter, we focus on the ASE noise introduced by the optical amplifiers, the other performance-limiting source in a long-haul optical fiber transmission system. The noise characteristics must be known in order to compute the bit error rate (BER) of the system. For most of the works on BER computation, the ASE noise is assumed to propagate through the fiber without the effect from fiber nonlinearity-induced interaction with the signal. In this case, the noise after transmission is still white Gaussian noise which simplifies the BER analysis significantly ([55] and references herein). However, the nonlinear interaction between the signal and the noise may significantly change the noise process. In this chapter, we present a method based on perturbative approach to analyze the nonlinear interaction between the noise and the signal when they co-propagate through the fiber.

The noise nonlinear amplification by the signal, or parametric gain due to the fiber nonlinearity has been found to change the noise significantly. First, the noise is no longer stationary or white after its interaction with the time-varying signal; second, there is an energy transfer between the in-phase and out-of-phase components of the noise. The noise

nonlinear amplification has been studied independently in [56] and [57] with the assumption of a continuous-wave (CW) pump signal; the noise process remains stationary at the output under this assumption. The result in [57] is further used to derive the BER performance of optical fiber communication systems where the transmitted signal pulse is long enough to be approximated by a CW signal [58]. The CW pump signal approximation limits the previous analysis to system of low bit rate. However, as stated in [58], the extension of CW approximation to arbitrarily modulated signal is a difficult problem. We present in this chapter the first result on the analysis of noise nonlinear amplification and coloring by an arbitrarily modulated signal based on perturbation theory.

The signal is assumed to be pseudo-random with finite period and arbitrary modulation in this chapter. Under this assumption, the output noise process is non-stationary and we characterize it by the use of correlation functions between its frequency components. We first find the nonlinear terms from the noise nonlinear amplification using a perturbative approach in the frequency domain. We then compute the noise correlation functions which are kept up to the second-order in fiber nonlinearity for computational simplicity. We check our results against an estimated noise correlation function using the SSF method with multiple realization of the amplifier noise. The valid range of our approximation is also discussed in this chapter. Our results give a new analytical method to compute the system performance with a more accurate noise model for ASE dominated optical fiber communication systems.

The chapter is organized as follows. Section 5.1 is a brief discussion of the system model we study in this chapter. We then introduce the first- and second-order nonlinear effects in Section 5.2. Section 5.3 is dedicated to the study of the noise correlation matrix

in the frequency domain and is the main result of this chapter. Section 5.4 extends our result to multi-span systems and it also includes a discussion on the computation complexity of our results. As a special case, we apply our method to the cases with CW signals and compare our results with the previous results in Section 5.5 We validate our results by numerical simulation and study the valid range of our method in Section 5.6. We give a simple application of our results to find the probability distribution function of the detector statistics in Section 5.7. Section 5.8 concludes the chapter.

5.1 System Model

The system that we study in this chapter is shown in Fig. 5.1. In this section, we focus on the noise nonlinear amplification in a single fiber span and we extend the results into multi-span systems later in Section V. Within each span, bi-end dispersion compensation (DC) [38] with both pre- and post-dispersion compensation is used and we model the pre- and post-dispersion compensators with transfer functions $\exp(jD_{pre}\omega^2/2)$ and $\exp(jD_{post}\omega^2/2)$ respectively, where ω is the baseband angular frequency. D_{pre} and D_{post} are the amount of dispersion compensation from the pre- and post-dispersion compensators respectively. In this study, we assume that $D_{pre}+D_{post}=\beta_2 L$ where β_2 is the second-order dispersion of the fiber and L is the span length, that is, the fiber dispersion is perfectly compensated by the pre- and post-dispersion compensation within each span as commonly used in practice. A lumped optical amplifier provides gain $G=\exp(\alpha L)$ to compensate for the fiber transmission loss. We have neglected the possible nonlinear effects in the dispersion compensation components in this simplified system model.

The input to the single-span system is the sum of a modulated signal and the ASE noise.

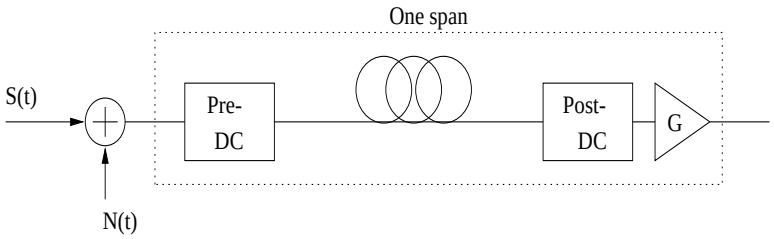


Figure 5.1: Simplified system model for single-span system. DC is the dispersion compensator.

The ASE noise is modeled as complex white Gaussian noise with power spectral density $\Psi_0 = (G-1) \times NF \times hv$, where h is Planck's constant, v is the optical carrier frequency and NF is the optical amplifier noise figure [52]. In this chapter, we characterize the noise process in the frequency domain and we write the observed noise process in the frequency domain as $N(\omega) = U(\omega) + jV(\omega)$, i.e., $U(\omega)$ and $V(\omega)$ are the real and imaginary parts of the noise process at frequency ω. The noise is also assumed to be much smaller than the signal to obtain a reasonably good system performance. The nonlinear effects from the noise on the signal are therefore neglected.

In previous studies on the nonlinear interaction between noise and signal, the signal was assumed to be CW [56, 57]. With a CW signal, the noise is still stationary, experiencing only a change in the power spectrum and a correlation between the in-phase and out-of-phase components. However, with a modulated signal, the noise is no longer stationary since the noise is signal-dependent after its interaction with the signal. In this case, the power spectrum can no longer be defined.

In this chapter, we use an arbitrarily modulated signal with a pseudo-random input bit sequence and finite window length as used in any SSF simulation. To fully characterize the output noise process, we find the correlation functions between the frequency components

of the noise at the output and we define the correlation functions of the noise frequency components as

$$S_{XY}(\omega_1, \omega_2) \quad = \quad E\{X(\omega_1)Y(\omega_2)\}. \tag{5.1}$$

$X(\omega)$ and $Y(\omega)$ could each be $U(\omega)$ or $V(\omega)$. $E\{\cdot\}$ is the expectation operation over the random noise. When $X = Y$, we call the correlation function an auto-correlation function; otherwise, we call it a cross-correlation function. In the above definition, we have used the fact that both $U(\omega)$ and $V(\omega)$ have zero means. Since the perturbative approach gives a natural description of the output signal in the frequency domain, we focus here on the noise correlation function in the frequency domain instead of in the time domain. Fourier transformation can be used if the noise correlation function in the time domain is needed.

In the above definition of noise frequency component correlation function, real and imaginary parts of the noise process are used instead of in-phase and out-of-phase components of the noise as in [57] due to the presence of a time-varying phase on the signal. At the output of the fiber, the signal gains a time-varying phase due to nonlinearity-induced self-phase modulation (SPM). Suppose that the signal at the output can be written as $\sqrt{P(t)}\exp(j\theta(t))$ where $P(t)$ is its time-varying power and $\theta(t)$ is its phase, then the in-phase and out-of-phase noise components are $\Re\{N(t)\exp(-j\theta(t))\}$ and $\Im\{N(t)\exp(-j\theta(t))\}$ respectively. $\Re$ stands for the real part and $\Im$ stands for the imaginary part. Unfortunately, $\theta(t)$ can not be written out in closed form due to the complicated nature of the SPM process; therefore, the correlation functions based on $U(\omega)$ and $V(\omega)$ can not be replaced by the in-phase and out-of-phase components of the noise. As a result, in this chapter we describe the noise by its real and imaginary parts so that the correlation functions can be written in closed-form.

5.2 Nonlinear Effects

In Chapter 3, we have defined the SPM, XPM and FWM terms based on the VSTF method. In this chapter, we follow a perturbative approach where we assume that the nonlinearity is small enough for the nonlinear term to be treated as a perturbation. Following the perturbation approach, we divide the fiber span into many infinitesimal segments of length Δz and the small contribution of nonlinear effect from a small piece of fiber at location z is found to be

$$\Delta A_{NL}(t,z) = -j\gamma |A(t,z)|^2 A(t,z)\Delta z. \tag{5.2}$$

This small contribution then passes though the remainder of the fiber from z to L without nonlinearity. It also passes through the post-DC and the optical amplifier as shown in Fig. 5.1. The total nonlinear term at the output is the sum of these small contributions.

To find $\Delta A_{NL}(t,z)$, we approximate $A(t,z) = A_s(t,z) + N(t,z)$ with

$$A_s(t,z) = A_{s,0}(t,z) + A_{s,1}(t,z) + A_{s,2}(t,z) + \ldots \tag{5.3}$$

$$N(t,z) = N_0(t,z) + N_1(t,z) + N_2(t,z) + \ldots \tag{5.4}$$

$A_{s,0}(t,z)$ is simply the linear transmission part of the signal up to distance z, that is, the signal that passes through the fiber up to distance z without consideration of nonlinearity. It is easy to show that

$$A_{s,0}(\omega,z) = \exp(-\alpha z/2)\exp(-j\beta_2\omega^2 z/2)\exp(jD_{pre}\omega^2/2)A_{s,i}(\omega) \tag{5.5}$$

where $A_{s,i}(\omega)$ is the signal input. $N_0(t,z)$ is similarly defined. $A_{s,1}(t,z)$ and $N_1(t,z)$ are the first-order nonlinear terms in γ for the signal and noise respectively. $A_{s,2}(t,z)$ and $N_2(t,z)$ are similarly defined as the second-order nonlinear terms in γ.

5.2.1 First-Order Noise Nonlinear Effects

Substituting $A_s(t,z)$ and $N(t,z)$ into (5.2), we get the first-order nonlinear term in γ as

$$\begin{aligned}\Delta A_{s,1}(t,z)+\Delta N_1(t,z) &= -j\Delta z[|A_{s,0}(t,z)|^2A_{s,0}(t,z)+2|A_{s,0}(t,z)|^2N_0(t,z)\\ &+ A_{s,0}^2(t,z)N_0^*(t,z)+H.O.T.]. \end{aligned} \tag{5.6}$$

H.O.T. stands for higher-order terms of $N_0(t,z)$ and they are neglected in the following based on the assumption that the noise is much smaller than the signal. The first term on the right side of (5.6) is $\Delta A_{s,1}(t,z)$, the SPM effect on the signal. The second- and third-term on the right side of (5.6) are the first-order noise nonlinear amplification from the signal due to cross-phase modulation (XPM) and four-wave mixing (FWM) effects, respectively, between the signal and noise. Accordingly, we define $\Delta N_X(t,z) = -j\gamma 2|A_{s,0}(t,z)|^2N_0(t,z)\Delta z$ and $\Delta N_F(t,z) = -j\gamma A_{s,0}^2(t,z)N_0^*(t,z)\Delta z$, thus $\Delta N_1(t,z) = \Delta N_X(t,z)+\Delta N_F(t,z)$. The small nonlinear terms generated at distance z then propagates linearly through the remainder of the fiber, the post-DC and the optical amplifier, and at the output we have in the frequency domain

$$\Delta N_{o,1}(\omega,z) = \exp(\frac{\alpha}{2}L)\exp(\frac{j}{2}D_{post}\omega^2)\exp(-\frac{\alpha}{2}(L-z))\exp(-\frac{j}{2}\beta_2\omega^2(L-z))\Delta N_1(\omega,z). \tag{5.7}$$

Defining L_D as $D_{pre} = \beta_2 L_D$, and $D_{post} = \beta_2(L-L_D)$, we can simplify (5.7) to $\Delta N_{o,1}(\omega,z) = \exp(\alpha z/2)\exp(-j\beta_2\omega^2(L_D-z)/2)\ \Delta N_1(\omega,z)$ since $D_{pre}+D_{post} = \beta_2 L$. Now, we obtain the first-order noise nonlinear amplification term at the output as $N_{o,1}(\omega) = \int \Delta N_{o,1}(\omega,z)$, where the integral is taken effectively over Δz from 0 to L.

Note that $N_{o,1}(\omega)$ consists of two terms. One of the terms comes from the XPM effect

and is found to be

$$
\begin{aligned}
& N_{o,X}(\omega) \\
&= \int_0^L \exp(\frac{\alpha}{2}z)\exp(-\frac{j}{2}\beta_2\omega^2(L_D-z))(-j\gamma)2\mathcal{F}\{A_{s,0}(t,z)A^*_{s,0}(t,z)N_0(t,z)\}dz \qquad (5.8) \\
&= \int_0^L \exp(\frac{\alpha}{2}z)\exp(-\frac{j}{2}\beta_2\omega^2(L_D-z))(-\frac{j\gamma}{4\pi^2})2\int\int A_{s,0}(\omega,z)A^*_{s,0}(\omega,z) \\
&\quad N_0(\omega-\omega_1+\omega_2,z)d\omega_1 d\omega_2 dz. \qquad (5.9)
\end{aligned}
$$

$\mathcal{F}(\cdot)$ stands for the Fourier transform. The other term comes from the FWM effect and is found to be

$$
N_{o,F}(\omega) = \int_0^L \exp(\frac{\alpha}{2}z)\exp(-\frac{j}{2}\beta_2\omega^2(L_D-z))(-j\gamma)\mathcal{F}\{A_{s,0}(t,z)N^*_0(t,z)A^*_{s,0}(t,z)\}dz \qquad (5.10)
$$

Note that in (5.9), if we do the integration over z first, we will get exactly the same expression as those obtained from the truncated third-order VSTF method as described in Chapter 2.

5.2.2 Second-Order Noise Nonlinear Effects

Similar to the first-order case, we get the noise nonlinear amplification of second-order term in γ as

$$
\begin{aligned}
\Delta N_2(t,z) &= -j\gamma\Delta z[A^2_{s,0}(t,z)N^*_1(t,z) + 2A_{s,0}(t,z)A^*_{s,1}(t,z)N_0(t,z) \\
&\quad +2A_{s,0}(t,z)N^*_0(t,z)A_{s,1}(t,z) + 2A_{s,0}(t,z)A^*_{s,0}(t,z)N_1(t,z) \\
&\quad +2A_{s,1}(t,z)A^*_{s,0}(t,z)N_0(t,z) + H.O.T.]. \qquad (5.11)
\end{aligned}
$$

Note that $A_{s,1}(t,z)$ is the SPM term for the signal generated in the fiber from 0 to z as defined above. Similarly, $N_1(t,z)$ is the first-order nonlinear term for the noise generated

in the fiber from 0 to z. Substituting $N_1(t,z) = N_X(t,z) + N_F(t,z)$ into (5.11), we have 7 terms in total and we define them as

$$\Delta N_H(t,z) = (-j\gamma)A_{s,0}(t,z)N_X^*(t,z)A_{s,0}(t,z)\Delta z \tag{5.12}$$

$$\Delta N_I(t,z) = (-j\gamma)A_{s,0}(t,z)N_F^*(t,z)A_{s,0}(t,z)\Delta z \tag{5.13}$$

$$\Delta N_J(t,z) = (-j\gamma)2A_{s,0}(t,z)A_{s,0}^*(t,z)N_X(t,z)\Delta z \tag{5.14}$$

$$\Delta N_K(t,z) = (-j\gamma)2A_{s,0}(t,z)A_{s,0}^*(t,z)N_F(t,z)\Delta z \tag{5.15}$$

$$\Delta N_L(t,z) = (-j\gamma)2N_0(t,z)A_{s,0}^*(t,z)A_{s,1}(t,z)\Delta z \tag{5.16}$$

$$\Delta N_M(t,z) = (-j\gamma)2A_{s,0}(t,z)N_0^*(t,z)A_{s,1}(t,z)\Delta z \tag{5.17}$$

$$\Delta N_N(t,z) = (-j\gamma)2A_{s,0}(t,z)A_{s,1}^*(t,z)N_0(t,z)\Delta z \tag{5.18}$$

The corresponding terms at the output of the fiber, $\Delta N_{o,H}(\omega,z)$ to $\Delta N_{o,N}(\omega,z)$ can be defined similar to (5.7).

Finally, we obtain the second-order noise nonlinear amplification term at the output through integration over Δz as

$$N_{o,2}(\omega) = N_{o,H}(\omega)+N_{o,I}(\omega)+N_{o,J}(\omega)+N_{o,K}(\omega)+N_{o,L}(\omega)+N_{o,M}(\omega)+N_{o,N}(\omega) \tag{5.19}$$

and

$$N_o(\omega) \approx N_{o,0}(\omega) + N_{o,1}(\omega) + N_{o,2}(\omega) \tag{5.20}$$

with $N_{o,0}(\omega)$ is the same as the input noise $N_i(\omega)$ due to the 100% dispersion compensation used.

Note that every term in (5.20) is composed of both real and imaginary parts, $U(\omega)$ and $V(\omega)$. In the next section, we focus on the computation of the auto-correlation functions $S_{UU}(\omega_1, \omega_2)$ and $S_{VV}(\omega_1, \omega_2)$, and the cross-correlation function $S_{UV}(\omega_1, \omega_2)$.

5.3 Noise Frequency Correlation Functions

To compute the noise frequency correlation function, we substitute $U_o(\omega) \approx U_{o,0}(\omega) + U_{0,1}(\omega) + U_{o,2}(\omega)$, $V_o(\omega) \approx V_{o,0}(\omega) + V_{o,1}(\omega) + V_{o,2}(\omega)$ into (5.1) and we have for example,

$$
\begin{aligned}
& E\{U_o(\omega_1)U_o(\omega_2)\} \\
\approx\ & E\{(U_{o,0}(\omega_1) + U_{o,1}(\omega_1) + U_{o,2}(\omega_1))(U_{o,0}(\omega_2) + U_{o,1}(\omega_2) + U_{o,1}(\omega_2))\} \\
=\ & E\{U_{o,0}(\omega_1)U_{o,0}(\omega_2)\} && (5.21) \\
+\ & E\{U_{o,0}(\omega_1)U_{o,1}(\omega_2)\} + E\{U_{o,1}(\omega_1)U_{o,0}(\omega_2)\} && (5.22) \\
+\ & E\{U_{o,0}(\omega_1)U_{o,2}(\omega_2)\} + E\{U_{o,1}(\omega_1)U_{o,1}(\omega_2)\} + E\{U_{o,2}(\omega_1)U_{o,0}(\omega_2)\} && (5.23) \\
+\ & H.O.T.
\end{aligned}
$$

(5.21) is the zeroth-order nonlinear term in γ in the noise frequency correlation functions, and we call it $S_{UU,0}(\omega_1, \omega_2)$. Similarly, we define the first- and second- order nonlinear terms in the noise frequency correlation functions $S_{UU,1}(\omega_1, \omega_2)$ and $S_{UU,2}(\omega_1, \omega_2)$ defined in (5.22) and (5.23) respectively.

With our assumptions that the ASE noise is Gaussian white noise with independent real and imaginary parts at the input, we have

$$S_{UU,0}(\omega_1, \omega_2) = E\{U_{o,0}(\omega_1)U_{o,0}(\omega_2)\} = E\{V_{o,0}(\omega_1)V_{o,0}(\omega_2)\} = S_{VV,0}(\omega_1, \omega_2)$$

$$= \frac{\sigma_\omega^2}{2}\delta(\omega_1 - \omega_2) \tag{5.24}$$

where $\sigma_\omega^2 = 2\pi\Psi_0$, because frequency components for the noise are independent and

$$S_{UV,0}(\omega_1, \omega_2) = E\{U_{o,0}(\omega_1)V_{o,0}(\omega_2)\} = 0 \tag{5.25}$$

because the real and imaginary parts are independent.

In the remainder of this section, we concentrate on studying the first-order and second-order nonlinear terms in the noise frequency correlation functions.

5.3.1 First-Order Nonlinear Term for the Noise Frequency Correlation Functions

To compute the first-order nonlinear term in the noise frequency correlation functions, we have $S_{UU,1}(\omega_1, \omega_2) = E\{U_{o,0}(\omega_1)U_{o,1}(\omega_2)\} + E\{U_{o,1}(\omega_1)U_{o,0}(\omega_2)\}$, $S_{VV,1}(\omega_1, \omega_2) = E\{V_{o,0}(\omega_1)V_{o,1}(\omega_2)\} + E\{V_{o,1}(\omega_1)V_{o,0}(\omega_2)\}$, $S_{UV,1}(\omega_1, \omega_2) = E\{\ U_{o,0}(\omega_1)\ V_{o,1}(\omega_2)\} + E\{U_{o,1}\ (\omega_1)\ V_{o,0}(\omega_2)\}$. After substituting $U_{o,k}(\omega) = (N_{o,k}(\omega) + N_{o,k}^*(\omega))/2$ and $V_{o,k}(\omega) = (N_{o,k}(\omega) - N_{o,k}^*(\omega))/2j$ for $k = 0, 1$ into the above equations, we find that $S_{UU,1}(\omega_1, \omega_2)$, $S_{VV,1}(\omega_1, \omega_2)$ and $S_{UV,1}(\omega_1, \omega_2)$ are all composed of the four terms $E\{N_{o,0}(\omega_1)N_{o,1}(\omega_2)\}$, $E\{N_{o,0}(\omega_1)\ N_{o,1}^*(\omega_2)\}$, $E\{N_{o,0}^*(\omega_1)N_{o,1}(\omega_2)\}$, and $E\{N_{o,0}^*(\omega_1)N_{o,1}^*(\omega_2)\}$. For example,

$$E\{U_{o,0}(\omega_1)U_{o,1}(\omega_2)\} = \frac{1}{4}E\{(N_{o,0}(\omega_1) + N_{o,0}^*(\omega_1))(N_{o,1}(\omega_2) + N_{o,1}^*(\omega_2))\} \tag{5.26}$$

For $E\{N_{o,0}(\omega_1)N_{o,1}(\omega_2)\}$, we have

$$E\{N_{o,0}(\omega_1)N_{o,1}(\omega_2)\} = E\{(N_{o,0}(\omega_1)(N_{o,X}(\omega_2) + N_{o,F}(\omega_2))\} = E\{(N_{o,0}(\omega_1)N_{o,F}(\omega_2)\} \tag{5.27}$$

with $E\{N_{o,0}(\omega_1)N_{o,X}(\omega_2)\} = 0$ because $E\{N_{o,0}(\omega_1)N_0(\omega_2, z)\} = 0$ for any z. After substituting in the expression for $N_{o,F}(\omega)$ and some mathematical manipulations, we find

that

$$E\{N_{o,0}(\omega_1)N_{o,1}(\omega_2)\} = \int_0^L \exp(-\frac{j}{2}\beta_2\omega_2^2(L_D-z))(-\frac{j\gamma}{4\pi^2})\sigma_\omega^2 \int A_{s,0}(v,z) A_{s,0}(\omega_1+\omega_2-v,z)dvdz \quad (5.28)$$
$$\triangleq T_1(\omega_1,\omega_2).$$

Similarly, we obtain that

$$E\{N_{o,0}(\omega_1)N_{o,1}^*(\omega_2)\} = \int_0^L \exp(-\frac{j}{2}\beta_2(\omega_1^2-\omega_2^2)(L_D-z))(\frac{j\gamma}{4\pi^2})\sigma_\omega^2 \int 2A_{s,0,z}(v) A_{s,0}^*(v+\omega_1-\omega_2,z)dvdz \quad (5.29)$$
$$\triangleq T_2(\omega_1,\omega_2),$$

$$E\{N_{o,0}^*(\omega_1)N_{o,1}(\omega_2)\} = T_2^*(\omega_1,\omega_2), \quad (5.30)$$

$$E\{N_{o,0}^*(\omega_1)N_{o,1}^*(\omega_2)\} = T_1^*(\omega_1,\omega_2). \quad (5.31)$$

With these expressions, we are now ready to obtain the first-order nonlinear term in the noise frequency correlation function. The first-order real noise frequency auto-correlation function $E\{U_o(\omega_1)U_o(\omega_2)\}$ is

$$S_{UU,1}(\omega_1,\omega_2) = \frac{1}{2}\{T_1(\omega_1,\omega_2)+T_1^*(\omega_1,\omega_2)\} \quad (5.32)$$
$$= \frac{1}{2}\frac{\gamma}{2\pi}\sigma_\omega^2\Im\{\int_0^L \exp(-\frac{j}{2}\beta_2(\omega_1^2+\omega_2^2)(L_D-z))C(\omega_1+\omega_2,z)dz\}. \quad (5.33)$$

In above equation, we have used the definition

$$C(\omega,z) = \frac{1}{2\pi}\int A_{s,0}(v,z)A_{s,0}(w-v,z)dv \quad (5.34)$$

or $C(t,z) = A_{s,0}^2(t,z)$ and the fact that $T_1(\omega_2,\omega_1) = T_1(\omega_1,\omega_2)$, $T_2(\omega_2,\omega_1) = -T_2^*(\omega_1,\omega_2)$.

Similarly, the first-order imaginary noise frequency auto-correlation function, $E\{V_o(\omega_1)$ $V_o(\omega_2)\}$, is

$$S_{VV,1}(\omega_1,\omega_2) = -\frac{1}{2}\left\{T_1(\omega_1,\omega_2) + T_1^*(\omega_1,\omega_2)\right\}. \tag{5.35}$$

The first-order real-imaginary noise frequency cross-correlation function, $E\{U_o(\omega_1)$ $V_o(\omega_2)\}$, is

$$S_{UV,1}(\omega_1,\omega_2) = -\frac{j}{2}\left\{T_1(\omega_1,\omega_2) - T_1^*(\omega_1,\omega_2)\right\}. \tag{5.36}$$

One interesting observation about the first-order nonlinear effect in the noise frequency correlation is that the its complex noise variance, $E\{N_o(\omega)N_o^*(\omega)\}$, is

$$E\{N_{o,0}(\omega)N_{o,1}^*(\omega) + N_{o,1}(\omega)N_{o,0}^*(\omega)\} = S_{UU,1}(\omega,\omega) + S_{VV,1}(\omega,\omega) = 0 \tag{5.37}$$

That is, the first-order nonlinear effect has no amplification over the complex noise, but it introduces an energy transfer between the real and imaginary parts plus correlations between the real and imaginary parts.

Another interesting observation is that the XPM effect does not contribute any effect to the noise frequency correlation functions up to the first-order since the term $T_2(\omega_1,\omega_2)$ that comes form the XPM effect cancels out.

5.3.2 Second-Order Nonlinear Terms for the Noise Frequency Correlation Functions

From (5.23), we have $S_{UU,2}(\omega_1,\omega_2) = E\{U_{o,0}(\omega_1)U_{o,2}(\omega_2)\} + E\{U_{o,1}(\omega_1)U_{o,1}(\omega_2)\} +$ $E\{U_{o,2}(\omega_1)\ U_{o,0}(\omega_2)\}$ and similar expressions for $S_{VV,2}(\omega_1,\omega_2)$ and $S_{UV,2}(\omega_1,\omega_2)$. After substituting $U_{o,k}(\omega) = (N_{o,k}(\omega) + N_{o,k}^*(\omega))/2$ and $V_{o,k}(\omega) = (N_{o,k}(\omega) - N_{o,k}^*(\omega))/2j$ for $k = 0, 1, 2$ into these equations, we find that $S_{UU,2}(\omega_1,\omega_2)$, $S_{VV,2}(\omega_1,\omega_2)$ and $S_{UV,2}(\omega_1,\omega_2)$ are all composed of two groups of terms. One group of these terms is of form $E\{U_{o,0}(\omega_1)$

$U_{o,2}(\omega_2)\}$ and the other group is of form $E\{U_{o,1}(\omega_1)U_{o,1}(\omega_2)\}$. We first study these two groups of terms separately in the following two subsections and we combine them in the third subsection to obtain the second-order terms in the correlation functions.

Terms of the form $E\{U_{o,0}(\omega_1)U_{o,2}(\omega_2)\}$

From (5.23), the first term of $S_{UU,2}(\omega_1,\omega_2)$ is $E\{U_{o,0}(\omega_1)U_{o,2}(\omega_2)\} = \frac{1}{4}E\{(N_{o,0}(\omega_1) + N^*_{o,0}(\omega_1))(N_{o,2}(\omega_2) + N^*_{o,2}(\omega_2))\}$. Note that $N_{o,2}(\omega)$ is composed of seven terms from $N_{o,H}(\omega)$ to $N_{o,N}(\omega)$. We list the results from each of these terms successively.

First, we derive $E\{N_{o,0}(\omega_1)N_{o,H}(\omega_2)\}$ and the result is

$$\begin{aligned} E\{N_{o,0}(\omega_1)N_{o,H}(\omega_2)\} &= 2\gamma^2\sigma_\omega^2 \int_0^L \int_y^L \exp(-\frac{j}{2}\beta_2\omega_1^2(L_D-y))\exp(-\frac{j}{2}\beta_2\omega_2^2(L_D-z)) \\ &\times \left(\frac{1}{4\pi^2}\int C(\omega_2+v,z)B(\omega_2-v,y)\exp(\frac{j}{2}\beta_2 v^2(z-y))dv\right)dzdy \\ &\triangleq T_H(\omega_1,\omega_2) \end{aligned} \tag{5.38}$$

where we have used the definition of $C(\omega,z)$ from (5.34) and a new definition of $B(\omega,z) = \frac{1}{2\pi}\int A_{s,0}(v,z)A^*_{s,0}(v-\omega,z)dv$ or $B(t,z) = |A_{s,0}(t,z)|^2$.

Equation (5.38) has a clear physical meaning. We would expect a factor of $C(t,z)$ from the FWM effect in (5.12). We also expect a factor of $B(t,y)$ from the XPM effect for $N_X(t,z)$. $C(t,z)$ and $B(t,y)$ are the nonlinear terms generated at positions z and y respectively and the factor $\exp(j\beta_2 v^2(z-y)/2)$ is the dispersion-induced phase interaction between them. The other two dispersion factors are simply the linear transmission after the nonlinear terms are generated.

The other term for $N_{o,H}(\omega)$ is of form $E\{N_{o,0}(\omega_1)N^*_{o,H}(\omega_2)\} = 0$ since $E\{N_{o,0}(\omega_1)N_0(\omega_2,z)\} = 0$ for any z.

$N_{o,H}(\omega_2)$: $T_H(\omega_1,\omega_2)$	$2\gamma^2\sigma_\omega^2\int_0^L\int_y^L \exp(-\frac{j}{2}\beta_2\omega_1^2(L_D-y))\exp(-\frac{j}{2}\beta_2\omega_2^2(L_D-z))$ $\times\left(\frac{1}{4\pi^2}\int C(\omega_2+v,z)B(\omega_2-v,y)\exp(\frac{j}{2}\beta_2 v^2(z-y))dv\right)dzdy$
$N^*_{o,I}(\omega_2)$: $T_I(\omega_1,\omega_2)$	$\gamma^2\sigma_\omega^2\int_0^L\int_y^L \exp(-\frac{j}{2}\beta_2\omega_1^2(L_D-y))\exp(\frac{j}{2}\beta_2\omega_2^2(L_D-z))$ $\times\left(\frac{1}{4\pi^2}\int C(v+\omega_1,y)C^*(v+\omega_2,z)\exp(-\frac{j}{2}\beta_2 v^2(z-y))dv\right)dzdy$
$N^*_{o,J}(\omega_2)$: $T_J(\omega_1,\omega_2)$	$-4\gamma^2\sigma_\omega^2\int_0^L\int_y^L \exp(-\frac{j}{2}\beta_2\omega_1^2(L_D-y))\exp(\frac{j}{2}\beta_2\omega_2^2(L_D-z))$ $\times\left(\frac{1}{4\pi^2}\int B(\omega_1-v,y)B^*(v-\omega_2,z)\exp(-\frac{j}{2}\beta_2 v^2(z-y))dv\right)dzdy$
$N_{o,K}(\omega_2)$: $T_K(\omega_1,\omega_2)$	$-2\gamma^2\sigma_\omega^2\int_0^L\int_y^L \exp(-\frac{j}{2}\beta_2\omega_1^2(L_D-y))\exp(-\frac{j}{2}\beta_2\omega_2^2(L_D-z))$ $\times\left(\frac{1}{4\pi^2}\int C(\omega_1+v,y)B^*(\omega_2-v,z)\exp(-\frac{j}{2}\beta_2 v^2(z-y))dv\right)dzdy$
$N^*_{o,L}(\omega_2)$: $T_L(\omega_1,\omega_2)$	$-2\gamma^2\sigma_\omega^2\int_0^L\int_y^L \exp(-\frac{j}{2}\beta_2\omega_1^2(L_D-z))\exp(\frac{j}{2}\beta_2\omega_2^2(L_D-z))$ $\times\left(\frac{1}{4\pi^2}\int D^*(v,y)A_{s,0}(\omega_1-\omega_2+v,z)\exp(\frac{j}{2}\beta_2 v^2(z-y))dv\right)$ $\times\exp(-\frac{\alpha}{2}(z-y))dzdy$
$N_{o,M}(\omega_2)$: $T_M(\omega_1,\omega_2)$	$-2\gamma^2\sigma_\omega^2\int_0^L\int_y^L \exp(-\frac{j}{2}\beta_2\omega_1^2(L_D-z))\exp(-\frac{j}{2}\beta_2\omega_2^2(L_D-z))$ $\times\left(\frac{1}{4\pi^2}\int D(v,y)A_{s,0}(\omega_1+\omega_2-v,z)\exp(-\frac{j}{2}\beta_2 v^2(z-y))dv\right)$ $\times\exp(-\frac{\alpha}{2}(z-y))dzdy$
$N^*_{o,N}(\omega_2)$: $T_N(\omega_1,\omega_2)$	$2\gamma^2\sigma_\omega^2\int_0^L\int_y^L \exp(-\frac{j}{2}\beta_2\omega_1^2(L_D-z))\exp(\frac{j}{2}\beta_2\omega_2^2(L_D-z))$ $\times\left(\frac{1}{4\pi^2}\int D(v,y)A^*_{s,0}(\omega_2-\omega_1+v,z)\exp(-\frac{j}{2}\beta_2 v^2(z-y))dv\right)$ $\times\exp(-\frac{\alpha}{2}(z-y))dzdy$

Table 5.1: Second-order nonlinear terms in the noise frequency correlation functions from the second-order nonlinear terms, $N_{o,H}(\omega)$ to $N_{o,N}(\omega)$ or their conjugates.

Similarly, we obtain other non-zero terms for $E\{U_{o,0}(\omega_1)U_{o,2}(\omega_2)\}$ and we list them in Table 5.1. A new expression $D(t,z)=A_{s,0}(t,z)A^*_{s,0}(t,z)A_{s,0}(t,z)$ is used and it originates from the SPM of the signal. All other terms are zero. It is easy to prove that $T^*_N(\omega_1,\omega_2)=-T_L(\omega_1,\omega_2)$ simply by substitution.

Terms of the form $E\{U_{o,1}(\omega_1)U_{o,1}(\omega_2)\}$

The second term of $S_{UU,2}(\omega_1,\omega_2)$ is of form $E\{U_{o,1}(\omega_1)U_{o,1}(\omega_2)\}$ or $\frac{1}{4}E\{(N_{o,1}(\omega_1)+N^*_{o,1}(\omega_1))\ (N_{o,1}(\omega_2)+N^*_{o,1}(\omega_2))\}$. Note that $N_{o,1}(\omega)$ is composed of two terms, $N_{o,X}(\omega)$ and $N_{o,F}(\omega)$. We list the non-zero terms in Table 5.2

Note that in the table, the integration for z and y are both from 0 to L. Now, we try to use $T_H(\omega_1,\omega_2)$ to $T_N(\omega_1,\omega_2)$ to express these three terms by dividing $\int_0^L\int_0^L$ into two parts

$N_{o,X}(\omega_1)N_{o,F}(\omega_2)$	$-2\gamma^2\sigma_\omega^2 \int_0^L \int_0^L \exp(-\frac{j}{2}\beta_2\omega_1^2(L_D-z))\exp(-\frac{j}{2}\beta_2\omega_2^2(L_D-y))$ $\times \left(\frac{1}{4\pi^2}\int B(\omega_1-v,z)C(\omega_2+v,y)\exp(-\frac{j}{2}\beta_2 v^2(z-y))dv\right)dzdy$
$N_{o,X}(\omega_1)N_{o,X}^*(\omega_2)$	$-4\gamma^2\sigma_\omega^2 \int_0^L \int_0^L \exp(-\frac{j}{2}\beta_2\omega_1^2(L_D-z))\exp(\frac{j}{2}\beta_2\omega_2^2(L_D-y))$ $\times \left(\frac{1}{4\pi^2}\int B(\omega_1-v,z)C(v-\omega_2,z)\exp(-\frac{j}{2}\beta_2 v^2(z-y))dv\right)dzdy$
$N_{o,F}(\omega_1)N_{o,F}^*(\omega_2)$	$\gamma^2\sigma_\omega^2 \int_0^L \int_0^L \exp(-\frac{j}{2}\beta_2\omega_1^2(L_D-z))\exp(\frac{j}{2}\beta_2\omega_2^2(L_D-y))$ $\times \left(\frac{1}{4\pi^2}\int C(\omega_1+v,z)C^*(\omega_2+v,y)\exp(\frac{j}{2}\beta_2 v^2(z-y))dv\right)dzdy$

Table 5.2: Second-order nonlinear terms in the noise frequency correlation functions from the beating of the first-order nonlinear terms, $N_{o,X}(\omega)$ and $N_{o,F}(\omega)$ and their conjugates.

$\int_0^L \int_y^L$ and $\int_0^L \int_0^y$. After careful mathematical manipulation, we obtain

$$E\{N_{o,X}(\omega_1)N_{o,F}(\omega_2)\} = T_K(\omega_2,\omega_1) - T_H(\omega_1,\omega_2) \tag{5.39}$$

$$E\{N_{o,X}(\omega_1)N_{o,X}^*(\omega_2)\} = -T_J(\omega_1,\omega_2) - T_J^*(\omega_2,\omega_1) \tag{5.40}$$

$$E\{N_{o,F}(\omega_1)N_{o,F}^*(\omega_2)\} = T_I(\omega_1,\omega_2) + T_I^*(\omega_2,\omega_1) \tag{5.41}$$

Combined second-order nonlinear terms in the correlations

Now we combine the results in the previous two sub-sections to obtain the final results on the second-order nonlinear effects on the noise frequency correlation functions. The results are the following:

$$\begin{aligned} S_{UU,2}(\omega_1,\omega_2) &= \Re\{T_I(\omega_1,\omega_2)+T_I(\omega_2,\omega_1)\} + \Re\{(T_K(\omega_1,\omega_2)+T_K(\omega_2,\omega_1)\} \\ &\quad +\frac{1}{2}\Re\{T_M(\omega_1,\omega_2)+T_M(\omega_2,\omega_1)\}, \end{aligned} \tag{5.42}$$

$$\begin{aligned} S_{VV,2}(\omega_1,\omega_2) &= \Re\{T_I(\omega_1,\omega_2)+T_I(\omega_2,\omega_1)\} - \Re\{T_K(\omega_1,\omega_2)+T_K(\omega_2,\omega_1)\} \\ &\quad -\frac{1}{2}\Re\{T_M(\omega_1,\omega_2)+T_M(\omega_2,\omega_1)\}, \end{aligned} \tag{5.43}$$

$$\begin{aligned} S_{UV,2}(\omega_1,\omega_2) &= -\Im\{T_I(\omega_1,\omega_2)-T_I(\omega_2,\omega_1)\} + \Im\{T_K(\omega_1,\omega_2)+T_K(\omega_2,\omega_1)\} \\ &\quad +\frac{1}{2}\Im\{T_M(\omega_1,\omega_2)+T_M(\omega_2,\omega_1)\}. \end{aligned} \tag{5.44}$$

Comparing with the first-order nonlinear effects, we find that the second-order effect for the complex noise frequency domain variance is

$$S_{UU,2}(\omega,\omega)+S_{VV,2}(\omega,\omega)=2[T_I(\omega,\omega)+T_I^*(\omega,\omega)]\neq 0. \tag{5.45}$$

That is, unlike the first-order effect, the second-order nonlinear effect has non-zero amplification over the complex noise. Moreover, it also introduces energy transfer and correlation between the real and imaginary parts.

5.4 Extension to Multiple Spans

In this section, we extend our results on the noise nonlinear amplification to multi-span systems. We also discuss the computational complexity associated with our method and compare it with the SSF method.

5.4.1 Noise Correlations for Multiple Spans

For usual multi-span systems, we have both multiple fiber spans and multiple optical amplifiers. Each optical amplifier introduces ASE noise independent of each other and the ASE noise from each optical amplifier experiences nonlinear amplification by the signal during the transmission through the following fiber spans after its generation.

As a first step to study the noise nonlinear amplification in multi-span systems, we study the following system shown in Fig. 5.2 where only one ASE noise source is included. The input to the multi-span system is still the sum of the modulated signal and the ASE noise from the previous optical amplifier. All following optical amplifiers are assumed to be noiseless. With the correlation functions found for this case, the correlation functions can be easily extended to the case of multi-span systems with multiple ASE sources due to the

fact that the ASE noise sources from different amplifiers are independent. The extension is included at the end of this subsection.

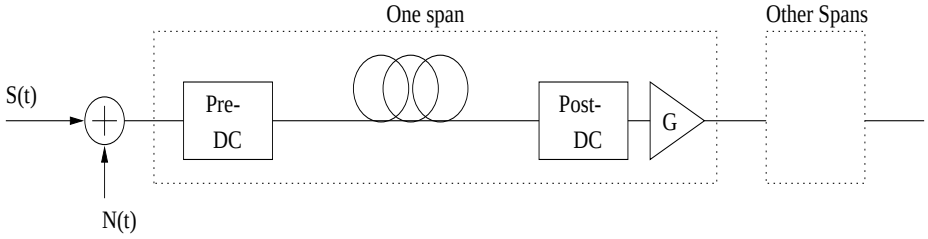


Figure 5.2: Simplified system model for multi-span system.

Since 100% dispersion compensation and perfect loss compensation are used within each span, we have the following relationship

$$A_{s,0}(t, z + m \times L) = A_{s,0}(t, z), \tag{5.46}$$

$$N_0(t, z + m \times L) = N_0(t, z), \tag{5.47}$$

for different span number $m = 1, 2, \ldots$ for both the modulated signal and the noise. It is then easy to show that

$$N_{o,X,m}(\omega) = mN_{o,X}(\omega), \qquad N_{o,F,m}(\omega) = mN_{o,F}(\omega) \tag{5.48}$$

where $N_{o,X,m}(\omega)$ and $N_{o,F,m}(\omega)$ are the first-order noise nonlinear amplification terms from the XPM and FWM effects respectively for m spans and $N_{o,X}(\omega)$ and $N_{o,F}(\omega)$ are the corresponding terms from the single-span case. As a result, we have the first-order nonlinear terms for the noise frequency correlation functions as

$$S_{UU,1,m}(\omega_1, \omega_2) = \frac{1}{2}m\{T_1(\omega_1, \omega_2) + T_1^*(\omega_1, \omega_2)\}, \tag{5.49}$$

$$S_{VV,1,m}(\omega_1, \omega_2) = -\frac{1}{2}m\{T_1(\omega_1, \omega_2) + T_1^*(\omega_1, \omega_2)\}, \tag{5.50}$$

$$S_{UV,1,m}(\omega_1, \omega_2) = -\frac{j}{2}m\{T_1(\omega_1, \omega_2) - T_1^*(\omega_1, \omega_2)\}. \tag{5.51}$$

The first-order noise nonlinear terms $N_{o,X}(\omega)$ and $N_{o,F}(\omega)$ contribute not only to the first-order nonlinear terms in the noise frequency correlation functions, i.e., terms of the form $E\{N_{o,0}(\omega_1)N_{o,1}(\omega_2)\}$, but also the second-order nonlinear terms for the noise frequency correlation functions through terms of the form $E\{N_{o,1}(\omega_1)N_{o,1}(\omega_2)\}$. It is easy to see that their contributions to the second-order nonlinear terms for the noise frequency correlation functions are the same as those from the single-span except a factor of m^2 from (5.48).

Now we study the second-order nonlinear terms in the noise frequency correlation functions (that is, terms of the form $E\{N_{o,0}(\omega_1)N_{o,2}(\omega_2)\}$) from the second-order noise nonlinear terms, $N_{o,H,m}(\omega)$ to $N_{o,N,m}(\omega)$. For multi-span systems, we have $N_{o,H,m}(\omega)$ as an integration over $[0, mL]$, that is,

$$\begin{aligned} N_{o,H,m}(\omega) &= \int_0^{mL} \exp(\frac{\alpha}{2}\tilde{z}) \exp(-\frac{j}{2}\beta_2\omega^2(L_D-\tilde{z}))(-\frac{j\gamma}{4\pi^2}) \int\int A_{s,0}(v_3,z) \\ &\quad A_{s,0}(\omega - v_3 + v_4, z) \int_0^z \exp(\frac{j}{2}\beta_2 v_4^2(\tilde{z}-\tilde{y})) \exp(-\frac{\alpha}{2}(\tilde{z}-\tilde{y}))\frac{j\gamma}{4\pi^2} \\ &\quad 2\int\int A_{s,0}(v_1,y)A_{s,0}(v_2,y)N_0^*(v_4-v_1+v_2,y)d^4v dy dz \end{aligned} \tag{5.52}$$

with $\tilde{z} = z - \text{floor}(z/L) \times L$, $\tilde{y} = y - \text{floor}(y/L) \times L$ as the transmission distance after the last optical amplifier. Using (5.46) and (5.47) and integrating first over the four frequency variables in (5.52), (5.52) can be separated into two terms with integration limits of $\int_0^L \int_0^z$ and $\int_0^L \int_z^L$ and the result is

$$N_{o,H,m}(\omega) = \frac{m^2+m}{2}\int_0^L\int_0^z \Delta N_H(\omega,y,z)dydz + \frac{m^2-m}{2}\int_0^L\int_z^L \Delta N_H(\omega,y,z)dydz \tag{5.53}$$

with $\Delta N_H(\omega, y, z)$ as the total integrand in (5.52), including integration over the frequency variables. For other terms from $N_{o,I,m}(\omega)$ to $N_{o,N,m}(\omega)$, similar things can be done.

The first double integration in (5.53) is simply $N_{o,H}(\omega)$; the second term is a new term,

which we name as $N'_{o,H}(\omega)$. We substitute (5.53) into $E\{N_{o,0}(\omega_1)N_{o,H,m}(\omega_2)\}$ for the second-order nonlinear effects for the noise frequency correlation, we obtain

$$E\{N_{o,0}(\omega_1)N_{o,H,m}(\omega_2)\} = \frac{m^2+m}{2}T_H(\omega_1,\omega_2) + \frac{m^2-m}{2}T'_H(\omega_1,\omega_2) \tag{5.54}$$

with $T'_H(\omega_1,\omega_2)$ defined similarly as $T_H(\omega_1,\omega_2)$ in (5.38) using the second term on the right side of (5.53). Note that in the above equation we have used the fact that $N_{o,0,m}(\omega) = N_{o,0}(\omega)$ for any span number m due to the 100% dispersion compensation per span. With some mathematical manipulation, we can show that

$$T'_H(\omega_1,\omega_2) = -T_K(\omega_2,\omega_1). \tag{5.55}$$

For other terms from $N_{o,I,m}(\omega)$ to $N_{o,N,m}(\omega)$, we obtain similar results summarized as

$$T'_I(\omega_1,\omega_2) = T^*_I(\omega_2,\omega_1) \tag{5.56}$$

$$T'_J(\omega_1,\omega_2) = T^*_J(\omega_2,\omega_1) \tag{5.57}$$

$$T'_K(\omega_1,\omega_2) = T'_H(\omega_2,\omega_1) \tag{5.58}$$

$$T'_L(\omega_1,\omega_2) = -T'^*_N(\omega_2,\omega_1) \tag{5.59}$$

Following the same procedures as that for the single-span case, we obtain

$$\begin{aligned} S_{UU,2,m}(\omega_1,\omega_2) &= \frac{1}{4}\times\{2m^2\tilde{T}_I(\omega_1,\omega_2) - (m^2-m)\tilde{T}_H(\omega_1,\omega_2) + (m^2-m)\tilde{T}_K(\omega_1,\omega_2) \\ &+ \frac{m^2+m}{2}\tilde{T}_M(\omega_1,\omega_2) + \frac{m^2-m}{2}\tilde{T}'_M(\omega_1,\omega_2)\} \end{aligned} \tag{5.60}$$

$$\begin{aligned} S_{VV,2,m}(\omega_1,\omega_2) &= \frac{1}{4}\times\{2m^2\tilde{T}_I(\omega_1,\omega_2) + (m^2-m)\tilde{T}_H(\omega_1,\omega_2) - (m^2-m)\tilde{T}_K(\omega_1,\omega_2) \\ &- \frac{m^2+m}{2}\tilde{T}_M(\omega_1,\omega_2) - \frac{m^2-m}{2}\tilde{T}'_M(\omega_1,\omega_2)\} \end{aligned} \tag{5.61}$$

$$\begin{aligned} S_{UV,2,m}(\omega_1,\omega_2) &= \frac{j}{4}\times\{2m^2\tilde{\tilde{T}}_I(\omega_1,\omega_2) + (m^2-m)\tilde{\tilde{T}}_H(\omega_1,\omega_2) - (m^2-m)\tilde{\tilde{T}}_K(\omega_1,\omega_2) \\ &- \frac{m^2+m}{2}\tilde{\tilde{T}}_M(\omega_1,\omega_2) - \frac{m^2-m}{2}\tilde{\tilde{T}}'_M(\omega_1,\omega_2)\} \end{aligned} \tag{5.62}$$

with the simplifying notations

$$\tilde{T}_i(\omega_1,\omega_2) \;=\; T_i(\omega_1,\omega_2)+T_i^*(\omega_1,\omega_2)+T_i(\omega_2,\omega_1)+T_i^*(\omega_2,\omega_1), i=I,H,K,M,M' \tag{5.63}$$

$$\tilde{\tilde{T}}_I(\omega_1,\omega_2) \;=\; T_I(\omega_1,\omega_2)-T_I^*(\omega_1,\omega_2)-T_I(\omega_2,\omega_1)+T_I^*(\omega_2,\omega_1) \tag{5.64}$$

$$\tilde{\tilde{T}}_i(\omega_1,\omega_2) \;=\; T_i(\omega_1,\omega_2)-T_i^*(\omega_1,\omega_2)+T_i(\omega_2,\omega_1)-T_i^*(\omega_2,\omega_1), i=H,K,M,M' \tag{5.65}$$

For the second-order nonlinear effects on the complex noise variance, $E\{N_o(\omega)\,N_o^*(\omega)\}$, we still have

$$S_{UU,2,m}(\omega,\omega)+S_{VV,2,m}(\omega,\omega)=m^2\tilde{T}_I(\omega,\omega). \tag{5.66}$$

That is, the noise nonlinear amplification is simply a factor of m^2 larger, but the noise energy transfer and correlations are more complicated.

Now we are ready to compute the noise frequency correlation functions with multiple ASE noise sources. With the fact that the ASE noise from each optical amplifier is independent of each other, we can show that

$$SS_{XY,i,N_T}(\omega_1,\omega_2)=\sum_{k=0}^{N_T-1} S_{XY,i,N_T-k}(\omega_1,\omega_2), i=0,1,2, \tag{5.67}$$

for a system with N_T spans in total. X and Y can be either U or V. Note that we have used notation $S_{XY,i,m}(\omega_1,\omega_2)$ as the i-th order nonlinear term for the noise frequency correlation functions with single ASE noise source passing through m spans of fibers with the signal. The double S notation $SS_{XY,i,m}(\omega_1,\omega_2)$ is for the noise frequency correlation functions with multiple ASE noise sources.

5.4.2 Computational Complexity

To build the correlation matrix for the noise frequency components, we need to compute the following terms: $T_1(\omega_1, \omega_2)$, $T_I(\omega_1, \omega_2)$, $T_H(\omega_1, \omega_2)$, $T_K(\omega_1, \omega_2)$, $T_M(\omega_1, \omega_2)$ and $T'_M(\omega_1, \omega_2)$. We discuss here the computational complexity associated with computing these terms.

We first re-write $T_1(\omega_1, \omega_2)$ as follows

$$T_1(\omega_1, \omega_2) = \int_0^L \exp(-\frac{j}{2}\beta_2\omega_2^2(L_D - z)) \times (-\frac{j\gamma}{4\pi^2})\sigma_\omega^2 \int A_{s,0}(v, z) A_{s,0}(\omega_1 + \omega_2 - v, z) dv dz. \tag{5.68}$$

Note that in the above equation, $A_{s,0}(\omega, z)$ is the linear propagation of the signal up to distance z and it can be easily computed in the frequency domain. In the following discussion, we assume that both $A_{s,0}(t, z)$ and $A_{s,0}(\omega, z)$ are available for every z we use in the numerical integration. Then $C(t, z) = A_{s,0}^2(t, z)$ can be easily computed. Note that the inner-integral in (5.68) is a convolution resulting in $C(\omega_1 + \omega_2)$. Now, for each ω_1 and z, we can write the integrand in (5.68) as

$$\mathcal{F}\{\mathcal{F}^{-1}\{\exp(-\frac{j}{2}\beta_2\omega_2^2(L_D - z))\} \times \mathcal{F}^{-1}\{C(\omega_1 + \omega_2)\}\} \tag{5.69}$$

with the help of the Fast Fourier transform (FFT) and inverse FFT (IFFT). As a result, we need $3N_zN_\omega$ FFT or IFFT to compute $T_1(\omega_1, \omega_2)$, with N_z and N_ω defined as the number of steps in the transmission distance and frequency respectively.

For the second-order nonlinear terms, we first look at $T_I(\omega_1, \omega_2)$. If we compute $T_I(\omega_1, \omega_2)$ in the same way as $T_1(\omega_1, \omega_2)$, we need about $N_z^2N_w$ FFT or IFFT due to the double integration over the transmission distance. However, we can do much better. For

each ω_1 and z, we integrate over y first, that is, we rewrite $T_I(\omega_1, \omega_2)$ as

$$\begin{aligned} T_I(\omega_1,\omega_2) &= \gamma^2\sigma_\omega^2 \int_0^L \exp(\frac{j}{2}\beta_2\omega_2^2(L_D-z))\frac{1}{4\pi^2}\int C^*(v+\omega_2,z)\exp(-\frac{j}{2}\beta_2 v^2 z) \\ &\times \int_0^z \exp(-\frac{j}{2}\beta_2\omega_1^2(L_D-y))\exp(\frac{j}{2}\beta_2 v^2 y)C(v+\omega_1,y)dydvdz \end{aligned} \quad (5.70)$$

Note that $\int_0^z \exp(-\frac{j}{2}\beta_2\omega_1^2(L_D-y))\exp(\frac{j}{2}\beta_2 v^2 y)C(v+\omega_1,y)dy$ is simply a function of v if ω_1 is fixed, then we can combine it with $\exp(-\frac{j}{2}\beta_2 v^2 z)$ and convolve the product with $C^*(v+\omega_2,z)$ as for the $T_1(\omega_1,\omega_2)$ term. Supposing that the FFT and IFFT are the most time-consuming manipulation, we only need the same number of FFT and IFFT as $T_1(\omega_1,\omega_2)$ and the computational cost is similar with some extra computation for addition.

For all other terms, we have confirmed that every second-order nonlinear term can be computed in a similar way with the number of FFT or IFFT about $3N_zN_\omega$.

Now, we take a look at how many FFT or IFFT are needed for the SSF simulation as a comparison. At each step, we need at least one FFT to transform the signal to the frequency-domain and compute the linear effect. Then, one IFFT is needed to transform the signal back to the time-domain and add the nonlinear effect. As a result, each SSF simulation needs $2N_z$ FFT or IFFT.

However, to obtain the noise frequency correlation function, simulation with different realization of the noise must be used. To get an idea of how many noise realizations are needed, we study the estimation confidence of estimating the variance of a zero-mean Gaussian random variable. Suppose the variance of the Gaussian random variable is σ^2, and the number of random observations is N_R, then the unbiased estimation for the variance is $\frac{1}{N_R-1}\sum y_k^2$, where y_k is one independent observation. For large N_R, $\frac{1}{N_R}\sum y_k^2$ can

	δ = 0.9	δ = 0.95	δ = 0.99
β = 0.1	542	768	2,170
β = 0.05	2,170	3,070	8,660
β = 0.01	54,000	77,000	216,000

Table 5.3: Number of random observations needed for different estimation confidence levels.

be used instead. Now, we need to solve for N_R from the bound

$$P\left(|\frac{1}{N_R}\sum y_k^2 - \sigma^2| \leq \beta\sigma^2\right) > \delta \tag{5.71}$$

with fixed (δ, β). The parameter β gives the tolerable fluctuation range of our estimation. The smaller β is, the less fluctuation is tolerable and more observations are needed. The parameter δ gives the estimation confidence. The larger δ is, the more confidence we have in our estimation and more observations are needed. The result is included in TABLE 5.3. From the table, we know that over three thousands observations are needed in order to get the variance estimation below 5% fluctuation with a confidence of only 95%. Practically, we need much more simulations to get a good enough estimation for the entire correlation function.

There are other ways to reduce the computational cost for our method. First, an adaptive step size can be used as in the adaptive SSF method [1]. Second, we don't need to compute the whole frequency range since the noise amplification is band-limited as in the CW pump case. As a result, we only need to compute the noise frequency correlation function in the frequency range where the nonlinear amplification is significant. For those frequency components outside the range, we can simply assume that they are not affected by the signal and they are still independent from all other frequency components.

Based on the above discussion on the computational complexity of our method, we con-

clude that our method provides a more practical way to estimate the correlation functions of the output noise process than using simulation based on the SSF method.

5.5 Special Case of CW-Pumped Nonlinear Noise

In this section, we apply our methods to the special case where the noise is co-propagating through the fiber with a CW pump signal. One difference between a CW pump signal and a modulated pump signal is that the CW pump signal assumption simplifies the expressions for the nonlinear terms significantly and a transfer matrix for the noise process can now be found. Another difference with a CW pump signal is that the SPM effect of the CW pump signal is simply a constant phase and consequently the in-phase and quadrature noise can be easily defined and studied. We first find the transfer matrix between the in-phase and quadrature noise at the input and the output. We then compare our results with the previous results on the CW pump signal in [57] and very good agreement is found while our method is much simpler to use than the previous method.

With the CW pump signal assumption, the signal is modeled as $A_{s,i}(t) = \sqrt{P_0}$ at the input and $A_{s,0}(t,z) = \sqrt{P_0}\exp(-\frac{\alpha}{2}z)$ because the fiber dispersion does not affect the CW pump signal. In this case, we have the output noise process after one fiber span as

$$\begin{aligned} N_o(\omega) &= \left\{ N_i(\omega)\left[1 - j\frac{2\gamma P_0}{\alpha} + \frac{\gamma^2 P_0^2}{2\alpha(\alpha - j\beta_2\omega^2)} - \frac{2\gamma^2 P_0^2}{\alpha^2}\right]\right. \\ &- \left. N_i^*(-\omega)\left[j\frac{\gamma P_0}{\alpha - j\beta_2\omega^2} + \frac{2\gamma^2 P_0^2}{\alpha(2\alpha - j\beta_2\omega^2)}\right]\right\}\exp(-\frac{\alpha}{2}L)\exp(-\frac{j}{2}\beta_2 L\omega^2) \\ &\quad \exp(\frac{j}{2}D_c\omega^2)G. \end{aligned} \tag{5.72}$$

$N_o(\omega)$ and $N_i(\omega)$ refer to the noise process observed in the frequency domain at the output and input respectively. Unlike previous analysis with modulated signals, we have included

explicitely the amplifier gain G and the fiber loss. We also assumed that only a optional post-DC with $D_{post} = D_c$ is used for this case in order to compare our results with previous results later.

Due to the CW pump signal assumption, the output noise at frequency ω is only related with the noise input at frequency ω and $-\omega$. Similarly, we can write $N_o^*(-\omega)$ as

$$\begin{aligned} N_o^*(-\omega) &= \left\{ N_i^*(-\omega)\left[1 + j\frac{2\gamma P_0}{\alpha} + \frac{\gamma^2 P_0^2}{2\alpha(\alpha + j\beta_2\omega^2)} - \frac{2\gamma^2 P_0^2}{\alpha^2}\right]\right. \\ &- \left. N_i(\omega)\left[-j\frac{\gamma P_0}{\alpha + j\beta_2\omega^2} + \frac{2\gamma^2 P_0^2}{\alpha(2\alpha + j\beta_2\omega^2)}\right]\right\} \exp(-\frac{\alpha}{2}L)\exp(\frac{j}{2}\beta_2 L\omega^2) \\ &\exp(-\frac{j}{2}D_c\omega^2)G. \end{aligned} \tag{5.73}$$

Now we can write (5.72) and (5.73) in matrix form as

$$\begin{bmatrix} N_o(\omega) \\ N_o^*(-\omega) \end{bmatrix} = \mathcal{M} \begin{bmatrix} N_i(\omega) \\ N_i^*(-\omega) \end{bmatrix} \tag{5.74}$$

with

$$\mathcal{M}_{11} = \exp(-\frac{\alpha}{2}L)G\exp(-\frac{j}{2}(\beta_2 L - D_c)\omega^2)\left(1 - j\frac{2\gamma P_0}{\alpha} + \frac{\gamma^2 P_0^2}{2\alpha(\alpha - j\beta_2\omega^2)} - \frac{2\gamma^2 P_0^2}{\alpha^2}\right) \tag{5.75}$$

$$\mathcal{M}_{12} = \exp(-\frac{\alpha}{2}L)G\exp(-\frac{j}{2}(\beta_2 L - D_c)\omega^2)\left(-j\frac{\gamma P_0}{\alpha - j\beta_2\omega^2} - \frac{2\gamma^2 P_0^2}{\alpha(2\alpha - j\beta_2\omega^2)}\right) \tag{5.76}$$

$$\mathcal{M}_{21} = \exp(-\frac{\alpha}{2}L)G\exp(\frac{j}{2}(\beta_2 L - D_c)\omega^2)\left(j\frac{\gamma P_0}{\alpha + j\beta_2\omega^2} - \frac{2\gamma^2 P_0^2}{\alpha(2\alpha + j\beta_2\omega^2)}\right) \tag{5.77}$$

$$\mathcal{M}_{22} = \exp(-\frac{\alpha}{2}L)G\exp(\frac{j}{2}(\beta_2 L - D_c)\omega^2)\left(1 + j\frac{2\gamma P_0}{\alpha} + \frac{\gamma^2 P_0^2}{2\alpha(\alpha + j\beta_2\omega^2)} - \frac{2\gamma^2 P_0^2}{\alpha^2}\right). \tag{5.78}$$

The above matrix shows clear physical meanings on dispersion and nonlinearity and their interaction up to second-order in fiber nonlinearity. The factor $\exp(-\frac{j}{2}\beta_2 L\omega^2)$ and $\exp(\frac{j}{2}D_c$

ω^2) and their conjugates come from the fiber dispersion and the dispersion compensation. Other factors are composed of the linear transmission part of the noise, and the nonlinear amplification of the noise up to first- and second-order in fiber nonlinearity or signal power. For example, the term $-j\frac{2\gamma P_0}{\alpha}$ and $-j\frac{\gamma P_0}{\alpha - j\beta_2\omega^2}$ are from the XPM and FWM effects respectively. From the noise transfer matrix, it is clear that the nonlinear amplification decreases as ω increases and the interaction between the fiber loss and the fiber dispersion determines the noise amplification bandwidth.

To find the in-phase and the quadrature noise at the fiber output, first note that the CW pump signal at the output can be written as $\sqrt{P_0}\exp(j\theta_{SPM})\exp(-\alpha L)G$ with

$$\theta_{SPM} = -\frac{\gamma P_0}{\alpha}(1 - \exp(-\alpha L)) \approx -\frac{\gamma P_0}{\alpha} \tag{5.79}$$

as $\exp(-\alpha L) \ll 1$ for most systems. Then the in-phase and quadrature noise are

$$N_{I,o}(t) = \Re\{N_o(t)\exp(-j\theta_{SPM})\} = \frac{1}{2}\{N_o(t)\exp(-j\theta_{SPM}) + N_o{}^*(t)\exp(j\theta_{SPM})\} \tag{5.80}$$

and

$$N_{Q,o}(t) = \Im\{N_o(t)\exp(-j\theta_{SPM})\} = \frac{1}{2j}\{N_o(t)\exp(-j\theta_{SPM}) - N_o{}^*(t)\exp(j\theta_{SPM})\} \tag{5.81}$$

respectively where $\Re$ and $\Im$ stand for real and imaginary parts. We can also write the above equations in matrix form as

$$\begin{bmatrix} N_{I,o}(t) \\ N_{Q,o}(t) \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\exp(-j\theta_{SPM}) & \frac{1}{2}\exp(j\theta_{SPM}) \\ \frac{1}{2j}\exp(-j\theta_{SPM}) & -\frac{1}{2j}\exp(j\theta_{SPM}) \end{bmatrix} \begin{bmatrix} N_o(t) \\ N_o{}^*(t) \end{bmatrix}. \tag{5.82}$$

At the input, we have

$$N_i(t) = N_{I,i}(t) + jN_{Q,i}(t), N_i^*(t) = N_{I,i}(t) - jN_{Q,i}(t), \tag{5.83}$$

or

$$\begin{bmatrix} N_i(\omega) \\ N_i^*(-\omega) \end{bmatrix} = \begin{bmatrix} 1 & j \\ 1 & -j \end{bmatrix} \begin{bmatrix} N_{I,i}(\omega) \\ N_{Q,i}(\omega) \end{bmatrix}. \tag{5.84}$$

Now we can write out the relationship between the in-phase and quadrature noise at the input and the output in the format of a transfer matrix as

$$\begin{bmatrix} N_{I,o}(\omega) \\ N_{Q,o}(\omega) \end{bmatrix} = T \begin{bmatrix} N_{I,i}(\omega) \\ N_{Q,i}(\omega) \end{bmatrix} \tag{5.85}$$

with

$$T = \begin{bmatrix} \frac{1}{2}\exp(-j\theta_{SPM}) & \frac{1}{2}\exp(j\theta_{SPM}) \\ \frac{1}{2j}\exp(-j\theta_{SPM}) & -\frac{1}{2j}\exp(j\theta_{SPM}) \end{bmatrix} \mathcal{M} \begin{bmatrix} 1 & j \\ 1 & -j \end{bmatrix}. \tag{5.86}$$

To extend the results into systems with m spans, we can simply write

$$\begin{bmatrix} N_{I,m}(\omega) \\ N_{Q,m}(\omega) \end{bmatrix} = \prod_{k=1}^{m} T_k \begin{bmatrix} N_{I,i}(\omega) \\ N_{Q,i}(\omega) \end{bmatrix}. \tag{5.87}$$

where the total transfer matrix is the product of the transfer matrix from each following span. With the noise transfer function defined, we can follow the same procedure as in [57] to find the power spectra of the enhanced ASE noise with the transfer matrix T.

Note that the above analysis assumes a single ASE noise source at the input. The result can be easily extended to the case with multiple ASE noise sources from different optical amplifiers. In this case, the output noise after N_T spans can be characterized by

$$\begin{bmatrix} N_{I,o,N_T}(\omega) \\ N_{Q,o,N_T}(\omega) \end{bmatrix} = \sum_{m=1}^{N_T} \bar{T}_m \begin{bmatrix} N_{I,i,m}(\omega) \\ N_{Q,i,m}(\omega) \end{bmatrix} \tag{5.88}$$

where $\bar{T}_m = \prod_{k=m+1}^{N_T+1} T_k$ is the total transfer matrix for the noise introduced by the m-th optical amplifier and T_k is the noise transfer matrix from the k-th fiber span. Note that the transfer matrix for the N_T-th optical amplifier is simply an identity matrix since it does not propogate through any fiber span, thus, $\bar{T}_{N_T} = T_{N_T+1} = I_{2\times 2}$ where $I_{2\times 2}$ is a 2×2 identity

matrix. Based on the fact that the noise from different amplifiers are independent of each other, the output noise spectrum matrix S_{N_T} after the N_T spans is

$$S_{N_T} = \sum_{m=1}^{N_T} \frac{\Psi_{0,m}}{2} \begin{bmatrix} |\bar{T}_{m,11}|^2 + |\bar{T}_{m,12}|^2 & \bar{T}_{m,11}\bar{T}_{m,21} + \bar{T}_{m,12}\bar{T}_{m,22} \\ \bar{T}_{m,11}\bar{T}_{m,21} + \bar{T}_{m,12}\bar{T}_{m,22} & |\bar{T}_{m,21}|^2 + |\bar{T}_{m,22}|^2 \end{bmatrix} \tag{5.89}$$

where $\Psi_{0,m}/2$ is the power spectral density of the m-th noise source and $\bar{T}_{m,ij}$ is the (i,j)-th entry of matrix $\bar{T}_m$.

Now, we compare the gain spectra from [57] and that from our method in Figs. 5.3 and 5.4. The same parameters are used as those for Fig. 1 and 2 in [57]. The system studied is made up of 60 spans of 50km each. The amplifiers have a noise figure of 6dB and their gains are individually adjusted so that the total power (noise + signal) is preserved along the chain. Other parameters are $\alpha = 0.22$dB/km, $\gamma = 2\text{W}^{-1}\text{km}^{-1}$, $\beta_2 = \pm 2\text{ps}^2/\text{km}$ and $P_0 = 1$mW. No dispersion compensation is used here, i.e., $D_c = 0$. From the plot, it is clear that our results agree with the previous results in [57] very well while our method is much simpler to use and has a clearer physical meaning than the previous method with Hankel's functions.

Dispersion compensation has been widely used to improve the system performance. However, its effect on the noise nonlinear amplification has not been shown before. We show here a comparison of the in-phase and quadrature ASE noise gain spectra between dispersion compensated systems with anomalous dispersion and normal dispersion in Fig. 5.5. The same parameters are used as above except 100% dispersion compensation is used after each fiber span to fully compensate the fiber dispersion within each span, i.e., $D_c = \beta_2 L$. For both cases, the noise gain spectra with dispersion compensation are significantly different from the corresponding ones without dispersion compensation where the dispersion compensation broadens the bandwidth of the gain spectrum.

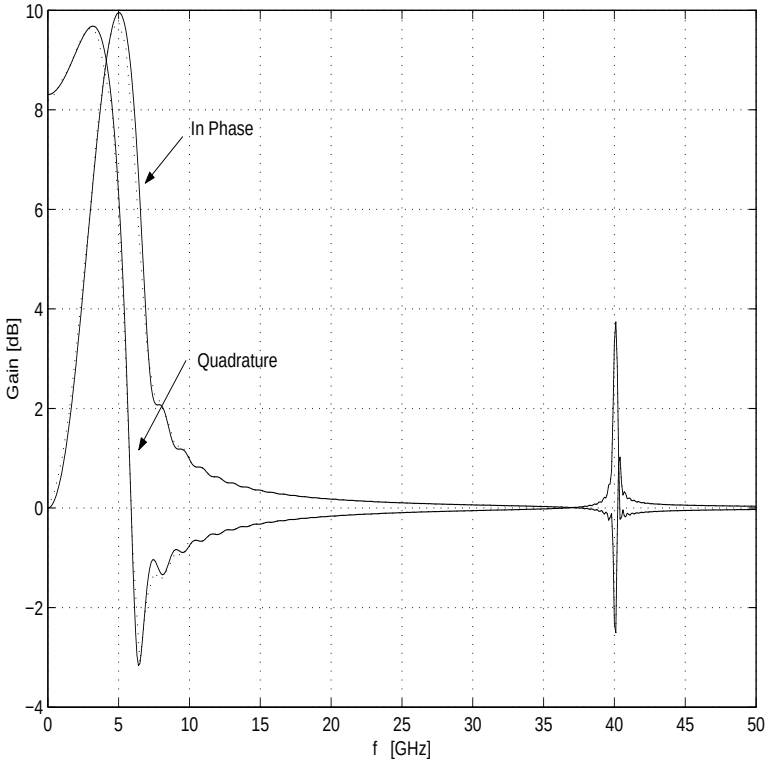


Figure 5.3: In-phase and Quadrature ASE noise gain spectra comparison between our results (solid) and the previous results (dotted) in anomalous dispersion.

5.6 Simulation Validation and Validation Range

5.6.1 Validation via Simulation

In this section, we compare our analytical results for the noise frequency correlation function with the one estimated using the SSF simulation. The simulation parameters we use are: $\beta_2 = -2.6\text{ps}^2/\text{km}$, $\gamma = 2.2\text{W}^{-1}\text{km}^{-1}$, $\alpha = 0.25\text{dB/km}$, and a two-span system with span length $L = 50\text{km}$. The signal we use is a stream of Gaussian pulses at a rate of 40Gb/s with pulse width $T_0 = 9\text{ps}$. The peak power is set to 5mW at the input. Here, a high input power is used to magnify the noise nonlinear amplification effect. A random sequence of 64 bits is used with 8 samples per bit; our results show that 8 samples per bit is enough to cover the noise nonlinear amplification range at this rate.

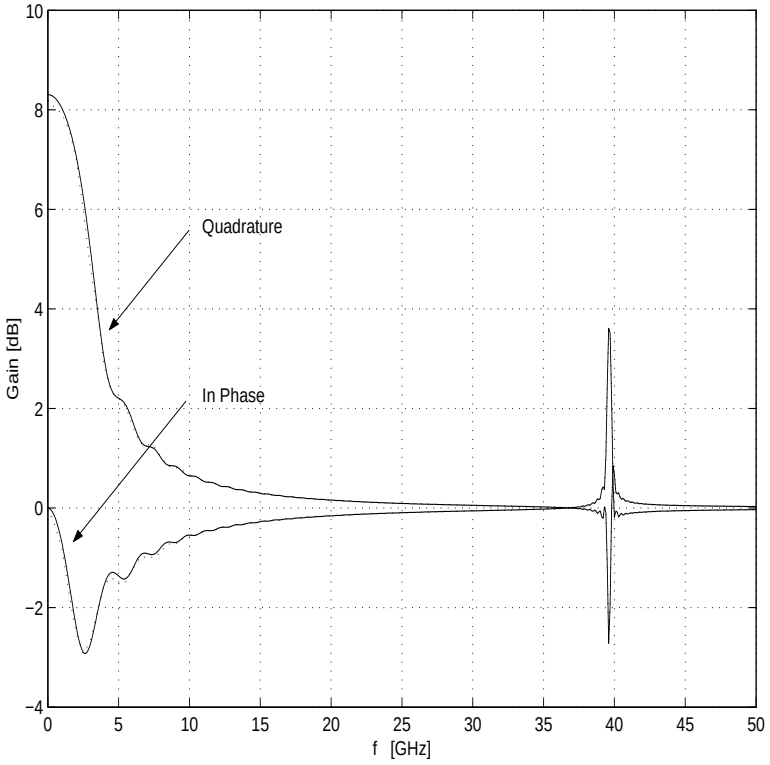


Figure 5.4: In-phase and Quadrature ASE noise gain spectra comparison between our results (solid) and the previous results (dotted) in normal dispersion.

The noise frequency correlation functions are estimated from simulation using

$$S_N(\omega_1,\omega_2) = \frac{1}{N_R}\sum N_i(\omega_1)N_i^*(\omega_2) - (\frac{1}{N_R}\sum N_i(\omega_1))(\frac{1}{N_R}\sum N_i(\omega_2))^* \quad (5.90)$$

for the complex noise processes, and

$$S_{XY}(\omega_1,\omega_2) = \frac{1}{N_R}\sum X_i(\omega_1)Y_i(\omega_2) - (\frac{1}{N_R}\sum X_i(\omega_1))(\frac{1}{N_R}\sum Y_i(\omega_2)) \quad (5.91)$$

for the real and imaginary noise process with X and Y representing U and/or V. Both of the above estimators for the noise frequency cross-correlation functions are then normalized with respect to σ_ω^2. In the above equations, $N_R = 12,000$ is the number of simulation trials used.

We first show $S_N(\omega,\omega)$ in Fig. 5.6 which is the normalized variance for the complex

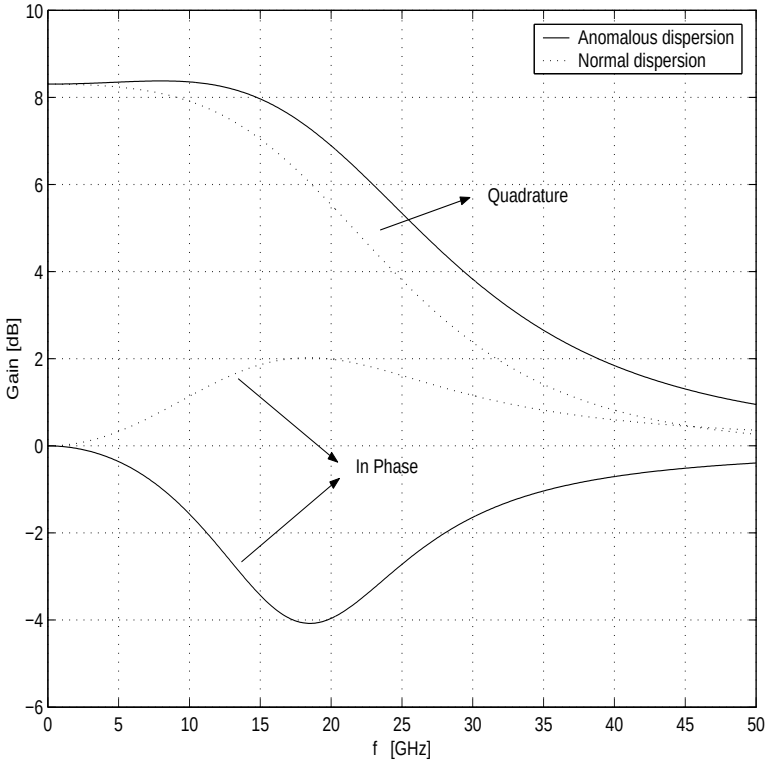


Figure 5.5: In-phase and Quadrature ASE noise gain spectra comparison between dispersion compensated systems in anomalous dispersion and in normal dispersion..

noise at each frequency. From our previous discussion, $S_N(\omega,\omega) = 4\tilde{T}_I(\omega,\omega)$, a second-order term in the nonlinearity. It can be seen from the plot that our analytical result matches the estimations quite well. The fluctuation of the estimation from the simulation is also clearly seen even for such a large number of different noise realizations.

We then show $S_{UU}(\omega,\omega)$ and $S_{VV}(\omega,\omega)$ in Fig. 5.7 and we clearly see the energy transfer between the real and imaginary parts of the noise and the nonlinear amplification on the noise.

In Fig. 5.8, we then focus on the offset auto-correlation functions, $S'_{UU}(\omega_1,\omega_2) = S_{UU}(\omega_1,\omega_2) - \frac{\sigma_\omega^2}{2}\delta(\omega_1-\omega_2)$, $S'_{VV}(\omega_1,\omega_2) = S_{VV}(\omega_1,\omega_2) - \frac{\sigma_\omega^2}{2}\delta(\omega_1-\omega_2)$ and the cross-correlation function $S_{UV}(\omega_1,\omega_2)$ with fixed $\omega_1 = 0$ and different ω_2. Note that we have

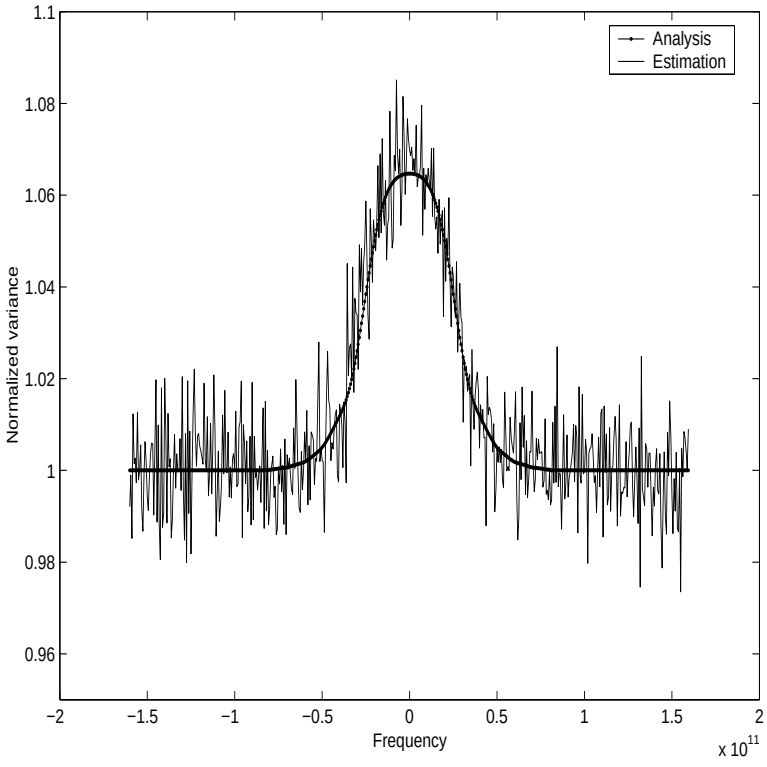
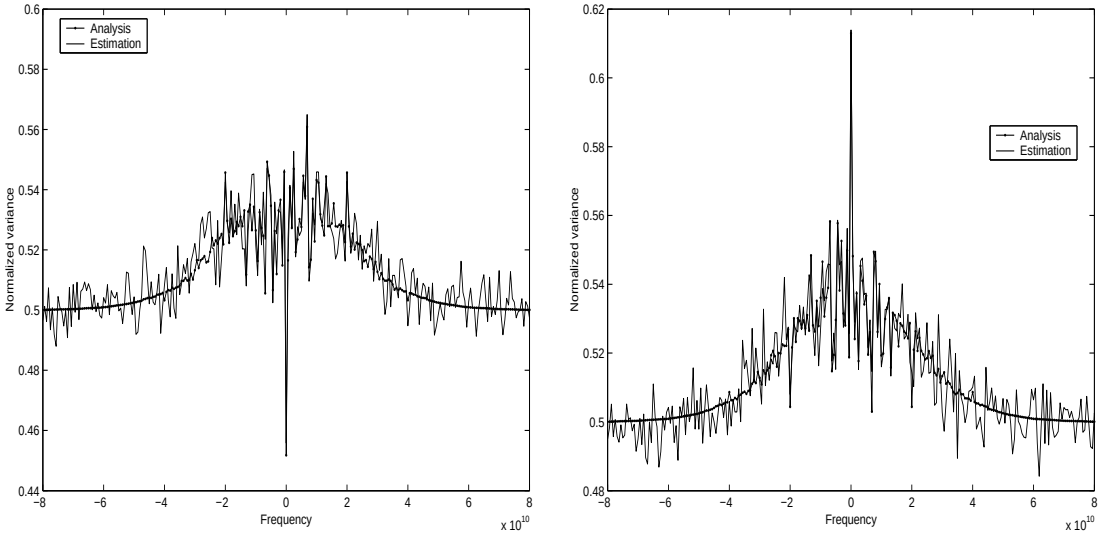


Figure 5.6: Complex noise variance at each frequency.

subtracted the linear contribution from $S_{UU}(\omega_1, \omega_2)$ and $S_{VV}(\omega_1, \omega_2)$ in order to emphasize the nonlinearity-induced correlation. Our analytical results fit the estimated results very well in the whole frequency range.

5.6.2 Parameter Range Validation

It is known that the auto-correlation matrix is always positive semi-definite for any noise, that is, the eigenvalues of this matrix should all be non-negative. However, as the span number or signal power increases, our approximated noise frequency correlation matrix might have negative eigenvalues. In Fig. 5.9 and 5.10, we show the eigenvalues of the approximated noise frequency correlation matrix for the previous 2-span case and its extension to 4-span. In the 2-span case, all the eigenvalues of the correlation matrix are



A. Normalized $S_{UU}(\omega, \omega)$ B. Normalized $S_{VV}(\omega, \omega)$

Figure 5.7: Noise variance at each frequency for real and imaginary parts.

non-negative and this agrees with our previous finding that our analytical results in this case match the estimates well from the previous sub-section. However, some of the eigenvalues are negative in the 4-span case and consequently our analytical results can not be used to approximate the real noise frequency correlation functions in this case.

The reason for this failure is our assumption that, as signal power or the number of spans increases, the nonlinear term can no longer be treated as a small, perturbative term. We show here a simple analysis on the valid range of perturbation theory for a simple case with $\beta_2 = 0$ where closed-form expressions for the signal at the fiber output exist. Assume that the input is $A_{s,i}(t) = \sqrt{P_0}g(t)$ where P_0 is the peak power at the input and $g(t)$ is the normalized pulse shape; then the output signal after m spans is

$$A_{s,m}(t) = \sqrt{P_0}g(t) \exp\left(j\frac{\gamma m P_0 |g(t)|^2}{\alpha}(1 - \exp(-\alpha L))\right). \tag{5.92}$$

Perturbation theory tries to approximate $A_{s,m}(t)$ by a linear expansion as

$$A_{s,m}(t) = A_{s,i}(t)\left\{1 + j\frac{\gamma m P_0}{\alpha}|g(t)|^2 - \frac{1}{2}(\frac{\gamma m P_0}{\alpha}|g(t)|^2)^2 - \frac{j}{6}(\frac{\gamma m P_0}{\alpha}|g(t)|^2)^3 + \ldots\right\} \tag{5.93}$$

where we have used the simplification that $\exp(-\alpha L) \ll 1$. When mP_0 is small, the approximation works very well while the approximation fails as mP_0 increases. Even though $\exp(j\theta)$ can be expanded into $\sum_{n=0}^{\infty}(j\theta)^n/n!$ for any θ, the dominating terms certainly are not the first two terms if θ is large enough. For our simulation parameters, we have $m\gamma P_0/\alpha = 0.6$ for the 4-span case, which is quite a large value.

Note that the above analysis is for the SPM effect. We know that the XPM and FWM effects are more efficient than the SPM effect, thus, the failure of the perturbation theory up to second-order for the noise nonlinear amplification analysis is entirely predictable for the input power that we are using.

As a simple rule, we will limit the valid range to where $m\gamma P_0/\alpha < 0.3$ holds for the systems studied in this chapter. Note that systems with 100% dispersion compensation per span represents the worst-case in nonlinear effects since the nonlinear terms accumulates most efficiently. For other types of systems, perturbation theory can be applied to more spans or higher powers as indicated from the discussion in Chapter 2 on the VSTF and MVSTF methods.

5.7 Application

We show a simple application of our analysis on the noise nonlinear amplification in this section. With the noise frequency correlation matrix on hand, it is possible for us to use Monte-Carlo (MC) simulation to find the probability density function (PDF) of the detector

statistics that is not possible to obtain by SSF-based MC simulation because a very large number of simulation trials would be necessary in order to estimate the tail behavior of the PDFs.

As an example, we show the PDFs for the detector samples when either bit 0 or 1 is send for a 2-span system. The parameters used are the same as before except a higher noise level. Our input is the sum of signal and noise and both in-line optical amplifiers in the 2-span system are assumed to be noise-free. The signal to noise ratio is given by $10 \log P_{AVE}/\Psi_0 = 5$dB. We picked a relatively high noise level in order to show the cross-over of the two PDFs for bit 0 and 1 respectively so that we can read the optimal threshold from the plot.

To do the MC simulation, we first use the SSF method to find the signal output without noise. Then, we generate Gaussian noise in the frequency domain according to the noise frequency correlation matrix and convert it into the time-domain. The sum of the output signal and noise is then passed through an optical filter, a photo-detector and an electrical filter. Both the optical and electrical filters are modeled as Butterworth filters of 5-th order with bandwidth of 50GHz and 32GHz respectively. About 1.5 million trials are used to generate the PDFs by histogram and a pseudo-random bit sequence of 64 bits is used for each trial. The result is shown in Fig. 5.11 together with the approximated PDFs based on a Gaussian approximation using a bit sequence of length $K = 1$ and 3. The Gaussian approximated PDF's for the detector statistics taking into account the effect of intersymbol interference (ISI) from a different number of adjacent bits can be written as

$$f_Y(y|b_0 = 0) = \frac{1}{2^{K-1}} \sum_{\underline{\tilde{b}}_0 \in \{0,1\}^{K-1}} \mathcal{N}(m(b_0 = 0, \underline{\tilde{b}}_j), \sigma^2(b_0 = 0, \underline{\tilde{b}}_j)) \tag{5.94}$$

$$f_Y(y|b_0=1)=\frac{1}{2^{K-1}}\sum_{\tilde{\underline{b}}_0\in\{0,1\}^{K-1}}\mathcal{N}(m(b_0=1,\tilde{\underline{b}}_j),\sigma^2(b_0=1,\tilde{\underline{b}}_j)) \tag{5.95}$$

for both bit 0 and 1 respectively. $\underline{b}=\ldots,b_{-1},b_0,b_1,\ldots\in\{0,1\}^K$ is the bit sequence of length K, b_0 is the central bit we are studying and $\tilde{\underline{b}}_0$ is the bit sequence excluding b_0. The notation $\mathcal{N}(m(b_0=1,\tilde{\underline{b}}_j),\sigma^2(b_0=1,\tilde{\underline{b}}_j))$ represents a Gaussian PDF with mean $m(\underline{b})$ and variance $\sigma^2(\underline{b})$ when $\underline{b}$ is send. The means and variances for the Gaussian approximation are estimated from the simulations.

The discrepancies between the true PDFs based on MC simulation and the Gaussian approximation are clearly seen especially near the tails of the PDFs. The thresholds for the maximum-likelihood (ML) detector associated with the different PDF approximations are also different. The MC-based PDFs give an optimal threshold of 1.1mW, while the Gaussian approximations with 1-bit and 3-bit give a threshold of 0.8 and 0.6mW respectively. This indicates that one must be careful in applying the Gaussian approximation in such types of fiber optic communication systems.

Another observation is that the MC-based PDFs look like Chi-square distributions as if the noise before the photo-detector was Gaussian white noise and no electrical filtering was present. However in our study, the noise before the photo-detector is modeled as non-stationary and colored due to the nonlinear interaction between the noise and the signal, and optical and electrical filtering are also present. A closer investigation is needed and is the subject of future work.

5.8 Chapter Summary

An accurate description of the noise is essential for the study of system performance. For this purpose, we have introduced a perturbation theory based method to analyze the

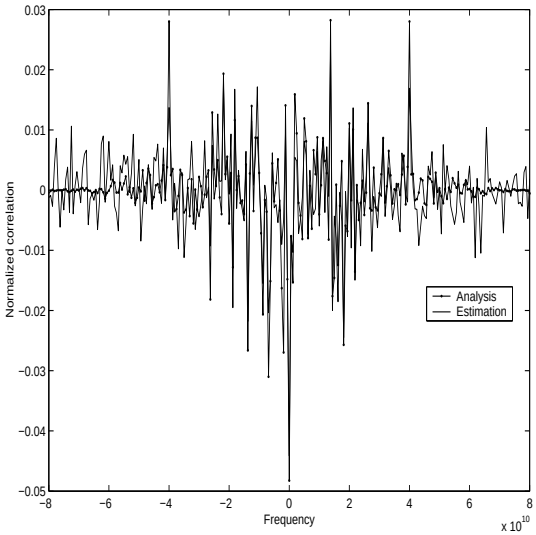
noise nonlinear amplification and coloring in an optical amplified fiber system used to transmit an arbitrarily modulated signal. Due to the nonlinear interaction between the noise and the signal, the noise becomes signal-dependent, thus not stationary. To describe such a noise process, we can no longer use the concept of a power spectrum. Instead, we use the noise correlation functions between frequency components; the noise correlation functions between any time samples can be found by Fourier transform accordingly.

The analytical results presented in this chapter are validated by comparing with numerical simulation and its range of validity is discussed. Our method is accurate within the range where perturbation theory holds but fails for large signal power or large number of spans. However, our method based on the perturbation theory still gives an idea of how the noise interacts with an arbitrarily modulated signal. A simple application is demonstrated to show the probability density function of the detector statistics, which is impossible to obtain by numerical simulation with the SSF method.

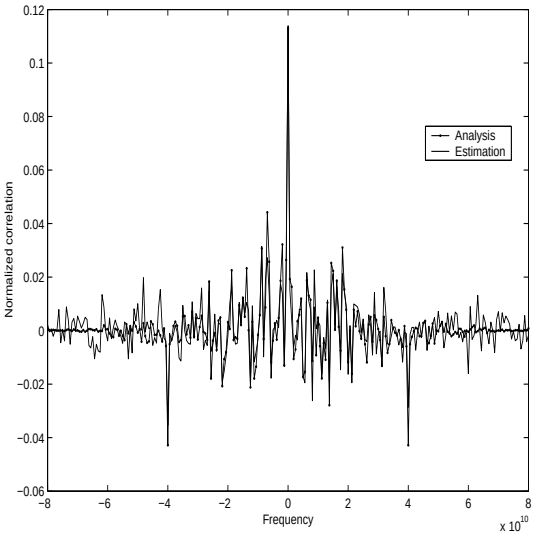
Even though this chapter focuses on a specific type of system with 100% dispersion compensation per span and lumped optical amplifiers, the analysis can be extended without much difficulty to other systems with different configurations of dispersion compensation and/or distributed amplification from Raman amplifiers. Both the dispersion and distributed amplification can be taken care of by modifying the integrands of the nonlinear terms. The analysis can also be applied to WDM systems and polarization multiplexed systems similar to the works in [59] and [60]. Other studies on the effect of different types of filtering can also be conducted efficiently using our results.

This study on noise correlation is a first step towards accurately evaluating fiber optic communication system performance. The next step is to understand the effect of the

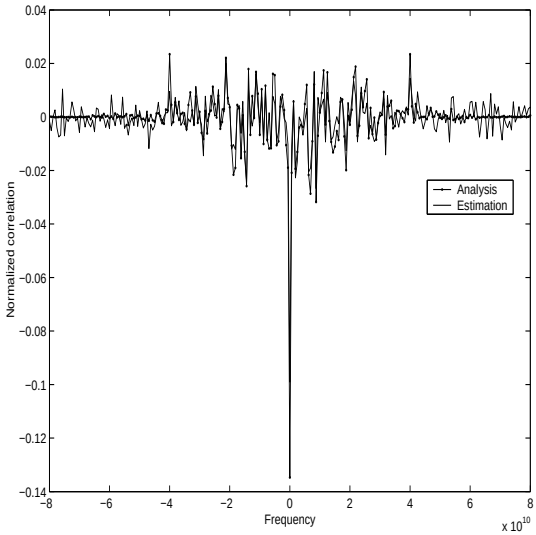
photo-detector on the non-stationary and colored noise. The square-law operation from the photo-detector introduces an interaction between the signal and the signal-dependent noise. However, the noise process after the photo-detector is no longer Gaussian distributed and thus can not be described uniquely by its correlation function. Another method should be used, such as the Karhunen-Loeve (K-L) expansion as used in [58].



A. Normalized $S'_{UU}(\omega_1 = 0, \omega_2)$



B. Normalized $S'_{VV}(\omega_1 = 0, \omega_2)$



C. Normalized $S_{UV}(\omega_1 = 0, \omega_2)$

Figure 5.8: Noise correlation between $\omega_1 = 0$ and different ω_2.

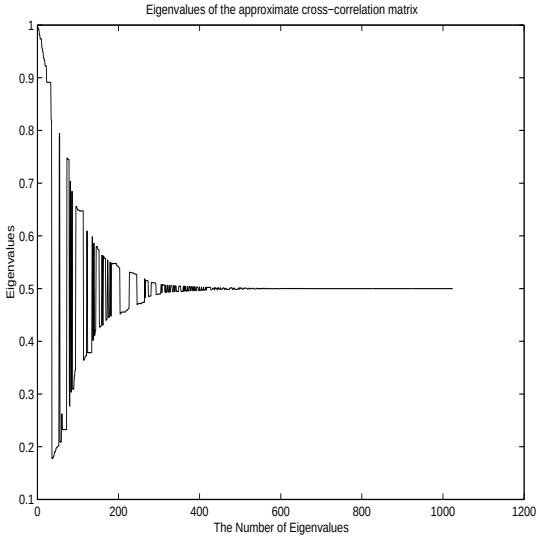


Figure 5.9: Eigenvalues of the approximate correlation matrix for 2 spans.

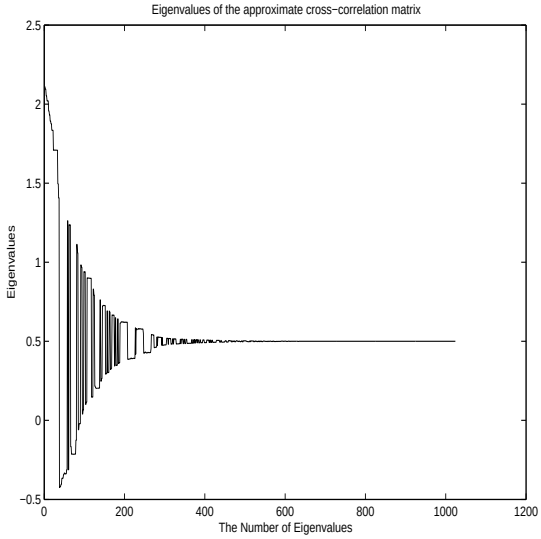


Figure 5.10: Eigenvalues of the approximate correlation matrix for 4 spans.

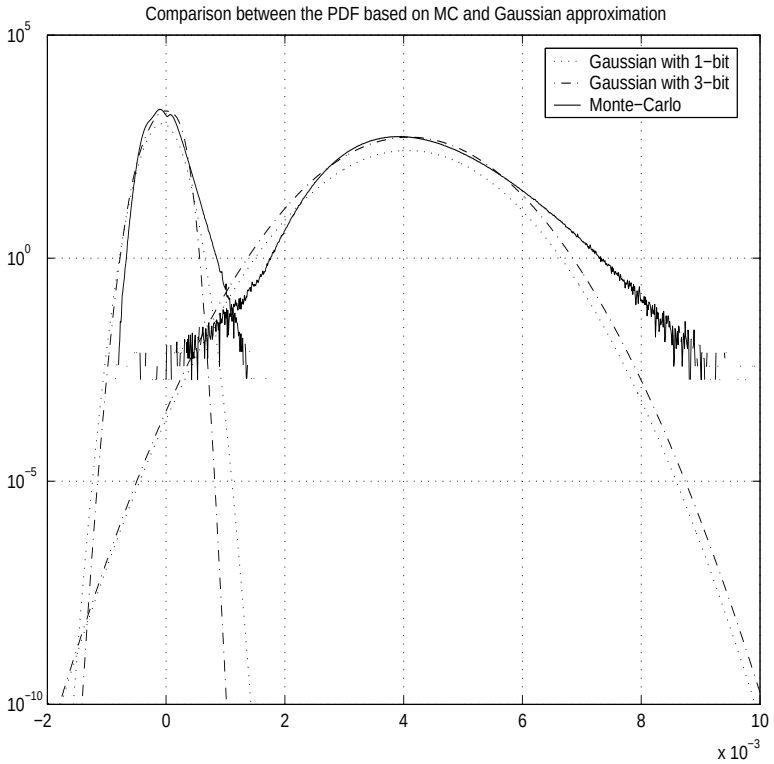


Figure 5.11: Comparison between PDFs based on MC simulation and Gaussian approximation.

Chapter 6

Optimal Dispersion Compensation

As fiber optic communication systems evolves from 10Gb/s to 40Gb/s, the nonlinear interaction between pulses introduces significant ISI. In this chapter, we study how to achieve better system performance by optimizing the dispersion compensation scheme of a system. We also show the variation of optimal dispersion compensation with the different system and device parameters.

Dispersion compensation (DC) or dispersion management has been studied and used for several years and it is one of the essential techniques that make 40Gb/s and higher-rate fiber-optic transmission possible. Much research has been done on the fabrication of better dispersion compensation devices and schemes [62, 63, 64] and on the effect of dispersion compensation on the system performance [21, 38, 22, 65]. However, there still exist important questions not yet answered for 40Gb/s systems due to the complexity of such systems.

Among the different dispersion compensation techniques, bi-end dispersion compensation introduced in [38] is one of the most comprehensive. In such systems, both a pre-dispersion compensator and a post-dispersion compensator are used to compensate the fiber dispersion within each span. As the bit rate of the systems evolves to 40Gb/s, we find that

significant system performance improvement can be achieved by optimizing the bi-end DC as we show in this chapter.

To characterize the performance of a digital communication system like the optical fiber system studied in this chapter, we usually need to know its bit error rate (BER). For complex systems where no analytical method exists to compute the BER, Monte-Carlo (MC) simulation is often used. However, one difficulty with the application of MC simulation to optical fiber communication systems is the low BER required for such systems; MC simulation can not estimate efficiently a BER of 10^{-9} or below even for the simplest system. Another difficulty with the application of MC simulation to optical fiber communication systems comes from the modeling of the fiber channel. The very time-consuming numerical method (SSF method) is the only widely-accepted method to solve the light signal transmission in a fiber channel if the complicated interaction between fiber dispersion and nonlinearity is included. As a result, BER analysis based on MC simulation and the SSF method is not feasible for systems studied in this chapter. There exist analytical methods to compute the BER of optical fiber communication systems with the assumption that the noise propagates through the fiber with the signal without nonlinear interaction between them ([66] and references therein). [58] extended the analysis in [66] by approximating the noise nonlinear interaction with a continuous-wave (CW) signal, thus limited the analysis to systems with low transmission rate. We show later in this chapter that the nonlinear interaction between the noise and the modulated signal must be carefully modeled for 40-Gbps systems and above.

Other figures of merit (FOM) instead of BER are used in practice as well. One of them is the eye-opening penalty (EOP), which indicates the signal waveform degradation after

signal transmission and processing. The EOP is fast to obtain since it neglects the effect of optical amplifier noise and the interaction between the noise and the signal. Consequently, it is only limited to systems where the error is dominated by the signal waveform degradation. Another commonly used FOM is the Q-factor based on the assumption that the receiver statistics is Gaussian distributed. Even though the validity of the Gaussian approximation is in question [58], many works still use the Q-factor to characterize the system performance as a compromise between accuracy and computational complexity.

Under a Gaussian approximation, only the mean and variance are needed to characterize the probability distribution functions (PDF) and they can be accurately estimated by a reasonable number of SSF simulation with random optical amplifier noise, random input bit sequence and random input channel phases etc. To take into account the important ISI effects between pulses for the systems studied here, a bit-sequenced Gaussian approximation is used in this work.

An optical fiber communication system consists of many components including transmitter, fiber, optical amplifier, receiver, synchronization circuits, coupler, etc; every device must be carefully modeled to obtain accurate performance analysis of such complex system. With the diverse techniques available, a device can often be modeled in different ways with different parameters. For example, the transmitter has been modeled to send out pulse trains of Gaussian, super-Gaussian or raised-cosine pulse shapes in previous works [1]. The receiver optical and electrical filters of different types and different bandwidths, and dispersion compensation fibers (DCF) with and without nonlinearity has also been used in different works [55, 66, 69, 70]. Our goal in this chapter is to study the robustness of the DC optimization with respect to systems and devices with different parameters.

This chapter is organized as follows. We first introduce the systems we study and the bit-sequenced Gaussian approximation in Section 6.1. In Section 6.2, we study three possible methods to model the optical amplifier noise and its interaction with the signal. We then study the effects of the different device parameters on the DC optimization in Section 6.3 . We summarize the conclusion of this chapter in Section 6.4.

6.1 System Model

A schematic of the systems that we study in this chapter is shown in Fig. 6.1. The transmitter sends out an on-off keying (OOK) intensity-modulated pulse train with the specified pulse shape. After the optical signal has propagated through multiple identical fiber spans with the addition of optical amplifier noise, an optical low-pass filter (LPF) is used to cut the noise level before the signal and the noise enter the photo-detector, which is simply modeled as a perfect square-law detector in this chapter. After the electrical signal from the photo-detector passes through an electrical low-pass filter and is sampled, a threshold device decides whether a 0 or 1 is sent based on the samples. The receiver is assumed to be synchronized and the samples are taken at the bit center.

In Fig. 6.1, bi-end dispersion compensation with both pre- and post-dispersion compensation is used within each span. The pre- and post-dispersion compensators are modeled as linear devices with transfer functions $\exp(j\epsilon\beta_2 L\omega^2/2)$ and $\exp(j(1-\epsilon)\beta_2 L\omega^2/2)$ respectively, where β_2 is the second-order dispersion of the fiber and L is the span length, that is, the fiber second-order dispersion $\exp(-j\beta_2 L\omega^2/2)$ is perfectly compensated by the pre- and post-dispersion compensation within each span. The parameter ϵ represents the percentage of fiber dispersion compensated by the pre-DC and we call it the *normal-*

ized pre-DC value. A lumped optical amplifier provides linear gain to compensate for the fiber transmission loss within each span while introducing amplified spontaneous emission (ASE) noise. The ASE noise is modeled as wide-band complex white Gaussian noise with power spectral density $\Psi_0 = (G-1)NF \times hv$ [71, 52]. $G = \exp(\alpha L)$ is the amplifier gain for the signal power with α as the fiber attenuation factor, h is Planck's constant, v is the optical carrier frequency and NF is the optical amplifier noise figure. The noise introduced by the different optical amplifiers are assumed to be independent. Note that other types of loss like connection loss and coupling loss have been neglected in this work. The system parameters are kept the same from span to span in our study. To find the optimal DC, we vary ϵ from 0 to 1 to achieve the best system performance as indicated by the lowest BER.

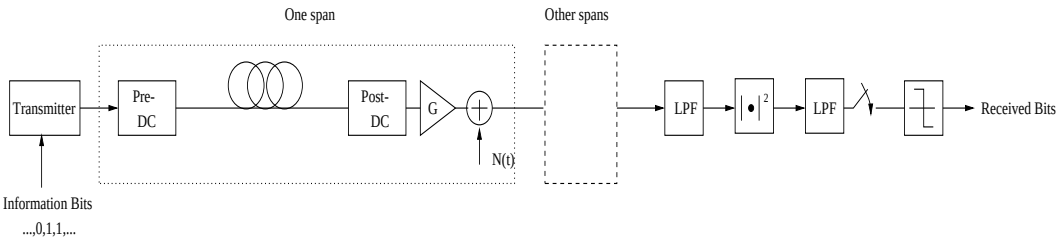


Figure 6.1: Simplified system models for multi-span system with bi-end dispersion compensation. DC is the dispersion compensator. G is the optical amplifier. LPF stands for low-pass filter.

To characterize the performance of an optical fiber communication system, in this work we use BER analysis based on a bit-sequenced Gaussian approximation of the receiver statistics as a compromise between accuracy and computational complexity. Note that the conventional Gaussian approximation is not suitable for 40Gb/s systems where the system performance might be seriously degraded by ISI from two main sources: the ISI from nonlinear interaction between pulses and the ISI from signal filtering. To take into

account the ISI, we introduce a bit-sequenced Gaussian approximation with a bit sequence of variable length K; we write the Gaussian approximated PDF's for the detector samples when bit $b_0 = 0$ and 1 are send as

$$f_Y(y|b_0 = 0) = \frac{1}{2^{K-1}} \sum_{\underline{\tilde{b}}_j \in \{0,1\}^{K-1}} \mathcal{N}(m(b_0 = 0, \underline{\tilde{b}}_j), \sigma^2(b_0 = 0, \underline{\tilde{b}}_j)), \tag{6.1}$$

$$f_Y(y|b_0 = 1) = \frac{1}{2^{K-1}} \sum_{\underline{\tilde{b}}_j \in \{0,1\}^{K-1}} \mathcal{N}(m(b_0 = 1, \underline{\tilde{b}}_j), \sigma^2(b_0 = 1, \underline{\tilde{b}}_j)) \tag{6.2}$$

where the interference from $K-1$ neighboring pulses are considered. The two PDF's are defined the same as in the Application section of the previous chapter. To compute the BER of the optimal threshold detector, we first find the optimal threshold T numerically, we then calculate the BER using

$$\begin{aligned} BER &= \frac{1}{2}P(y > T|b_0 = 0) + \frac{1}{2}P(y < T|b_0 = 1) && (6.3) \\ &= \frac{1}{2^K} \sum_{\underline{\tilde{b}}_j \in \{0,1\}^{K-1}} \mathrm{Q}\left(\frac{T - m(b_0 = 0, \underline{\tilde{b}}_j)}{\sigma(b_0 = 0, \underline{\tilde{b}}_j)}\right) + \mathrm{Q}\left(\frac{m(b_0 = 1, \underline{\tilde{b}}_j) - T}{\sigma(b_0 = 1, \underline{\tilde{b}}_j)}\right). && (6.4) \end{aligned}$$

where $\mathrm{Q}(x)$ is the Gaussian tail integral function. Note that the approximated BER from the conventional Q-factor is very close to the BER from an optimal receiver with a Gaussian approximation based on a single bit.

As a first step, we need to know what K should be used in our BER performance analysis. To obtain a reasonably accurate estimates of the mean and variance, we also need to know the number of SSF simulation realizations and the number of bits per simulation. The larger K, the more simulation realizations and the larger number of bits per simulation we use, the more accurate the computed BER performance is, but the more computation time is necessary. One should bear in mind that since the mean and variance are estimated from the simulation, the BER performances computed based on these estimations fluctuate

from trial to trial. In the following study, we have used a total of 100 realizations of different random amplifier noise with a fixed pseudo-random input signal bit sequence of length 256. Other pseudo-random input signal bits and higher numbers of simulation realization are also used to confirm that the input bit length and the number of different realizations of noise that we use are large enough to avoid large fluctuations in the BERs obtained.

To find out what K is enough for our BER analysis, we show the BER's approximated based on a Gaussian distribution with variable bit sequence lengths of 1, 3, 5, and 7 in Fig. 6.2. The system we study is a 40-span, 40Gb/s system with a pulse train of Gaussian pulse shape at the input. The pulse width T_0 is 8.8ps and the peak power of the pulse is set to 5mW. Other parameters used are fiber span length $L = 50$km, fiber attenuation factor $\alpha = 0.25$dB/km, second-order dispersion $\beta_2 = -2.6$ps^2/km and nonlinearity coefficient $\gamma = 2.2\text{W}^{-1}\text{km}^{-1}$. The optical and electrical filters are both modeled as fourth-order Butterworth filters with bandwidth of 50GHz and 32GHz respectively. The receiver thermal and shot noise are neglected here since the accumulated optical amplifier ASE noise dominates for systems of such length. From the plot, it is clear that significant BER performance improvement can be achieved by optimizing the normalized pre-DC value. However, the optimal pre-DC value might be different for different K. The huge error with the single-bit sequence Gaussian approximation is also clearly seen in the plot, consequently, the single-bit sequence Gaussian approximation can not be used to predict the system performance. While the 3-bit sequence Gaussian approximation is good enough for BER analysis at small normalized pre-DC value, it losses its accuracy as the normalized pre-DC value increases because high normalized pre-DC value introduces large pulse broadening and overlapping at the fiber input where the pulses have the highest power and consequently

the nonlinearity-induced pulse interaction and ISI increases. The results from the 5-bit and 7-bit Gaussian approximation are very close to each other for the whole range of the normalized pre-DC value. Thus, we use the 5-bit sequence Gaussian approximation in the following discussions.

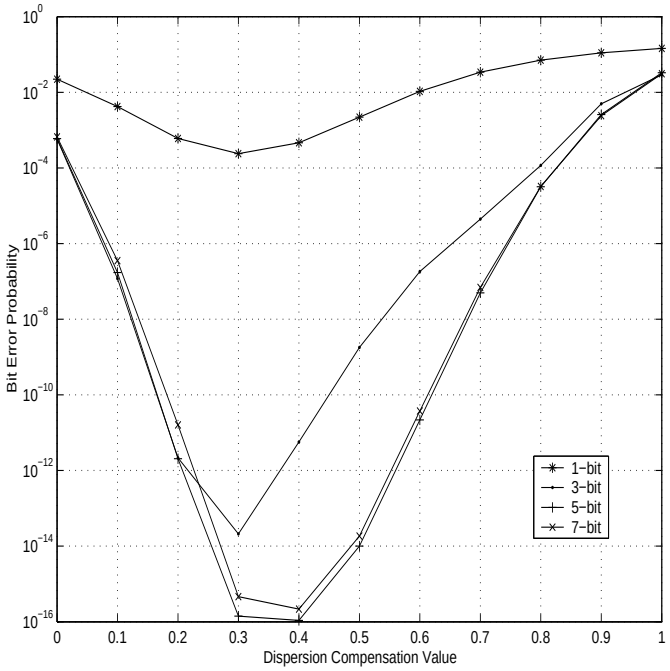


Figure 6.2: BER comparison based on Gaussian approximation with different bit sequence length of 1, 3, 5 and 7.

6.2 Noise Models

We have discussed the effect from ISI in the previous section. In this section, we focus on the other phenomenon that limits the system performance, the ASE noise. Thermal noise and shot noise from the photo-detector are neglected here as the ASE noise dominates for the systems studied in this chapter.

There exist different models for the ASE noise process after its transmission with the

signal through the fiber. The first and the simplest way to model the output noise after its co-propagation with the signal is to neglect the nonlinear interaction between the ASE noise and the signal; the noise at the output is then simply the accumulation of independent noise from the different optical amplifiers. With this model, we can write the fiber output signal as the sum of two parts, $A_s(t)$ and $N(t)$ respectively. $A_s(t)$, which is the output of the signal when the signal is transmitted alone without the noise, can be computed by the SSF method. $N(t)$ is the linearly accumulated ASE noise and its power spectral density can be easily computed. We call this noise model the linear noise model. Note that only one SSF simulation for the signal is needed with this noise model.

To obtain more accurate results on the BER performance, the nonlinear interactions between the signal and the noise must be carefully modeled. However, this is in itself a hard problem not fully solved yet. Analytical results on the interaction between the noise and the signal, or the noise nonlinear amplification, usually assume that the signal is a continuous-wave (CW) pump [56, 57, 73]. In order to use these analytical results, we propose to approximate the output noise process as if the the ASE noise was transmitted through the fiber spans with a CW signal of the same average power as the modulated signal. With this model, we can again write the fiber output as the sum of two parts, $A_s(t)$ and $N(t)$ where $A_s(t)$ is the average signal of the fiber output. Because the noise is small in order to obtain a reasonably good system BER performance, $A_s(t)$ can be safely approximated by the output of the signal as if the signal is transmitted alone without the noise and it can be computed by the SSF method. The power spectrum of $N(t)$ is computed analytically with the CW signal approximation. We call this noise model the CW pump approximation and it needs only one SSF simulation as for the linear noise model.

The most comprehensive noise model is to include the nonlinear interaction between the noise and the real modulated signal. We have proposed an analytical noise model with modulated signal in the previous chapter and intended to apply it here, but the proposed model failed for systems studied in this chapter due to the high power and large number of spans used. Without any suitable analytical noise model in hand, we have to study the noise process by simulating the transmission of the signal and the noise together using the SSF method. Note that the output noise is no longer stationary after its interaction with the modulated signal. It is clear that this noise model needs much more computation time than the other two noise models as multiple SSF simulations with different realizations of the ASE noise are needed to calculate the statistics at the photo-detector.

To compute the BER performance using the bit-sequenced Gaussian approximation, we need to estimate the mean and variance of the receiver samples for each bit pattern. We first generate multiple realization of the random ASE noise at the input and we then find the output noise using the three different noise models. The sum of the output signal and the output noise then passes through the optical filter, the photo-detector, the electrical filter and is sampled at the bit center. With all the samples that are available, the mean and variance are estimated for each bit pattern and the BER is computed using the 5-bit Gaussian approximation as discussed in the previous section.

In Fig. 6.3, we compare the BER performance based on the three noise models as the normalized pre-DC value varies. Unfortunately, the BER performances from both the linear noise model and the nonlinear noise model based on CW pump approximation are quite different from the BER performance from the full simulation model. The optimal normalized pre-DC value also varies significantly for the three noise models. As a result,

we conclude that the nonlinear interaction between the noise and the modulated signal is significant and should be carefully modeled for 40Gb/s systems. The full simulation model is adopted in the following discussion.

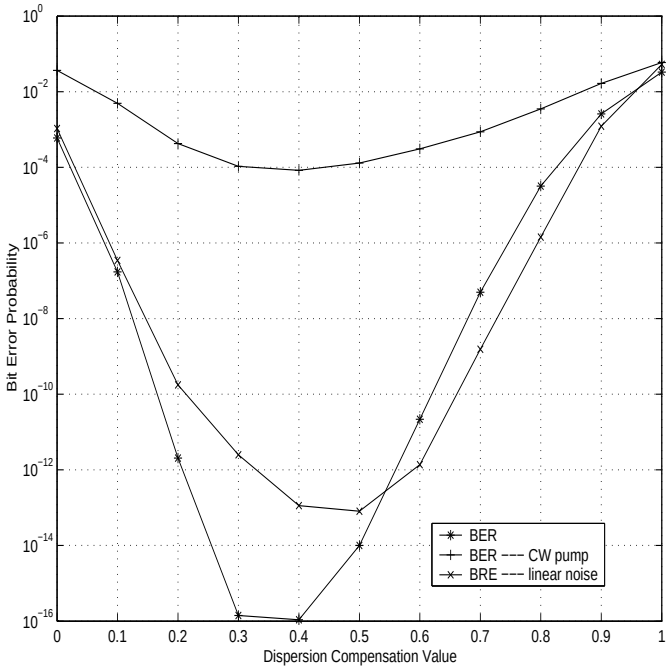


Figure 6.3: BER comparison with different noise models.

Due to the nonlinear interaction with the signal, the noise is expected to be amplified by the signal and the noise output with the full simulation model is expected to have larger variance than that from the linear noise model. Consequently, we expect that the BER performance from the full simulation model should be worse than that from the linear noise model. On the contrary, the plot indicates that the BER from the full simulation model is lower than the BER from the CW pump noise model for all normalized pre-DC values and is even lower than the BER from the linear noise model at low normalized pre-DC values. The reason for this behavior might be the noise squeezing where the in-phase part of the noise has reduced power after the transmission. However, the interaction between the noise

and the signal is so complicated that no analytical expression has yet been derived for the optimal normalized pre-DC value and the BER performance.

6.3 Robustness of Optimal Dispersion Compensation

In this section, we study the effects of different device parameters on the DC optimization using the noise full simulation model together with the 5-bit Gaussian approximation. The devices studied here include transmitters with different pulse shapes, receivers with different filter bandwidths and filter types, fibers with different dispersion, fibers with different uncompensated third-order dispersion and DCFs with and without nonlinearity.

6.3.1 Transmitter Pulse Shape

We first consider the effect from using transmitters with different pulse shapes. We have considered the following four models on the input pulse shapes: (1) Raised-Cosine; (2) Gaussian with $T_0 = 8.8$ps; (3) Gaussian with $T_0 = 7$ps; and (4) super-Gaussian of order 3 with $T_0 = 7$ps. The first two input pulse shapes are commonly used to represent the non-return-to-zero (NRZ) pulses while the last two input pulse shapes are commonly used to represent the return-to-zero (RZ) pulses. To give an idea on how the pulse shapes from these models differ in time, we show pulse trains with each pulse shape in Fig. 6.4. Note that we have chosen to use the same peak power of 5mW at the input for each pulse shape instead of the same average power. All other parameters are kept the same as those in Fig. 6.2 and we compare the BER performance as the normalized pre-DC varies in Fig. 6.5 for these four different pulse shapes. From the plot, we see significant differences in both the BER performance and the optimal dispersion compensation among the chosen pulse shapes. We believe that one major source of the differences stems from the different

bandwidths of these pulse shapes. Among the four pulse shapes, the Raised-Cosine has the smallest bandwidth, followed by Gaussian with $T_0 = 8.8$ps, Gaussian with $T_0 = 7$ps and super-Gaussian. The higher bandwidth the signal has, the higher ISI the signal experiences from filtering. In order to keep the total ISI down, the signal with higher bandwidth prefers a lower normalized pre-DC value to lower the ISI from the nonlinear interactions between the overlapping pulses. From the plot and the above discussion, we conclude that the optimization of DC must be studied in practice with careful attention to transmitter pulse shape. The plot also shows that the BER performance for the super-Gaussian pulse has multiple local minimum as the normalized pre-DC value varies. In order to explain this behavior, a thorough study should be conducted including the effects from average power, pulse width, rise time, extinction ratio and their interactions; this is beyond the scope of this chapter.

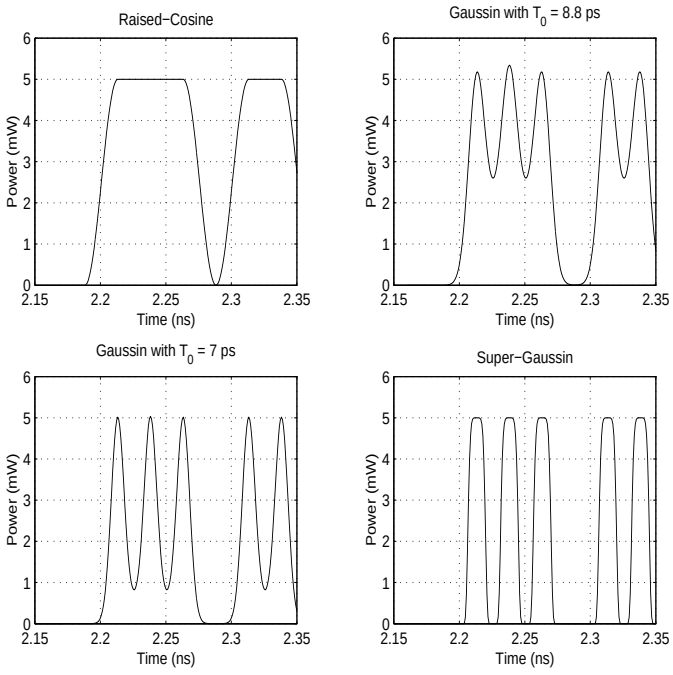


Figure 6.4: Pulse train in the time domain with different pulse shapes.

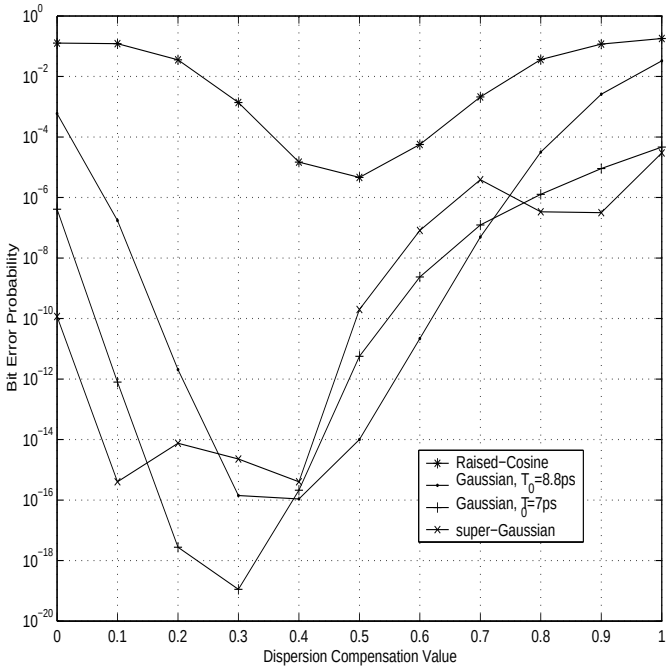


Figure 6.5: BER comparison with different input pulse shapes.

6.3.2 Receiver Design

We consider the effect from receivers with different filter bandwidths. As indicated in our system schematic shown in Fig. 6.1, a wide-band optical filter is used followed by a narrow-band electrical filter. The bandwidths of both the optical and electrical filters are important parameters in the system design. Generally, a narrow filter can cut more noise while introducing more ISI and signal power loss; a compromise must be made to achieve the optimum system performance. Filters with optimal bandwidths have been found to improve the system performance significantly for 10Gb/s systems [72]. Our focus here is on the robustness of DC optimization to the choice of filter bandwidths, so we compare the BER performance versus the normalized pre-DC value for different optical and electrical filter bandwidths in Fig. 6.6. In each of the three plots in Fig. 6.6, the optical filter band-

width varies from 30 GHz to 100 GHz while the electrical filter bandwidth is fixed at 26 GHz, 32 GHz, and 40 GHz for plots (A), (B), and (C) respectively. Both filters are modeled as fourth-order Butterworth filters as before and other parameters for the simulation are kept the same as those used in Fig. 6.2. From the comparison in the plots, it is clear that different optimal normalized pre-DC value is needed for different filter bandwidths. One preliminary guideline is that for a fixed electrical filter bandwidth, the optical filter with smaller bandwidth usually needs a smaller optimal normalized pre-DC value to lower the ISI contribution from the nonlinear interaction between overlapping pulses and thus keep the total ISI low. Note that the best performance of BER$< 10^{-18}$ is obtained using optical and electrical filters with bandwidths of 45 GHz and 26 GHz respectively.

The second receiver parameter we consider is the order of the optical filter, where the optical filter is assumed to be a Butterworth filter of varying orders with a fixed 3-dB bandwidth of 50GHz. Note that a Butterworth filter of first order is simply a Lorentz filter which is often used to model optical filters [52]. Fig. 6.7 shows the BER performance as the normalized pre-DC value varies for filters with orders varying from 1 to 5 and other parameters the same as before. From the plot, it is clear that the optimal pre-DC value is much more robust with respect to the filter order than with respect to the filter bandwidth. Even though we expect a best BER performance for the fifth-order filter since the higher the order is, the sharper transition the filter has and the more noise is cut by the filter, the second-order filter is found to give the best BER performance among the five filters. This indicates that the interaction between the filter response and the signal also plays an important role in the system performance. This is fortuitous since optical filters of higher order are difficult to design.

6.3.3 Fiber Dispersion Parameters

Fiber dispersion is an important effect in the system design and it is known to affect the system performance significantly. To give an idea of how fiber dispersion affects the optimal pre-DC value, we show in Fig. 6.8 the system performance vs. the normalized pre-DC value for $\beta_2 = \pm 2.6$, and ± 5ps^2/km. Other parameters are kept the same as before. From the plot, we find that the sign of the fiber dispersion determines the optimal pre-DC value; the systems with normal dispersion fibers with $\beta_2 > 0$ needs an optimal pre-DC value about 0.1 while the systems with abnormal dispersion fibers with $\beta_2 < 0$ needs an optimal pre-DC value about 0.4. The difference between the BER performance behavior for normal and abnormal fibers originates from the interaction between the fiber dispersion and the nonlinearity.

In all previous simulations, we have not included the effects from uncompensated third-order dispersion which might be important after full compensation of the second-order dispersion. The results are shown in Fig. 6.9 with several different third-order dispersion values, $\beta_3 = 0$, 0.066, 0.132 and 0.264ps^3/km respectively, with a fixed second-order dispersion of $\beta_2 = -2.6$ps^2/km. As the signal propagates from span to span, the uncompensated third-order dispersion accumulates and the signal pulse gets wider accordingly. As a result, a lower pre-DC value is expected with uncompensated third-order dispersion to lower the nonlinearity-induced ISI. From the plot, the uncompensated third-order dispersion is also found to degrade the system performance as a result of the pulse broadening.

6.3.4 Nonlinearity of DCF

The last device we consider is the dispersion compensator with and without nonlinearity. Two main types of dispersion compensators have been demonstrated experimentally. One of them is based on the design of optical filters with the correct phase properties and this type of dispersion compensator can be modeled as a linear device within the appropriate frequency range. DCF is another widely-used type of dispersion compensator. Even though the DCF is usually much shorter than the normal fiber, its smaller dimension enhances its nonlinearity which might be important to system design. We compare the BER performance as pre-DC value varies for several dispersion compensators with and without nonlinearity in Fig. 6.10. Same parameters as in Fig. 6.2 are used and no third-order dispersion is included. From the plot, the nonlinearity of the DCF not only lowers the optimal pre-DC value, but also degrades the system performance. The higher nonlinearity the DCF has, the lower the optimal pre-DC value is and the worse the system performance is.

6.4 Chapter Summary

In this chapter, we have studied the effects of different device parameters on the optimal DC for a 40Gb/s bi-end dispersion compensated optical fiber communication system up to 40 spans.

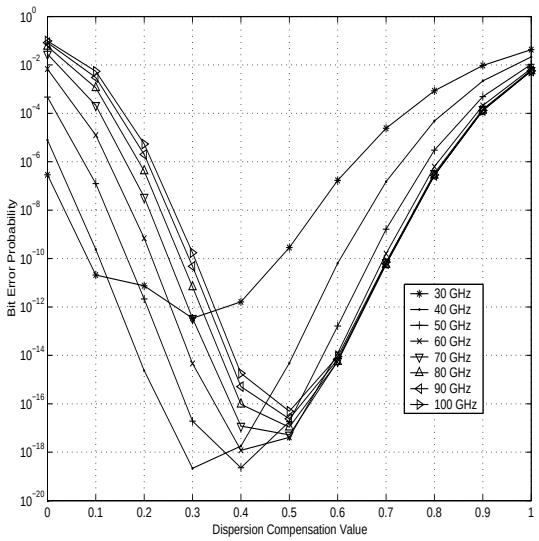
We start by checking the accuracy of BER computation based on Gaussian approximation with variable bit length where we have found that the conventional Q-factor is not accurate enough for a 40Gb/s system due to the large ISI associated with such systems. Instead, we use the bit-sequenced Gaussian approximation with 5-bits where the ISI from the 4 neighboring pulses are taken into account.

Different noise models are also checked in this chapter. Both the noise model neglecting noise nonlinear interaction with the signal and the noise model where the noise is propagating through the fiber with CW signal are found to be inaccurate for 40Gb/s systems even though they require much less computation time. Full simulation by the SSF method with multiple realization of the ASE noise is adopted in this work.

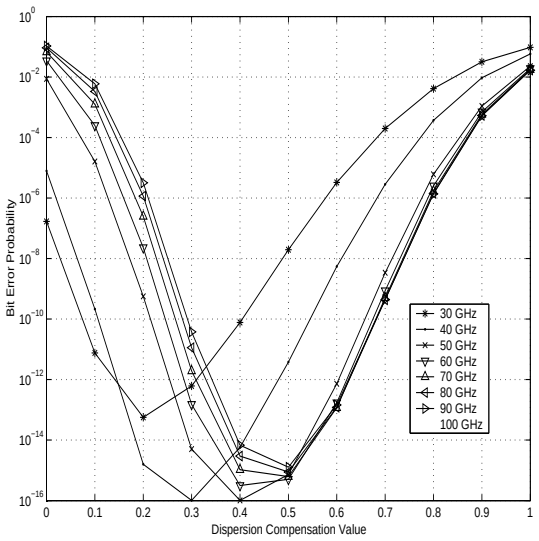
The robustness of optimal dispersion compensation is studied for several different device: (1) transmitters with different pulse shapes: (2) receivers with different optical and electrical filter bandwidth; (3) receivers with different optical filter orders; (4) fibers with different second- and third-order dispersion; (5) dispersion compensators with and without nonlinearity. In most cases, the optimal DC is not robust and varies significantly with device parameters. Consequently, one must be careful when extending the results from one study to another with different device parameters.

The study also shows some unexpected result. An optical filter with higher-order does not necessarily give the best system BER performance (at least for Gaussian pulses). The reason for this unexpected behavior need to be studied and the result might have significant impact on system design.

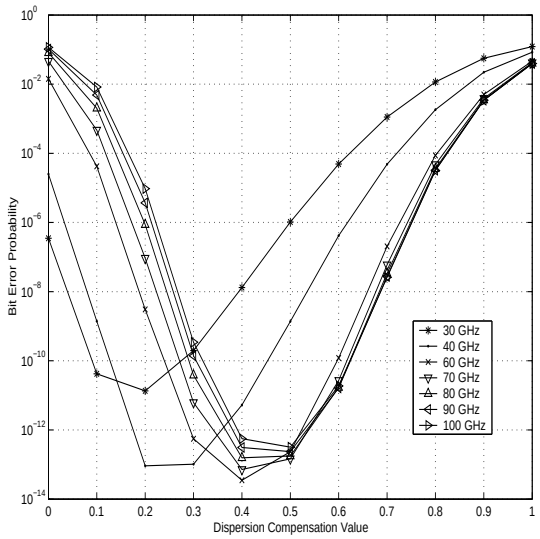
Numerical simulation with the SSF method and multiple realizations of amplifier ASE noise is used in this work even though it is very time-consuming and difficult to be applied to more complex systems. On the other hand, analytical studies available all have stringent limitations as the bit rate of the systems increases to 40Gb/s or higher. New analytical tools or extension of old analytical tools to systems with 40Gb/s or higher bit rate are essential for efficient system performance analysis and system design in the future as all optical networks become more and more complex.



A. Electrical filter bandwidth of 26 GHz.



B. Electrical filter bandwidth of 32 GHz.



C. Electrical filter bandwidth of 40 GHz.

Figure 6.6: BER comparison with different optical and electrical filter bandwidth.

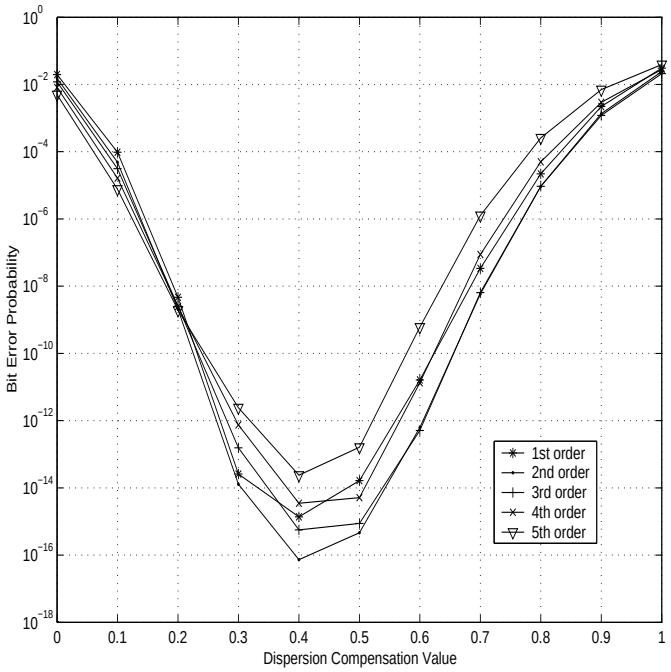


Figure 6.7: BER comparison with different optical filter models.

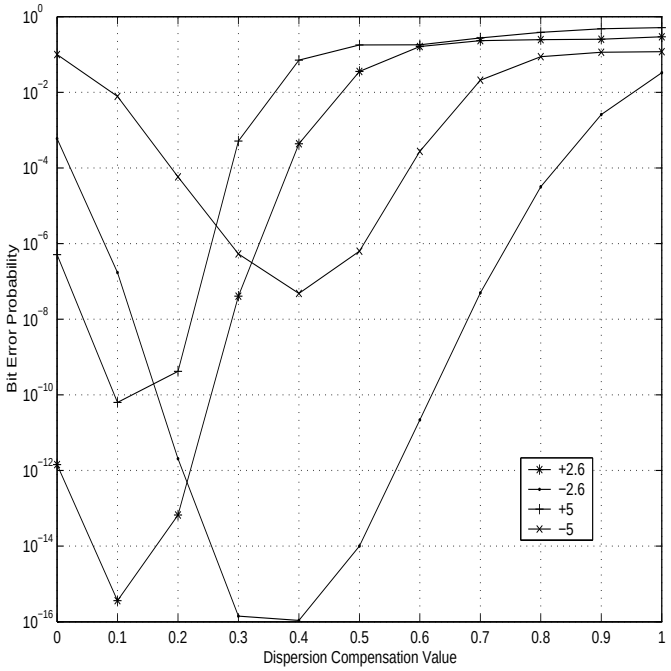


Figure 6.8: BER comparison with on different fiber second-order dispersion.

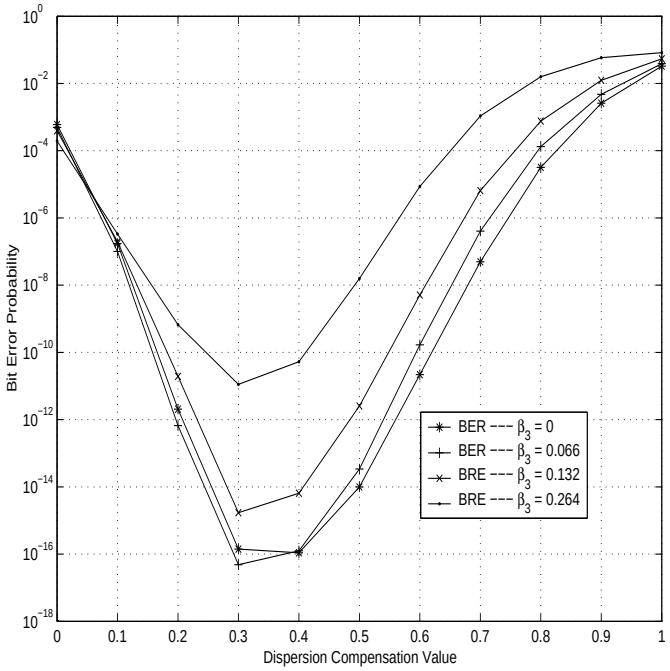


Figure 6.9: BER comparison with on uncompensated fiber third-order dispersion with different values.

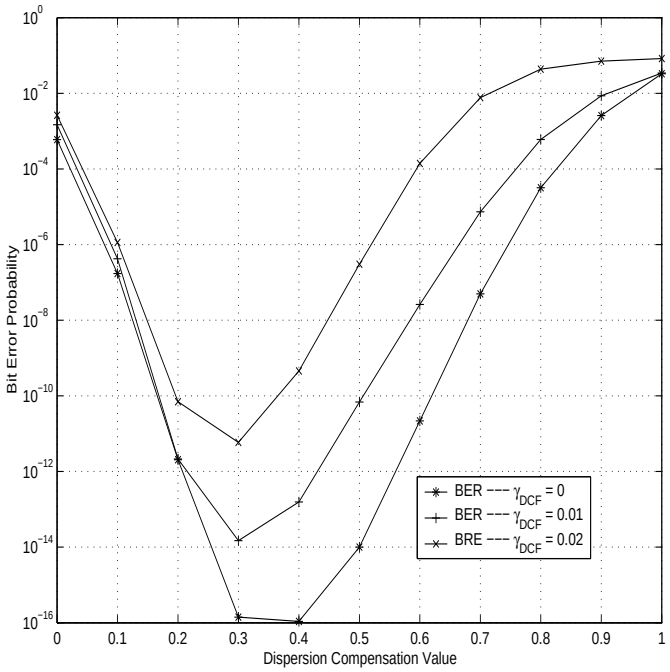


Figure 6.10: BER comparison with the nonlinearity coefficient from the DCF.

Chapter 7

Conclusions and Future Work

Fiber nonlinearity has been found to be one of the limiting factors for optical fiber communication systems as channel capacity requirement continues to grow. This dissertation contributes to understanding the effect of fiber nonlinearity on the performance of optical fiber communication systems. For this purpose, we have studied the following different nonlinear effects in detail:

- NLS equation and its analytical solutions;
- XPM- and FWM-induced intensity fluctuations in WDM systems;
- Application of multiuser detection in nonlinearity-limited WDM systems for better performance;
- Nonlinear interaction between signal and optical amplifier noise during co-propagation through fibers;
- Interaction between dispersion and nonlinearity for 40Gb/s systems and the optimization of dispersion compensation.

We summarize our results in Section 7.1 and then propose some future work in Section 7.2.

7.1 Summary

To efficiently study optical fiber communication systems, the fiber channel model based on the NLS equation must be correctly and efficiently solved. However, the complicated interaction between fiber dispersion and fiber nonlinearity make it a challenging task to solve the NLS equation. Despite the fact that the SSF method is the most widely-used tool in present studies on optical fiber systems, an analytical solution to the NLS equation is preferable when stochastic nature of the system is taken into account. To take advantage of the analytical nature of the truncated third-order VSTF method and keeping its computation efficiency while reducing its energy divergence, the modified VSTF method is proposed and used to study the signal propagation. Compared with the truncated third-order VSTF method, the modified VSTF method significantly reduces the energy divergence and thus extends the application range of such analytical tools.

WDM has been an efficient way to increase the system capacity with channels being more and more tightly packed. In such systems, XPM- and FWM-induced intensity fluctuation are known to be two limiting degradation sources. In this dissertation, we have applied the VSTF method to study thoroughly these two nonlinear effects, taking advantage of its efficiency on modeling stochastic system parameters including random information bits and random channel phases. Our work also extends previous theoretical works on XPM and FWM effects significantly. It leads to a better model for the study of the two nonlinear effects under which the two nonlinear effects can be correctly compared. The other important extension is in the analysis of the effects of XPM and FWM on complex optical networks where both networks with and without synchronization can be studied by the variance of the intensity fluctuations. Several methods to reduce the intensity fluctuation

are also discussed and validated in this dissertation.

As XPM and FWM effects introduce interference from channel to channel, the correlation between the channels also allows the possibility of achieving better error performance by using multiuser detection. Consequently, we propose the use of multiuser detection in WDM systems where the error performance of such systems is studied using a validated Gaussian approximation with lower- and upper-bounds. Significant differences are found between our proposed multiuser detector for optical WDM system and a conventional linear detector due to the presence of the photodetector which is a nonlinear device with square-law operation on the signal. Asymptotic behavior of the multiuser square-law detector is also carefully studied and compared with the linear detector case. Multiuser square-law detection is also applied to a practical WDM system with very narrow channel spacing and the multiuser detector shows promising performance that is impossible for a single-user detector with such narrow channel spacing.

Error performance of optical fiber communication systems has been studied since the advent of these systems with significant improvement made throughout these years. But most of the theoretical analyses on the error performance are based on the omission of the nonlinear interaction between the ASE noise and the modulated signal during their co-propagation through the fiber. In this dissertation, we have used perturbation theory to study this nonlinear interaction between the ASE noise and the signal. Unlike for previous studies which use the assumption of a CW signal, the output noise in our study is signal-dependent and non-stationary. In this case, we have found the noise frequency correlation functions using perturbation theory, keeping up to the second order in nonlinearity for computation simplicity. Our results are shown to be accurate wherever the perturbative

approach is applicable to solve the NLS equation. With the noise characteristics in hand, the detector statistics can be obtained by the Monte Carlo simulation. The results show PDF's that resemble χ^2 distribution, suggesting that the noise is near Gaussian before the photodetector.

As an alternative to more and more channels being used in WDM systems, we see significant improvement on the transmission capacity of a single channel with the optical fiber communication systems evolving from 10Gb/s to 40Gb/s and above per channel. At a bit rate as high as 40Gb/s, the effects from both fiber dispersion and fiber nonlinearity are pronounced due to the high power and large signal bandwidth used. As a result, the complicated interaction between the fiber dispersion and fiber nonlinearity must be carefully studied and the nonlinear interaction between pulses is found to be the limiting system degradation source. To study this nonlinear effect, previous studies have assumed different system and device parameters. We focus on the robustness of system performance and dispersion compensation optimization with respect to different device parameters in Chapter 6. The different devices we consider include transmitters with different pulse shapes; receivers with different optical filter and electrical filter bandwidth; receivers with optical filters of different orders; fibers with different second-order dispersion; fibers with different uncompensated third-order dispersion; and dispersion compensators with and without nonlinearity. First, we find that significant system performance can be achieved by optimizing the bi-end DC. Second, a significant difference is found for both system error performance and optimal dispersion compensation for different device parameters in most cases. The results indicate that the dispersion compensation must be carefully optimized using device parameters as close to the practical systems as possible.

With these separate studies on several different nonlinear effects, this dissertation has identified and studied the most important nonlinear effects for modern long-haul optical fiber communication systems. The goal of this dissertation is to find accurate and efficient tools for the system performance study instead of studying a specific system. The proposed analytical tools have all been carefully validated and numerous applications to specific systems have also been given as examples.

7.2 Future Work

Approximations have been used throughout the works in this dissertation to simplify the problems. For example, in the study of XPM- and FWM-induced intensity fluctuations, we have used a CW probe channel reduce the computational complexity which assumes that the probe channel stays at its peak power. Even though the VSTF method can be applied to the case when a modulated probe channel is needed, it is impractical for the computation of the fluctuation variance due to the multiple-dimension integration in the frequency domain. A computationally efficient way to do the frequency-domain integration is the key to solve this problem.

Another approximation used in this dissertation is the Q-factor and Gaussian approximation used in the study of multiuser square-law detection in Chapter 4 and the BER performance in Chapter 6. The Gaussian approximation is popular in research, but its accuracy has been questioned and more accurate techniques are always preferred. Recently, an extension of the K-L expansion with colored noise has been used to find the BER performance for optical fiber communication systems. To move forward, the colored, but still stationary noise should be replaced by the more realistic signal-dependent noise. On the other hand,

all known techniques to approximate the system BER performance are very complicated to use and have stringent limitations on the system channel models. Consequently, we need to compromise between accuracy and computational efficiency. One possible way is to replace the Gaussian approximation by a χ^2 distribution where the parameters for the χ^2 distribution are estimated based on the SSF simulation.

In this dissertation, we have studied several important nonlinear effects separately. Most of the works are done by the VSTF method or its equivalent, the perturbation method. All of these nonlinear effects contribute to the system performance degradation and should be considered together for a practical system. A systematic tool including each of these nonlinear effects should be developed.

From our study on the multiuser square-law detector, we see significant system performance improvement for WDM systems based on the Gaussian approximation. A similar study with the χ^2 distribution should also be conducted. Moreover, the Gaussian approximation can not be used to design the detector with the correct threshold or boundary. In order to apply the multiuser square-law detector in practical systems, a method to design the optimal or sub-optimal detector and study its performance under χ distribution is essential. Other than the nonlinearity-induced correlations, other sources of correlation should be included which might include the filtering leakage and correlated noise etc. Similar to the multiuser detector in a ICI-limited systems, a ISI-limited system can also benefit from a group detector where multiple bits are detected as a group.

As optical fiber communication systems shift from point-to-point long-haul transmission to more intelligent transport networks, many new problems need to be studied and answered. One especially important one is how to design an optical network efficiently.

The present most widely-used network design criteria is based on the worst-case scenario where the worst-case path is identified first and then designed to meet the specification requirement. However, this method is clearly inefficient if this path is only a very small portion of the whole network. To achieve better use of the scarce sources we have, better design criteria should be used while meeting the performance requirement for every channel. For this purpose, something similar to the channel capacity could be defined for the network as the network capacity, then we try to achieve this network capacity by allocating the channel wavelength and channel power etc. Other than the static network, adaptive network is another option which can usually make better use of the network sources than the static network. In this case, to find the worst-case path might itself be a challenging task.

Bibliography

[1] G. P. Agrawal, “Nonlinear Fiber Optics,” *Academic Press*, 1995.

[2] K. V. Peddanarappagari and M. Brandt-Pearce, “Volterra series transfer function of single-mode fibers,” *J. Lightwave Technol.*, vol. 15, pp. 2232–2241, 1997.

[3] M. Shtaif and M. Eiselt, “Analysis of intensity interference caused by cross-phase modulation in dispersive optical fibers,” *IEEE Photonics Technol. Lett.*, vol. 7, 979-981 (1998).

[4] G. Bellotti, M. Varani, C. Francia and A. Bononi, “Intensity distortion induced by cross-phase modulation and chromatic dispersion in optical-fiber transmissions with dispersion compensation,” *IEEE Photonics Technol. Lett.*, vol. 10, 1745-1747 (1998).

[5] R. Hui, K. R. Demarest and C. T. Allen, “Cross-phase modulation in multispan WDM optical fiber systems,” *J. Lightwave Technol.*, vol. 17, pp. 1018–1026, 1999.

[6] M. Eiselt, M. Shtaif, and L. D. Garrett, “Contribution of timing jitter and amplitude distortion to XPM system penalty in WDM systems,” *IEEE Photonics Technol. Lett.*, vol. 11, 748-750 (1999).

[7] R. I. Killey, H. J. Thiele, V. Mikhailov and P. Bayvel, “Prediction of transmission penalties due to cross-phase modulation in WDM systems using a simplified technique,” *IEEE Photonics Technol. Lett.*, vol. 12, 804–806 (2000).

[8] N. Shibata, R. P. Braun, and R. G. Waarts, “Phase-mismatch dependence of efficiency of wave generation through four-wave mixing in a single-mode optical fiber,” *J. Quantum Electron.*, QE-23, pp. 1205–1210, 1987.

[9] S. Song, C. T. Allen, K. R. Demarest, and R. Hui, “Intensity-dependent phase-matching effects on four-wave mixing in optical fibers,” *J. Lightwave Technol.*, vol. 17, pp. 2285–2290, 1999.

[10] M. Eiselt, “Limits on WDM systems due to four-wave mixing: a statistical approach,” *J. Lightwave Technol.*, vol. 17, pp. 2261–2267, 1999.

[11] S. Ten, K. M. Ennser, J. M. Grochocinski, S. P. Burtsev and V. L. daSilva, “Comparison of four-wave mixing and cross phase modulation penalties in dense WDM systems,” *OFC*, 1999, ThC4.

[12] J. H. Lee, “Analysis and characterization of fiber nonlinearities with deterministic and stochastic signal sources,” Ph. D. dissertation, Virginia Tech., Blacksburg, Virginia, 1999.

[13] Q. Zhang, “Performance analysis and design of WDM based optical communication systems using a Volterra series method,” Ph. D. dissertation, University of Virginia, Charlottesville, Virginia, 2001.

[14] K. V. Peddanarappagari and M. Brandt-Pearce, “Volterra series approach for optimizing fiber-optic communication system designs,” *J. Lightwave Technol.*, vol. 16, pp. 2046–2055, 1998.

[15] B. Xu and M. Brandt-Pearce, “Analysis of XPM-induced intensity distortion using the VSTF method,” *LEOS*, 2001, pp. 279–280.

[16] A. Sano, Y. Miyamoto, S. Kuwahara and H. Toba, “A 40-Gb/s/ch WDM transmission with SPM/XPM suppression through prechirping and dispersion management,” *J. Lightwave Technol.*, vol. 18, pp. 1519–1527, 2000.

[17] F. Matera, M. Settembre, M. Tamburrini, M. Zitelli and S. Turitsyn “Reduction of four wave mixing in optically amplified links by reducing pulse overlapping,” *Optics Comm.*, vol. 181, pp. 407–411, 2000.

[18] A. Okada, V. Curri, S. M. Gemelos and L. G. Kazovsky, “Reduction of four-wave mixing crosstalk using a novel hybrid WDM/TDM technique,” *Proceedings to ECOC 1998.*, pp. 289–290.

[19] A. Chraplyvy, A. Gnauck, R. Tkach, R. Derosier, C. Giles, B. Nyman, G. Ferguson, J. Sulhoff and J. Zyskind, “One-third Terabit/s transmission through 150 km of dispersion-managed fiber,” *IEEE Photonics Technol. Lett.*, vol. 7, 98-100 (1995).

[20] T. Chiang, N. Kagi, M, Marhic and L. Kazovsky, “Cross-phase modulation in fiber links with multiple optical amplifiers and dispersion compensators,” *J. Lightwave Technol.*, vol. 14, pp. 249–259, 1996.

[21] M. Hayee and A. Willner, “Pre- and post-compensation of dispersion and nonlinearities in 10-Gb/s WDM systems,” *IEEE Photonics Technol. Lett.*, vol. 9, 1271-1273 (1997).

[22] S. Wen and T. Lin, “Ultralong lightwave systems with incomplete dispersion compensations,” *J. Lightwave Technol.*, vol. 19, pp. 471–479, 2001.

[23] B. Dany, O, Leclerc, F. Neddam and P. Lourec, “Optimization of 40Gbit/s dispersion maps for long-haul WDM transmissions with up to 0.4bit/s/Hz spectral efficiency,” *OFC*, 2001, TuN5.

[24] K. Sekine, S. Sasaki and N, Kikuchi, “10 Gbit/s four-channel wavelength- and polarization-division multiplexing transmission over 340km with 0.5nm channel spacing,” *Electron. Lett.*, vol. 31, pp. 49–50, 1995.

[25] F. Forghieri, R. Tkach, A. Chraplyvy and D. Marcuse, “Reduction of four-wave mixing crosstalk in WDM systems using unequally spaced channels,” *IEEE Photonics Technol. Lett.*, vol. 6, 754-756 (1994).

[26] R. Tkach, A. Chraplyvy, F. Forghieri, A. Gnauck and R. Derosier, “Four-photon mixing and high-speed WDM systems,” *J. Lightwave Technol.*, vol. 13, pp. 841–849, 1995.

[27] H. Suzuki, S. Ohteru and N. Takachio, “22 $\times$ 10 Gb/s WDM transmission based on extended method of unequally spaced channel allocation around the zero-dispersion wavelength region,” *IEEE Photonics Technol. Lett.*, vol. 11, 1677-1679 (1999).

[28] S. Verdú, *Multiuser Detection*. Cambridge Univ. Press, 1999.

[29] A. Kaveie and J. M. F. Moura, “The Viterbi Algorithm and Markov Noise Memory,” *IEEE Trans. Information Theory*, vol. 46, 291–301 (2000).

[30] J. Moon and J. Park, “Pattern-Dependent Noise Prediction in Signal-Dependent Noise,” *IEEE J. Selected Areas in Commun.*, vol. 19, 730–742 (2001).

[31] E. E. Narimanov and P. Mitra, “The Channel Capacity of a Fiber Optics Communication System: Perturbation Theory,” *J. Lightwave Technol.*, vol. 20, pp. 530–537, 2002.

[32] B. Xu and M. Brandt-Pearce, “Modified Volterra series transfer function method and applications to fiber-optic communications,” *Conf. Record of the Asilomar Conf. on Signals, Systems and Computers*, 2001, pp. 23-27.

[33] B. Xu and M. Brandt-Pearce, “Modified Volterra series transfer function method,” *IEEE Photonics Technol. Lett.*, 2002, pp. 47-49.

[34] B. Xu and M. Brandt-Pearce, “Comparison of FWM- and XPM-induced crosstalk using the Volterra series transfer function method,” *J. Lightwave Technol.*, vol. 21, pp. 40-53, Jan. 2003.

[35] B. Xu and M. Brandt-Pearce, “Multiuser Detection for Asymmetric Gaussian and NonGaussian Statistics with Applications to Fiber Optic Communications,” *Proc. Conf. on Information Sciences and Systems*, 2002.

[36] B. Xu and M. Brandt-Pearce, “Multiuser Detection for WDM Fiber-Optic Communication Systems,” accepted for presentation in *ISIT*, 2002.

[37] R. J. Nuyts, Y. K. Park and P. Gallion, “Dispersion Equalization of a 10 Gb/s Repeated Transmission System Using Dispersion Compensating Fibers,” *J. Lightwave Technol.*, vol. 15, pp. 31–42, 1997.

[38] S. Wen, “Bi-end dispersion compensation for ultralong optical communication system,” *J. Lightwave Technol.*, vol. 17, pp. 792-798, May 1999.

[39] M. Florjanczyk and R. Tremblay, “RMS Width of Pulses in Nonlinear Dispersive Fibers: Pulses of Arbitrary Initial Form with Chirp,” *J. Lightwave Technol.*, vol. 13, pp. 1801–1806, 1995.

[40] Q. Yu and C. Fan, “Analytical Study on Pulse Broadening in Chained Optical Amplifier Systems,” *J. Lightwave Technol.*, vol. 15, pp. 444–451, 1997.

[41] M. Midrio, “Analytical Performance Evaluation of Nonreturn-to-Zero Transmission Systems Operating in Normally Dispersive Nonlinear Fibers,” *J. Lightwave Technol.*, vol. 15, pp. 2038–2050, 1997.

[42] T. Numai and O. Kubota, “Analysis of repeated unequally spaced channels for FDM lightwave systems,” *J. Lightwave Technol.*, vol. 18, pp. 656–664, 2000.

[43] K.-D. Chang, G.-C. Yang and W. Kwong, “Determination of FWM products in unequally-spaced channel WDM lightwave systems,” *J. Lightwave Technol.*, vol. 18, pp. 2113–2122, 2000.

[44] R. Killey, H. Thiele, V. Mikhailov and P. Bayvel, “Reduction of intrachannel nonlinear distortion in 40-Gbps-based WDM transmission over standard fiber,” *IEEE Photonics Technol. Lett.*, vol. 12, 1624–1626 (2000).

[45] M. Hayee and A. Willner, “NRZ versus RZ in 10–40-Gb/s dispersion-managed WDM transmission systems,” *IEEE Photonics Technol. Lett.*, vol. 11, 991–993 (1999).

[46] B. Konrad and K. Pertermann, “Optimum fiber dispersion in high-speed TDM systems,” *IEEE Photonics Technol. Lett.*, vol. 13, 299–301 (2001).

[47] M. K. Varanasi and B. AaZhang, “Optimally near-far resistant multiuser detection in differentially coherent synchronous channels,” *IEEE Trans. Information Theory*, vol. 37, 1006–1018 (1991).

[48] M. K. Varanasi, “Noncoherent detection in asynchronous multiuser channels,” *IEEE Trans. Information Theory*, vol. 39, 157–176 (1993).

[49] M. K. Varanasi and A. Russ, “Noncoherent decorrelative detection for nonorthogonal multipulse modulation over the multiuser Gaussian channel,” *IEEE Trans. Commun.*, vol. 46, 1675–1684 (1998).

[50] M. K. Varanasi and D. Das, “Noncoherent decision-feedback multiuser detection,” *IEEE Trans. Commun.*, vol. 48, 259–269 (2000).

[51] A. Russ and M. K. Varanasi, “Noncoherent multiuser detection for nonlinear modulation over the Rayleigh-fading channel,” *IEEE Trans. Information Theory*, vol. 47, 295–306 (2001).

[52] E. Iannone, F. Matera, A. Mecozzi and M. Settembre, *Nonlinear Optical Communication Networks*, John Wiley & Sons, 1998.

[53] P. A. Humblet and M. Azizoglu, "On the bit Error rate of lightwave systems with optical amplifiers," *J. Lightwave Technol.*, vol. 11, pp. 1576–1582, 1991.

[54] M. Schwartz, W. Bennett and S. Stein, *Communication Systems and Techniques*. New York: McGraw-Hill, 1966.

[55] I. T. Monroy and G. Einarsson, "Bit error evaluation of optially preamplified direct detection receivers with Fabry-Perot optical filters," *IEEE J. Lightwave Technol.*, vol. 15, 1546-1553 (1997).

[56] R. Hui, D. Chowdhury, M. Newhouse, M. O'Sullivan, and M. Poettcker, "Nonlinear amplification of noise in fibers with dispersion and its impact in optically amplified systems," *IEEE Photonics Technol. Lett.*, vol. 9, 392-394 (1997).

[57] A. Carena, V. Curri, R. Gaudino, P. Poggiolini and S. Benedetto, "New analytical results on fiber parametric gain and its effects on ASE noise," *IEEE Photonics Technol. Lett.*, vol. 9, 535-537 (1997).

[58] G. Bosco, A. Carena, V. Curri, R. Gaudino, P. Poggiolini, and S. Benedetto, "A novel analytical approach to the evaluation of the impact of fiber parametric gain on the bit error rate," *IEEE Trans. on Communications* , vol. 49, 2154-2163 (2001).

[59] G. Bosco, A. Carena, V. Curri, R. Gaudino, P. Poggiolini and S. Benedetto, "Parametric gain in multiwavelength systems: a new approach to noise enhancement analysis," *IEEE Photonics Technol. Lett.*, vol. 12, 152-154 (2000).

[60] A. Carena, V. Curri, R. Gaudino, P. Poggiolini and S. Benedetto, "On the joint effects of fiber parametric gain and birefringence and their influence on ASE noise," *IEEE J. Lightwave Technol.*, vol. 16, 1149-1157 (1998).

[61] B. Xu and M. Brandt-Pearce, “Analysis of nonlinear amplification of noise in optical fibers by Volterra series method,” *Proc. Conf. on Information Sciences and Systems*, 2001.

[62] C. K. Madsen, and G. Lenz, “Optical all-pass filters for phase response design with applications for dispersion compensation,” *IEEE Photonics Technol. Lett.*, vol. 10, 994-996, Jul. 1998.

[63] A. H. Gnauck, J. M. Wiesenfeld, L. D. Garret, M. Eiselt, F. Forghieri, L. Arcangeli, B. Agogliata, V. Gusmeroli, and D. Scarano, “16 $\times$ 20-Gb/s, 400-km WDM transmission over NZDSF using a slope-compensating fiber-grating module,” *IEEE Photonics Technol. Lett.*, vol. 12, 437-439. Apr. 2000.

[64] L. Gruner-Nielsen, T. Veng, S. N. Knudsen, C. C. larsen, and B. Edvold, “New dispersion compensating fibers for simultaneous compensation of dispersion and dispersion slope of non-zero dispersion shifted fibers in the C or L band,” *OFC*, 2000, TuG6, pp. 101-103.

[65] T. Schafer, E. W. Laedke, M. Gunkel, C. Karle, A. Posth, K. H. Spatschek, and S. K. Turitsyn, “Optimization of dispersion-managed optical fiber lines,” *J. Lightwave Technol.*, vol. 20, pp. 946-952, June 2002.

[66] E. Forestieri, “Evaluating the error probability in lightwave systems with chromatic dispersion, arbitrary pulse shape and pre- and postdetection filtering,” *J. Lightwave Technol.*, vol. 18, pp. 1493-1503, Nov. 2000.

[67] M. Zitelli, F. Matera, and M. Settembre, “Single-channel transmission in dispersion managed links in conditions of very strong pulse broadening: application to 40 Gb/s signals on step-index fibers,” *J. Lightwave Technol.*, vol. 17, pp. 2498-2505, Dec. 1999.

[68] T. L. Lakoba, and G. P. Agrawal, “Optimization of the average-dispersion range for long-haul dispersion-managed soliton systems,” *J. Lightwave Technol.*, vol. 18, pp. 1504-1512, Nov. 2000.

[69] A. Naka and S. Saito, “Transmission distance of in-line amplifier systems with group-velocity-dispersion compensation,” *J. Lightwave Technol.*, vol. 13, pp. 862-867, May 1995.

[70] C. M. Weinert, R. Ludwig, W. Pieper, H. G. Weber, D. Breuer, K. Petermann and F. Kuppers, “40 Gb/s and 4 $\times$ 40 Gb/s TDM/WDM standard fiber transmission,” *J. Lightwave Technol.*, vol. 17, pp. 2276-2284, Nov. 1999.

[71] J-P Laude, *DWDM Fundeamentals, Components, and Applications*, Artech House, 2002.

[72] E. W. Laedke, N. Goder, T. Schaefer, K. H. Spatschek, and S. Turitsyn, “Improvement of optical fiber systems performance by optimization of receiver filter bandwidth and use of numerical methods to evaluate Q-factor,” *Electron. Lett*, vol. 35, no. 24, 2131-2133, Apr. 1999.

[73] B. Xu and M. Brandt-Pearce, “Analysis on noise amplification by a CW pump signal due to fiber nonlinearity,” has been submitted to *IEEE Photonics Technol. Lett.*, May, 2003.