CHALMERS

Stochastic Backpropagation for Coherent Optical Communication

Master's Thesis in communication systems and information theory

YAN GONG
NAN JIANG

Department of Signals & Systems
Communication Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2011
EX081/2011

Abstract

We present stochastic backpropagation, a novel maximum a posteriori detection method for coherent optical communications with dual polarization and Multilevel quadrature amplitude modulation formats (M-QAM) transmission. The proposed detector is shown to outperform conventional backpropagation in a scenario where nonlinear phase noise is the dominant impairment and polarization dependent loss (PDL) channels are existed. We also provide a novel channel estimation method for unitary channel effect and a ML-based channel estimation method for PDL channel.

Acknowledgements

This thesis work has been performed at Chalmers in Gothenburg, Sweden, under the supervision of Assistant Professor Henk Wymeersch. We would like to express our gratitude to people who has been involved to sponsor and help us to complete our work.

First we would like to thank Assistant Professor Henk Wymeersch, who has guided us with his brilliant advices and kept pushing us to further study. We would also like to thank PhD student Johnny Karout, who helped us a lot to solve all kinds of problems in this thesis. Thanks to the help of Henk and Johnny, we released a paper in ECOC during the thesis study and we are very proud of working with them. Another important person we would like to thank is master student Mehrnaz Tavan, who shared her research achievement generously with us and helped us a lot in our later study.

Last but not least, we would like to thank all the friends of ours, who helped us a lot not only in thesis study but also during the 2 years master study. We shared a lot of good time together here in Chalmers, thanks to all of you.

Yan Gong and Nan Jiang, Gothenburg 2011/9/11

Contents

1

Introduction

Multilevel quadrature amplitude modulation formats (M-QAM) is widely used in the coherent optical communication due to its ability to increase the spectral efficiency comparing to the traditional optical communication. However, as the minimum distance decreases on the constellation, M-QAM is more sensitive to the noise and requires more power to reach a certain bit-error-rate (BER). Some other characters of optical system will also affect the results during the transmission, like high power will lead to increasing intrachannel four-wave mixing which causes the nonlinear inter-symbol interference (ISI), and self-phase modulation (SPM) [1].

This thesis is aimed to build a novel detector to solve both SPM noise and PDL channel effect, the detector is built based on some techniques such as back propagation, factor graph (FG) and Monte Carlo technique. In this thesis, we focus on a dual polarization 16-QAM system. The key components will be studied separately before the whole system is taken into consider.

1.1 Coherent optical communication

Coherent technologies have been as a hot topic as the transmission-capacity increases in wavelength-division multiplexed (WDM) systems [2]. In coherent optical communication, signal is encoded onto the electrical filed of the light wave. Polarization is one of the main issues may take into consideration. A polarized signal is a wave which all the electric components are in a fixed phase relationship. So phase and polarization are the key obstacles for the implementation of the coherent receivers.

1.1.1 Nonlinear impairments

Both linear and nonlinear impairments are contained in optical communication model. Linear impairments include chromatic dispersion (CD) and polarization mode dispersion (PMD). Nonlinear impairments originate from Kerr effect, which is named after the

Scottish Physicist John Kerr. It refers to the phenomenon that the refractive index of the optical fiber varies as the launched power changes. Nonlinear impairments include self phase modulation (SPM), cross phase modulation (XPM) and four wave mixing (FWM).

Self-phase modulation (SPM) is interesting to study since it only can be partially compensated while nonlinear inter-symbol interference (ISI) can be compensated for, as shown in [3]. SPM causes the signal to change its own phase while travelling through the fiber. In combination with amplified spontaneous emission (ASE) noise, SPM leads to nonlinear phase noise (NLPN) [4], which in signal space can be viewed as symbols elongated in the phase direction [1]. In this thesis work, the chromatic dispersion (CD) is neglected so that NLPN is the dominant impairment for the communication system [1, 5, 6].

An experimental investigation has showed that NLPN is a limiting factor [7] and recent results indicate that NLPN is important up to 40 Gbaud [8]. To mitigate the effect caused by NPLN, compensation have to be made at the receiver side. A stochastic back propagation method based on factor graph and Monte carlo technique is studied in this thesis to achieve the goal.

1.1.2 Memoryless linear impairments

Polarization dependent loss (PDL) channel is another dominant component which limits the system performance. Polarization dependent loss is the ratio of the maximum and the minimum power amplitude to the all polarization states. It is usually caused by circular dichrosim, fiber bending, angled optical interfaces or oblique reflection [9].

Each optical device exhibits a polarization dependent transmission, the polarization changes randomly over each fiber span. Generally, the PDL of each span cannot be determined by simply adding all the PDL components together in this case which is the worst situation. The total PDL depends on the polarization transformation of each span. The BER of the system is highly affected by these. PDL is also the main source of pulse distortion [10].

In fiber optical transmission, the devices have PDL are fibers, optical couplers, isolators, WDM and photonic-detectors. The PDL of each component are different, it may depend on the input wavelength of the source. The polarization along the fiber is unpredictable and uncontrolled which will lead to the decrease of the quality of fiber optical transmission. Therefore, the systems always require the device with the low PDL.

Consequently, the measurement of PDL has become a hot topic and attracted lots of attention from manufacturers and researchers. The perfect estimation can highly improve the quality of transmission.

Erbium-doped fiber amplifier (EDFA) is the most common amplifier that is used in the fiber optical communication. EDFAs use a doped optical fiber as a gain medium to amplify an optical signal. EDFAs have two major pumping bands 980 nm and 1480 nm. The 980 nm band has a higher absorption cross-section and is used in the low noise situation, on the contrary, the 1480 nm band has a lower absorption cross-section and used when high power is required.

Additive white Gaussian noise (AWGN) is an ideal channel model in which the added noise is wideband with constant spectral density and a Gaussian distribution of amplitude. The AWGN channel is a good model deep space communication links. It is not good for most terrestrial links because of multipath or interference. But AWGN is commonly used to simulate background noise of the channel.

1.2 Thesis structure

After the introduction, we will give the system model we studied in this thesis, also with the abstract of previously research achievements of coherent optical communication. The details of factor graph and Monte Carlo technique are also contained in chapter 2.

Chapter 3 includes most parts of achievements in this thesis. We divided the whole system into several key components and solve the problems separately. In the end, a detector designed for the entire system will be built.

We present the results in chapter 4, analyze the performance by comparing the results with some known methods and different simulation parameters.

Chapter 5 is the conclusion of this thesis study, followed by chapter 6, the future work.

2

Literature Review

In this chapter we will give all the basic knowledge for our study, including system model, channel character, introduction of backpropagation, factor graph and Monte Carlo technique, along with some previously research achievements of SPM and PDL channel.

2.1 System model

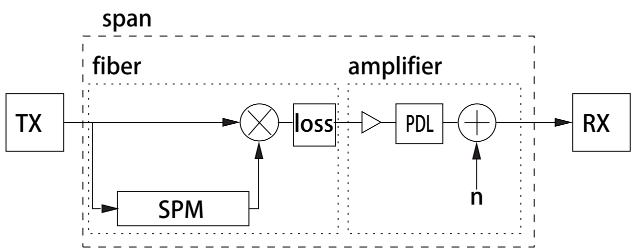

Figure 2.1: Optical transmission model

We consider a discrete-time multi-span polarization multiplexed coherent optical communication system with optical dispersion compensation. The system model is given in Figure 2.1. The long-haul transmission system contains many spans like this. At each symbol period, we transmit a two-dimensional complex data vector $\mathbf{a}$, $\mathbf{a}$ is drawn uniformly from a constellation Ω^2 with average energy proportional to the input power P_{in} per polarization.

Optical communication system consists of a single mode fiber (SMF) followed by a dispersion compensating fiber (DCF) and an amplifier in each span. SMF and DCF have two Kerr nonlinear parameters and two attenuation factors α_{smf}, α_{dcf}. The DCF is assumed to ideally compensate for the chromatic dispersion [1]. The signal power is attenuated by $e^{-\xi} = e^{-(\alpha_{\text{smf}} L_{\text{smf}} + \alpha_{\text{dcf}} L_{\text{dcf}})}$ where L_{smf} and L_{dcf} are the lengths of SMF and DCF. The amplifier has a power gain $G = e^{-\xi}$, in order to restore the signal power to the levels in the transmitter side. The amplifier generates an amplified spontaneous emission (ASE) noise, which can be modeled as an additive white Gaussian noise (AWGN) process. The PDL channel exists between SPM and AWGN.

SPM, PDL channel and AWGN are three major components in this system. As we are familiar to deal with additive white Gaussian noise, we focus our study on SPM and PDL channel. To make the problem more clear, we isolate the nonlinear part and the channel, then discuss them separately.

2.1.1 Non-linear optical communication system

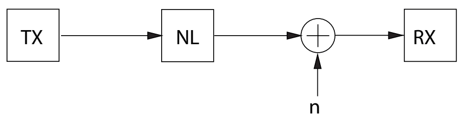

Figure 2.2: Simplified transmission model for 1 span

If we only take SPM and AWGN into account, we can redraw the transmission system for one span like Figure 2.2. The NL block represent the SPM effect of the fiber and n is the ASE noise. The discrete-time signal at the output of the i-th span will be given by

$$\mathbf{r}_i(t) = \mathbf{r}_{i-1}(t) \exp\left(j\gamma L_{\text{eff}} \left\|\mathbf{r}_{i-1}(t)\right\|^2 \right) + \mathbf{n}_i(t), \tag{2.1}$$

where $\mathbf{r}_0(t) = \mathbf{s}(t)$. For clarity, SPM is neglected in the DCF[1], so $\gamma = \gamma_{\text{smf}}$ as γ_{smf} is the nonlinearity parameter of SMF, L_{eff} is the effective length of the fiber, which is written as $L_{\text{eff}} = \left(1 - e^{-\alpha L}\right)/\alpha$, in which attenuation factor $\alpha = \alpha_{\text{smf}}$ and length $L = L_{\text{smf}}$, the operator $\|\cdot\|$ represents the norm given by $\|\mathbf{x}\| := \sqrt{\mathbf{x}^{\text{H}}\mathbf{x}}$ and $\mathbf{n}_i(t)$ is ASE noise, which is modeled as a zero-mean and has power spectral density of N_0 per polarization, the noise is bandlimited to a bandwidth B.

2.1.2 PDL channel

Polarization dependent loss is defined as:

$$\text{PDL}_{\text{dB}} = 10 \times \log(\frac{\text{P}_{\text{Max}}}{\text{P}_{\text{Min}}}), \tag{2.2}$$

[1]The input power to the DCF is low, so SPM is neglected.

In practical PDL channel is defined as:

$$\mathrm{H_{PDL}} = \mathbf{U} \times \begin{bmatrix} 1 & 0 \\ 0 & \tilde{\gamma} \end{bmatrix} \times \mathbf{U}^{\mathrm{H}}, \tag{2.3}$$

where $\mathbf{U}$ is a random unitary matrix[2], $\tilde{\gamma}$[3] is the polarization dependent attenuation factor, it is in the range of 0 to 1, usually around 1.

This channel has several properties. The $\tilde{\gamma}$ matrix is a diagonal matrix. When $\tilde{\gamma}$ equals to 1, the PDL channel will no longer exist, as unitary matrix times its hermitian is an identity matrix. This can be used to check if the channel estimation works well. Another property is that $\tilde{\gamma}$ cannot be 0, in this thesis, we focus on the dual polarization, there are always two lines of signals transporting together, if $\tilde{\gamma}$ is zero, there will be unexpected mistake.

The PDL channel is added in each span of the fiber. Our task is to estimate and equalize the channel through numbers of spans with the nonlinear phase noise effect. In our study, $\tilde{\gamma}$ is considered as a constant for each span which value is near 1.

2.2 Previously research

As the new technology like erbium-doped fiber amplifier (EDFA), WDM and fiber non-linearity management breakthrough, the capacity of the fiber has been growing fast in past years. Meanwhile, the spectral efficiency for optical communication has also been increasing.

In order to achieve high spectral efficiency, the study of coherent optical communications has attracted the attention from all over the world. The coherent optical receivers have following advantages: by using phase modulation, the multilevel modulation format can be introduced; the receiver sensitivity has improved since the ability of the phase detection increased; the closely spaced WDM channels can be separated at the electrical level as the frequency resolution at baseband is high; the local oscillator power can help to achieve the shot-noise limited receiver sensitivity.

These are all attribute to the rapid growth of high-speed digital signal processing (DSP). By using DSP, both the phase and polarization management can be done.

2.2.1 Previously research of SPM

Many studies have been done to analyze the effect of nonlinear phase noise and different methods have been brought to mitigate the effect of NLPN at the receiver, including different kind of compensation technique, MAP and ML detection and backpropagation. Here we give a brief review of several typical methods.

[2] A unitary matrix has the character that $\mathbf{U} \times \mathbf{U}^{\mathrm{H}} = \mathbf{U}^{\mathrm{H}} \times \mathbf{U} = I$, where I is an identity matrix and $\mathbf{U}^{\mathrm{H}}$ is the Hermitian adjoint of $\mathbf{U}$, both $\mathbf{U}$ and $\mathbf{U}^{\mathrm{H}}$ are square matrix

[3] For clarity, in this thesis γ represents the nonlinear parameter and $\tilde{\gamma}$ represents the PDL factor

Electronic compensation technique

Optimal compensation using electronic circuits to mitigate NLPN for binary phase-shift keying (BPSK) and differential quadrature phase-shift keying (DQPSK) was proposed in [5]. In their study, the overall nonlinear phase shift is equal to:

$$\Phi_{NL} = \gamma L_{\text{eff}}\{|E_0 + \text{n}_1|^2 + |E_0 + \text{n}_1 + \text{n}_2|^2 + \cdots + |E_0 + \text{n}_1 + \text{n}_2 + \cdots + \text{n}_\text{N}|^2\}, \quad (2.4)$$

where E_0 is the transmitted signal, n_k is the complex amplifier noise at the k_{th} span, γ is the nonlinear coefficient of the fiber and L_{eff} is the effective length of the fiber. They use a an optical phase-locked loop (PLL) in PSK system and a pair of interferometers in DPSK to receive the in-phase and quadrature components of the received electric field. They derived optimal compensator by finding a scale factor β which can minimize the variance of the residual nonlinear phase shift $\Phi_{NL} + \beta P_N$. The corrected phase estimate is $\Phi_R - \beta P_N$, where Φ_R is the phase of the received electric field E_R [5]. The optimal scale factor is found to be

$$\beta = -\gamma L_{\text{eff}} \frac{N+1}{2} \cdot \frac{|E_0|^2 + (2N+1)\sigma^2/3}{|E_0|^2 + N\sigma^2} \approx -\gamma L_{\text{eff}} \frac{N+1}{2}. \quad (2.5)$$

In their research, optimal compensation can halve the standard deviation (STD) of the nonlinear phase noise, doubling the transmission distance in systems whose dominant impairment is nonlinear phase noise [5]. Their method was extended to M-ary phase-shift keying in [6].

Close-form detector

In [1], a close-form detector was derived for a polarization-multiplexed M-QAM system with discrete amplification, limited by NLPN. The detector is based on maximum likelihood (ML) detection. They use the same system model as we mentioned in 2.1.1 and get the receive signal for N_a span by recursive calculation:

$$\mathbf{r}_k = \mathbf{s}_k \exp\left(j\gamma L_{\text{eff}} \sum_{i=0}^{N_\text{a}-1} \|\mathbf{r}_{i,k}\|^2\right) + \mathbf{w}_k, \quad (2.6)$$

where $\mathbf{s}_k = \sqrt{P_{\text{in}}}\mathbf{a}_k$ (assuming negligible inter-symbol interference), $\mathbf{w}_k \sim \mathcal{CN}(\mathbf{0}, N_\text{a}\sigma^2_{\text{ASE}}\mathbf{I})$, and $\sigma^2_{\text{ASE}} = BN_0$ [1]. They derived the approximate likelihood function and used it as a detector, which is given by

$$\hat{\mathbf{a}}_k = \arg\max_{\mathbf{a}_k \in \Omega^2} \; p(\mathbf{r}_k|\mathbf{a}_k) \quad (2.7)$$

$$= \arg\max_{\mathbf{a}_k \in \Omega^2} \; \exp\left(-\frac{\|\mathbf{s}_k\|^2}{N_\text{a}\sigma^2_{\text{ASE}}}\right) \frac{\text{I}_0(|\beta_k|)}{\text{I}_0(1/\sigma^2_\psi)}. \quad (2.8)$$

where $\text{I}_0(.)$ is the zeroth order modified Bessel function of the first kind, and

$$\beta_k = \frac{2\mathbf{r}_k^\text{H}\mathbf{s}_k e^{\left(j\gamma L_{\text{eff}} N_\text{a} \|\mathbf{s}_k\|^2 + j\bar{\phi}\right)}}{N_\text{a}\sigma^2_{\text{ASE}}} + \frac{1}{\sigma^2_\psi}. \quad (2.9)$$

where

$$\bar{\phi} = \gamma L_{\text{eff}} \sigma^2_{\text{ASE}} (N_{\text{a}} - 1) N_{\text{a}}/2, \tag{2.10}$$

and

$$\sigma^2_{\psi} = 2\gamma^2 L^2_{\text{eff}} \|\mathbf{s}_k\|^2 \sigma^2_{\text{ASE}} (N_{\text{a}} - 1) N_{\text{a}} (2N_{\text{a}} - 1)/6, \tag{2.11}$$

In their research, the closed-form detector provides the best complexity-performance trade-off compared with a number of suboptimal detectors and a complex non-parametric detector.

Adaptive MAP detection

An adaptive maximum a posteriori (MAP) detection scheme was investigated in [11]. They studied the performance of a MAP detection scheme with a look-up table for non-linear inter-symbol interference compensation and showed that MAP detection can help to increase spectral efficiency and improve nonlinearity tolerance in long-haul transmissions [11].

Digital backpropagation

Backpropagation was proposed as a universal technique for jointly compensating linear and nonlinear impairments in [12] and [13]. The use of digital backpropagation in conjunction with coherent detection to mitigate fiber nonlinearity and dispersion jointly is studied in [3]. They proposed a solution of the inverse nonlinear Schröinger equation (NLSE) by using noniterative asymmetric split-step Fourier method (SSFM). BP has been shown to enable higher launched power and longer system reach in dense wavelength-division-multiplexed (WDM) transmission over zero-dispersion fiber [12].

Backpropagation learning algorithm contains two parts: propagation and weight update. Each propagation involves forward propagation and backward propagation. In this thesis, forward propagation is the transmission process and the backward propagation is the detection process. By using backpropagation, we can detect the original transmitted signals from the receive signals. To achieve this goal, we need to analyze the transmission system and calculate the update weight. This is done by using factor graph and Monte Carlo techniques.

In this thesis, we compare our detector with a backpropagation detector. The backpropagation detector only considers the deterministic part of the nonlinearity while ignore the effects in nonlinear part caused by AWGN noise. The system model is $\mathbf{r} = \mathbf{s} \exp(j\gamma L_{\text{eff}} N_{\text{a}} \|\mathbf{s}\|^2) + \mathbf{n}$, the detector is given by

$$\hat{\mathbf{a}}_{\text{k}} = \arg \min_{\mathbf{a}_{\text{k}} \in \Omega^2} \left\| \mathbf{r}_{\text{k}} - \mathbf{s}_{\text{k}} \exp(j\gamma L_{\text{eff}} N_{\text{a}} \|\mathbf{s}_{\text{k}}\|^2) \right\|^2 . \tag{2.12}$$

2.2.2 Previously research of PDL channel

The research and development in optical fiber communication started around 1970s [2]. There are two basic methods to measure PDL of the device, the polarization scanning technique and the Mueller method. At some specific wavelengths polarization scanning technique is more suitable. For a wavelength range, the Mueller method gives a better performance.

Polarization scanning technique

The polarization scanning technique is a non-deterministic method based on the measurement of the maximum and minimum power amplitude. The device under test (DUT) is exposed to all stated of polarization state. The polarization is controlled by a polarization controller that can be deterministically or randomly. Then the maximum and minimum is measured over time. And the PDL is calculated by its definition. But we cannot tell when there is a change, whether it caused by the PDL of the DUT or the input power of the device.

In order to get a precise measurement, the input power must be stable, the source must have high degree of polarization, the loss in the polarization controller must be small, also the detector must have low polarization dependent responsitivity, as these will largely influence the detection of the PDL.

In practical, the errors from the scanning time and the measuring time are limited, which means the longer the polarization scanning takes, the more accurate the measurement of PDL is.

Besides setting the scanning time, it is important to set a correct scanning rate as well. The faster scanning rate will generate more polarization state then the measuring time will decrease. But if the rate is too fast, the results may get worse.

Mueller method

Mueller method is a different approach from polarization scanning technique, as it is a deterministic method and measures the PDL of a DUT from its Mueller matrix.

Mueller method only exposes the source to four polarization states, linear horizontal polarized (LHP), linear vertical polarized (LVP), linear +45 degrees (L +45), right hand circular (RHC). PDL is calculated from the information obtained by these states.

The Mueller matrix indicates the polarization and power properties of the DUT. The relationship between input Stokes vector, output Stokes vectors and the Mueller matrix is

$$\mathrm{S_{out}} = \mathrm{M} \times \mathrm{S_{in}}, \tag{2.13}$$

The Mueller matrix is a 4 by 4 matrix. The first row coefficients of the Mueller matrix describe the total optical power are required in calculating the PDL of the DUT.

$$\mathrm{S} = \mathrm{m_{11}S_{0in}} + \mathrm{m_{12}S_{1in}} + \mathrm{m_{13}S_{2in}} + \mathrm{m_{14}S_{3in}}, \tag{2.14}$$

where $m_{11}...m_{14}$ are the first row coefficients, and input Stokes Vectors are:

$$\mathrm{LHP}: S_{0in} = \begin{bmatrix} P_{in1} \\ P_{in1} \\ 0 \\ 0 \end{bmatrix} \mathrm{LVH}: S_{1in} = \begin{bmatrix} P_{in2} \\ -P_{in2} \\ 0 \\ 0 \end{bmatrix} \mathrm{L+45}: S_{2in} = \begin{bmatrix} P_{in3} \\ 0 \\ P_{in3} \\ 0 \end{bmatrix} \mathrm{RHC}: S_{3in} = \begin{bmatrix} P_{in4} \\ 0 \\ 0 \\ P_{in4} \end{bmatrix}, \quad (2.15)$$

Output Stokes Vectors for four polarizations are:

$$\begin{aligned} P_{out1} &= m_{11}P_{in1} + m_{12}P_{in2} \\ P_{out2} &= m_{11}P_{in1} - m_{12}P_{in2} \\ P_{out3} &= m_{11}P_{in3} + m_{13}P_{in3} \\ P_{out4} &= m_{11}P_{in4} + m_{14}P_{in4}. \end{aligned} \quad (2.16)$$

Determining all the power values needs following steps. First, four polarization states must be taken measurement individually, in order to get the different input powers. Second, the DUT is inserted and the output powers corresponding to four polarization states must be recorded.

It should be noticed that the power and wavelength for PDL measurement must be in the same condition.

Solving the equation system for $m_{11}...m_{14}$ of the Mueller matrix yields:

$$\begin{bmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\left(\frac{P_{out1}}{P_{in1}} + \frac{P_{out2}}{in2}\right) \\ \frac{1}{2}\left(\frac{P_{out1}}{P_{in1}} - \frac{P_{out2}}{in2}\right) \\ \frac{P_{out3}}{P_{in3}} - m_{11} \\ \frac{P_{out4}}{P_{in4}} - m_{11} \end{bmatrix}, \quad (2.17)$$

Then the maximum and minimum transmissions through DUT can be derived as:

$$\begin{aligned} P_{Max} &= m_{11} + \sqrt{m_{12}^2 + m_{13}^2 + m_{14}^2} \\ P_{Min} &= m_{11} - \sqrt{m_{12}^2 + m_{13}^2 + m_{14}^2}, \end{aligned} \quad (2.18)$$

By the definition, the PDL of the DUT can be calculated.

The Mueller method has some specific requirements on the polarization controller. It helps to generate four polarization states. It contains a polarizer and a retarder. The polarizer generates the linear polarization state, and the rotation of the retarder generates other polarization states. The angular of the controller will affect the accurate of the measurement.

The polarization controller also provides some loss. It will affect the measurement of the reference power.

Comparison

Generally, the polarization scanning technique is suitable for PDL measurement for single wavelength, and the Mueller method is capable for a wavelength range. The polarization scanning technique exposes the source to all the polarization states, so it can measure the PDL at a certain wavelength at one time. And it is easy to implement and doesn't take too complicated math work. On the contract, the Mueller method requires many wavelength points, it suitable for several channels.

In this thesis, we only have to take care of the PDL channel, which means the PDL is known to us. And we neglect all the noise introduced by the devices to simplify the simulation.

2.3 Factor Graph and Monte Carlo technique

As we mentioned above, factor graph and Monte Carlo technique are two very important techniques in this thesis, almost all the research works are based on them. We use factor graph to analysis the transmission system and use Monte Carlo technique to calculate the results of backpropagation. Before we can start derive the detector, we need to give some introductions for these two techniques.

2.3.1 Factor Graph

Factor graph is a particular type of graphical model that represents the factorization of a function. Suppose there is a function $f\colon \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_N \to \mathbb{R}$, which is factorized to K factors,

$$f(\mathrm{x}_1,\mathrm{x}_2,...,\mathrm{x}_N) = \prod_{k=1}^{K} f_k(\mathrm{s}_k), \tag{2.19}$$

where $f_k(\cdot)$ is a real-valued function and $\mathrm{s}_k \subseteq \{\mathrm{x}_1,\mathrm{x}_2,...,\mathrm{x}_N\}$ is the kth variable subset. The factor graph of this factorization is created as follows. First, We create an edge (drawn as a line) for every variable, and a node (drawn as a circle or a square) for every factor. If a certain variable appears in a certain factor, we attach the corresponding edge to the corresponding node. In our study, an edge can be attached only to two nodes, which means there is no branch in the factor graph in this thesis.

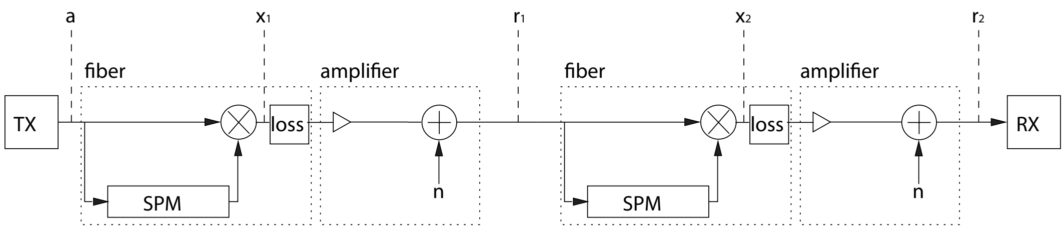

Figure 2.3: Transmission model for 2 spans.

If we substitute $f(\cdot)$ and $f_k(\cdot)$ with probability density functions, we can get the corresponding factorized representation of a certain distribution. For the system model given by Figure 2.3,(where the PDL channel is ignored) the factorization representation of the distribution $p(\mathbf{a},\mathbf{x}_1,\mathbf{r}_1,\mathbf{x}_2|\mathbf{r}_2)$ is given by

$$p(\mathbf{a},\mathbf{x}_1,\mathbf{r}_1,\mathbf{x}_2|\mathbf{r}_2) \propto p(\mathbf{a})p(\mathbf{x}_1|\mathbf{a})p(\mathbf{r}_1|\mathbf{x}_1)p(\mathbf{x}_2|\mathbf{r}_1)p(\mathbf{r}_2|\mathbf{x}_2), \tag{2.20}$$

where $\propto$ denotes quality up to a multiplicative constant. The corresponding factor graph is shown in Figure 2.4.

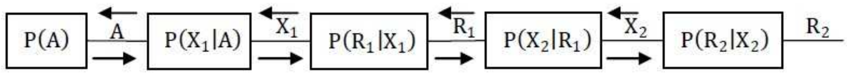

Figure 2.4: Factor graph of $p(\mathbf{a},\mathbf{x}_1,\mathbf{r}_1,\mathbf{x}_2|\mathbf{r}_2)$.

The arrow refers to the messages which are computed as follow: given a factor $f(\cdot)$ with variables x and y, and an incoming message $m(x)$, then the out going message is given by

$$m(y) = \mathrm{C}\int f(x,y)m(x)\mathrm{d}x, \tag{2.21}$$

where C is a normalization constant and the integration interval is the domain of x. Some factors may only have one variable (the factors on the boundary), if so, $m(y) = \mathrm{C}f(y)$ (if y is the only variable).

2.3.2 Monte Carlo technique

Monte Carlo techniques gives an innovation method to solve complex integration and optimization compare to traditional Bayesian estimation. The essential idea is representing distributions as a list of samples on which integration and optimization are based. The definition of particle representation is given as follows: suppose we have a random variable $\mathbf{Z}$ which defined over a set $\mathcal{Z}$. A particle representation of a distribution$p_{\mathbf{Z}}(\mathbf{z})$ is a set of L couples $(w^{(l)},\mathbf{z}^{(l)})$, with $\Sigma_l w^{(l)} = 1$, for any integrable function $f(\mathbf{z})$ from $\mathcal{Z} \to \mathbb{C}$,

$$I = \mathbb{E}_{\mathbf{Z}}\{f(\mathbf{Z})\}, \tag{2.22}$$

can be approximated by

$$I_L = \sum_{l=1}^{L} w^{(l)} f(\mathbf{z}^{(l)}), \tag{2.23}$$

where $\mathbf{z}^{(l)}$ is named a properly weighted sample with weight $w^{(l)}$.

Then we introduce an alternative notation for continuous $\mathbf{Z}$:

$$p_{\mathbf{Z}}(\mathbf{z}) \approx \sum_{l=1}^{L} w^{(l)} \delta(\mathbf{z} - \mathbf{z}^{(l)}), \tag{2.24}$$

where $\delta(\cdot)$ is the Dirac distribution. On the contrary, the notation of discrete $\mathbf{Z}$ is given by

$$p_{\mathbf{Z}}(\mathbf{z}) \approx \sum_{l=1}^{L} w^{(l)} \mathbb{I}\{\mathbf{z} = \mathbf{z}^{(l)}\}, \tag{2.25}$$

where, for a proposition X, $\mathbb{I}\{X\}$ is the indicator function, defined as $\mathbb{I}\{X\} = 1$ when X is true and $\mathbb{I}\{X\} = 0$ when X is false. With Mote Carlo technique, we can easily compute the expectation $\mathbb{E}_{\mathbf{Z}}\{f(\mathbf{Z})\}$.

3

Data detection

In this thesis, we simulate a coherent optical communication system of dual polarization 16-QAM by using FG, Monte-Carlo technique and back propagation. We simplify the problem by dividing the complicated system into separate components and organized the process in four steps.

3.1 Linear unitary channel

First, we build a simple transmission system by ignoring both the nonlinear noise and PDL channel effect. This can considered as the nonlinear parameter $\gamma = 0$ and the PDL channel is simplified to a unitary channel. Using this simplified model, we give a solution of channel estimation and data detection by factor graph and Monte Carlo technique. These estimation and detection methods are the bases of this thesis. The model is a transmission system with linear noise and unitary channel effects, which is shown in Figure 3.1.

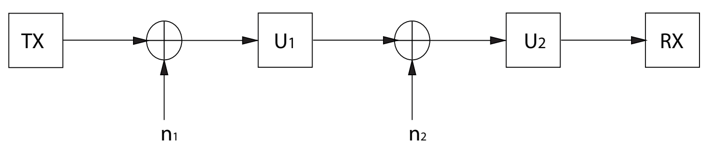

Figure 3.1: System model for 2 spans

The U block represent the unitary channel and n is the ASE noise of the fiber. For each span, the unitary channels are independent and different. Since the optical channel varies slowly, we treat the unitary channel as constant matrix in our simulation.

Assuming that the original signal is $\mathbf{s}_0$, the output of the first span $\mathbf{x}_1$ is given by

$$\mathbf{x}_1 = \mathbf{U}_1(\mathbf{s}_0 + \mathbf{n}_1) \quad (3.1)$$
$$= \mathbf{U}_1\mathbf{s}_0 + \mathbf{U}_1\mathbf{n}_1, \quad (3.2)$$

where $\mathbf{U}_1$ is the unitary channel of the first span and $\mathbf{n}_1$ is the corresponding ASE noise. In 3.2, we can calculate the mean and variance of $\mathbf{U}_1\mathbf{n}_1$. Since $\mathbf{n}_1$ has zero mean and a variance $\mathbf{var}\{\mathbf{n}_1\}$, the expectation of $\mathbf{U}_1\mathbf{n}_1$ is $\mathbf{E}\{\mathbf{U}_1\mathbf{n}_1\} = \mathbf{E}\{\mathbf{U}_1\}\mathbf{E}\{\mathbf{n}_1\} = 0$, the variance of $\mathbf{U}_1\mathbf{n}_1$ is $\mathbf{var}\{\mathbf{U}_1\mathbf{n}_1\} = \mathbf{E}\{(\mathbf{U}_1\mathbf{n}_1)^2\} - (\mathbf{E}\{\mathbf{U}_1\mathbf{n}_1\})^2$, where $\mathbf{E}\{\mathbf{U}_1\mathbf{n}_1\} = 0$ and $(\mathbf{U}_1\mathbf{n}_1)^2 = \mathbf{n}_1^{\mathrm{H}}\mathbf{U}_1^{\mathrm{H}}\mathbf{U}_1\mathbf{n}_1 = \mathbf{n}_1^2$, so $\mathbf{var}\{\mathbf{U}_1\mathbf{n}_1\} = \mathbf{E}\{\mathbf{n}_1^2\} = \mathbf{var}\{\mathbf{n}_1\}$. After the derivation above, we can conclude that $\mathbf{U}_1\mathbf{n}_1$ is still an additive white Gaussian noise with the same mean and variance of $\mathbf{n}_1$, then 3.2 can be simplified to

$$\mathbf{x}_1 = \mathbf{U}_1\mathbf{s}_0 + \mathbf{n}_1, \quad (3.3)$$

Then the output of the second span $\mathbf{x}_2$ is given by

$$\mathbf{x}_2 = \mathbf{U}_2(\mathbf{x}_1 + \mathbf{n}_2) \quad (3.4)$$
$$= \mathbf{U}_2(\mathbf{U}_1\mathbf{s}_0 + \mathbf{n}_1 + \mathbf{n}_2) \quad (3.5)$$
$$= \mathbf{U}_2\mathbf{U}_1\mathbf{s}_0 + \mathbf{n}_1 + \mathbf{n}_2, \quad (3.6)$$

where we also use the conclusion above. In 3.6, $\mathbf{U}_2\mathbf{U}_1$ is still a unitary matrix since $(\mathbf{U}_2\mathbf{U}_1)^{\mathrm{H}}(\mathbf{U}_2\mathbf{U}_1) = \mathbf{U}_1^{\mathrm{H}}\mathbf{U}_2^{\mathrm{H}}\mathbf{U}_2\mathbf{U}_1 = I$. So 3.6 can be rewrite as

$$\mathbf{x}_2 = \mathbf{U}\mathbf{s}_0 + \mathbf{n}_1 + \mathbf{n}_2, \quad (3.7)$$

where $\mathbf{U} = \mathbf{U}_2\mathbf{U}_1$ is a unitary matrix. We can easily extend the result to k-th span:

$$\mathbf{x}_k = \mathbf{U}\mathbf{s}_0 + \mathbf{n}_1 + \mathbf{n}_2 + \cdots + \mathbf{n}_k \quad (3.8)$$
$$= \mathbf{U}(\mathbf{s}_0 + \mathbf{U}(\mathbf{n}_1 + \mathbf{n}_2 + \cdots + \mathbf{n}_k)) \quad (3.9)$$
$$= \mathbf{U}(\mathbf{s}_0 + \mathbf{n}_1 + \mathbf{n}_2 + \cdots + \mathbf{n}_k), \quad (3.10)$$

where $\mathbf{U} = \mathbf{U}_k \ldots \mathbf{U}_2\mathbf{U}_1$, here we use the character that $\mathbf{U}^{\mathrm{H}}$ is also a unitary matrix to derive 3.10. Now the system model for k spans is reorganized as Figure 3.2.

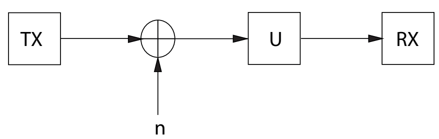

Figure 3.2: Simplified system model for k spans

The system model for k spans is exactly the same as for 1 span, the only differences are the values of noise $\mathbf{n}$ and channel $\mathbf{U}$. In this model, $\mathbf{n} = \mathbf{n}_1 + \mathbf{n}_2 + \cdots + \mathbf{n}_k$ and $\mathbf{U} = \mathbf{U}_k \ldots \mathbf{U}_2\mathbf{U}_1$.

3.1.1 Channel estimation

Assume that the input signal is $\mathbf{s}$, the output variable of the system is $\mathbf{r}$ and the variable $\mathbf{x}$ is given by $\mathbf{x} = \mathbf{s} + \mathbf{n}$. The receive signal is given by

$$\mathbf{r} = \mathbf{U}(\mathbf{s} + \mathbf{n}) \tag{3.11}$$
$$= \mathbf{U}\mathbf{x}, \tag{3.12}$$

where we can get the channel estimation method since

$$\mathbf{U} = \mathbf{r}\mathbf{x}^{-1}, \tag{3.13}$$

where we assumed that both $\mathbf{r}$ and $\mathbf{x}$ are 2×2 square matrices, $\mathbf{x}^{-1}$ is the inverse matrix of $\mathbf{x}$. We send a number of fixed signals which is known as training symbols, use these training symbols estimate the channel. In 3.13, $\mathbf{r}$ is known as the receive signal, but $\mathbf{x}$ is a variable that cannot be measured. However, we can compute the estimation of $\mathbf{x}$ by using Monte Carlo technique instead of using the exact value. Then the channel estimation equation is given by

$$\widehat{\mathbf{U}} = \mathbf{E}(\mathbf{r})\mathbf{E}(\mathbf{x})^{-1}, \tag{3.14}$$

where $\mathbf{E}(\mathbf{r})$ and $\mathbf{E}(\mathbf{x})$ are the estimation of $\mathbf{r}$ and $\mathbf{x}$, $\widehat{\mathbf{U}}$ is the estimated channel. $\mathbf{E}(\mathbf{x}) = \mathbf{E}(\mathbf{s}) + \mathbf{E}(\mathbf{n})$, where $\mathbf{E}(\mathbf{n}) = 0$. To make sure 3.14 works well, the determinant of $\mathbf{E}(\mathbf{s})$ should be a non-zero value.

The zero-forcing equalizer is given by

$$\mathbf{w} = \widehat{\mathbf{U}}^{\mathrm{H}}, \tag{3.15}$$

If the estimation of the channel is perfect, which means $\widehat{\mathbf{U}} = \mathbf{U}$ then the output of the equalizer is given by

$$\hat{\mathbf{x}} = \widehat{\mathbf{U}}^{\mathrm{H}}\mathbf{r} \tag{3.16}$$
$$= \widehat{\mathbf{U}}^{\mathrm{H}}\mathbf{U}\mathbf{x} \tag{3.17}$$
$$= \mathbf{x}. \tag{3.18}$$

3.1.2 Data detection

Assume that we have perfect knowledge of the channel, the transmission system will become to a simple AWGN model, backpropagation is then used to perform data detection. It works as follow: first, for each received signal $\mathbf{r}_i$ (after equalization), M samples(particles) are generated as $\mathbf{r}_i^{(l)}, l = 1{,}2 \ldots M$, then the samples are sent to backward transmission in the system. Since the system only contains AWGN noise, the output of the backward transmission is $\mathbf{a}_i^{(l)} = \mathbf{r}_i^{(l)} + \mathbf{n}$ where $\mathbf{a}_i^{(l)}$ is a bunch of samples which can be considered as the prediction of the original transmitted signal $\mathbf{s}_i$. We calculate the distribution (mean and variance) of the samples and substitute all the possible original signal into the distribution. The results are in direct proportion to the probability

$p(\mathbf{s}_i|\mathbf{r}_i)$ which means the signal gives the maximum value of the distribution function is the result of the detection. Our detector is called stochastic backpropagation.

Here we discuss a little bit of the number M, generally speaking, the bigger M is the better result we get since more information of the function are collected from the particles. However, the result varies very slowly after M reaches a certain number[1]. The system will be become less efficiency if we continue increase M and will not give good performance if M is not big enough.

3.2 Non-linear unitary channel

After simulated the system of linear noise and unitary channel effect, we built a system contains the nonlinear part and white Gaussian noise with the unitary channel effect. It can be considered as we change the nonlinear parameter γ from 0 to a non-zero constant. The system model is shown in Figure 3.3.

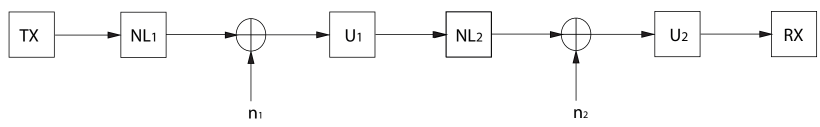

Figure 3.3: System model for 2 spans

Assuming that the original signal is $\mathbf{s}_0$, the output of the first span $\mathbf{x}_1$ is given by

$$\mathbf{x}_1 = \mathbf{U}_1(\mathbf{s}_0 \exp(j\gamma L_{\text{eff}} \|\mathbf{s}_0\|^2) + \mathbf{n}_1) \tag{3.19}$$
$$= \mathbf{U}_1\mathbf{s}_0 \exp(j\gamma L_{\text{eff}} \|\mathbf{s}_0\|^2) + \mathbf{n}_1, \tag{3.20}$$

In 3.20, we used the conclusion from the above section that $\mathbf{U}_1\mathbf{n}_1 = \mathbf{n}_1$. The output of 2 spans is then given by

$$\mathbf{x}_2 = \mathbf{U}_2\mathbf{x}_1 \exp(j\gamma L_{\text{eff}} \|\mathbf{x}_1\|^2) + \mathbf{n}_2, \tag{3.21}$$

If $\mathbf{s}_1 \triangleq \mathbf{s}_0 \exp(j\gamma L_{\text{eff}} \|\mathbf{s}_0\|^2) + \mathbf{n}_1$, then $\mathbf{x}_1 = \mathbf{U}_1\mathbf{s}_1$ and $\|\mathbf{x}_1\|^2 = \|\mathbf{U}_1\mathbf{s}_1\|^2 = (\mathbf{s}_1)^{\mathrm{H}}\mathbf{U}_1^{\mathrm{H}}\mathbf{U}_1\mathbf{s}_1 = (\mathbf{s}_1)^{\mathrm{H}}\mathbf{s}_1 = \|\mathbf{s}_1\|^2$. So 3.21 can be rewrite as

$$\mathbf{x}_2 = \mathbf{U}_2\mathbf{U}_1\mathbf{s}_1 \exp(j\gamma L_{\text{eff}} \|\mathbf{s}_1\|^2) + \mathbf{n}_2 \tag{3.22}$$
$$= \mathbf{U}\left(\mathbf{s}_1 \exp(j\gamma L_{\text{eff}} \|\mathbf{s}_1\|^2) + \mathbf{n}_2\right), \tag{3.23}$$

where $\mathbf{U} = \mathbf{U}_2\mathbf{U}_1$. From the derivation above, we find that unitary channel has no effect on nonlinear noise, so we reorganize the system model as Figure 3.4.

U is the only channel effect in the system and we use the same method as above to estimate it.

[1] In this thesis, M = 100 is considered to be sufficient

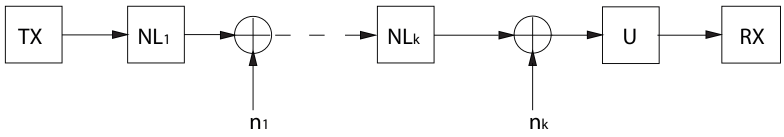

Figure 3.4: System model for k spans

3.2.1 Data detection

Assume that we have the perfect knowledge of the unitary channel, the only problem left is how to deal with the nonlinear noise. With stochastic backpropagation technique, we compensate the nonlinear noise for each span. The nonlinear noise can be considered as a rotation in constellation without changing the signal energy. For a nonlinear component $\mathbf{r}_k = \mathbf{s}_k \exp(j\gamma L_{\text{eff}} \|\mathbf{s}_k\|^2)$, the original signal $\mathbf{s}_k$ can be regenerated as $\mathbf{s}_k = \mathbf{r}_k \exp(-j\gamma L_{\text{eff}} \|\mathbf{r}_k\|^2)$. As we said above, the nonlinear noise will not change the signal energy which means $\|\mathbf{s}_k\|^2 = \|\mathbf{r}_k\|^2$.

Combining the methods we gave in linear noise system, we can describe the back propagation method as follow. After channel equalization, we generate particles for each received signal $\mathbf{r}_k$ then send the particles to backward transmission. For each span, the transmission model is $\mathbf{x}_k = \mathbf{r}_k + \mathbf{n}$, $\mathbf{a}_k = \mathbf{x}_k \exp(-j\gamma L_{\text{eff}} \|\mathbf{x}_k\|^2)$. After the transmission process is finished, we can get the prediction of the original transmitted signal. The detection method is same as we provide in the first section.

In this step, we compare our method with two existed techniques, the regular maximum likelihood (RML) detector and the back propagation method we provided in chapter 2. The RML detector simply ignores the nonlinear noise, considers the system as $\mathbf{r}_k = \mathbf{s}_k + \mathbf{n}$. The detector is given by

$$\hat{\mathbf{a}}_{\mathrm{k}} = \arg \min_{\mathbf{a}_{\mathrm{k}} \in \Omega^2} \|\mathbf{r}_{\mathrm{k}} - \mathbf{s}_{\mathrm{k}}\|^2 , \tag{3.24}$$

3.3 Linear non-unitary channel

The third step is mainly about the estimation for PDL channel. The system we discussed here only consist channels and white Gaussian noise. For each span the signal goes through the channel first then the noise is added. The signal is dual polarization QPSK. The original signal is $\mathbf{s}_0$, the polarization dependent attenuation factor $\tilde{\gamma}$ for PDL is fixed and all the channels have the same value. Figure 3.5 shows the system model for two spans.

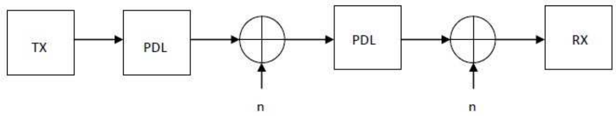

Figure 3.5: System model for 2 span

3.3.1 Channel estimation

The signal passed the first channel is

$$\mathbf{r}_1 = \mathrm{H}_1\mathbf{s}_0 + \mathbf{n}_1, \tag{3.25}$$

where H_1 is the first channel and $\mathbf{n}_1$ is the added noise with noise covariance matrix N_0/L, N_0 is the noise variance, L is number of spans. In our simulation, N_0 is 1, L is 22.

The signal passed the second channel is

$$\mathbf{r}_2 = \mathrm{H}_2\mathbf{r}_1 + \mathbf{n}_2 = \mathrm{H}_2\mathrm{H}_1\mathbf{s}_0 + (\mathrm{H}_2\mathbf{n}_1 + \mathbf{n}_2), \tag{3.26}$$

The noise covariance matrix for this case is

$$\begin{aligned}&(\mathrm{H}_2\mathbf{n}_1 + \mathbf{n}_2) \times (\mathrm{H}_2\mathbf{n}_1 + \mathbf{n}_2)^{\mathrm{H}} \\ &= (\mathrm{H}_2\mathbf{n}_1 + \mathbf{n}_2) \times \left(\mathbf{n}_1^{\mathrm{H}}\mathrm{H}_2^{\mathrm{H}} + \mathbf{n}_2^{\mathrm{H}}\right) \\ &= \mathrm{H}_2\left(\frac{\mathrm{N}_0}{\mathrm{L}}\right)\mathrm{H}_2^{\mathrm{H}} + \frac{\mathrm{N}_0}{\mathrm{L}},\end{aligned} \tag{3.27}$$

So the signal after passing i-th span is

$$\mathbf{r}_i = \mathrm{H}_i\mathbf{r}_{i-1} + \mathbf{n}_i, \tag{3.28}$$

and the noise covariance matrix is

$$\Sigma = \mathrm{H}_i ... \mathrm{H}_2\left(\frac{\mathrm{N}_0}{\mathrm{L}}\right)\mathrm{H}_2^{\mathrm{H}} ... \mathrm{H}_i^{\mathrm{H}} + ... + \mathrm{H}_i\left(\frac{\mathrm{N}_0}{\mathrm{L}}\right)\mathrm{H}_i^{\mathrm{H}} + \frac{\mathrm{N}_0}{\mathrm{L}}, \tag{3.29}$$

the total channel we wish to estimate is the multiplication of all the channel per span, the total channel is

$$\mathrm{H}_{\mathrm{total}} = \mathrm{H}_i\mathrm{H}_{i-1} ... \mathrm{H}_2\mathrm{H}_1, \tag{3.30}$$

The noise covariance matrix has an interesting property. After passing 22 of spans, it is nearly stable. After simulation, when $\tilde{\gamma}$ equals to 0.99, Σ is around $\begin{bmatrix} 0.9 & 0 \\ 0 & 0.9 \end{bmatrix}$, when $\tilde{\gamma}$ equals to 0.9, Σ is around $\begin{bmatrix} 0.4 & 0 \\ 0 & 0.4 \end{bmatrix}$.

We use maximum likelihood function of the posteriori probability to estimate the channel. The original dual polarization signal is jointly Gaussian, so the maximum likelihood function is

$$\hat{\mathbf{r}} = \arg\max\ p(\mathbf{r}|\mathrm{H}_{\mathrm{total}}\mathbf{s}_0) = (\mathbf{r} - \mathrm{H}_{\mathrm{total}}\mathbf{s}_0)^{\mathrm{H}}\Sigma^{-1}(\mathbf{r} - \mathrm{H}_{\mathrm{total}}\mathbf{s}_0), \tag{3.31}$$

where $\mathbf{r}$ is the received signal, $\mathbf{s}_0$ is the original signal, Σ is the noise covariance matrix, $\mathrm{H}_{\mathrm{total}}$ is the total channel we have to estimate. The function derives:

$$\begin{aligned} \hat{\mathbf{r}} &= \arg\max\ p(\mathbf{r}|\mathrm{H}_{\mathrm{total}}\mathbf{s}_0) \\ &= \mathbf{r}^{\mathrm{H}}\Sigma^{-1}\mathbf{r} + \mathbf{s}_0^{\mathrm{H}}\mathrm{H}_{\mathrm{total}}^{\mathrm{H}}\Sigma^{-1}\mathrm{H}_{\mathrm{total}}\mathbf{s}_0 - \mathbf{s}_0^{\mathrm{H}}\mathrm{H}_{\mathrm{total}}^{\mathrm{H}}\Sigma^{-1}\mathbf{r} - \mathbf{r}^{\mathrm{H}}\Sigma^{-1}\mathrm{H}_{\mathrm{total}}\mathbf{s}_0, \end{aligned} \tag{3.32}$$

In order to maximum the function, we take the derivatives of this function, and set it equals to 0 to calculate the estimated channel. And according to the simulation, the noise covariance matrix Σ can be considered as a fixed matrix β. So the function can be written as

$$\begin{aligned} \frac{\partial \arg\max\ p(\mathbf{r}|\mathrm{H}_{\mathrm{total}}\mathbf{s}_0)}{\partial \mathrm{H}_{\mathrm{total}}} &= 0 \\ \left((\beta^{-1})^{\mathrm{H}} + \beta^{-1}\right)\mathrm{H}_{\mathrm{total}}\mathbf{s}_0\mathbf{s}_0^{\mathrm{H}} &= \left((\beta^{-1})^{\mathrm{H}} + \beta^{-1}\right)\mathbf{r}\mathbf{s}_0^{\mathrm{H}}, \end{aligned} \tag{3.33}$$

So the estimated channel is

$$\mathrm{H}_{\mathrm{esti}} = \mathbf{r}\mathbf{s}_0(\mathbf{s}_0\mathbf{s}_0^{\mathrm{H}})^{-1}, \tag{3.34}$$

It is just a function of received signal and original signal, doesn't relate to the noise. And the estimated channel is a scaled identity matrix.

3.3.2 Data detection

We presented two kinds of equalizers, zero forcing (ZF) equalizer and minimum mean square error (MMSE) equalizer in detecting the data.

Zero forcing (ZF) equalizer simply applies the inverse of the channel to the received signal. It is a form of linear equalization which is widely used in the communication systems[14]. For the estimated channel $\mathrm{H}_{\mathrm{esti}}$, C_{ZF} is the equalization matrix, it is defined as

$$\mathrm{C}_{\mathrm{ZF}} = {\mathrm{H}_{\mathrm{esti}}}^{-1}, \tag{3.35}$$

It should be noted that, ZF equalization is not suitable for most applications, as even the channel impulse has finite length, the impulse response of the ZF equalizer is infinite; and ZF equalizer neglects the effects of the noise.

Minimum mean square error (MMSE) equalizer may be a better solution. MMSE equalizer minimizes the mean square error (MSE) of the transmitted data. MSE is a risk function. It is related to the expectation of the squared error loss, and measures the average of the square errors. It is defined as[14]

$$\mathrm{C}_{\mathrm{MMSE}} = \frac{\mathrm{H}_{\mathrm{esti}}^{\mathrm{H}}}{\|\mathrm{H}_{\mathrm{esti}}\|^2 + \frac{\Sigma^2}{{\sigma_s}^2}}, \tag{3.36}$$

where Σ is the noise covariance matrix, which is considered to be a fixed matrix and can be obtained from the simulation. ${\sigma_s}^2$ is the variance of the modulated symbols.

In the general cases, the MMSE equalizer should have a better performance than the ZF equalizer, as the MMSE equalizer takes the effects of the noise into consideration. But in the multi-spans PDL channel, these two equalizers have nearly the same performance when estimating the same channel under the same circumstance. This may due to the noise.

We built a similar system to check this assumption. Instead of passing the channel first, the signal is firstly added the noise then passing the PDL channel. So the received signal after passing the first span is rewritten as

$$\mathbf{r}_1 = \mathrm{H}_1(\mathbf{s}_0 + \mathbf{n}_1), \tag{3.37}$$

so now the noise covariance matrix is $\mathrm{H}_1 \left(\frac{\mathrm{N}_0}{\mathrm{L}}\right) \mathrm{H}_1^{\mathrm{H}}$.

The signal passing the second span is

$$\mathbf{r}_2 = \mathrm{H}_2(\mathbf{r}_1 + \mathbf{n}_2) = \mathrm{H}_2\mathrm{H}_1\mathbf{s}_0 + (\mathrm{H}_2\mathrm{H}_1\mathbf{n}_1 + \mathrm{H}_2\mathbf{n}_2), \tag{3.38}$$

and the noise covariance matrix is

$$\begin{aligned} &(\mathrm{H}_2\mathrm{H}_1\mathbf{n}_1 + \mathrm{H}_2\mathbf{n}_2) \times (\mathrm{H}_2\mathrm{H}_1\mathbf{n}_1 + \mathrm{H}_2\mathbf{n}_2)^{\mathrm{H}} \\ &= (\mathrm{H}_2\mathrm{H}_1\mathbf{n}_1 + \mathrm{H}_2\mathbf{n}_2) \times (\mathbf{n}_1^{\mathrm{H}}\mathrm{H}_1^{\mathrm{H}}\mathrm{H}_2^{\mathrm{H}} + \mathbf{n}_2^{\mathrm{H}}\mathrm{H}_2^{\mathrm{H}}) \\ &= \mathrm{H}_2\mathrm{H}_1 \left(\frac{\mathrm{N}_0}{\mathrm{L}}\right) \mathrm{H}_1^{\mathrm{H}}\mathrm{H}_2^{\mathrm{H}} + \mathrm{H}_2 \left(\frac{\mathrm{N}_0}{\mathrm{L}}\right) \mathrm{H}_2^{\mathrm{H}}, \end{aligned} \tag{3.39}$$

The received signal after passing the i-th span is

$$\mathbf{r}_i = \mathrm{H}_i(\mathbf{r}_{i-1} + \mathbf{n}_i), \tag{3.40}$$

the noise variance matrix Σ for the total system is

$$\Sigma = \mathrm{H}_i ... \mathrm{H}_1 \left(\frac{\mathrm{N}_0}{\mathrm{L}}\right) \mathrm{H}_1^{\mathrm{H}}\mathrm{H}_i^{\mathrm{H}} + ... + \mathrm{H}_i\mathrm{H}_{i-1} \left(\frac{\mathrm{N}_0}{\mathrm{L}}\right) \mathrm{H}_{i-1}^{\mathrm{H}}\mathrm{H}_i^{\mathrm{H}} + \mathrm{H}_i \left(\frac{\mathrm{N}_0}{\mathrm{L}}\right) \mathrm{H}_i^{\mathrm{H}}, \tag{3.41}$$

The channel is estimated in the same way by using the same calculation, the estimated channel is still

$$\hat{\mathrm{H}_{\mathrm{esti}}} = \mathbf{r}\mathbf{s}_0(\mathbf{s}_0\mathbf{s}_0^{\mathrm{H}})^{-1}, \tag{3.42}$$

and the Σ is still nearly a fixed matrix after passing several spans.

In this case, the performance of the MMSE equalizer has a clear improvement, is much better than the ZF equalizer. Comparing with the previous case, we can easily figure out that the difference in the noise covariance matrix Σ is the key issue.

In the second case, each component of the Σ matrix is affected by the first the channel. The noise will contribute more in the transmission. Since the MMSE equalizer is an estimation dealing with the noise. So the estimation in the second case for MMSE equalizer is better.

Since the ZF and MMSE equalizer has the same performance, we may implement ZF equalizer in practical, as it is easier to apply and has less computing complexity.

In both cases, the regular maximum likelihood (ML) detector is implemented to check the BER of the transmission.

Besides, the quality of transmission largely depends on the value of polarization dependent attenuation factor $\tilde{\gamma}$. When $\tilde{\gamma}$ is high, near 1, the BER can easily reach 0.001, that is the reason that we all wish the attenuation of the channel to be less.

3.4 Non-linear non-unitary channel

In the end, the nonlinear noise is added to the non-unitary system, the system model for 1 span is shown in Figure 3.6.

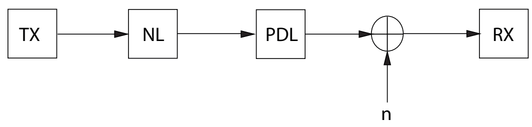

Figure 3.6: System model for 1 span

The NL block is the nonlinear phase noise, the PDL block represents the PDL channel effect, n is the AWGN noise. The PDL channel is the multiplication of a random unitary matrix, an attenuation Γ matrix and the hermitian of the unitary matrix.

3.4.1 Data detection

The channel estimation here is a little bit different here. since the nonlinearity and the Γ matrix is existed, we can not apply the method we provided in the previous section. Due to the limitation of research period, we only provide the data detection method here.

Assuming that we have the prefect knowledge of the PDL channel, which means we know not only the value of Γ, but also every unitary matrix in each span. With the stochastic backpropagation technology, the detection part is almost the same with the non-linear unitary channel, the only difference is there is a PDL channel component in each span. We generate bunch of samples from the backward transmission, and then the distribution of these samples were calculated. The original signals are then substituted to the distribution. The maximum value indicates the most probably prediction.

4

Performance analysis

Here we discuss the numerous results, like the lowest BER we can achieve, the performance of the detector comparing with others, something interesting on the constellation, the least number of training symbols we need to estimate the channel, the components of the training symbols and so on.

There are some common parameters for all the simulation, we simulate at 14 Gbaud per polarization, with $N_a = 22$ spans, $M = 100$ particles, $N_0 = 4.9 \times 10^{-7}\ W/Hz$, the nonlinear parameter $\gamma = 1.25\ W^{-1}\text{km}^{-1}$, $L_{\text{eff}} = 17.36$ km.

4.1 Linear unitary channel

We assume that transmitted signals are dual polarization 16-QAM. The input power is from -10 dBm to -2 dBm. According to the analysis in chapter 3, the system is a normal AWGN model if we gain perfect knowledge of the unitary channels. So we can compare our simulation result with the theoretical value of symbol error rate (SER) for 16-QAM. If the two results are approximate to each other, then we can conclude that the channel estimation is good enough and stochastic backpropagation method with FG and Monte Carlo technique is a correct solution.

4.1.1 Simulation results and performance analysis

The theoretical SER performance for M-QAM is given by

$$P_s = 1 - \left(1 - \frac{2(\sqrt{M}-1)}{\sqrt{M}} Q\left(\sqrt{\frac{3\text{SNR}}{M-1}}\right)\right)^2, \tag{4.1}$$

where M is the modulation factor which in this case is 16, SNR is the Signal-to-Noise Ratio which is SNR $= E_s/N_0$, where E_s is the average symbol energy and N_0 is the noise variance, Q is the Q-function given by

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{x^2}{2}} dx. \tag{4.2}$$

Assume we have perfect knowledge of the channel, we get the simulation results as follow:

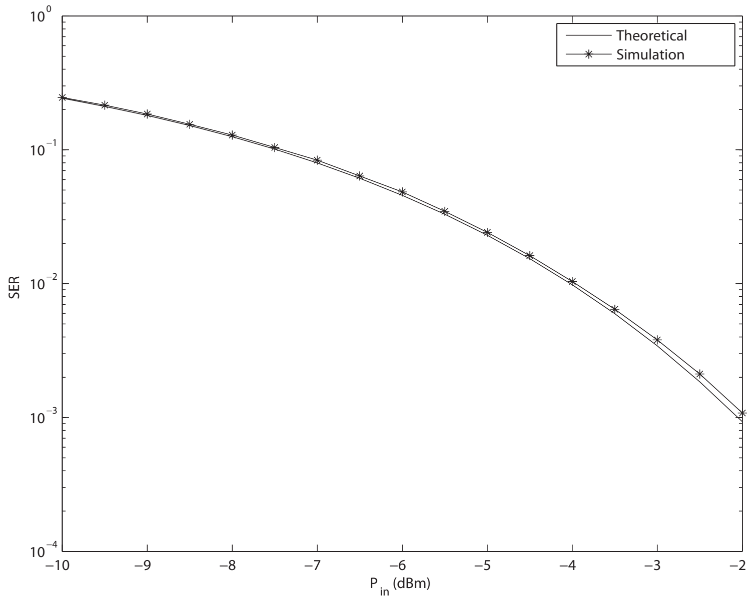

Figure 4.1: SER as a function of P_{in} with perfect channel knowledge

We can find that the two curves almost overlap each other, which mean the stochastic backpropagation method performs very well when we ignore the channel effect. When $P_{\text{in}} = -2$ dBm, the theoretical SER is about 9.3×10^{-4} and the stochastic backpropagation SER is 1.1×10^{-3}. These simulation results proved that we can use the stochastic backpropagation method as the data detection solution. The reason why the stochastic backpropagation is a little worse than the theoretical value is because the particle representation does not contains all the information of the transmission function since the number of particles is limited.

The simulation results with channel effect is shown in Figure 4.2

Since the stochastic backpropagation results are still very close to the theoretical curve, we can conclude that the method we provided in chapter 3 is a right solution for linear noise and unitary channel system. When $P_{\text{in}} = -2$ dBm, the stochastic backpropagation SER is still 1.1×10^{-3}. The simulation result is same as before, which mean the channel estimation is almost perfect in this situation.

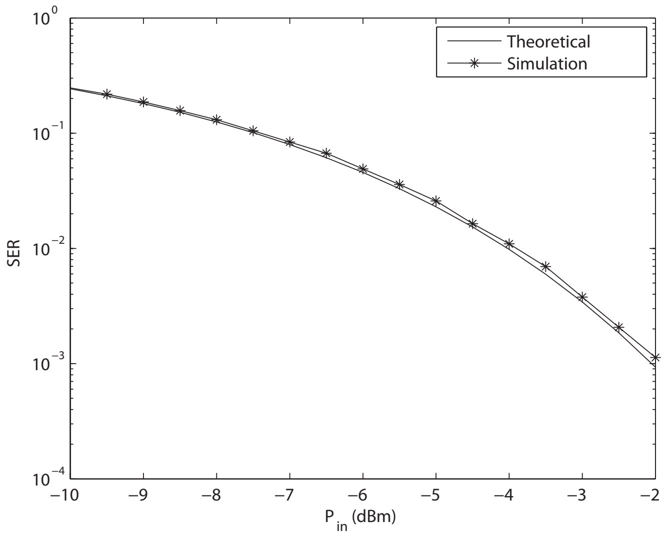

Figure 4.2: SER as a function of P_{in} with unitary channel effect

4.1.2 Channel estimation results

Since we transmit dual-polarization 16-QAM signals, the unitary channel is a 2×2 matrix. If the transmitted signals and the receive signals are also 2×2 matrices, we can easily compute $\widehat{\mathbf{U}}$ from equation 3.14. This gives us the method of how to choose training symbols, assume the number of training symbols is 1000, the training symbols should be chosen like

$$\begin{bmatrix} \mathbf{s}_{1\text{x}} \cdots \mathbf{s}_{1\text{x}} & \mathbf{s}_{2\text{x}} \cdots \mathbf{s}_{2\text{x}} \\ \mathbf{s}_{1\text{y}} \cdots \mathbf{s}_{1\text{y}} & \mathbf{s}_{2\text{y}} \cdots \mathbf{s}_{2\text{y}} \end{bmatrix}$$

There are 1000 dual polarization training symbols, contain 500 $\mathbf{s}_1$ and 500 $\mathbf{s}_2$. $\mathbf{s}_{1\text{x}}$, $\mathbf{s}_{1\text{y}}$, $\mathbf{s}_{2\text{x}}$ and $\mathbf{s}_{2\text{y}}$ are chosen from the 16-QAM constellation separately. In channel estimation process, we calculate the mean value of the first 500 symbols and the last 500 symbols, then we can get $\mathbf{E}(\mathbf{x})$, which is a 2 dimensional square matrix. Using the similar method, we get $\mathbf{E}(\mathbf{r})$. Then we estimate the channel by equation 3.14.

In the simulation above, we use 1000 training symbols, which is considered to be sufficient to get good estimation of the channel. However, we want to use as less training symbols as we can to improve the efficiency of the system in practise. For this purpose, we need to find out the relationship between the number of training symbols N_{t} and the estimation results. Here we fix the input power to -2 dBm, which gives the theoretical SER $P_s = 9.3 \times 10^{-4}$. We change the number of training symbols from 10 to 1000, calculate SER for these values, the results is given in Figure 4.3.

We can find that SER varies slow after N_{t} reaches 600. When N_{t} is bigger that 800, the results are less than 2×10^{-3}. For practical use, 700 or 800 should be a good choice if we take the transmission energy and time cost into account.

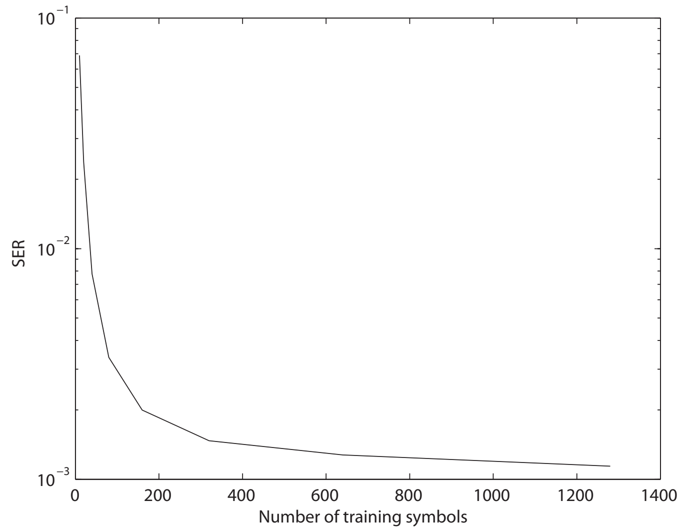

Figure 4.3: SER as a function of N_t

4.2 Non-linear unitary channel

4.2.1 Simulation results and performance analysis

The simulation environment is the same as linear noise, unitary channel system. By applying the methods in chapter 3, we get a result of SER against the input power of dBm for RML, backpropagation and stochastic backpropagation. Same as the above section, we give the results with prefect channel estimation first.

From Figure 4.4 we can tell that RML detector gives the worst performance, this is predictable since it does not take SPM into account. The backpropagation detector is much better that RML, because its post-compensation of the phase shift. As the input power becomes bigger, the nonlinear phase noise effect increases while AWGN noise effect decreases. When SNR is low, the AWGN noise is the dominant component in the system, so the SER becomes lower when input power increases. When SNR is high, the nonlinear phase noise becomes the dominant components and cause the performance degrades. The oscillation in the curve is caused by the interaction between linear noise and SPM[1]. The stochastic backpropagation detector we provided in this thesis gives the best performance of these three detector. It considers not only the nonlinearity but also the linear noise interaction. The lowest SER we achieved is about 3.98×10^{-4}, when $P_{\text{in}} = 4$ dBm.

The results with channel estimation by 1000 training symbols are shown in Figure 4.5 We can find that the results of RML and backpropagation detector are almost same as before, while the performance of stochastic backpropagation detector degraded. Especially when SNR is high, the SER of backpropagation detector increases very fast. However, the stochastic backpropagation detector still gives the lowest SER compare to

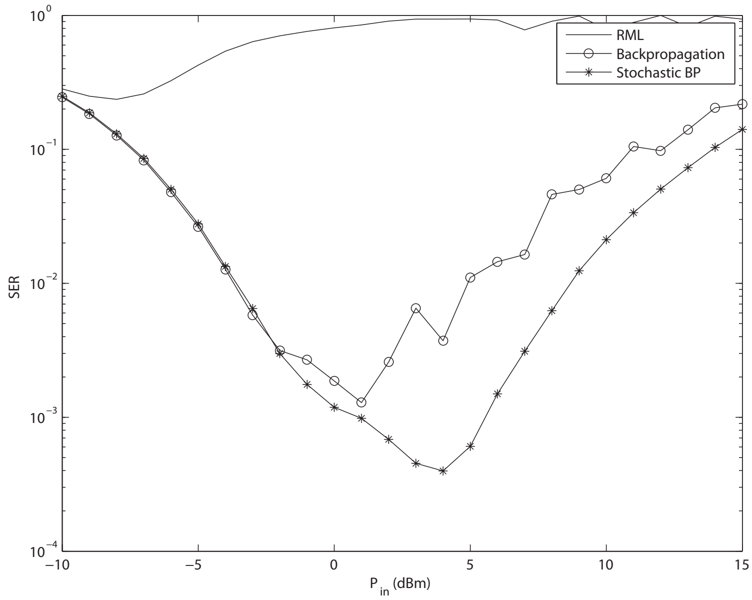

Figure 4.4: SER as a function of P_{in} without channel effect

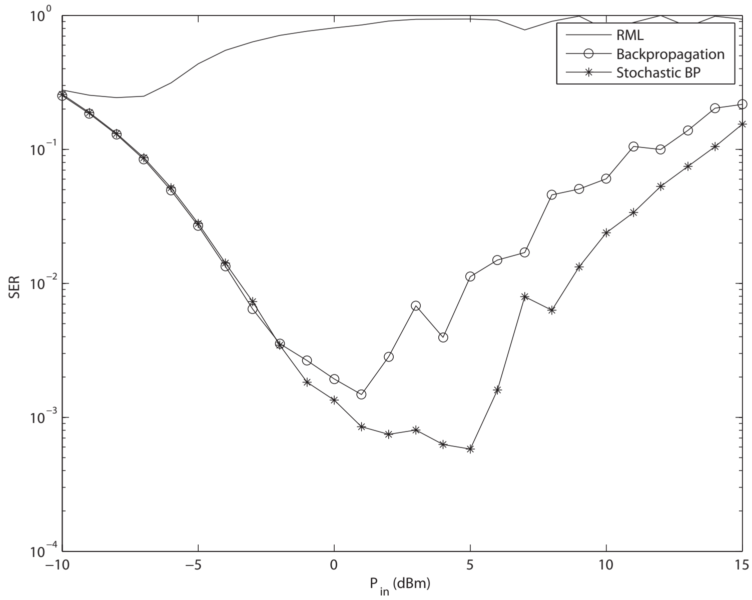

Figure 4.5: SER as a function of P_{in} for nonlinear noise unitary channel system

other detector. The minimum value is 5.8×10^{-4} when $P_{\text{in}} = 5$ dBm

4.2.2 Channel estimation results

As we discussed in the above section, the channel estimation method we provided in chapter 3 works very well in linear noise system, so the possible reason that the performance becomes worse is the nonlinear phase noise effect the channel estimation results. To verify this assumption,we simulate the channel estimation results. We fix the input power to 5.5 dBm, which gives the SER $P_s = 9.2 \times 10^{-4}$ when there is no channel effect. We change the number of training symbols from 10 to 1000, calculate SER for these values, the results is given in Figure 4.6.

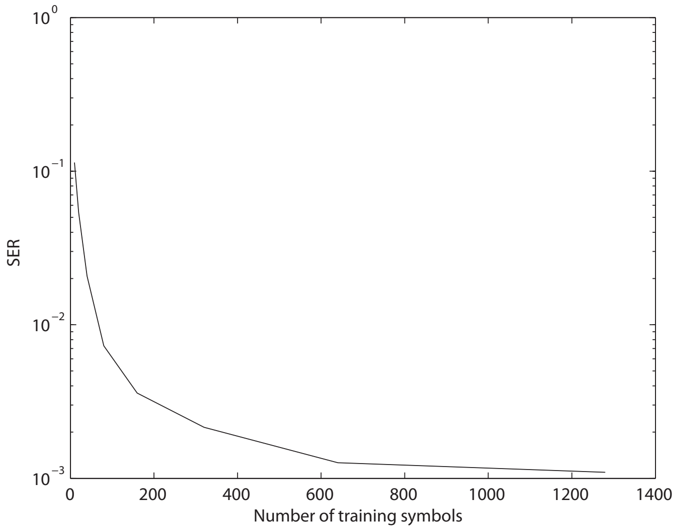

Figure 4.6: SER as a function of N_t

If we compare Figure 4.6 and Figure 4.3, we can find that the results are almost the same, especially when $N_t > 600$, which means the channel estimation performance is still good. So the real reason cause the detection performance degradation is the detector is quite sensitive to the minor errors of channel estimation. Since we use the stochastic backpropagation method, the minor errors will be enlarged in every span, after transmitted through 22 spans, the errors are big enough to effect the detection. This effect is more obvious when nonlinear phase noise dominant the system, because the nonlinear phase noise is a deterministic function, so the error cannot be eliminate. That is why the SER decreases fast when SNR becomes bigger.

4.3 Linear non-unitary channel

We will discuss performances of the MMSE and ZF equalizers of two different channels. Besides the simulation results of different $\tilde{\gamma}$ will also be mentioned.

4.3.1 simulation results

The following five figures show the simulation results of linear PDL channel. The communication system is set up with 22 fiber spans, the input power is from −10dB to 25dB, and the detector is the regular ML detector. Figure 4.7 shows the result for $\tilde{\gamma}$ equals to 0.99, the signal is dual polarization QPSK signal, the perfect knowledge of the channel is known and both ZF and MMSE equalizer are applied.

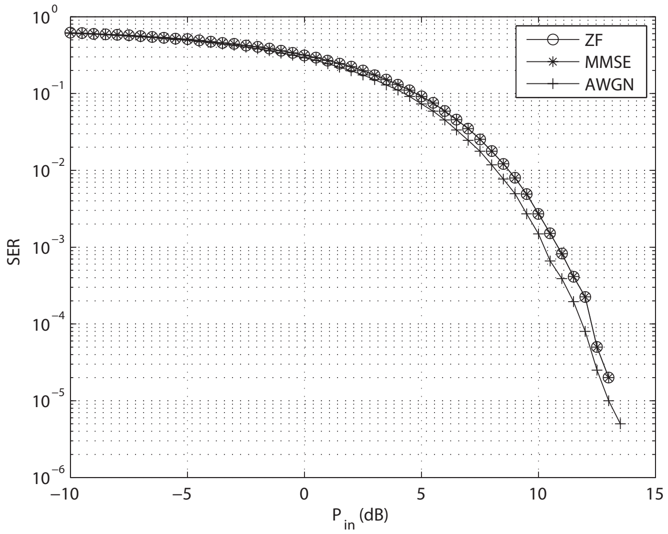

Figure 4.7: SER as a function of P_{in} for dual polarization QPSK, $\tilde{\gamma} = 0.99$.

Figure 4.8 shows the result for $\tilde{\gamma}$ equals to 0.9, the signal is dual polarization QPSK signal, the perfect knowledge of the channel is known and both ZF and MMSE equalizers are applied, but the channel is the noise-added first PDL channel which mentioned in section 3.2.

Figure 4.9 and Figure 4.10 show the results for $\tilde{\gamma}$ equals to 0.99, the signal is dual polarization QPSK signal, ZF and MMSE equalizers are applied individually with the estimation of the channel.

Figure 4.11 shows the result for $\tilde{\gamma}$ equals to 0.99, the signal is dual polarization 16-QAM signal, the perfect knowledge of the channel is unknown and ZF equalizer is applied.

4.3.2 Performance analysis

Figure 4.7 and Figure 4.8 show the performances of ZF and MMSE equalizers of the two different channels, one is the channel we use, the other is a channel for comparison. Figure 4.7 shows that MMSE and ZF equalizers have the same performance. The attenuation fact $\tilde{\gamma}$ is 0.99, so comparing with the Figure 4.8 the AWGN curve is little higher, in Figure 4.7 to maintain SER of 10^{-3} the power we need is 10.16dB in Figure 4.8 is

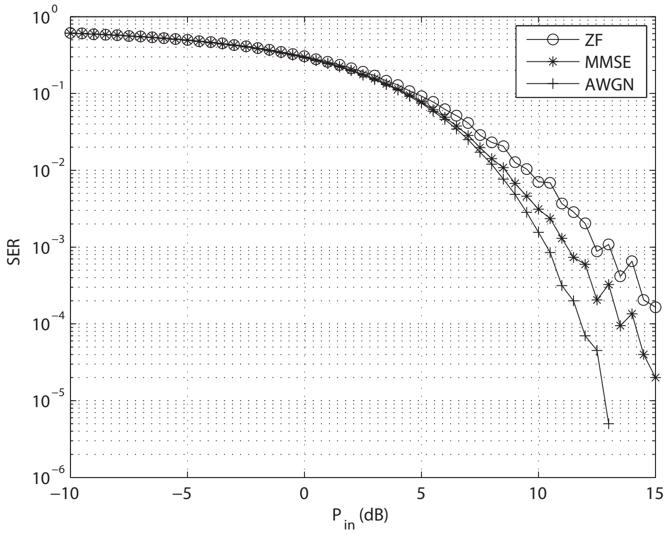

Figure 4.8: SER as a function of P_{in} for dual polarization QPSK, $\tilde{\gamma} = 0.9$.

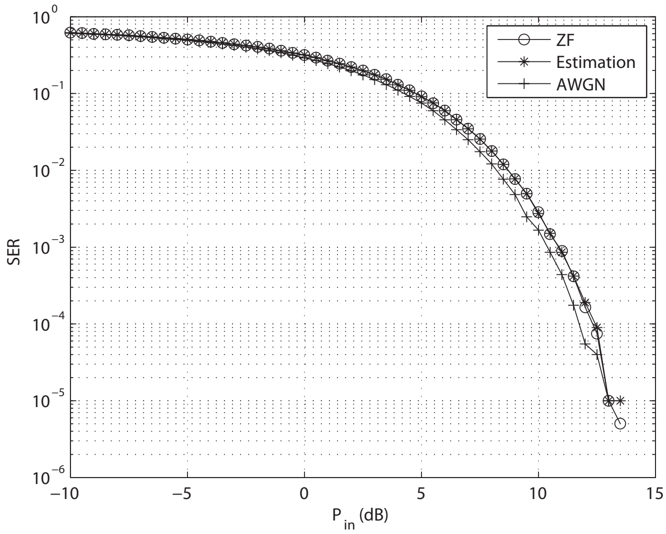

Figure 4.9: SER as a function of P_{in} for dual polarization QPSK, $\tilde{\gamma} = 0.99$.

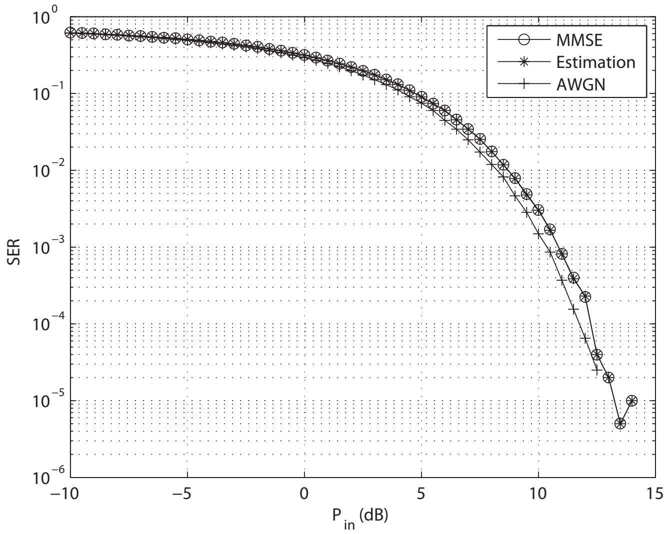

Figure 4.10: SER as a function of P_{in} for dual polarization QPSK, $\tilde{\gamma} = 0.99$.

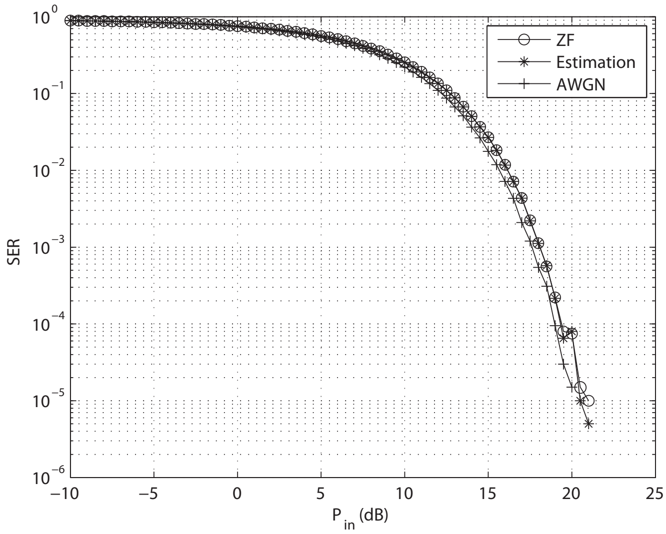

Figure 4.11: SER as a function of P_{in} for dual polarization 16-QAM, $\tilde{\gamma} = 0.99$.

10.39dB. In Figure 4.8 the MMSE equalizer has a much better performance than the ZF equalizer, this is due to the noise as we have proved in the data detection section. In both these figures, the channel is known, the equalizer is built on the perfect knowledge of the channel.

Figure 4.9 and Figure 4.10 show that on the same channel with the same estimation, the ZF equalizer and MMSE equalizer has nearly the same performance. In these situations, the channels are unknown.

Figure 4.11 shows the performance of ZF equalizer for 16-QAM signal. 16-QAM requires more power to maintain low SER, as it is more sensitive to the channel losses, the points on the constellation are more dense than QPSK. The SER of 10^{-3} requires input power of 18dB. As the ZF equalizer and the MMSE equalizer has the same performance, in the estimation, ZF equalizer is used, as it is easier to fulfill.

4.4 Non-linear non-unitary channel

In this section, the modulation format is dual polarization 16-QAM.

4.4.1 Simulation results and performance analysis

The value of $\tilde{\gamma}$ will effect the simulation results. We expect that the bigger $\tilde{\gamma}$ will give us the better performance. So we set $\tilde{\gamma} = 0.9$, $\tilde{\gamma} = 0.99$, $\tilde{\gamma} = 0.999$ and $\tilde{\gamma} = 1$ and compare the simulation results. Figure 4.12 is the SER against the input power for the system with prefect channel knowledge.

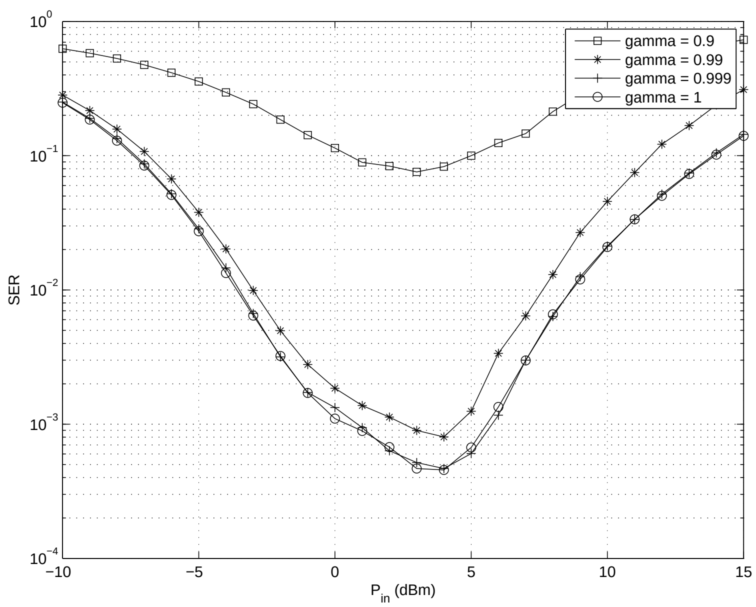

Figure 4.12: SER as a function of P_{in} with prefect channel knowledge

It is shown in the figure that our prediction is correct, when $\tilde{\gamma}$ is bigger, the performance is obviously better. When $\tilde{\gamma} = 0.9$, the performance of the detector is really bad, the lowest SER is 0.0757. When $\tilde{\gamma} = 0.99$, the lowest SER we reached is 8×10^{-4} and when $\tilde{\gamma} = 0.999$ the lowest SER is 4.7×10^{-4}, the corresponding input power is -4 dBm. When $\tilde{\gamma} = 1$, the lowest SER is 4.6×10^{-4}. Which means the performance of the detector will not increase obviously after $\tilde{\gamma}$ reaches 0.999. If we compare with the results for nonlinear noise, unitary channel system, the results of nonlinear non-unitary system is still acceptable when $\tilde{\gamma} = 0.999$.

5

Conclusion

We provide a detector of coherent optical communication system with SPM and PDL channel effect. The detection process contains two essential parts: channel estimation and data detection. To simplify the problem we also considered several transmission model with different combination of noise and channel.

5.1 Unitary channel estimation

The estimation of unitary channel is based on the channel character and Monte Carlo technique. We send training symbols to gain knowledge of the channel, use a simple method to calculate the estimation. As the results shows, the estimation method works well for both linear and nonlinear noise environments. We also discussed about the number of training symbols and the results for perfect channel estimation.

5.2 Non-unitary channel estimation

The channel estimation is suitable for a memoryless channel with AWGN noise. We have also explained the reason that why the performance of the ZF equalizer and MMSE equalizer are the same. From the simulation, we can get a reasonable result of the SER against the input power.

5.3 Data detection with backpropagation

We use factor graph to analyze the system, use particle representation to collect information, use Monte Carlo method to generate stochastic distribution. We start our research with the traditional AWGN system and then extend our method to the communication system with nonlinear phase noise and PDL channel effect. The backpropagation method gives better performance comparing to other detection method provided in this thesis when nonlinearity is considered and still has a good performance even we combine the PDL channel effect.

6

Future Work

In the previous sections, we have discussed the basic principle of backpropagation detector for a coherent optical communication system. In order to simplify the simulation, we neglect the effects such as intersymbol interference (ISI) and the noise introduced by the components of the optical fiber. In the future work, such affects should be taken into the consideration.

6.1 Channel estimation for nonlinear non-unitary system

In the last part of our thesis, we only provide the data detection method without channel estimation. In practical work, the channel information is unknown so an estimation method must be applied. Since the attenuation fact $\tilde{\gamma}$ is easy to measure, the major problem is how to estimate the unitary channels in each span.

6.2 ISI

Intersymbol interference (ISI) is a kind of signal distortion, in which the signal interferes with the subsequent signals. ISI add noise during the transmission, so the communication will be less reliable. ISI is usually caused by the multipath propagation and bandlimited channels. In this thesis, the ISI is mainly caused by the multipath propagation, as the dual polarization sends two different signals through the same fiber. The cause of this is reflection. Since these paths have the different lengths, the signal will be delayed by the reflection, and the baseband signal will spread to the sub-band. The phase of the received signal will be highly affected by these. In practice the ISI is also caused by dispersive effects such as Circular Dichroism (CD) and PMD.

The basic concept to prevent the ISI is adding the guard-band during the transmission, but these will reduce the spectrum efficiency. Another ways to fight against ISI are applying adaptive equalization and using error correcting codes. An adaptive equalizer is an equalizer can adapt the properties of the time-varying of a channel. Error correction is the correction and reconstruction of the original signal.

6.3 Other problems

As mentioned in the previous section, all the noise introduced by the components of the transmission line were neglected. So in the future work, these noises should be added. For instance, the receiver noise, there are three basic sources of noises added to a received signal, shot noise (photoelectron noise) which arises from the particle properties if the light, thermal noise (circuit noise) which aries from the random movement of the electron due to the temperature and optical noise (photon noise). All these will affect the results of the transmission.

Bibliography

[1] A. S. Tan, H. Wymeersch, P. Johannisson, E. Agrell, P. Andrekson, M. Karlsson, “An ML-based detector for optical communication in the presence of nonlinear phase noise, ”*International Conference on Communications* (2010) .

[2] I. P. Kaminow, T. Li, A. E. Willner, “Coherent optical communicaion systems, ”*Optical fiber telecommunications* 1 (5) (2010) 95.

[3] E. Ip, J. Kahn, “Compensation of dispersion and nonlinear impairments using digital backpropagation, ”*Journal of Lightwave Technology* 26 (20) (2008) 3416–3425.

[4] J. Gordon, L. Mollenauer, “Phase noise in photonic communications systems using linear amplifiers, ”*Optics Letters* 15 (23) (1990) 1351–1353.

[5] K. Ho, J. Kahn, “Electronic compensation technique to mitigate nonlinear phase noise, ”*Journal of Lightwave Technology* 22 (3) (2004) 779–783.

[6] A. Lau, J. Kahn, “Signal design and detection in presence of nonlinear phase noise, ”*Journal of Lightwave Technology* 25 (10) (2007) 3008–3016.

[7] H. Kim, A. Gnauck, “Experimental investigation of the performance limitation of DPSK systems due to nonlinear phase noise, ”*IEEE Photonics Technology Letters* 15 (320–322).

[8] L. Coelho, O. Gaete, E. Schmidt, B. Spinnler, N. Hanik, “Impact of PMD and nonlinear phase noise on the global optimization of DPSK and DQPSK systems, ”*Optical Fiber Communication Conference (OFC)* (2010) OWE5.

[9] A. Technologies, “Polarization Dependent Loss Measurement of Passive Optical Components” (5988-1232EN) (2002) 3.
URL `http://cp.literature.agilent.com/litweb/pdf/5988-1232EN.pdf`

[10] A. Technologies, “Measuring Polarization Dependent Loss of Passive Optical Components” (5990-3281EN) (2008) 2.
URL `http://cp.literature.agilent.com/litweb/pdf/5990-3281EN.pdf`

[11] Y. Cai, "MAP detection for linear and nonlinear ISI mitigation in long-haul coherent detection systems, "*Photonics Society Summer Topical Meeting Series* (2010) 42–43.

[12] X. Li, X. Chen, G. Goldfarb, E. Mateo, I. Kim, F. Yaman, G. Li, "Electronic post-compensation of WDM transmission impairments using coherent detection and digital signal processing, "*Opt. Expr.* 16 (2) (2008) 881–888.

[13] W. Shieh, H. Bao, Y. Tang, "Coherent optical OFDM: Theory and design, "*Opt. Expr.* 16 (2) (2008) 841–859.

[14] L. S. Muppirisetty, J. Karout, "Coherent optical OFDM: Theory and design (EX32) (2009) 43.