Solutions to Mitigate Cross-Phase Modulation Impairment in Coherent Fiber-Optic Communications

by

Nima Abbasi Firoozjah

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2019

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In this thesis, methods of exploiting and consequently, mitigating the nonlinearities in Cross-Phase Modulation (XPM) dominated fiber-optic communication is suggested using the model provided by Ciena Corporation. An enhanced Bit Error Rate (BER) per Signal to Noise Ratio (SNR) is obtained using the solutions provided in this work compared to the state of the art methods used in Ciena Corporation. Solutions include exploiting the properties of XPM in order to devise methods to detect the received data with less probability of error in optical communication and also to come up with new constellation designs, suitable for XPM-dominated optical channels.

Coherent optical communication and specifically, phase-modulated optical communication has received much attention due to its higher spectral efficiency and also better receiver sensitivity. However, due to the optical Kerr effect, two major phase nonlinearities are added to the phase of the received signal at the end of an optical fiber link called Self-Phase Modulation (SPM) and XPM.

This thesis is focused on XPM dominated fiber-optic links and tries to exploit the properties of XPM in order to obtain a better performance compared to the methods in which the cancellation of XPM noise is the main objective. Finally, a constellation design is discussed, considering the model we use in this thesis. Constellation design suitable for optical communication and in particular, XPM dominated fiber-optic channels, has already been a topic of vast research, but the outcome of such efforts is highly dependant on the model used to produce XPM values. Thus, our work is unique in its own way.

Acknowledgements

I must thank my supervisor, Prof. Amir Keyvan Khandani for all his help and fatherly support and the insights he has provided me with during my program. Also, I would like to thank Ciena Corporation for their technical support and for the models used in this thesis. Special thanks to Hamid Ebrahimzad and Masoud Ebrahimi for their availability to answer my questions throughout this work.

Many thanks to my friends, Takin, Shayan, Ali, Saber, and Alireza for their advice, help, and true friendship.

Last but not least, I would like to thank my amazing family for their heartwarming presence and unconditional love.

Dedication

To my caring and supportive family and my true friends.

Table of Contents

List of Tables

List of Figures

Abbreviations

16QAM 16-Quadrature Amplitude Modulation 2, 7, 11, 17, 21, 30–32, 39, 41

AGM Anti-Gray Mapping 16

APP A Posteriori Probability 16, 17

ASE Amplified Spontaneous Emission 1, 27

AWGN Additive White Gaussian Noise 2, 3, 6, 12, 27, 28, 45

BCJR Bahl-Cocke-Jelinek-Raviv 16

BER Bit Error Rate iii, 3, 11, 13, 14, 17, 22–25, 28–30

BICM Bit-Interleaved Coded Modulation 16, 17

CD Chromatic Dispersion 1

CND Check Node Decoder 16

DPSK Differential Phase-Shift Keying 1

DWDM Dense Wavelength Division Multiplexing 26

E-LEAF Enhanced-Large Effective Area Fiber 11, 22

EDFAs Erbium-Doped Fiber Amplifiers 1

FEC Forward Error Correction 26

FWM Four-Wave Mixing 1

LDPC Low-Density Parity-Check 16, 17

LLR Log-Likelihood Ratio 16, 17

MAP Maximum A Posteriori 15

MSE Mean Square Error 27, 28

NLPN Non-Linear Phase Noise 26, 27

NLSE Non-Linear Shrödinger Equation 3, 4, 28

PA Product Accumulate 20, 24

PAM Pulse-Amplitude Modulation 27

PM Phase Modulation 1, 2

PN Phase Noise 28

RS Reed-Solomon 17

SE Spectral Efficiency 16

SER Symbol Error Rate 3, 26–31

SNR Signal to Noise Ratio iii, 6, 13, 14, 22–24, 28, 30–32

SPC Single Parity Check 20

SPM Self-Phase Modulation iii, 2, 4–6, 27

SSFM Split-Step Fourier Method 4

TCM Trellis-Coded Modulation 16

TPC Turbo Product Codes 20

VND Variable Node Decoder 16

WDM Wavelength Division Multiplexing 1–3, 5

XPM Cross-Phase Modulation iii, 1–10, 12–15, 17–19, 22, 24, 26–30, 32, 33, 39, 41, 45

Chapter 1

Introduction

1.1 Overview

Currently, coherent fiber-optic signaling is widely used in the area of communication regarding the subtle fact that the word "coherent" has a different meaning here than that used in digital communication. In the literature, an optical communication system is regarded as coherent as long as there is optical signal mixing at the receiver end, even without carrier recovery at the receiver which is unlike the convention in digital communication. For example, in a DPSK optical system where there is no carrier recovery, we still consider the system to be coherent. Moreover, Phase Modulation (PM) optical communication is an example of coherent fiber-optic communication which is the case considered throughout this thesis [1].

Currently, the main technology used in high-capacity long-haul and ultra-long-haul fiber-optic links is Wavelength Division Multiplexing (WDM) [1]. To further enhance the capacity, narrower frequency spacing is used in modern WDM optical systems which gives rise to multiple nonlinear effects on the output of the fiber alongside other linear and nonlinear effects such as Chromatic Dispersion (CD) and Amplified Spontaneous Emission (ASE) [1].

Erbium-Doped Fiber Amplifiers (EDFAs) are used at the end of each fiber span in order to boost the optical signal power, making high-rate data transmission possible in optical communications. On the other hand, EDFAs affect the performance of the system in that they can intensify the nonlinear effects. ASE, Four-Wave Mixing (FWM) and XPM are among these nonlinearities [2].

XPM, which is the subject of this thesis, is generated due to the intensity dependency of the refractive index in an optical medium (the optical Kerr effect) [1]. This phenomenon is also the main culprit for SPM. XPM is then an intensity-dependent phase-shift of the signal propagating along the fiber link. When using WDM, this phase-shift depends on the intensity of the signal on neighboring channels, too. Obviously, in a PM optical communication system, it becomes more vital to account for XPM and try to mitigate it.

In the literature reviewed in the next section and in this thesis, these nonlinearities and specifically, XPM should be somehow modelled and then be analyzed. What follows, is a summary of works done in this area.

The rest of this thesis is organized as follows: In chapter 2, the XPM model that is used throughout this thesis is presented and analysed. In chapter 3 the effect of considering the shapes of the XPM clouds is investigated which is followed by the work in chapter 4 on considering the memory and also inter-polarization correlation of XPM values. Also, in section 4.1 from chapter 4, a thorough literature review on iterative decoding in digital and optical communication domain is presented and further in that chapter, a channel coding and an iterative decoding method is employed and the results are obtained and presented. Finally, in chapter 5, works in the area of constellation design for optical communication are mentioned and some constellation designs are suggested to reduce the error rate compared to our benchmark which is 16QAM. A conclusion and insights for future works are presented in chapter 6.

1.2 Literature Review

In WDM coherent optical communications, after compensation of fiber dispersion, XPM-induced nonlinear phase is summed coherently after each span which makes it the dominant phase nonlinearity in WDM systems [3, 4]. Besides, it can be interpreted intuitively and also observed analytically that unlike the case for SPM, XPM-induced nonlinear phase noise has a Gaussian distribution, statistically speaking [3, 5, 6].

One of the most commonly used methods to model XPM, as described in [1], is the pump-probe model. In this simple model, only two WDM channel are considered. This model is obtained by observing the main channel (probe) after perturbing the neighboring channel (probe). This model is used in [7, 8] to account for XPM nonlinearities and an optimum 16-ary constellation design was proposed for XPM-dominant optical channels in the sense that 16QAM is only optimum for AWGN channels.

For a large number of spans, the nonlinear phase noise can be modeled using a transformation of a Wiener process which results in different optimum constellation designs output [9], possibly not the best result for an XPM-dominated optical channel. In that work, the performance is evaluated based on Symbol Error Rate (SER) since simple Gray mapping is not going to be the optimum choice anymore. Nevertheless, other mappings can results in a better BER using channel coding techniques and soft decoding methods. In other works, the objective is to model XPM phase nonlinearity using different approaches. In [5], starting from Non-Linear Shrödinger Equation (NLSE) for optical wave propagation in an optical fiber, Beygi et al. (2012) came up with continuous-time and discrete-time models for an optical channel concluding that an optical channel can be modeled as a linear dispersive channel with AWGN and a complex scaling factor. Note that based on the assumption that XPM-induced phase noise has a Gaussian distribution and also assuming that the average phase nonlinearity is compensated, the rest can be approximated using the first term in the Taylor expansion and as a result the nonlinearity generated by XPM can be modeled as a zero-mean additive Gaussian noise. This interpretation is used in [6] and the P. Poggiolini et al. attempted to come up with an expression for the variance of the Gaussian noise that accounts for XPM. It was concluded that the variance is proportional to P_{ch}^3 where $P_{ch} = \frac{P_{tot}}{N_{ch}}$ and P_{tot} is the total power of WDM transmitted signal over all channels an N_{ch} is the number of channels.

Due to the nature of propagation in fiber-optics, XPM has memory and also the XPM on both x and y polarizations of a WDM channel show high correlation. In all previously mentioned works, the models presented for XPM lacks these properties and therefore a part of data is dismissed in the decoding process. In the next chapter, the modeled used for XPM in this thesis is presented as derived by Ciena Corporation and followed by that is an analysis on the properties of this model.

Chapter 2

Modelling SPM and XPM Nonlinearities

2.1 Discrete-time Memory Model by Ciena

Nonlinearities model in fiber as used and provided by Ciena Corporation [10, 11], in summary, is derived as described in what follows [12].

A Split-Step Fourier Method (SSFM) is used to model the nonlinearities in a fiber in an easier way than before [13]. In SSFM, a fiber with length L (in km) is subliminally divided into M segments with length $\delta = L/M$ and as a result, the whole nonlinearity introduced at the end of the fiber link is a function of nonlinearities generated at each section. The nonlinearity at the end of each section then could be obtained by a solving a simplified version of NLSE, thanks to SSFM. The length of each segment (step size, i.e., δ) should be small enough so it can be ensured that the linear and nonlinear effects affect the output, independently [12, 11].

As organized and explained in [12], an integration based Equation in frequency domain is derived for nonlinear phase noise and a discrete-time model for SPM and XPM is obtained based on that. The "additive" XPM-induced nonlinearity on the x-polarization of the transmitted data at time $t = 0$ (i.e., $A_x(0)$) can be represented as the following equation

$$\Delta A_x = SPM_1 + SPM_2 + \sum_w \left(XPM_{1w} + XPM_{2w} + XPM_{3w} + XPM_{4w}\right) \tag{2.1}$$

in which the summation is over all neighboring channels in WDM (i.e., w is any arbitrary neighboring channel) and SPM and XPM terms has the following form [10]

$$
\begin{aligned}
SPM_1 &= \sum_{m,n} C^{spm}_{m,n} A_x(m) A_x(n) A^c_x(m+n) \\
SPM_2 &= \sum_{m,n} C^{spm}_{m,n} A_x(m) A_y(n) A^c_y(m+n) \\
XPM_{1w} &= \sum_{m,n} C^{xpmw}_{m,n} A_x(m) B_x(n) B^c_x(m+n) \\
XPM_{2w} &= \sum_{m,n} C^{xpmw}_{m,n} A_x(m) B_y(n) B^c_y(m+n) \\
XPM_{3w} &= \sum_{m,n} C^{xpolmw}_{m,n} B_x(m) A_x(n) B^c_x(m+n) \\
XPM_{4w} &= \sum_{m,n} C^{xpolmw}_{m,n} B_x(m) A_y(n) B^c_y(m+n)
\end{aligned}
$$

where A_x and A_y are the streams of symbols (complex numbers corresponding to the transmitted constellation points) on the x and y polarizations of the channel under observation and B_x and B_y stand for the streams of symbols on the x and y polarizations of the neighboring channel (namely, w). Moreover, c is the complex conjugate operator. The size and entries of the C matrices used in above Equations is computed and provided by Ciena Corporation which depend on the type and length of the fiber, the level of phase noise compensation, the number of neighboring channels considered, etc. However, in order for this model to be consistent with the results of carrier recovery, we need to make sure that the average power of transmitted signal on each channel is set to 0.5 W and also remove the "conditional mean" of SPM and XPM values in order to obtain the effective values. This conditional mean can be derived analytically (will mention later on) and also be computed using Monte Carlo simulations. The procedure of deducting the conditional mean is as follows

$$
\Delta A_{ex} = \Delta A_x - E[\Delta A_x | A_x(0)] \tag{2.2}
$$

Carefully note that in Equation 2.2, $E[\Delta A_x | A_x(0)]$ is the conditional mean which is conditioned on the data transmitted at time $t = 0$. As a result, the model used for the channel, throughout this work, is as follows

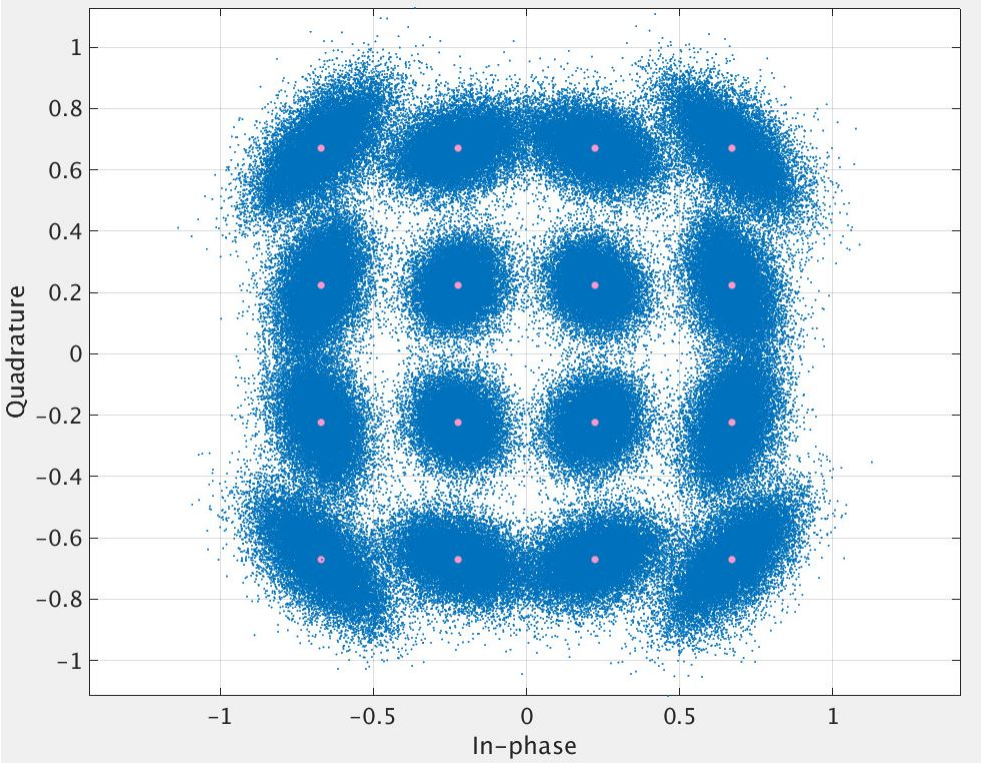


Fig. 2.1. A typical output of a 25-span long optical link at SNR $= \infty$ using 16QAM constellation design

$$Y = X + Z + \Delta A_{ex} \tag{2.3}$$

where Y, X, Z, and ΔA_{ex} are received signal, transmitted signal, AWGN, and effective phase nonlinearity added to the signal, respectively.

Since XPM-dominated optical communication is the subject of this thesis, we only consider XPM terms and discard SPM terms for the sake of simplicity. As mentioned before, conditional means can be computed both analytically and experimentally. For example, in the case of using 16-ary constellation design, we need to compute 16 conditional mean values. All told, according to Equation 2.3, a typical output of an XPM-dominated optical link at SNR $= \infty$ resembles Fig. 2.1.

Back to the analytic way of computing conditional means, by fixing $A_x(0)$ and considering only XPM terms, we have (for the sake of demonstration only XPM_{1w} is considered)

$$E[XPM_{1w}|A_x(0)] = \sum_{m,n} C_{m,n}^{xpmw} E[A_x(m)B_x(n)B_x^c(m+n)]$$

However, since $A_x(0)$ is fixed and every data point on the neighboring channel is chosen uniformly randomly from a 16QAM constellation with the average power of 0.5 W, we can conclude that $E[B_x(n)B_x^c(m+n)] = 0$ for all $m \neq 0$. This is correct due to the assumption that any constellation design used in here is symmetrical with respect to the origin and different data point are chosen independently. As a result, we have

$$\begin{aligned} E[XPM_{1w}|A_x(0)] &= A_x(0) \sum_{0,n} C_{0,n}^{xpmw} E[|B_x(n)|^2] \\ &= \frac{A_x(0)}{2} \sum_{0,n} C_{0,n}^{xpmw} \end{aligned} \tag{2.4}$$

The last equality holds, since the average power of the transmitted constellation is assumed to be 0.5 W. Finally, a similar approach can be used to come up with an expression corresponding to other XPM terms and as a result we have

$$\begin{aligned} E[\Delta A_x|A_x(0)] &= E\Big[\sum_w \big(XPM_{1w} + XPM_{2w} + XPM_{3w} + XPM_{4w}\big)\Big] \\ &= \sum_w \big(E[XPM_{1w}] + E[XPM_{2w}] + E[XPM_{3w}] + E[XPM_{4w}]\big) \\ &= A_x(0) \sum_w \Big(\sum_{0,n} C_{0,n}^{xpmw} + \frac{1}{2} \sum_{m,0} C_{m,0}^{xpolmw} \Big) \end{aligned} \tag{2.5}$$

Note that the above expression depends only on $A_x(0)$, even though we did not make any assumptions about the transmitted symbols on A_x or A_y at different time-stamps which means that fixing $A_x(m)$ for $m \neq 0$ would not have any effect on the conditional mean XPM added to $A_x(0)$. Also, since A_y terms only appear in XPM_{4w} which also have independent B_x and B_y multiplicative terms, fixing $A_y(m)$ for any m would not have any effect on the conditional mean of XPM on $A_x(0)$. However, as it will be shown later on, considering the data on the neighboring time-stamps or on the other polarization, could significantly enhance the decoding error rate.

2.2 Summary

A discrete-time, memory model for XPM-induced nonlinearity provided by Ciena Corporation is employed throughout this thesis and we model the output of the channel as expressed in Equation 2.3. The "effective" XPM added to the transmitted constellation is computed as in Equation 2.2 using the formula for the conditional mean in Equation 2.5.

This model enables us to exploit properties that previously have been overlooked, such as the memory and inter-polarization correlation of the XPM values.

Chapter 3

The Effect of Considering the Shapes of the XPM Clouds

3.1 Introduction

As mentioned before, XPM clouds are Gaussian distributed. However, the shapes of the clouds are not circular and are tilted with respect to x and y axes, meaning that the in-phase and quadrature components of the XPM nonlinearities are not necessarily equal in power and there is a correlation between them. This is even evident by the looks of the XPM clouds as shown in Fig. 2.1 and Fig. 3.1. It can be observed that the clouds of XPM corresponding to the points far from the origin are more stretched. Also, using 2D heat map shown in Fig. 3.1, we can see that the concentration of received points is higher closer to the transmitted constellation point. Another point to be mentioned is the direction of maximum variance of each cloud which is going to be exploited later on to come up with a heuristic-based constellation design.

3.2 Conditional Covariance Matrices

It is practically unjustifiable to attempt to find the mathematical expressions for the 2-by-2 covariance matrices of in-phase and quadrature components of XPM values conditioned on the data transmitted (namely, conditional covariance matrices). Alternatively, what can be done is to find these values using a Monte Carlo simulation in an offline manner. The result will be sufficiently accurate values for the constants we were looking for, quite same as the

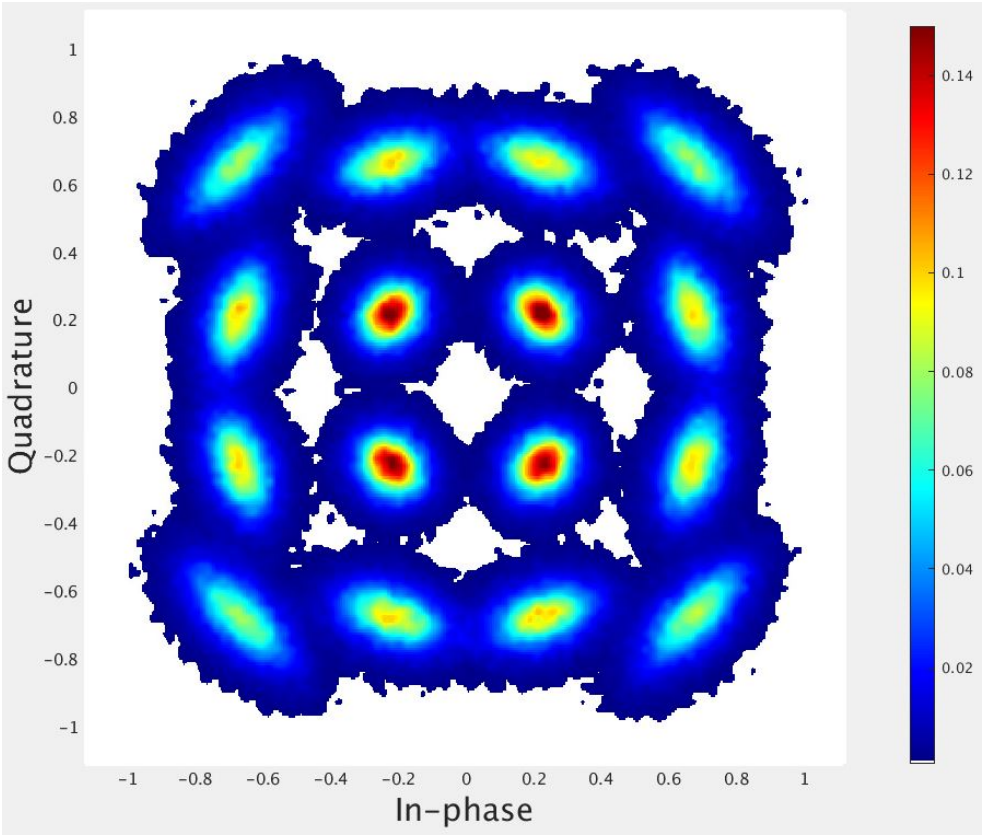


Fig. 3.1. Heat scatter plot of the output of a 25-span long optical link at SNR = ∞ using 16QAM constellation design

values for the conditional means with the mere difference that conditional mean values are easier to be computed analytically. Table 3.1 contains the conditional covariance matrices. As expected by the looks of clouds, there is a good symmetry between the conditional covariance matrices and the minor difference which could be ignored is due to the fact that in Monte Carlo simulation there is only a limited number of samples. Additionally, we can observe that the entries of the covariance matrices properly reflect the properties of the XPM clouds as seen in Fig. 3.1. For instance, the constellation points at the outer corners have higher in-phase and quadrature XPM power and also a higher covariance compared to those of the inner constellation points. Also, as mentioned before, by examining the covariance matrices in Table 3.1, it can be verified that the entries are accurate enough for the our purpose since the numbers that are supposed to be theoretically equal, are less than 5% off.

$\begin{bmatrix} 0.0105 & 0.0065 \\ 0.0065 & 0.0104 \end{bmatrix}$	$\begin{bmatrix} 0.0102 & 0.0020 \\ 0.0020 & 0.0047 \end{bmatrix}$	$\begin{bmatrix} 0.0105 & -0.0022 \\ -0.0022 & 0.0047 \end{bmatrix}$	$\begin{bmatrix} 0.0104 & -0.0064 \\ -0.0064 & 0.0104 \end{bmatrix}$
(a) Conditioned on $A_x(0) = (-3, 3)$	(b) Conditioned on $A_x(0) = (-1, 3)$	(c) Conditioned on $A_x(0) = (1, 3)$	(d) Conditioned on $A_x(0) = (3, 3)$
$\begin{bmatrix} 0.0046 & 0.0021 \\ 0.0021 & 0.0104 \end{bmatrix}$	$\begin{bmatrix} 0.0046 & 0.0007 \\ 0.0007 & 0.0047 \end{bmatrix}$	$\begin{bmatrix} 0.0048 & -0.0007 \\ -0.0007 & 0.0047 \end{bmatrix}$	$\begin{bmatrix} 0.0048 & -0.0021 \\ -0.0021 & 0.0103 \end{bmatrix}$
(e) Conditioned on $A_x(0) = (-3, 1)$	(f) Conditioned on $A_x(0) = (-1, 1)$	(g) Conditioned on $A_x(0) = (1, 1)$	(h) Conditioned on $A_x(0) = (3, 1)$
$\begin{bmatrix} 0.0047 & -0.0021 \\ -0.0021 & 0.0104 \end{bmatrix}$	$\begin{bmatrix} 0.0047 & -0.0007 \\ -0.0007 & 0.0047 \end{bmatrix}$	$\begin{bmatrix} 0.0047 & 0.0007 \\ 0.0007 & 0.0048 \end{bmatrix}$	$\begin{bmatrix} 0.0047 & 0.0022 \\ 0.0022 & 0.0105 \end{bmatrix}$
(i) Conditioned on $A_x(0) = (-3, -1)$	(j) Conditioned on $A_x(0) = (-1, -1)$	(k) Conditioned on $A_x(0) = (1, -1)$	(l) Conditioned on $A_x(0) = (3, -1)$
$\begin{bmatrix} 0.0104 & -0.0063 \\ -0.0063 & 0.0103 \end{bmatrix}$	$\begin{bmatrix} 0.0104 & -0.0020 \\ -0.0020 & 0.0048 \end{bmatrix}$	$\begin{bmatrix} 0.0104 & 0.0022 \\ 0.0022 & 0.0048 \end{bmatrix}$	$\begin{bmatrix} 0.0101 & 0.0062 \\ 0.0062 & 0.0103 \end{bmatrix}$
(m) Conditioned on $A_x(0) = (-3, -3)$	(n) Conditioned on $A_x(0) = (-1, -3)$	(o) Conditioned on $A_x(0) = (1, -3)$	(p) Conditioned on $A_x(0) = (3, -3)$

Table 3.1. Conditional Covariance Matrices: Covariance matrices of real and imaginary components of the XPM values on a fixed $A_x(0)$

3.3 Simulations

Simulations are based on the model provided by Ciena Corporation for a 25-span long Enhanced-Large Effective Area Fiber (E-LEAF) optical links considering three channels (two neighboring channels) with 90 percent dispersion compensation. Considering the case of using 25-span long fiber link, and the values computed previously for conditional means and covariance matrices, we can compare two decoding methods; closest decoding (as the benchmark) and decoding using covariance matrices. Furthermore, it should be noted that the simulations are done for an uncoded communication and in order to obtain a BER, we only consider a simple Gray mapping for 16QAM constellation as in Figure 3.2.

As mentioned earlier, the channel model is as follows

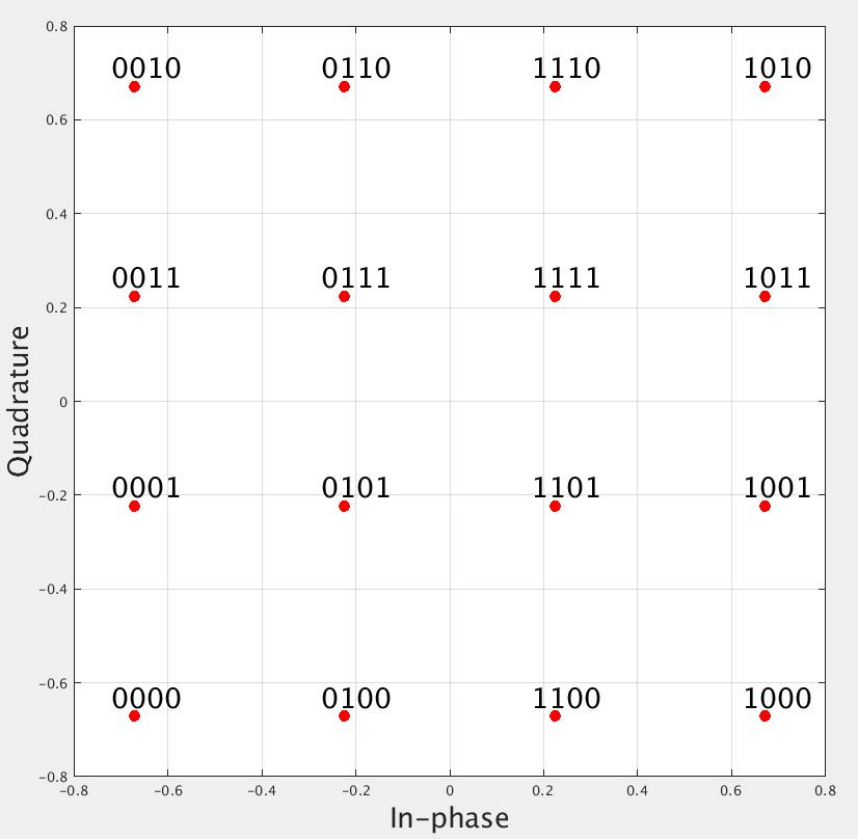


Fig. 3.2. Gray mapping used in uncoded communication for 16QAM constellation design

$$Y = X + Z + XPM \tag{3.1}$$

where Y, X, Z, and XPM are received signal, transmitted signal, AWGN, and XPM nonlinearity added to the signal, respectively. In neither of the decoding methods used here, the memory of the XPM values is considered and only the shapes of the Gaussian-distributed clouds is considered. Closest decoding is based on the assumption that the total noise is a Gaussian one with a diagonal matrix as its covariance matrix. The other method considers the actual covariance matrices and then computes ML probabilities under the assumption of Gaussian distribution of the overall noise for which the covariance matrix can be computed as follows

$$\Sigma_t = \Sigma_{cnst} + \frac{N_0}{2} I_2 \tag{3.2}$$

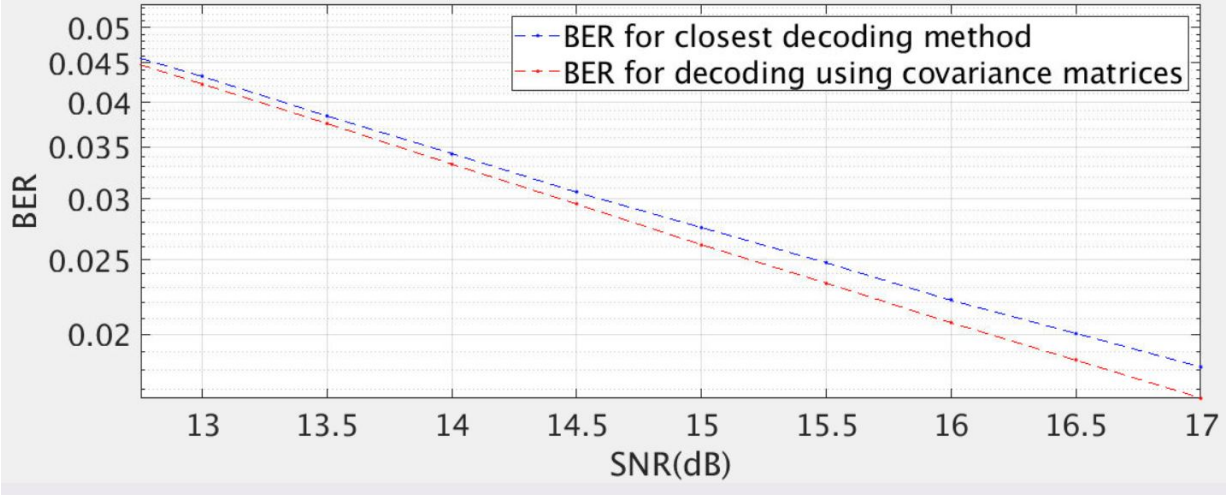


Fig. 3.3. Comparison of BER vs SNR graphs corresponding to two methods of decoding (blue curve: closest decoding, red curve: decoding using covariance matrices)

Where Σ_t is the total covariance matrix used in decoding corresponding to one of the 16 possibilities (assuming a 16-ary constellation) and Σ_{cnst} is the covariance matrix computed for the cloud corresponding to one of the 16 possibilities of the constellations at SNR $= \infty$ and I_2 is the 2-by-2 identity matrix. $N_0/2$ is the power of the additive white Gaussian noise per polarization per carrier phase (in-phase/quadrature). Due to the independence of the Gaussian noise and XPM, the total covariance matrix would be the summation of the covariance matrices of these two additive impairments.

3.4 Result

As shown in Fig. 3.3, using computed covariance matrices yields an enhanced BER for the uncoded optical communication. Especially of interest, is the BER of uncoded communication of about 0.03. This is important to us because after using channel coding and soft decoding in Ciena Corporation, the final BER will drop to the error floor at this approximate uncoded BER. The channel coding and the soft decoding method used in this thesis will be discussed in the next chapter.

Compared to closest decoding, exploiting the shapes of the XPM clouds in decoding yields, approximately, 0.15 dB improvement in the SNR to obtain an uncoded BER of 0.03 which in practice, brings about a noticeable power saving in the transmission end.

3.5 Summary

Additive XPM nonlinearity form a Gaussian distributed cloud around the transmitted constellation point. The shape and the orientation of the XPM clouds is a function of the signal transmitted. It was shown in this chapter that taking the shapes of the XPM clouds into account improves the SNR in the uncoded BER of interest by 0.15 dB which is substantial in terms of the power saved in the transmitter.

Chapter 4

Exploiting Inter-Polarization and Time Correlation of XPM Values

4.1 Introduction

Both conventional and state-of-the-art research in optical communications revolve around models that neglect the memory of XPM values and also the correlation of XPM values on the two polarizations of a channel. One of the merits of the model provided by Ciena Corporation is that it reflects the properties that could be exploited for better decoding of the received signals. The intuition behind this is based on the formula derived using a simplified model for the correlation of XPM on neighboring time-stamps (time correlation) and polarizations (inter-polarization correlation). The details are provided further in this chapter and the proofs are provided in Appendix A. This intuition led us to the suspicion that there must be a high correlation between the "effective" XPM added to the data on the two polarizations and neighboring time-stamps. In the second and third sections of this chapter (sections 4.2 and 4.3), a demonstration of these properties is provided.

Moreover, channel coding and an iterative demapping/decoding method is used to further improve the error rate. Iterative decoding has been subject to extensive research in the area of digital communication, especially after the discovery of turbo codes [14]. Iterative decoding is based on the turbo principle [15] and that's why it is also called turbo decoding. Turbo principle, in general, is formulated as follows: the receiver system, iteratively performs Maximum A Posteriori (MAP) given a priori probabilities available at each iteration and after a certain number of iterations, the hard decision output is made, based on the calculated MAP in the last iteration.

In [16], the process of iterative decoding is well explained and also, a coding and decoding system for multilevel modulation schemes is suggested. The coding system used in [16] comprises of two coding systems, i.e., a memory-4 half-rate convolutional channel encoder and a mapper (e.g., 16QAM), concatenated serially with a random interleaver in between. In each iteration, Log-Likelihood Ratio (LLR) values are computed for each bit by a demapper block and extrinsic information is then gets fed to the decoder after going through a deinterleaver matched to that used in the transmitter and then the extrinsic information from the decoder goes through an interleaver matched to the one used in the transmitter and the output is then used by the demapper as a priori information on the bits. Afterwards, the new extrinsic information is then calculated by the demapper and gets fed to the decoder again. It is insightful to mention that thanks to the random interleaver, bits in a symbol can be considered independent random binary variables. The decoder (APP calculator), employs BCJR algorithm in order to compute a posteriori probabilities which is a well established algorithm and is also used in other works including the state of the art. An abstract yet sufficient explanation on BCJR algorithm can be found in [15].

Authors in [16] and some other works that will be mentioned in the following, assume non-Gray mappings and also, Anti-Gray Mapping (AGM) which result in superior performances when employed in junction with iterative decoding, compared to Gray mapping schemes.

Additionally, an example of iterative coding for a parallel concatenated (turbo) codes can be found in [15]. In [17], LDPC is used for the purpose of channel coding and iterative decoding then is employed to decode the signal received from the channel. Again, in each iteration, extrinsic data gets passed between Variable Node Decoder (VND) and Check Node Decoder (CND) blocks as a priori information for the destination block.

As the need for faster data transmission rates have been increasing over the last two decades, high-SE modulation schemes have been used in optical communication and consequently, with the same power usage in the transmitter, the ability to decode the channel-distorted signal in the receiver with the same error rate has become more vital. Optical communication could also benefit from the achievements in digital communications mentioned earlier. During the last two decades, computationally reasonable iterative methods have been vastly investigated by scholars alongside coded modulations schemes [18] in order to achieve lower error rates in fiber-optic communications. Coded modulation helps better achieve the Shannon limit of a channel by removing the condition of separate coding and modulation blocks. A comprehensive survey on coded modulation schemes such as Bit-Interleaved Coded Modulation (BICM) and Trellis-Coded Modulation (TCM) is provided in [19] which also includes some of the publications reviewed in this thesis.

In [20], I. B. Djordjevic et al. propose a bit-interleaved LDPC-coded modulation scheme for ultra high-speed optical communication (i.e., 100 Gb/s and above) which combines the coding and modulation steps and also the multiplexing step in the case of multi-input applications. This task is done by writing (n, k)-LDPC rows of input codes from m source channels into an $m \times n$ interleaver matrix followed by a mapper fed by m column-wise chosen bits from the interleaver matrix. Decoding is done by iterating soft extrinsic LLR information between LDPC decoder and APP demapper and using perpetually refined a priori information provided in each iteration. The results of this work are compared to that of the work in [21] which are considered the benchmark in [20]. In [21], authors use direct detection as opposed to coherent detection used in [20]. Also, I. B. Djordjevic et al., in [22] have previously investigated LDPC coding for lower rate optical systems.

As mentioned earlier, higher modulation formats such as 16QAM are used in modern optical systems due to the need to exploit the capacity of the channel. However, these modulation schemes are more vulnerable to noise and nonlinearities in fiber channels. Authors in [23], propose a channel coding and an iterative decoding method similar to that in [16]. The outer encoder in this work is a $(255, 239)$-RS code and the inner encoder is a high-rate ($R_{in} = 0.8$) bit-interleaved coded modulation block, i.e., BICM which result in a total rate of $R_t \approx 0.75$. Here, the inner decoder itself is a serially concatenated encoding system consisting of a convolutional encoder, a random interleaver, and a mapping block. As a result, the decoding process is similar to that of in [16] followed by the outer decoder based on RS codes. This procedure results in a final BER of less than 10^{-15}.

A recent work in the area of optical communications employs machine learning techniques to come up with a neural network-aided BICM decoder. Authors in [19] use a feed forward neural network to obtain the channel conditional probabilities and subsequently, modify the computed LLR values for iterative decoding.

As it will be explained in details in section 4.4, channel coding and iterative decoding used in this thesis is base on the work in [24].

4.2 Correlation of XPM on $A_x(0)$ and $A_y(0)$

Due to the rigorous nature of the mathematical model we use for the XPM nonlinearity and hence for the sake of simplicity and demonstration, we only consider XPM_{1w} and XPM_{2w} representing the additive XPM terms on x and y polarizations out of the four equations (see chapter 2). For the variance of the XPM on $A_x(0)$, we have

$$\begin{bmatrix} 0.0131 & 0.0064 & 0.0045 & 0.0107 \\ 0.0064 & 0.0133 & 0.0108 & 0.0043 \\ 0.0045 & 0.0108 & 0.0133 & 0.0065 \\ 0.0107 & 0.0043 & 0.0065 & 0.0132 \end{bmatrix} \qquad \begin{bmatrix} 0.0073 & 0.0063 & -0.0034 & 0.0015 \\ 0.0063 & 0.0075 & -0.0014 & 0.0036 \\ -0.0034 & -0.0014 & 0.0075 & -0.0008 \\ 0.0015 & 0.0036 & -0.0008 & 0.0075 \end{bmatrix}$$

(a) $A_x(0) = (-3, 3)$, $A_y(0) = (-3, 3)$ (b) $A_x(0) = (-3, 3)$, $A_y(0) = (-1, -1)$

Table 4.1. 4-by-4 Conditional Covariance Matrices: Covariance matrices of real and imaginary components of the XPM values corresponding to two instances of fixed $A_x(0)$ and $A_y(0)$ (first and last two columns and rows of the covariance matrices correspond to the real and imaginary parts of the XPM on $A_x(0)$ and $A_y(0)$, respectively)

$$\sigma^2 = 0.25 \sum_{m \neq 0, n} |C_{m,n}^{xmpw}|^2 + 0.16 |A_x(0)|^2 \sum_n |C_{0,n}^{xpmw}|^2 \tag{4.1}$$

Also shown in Appendix A is that for the covariance of the XPM on x and y polarizations we have

$$\rho_{xy}^2 = 0.16 A_x(0) \overline{A_y}(0) \sum_n |C_{0,n}^{xpmw}|^2 \tag{4.2}$$

It is evident from Equation 4.1 and Equation 4.2 that the magnitude of the covariance of the XPM values on the two polarizations is relatively noticeable compared to the variance of the XPM values and depending on the data on x and y polarizations, the covariance could be even larger than the variance, itself. This means that, involving the data on y polarization must be helpful in decoding the received signal on the x polarization, hence the simultaneous decoding of the data on x and y polarizations. This fact is also conceivable from the computed values for the 4-by-4 covariance matrices by simulations for a fixed $A_x(0)$ and $A_y(0)$.

As shown in Table 4.1, the magnitude of the covariance between the real and imaginary parts of the XPM added to $A_y(0)$ and the real and imaginary parts of the XPM added to $A_x(0)$, is noticeable relative to variance of these values. As an example, we can see from both matrices in Table 4.1 that there is a high covariance between the real part of the XPM on $A_x(0)$ and the real and imaginary parts of the XPM on $A_y(0)$, almost as high as the variance of the XPM on $A_x(0)$ and $A_y(0)$. This means that the covariance matrices are far from a diagonal one and hence involving the 4-by-4 covariance matrices including the covariance between the real and imaginary parts of the XPM on $A_x(0)$ and

$A_y(0)$ could be much more beneficial compared to closest decoding or decoding using only the shapes of the XPM clouds (2-by-2 covariance matrices in the previous chapter). Also, comparing the entries of the two covariance matrices in Table 4.1, we can see that the signal transmitted on y polarization has a substantial impact on the covariance matrices. As we can see in these tables, for a fixed $A_x(0)$, we obtain two different 2-by-2 conditional covariance matrices given two different transmitted constellation point on $A_y(0)$. These observations are consistent with our expectation to see high inter-dependency between the XPM values on the two polarizations of a channel.

4.3 Time Correlation

Another property of XPM phase nonlinearity to be exploited, is time correlation. Similar to what we had for the correlation of the XPM values on synchronous transmitted data on x and y polarizations, we can inspect the correlation of XPM values on consecutive time-stamps. As we were able to come up with an expression for the variance of XPM on $A_x(0)$ (see Equation 4.1), we can derive a formula for the covariance of XPM values on consecutive transmitted data, as follows, under the assumption that only $A_x(0)$ and $A_x(1)$ are fixed

$$\rho^2_{time} = 0.16 A_x(0)\overline{A_x}(1) \sum_n C^{xpmw}_{0,n} \overline{C^{xpmw}_{0,n-1}} \tag{4.3}$$

Again, the proof for Equation 4.3 can be found in Appendix A. Similar to what we had in the case of x and y polarizations, we can compute the covariance and variances, using Monte Carlo simulations. Table 4.2 contains the 4-by-4 covariance matrices for a fixed $A_x(0)$ and $A_x(1)$. As suspected, high correlation between the XPM values on $A_x(0)$ and $A_x(1)$ can be observed. As it can be seen from the covariance matrices in Table 4.2, there is a good correlation between real and imaginary parts of the XPM on $A_x(0)$ and $A_x(1)$. Also, comparing these two tables, we can see how the constellation point transmitted on $A_x(1)$ changes the covariance matrix. This is also in full consistency with Equation 4.3 which shows the direct relationship of the covariance of the XPM values on $A_x(0)$ and $A_x(1)$ with the constellation point transmitted on A_x at these time-stamps.

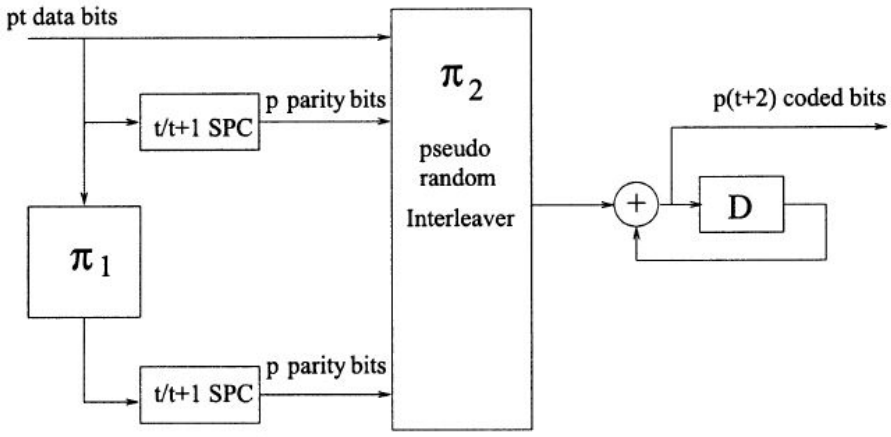


Fig. 4.1. Block diagram for PA channel coding (Credit for the image: [24])

$$\begin{bmatrix} 0.0101 & 0.0065 & 0.0061 & 0.0064 \\ 0.0065 & 0.0108 & 0.0066 & 0.0064 \\ 0.0061 & 0.0066 & 0.0101 & 0.0062 \\ 0.0064 & 0.0064 & 0.0062 & 0.0101 \end{bmatrix} \qquad \begin{bmatrix} 0.0109 & 0.0068 & -0.0025 & 0.0023 \\ 0.0068 & 0.0109 & -0.0021 & 0.0025 \\ -0.0025 & -0.0021 & 0.0048 & -0.0007 \\ 0.0023 & 0.0025 & -0.0007 & 0.0051 \end{bmatrix}$$

(a) $A_x(0) = (-3, 3)$, $A_x(1) = (-3, 3)$ (b) $A_x(0) = (-3, 3)$, $A_x(1) = (-1, -1)$

Table 4.2. 4-by-4 Conditional Covariance Matrices: Covariance matrices of real and imaginary components of the XPM values corresponding to two instances of fixed $A_x(0)$ and $A_x(1)$ (first and last two columns and rows of the covariance matrices correspond to the real and imaginary parts of the XPM on $A_x(0)$ and $A_x(1)$, respectively)

4.4 Channel Coding and Soft Decoding

A potentially effective channel coding used in this thesis is based on the work in [24]. The coding of the information bits in this work consists of a 2-D product code consisting of on Single Parity Check (SPC) coding, a pseudo-random interleaver, and a rate-1 recursive convolutional code. Product codes are also called Turbo Product Codes (TPC) since iterative (turbo) decoding has been employed to decode product codes and as a result, the coding system up until the interleaver is regarded as TPC/SPC in [24]. The block diagram of the overall encoding procedure known as Product Accumulate (PA) code is shown in Fig. 4.1 [24].

The name PA for the coding system is assigned regarding the fact that using a rate-1

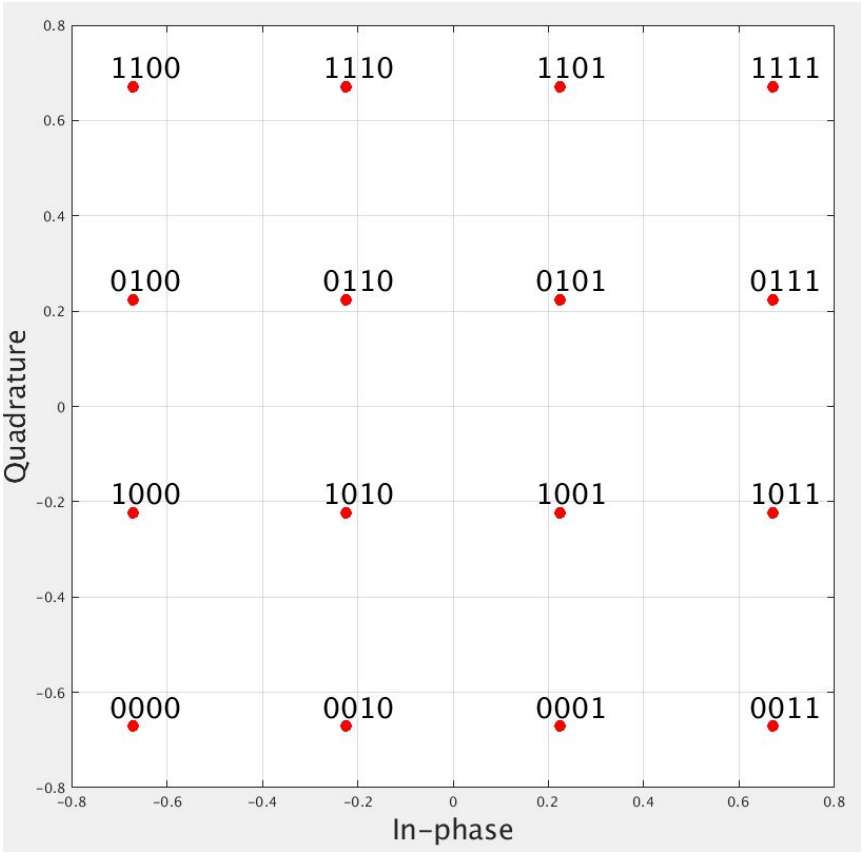


Fig. 4.2. Set-partitioning mapping used in the coded communication for 16QAM constellation design

accumulator ($1/(D+1)$ operator) after the product codes makes this system quite unique and efficient.

After the channel coding stage, the output bits will be mapped to the 16QAM symbols according to the mapping in Fig. 4.2. Set-partitioning mapping is particularly used in here because of its superior performance compared to other mappings, e.g., Gray mapping, when iterative decoding is applied.

The decoding system is performed in an iterative manner as shown in Fig. 4.3. The procedure is explained in details in [24]. The information obtained from the channel, soft L_{ch} values, are computed as follows

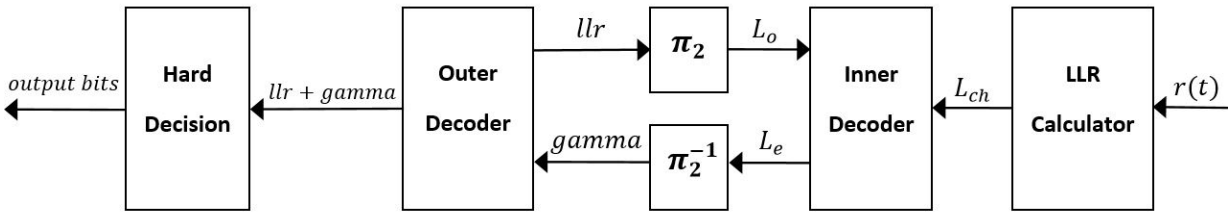


Fig. 4.3. Turbo decoder of PA codes

$$L_{ch}(y) = \ln \frac{Pr(r|y=0)}{Pr(r|y=1)} \tag{4.4}$$

However, in the light of the insight we obtained from previous sections, in order to obtain better BER-per-SNR performance, we can compute L_{ch} values more accurately with the assistance of the conditional covariance matrices. It should be noted that since we use a modulation with 4 bit/symbol spectral efficiency, each complex value r corresponds to 4 bits of channel-coded data, meaning that each received point is used 4 times in the procedure of computing L_{ch} values.

4.5 Simulation

Similar to the previous chapter, simulations are done using the model provided by Ciena Corporation for a 25-span long E-LEAF optical links considering three channels (two neighboring channels) with 90 percent dispersion compensation. Results are obtained for both uncoded and coded communications. In order to exploit the correlation in both time and polarization, the 8-by-8 covariance matrices corresponding the real and imaginary parts of the XPM on fixed $A_x(0)$, $A_x(1)$, $A_x(-1)$, and $A_y(0)$ are used in demapping. Also, to observe the comparatively higher effect of the XPM added to $A_y(0)$ on the XPM added to $A_x(0)$, the 4-by-4 covariance matrices corresponding to the real and imaginary parts of the XPM on fixed $A_x(0)$ and $A_y(0)$ is also computed and used for decoding. The model used for the channel is the same as before in Chapter 3 (see Equation 3.1). The total covariance matrices are computed in the same manner as Equation 3.2, except that instead of the 2-by-2 Identity matrix (I_2), we use the 8-by-8 or 4-by-4 identity matrix (I_8 for fixed $A_x(0)$, $A_x(1)$, $A_x(-1)$, and $A_y(0)$, and I_4 for fixed $A_x(0)$ and $A_y(0)$). Again, for uncoded communication, a simple Gray mapping is employed to obtain the BER vs SNR graph as shown in 3.2.

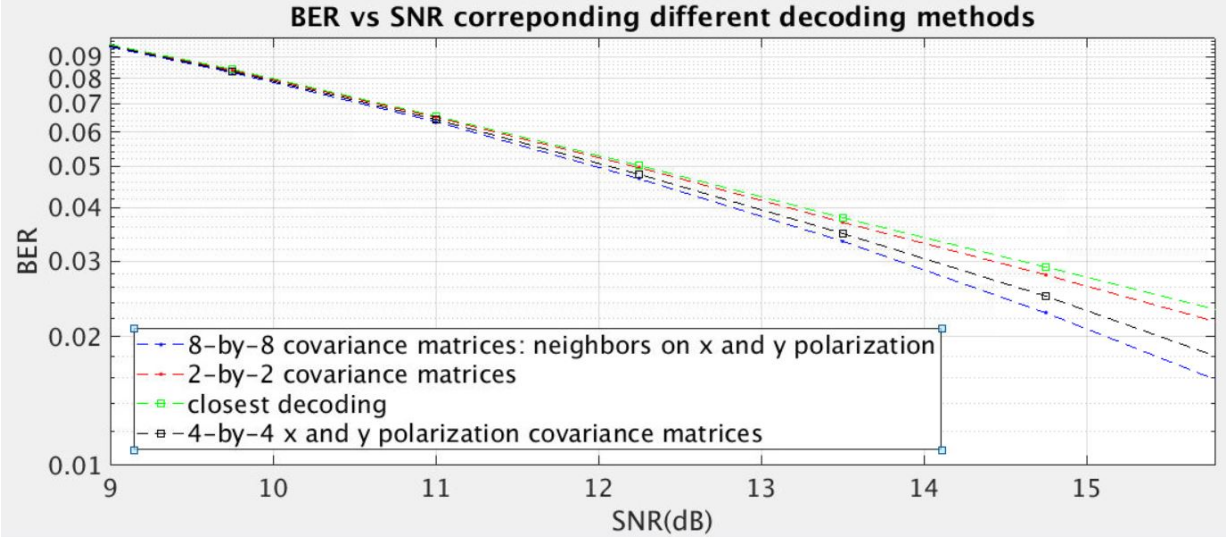


Fig. 4.4. Comparison of BER vs SNR graphs corresponding to four methods of demapping

Due to the high computational complexity of computing L_{ch} values with the assistance of 8-by-8 covariance matrices and its relatively low gain in BER, it is discarded in channel-coded communication results.

Also, in order to have a moderate time complexity for the coding and decoding processes, the number of iterations between the inner and outer decoder is set to be 10. Furthermore, based on the fact that 130 number of simulations is done, each for a 5×2^{15}-bits long information sequence, the error rate that is especially of interest is about 7×10^{-6}. This is because of the fact that at this BER, there are approximately 150 occurrences of error in the output bit sequence which is large enough to make the computed BER to be stable. The parameter t, introduced in 4.1 is set to be 5 which in result, yields a coding rate of $R = t/(t+2) \approx 0.714$ and enables us to categorize the coding system used in here as a high-rate one [24].

4.6 Results

For the uncoded communication, a noticeable improvement in the required SNR is obtained as much as 0.75 dB compared to the closest decoding by using the 8-by-8 covariance matrices at the BER of target ($BER \approx 0.03$).

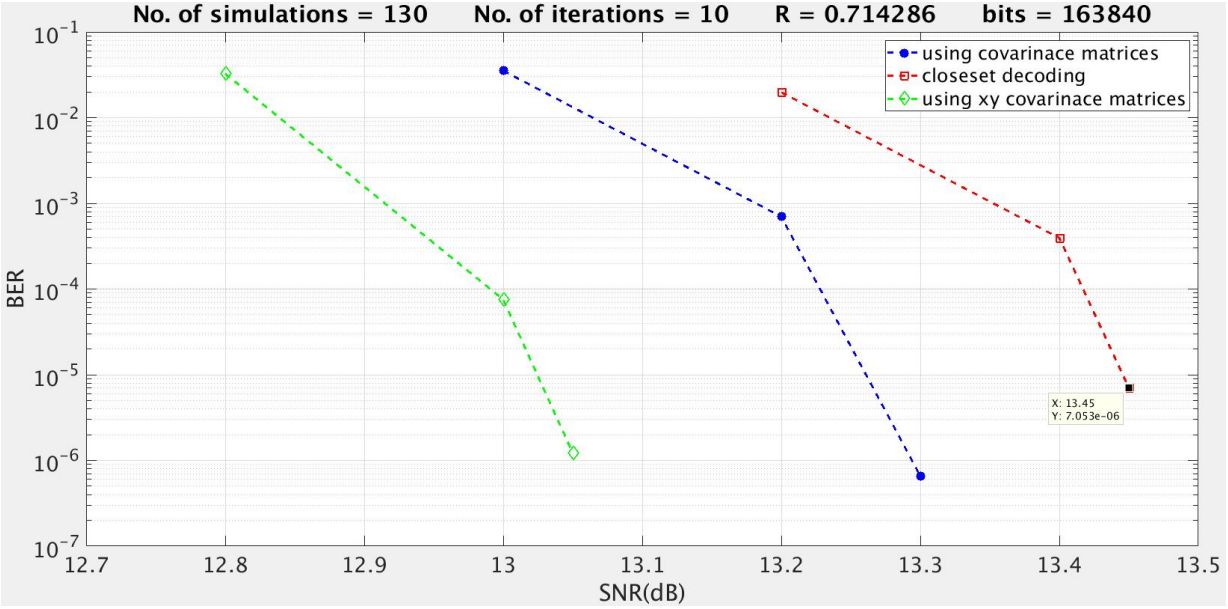


Fig. 4.5. Comparison of BER vs SNR graphs corresponding to three methods of demapping in PA-coded communication

However, for the PA-coded communication and iterative decoding results, we can observe an enhanced required SNR to obtain the error rate of 7×10^{-6} as shown in Fig. 4.5. Also, by comparing curves in Fig. 4.4 against those in Fig. 4.5, it can be observed that the SNR required in the uncoded communication to achieve the BER of 0.03 is close to that of required in PA-coded communication to achieve the BER of the error floor. Also, we can observe that at the BER of observation, there is a noticeable improvement in the required SNR by using 4-by-4 x/y-polarization covariance matrices (0.4 dB compared to closest decoding method, to say the least).

4.7 Summary

In this chapter, the correlations between consecutive samples of XPM in time and XPM samples on x and y polarizations were investigated. It was theoretically showed that there must be a high correlation of the aforementioned sort. Experimental results confirmed the suspicion and showed improvements in terms of uncoded BER-vs-SNR performance. Additionally, PA channel coding and turbo decoding were applied and alongside the findings from previous sections of this chapter, showed a significant enhance in the SNR required to

achieve a desired coded BER, compared to the benchmark, i.e., minimum-distance (closest) decoding.

Chapter 5

Constellation Design for XPM-Dominated Optical Channel

5.1 Introduction

Due to the demand of a yet higher bit rate in fiber-optic communications, spectral efficiency needs to be maximized and one way to do so is to employ constellation designs with high bit/symbol spectral efficiency (e.g., 4 or higher), alongside other methods such as Dense Wavelength Division Multiplexing (DWDM). This will result in higher susceptibility to noise and an increase in the error rate. A part of the solution to the aforementioned issue is to find constellation designs with less probability of error as well as FEC techniques. Many efforts have been made in this area and different constellation design have been suggested considering different communication channel models. Nevertheless, due to the unique nature of XPM in fiber-optics, it is essential to attempt to find an XPM-specific constellation design given any bit/symbol spectral efficiency.

Constellation design in the area of optical communication has received much attention by scholars, especially during the past decade. Different scenarios of optical channels have been considered and accordingly, various constellation designs have been obtained for 2D, 3D, and higher dimensional modulation schemes. Below, a number of the works in this area are mentioned and their assumptions and approaches are discussed.

L. Beygi et al. employed a probabilistic model of the phase and amplitude of the received signal to obtain an expression describing SER of optical channels dominated by Non-Linear Phase Noise (NLPN) [25]. Annular constellation formats are assumed beforehand and the radius and phase-shift of the rings are then optimized numerically, considering

SER as the optimization criterion. However, due to the two-stage decoding method introduced in [25], the phase-shift of the constellation rings do not affect the performance of the optical system and therefore they are not subject to optimization.

The approach introduced in [26] is employed by T. Pfau et al. in [9] to optimize different 16-ary constellation designs suitable for an AWGN channel impaired by linear phase noise. An expression for SER is derived using the channel model introduced above and then minimized numerically under the constraint of constant-power 16-ary constellation [26]. Although the method in [9] does not consider NLPN such as SPM and XPM, but the outcome is insightful and will prove beneficial for XPM-dominated optical channel, as shown later in this thesis.

In [27], ASE-dominated optical channel is considered by I. B. Djordjevic et al. and an optimum signal constellation is then obtained by minimizing the MSE of the signal constellation representing the source. Their method to achieve the optimal constellation design, in minimum-MSE sense, is an iterative one. The process starts with an arbitrary constellation of the target size (M), say M-ary QAM. Then the optimum source distribution is obtained using Arimoto-Blahut algorithm. After that, a long sequence of samples from the optimum source distribution is generated and grouped together in M clusters based on the Euclidean distance of their distance to M constellation points from the previous iteration. A point belongs to the cluster with the minimum distance from the constellation point representing that cluster. The new constellation points are obtained by computing the center of mass of the points belonging to each of the M clusters. This process is repeated until convergence or reaching a predefined maximum number of iterations. As claimed in [27], the resulting constellation point will be optimum in the minimum-MSE sense.

I. B. Djordjevic and Tao Liu et al., in [28, 29], propose a constellation design algorithm similar to that in [27]. However, in [28, 29], the design algorithm is modified to suit multidimensional modulation schemes. The difference is in the initiation method. In order to initiate the process, D-dimensional Cartesian product of PAM constellation designs is employed [28, 29].

A new approach to signal constellation design is introduced in [30]. Authors in [30], used the analogy between vector quantization of a Gaussian source and the D-dimensional constellation design for a ASE-dominated scenario.

A slightly different version of the approaches in [27] is presented in [7, 8, 31]. Considering XPM-dominated optical channel, Tao Liu et al. proposed an iterative algorithm to obtain an optimum 2D-16ary constellation design in maximum log-likelihood sense, in [7, 8]. The rest of the algorithm is similar to that proposed in [27], i.e., grouping training samples and computing new constellation points by computing the center of mass of each

group (cluster), iteratively. Moreover, in [31], authors also investigated optical channels dominated by linear phase noise and applied the same constellation design method.

Before any attempt to find a better alternative for constellation design, we need to find an easily manageable mathematical model for the output of an XPM-dominated optical channel. In [7, 8], a discrete-time memory-less *pump-probe* model is used which is introduced and explained in detail in [1]. Moreover, other discrete-time models have been obtained starting from the very Equations defining the transmission of optical wave through the fiber link, i.e., NLSE [5, 3]. Other works used a model defining a channel impaired with Phase Noise (PN) for the purpose of constellation design optimization in which the results are applicable to XPM-dominated channels and show improvement in terms of BER and SER [25, 9, 26].

Secondly, we need to decide an optimization method and criterion (an objective function). In [7, 8], received data are clustered based on maximum cumulative log-likelihood criterion and the new constellation points are computed as the average of the transmitted constellation points corresponding to the received data points in each cluster. This procedure is then repeated as many times as needed for the method to converge, up to a threshold or for the iteration reaching a predefined maximum number. More easily, the clustering method could be simply based on minimum mean square error Mean Square Error (MSE) as in [32]. Nonetheless, analytical methods could also be used. In [9], based on the work in [26], an expression for the probability of the error is derived in the presence of phase-noise and AWGN. In [25], a joint probability density function (pdf) of the normalized amplitude and compensated phase is employed to derive an expression for SER and then the constellation design is numerically optimized by minimizing SER. Also, in that work, based on the level of vulnerability of each constellation point at each SNR, a predefined constellation format is used, e.g., it is assumed that the final constellation design is supposed to be in rings of 4/8/4 points and the radius distribution of the rings is the subject of the optimization (not considering the relative phase-difference between the rings due to the robustness of the proposed decoding method).

In this work, the model provided by Ciena Corporation is used to model an XPM-dominated channel for the purpose of constellation design as elaborated in the next section.

5.2 Optimization Algorithm

The optimization algorithm is based on the intuition obtained by observing the properties of XPM clouds and the previous works mentioned in section 5.1. It is assumed that the final

constellation design is supposed to be in a predetermined multiple-rings annular format (e.g., 1/6/9 rings). It is insightful to observe that fixing one constellation point at (0,0) could be beneficial to the obtained SER under the constraint of 0.5 W average constellation power. As a result, a number of strong candidates (to be mentioned later) are chosen to be the format of the final constellation design and then the normalized radius of the second ring and the phase-difference between the two rings are the subjects of optimization.

The optimization method, however, is simple. It is based on a numerical optimization method suitable for non-differentialble objective functions which is empirical SER in this case. Since simple Gray coding is no longer the best option, which is not necessarily a disadvantage, empirical SER is chosen over BER as the optimization criterion, for the time being. Nelder-Mead Simplex method could be used as the optimization method [33]. However, low convergence-rate and getting trapped in local minima are drawbacks of this method.

Also, simulations show that the dependence between effects of the phase-difference and the normalized-radius on empirical SER is almost negligible in the optimization process. Assuming 2-rings annular constellation designs, one can justify the aforementioned observation by arguing that adding white Gaussian noise to the XPM clouds, will not change their orientations (proof in Appendix B). Additionally, since the phase-difference has no effect on the average power, it can only impact the inter-ring symbol errors (an inter-ring symbol error happens when a point from one of the rings is sent and a point from the other ring is detected at the receiver, due to the impairments). Therefore, given that $\Delta\theta_0$ is an optimum phase-difference for the normalized radius of $r_0 > 1$, it can be concluded that it is also the optimum phase-difference for any other normalized radii. Consequently, we can optimize them separately, using a uni-variant numerical optimization method such as Golden-Section search [34] which is empirically proved to be faster and more likely to converge to the optimum point.

After simulations, it was brought to our notice that in some cases, the phase-difference does not play a significant role in computing the empirical SER and henceforth, the normalized radius of the second ring is the only parameter left to be optimized. Also, it was mentioned earlier on that empirical SER is chosen as the optimization criterion. Nevertheless, we can use BER as the optimization criterion accompanied by a proper mapping rule since in practice, uncoded BER is more important than SER.

5.3 Simulation

Optimization is done for a channel model described in Equation 3.1, at SNR = 14 dB, chosen heuristically based on the fact that 16QAM constellation design yields an uncoded BER of 0.03 at an SNR close to 15 dB. Moreover, as mentioned before, empirical SER is used as the optimization criterion and approximately, 0.1 is considered as the SER corresponding to an uncoded BER of 0.03 for 16QAM constellation design. Therefore, any improvements in terms of SNR at SER = 0.1 is imperative.

Based on the works surveyed previously in this chapter, and also in order to manage the time complexity of optimization algorithm the annular formats considered for the simulations are the following two and three rings designs: 1/6/9 (1/9/6), 1/7/8 (1/8/7), 4/8/4, and 4/12. Also, another reason for choosing only two and three rings formats is the fact that the multivariate optimization algorithm used here (Nelder-Mead Simplex method) is likely to get trapped in local minima, although those local minima could be as good as the optimum point. Finally, we initialize the optimization process with reasonable values, and then the second and also maybe the third ring's normalized radii and phase-differences are optimized.

The decoding performed in computing empirical SER at each iteration of optimization is done using 2-by-2 conditional matrices introduced in chapter 3. This method was chosen over more sophisticated alternatives, due to the intuition that the shape of the constellation point has little to do with inter-polarization and time correlation of XPM values (see chapter 4).

5.4 Results

The SER of the obtained constellations design at the SNR of 14 dB is summarized in Table 5.1. As shown in the following table, 1/6/9 format has the best performance and 1/7/8 format is almost as good as that. Also, the fact that the performance-gap between 1/6/9 and 1/9/6 formats is more than that between 1/7/8 and 1/8/7 formats is in consistency with our intuition. This is because 1/7/8 format is almost similar to 1/8/7 format but on the other hand, 1/6/9 and 1/9/6 formats are more distant in appearance. Finally, we can see other annular formats have inferior performance, compared to the ones mentioned earlier. One possible reason could be the fact that 4/8/4 and 4/12 formats do not have the advantage of a constellation point in the origin.

Moreover, since the calculation of empirical SER is done using Monte Carlo simulations and the number of samples is limited, the outcome could be inaccurate. As a result, the

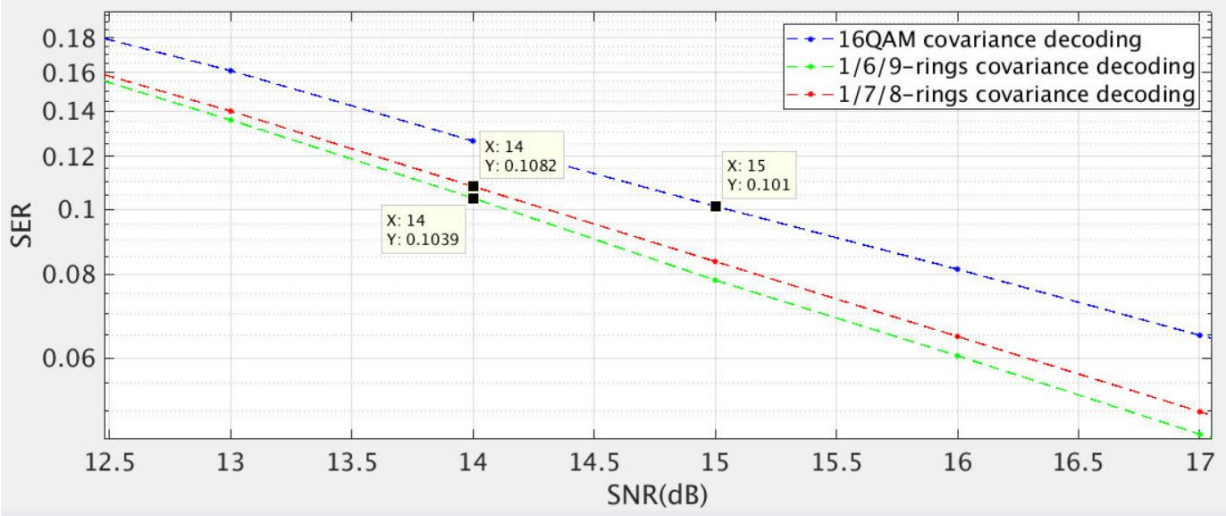


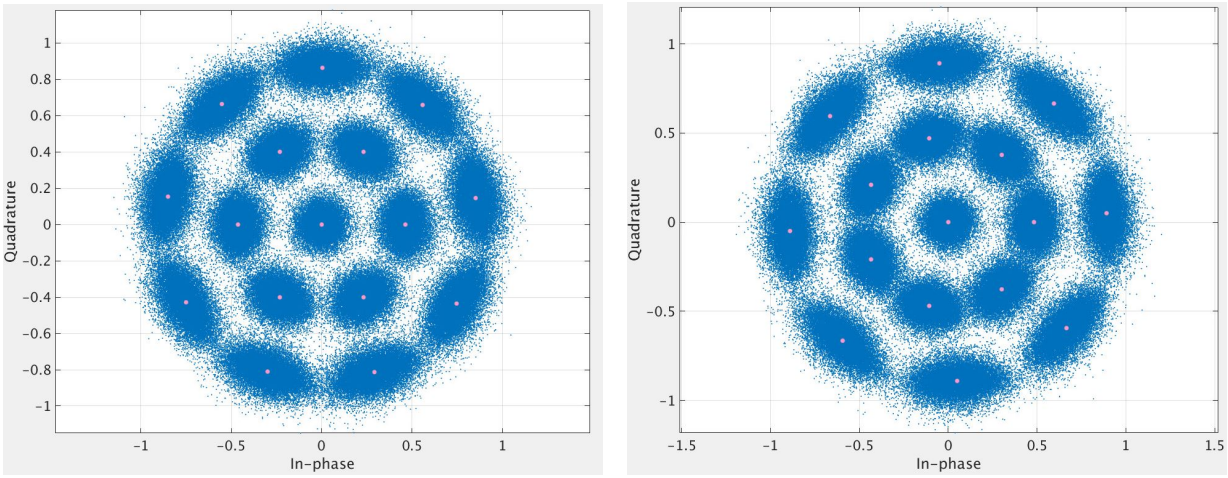
Fig. 5.1. Comparison of SER vs SNR corresponding to optimum 1/6/9, 1/7/8, and 16QAM constellation designs

Annular Constellation	Symbol Error Rate
1/6/9	0.102
1/9/6	0.135
1/7/8	0.106
1/8/7	0.116
4/8/4	0.128
4/12	0.133

Table 5.1. BER of different optimum annular constellation designs at the SNR of 14 dB

optimization process is done multiple times for each annular format and the numbers in Table 5.1 are an average of multiple outputs. Therefore, we can say the marginal difference in the performance of 1/6/9 and 1/7/8 formats is meaningful. This can also be observed in the SER-vs-SNR curves corresponding to these two annular formats.

Fig. 5.1 compares the performance of the two strongest candidates according to Table 5.1 versus the benchmark, i.e., 16QAM constellation. It can be observed in Fig. 5.1 that using the optimum 1/6/9 constellation design results in an improvement of about 1 dB in the SNR required to obtain an SER of 0.1, compared to 16QAM constellation design. Also, it can be observed that Optimum 1/6/9 constellation design marginally outperforms Optimum 1/7/8 constellation design.



(a) Optimum 1/6/9 constellation design

(b) Optimum 1/7/8 constellation design

Fig. 5.2. Output of a 25-span long optical link at SNR = ∞ using the two annular constellation designs

Fig. 5.2a and 5.2b show the output of a 25-span long fiber-optic link operating at SNR=∞ and employing 1/6/9 and 1/7/8 signal constellation designs, respectively.

5.5 Summary

In this chapter the effect of constellation design on the error rate of the optical communication system was investigated and some potentially better constellations were suggested to be used in XPM-dominated fiber-optics signaling. Based on both intuition and analytical reasoning, some annular constellation designs were considered as the format of our final solutions and then the radii of their rings and phase-differences between the rings were optimized using two numerical optimization methods, i.e., Golden-Section search and Nelder-Mead Simplex method. The resulting constellation designs show a noticeable improvement in the SNR required for the optical system to operate in accordance with standards. Finally, it should be mentioned that the 1/6/9 and 1/7/8 constellation formats, showed better performance compared to other two and three rings annular formats and also the reference 16QAM constellation design.

Chapter 6

Conclusion and Future Works

In this thesis, mitigating Cross-Phase Modulation (XPM) in fiber-optic communication was investigated and solutions have been proposed form exploiting the correlation of XPM values in time and between polarizations to suggesting constellation designs in order reduce the error rate in an XPM impaired optical channel.

Depending on the computational capacity of the optical system, exploiting memory of the XPM values of time-stamps other than immediate neighboring ones could be investigated, although it is suspected that it would not improve the performance considering the computational complexity overhead.

Additionally, finding an optimum mapping rules for different modulation schemes could be an interesting future work. Also, we performed simultaneous decoding of data on x and y polarizations in chapter 4. This could give a new impetus to 4D constellation design comprising in-phase and quadrature components of signal on the two polarizations.

References

[1] Keang-Po Ho. *Phase-modulated optical communication systems*. Springer Science & Business Media, 2005.

[2] Ajay Singh, J. M. Khurana, M tech Scholar, Shri Baba, and Mast Nath. Performance analysis of amplified spontaneous emission (ASE) noise and cross phase modulation (XPM) in optical system. 2014.

[3] Xiaojun Liang and Shiva Kumar. Analytical modeling of xpm in dispersion-managed coherent fiber-optic systems. *Opt. Express*, 22(9):10579–10592, May 2014.

[4] Shiva Kumar. *Impact of nonlinearities on fiber optic communications*, volume 7, pages 325–341. Springer Science & Business Media, NYC, 2011.

[5] L. Beygi, E. Agrell, P. Johannisson, M. Karlsson, and H. Wymeersch. A discrete-time model for uncompensated single-channel fiber-optical links. *IEEE Transactions on Communications*, 60(11):3440–3450, November 2012.

[6] Pierluigi Poggiolini, Andrea Carena, Vittorio Curri, Gabriella Bosco, and Fabrizio Forghieri. Analytical modeling of nonlinear propagation in uncompensated optical transmission links. *IEEE Photonics technology letters*, 23(11):742–744, 2011.

[7] Tao Liu and Ivan B Djordjevic. Signal constellation design for cross-phase modulation dominated channels. *IEEE Photonics Journal*, 7(4):1–8, 2015.

[8] Tao Liu, Ivan Djordjevic, and Mo Li. Signal constellation design for cross-phase modulation dominated channels. pages 1–4, 07 2015.

[9] Timo Pfau, Xiang Liu, and S Chandrasekhar. Optimization of 16-ary quadrature amplitude modulation constellations for phase noise impaired channels. In *2011 37th European Conference and Exhibition on Optical Communication*, pages 1–3. IEEE, 2011.

[10] H. Ebrahimzad. Nonlinear variance formula. Technical report, Ciena Corporation, 2017.

[11] Shahab Oveis Gharan. Nonlinearity model and compensation. Technical report, Ciena Corporation, 2017.

[12] Ali Saheb Pasand. Methods for nonlinear impairments compensation in fiber-optic communication systems. Master's thesis, University of Waterloo, 2018.

[13] Jing Shao, Xiaojun Liang, and Shiva Kumar. Comparison of split-step fourier schemes for simulating fiber optic communication systems. *IEEE Photonics Journal*, 6(4):1–15, 2014.

[14] Claude Berrou, Alain Glavieux, and Punya Thitimajshima. Near shannon limit error-correcting coding and decoding: Turbo-codes. 1. In *Proceedings of ICC'93-IEEE International Conference on Communications*, volume 2, pages 1064–1070. IEEE, 1993.

[15] Joachim Hagenauer. The turbo principle: Tutorial introduction and state of the art. In *Proc. International Symposium on Turbo Codes and Related Topics*, pages 1–11, 1997.

[16] Stephan Ten Brink, Joachim Speidel, and Ran-Hong Yan. Iterative demapping and decoding for multilevel modulation. In *IEEE GLOBECOM 1998 (Cat. NO. 98CH36250)*, volume 1, pages 579–584. IEEE, 1998.

[17] Stephan Ten Brink, Gerhard Kramer, and Alexei Ashikhmin. Design of low-density parity-check codes for modulation and detection. *IEEE transactions on communications*, 52(4):670–678, 2004.

[18] Henning Bülow and Ekaterina Masalkina. Coded modulation in optical communications. In *2011 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference*, pages 1–3. IEEE, 2011.

[19] Yuan He, Ming Jiang, Xintong Ling, and Chunming Zhao. A neural network aided approach for ldpc coded dco-ofdm with clipping distortion. In *ICC 2019-2019 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2019.

[20] Ivan B Djordjevic, Milorad Cvijetic, Lei Xu, and Ting Wang. Using ldpc-coded modulation and coherent detection for ultra highspeed optical transmission. *Journal of Lightwave Technology*, 25(11):3619–3625, 2007.

[21] Ivan B Djordjevic, Milorad Cvijetic, Lei Xu, and Ting Wang. Proposal for beyond 100-gb/s optical transmission based on bit-interleaved ldpc-coded modulation. *IEEE Photonics Technology Letters*, 19(12):874–876, 2007.

[22] Ivan B Djordjevic, Sundararajan Sankaranarayanan, Shashi Kiran Chilappagari, and Bane Vasic. Low-density parity-check codes for 40-gb/s optical transmission systems. *IEEE Journal of Selected Topics in Quantum Electronics*, 12(4):555–562, 2006.

[23] Student T Lotz, W Sauer-Greff, and R Urbansky. Iterative demapping and decoding for bit-interleaved coded modulation in optical communication systems. In *2010 12th International Conference on Transparent Optical Networks*, pages 1–4. IEEE, 2010.

[24] Jing Li, Krishna R Narayanan, and Costas N Georghiades. Product accumulate codes: a class of codes with near-capacity performance and low decoding complexity. *IEEE Transactions on Information Theory*, 50(1):31–46, 2004.

[25] Lotfollah Beygi, Erik Agrell, and Magnus Karlsson. Optimization of 16-point ring constellations in the presence of nonlinear phase noise. 03 2011.

[26] Yang Li, Shuzheng Xu, and Huazhong Yang. Design of signal constellations in the presence of phase noise. In *2008 IEEE 68th Vehicular Technology Conference*, pages 1–5. IEEE, 2008.

[27] Ivan B Djordjevic, Tao Liu, Lei Xu, and Ting Wang. Optimum signal constellation design for high-speed optical transmission. In *OFC/NFOEC*, pages 1–3. IEEE, 2012.

[28] Ivan B Djordjevic, Tao Liu, Lei Xu, and Ting Wang. On the multidimensional signal constellation design for few-mode-fiber-based high-speed optical transmission. *IEEE Photonics Journal*, 4(5):1325–1332, 2012.

[29] Tao Liu and Ivan B Djordjevic. Multidimensional optimal signal constellation sets and symbol mappings for block-interleaved coded-modulation enabling ultrahigh-speed optical transport. *IEEE Photonics Journal*, 6(4):1–14, 2014.

[30] Ivan B Djordjevic, Aleksandra Z Jovanovic, Milorad Cvijetic, and Zoran Peric. Multidimensional vector quantization-based signal constellation design enabling beyond 1 pb/s serial optical transport networks. *IEEE Photonics Journal*, 5(4):7901312–7901312, 2013.

[31] Tao Liu and Ivan B Djordjevic. Optimal signal constellation design for ultra-high-speed optical transport in the presence of nonlinear phase noise. *Optics Express*, 22(26):32188–32198, 2014.

[32] Tao Liu and Ivan B Djordjevic. On the optimum signal constellation design for high-speed optical transport networks. *Optics express*, 20(18):20396–20406, 2012.

[33] Wikipedia contributors. Nelder–mead method — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Nelder%E2%80%93Mead_method&oldid=915181932, 2019. [Online; accessed 18-October-2019].

[34] Wikipedia contributors. Golden-section search — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Golden-section_search&oldid=921719252, 2019. [Online; accessed 18-October-2019].

APPENDICES

Appendix A

Proof for Variance and Covariance Formulas

A.1 Proof for the variance formula

Here, we assume only $A_x(0)$ is fixed and the data on other data stamps on every channel is chose uniformly randomly from a 16QAM constellation design with the average power of 0.5 W. Also, as mentioned in section 4.2, the proof will be provided for the case where only XPM_{1w} and XPM_{2w} is considered as the main sources of XPM on both polarizations. Therefore, we have

$$\sigma^2 = E[XPM_x\overline{XPM_x}] - E[XPM_x]E[\overline{XPM_x}]$$

where XPM_x is assumed to be the total XPM on $A_x(0)$. However, since we know $XPM_x = XPM_{x1w} + XPM_{x2w}$, we have

$$\begin{aligned}\sigma^2 = E[XPM_{x1w}\overline{XPM_{x1w}}] + E[XPM_{x2w}\overline{XPM_{x2w}}] \\ +E[XPM_{x1w}\overline{XPM_{x2w}}] + E[XPM_{x2w}\overline{XPM_{x1w}}] \\ - E[XPM_x]E[\overline{XPM_x}]\end{aligned} \tag{A.1}$$

Note that the first two terms in equation above are basically identical as a result of the form of XPM terms. The same goes for the last two terms of the Equation A.1. For the first term (also the second one) we have

$$E[XPM_{x1w}\overline{XPM_{x1w}}] =$$

$$\sum_{m,n}\sum_{m',n'} C_{m,n}^{xpmw}\overline{C_{m',n'}^{xpmw}} E[A_x(m)\overline{A_x(m')}]E[B_x(n)\overline{B_x(n')}B_x(m'+n')\overline{B_x(m+n)}] =$$

$$\sum_{n\neq n'} C_{0,n}^{xpmw}\overline{C_{0,n'}^{xpmw}}|A_x(0)|^2 E\big[|B_x(n)|^2\big]E\big[|B_x(n')|^2\big] +$$

$$\sum_{m\neq 0,n} |C_{m,n}^{xpmw}|^2 E\big[|A_x(m)|^2\big]E\big[|B_x(n)|^2\big]E\big[|B_x(m+n)|^2\big] +$$

$$\sum_{n} |C_{0,n}^{xpmw}|^2|A_x(0)|^2 E\big[|B_x(n)|^4\big] + \dashrightarrow 0$$

Note that the rest of the terms are equal to be zero due to the symmetries of the constellation design. Given the constellation design and the average power of 0.5 W, we have

$$E[XPM_{x1w}\overline{XPM_{x1w}}] = 0.25|A_x(0)|^2\sum_{n}\sum_{n'} C_{0,n}^{xpmw}C_{0,n'}^{xpmw} + 0.125\sum_{m\neq 0,n}|C_{m,n}^{xpmw}|^2 + 0.33|A_x(0)|^2\sum_{n}|C_{0,n}^{xpmw}|^2$$

Now, using the identity equation $\sum_n\sum_{n'} C_{0,n}^{xpmw}C_{0,n'}^{xpmw} = |\sum_n C_{0,n}^{xpmw}|^2 - \sum_n |C_{0,n}^{xpmw}|^2$, we have

$$E[XPM_{x1w}\overline{XPM_{x1w}}] = 0.25|A_x(0)|^2\Big[\Big|\sum_{n} C_{0,n}^{xpmw}\Big|^2 - \sum_{n}|C_{0,n}^{xpmw}|^2\Big] + 0.125\sum_{m\neq 0,n}|C_{m,n}^{xpmw}|^2 + 0.33|A_x(0)|^2\sum_{n}|C_{0,n}^{xpmw}|^2$$

And with proper simplifications, we have

$$E[XPM_{x1w}\overline{XPM_{x1w}}] = |A_x(0)|^2\Big[0.25\Big|\sum_{n} C_{0,n}^{xpmw}\Big|^2 + 0.08\sum_{n}|C_{0,n}^{xpmw}|^2\Big] + 0.125\sum_{m\neq 0,n}|C_{m,n}^{xpmw}|^2 \tag{A.2}$$

Moreover, for the third term (also the fourth term), we have

$$E[XPM_{x1w}\overline{XPM_{x2w}}] =$$
$$\sum_{m,n}\sum_{m',n'} C_{m,n}^{xpmw}\overline{C_{m',n'}^{xpmw}} E[A_x(m)\overline{A_x(m')}]E[B_x(n)\overline{B_y(n')}B_y(m'+n')\overline{B_x(m+n)}] =$$
$$\sum_{n,n'} C_{0,n}^{xpmw}\overline{C_{0,n'}^{xpmw}}|A_x(0)|^2 E\big[|B_x(n)|^2\big]E\big[|B_y(n')|^2\big] + \cancelto{0}{\cdots}$$

Following the same procedure as seen previously, we get

$$E[XPM_{x1w}\overline{XPM_{x2w}}] = 0.25|A_x(0)|^2\Big|\sum_n C_{0,n}^{xpmw}\Big|^2 \tag{A.3}$$

However, we can easily derive the following equation from Equation 2.4, we have

$$E[XPM_{x1w} + XPM_{x2w}] = A_x(0)\sum_n C_{0,n}^{xpmw} \tag{A.4}$$

Finally, combining Equations A.1, A.2, A.3, and A.4, we get

$$\sigma^2 = 0.25\sum_{m\neq 0,n}|C_{m,n}^{xmpw}|^2 + 0.16|A_x(0)|^2\sum_n|C_{0,n}^{xpmw}|^2$$

which is the same as Equation 4.1. □

A.2 Proof for inter-polarization correlation formula

Again, the proof will be provided for the case where only XPM_{1w} and XPM_{2w} are considered. Assuming only $A_x(0)$ and $A_y(0)$ are fixed and every other data is chosen uniformly randomly from a 16QAM constellation with the average power of 0.5 W, the formula for the covariance of XPM on $A_x(0)$ and $A_y(0)$ is computed as follows

$$\rho_{xy}^2 = E[XPM_x\overline{XPM_y}] - E[XPM_x]E[\overline{XPM_y}]$$

following the same notation as in previous section, we have

$$\begin{aligned}\rho_{xy}^2 = E[XPM_{x1w}\overline{XPM_{y1w}}] + E[XPM_{x2w}\overline{XPM_{y2w}}] \\ +E[XPM_{x1w}\overline{XPM_{y2w}}] + E[XPM_{x2w}\overline{XPM_{y1w}}] \\ - E[XPM_x]E[\overline{XPM_y}]\end{aligned} \tag{A.5}$$

Again, for the first term we have

$$E[XPM_{x1w}\overline{XPM_{y1w}}] =$$
$$\sum_{m,n}\sum_{m',n'} C_{m,n}^{xpmw}\overline{C_{m',n'}^{xpmw}} E[A_x(m)\overline{A_y(m')}]E[B_x(n)\overline{B_y(n')}B_y(m'+n')\overline{B_x(m+n)}] =$$
$$\sum_{n\neq n'} C_{0,n}^{xpmw}\overline{C_{0,n'}^{xpmw}} A_x(0)\overline{A_y(0)}E\left[|B_x(n)|^2\right]E\left[|B_y(n')|^2\right] + \overset{\rightarrow}{0}$$

Similar to what we had before, we can conclude that

$$E[XPM_{x1w}\overline{XPM_{y1w}}] = 0.25A_x(0)\overline{A_y(0)}\Big|\sum_n C_{0,n}^{xpmw}\Big|^2 \tag{A.6}$$

Now, similarly, for the third term in Equation A.5, we can derive what follows

$$E[XPM_{x1w}\overline{XPM_{y2w}}] = A_x(0)\overline{A_y(0)}\Big[0.25\Big|\sum_n C_{0,n}^{xpmw}\Big|^2 + 0.08\sum_n |C_{0,n}^{xpmw}|^2\Big] \tag{A.7}$$

Finally, based on Equations A.5, A.6, A.7, and A.4, we get

$$\rho_{xy}^2 = 0.16|A_x(0)|^2\sum_n |C_{0,n}^{xpmw}|^2$$

which is the same as Equation 4.2. □

A.3 Proof for time correlation formula

In order to come up with the formula in Equation 4.2, we assume only $A_x(0)$ and $A_x(1)$ are fixed. Similar to what we had before, we have the following equations

$$\rho^2_{time} = E[XPM_x(0)\overline{XPM_x(1)}] - E[XPM_x(0)]E[\overline{XPM_x(1)}]$$

$$\begin{aligned}\rho^2_{time} = E[XPM_{x1w}(0)\overline{XPM_{x1w}(1)}] + E[XPM_{x2w}(0)\overline{XPM_{x2w}(1)}] \\ +E[XPM_{x1w}(0)\overline{XPM_{x2w}(1)}] + E[XPM_{x2w}(0)\overline{XPM_{x1w}(1)}] \\ - E[XPM_x(0)]E[\overline{XPM_x(1)}]\end{aligned} \tag{A.8}$$

And for the first term, we have

$$E[XPM_{x1w}(0)\overline{XPM_{x1w}(1)}] =$$

$$\sum_{m,n}\sum_{m',n'} C^{xpmw}_{m,n}\overline{C^{xpmw}_{m',n'}}E[A_x(m)\overline{A_y(m'+1)}]E[B_x(n)\overline{B_x(n'+1)}B_x(m'+n'+1)\overline{B_x(m+n)}] =$$

$$\sum_{n\neq n'+1} C^{xpmw}_{0,n}\overline{C^{xpmw}_{0,n'}}A_x(0)\overline{A_x(1)}E\big[|B_x(n)|^2\big]E\big[|B_x(n'+1)|^2\big] +$$

$$\sum_{n} C^{xpmw}_{0,n}\overline{C^{xpmw}_{0,n-1}}A_x(0)\overline{A_x(1)}E\big[|B_x(n)|^4\big] + \dashrightarrow 0$$

which after simplifications, equals to

$$\begin{aligned}E[XPM_{x1w}(0)\overline{XPM_{x1w}(1)}] = A_x(0)\overline{A_x(1)}\Big[0.25\Big|\sum_n C^{xpmw}_{0,n}\Big|^2 \\ + 0.08\sum_n C^{xpmw}_{0,n}\overline{C^{xpmw}_{0,n-1}}\Big]\end{aligned} \tag{A.9}$$

And similarly, for the third term in Equation A.8 we have

$$E[XPM_{x1w}(0)\overline{XPM_{x2w}(1)}] = 0.25A_x(0)\overline{A_x(1)}\Big|\sum_n C^{xpmw}_{0,n}\Big|^2 \tag{A.10}$$

And combining Equations A.8, A.9, A.10, and A.4 yields the following

$$\rho^2_{time} = 0.16 A_x(0)\overline{A_x}(1) \sum_n C^{xpmw}_{0,n} \overline{C^{xpmw}_{0,n-1}}$$

which is the same Equations as 4.3. □

Appendix B

Proof for The Effect of AWGN on the Orientation of XPM Clouds

Given that Σ is the covariance matrix of Gaussian-shaped cloud of XPM values added to a certain constellation point, the cloud's orientation is determined by the its perpendicular eigenvectors, say v_1 and v_2 with e_1 and e_2 as the corresponding eigenvalues (any real symmetric matrix has an orthonormal base of eigenvectors). Assuming that the AWGN has the power of $N_0/2$ per in-phase/quadrature component, the covariance matrix of the total noise added to the constellation point is computed as follows

$$\Sigma_t = \Sigma + \frac{N_0}{2} I_2 \tag{B.1}$$

where I_2 is the 2-by-2 identity matrix. Following the Equation B.1, we have

$$\Sigma_t v_1 = \Sigma v_1 + \frac{N_0}{2} I_2 v_1 = e_1 v_1 + \frac{N_0}{2} v_1 = (e_1 + \frac{N_0}{2}) v_1$$

which means v_1 is also an eigenvector of Σ_t and the same statements is true for v_2 meaning that the orientation of the clouds does not change with AWGN. □