Study of Phase Noise in Optical Coherent Systems

Thèse

Ramtin Farhoudi

Doctorat en génie électrique
Philosophiæ doctor (Ph.D.)

Québec, Canada

Résumé

Le bruit de phase est un problème important dans la conception de systèmes cohérents optiques. Bien que le bruit de phase soit étudié énormément dans les communications sans fil, certains aspects de bruit de phase sont nouveaux dans des systèmes cohérents optiques. Dans cette thèse, nous explorons les statistiques de bruit de phase dans les systèmes optiques cohérentes et proposons une nouvelle technique pour améliorer la robustesse du système envers le bruit de phase.

Notre première contribution traite de l'étude des statistiques de bruit de phase en présence de compensation électronique de la dispersion chromatique (CD) dans des systèmes cohérents. Nous montrons que le modèle proposé précédemment pour l'interaction de CD avec bruit de phase doit être modifié à cause d'un modèle trop simple pour la récupération de phase. Nous dérivons une expression plus précise pour le bruit de phase estimé par la récupération de phase avec décision dirigée (DD), et utilisons cette expression pour modifier les statistiques de décision pour les symboles reçus. Nous calculons le taux d'erreur binaire (BER) pour le format de transmission DQPSK semi-analytiquement en utilisant nos statistiques de décision modifiées et montrons que pour la récupération de phase idéale, le BER semi-analytique est bien assorti avec le BER simulé avec la technique Monte-Carlo (MC).

Notre deuxième contribution est l'adaptation d'une technique de codage MLCM pour les systèmes cohérents limités par le bruit de phase et le bruit blanc additif Gaussien (AWGN). Nous montrons que la combinaison d'une constellation optimisée pour le bruit de phase avec MLCM offre un système robuste à complexité modérée. Nous vérifions que la performance de MLCM dans des systèmes cohérents avec constellations 16-aires se détériorés par le bruit de phase non-linéaire et de Wiener. Pour le bruit de phase non-linéaire, notre conception de MLCM démontre une performance supérieure par rapport àune conception de MLCM déjà présente dans la littérature. Pour le bruit de phase de Wiener, nous comparons deux format de transmission, constellations carrées et optimisée pour bruit de phase, et deux techniques de codage, MLCM et codage à débit uniforme. Nos résultats expérimentaux pour BER après codage suivent les mêmes tendances que le BER simulé et confirment notre conception.

Abstract

Phase noise is an important issue in designing today's optical coherent systems. Although phase noise is studied heavily in wireless communications, some aspects of phase noise are novel in optical coherent systems. In this thesis we explore phase noise statistics in optical coherent systems and propose a novel technique to increase system robustness toward phase noise.

Our first contribution deals with the study of phase noise statistics in the presence of electronic chromatic dispersion (CD) compensation in coherent systems. We show that previously proposed model for phase noise and CD interaction must be modified due to an overly simple model of carrier phase recovery. We derive a more accurate expression for the estimated phase noise of decision directed (DD) carrier phase recovery, and use this expression to modify the decision statistics of received symbols. We calculate bit error rate (BER) of a differential quadrature phase shift keying (DQPSK) system semi-analytically using our modified decision statistics and show that for ideal DD carrier phase recovery the semi-analytical BER matches the BER simulated via Monte-Carlo (MC) technique. We show that the semi-analytical BER is a lower bound of simulated BER from Viterbi-Viterbi (VV) carrier phase recovery for a wide range of practical system parameters.

Our second contribution is concerned with adapting a multi-level coded modulation (MLCM) technique for phase noise and additive white Gaussian noise (AWGN) limited coherent system. We show that the combination of a phase noise optimized constellation with MLCM offers a phase-noise robust system at moderate complexity. We propose a numerical method to design set-partitioning (mapping bits to symbols) and optimizing code rates for minimum block error rate (BLER). We verify MLCM performance in coherent systems of 16-ary constellations impaired by nonlinear and Wiener phase noise. For nonlinear phase noise, superior performance of our MLCM design over a previously designed MLCM system is demonstrated in terms of BLER. For Wiener phase noise, we compare optimized and square 16-QAM constellations assuming either MLCM or uniform rate coding. We compare post forward error correction (FEC) BER in addition to BLER by both simulation and experiment and show that superior BLER performance is translated into post FEC BER. Our experimental post FEC BER results follow the same trends as simulated BER, validating our design.

Table des matières

Liste des tableaux

Liste des figures

Acronyms and Abbreviations

ADC	Analog-to-Digital Converter
ASE	Amplified Spontaneous Emission
AWG	Arbitrary Waveform Generator
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BICM	Bit-Interleaved Coded Modulation
BLER	Block Error Rate
BPSK	Binary Phase Shift Keying
BS	Beam Splitter
CD	Chromatic Dispersion
CM	Coded Modulation
CO-OFDM	Coherent Orthogonal Frequency Division Multiplexing
CPE	Carrier Phase Estimation
DAC	Digital-to-Analog Converter
DCE	Digital Coherent Enhancement
DCF	Dispersion Compensating Fiber
DD	Decision Directed
DQPSK	Differential Quadrature Phase Shift Keying
DSP	Digital Signal Processing
ECL	External Cavity Laser
EEPN	Equalization Enhanced Phase Noise
EVM	Error Vector Magnitude
FEC	Forward Error Correction
FIR	Finite Impulse Response
FPGA	Field-Programmable Gate Array
GLDPC	Generalized Low Density parity Check Code
HD-MSD	Hard-Decision MSD
IMDD	Intensity Modulation with Direct Detection

IIR	Infinite Impulse Response
ISI	Inter-Symbol Interference
LDPC	Low Density Parity Check Code
LO	Local Oscillator
LPF	Low Pass Filter
LW	Linewidth
MAP	Maximum A Posteriori
MC	Monte-Carlo
MLCM	Multi-Level Coded Modulation
MLSE	Maximum Likelihood Sequence Estimation
MMSE	Minimum Mean Square Error
MSD	Multi-Stage Decoder
MZM	Mach-Zehnder Modulator
NLPN	Nonlinear Phase Noise
NLSE	Nonlinear Schrodinger Equation
NRZ	Non-Return to Zero
OSA	Optical Spectrum Analyzer
OSNR	Optical Signal-to-Noise Ratio
PBS	Polarization Beam Splitter
PC	Polarization Controller
PD	Photo-Detector
PDF	Probablity Density Function
PDM-QPSK	Polarization Division Multiplexing QPSK
PLL	Phase-Locked Loop
PM	Polarization Multiplex
PMD	Polarization Mode Dispersion
POSK	Polarization Shift Keying
PT	Pilot Tone
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RF	Radio Frequency
RS	Reed-Solomon
RTO	Real-Time Oscilloscope
RZ	Return-to-Zero
SD-MSD	Soft Decision MSD

SER	Symbol Error Rate
SNR	Signal-to-Noise Ratio
TCM	Trellis Coded Modulation
VOA	Variable Optical Attenuator
VV	Viterbi-Viterbi

to my parents

Acknowledgements

During my Ph.D., I enjoyed the help, support and encouragement from many people. First, I should acknowledge invaluable guidance of my supervisor Leslie. A. Rusch during my study. We had a lot of fruitful discussions which improved a lot the research presented in this thesis. I also admire her spirit for supporting students generously.

My old friend, Mehrdad helped me a lot during my initial settlement in Quebec and our discussions contributed to my research work. I am always grateful for his help and never forget it. I should also thank Amir from whom I learnt many things and we had many fruitful discussions. I am grateful to other members of Leslie's group including A. T. Nguyen, Chul Soo Park, Siamak and Yousra. My special thanks go to A. T. Nguyen for the experimental works that we did together.

I found many friends in Quebec City and I had many memorable moments with them. Particularly I never forget memorable moments that I had with Samim. My special thanks go to him for all the joyful moments.

My deepest gratitude goes to my parents. I am so lucky to have such caring, supportive and cultivated parents. I owe all of my achievements, if any, to them. My final gratitude goes to my brother whose appetite for learning is always admirable to me.

Foreword

Three chapters of this thesis are based on three journal papers. Most of the content in these three chapters is the same as papers. However, some modifications are made in the introduction of the papers to make the chapters more coherent. Occasionally, section titles are changed and some figures are modified.

Paper 1 : R. Farhoudi, A. Ghazisaeidi, and L. A. Rusch, “Performance of carrier phase recovery for electronically dispersion compensated coherent systems," *Optics Express*, vol. 20, no. 24, pp. 26568-26582, Nov 2012.

This is a paper for analysis of phase noise statistics in presence of chromatic dispersion. The problem was defined by me, and was elaborated in discussions with L. A. Rusch. I did theoretical analysis by myself and found the modified PDF. I also did extensive simulations to validate the analysis. In doing simulations, I occasionally consulted with A. Ghazisaeidi for help. I wrote the manuscript by myself and it was reviewed by the other authors before submission. My supervisor, L. A. Rusch, also provided valuable suggestions in modification of the final manuscript.

Paper 2 : R. Farhoudi, and L. A. Rusch, “Multi-level coded modulation for 16-ary constellations in presence of phase noise," *Journal of Lightwave Technology*, vol. 32, no. 6, pp. 1159-1167, Nov 2013.

Application of multi-level coded modulation (MLCM) to phase noise limited coherent systems is examined in this paper. The idea of applying MLCM to phase noise optimized constellations was suggested by L. A. Rusch. I did the initial study and proposed the method for finding optimal set-partitioning and code rates. I performed all simulations including both Wiener and nonlinear phase noise limited systems. The idea of applying our method for designing MLCM system to a system with nonlinear phase noise was suggested by Prof. A. Bononi. I wrote the manuscript myself and my supervisor, L. A. Rusch, reviewed it and suggested modifications.

Paper 3 : R. Farhoudi, A. T. Nguyen, and L. A. Rusch, “Experimental verification of multi-level coded modulation for 16-ary constellations," vol. 26, no. 17, pp. 1774-1777, *Photonics Technology Letters*, Jul 2014.

Experimental verification of the MLCM design proposed in the previous paper is examined in this paper. I did the measurements with A. T. Ngyuen. I used a set-up prepared and maintained by A. T. Ngyuen. I designed the MLCM and uniform rate codes applied to both phase noise optimized and square 16 quadrature amplitude modulation (16-QAM) using my MATLAB code. I also wrote the code for offline processing of the experimental data and found post FEC BER curves. I wrote the manuscript and it was reviewed by the authors before submission.

Introduction

0.1 Motivation

The optical fiber communication has witnessed tremendous advances since its deployment in 80's. Increasing data traffic and demand for faster and more secure communication play an important role in pushing this advancement. Today, demand for higher bit rates is overwhelming due to the emergence of broadband application services like video streaming and huge increase in mobile data traffic. As an example, the exponential growth of data network traffic in US is shown in Fig. 0.1. It can be seen that the rate of growth in bit rate is about 60 percent each year. According to Cisco visual networking index (VNI) forecast, shown in Fig. 0.2, there will be thirteen times more data traffic around the globe in 2017. Last year's global mobile data traffic was twelve times the size of the entire global Internet in 2000 [2].

In order to meet this demand for the increasing traffic, novel solutions in optical communications are needed. Novel technologies are being developed currently, and some are already commercialized. Optical coherent detection with the aid of digital signal processing (DSP) is a promising solution. Successful deployment of 100 Gb/s coherent systems using polarization-division multiplexing quadrature phase shift keying (PDM-QPSK) confirms the trend toward coherent systems in future optical communication systems. Fig. 0.3 shows the required optical signal to noise ratio (OSNR) for 10 Gb/s intensity modulation with direct detection (IMDD), 40 Gb/s differential QPSK (DQPSK), 100 Gb/s PDM-QPSK coherent system and future 400 Gb/s systems. The upgrade from 10/40 Gb/s systems by increasing baud rate to the 100 Gb/s system can be well understood from this figure as this upgrade would require much larger required OSNR than 100 Gb/s coherent system. Today, one of the trends in long-haul optical communications is to upgrade the existing 10 Gb/s systems with 100 Gb/s while maintaining the already installed links. This upgrade should also consider the limitations imposed by the already installed links. For example, the upgrade needs 10 times improvement in OSNR sensitivity and 100 times improvement in chromatic dispersion (CD) tolerance [12]. From Fig. 0.3, it can be seen that the tolerable penalty due to transmission effects is decreased by upgrading to 100 or 400 Gb/s systems. In this regard, DSP techniques play crucial role in the future optical communication systems to have acceptable penalty in the system. The objective of the research in this thesis is to study and develop DSP techniques for two important impairments

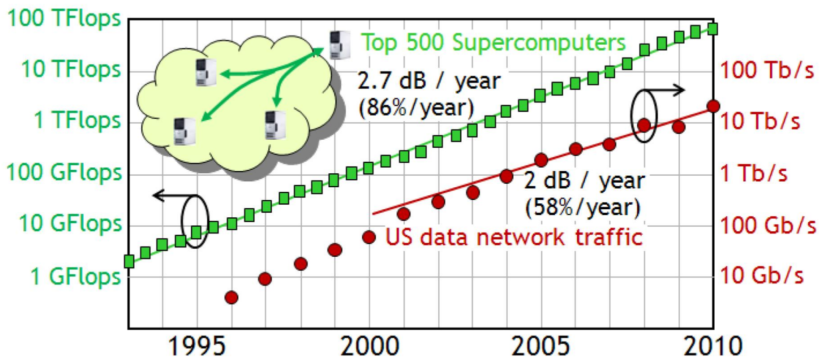


FIGURE 0.1 – Exponential growth of data network traffic in US (red circle) and the processing power (average of top 500 supercomputers). Flop designates for floating point operation per second [1].

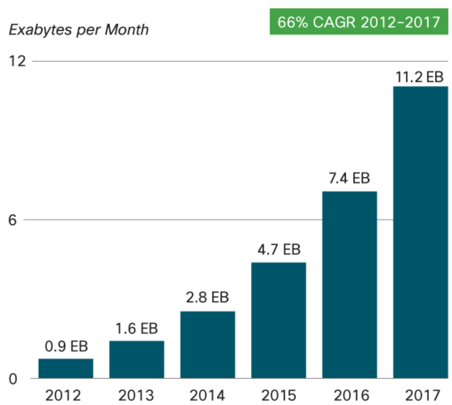


FIGURE 0.2 – Forecast for growth in mobile data traffic [2].

in optical coherent systems : CD and phase noise.

Although both CD and phase noise exist in IMDD systems and they have been the subject of heavy research, their analysis and compensation is different in coherent systems. Compensation of huge amounts of CD using DSP is unique to coherent systems as in IMDD systems optical compensation typically is employed. Phase noise is much more important in coherent systems as phase of the electrical field could carry information bits. Both CD and phase noise are estimated and compensated in the DSP of the coherent system. For phase noise, there exist several techniques to estimate and compensate as phase noise has been studied a lot in radio communications. However, approaches to cope with phase noise in fiber optics are different to some extent. This is due to the higher levels of phase noise due to lasers. In addition, interaction of phase noise with dispersion and nonlinearity is something unique to optical coherent systems. Study of these unique aspects of phase noise and dispersion is the motivation for the research in this thesis. In the following sections, basic concepts relating

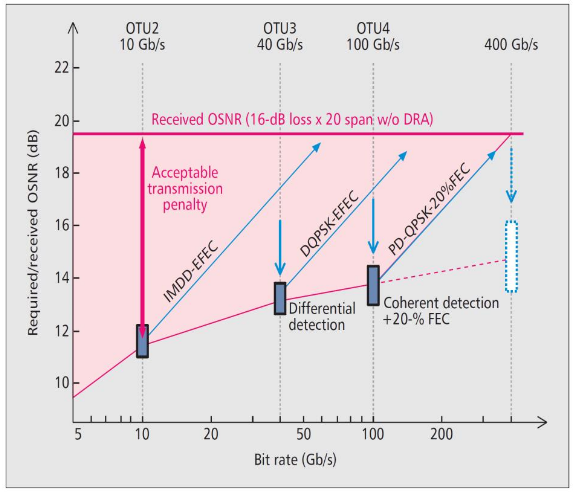


FIGURE 0.3 – Required OSNR vs. bit rate for different systems [3].

to coherent systems and DSP techniques for compensating phase noise and dispersion are reviewed. Then, major contributions and structure of the thesis are presented.

0.2 Background

In this section, first concepts of optical coherent systems are reviewed and its system model is presented. Then, two DSP functions, namely phase noise estimation and CD compensation, playing an important role in the thesis, are described in detail. Finally, a literature review of coded modulation techniques which is related to the subject of chapters 4 and 5 is presented.

0.2.1 Model of an optical coherent system

A schematic of the optical coherent system is shown in Figure 0.4. In the transmitter, the transmit laser is modulated by using Mach-Zehnder modulators (MZMs), pulse carvers and beam splitters (BS). In Figure 0.4, the modulation is assumed to be polarization multiplex (PM) binary phase shift keying (BPSK) so that the orthogonal polarizations of the input electrical field are separated by polarization BS (PBS) and then each signal polarization is phase modulated by the data using an MZM. Pulse carvers employ MZMs modulated with a sinusoid at the electrical input. They are used to generate return-to-zero (RZ) pulses with a desired pulse width (determined by the frequency of the sinusoid). The modulator of higher modulation formats like QPSK is sometimes called an IQ modulator. It consists of two modulators in two arms, with a 90 degree phase shift between them. These transmitters are shown in Figure 0.5. The electrical field of the transmitted signal can be written as :

$$\mathbf{E}_{tx} = \sum_k \mathbf{x}_k p(t - kT_s) e^{(\omega_s t + \phi_t(t))} \tag{1}$$

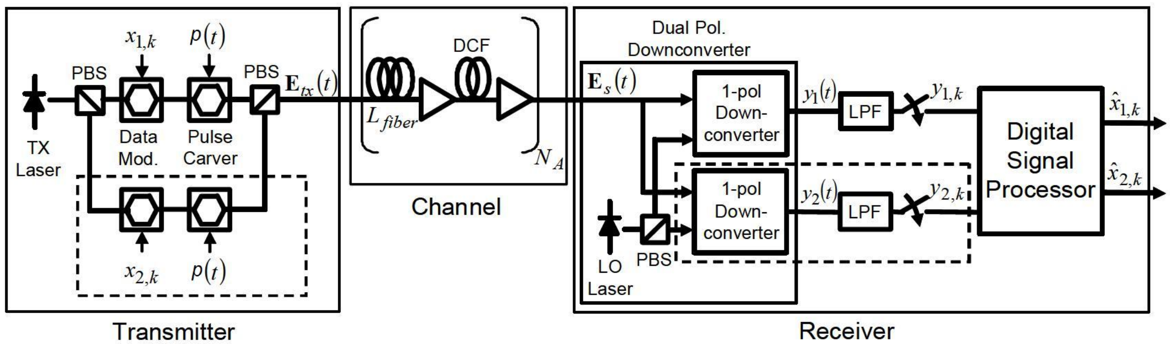


FIGURE 0.4 – Schematic of an optical coherent system [4]

where $\mathbf{x}_k$ is a 2×1 vector representing the k^{th} symbol (one symbol for each of the polarization states) and $p(t), T_s, \omega_s$ and $\phi_t(t)$ are the pulse shape, symbol interval, optical carrier frequency and the phase noise of the transmit laser respectively. The signal is passed through the fiber channel having N_A spans. Propagation of the electric field in the fiber is modelled by solving the vectorial nonlinear Schrodinger equation (NLSE) which can take into account all the effects of dispersion, loss, nonlinearity and polarization mode dispersion (PMD) together [13]. While not shown in Figure 0.4, the amplified spontaneous emission (ASE) noise due to the amplifiers is added after each span. This noise should be added as a complex additive white Gaussian noise (AWGN). Although addition of the noise after each span is the most accurate approach, the computational burden for solving NLSE is increased considerably. It is possible to use the noise loading approximation where the total ASE is added at the end of the spans [14].

The propagated signal is fed into the receiver where the incoming electrical field beats with the local oscillator (LO) laser and is down converted to baseband. The down conversion can be performed using a homodyne or heterodyne structure [4]. In this thesis, we use a homodyne receiver in which LO and transmit laser have the same frequencies and their beating results in a baseband electrical signal. In the homodyne structure, after beating of the in-phase and quadrature components of the incoming electrical field by the LO using 3 dB couplers, the signals are detected using balanced photodetectors (PDs). In Figure 0.4, each output of the down converter, $y_i(t)$, is a complex valued signal resulting from the addition of the in-phase and quadrature phase signals. We also assume an asynchronous coherent receiver in this thesis, i.e., a receiver detecting the signal based on the envelope of the signal and without any optical phase tracking [15].

After down conversion, the signal is low pass filtered and then sampled by ADCs and the digital data is sent to the DSP unit for processing and decision. The schematic of compensation of different impairments happening during the propagation. The schematic of

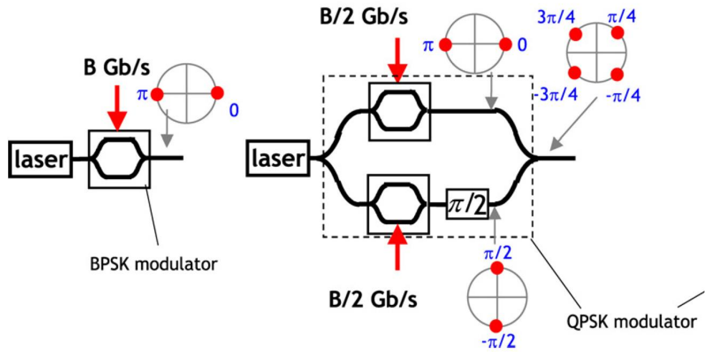


FIGURE 0.5 – BPSK and QPSK transmitters [5]

the DSP unit for a coherent receiver is shown in Figure 0.6. Only the main units of DSP are shown and given in their order of application, other DSP units may be added to enhance performance. Various algorithms exist in the literature for DSP functions in Figure 0.6 [16]. As an example, digital filters for compensation of chromatic dispersion (CD) and PMD are explained in [6, 17]. Phase estimation methods to compensate for phase noise are discussed in [18, 7]. We will discuss more about the algorithms for dispersion compensation and phase estimation in the following sections. Even compensation of nonlinearities is possible in DSP using a method called back propagation discussed in [19, 20]. An important aspect of all algorithms is that they should be parallelizable. This is due to the mismatch between ADC sampling rate and clock frequency of available DSPs. Frequency of the digital samples coming from ADC is in the range of 10 to 40 GS/s while the clock frequency of the DSP is in the range of 100 to 800 MHz. Due to this mismatch, we must use multiplexers and demultiplexers to do parallel processing in DSP. For example, if one uses a demultiplexer of $1 : m$ (so that the DSP clock frequency is $1/m$ times the sampling frequency), in the n^{th} DSP operation only results of operations $n - m$ and after can be used.

This section provides background on optical coherent systems. Specifically, the system model of a coherent system is used in chapter 2 where analysis is based on a simplified model of the system described in this section.

0.2.2 Compensation algorithms for dispersion and phase noise

In this section, we present a literature review of the algorithms used in two important units of DSP in Figure 0.6 ; CD and carrier phase estimation.

CD Compensation

Fiber dispersion is a linear impairment that must be compensated in fiber-optic communications. Fiber dispersion can be separated as CD and PMD. The transfer function of the CD

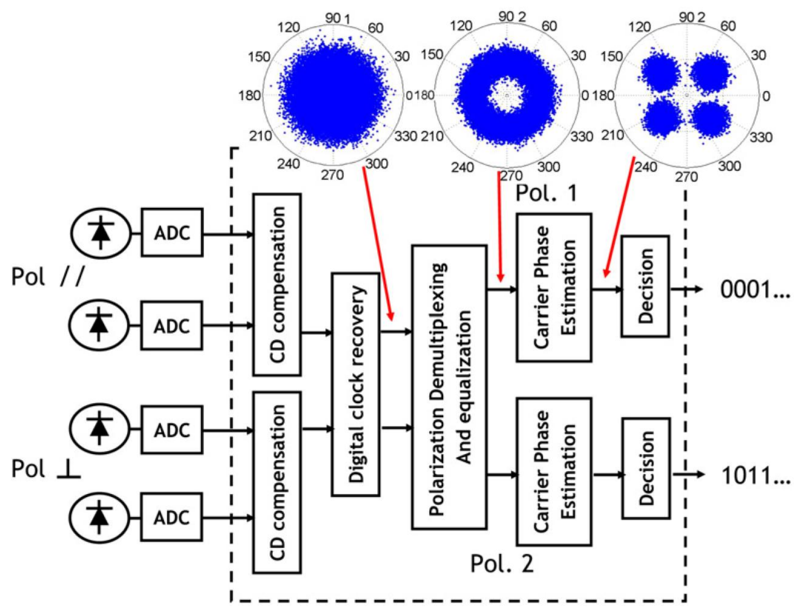


FIGURE 0.6 – Schematic of the DSP unit in a coherent receiver. Here there are two polarizations [5]

in the frequency domain can be written as :

$$H_{CD}(j\omega) = \exp\left(-j\frac{1}{2}\beta_2 L\omega^2 - j\frac{1}{6}\beta_3 L\omega^3\right) \tag{2}$$

where β_2 and β_3 are the second and third order dispersion coefficients, L is the fiber length and ω is the angular frequency (here carrier frequency ω_s is set to zero). The CD broadens the pulses so that it causes intersymbol interference (ISI). Another dispersion impairment, PMD, is caused by the random birefringence of the fiber. In contrast with CD, PMD is not a deterministic dispersion. It can be shown that PMD can be modelled by dividing the fiber into segments whose birefringence is constant but random ; the total transfer function of the fiber is given by the product of the matrices of each segment [13]. PMD can be modelled as a random delay in the output pulses. The probability density function (pdf) of this delay is Maxwellian and its mean is proportional to the square root of the fiber length L.

Traditionally, optical techniques for CD compensation are used in the 10 Gb/s and 40 Gb/s (direct detection) systems by employing dispersion compensating fibers (DCFs), fiber Bragg gratings and other techniques. In coherent receivers, DSP compensation can be employed so that these optical procedures can be reduced or totally removed. We begin by discussing CD compensation when ASE is absent and then discuss CD equalization in the presence of ASE. When there is no ASE noise, the equalizer for compensating CD should invert the transfer function given in (2). If we neglect third order dispersion ($\beta_3 = 0$), filter taps can be found analytically. The continuous time filter can be found using the inverse Fourier transform of

(2) which gives the impulse response of the channel. This is given by [17] :

$$h_{CD}(t) = \sqrt{\frac{c}{jD\lambda^2 L}} \exp\left(j\frac{\pi c}{D\lambda^2 L}t^2\right) \tag{3}$$

where c and λ are the light velocity and wavelength and $D = -(2\pi c/\lambda^2)\beta^2$. We should inverse the sign of β_2 and adopt this continuous time impulse response for use in the discrete time domain (assuming proper sampling to prevent aliasing). The resulting discrete finite impulse response (FIR) filter taps are given by [17] :

$$w_n = \sqrt{\frac{jc{T_s}^2}{D\lambda^2 L}} \exp\left(-j\frac{\pi c{T_s}^2}{D\lambda^2 L}n^2\right), -\left\lfloor\frac{N}{2}\right\rfloor \leq n \leq \left\lfloor\frac{N}{2}\right\rfloor \tag{4}$$

$$N = 2\left\lfloor\frac{|D|\,\lambda^2}{2c{T_s}^2}\right\rfloor + 1$$

where T_s is the sampling time and $\lfloor x \rfloor$ represent the largest integer value smaller than x. The value of N specified avoids aliasing. We can further reduce the number of taps in this filter by using windowing, thus reducing complexity but also reducing performance. Infinite impulse response (IIR) filters are also proposed for CD compensation [21]. These filters have considerably smaller number of taps compared with FIR filters, but their implementation is harder due to feedback.

In the presence of ASE noise, the filter given by (4) is not optimal. In this case, the equalization methods based on minimum mean square error (MMSE) can be exploited [22]. The equalization can be done either in the time or frequency domain [16], [6]. Although in principle it is possible to do equalization having one sample per symbol, this requires a matched filter before the ADC which is difficult to realize and in addition it is vulnerable to sampling time jitter. Due to these reasons, in practice oversampling of the optical signal is done by the ADCs. It can be shown that the minimum oversampling rate to compensate arbitrary amount of dispersion is 3/2 [6].

We now explain the time-domain MMSE equalizer. For simplicity of the formulation, we only consider one polarization but the extension to two polarizations is straightforward. The m^{th} sample of the received signal can be written as (in the absence of PMD and nonlinearity) :

$$y_m = \sum_k x_k q\left(mT - kT_s\right) + n'(mT) \tag{5}$$

where $q(t) = p(t) * h_{ADC}(t) * h_{CD}(t)$ and $n'(t) = h_{ADC}(t) * n(t)$ are the received dispersed pulse and filtered noise ($h_{ADC}(t)$ is the impulse response of the filter modelling the ADC front end). $T = T_s/S$ where S is the oversampling rate. A vector containing N adjacent samples (N must be larger than the system memory) of the m^{th} sample is formed to be used for filtering :

$$\mathbf{y}_m = \begin{bmatrix} y_{Sm+\lfloor\frac{N}{2}\rfloor} & y_{Sm+\lfloor\frac{N}{2}\rfloor-1} & \cdots & y_{Sm-\lfloor\frac{N}{2}\rfloor+1} & y_{Sm-\lfloor\frac{N}{2}\rfloor} \end{bmatrix}^T \tag{6}$$

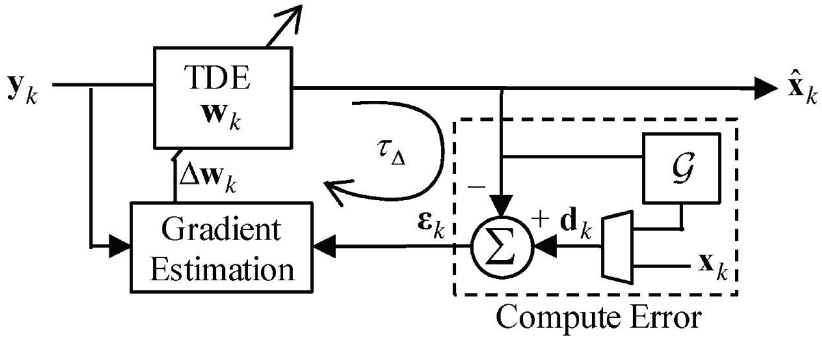


FIGURE 0.7 – Schematic of the adaptive filter [6]

The estimated symbol is given by :

$$\hat{x}_k = \mathbf{w}_{opt}^T \mathbf{y} \tag{7}$$

where $\mathbf{w}_{opt}$ is the optimal filter obtained by the MMSE criterion. It can be shown that this filter is given by [16] :

$$\mathbf{w}_{opt} = \left(E\left[\mathbf{y}^* \mathbf{y}^T \right] \right)^{-1} E\left[x_k \mathbf{y}^* \right] \tag{8}$$

where $E[.]$ represents the expectation value. The filter taps W_{opt} can be derived analytically for this system in terms of $q(t)$ and $h_{ADC}(t)$. When the number of filter taps is large, the filtering can be done more efficiently in the frequency domain using overlap-and-add method [6].

The time-domain CD equalizer presented in this section is used in the analysis of phase noise and CD interaction in chapter 2. In addition, an MMSE filter is used in offline processing of experimental data in chapters 3 and 5. The MMSE in those chapters is employed to reduce the impact of front-end low-pass fiter.

Phase noise estimation methods

One of the most important challenges for optical coherent systems is to estimate and compensate phase noise of the signal. Phase noise mainly originates from the non-ideal transmit and receive lasers having non-zero linewidth. In early optical coherent systems, an optical PLL was used to synchronize the signal phase to that of the LO. However, optical PLLs contain a feedback loop which is sensitive to the loop delay [23]. This delay must be very small for high bit rates and it is difficult to realize ; for instance, at 10 Gb/s this delay should be less than a few tens of nanoseconds [24]. In modern optical coherent systems, the task of the optical PLL is accomplished by the carrier phase estimation algorithms in DSP, enabling the use of coherent detection.

Phase estimation methods are heavily studied in radio frequency communications and considered a classical topic [25]. Although most of these methods can be applied to the optical domain, one should note that the typical linewidth of lasers used in the optical systems is higher than that of radio frequency, so higher levels of phase noise must be handled. Several phase estimation methods have been proposed for optical coherent systems. In this subsection, we study the following methods :

1. Maximum *a posteriori* (MAP) phase estimation [24]
2. Decision-directed (DD) phase estimation [26, 27]
3. Power-law phase estimation [7]

The idea of MAP phase estimation is to find the most probable pair of phase noise and data given the received signal. The MAP estimate can be formulated as one shot, that is, symbol by symbol, or sequential. We examine the one shot approach. Assuming back to back operation (neglecting dispersion and nonlinearity), the received symbol y_k at the k symbol interval can be written as :

$$y_k = x_k e^{j\phi_k} + n_k \tag{9}$$

where n_k and ϕ_k are AWGN and total phase noise (receive and transmit lasers) respectively. The phase noise of both transmit and receive lasers are included in ϕ_k. Given n_k is zero mean, σ_n^2 variance AWGN and ϕ_k is a Wiener process, $\hat{x}_k$ and $\hat{\phi}_k$ are MAP estimates if they maximize the joint pdf of y_k, x_k and ϕ_k. The joint pdf $f(y_k, x_k, \phi_k)$ can be written as :

$$f(y_k, x_k, \phi_k) = f(y_k|x_k, \phi_k) f(x_k) f(\phi_k) \tag{10}$$

where $f(x_k)$ is the probability mass function (pmf) of symbols which is generally assumed to be uniform and $f(\phi_k)$ is the pdf of phase noise for a Wiener process. Because increments of Wiener process are iid Gaussian random variables, we can write :

$$f(\phi_k) = \frac{1}{\sqrt{2\pi\sigma_p^2}} \exp\left(-\frac{(\phi_k - \phi_{k-1})^2}{2\sigma_p^2}\right) f(\phi_{k-1}) \tag{11}$$

where $\sigma_p^2 = 2\pi\Delta f T_s$, in which Δf is the total linewidth (receive and transmit lasers).

As n_k is AWGN, the conditional pdf $f(y_k|x_k, \phi_k)$ is :

$$f(y_k|x_k, \phi_k) = \frac{1}{2\pi\sigma_n^2} \exp\left(-\frac{\left|y_k - x_k e^{j\phi_k}\right|^2}{2\sigma_n^2}\right) \tag{12}$$

We should note that this conditional pdf is a two dimensional pdf because the AWGN is complex noise having two iid Gaussian components with mean zero and variance σ_n^2. The joint pdf is given by :

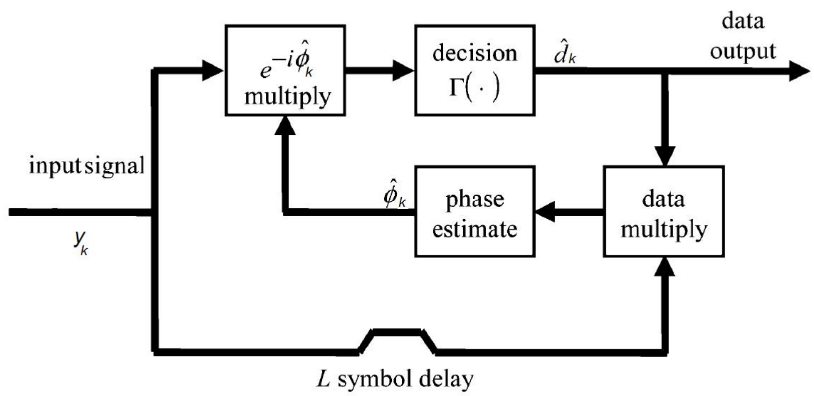


FIGURE 0.8 – Decision-directed phase estimator [7]

$$f(y_k, x_k, \phi_k) = \frac{1}{M}\frac{1}{2\pi\sigma_n^2}\exp\left(-\frac{\left|y_k - x_k e^{j\phi_k}\right|^2}{2\sigma_n^2}\right)\prod_{i=1}^{k}\frac{1}{\sqrt{2\pi\sigma_p^2}}\exp\left(-\frac{(\phi_i - \phi_{i-1})^2}{2\sigma_p^2}\right) \quad (13)$$

where M is the number of constellation points and the symbols are assumed to be equiprobable. Although the MAP phase estimation is optimal, the maximization problem has no analytical solution and numerical solution is computationally expensive for real-time implementations. Clearly, the MAP sequential estimate would be even more complex than this one shot MAP estimate.

In DD phase estimation, symbol detection and phase estimation are not performed jointly. Symbol detection generates $\hat{x}_k$ and this estimate is passed to the phase estimation algorithm. This method is near optimal when the system BER is low, as the symbol values are valid most of the time. A block diagram of this method is shown in Figure 0.8. As shown in this figure, the output of the decision block is used to remove the data modulation by multiplying the conjugate of the decision symbol (shown here as $\hat{d}_k$) by the received signal y_k. It is evident that this method is vulnerable to burst errors because these errors are fed back to the estimator. A training sequence is sequenced to bootstrap operation. Some modifications of this method are proposed in [26]. Decision feedback introduces delay, and is therefore difficult to implement in parallel DSP architectures. Given the high bit rates of interest in coherent optical communication, parallel architectures are essential. For this reason, decision feedback is mostly applied to modulation formats where the symbol amplitude is not constant, e.g., QAM. For constant amplitude modulations such as M-PSK, the power-law estimates are effective and offer easy parallelization.

Power-law phase estimation consists of two stages. In the first stage, data modulation is

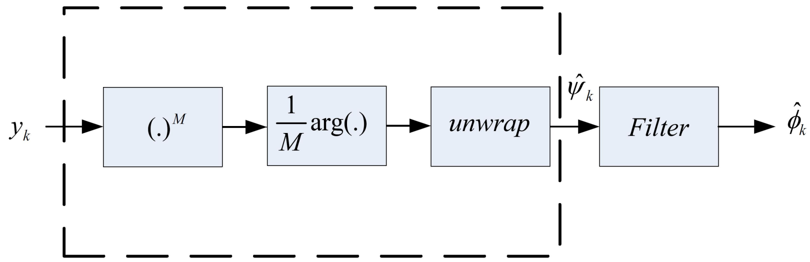


FIGURE 0.9 – A schematic of power-law phase estimator

eliminated to have a soft estimate of the phase noise at each symbol interval. In the second stage, these soft estimates are filtered in order to minimize their noise. The filter to optimally reduce noise in mean square sense is the Wiener filter. Suboptimal filters are also used in practice because of their simplicity. A schematic of this phase estimator is shown in Figure 0.9. Suppose that the modulation format in (9) is M-PSK, the received signal raised to the M^{th} power can be written as

$$y_k^M = (x_k e^{j\phi_k} + n_k)^M = x_k^M e^{jM\phi_k} + M x_k^* e^{j(M-1)\phi_k} n_k + O\left(n_k^2\right) \tag{14}$$

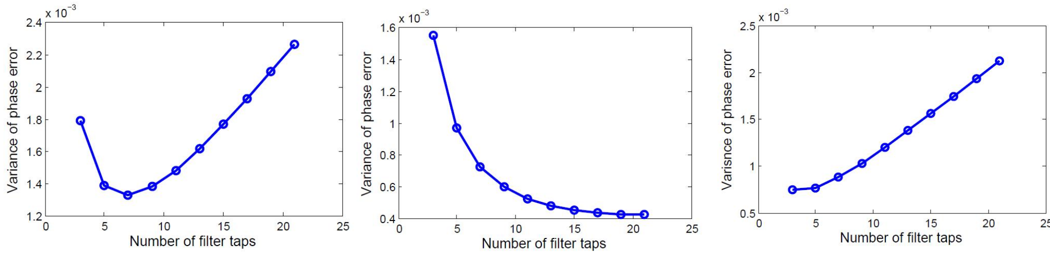
where $O\left(n_k^2\right)$ is high-order noise terms. If the modulation format in (9) is M-PSK, x_k^M equals one (i.e., data modulation is eliminated).

Let ψ_k be the total phase error due to phase noise ϕ_k and ASE n_k, i.e., $\psi_k = \arg(y_k)$ where arg(.) function takes the phase values between $-\pi$ and π. In order to find the soft estimate of the phase error ψ_k, we find the argument of y_k^M and divide it by M

$$\hat{\psi}_k = \frac{1}{M}\arg\left(y_k^M\right) \approx \phi_k + n'_k \tag{15}$$

where n'_k is the residual noise. The phase noise ϕ_k is modelled as a Wiener process, a summation of iid Gaussian random variables whose range is $-\infty$ and ∞. To resolve this mismatch in the ranges, the phase unwrapping must be done on the estimates $\hat{\psi}_k$. In order to smooth variations in the phase estimate over time phase unwrapping is used. Phase unwrapping is often accomplished by comparing the phase estimate $\hat{\psi}_k$ with the previous $\hat{\psi}_{k-1}$ and adding $\pm 2\pi/M$ when the difference $\left|\hat{\psi}_k - \hat{\psi}_{k-1}\right|$ is greater than π/M. This phase unwrapping is a nonlinear function and it can cause cycle slips (discontinuities in multiples of 2π).

The next step is to filter the soft estimates $\hat{\psi}_k$ and smooth n'_k the noise term. It is possible to find a filter that minimizes the mean square error $\epsilon_k = E[|\hat{\phi}_k - \phi_k|^2]$ where $\hat{\phi}_k$ is the output of the filter. The transfer function of the Wiener filter that minimizes ϵ_k is often derived using



(a) (b) (c)

FIGURE 0.10 – Variance of phase error in rad versus number of filter taps N_B (a) Linewidth (LW) = 2 MHz and SNR = 15 dB (b)LW = 200 kHz and SNR = 15 dB (c) LW = 2 MHz and SNR = 20 dB. Modulation format is QPSK with symbol rate 10.7 Gb/s.

z-transform. The impulse response of the filter is given by [24] :

$$w_n = \frac{\alpha r}{1-\alpha^2}\alpha^n u(n) + \frac{\alpha r}{1-\alpha^2}\alpha^{-n} u(-n) \tag{16}$$

$$\alpha = (1 + r/2) - \sqrt{(1 + r/2)^2 - 1}$$

where $u(n)$ is the discrete-time step function and $r = \sigma_p^2/\sigma_{n'}^2$ is the ratio of the phase noise and n'_k variances. The filter is non-causal. Causality can be achieved by buffering and delay in order to access samples after the current sample. As the taps of the Wiener filter decrease exponentially around $n = 0$, the effects in truncating the coefficients and forming an FIR filter with finite number of taps are tolerable. Instead of a truncated Wiener filter, an FIR filter, optimal in MMSE sense, [24] or a filter with uniform taps $w_n = 1/N_B$ (N_B is the number of filter taps) can be used. The latter is sometimes called a moving average filter.

The power-law phase estimator with a moving average filter is commonly used with some minor modifications. This algorithm of phase estimation is called Viterbi-Viterbi algorithm [28]. This approach is shown to be well-suited for practical implementation [29] and is widely employed. In practice, the incoming data is split into a number of blocks to be processed in parallel. The length of each block is equal to the length of the moving average filter. To reduce complexity, phase estimate calculated for the center symbol of each block is used for the whole block. In practice, a modification should be done for the symbols corresponding to the data at the edges of each block as the phase estimate of the center symbol may not be good for these edge symbols [30].

In order to quantify performance of different estimators, a phase error is defined as the difference between the estimated phase and the true phase

$$\Delta\phi_k = \hat{\phi}_k - \phi_k \tag{17}$$

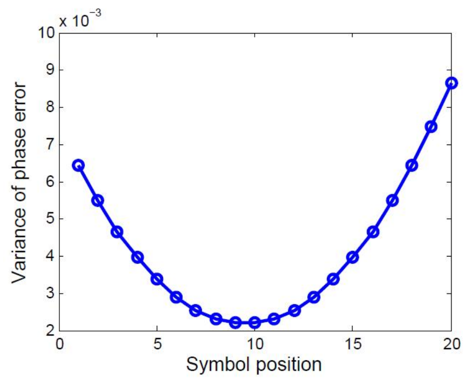


FIGURE 0.11 – Variance of phase error versus symbol position. $N_B = 20$ and other system specifications are the same as part (a) of Figure 0.10.

This phase error is a random variable whose variance quantifies the quality of the estimator. As we might expect, the suboptimum FIR filters with a finite number of taps have larger phase error variance compared with the Wiener filter. Here, we study the variance of the phase error for the moving average filter. The non causal Wiener filter tracks changes in phase noise by exploiting the known dynamics of Brownian motion process. The moving average filter is optimal only for a fixed phase offset in AWGN. When using the moving average filter, larger N_B decreases AWGN by averaging. However within large N_B, the phase has more random variation. Thus variance of phase error changes when the filter length N_B is changed. When the dominant noise is the AWGN (low SNR regime), larger N_B gives smaller phase error variance because larger N_B means a better averaging over Gaussian noise whose mean is zero. In the case where the phase noise is the dominant noise source, smaller N_B gives better performance as variation of phase noise is large from sample to sample and it is better to have less neighbouring samples contribute in the phase noise of a specific symbol. There is an optimum filter length where the phase error variance is minimum. This effect is sometimes called block length effect.

In Fig. 0.10, the variance of $\Delta\phi_k$ is shown versus the filter length for different values of SNR and laser linewidth. An analytical equation in [24] for phase noise variance is used in plotting Fig. 0.10.

Using the moving average filter means that the we average over the soft estimates of a block of N_B symbols to find an estimate of phase noise. Depending on the position of the symbol in the block the variance of $\Delta\phi_k$ is different. Intuitively, having the symbol at the center of the block results in the smallest phase error variance. This effect is shown in Figure 0.11 where the block length (or equivalently the filter length) is 20.

Concepts of phase estimation presented in this section will be used through this thesis. In

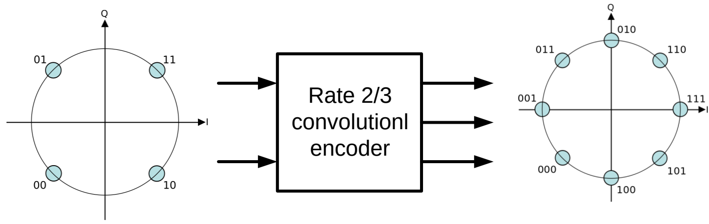


Figure 0.12 – Example of a TCM encoder.

chapter 2, DD phase estimation is assumed for the analysis of phase noise and CD interaction and the lower bound of BER is calculated based on this assumption. In the simulation results, BER performance of DD and power-law phase estimation methods is compared. In chapters 3 and 4, a DD method is employed for carrier phase recovery of 16 quadrature amplitude modulation (16-QAM) in offline processing of experimental data.

0.2.3 An introduction to coded modulation techniques

Coded modulation techniques refer to the general idea of combining modulation and coding. These techniques were first used in band-limited wired communication to get coding gain without bandwidth expansion. In traditional coding techniques, an overhead is added to the information bits and then encoded bits are converted to symbols and transmitted with a certain baud rate, which determine signal bandwidth. For a fixed bit rate, adding overhead leads to higher baud rate or bandwidth expansion. This means that spectral efficiency is reduced. Concept of coded modulation techniques is to accommodate added bits of overhead for a higher modulation format and keep the baud rate fixed. In this way, if the coding gain due to adding overhead surpasses the degradation due to smaller Euclidian distance of higher modulation format, a more spectrally efficient transmission is possible compared to traditional coding techniques where coding and modulation are separate entities. For example in Fig. 0.12, two bits of QPSK modulation are first converted to three bits in a convolutional encoder of rate 2/3 and then an 8-PSK signal is transmitted with the same baud rate. It can be shown that if a Viterbi decoder with enough number of states is used for decoding, a coding gain can be achieved despite reduction in Euclidian distance from QPSK to 8-PSK.

There are many works on satellite and wireless communications regarding combination of coding and modulation. Some pioneering work can be found in [31, 32, 33, 8]. The coded modulation methods can be divided into three general categories : 1) Trellis coded modulation (TCM) 2) Multilevel-coded modulation (MLCM) and 3) Bit-interleaved coded modulation (BICM). Set-partitioning, first introduced by Ungerboeck in [8] for a TCM system, is generally used to assign bits to symbols in the TCM and MLCM. An example of set-partitioning for square 16-QAM is shown in Fig. 0.13. It can be observed that Euclidian distance is increased from top to down layers. The coding strategy is to use strongest coding protection for upper

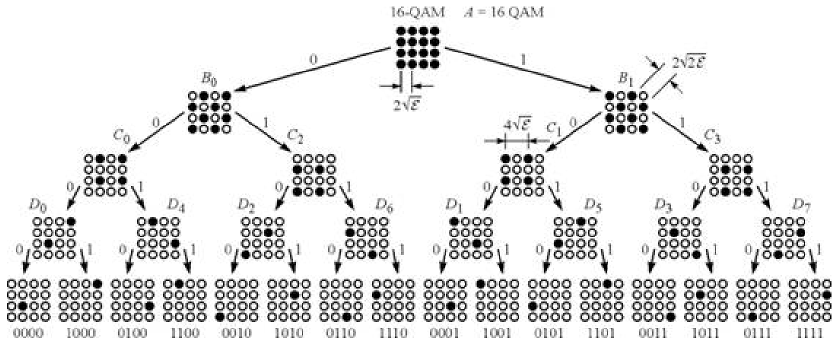


FIGURE 0.13 – Example of a 16-QAM set-partitioning [8].

layers and the weakest (or even no coding) for lower layers. We will discuss more about these concepts in chapter 4. It should be mentioned that set-partitioning is not optimal for iterative demapping and decoding used in BICM.

In most published work on coded modulation in wireless and satellite communication, the channel is assumed to be an AWGN or a fading channel. However, there are a few papers which consider phase noise in the channel [34, 35]. Recently, coded modulation is also studied for use in optical communication. There are a number of works which study coded modulation in this context but assuming an AWGN model for the channel [36, 37, 38, 39]. In two published works phase noise is taken into account ; one is presented for nonlinear phase noise [11] and one for actual laser phase noise in the TCM systems used in optical communication [40]. Recently in [41] nonlinear Schrodinger equation (without assuming phase noise) is used to study coded modulation in optical communication systems. Although TCM is studied well for optical communication, in almost all the channel is assumed to be AWGN [42, 43, 44, 45, 46].

More work is needed to apply current proposals for MLCM coding to optical communications as in most of the works a realistic channel model taking into account phase noise, nonlinearity, etc is not considered. In chapter 4 of this thesis, an MLCM coding is proposed for an optical coherent system impaired by phase noise and AWGN.

0.3 Structure of the thesis

The background presented in previous sections is useful in better understanding of two specific problems which are focus of the research in this thesis. In this section, the structure of the thesis is presented and material of each chapter is briefly described. Contributions in each chapter are outlined to highlight the novelty of the research.

In chapter 2, we study phase noise and CD interaction in optical coherent systems through analysis and simulation. We show that a previously proposed model for this interaction can be improved. We verify validity of the proposed modification by simulation. Our contributions in this chapter are the following :

1. Derivation of a more accurate expression for the estimated phase noise by carrier phase recovery
2. Derivation of a more accurate correlation matrix for the probability density function (PDF) of received samples after carrier phase recovery
3. Demonstration that our BER prediction of QPSK and differential QPSK (DQPSK) using the modified pdf matches the BER of DD carrier phase recovery and is a lower bound of BER from VV carrier phase recovery
4. Study extension of the modified pdf to 16-QAM modulation

In chapter 3, we study MLCM for optical coherent systems impaired by phase noise. We propose a numerical method for finding optimal set-partitioning and code rates. We apply the method to a phase noise optimized and square 16-QAM constellation and explore the performance by simulation. The contributions in this chapter include :

1. Proposing a numerical method to find optimal set-partitioning and code rates for minimum block error rate (BLER) which is applicable to an arbitary constellation and pdf of received samples
2. Applying our MLCM design method to 16-QAM ring constellation in a coherent system impaired by nonlinear phase noise ; showing that our designed MLCM coding outperforms an MLCM coding designed previously in the literature
3. Applying our MLCM design method to phase noise optimized and square 16-QAM constellation in a coherent system impaired by phase noise ; showing by Monte-Carlo (MC) simulations that our MLCM coding performs better than uniform rate coding with the same overhead in terms of post forward error correction (FEC) BER
4. Showing by simulation that minimization of BLER translates into minimization of BER
5. Exploring phase noise regime where post FEC BER of MLCM coding for phase noise optimized constellation performs better than square 16-QAM

In chapter 4, we present an experimental demonstration using MLCM coding, designed according to previous chapter, in an optical coherent system. We explore experimental post FEC BER of MLCM and uniform rate coding using either phase noise optimized or square 16-QAM constellations. Our contributions in this chapter are the following :

1. Experimental demonstration of MLCM coding for 16-ary constellations in an optical coherent system
2. Finding experimentally phase noise regimes where the optimized constellation exceeds square 16-QAM in terms of experimental post FEC BER

3. Showing by experiment that MLCM coding designed with our approach performs better than uniform rate coding in terms of post FEC BER
4. Finding experimentally phase noise regimes where the optimized constellation combined with MLCM coding performs better than uniform rate coding

Chapitre 1

Interaction of LO phase noise and chromatic dispersion

1.1 Introduction

Electronic equalization of fiber CD using DSP is an important advantage [6] of optical coherent systems over IMDD systems among other [4, 47, 48]. The cost and complexity of optical dispersion compensating devices are avoided and, equally important, the system becomes more robust to nonlinearity [49, 50, 51, 52]. In addition, adaptive equalization can be fully exploited [53].

Electronic CD compensation, in contrast to optical dispersion compensation, suffers from equalization enhanced phase noise (EEPN) which stems from the interaction of receive laser (or local oscillator, LO) phase noise with the taps of the dispersion equalizer. The penalty induced by EEPN increases with linewidth, symbol rate and dispersion, and limits the maximum allowable LO linewidth or system reach [54, 55]. The EEPN impact is more serious for high spectral efficiency modulation formats like QAM having a compact constellation [56].

The impact of EEPN on coherent systems has been studied in [55, 9, 54, 57, 58, 56, 59, 60, 61, 62]. Variance of EEPN was studied theoretically in [55] and a power penalty expression was suggested by adding the EEPN variance to the ASE noise variance. It was found in [55] that this penalty increases linearly with linewidth, fiber length and symbol rate. Simulation results reported in [56, 54] suggest that this penalty should instead increase exponentially with fiber length.

A two dimensional PDF of received symbols before decision was derived analytically in [9] and an elliptically shaped PDF was predicted in presence of EEPN. However, the impact of the carrier phase estimator (CPE) was not considered in the analysis. The sum of transmitter and receiver phase noises was used for phase tracking, essentially ignoring EEPN. The impact of EEPN on VV algorithm was also studied briefly by simulation in [9] which indicated that time-variation of VV phase estimate does not follow sum of the true transmitter and receiver phase noises closely.

The analytical expression for the BER floor due to EEPN was found in [55], assigning an effective laser linewidth in the presence of EEPN. A CPE based on one-tap normalized least mean square filter was used to validate the analysis. The impact of EEPN on reduced-guard-interval coherent optical orthogonal frequency-division multiplexing (CO-OFDM) systems was investigated in [60].

As described above, almost all the analytical results concerning EEPN are based on the assumption that the estimated phase by the CPE is approximately the sum of transmitter and receiver phase noises. This is certainly the case for an ideal CPE in optically dispersion-compensated systems. However, this model is no longer valid when the EEPN contribution becomes significant.

An analysis of the EEPN was also presented in [63] with and without a digital coherence enhancement (DCE) technique developed in [63, 64, 65] to reduce EEPN. It was shown that the EEPN contribution can be considered as an additive zero-mean Gaussian noise whose variance can be approximated with the variance derived in [55]. However, the assumption on the CPE is the same as [9].

An experimental verification of EEPN simulation results is reported in [66]. A reduction in the phase error variance of the received symbols in the presence of uncompensated dispersion is reported and it is shown that this reduction is proportional to the fiber dispersion.

In this chapter, we examine estimated phase noise in the presence of EEPN with a more realistic model for the CPE. The CPE in our derivation has perfect removal of data modulation, but retains additive noise that can be reduced by means of a moving average filter. An expression for the estimated phase noise as a function of transmitter and receiver phase noises is provided. The derivation corresponds to the performance of a CPE with ideal DD method. The expression that we provide for CPE phase shows the same behavior reported experimentally in [66] ; the phase variance is reduced by increasing dispersion.

With the expression for the DD CPE phase, we derive the semi-analytical PDF of the decision statistic in this chapter after the CPE. We show that a simple modification of the covariance matrix in the conditional PDF reported in [9] provides a much better approximation of the PDF of a practical CPE such as VV. For example our analysis correctly predicts the circularly shaped PDF following CPE rather than the elliptically shaped PDF predicted in [9]. Using our semi-analytical PDF, we calculate BER for a DQPSK system and compare it with BER from Monte-Carlo (MC) simulation of the system employing either a VV or ideal DD algorithm. We show that our semi-analytical PDF provides an accurate estimate of the system BER

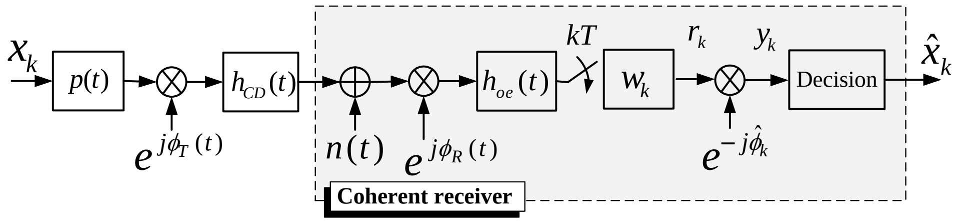


FIGURE 1.1 – Block diagram of a coherent system.

and power penalty for the DD CPE method. In addition, simulation results suggest that this semi-analytical BER can be considered as a lower bound for the BER of VV algorithm.

We also investigate the extension of our results to QAM. Our simulation results show that the same analytical expression for the CPE phase estimate accurately predicts performance for ideal DD CPE. The semi-analytical PDF of the decision statistic is also applicable to QAM, as no assumption on the modulation format is made in deriving this PDF. The accuracy of this BER needs to be investigated further in the case of other CPEs.

Using our analytical PDFs for evaluating BER via numerical integration is faster than MC simulation for BERs by several orders of magnitude, even at low BER, e.g. 10^{-4}. In our semi-analytical method, a number of covariance matrices must be evaluated for a few two dimensional Gaussian PDFs, which can be performed at low complexity. Although we present the analysis for an ideal DD carrier recovery, we can approximate the BER for another carrier recovery if we know the SNR penalty of that specific carrier recovery compared to the ideal DD. In addition to finding BER, having an accurate approximation of the PDFs is useful to implement soft FEC algorithms, as log-likelihood ratios must normally be provided to the decoder.

This chapter is organized as follows. In section 1.2, we consider the system model and notations for the analysis in the following sections. Our first contribution, an expression for the CPE phase estimate and the analytical PDF of the decision statistic, is presented in section 1.3. Our second contribution, modification of the correlation matix in the PDF of decision statistics, is covered in section 1.3.2. Our third contribution, proving by simulation that our BER prediction corresponds to BER from DD carrier phase recovery being a lower bound to VV carrier recovery BER, is presented in section 1.4. Our fourth contribution, extension to QAM modulation is introduced in section 1.5. Finally, we draw the conclusions in section 1.6.

1.2 System model for analysis of EEPN

In this section, we introduce the coherent system model to be used in our analysis. A simplified block diagram of a coherent system is shown in Fig. 1.1 where the input symbols x_k are transmitted using pulse shape $p(t)$. Single polarization transmission is considered here, but it is possible to extend the analysis to the dual-polarization case. The fiber is assumed to show only second order CD, of coefficient β_2 (transfer function $H_{CD}(j\omega)$ in Eq. (2) where $\beta_3 = 0$). In addition to the CD effects introduced by impulse response $h_{CD}(t)$, the transmitted signal is corrupted by noise sources $\phi_T(t)$, $\phi_R(t)$ and $n(t)$, respectively the transmitter and receiver laser phase noises and additive noise due to optical amplifiers. We have not considered the nonlinear effects in our model. Nonlinear effects can be taken into account by modifying the noise sources. For example in [67], it is shown that the phase noise can still be modeled as a Wiener process whose parameters are calculated based on a empirical model.

The received signal is mixed with the local oscillator and photo-detected. Mixing is modeled by multiplication by a phase noise process (*i.e.*, zero frequency offset is assumed), and the photodetection and RF front end are modeled by a single impulse response $h_{oe}(t)$. The signal is sampled at arbitrary rate $1/T$ (typically T is half the symbol duration).

Electronic dispersion equalization is performed using a finite-impulse response (FIR) filter corresponding to the inverse transfer function of the fiber. The N taps of the equalizer w_n are given in Eq. (4) [17].

After CD compensation, carrier phase recovery is performed by derotating the signal using the estimated phase $\hat{\phi}_k$. This estimated phase can be obtained by several different methods, as we discuss in the following sections.

1.3 Analysis in the case of perfect data remodulation

In this section we first provide an expression for the estimated phase $\hat{\phi}_k$ using remodulation with perfect knowledge of the transmitted symbols (*i.e.*, ideal decision feedback). We use this expression to find the PDF of the decision statistic y_k in Fig. 1.1. The details of the analysis can be found in the appendix.

1.3.1 Analytical expression for CPE phase estimate

An essential part of a coherent receiver is the CPE where the phase offset is estimated and the equalized signal is derotated with the estimated phase. Algorithms used in the CPE typically consist of two stages ; in the first stage, data modulation is eliminated and in the second stage the obtained raw phase estimate is filtered to reduce noise [24, 7, 29]. For example, data modulation can be eliminated by using decision feedback or raising to the power of M (for M-PSK signal) in the VV algorithm.

In this chapter, we estimate the best possible performance of the CPE assuming ideal removal of modulation. When modulation is completely removed, the residual signal consists of equalization enhanced phase noise (EEPN) and noise stemming from ASE. To remove the additive noise, a moving average filter is used. While the Wiener filter would give lowest mean square error for the estimate, the suboptimal moving average filter is analytically tractable and gives a reasonable approximation of the performance of the phase estimate with perfect data remodulation.

The number of taps in the CD compensating filter, $2N+1$, can be considerable for long fiber runs. However, the block length $2N_B+1$ of the moving average filter is chosen large enough to average additive noise, yet short enough to assure that the tracked laser phase noise remains relatively constant. Let r_m be the output of the CD compensating filter for the m^{th} symbol (see appendix 1 for details). The phase estimate is

$$\hat{\phi}_k = \text{Arg}\left\{\sum_{m=k-N_B}^{k+N_B} r_m x_m^*\right\} \tag{1.1}$$

We define

$$\Delta_{n,k} = \phi_R(kT - nT) - \phi_R(kT) \tag{1.2}$$

to be the phase increment in the receiver LO phase noise for the k^{th} sample at the n^{th} filter tap in the CD filter. In section 1.7 (Appendix 1), we show that the output of the moving average filter for ideal data remodulation is

$$\hat{\phi}_k \approx \phi_R(kT) + \phi_T(kT) + \sum_n \text{Re}\{w_n p_d(-nT)\}\Delta_{n,k} + n'_k + n''_k \tag{1.3}$$

where Re denotes real part, $p_d(t) = p(t) \otimes h_{CD}(t) \otimes h_{oe}(t)$ ($\otimes$ being the convolution) is the overall system impulse response. n'_k and n''_k are the noise contribution from intersymbol interference (ISI) attributable to EEPN and the residual ASE noise respectively . The former noise source n'_k contains zero-mean independent symbols that add incoherently and is thus averaged out in the moving average filter ; see appendix for more details. This term is an unpredictable part of the EEPN as it contains symbols unknown at the time of applying the CPE, despite ideal decision feedback.

The summation over n in Eq. (1.3) has non-zero mean and can be tracked by the CPE, along with the laser phase drifts $\phi_R(kT)$ and $\phi_T(kT)$. We call this summation the predictable part of the EEPN as it varies slowly as compared with n'_k. Note that when there is no electronic dispersion equalization in the system, the equalizer can be imagined as a one tap filter of $w_0 = 1$. As $\Delta_{0,k} = 0$, the EEPN contribution is zero and the estimated phase reduces to $\hat{\phi}_k = \phi_T + \phi_R$.

1.3.2 Modifying the symbol error PDF

In this section, we find the PDF of the decision statistic $y_k = r_k e^{-j\hat{\phi}_k}$ (shown in Fig. 1.1). We neglect transmitter phase noise, *i.e.*, $\phi_T(t) = 0$, as the contribution of the transmitter phase noise in the presence of EEPN is negligible. In section 1.8 (Appendix 2) we use Eq. (1.3) to show that the decision statistic y_k, in the case of perfect remodulation and a moving average filter to suppress noise, can be written as

$$y_k \approx x_{[k]} + \sum_{n=-N}^{N} s_{n,k}\Delta_{n,k} + \tilde{n}_k \tag{1.4}$$

where $\tilde{n}_k$ with variance $\tilde{\sigma}^2$ is a combination of ASE noise and the noise in phase estimation process and

$$s_{n,k} = j\sum_i x_i w_n p_d\left[(k-n)T - iT_s\right] - jx_{[k]}\mathrm{Re}\{w_n p_d(-nT)\} \tag{1.5}$$

The index of the symbol at the k^{th} sample time is given by $[k]$.

In [55, 9, 63] they assume perfect knowledge of the true phase noise $\phi_R(kT)$ and use this as the phase estimate $\hat{\phi}_k$ although it is highly suboptimal and essentially ignores EEPN. While the expression for y_k in Eq. (1.4) remains unchanged, $s_{n,k}$ is the case of $\hat{\phi}_k = \phi_R(kT)$

$$s_{n,k} = j\sum_i x_i w_n p_d\left[(k-n)T - iT_s\right] \tag{1.6}$$

which matches the results in [9].

Using Eq. (1.4), we can find the PDF of y_k following the same procedure in [9], but with the more realistic and higher performance estimate $\hat{\phi}_k$ in Eq. (1.3). Hence our calculations are based on the analysis leading to Eq. (1.5) that includes the EEPN term depending on the desired bit $x_{[k]}$, and not exclusively the ISI term as in previous analysis. We write $\Delta_{n,k}$ as a sum of independent and identically distributed (iid) Gaussian random variables $\Delta_{n,k} - \Delta_{n\pm1,k}$ each having zero mean and variance $\sigma^2 = 2\pi\Delta\nu_{LO}T$ ($\Delta\nu_{LO}$ being the LO linewidth)

$$\begin{aligned} y_k = x_{[k]} + \sum_{n=-N}^{-1}\left[(\Delta_{n,k} - \Delta_{n+1,k})\sum_{m=-N}^{n} s_{m,k}\right] + \\ \sum_{n=1}^{N}\left[(\Delta_{n,k} - \Delta_{n-1,k})\sum_{m=n}^{N} s_{m,k}\right] + n_k \end{aligned} \tag{1.7}$$

Suppose $2M$ is the system memory multiplied by the oversampling rate (2 in this chapter). Given symbol pattern $\mathbf{X} = [x_{[M-k]}, ..., x_{[k-1]}, x_{[k+1]}, ..., x_{[M+k]}]$ and desired symbol $x_{[k]}$, we can calculate $s_{n,k}$ in Eq. (1.5). The conditional PDF $f(y_k|\mathbf{X})$ is a complex Gaussian PDF with mean vector $[\mathrm{Re}\{x_{[k]}\}, \mathrm{Im}\{x_{[k]}\}]^{tr}$ and covariance matrix C which is given in section 1.8 (Appendix 2). This covariance matrix is a function of $s_{n,k}$, $\tilde{\sigma}^2$ and σ^2. The conditional pdf $f(y_k|x_{[k]})$ is a Gaussian mixture, that is, the normalized sum over PDFs $f(y_k|\mathbf{X})$ for different symbol patterns $\mathbf{X}$

$$f(y_k|x_{[k]}) = \frac{1}{|\mathrm{P}|}\sum_{\mathbf{X}\in\mathrm{P}} f(y_k|\mathbf{X}) \tag{1.8}$$

where P is the set of all possible symbol patterns $\mathbf{X}$ and $|\mathrm{P}|$ is its cardinality. The total PDF $f(y_k)$ is an average of PDFs $f(y_k|x_{[k]})$ over all possible symbols $x_{[k]}$ (or constellation points).

Although the number of symbol patterns could be very large, we show by simulation that a summation over a limited number of patterns chosen randomly is sufficient to find a relatively accurate estimate of the PDF. For calculating $s_{n,k}$, we need to calculate an infinite sum in Eq. (1.5) which can be truncated in practice to the finite system memory $2M$. It can be shown that $M \approx N$. Thus, for $L = 3000$ km, $R_s = 28$ Gbaud and two samples per symbol, M and N are around 640.

1.4 Simulation results

We consider the performance of three phase estimation methods using both MC and semi-analytical techniques. The methods are

- ideal decision feedback for removal of modulation and a moving average filter to remove noise (ideal DD)
- Viterbi-Viterbi for removal of modulation and a moving average filter to remove noise (VV)
- Using an ideal pilot tone (PT) to extract perfect knowledge of $\phi_R(t) + \phi_T(t)$

In the case of ideal DD we use our previous analysis to generate a semi-analytical prediction of the BER based on Eq. (1.8) using Eq. (1.5), as well as an MC simulation where ideal removal of modulation is followed by a moving average filter. In the case of VV, we only generate MC BER curves using VV algorithm. In the PT case, a continuous wave PT provides a separable signal at the receiver input whose phase is the sum of transmitter and receiver phase noise, $\phi_R(t) + \phi_T(t)$. We compare BER from MC for this PT method with an analytical formula where $\phi_R(t) + \phi_T(t)$ is assumed for phase noise cancelation of received symbols [55, 63]. Note that the PT is clearly suboptimal, but was examined in [55, 9] due to tractability of the analysis. Unfortunately the known total phase noise at the input provided by PT is **before** CD compensation, while the phase noise derotation is **after** CD compensation when the filtered phase noise no longer resembles the input phase noise.

1.4.1 System parameters

We perform numerical simulations of a coherent system and compare the analytical prediction of BER with the MC simulation results. We consider simulation of QPSK in this section and study the QAM modulation in the next section. In the case of QPSK transmitter differential and Gray coding are used. A root-raised cosine pulse having roll-off factor 1 is used as the transmitted pulse $p(t)$.

The fiber span of length L is assumed to have only second-order CD with dispersion coefficient $\beta_2 = 21.6$ ps^2/km. It is assumed that fiber loss is perfectly compensated with amplifier gains and the additive Gaussian noise due to amplifiers is added at the end of the fiber span. The transmitter and receiver laser phase noises (each having linewidth $\Delta\nu/2$) are modeled as a Wiener process.

Temporal resolution of the simulation is $T' = T_s/8$ which is also used in generating the Wiener process. In the coherent receiver, the received signal is first filtered by a filter matched to the transmitted pulse shape and then sampled at 2 ($T = T_s/2$). This is followed by the CD compensation which is performed using the filter in Eq. (4) ; filter length varies as a function of fiber length examined.

After CD compensation, carrier phase recovery is performed per one of the three methods discussed previously. In the DD case, data modulation is first removed using perfect knowledge of the transmitted symbols. Next the signal is filtered to remove noise using an 11-tap moving average filter. The ideal DD provides an upper-bound of performance that can be achieved by any CPE.

The VV case has data modulation removed by raising to the appropriate power (2 for BPSK, 4 for QPSK, etc.). The phase is then unwrapped (exploiting the differential encoding) and filtered for noise using the same 11-tap moving average filter for the DD case. Finally, in the PT case no CPE is used, the phase estimate is simply $\phi_R(t) + \phi_T(t)$ using perfect knowledge of the laser phase noise at transmitter and receiver.

1.4.2 Accuracy of semi-analytical PDFs

Two dimensional PDFs of the decision statistic $f(y_k)$ are estimated using MC simulation (assuming ideal DD as the CPE) and calculated semi-analytically for the DD and PT cases. MC estimates of the PDF are obtained by transmitting 5×10^6 symbols. In order to calculate the semi-analytical PDFs, 50 symbol patterns $\mathbf{X}$ chosen randomly are used in Eq. (1.8) for each of the four possible symbols x_k to calculate $f(y_k|x_k)$.

The estimated PDF and the semi-analytical PDF using DD are both circularly-shaped while the PDF in [9] is ellipticity-shaped ; the tails of each elliptic constellation points extends toward the decision boundary which explains why we will observe worse BER performance for the PT case. The one-dimensional cross section of these PDFs at $\mathrm{I}m\{\hat{y}_k\} = 0$ is shown in Fig. 1.2(a).

The semi-analytical PDF in Eq. (1.8) can be calculated very fast as the number of patterns required to get an accurate approximate for the true PDF is small. This is investigated in Fig. 1.2 where the error in PDF is defined as

$$\varepsilon = \int\limits_{y_k \in S} |f_{MC}(y_k) - f(y_k)| dy_k \tag{1.9}$$

where $f_{MC}(y_k)$ stands for the PDFs obtained by MC. The integration is over entire complex plane S. It can be observed from Fig. 1.2(b) that the error ε converges to an error floor very fast by increasing the number of patterns.

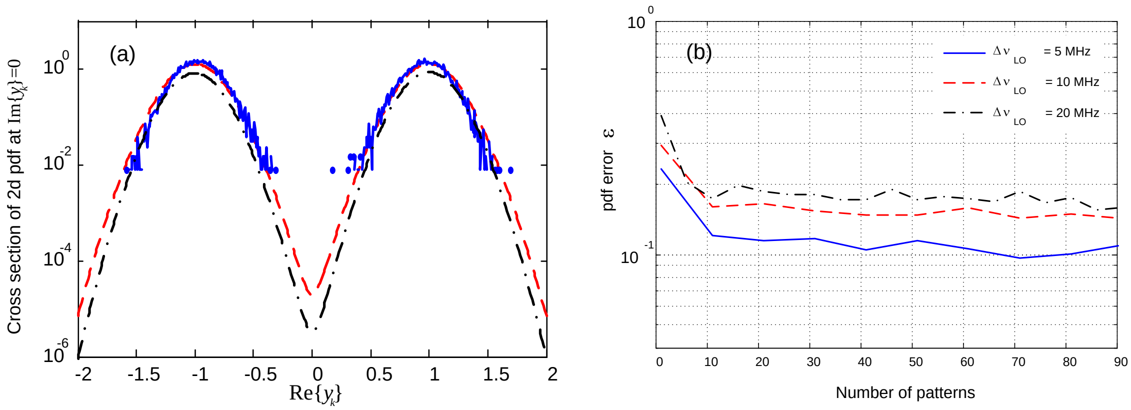


FIGURE 1.2 – (a) Cross section at $Im\{\hat{y}_k\} = 0$ of two-dimensional PDFs i) obtained by MC simulation (solid), ii) as developed in Eq. (1.8) (dashed) and iii) as reported in [9] (dot-dashed) ;$L = 3000$ km and $\Delta\nu = 10$ MHz, (b) error in PDF as a function of number of patterns for averaging in Eq. (1.8) ; PDF plotted in (a) uses 50 patterns.

1.4.3 BER for three phase estimates

In this section, we compare BER estimates found with MC simulations and analytical expressions. Semi-analytical BER is calculated for DD using the conditional PDFs $f(y_k|x_k)$ by Eq. (1.8). Each of the four conditional PDFs (two dimensional Gaussian PDFs) is calculated by averaging over 50 random pattern vectors $\mathbf{X}$, where the length of the vector corresponds to system memory $2M$. We use 50 patterns to find the sum in Eq. (1.8) as the total number of patterns is excessively large - 4^{2M-1} for the case of QPSK. We have confirmed via simulation that this number is sufficient to obtain accurate results. The conditional PDFs are integrated over QPSK decision regions (quarters of the complex plane) to find the symbol error rate (SER). As the SER of a Gray-coded QPSK system is twice its BER, and this BER is half of the BER of differential QPSK system, the calculated SER is equal to the BER of differential QPSK. We have generated MC estimates of the BER of the differential QPSK system with ideal DD.

We assume symbol rate of 28 Gbaud, $L = 3000$ km and total linewidth of $\Delta\nu = 10$ MHz (5 MHz for LO and transmitter laser). In Fig. 1.3(a) we present BER versus signal-to-noise ratio (SNR), including the AWGN case (an optically compensated with no EEPN) given by the solid line. For the assumed parameters, the main source of SNR penalty (as compared with AWGN case) is EEPN. The BER obtained from the semi-analytical PDF Eq. (1.8), where Eq. (1.5) is used in the covariance matrix, is given in the dashed curve in red. The accuracy of the semi-analytic result is confirmed by the MC BER given by the blue curve with square markers. Having confirmed our semi-analytic BER result for ideal DD, we next examine the utility of this result as a lower bound for the VV carrier recovery. An MC simulation of BER for VV carrier recovery is shown in Fig. 1.3(a) with diamonds markers in green. The

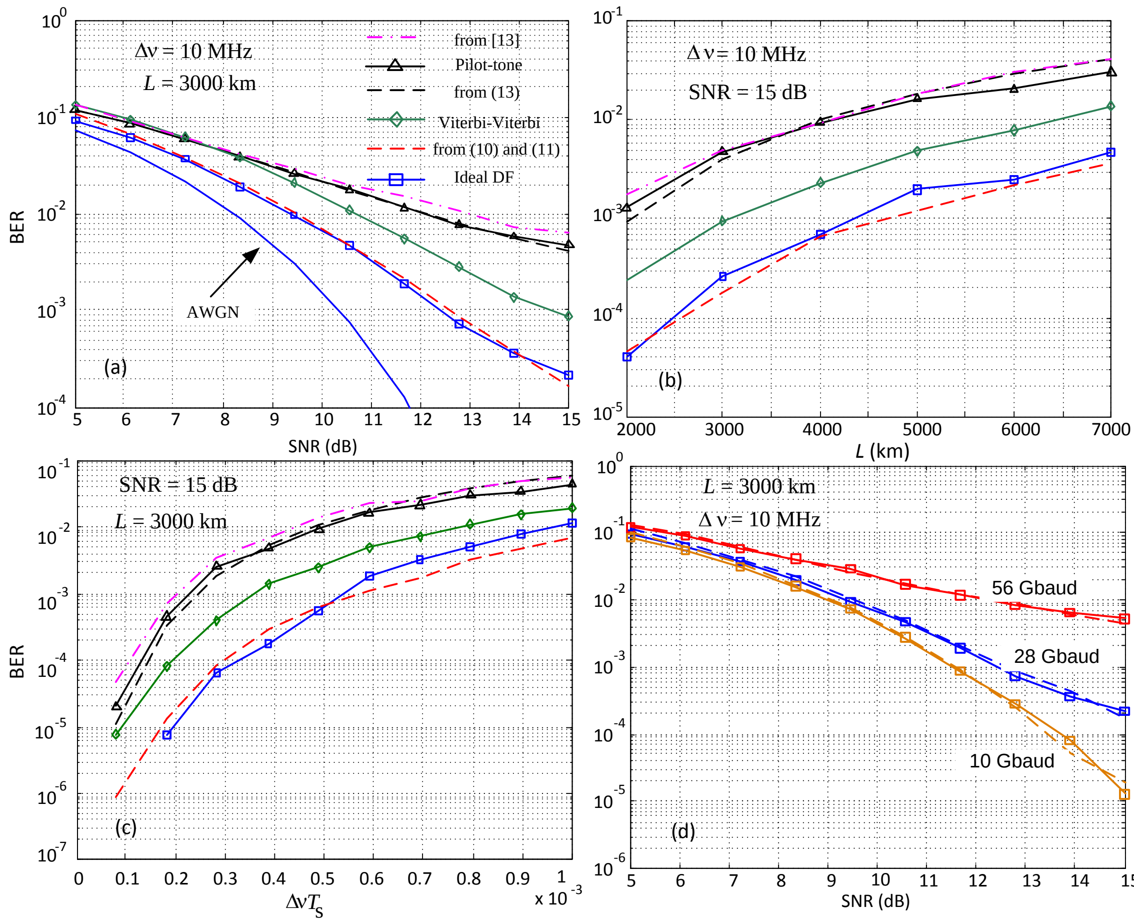


FIGURE 1.3 – BER versus (a) SNR (b) fiber length L (c) linewidth-symbol time product $\Delta\nu T_s$ (d) SNR for different baud rates from MC simulation (markers) and semi-analytical PDFs (dashed). Baud rate in (a)-(c) is 28 Gbaud.

non-ideal modulation removal of the VV technique leads to a clearly visible penalty. Finally, we reproduce the BER found in [63] by integrating the two dimensional PDF in [9] over the decision boundaries. This semi-analytical BER, shown in black triangles, assumed a pilot tone (PT) provided the true total phase noise at the receiver input that was then used for derotating the symbol after CD compensation. Note that the PT results allowed an analytical attack leading to the following simple expression (plotted in the black dashed curve) that closely matched the semi-analytical PT BER results

$$\mathrm{BER} = 2Q\left(1/\sqrt{\sigma_n^2 + \sigma_{eepn}^2}\right) \tag{1.10}$$

where $\sigma_{eepn}^2 = \pi^2 L \Delta\nu_{LO} R_s$, $Q(\cdot)$ is the Q-function and signal power is normalized to one. The factor of two accounts for differential detection. Despite the perfect side information, the BER found (either analytical or semi-analytical) is not a bound, and is in fact a pessimistic predictor of VV performance. This is due to the highly non-optimal strategy of derotating the received symbols with this side information.

We next fix the SNR to 15 dB and examine the impact of other system parameters on BER : BER versus fiber length, linewidth and baud rate are shown in Figs. 1.3(b)-(d), respectively. In Figs. 1.3(b)-(c), where a fixed baud rate of 28 Gbaud is assumed, a good match between our semi-analytical predication and MC validation is evident. In addition, changing the baud rate in Fig. 1.3(d) does not affect accuracy of the results. The prediction provides a lower bound for VV performance over the range of values considered.

We use the semi-analytical BER to calculate the penalty in SNR at BER of 3.8×10^{-3} and compare it with the penalty obtained by MC simulation. The penalty is measured with respect to the system having just additive Gaussian noise whose theoretical BER is known. The result is shown in Fig. 1.4 for different fiber lengths and laser linewidths.

It can be observed from Fig. 1.3 that our semi-analytical BER is a lower bound for the BER of the VV carrier recovery so it can be useful in practice to estimate a lower bound for the required SNR, minimum LO laser linewidth and system reach in presence of EEPN. In addition, having a closed-form expression for conditional PDFs, gives more insight into the stochastic properties of the decision statistic that can be exploited in taking measures to reduce EEPN impact, for example, via employing soft decision FEC.

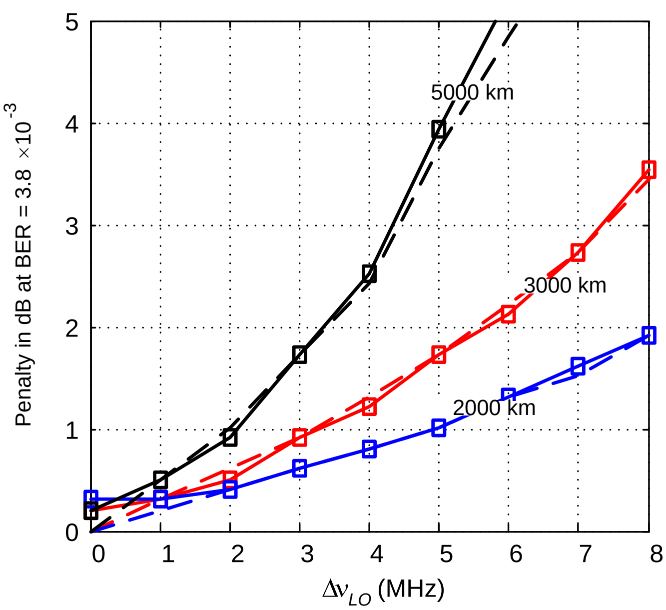


FIGURE 1.4 – Penalty in dB at the BER of 3.8×10^{-3} from MC simulation (markers) and analytical PDF (dashed). DD is assumed for carrier recovery and baud rate is 28 Gbaud.

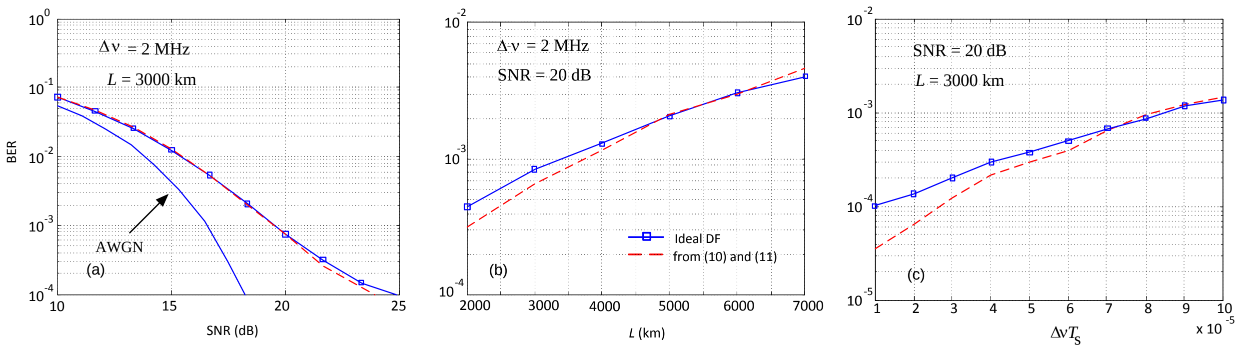


FIGURE 1.5 – BER for 16QAM at 28 Gbaud for ideal DD and for our prediction of ideal DD performance from (10) and (11). BER versus (a) SNR (b) fiber length L (c) linewidth-symbol time product $\Delta\nu T_s$.

1.5 Extension to QAM

The derivation of the analytical expression for $\hat{\phi}_k$ in Eq. (1.3) is presented in section 1.7 (Appendix 1) for the case of M-PSK signaling. The constant amplitude of M-PSK signaling was exploited in the derivation. While 16-QAM and 64-QAM are not constant amplitude, we nonetheless examined via simulation the accuracy of Eq. (1.3) for these cases over the parameter ranges on interest. We ran two separate MC simulations. In one case the CPE estimate $\hat{\phi}_k$ was generated with ideal DD remodulation followed by a moving average filter of the unwrapped phase. In the second case the MC values for phase noise and AWGN were used in Eq. (1.3) to produce the CPE estimate $\hat{\phi}_k$. While not presented here, the BER predictions of these two simulations were very close for a wide range of system parameters in the case of both 16 and 64-QAM.

Having verified the accuracy of Eq. (1.3) by simulation for QAM modulation, we adopt this expression and follow the same derivation for the conditional PDFs. Note that the constant amplitude of QPSK was not needed in deriving Eq. (1.7). The mean vectors $[\mathrm{R}e\{x_{[k]}\}, \mathrm{I}m\{x_{[k]}\}]^{tr}$ and the covariance matrix depending on Eq. (1.5) are calculated for 16-QAM symbols to produce the Gaussian PDFs $f(y_k|\mathbf{X})$. The analytical PDFs are integrated to find the BER.

We again assume symbol rate of 28 Gbaud, $L = 3000$ km and total linewidth of total linewidth of 2 MHz (1 MHz for LO or transmitter laser). In Fig. 1.5(a) we present BER versus SNR, including the AWGN case. The dashed curve is the BER from semi-analytical PDFs, which shows a good prediction of the BER by MC simulation (square markers). Again the SNR penalty compared with the AWGN curve is mainly due to EEPN here. Accuracy of the semi-analytical BER is verified versus fiber length and linewidth in Figs. 1.5(b)-(c). An SNR of 20 dB is used for the 16-QAM system. A good match between theory and simulation can be observed at large distances or linewidth, however, there is deviation from theory in small distances or narrow linewidth at this SNR. Future investigations should examine the applicability of these equations for a lower bound for realistic carrier recovery techniques for QAM.

1.6 Conclusion

Interaction of dispersion and phase noise in coherent systems is studied analytically in presence of CPE. Based on this analysis, a more realistic model for the CPE in terms of transmitter and receiver phase noise is proposed. Based on this CPE model, an analytical PDF of decision statistic previously studied in the literature is improved. It is shown than the BER prediction of previously presented PDF is pessimistic. The analysis is validated by MC simulations for a QPSK system using ideal DD CPE and it is shown that our model predicts well the system BER employing this ideal CPE. It is shown by simulation that the proposed semi-analytical BER can be served as a lower bound for the BER by using VV carrier recovery. It is also shown that the proposed model can be extended to the QAM and it provides fairly accurate prediction of the BER in the case of 16-QAM employing ideal DD. However, accuracy of our semi-analytical BER to estimate BER for other carrier recovery techniques should be studied further.

Using the CPE model, the link between the LO phase noise and the estimated phase by the CPE can be mathematically described. We are currently working to exploit this link to devise algorithms that compensate EEPN in the system. On the other hand, the analytical PDF of decision statistic is useful in practice when maximum-likelihood sequence estimation (MLSE) algorithms or soft decision FEC are employed to reduce EEPN impact. In addition, prediction of the system penalty without resorting to MC simulations could be of interest in the system design where connection between this penalty and dispersion, linewidth and electrical filter bandwidth needs to be determined.

1.7 Appendix 1

We consider an M-PSK signal for which $|x_i| = 1$ and assume that the transmitter phase noise is zero. The output of the CD filter is given by

$$r_k = \sum_{n=-N}^{N} w_n \sum_{i=-\infty}^{\infty} x_i p_d \left[(k-n)T - iT_s\right] e^{j\phi_R[(k-n)T]} + n_k \tag{1.11}$$

where T_s and T are symbol and sampling time respectively and $p_d(t) = p(t) \otimes h_{CD}(t) \otimes h_{oe}(t)$ ($\otimes$ being convolution). We assume the CD compensating filter $\{w_n\}$ perfectly compensates the chromatic dispersion, *i.e.*,

$$\sum_{n=-N}^{N} w_n \sum_{i=-\infty}^{\infty} x_i p_d \left[(k-n)T - iT_s\right] = x_{[k]} \tag{1.12}$$

where $[k]$ gives the index of the desired symbol at the k^{th} sample. Under this assumption, the output of the CD compensating filter is

$$r_k = x_{[k]} \left(1 + q_k\right) e^{j\phi_R(kT)} + n_k \tag{1.13}$$

where

$$q_k = \sum_{n=-N}^{N} w_n \left(e^{j\Delta_{n,k}} - 1\right) \sum_{i=-\infty}^{\infty} \frac{x_i}{x_{[k]}} p_d \left[(k-n)T - iT_s\right] \tag{1.14}$$

and where we define

$$\Delta_{n,k} = \phi_R(kT - nT) - \phi_R(kT) \tag{1.15}$$

to be the phase increment in the receiver LO phase noise for the k^{th} sample at the n^{th} filter tap in the CD compensating filter. Consider the product $w_n \left(e^{j\Delta_{n,k}} - 1\right)$. For small n the phase increment is small and we use the small angle approximation $w_n \left(e^{j\Delta_{n,k}} - 1\right) \approx j\Delta_{n,k} w_n$; for larger n the coefficient w_n is decaying exponentially and the product will be close to zero. Hence

$$q_k \approx \sum_{n=-N}^{N} j\Delta_{n,k} w_n \sum_{i=-\infty}^{\infty} \frac{x_i}{x_{[k]}} p_d \left[(k-n)T - iT_s\right] \tag{1.16}$$

We assume the filter output r_k is ideally remodulated, removing $x_{[k]}$. We further assume that the phase noise $\phi_R(kT)$ is nearly constant within the block length $2N_B + 1$ of the moving average filter, to obtain

$$\begin{aligned} \tfrac{1}{2N_B+1} \sum_{m=k-N_B}^{k+N_B} r_m x^*_{[m]} \approx \tfrac{1}{2N_B+1} e^{j\phi_R(kT)} \sum_m q_m \\ + e^{j\phi_R(kT)} + \tilde{n}''_k \end{aligned} \tag{1.17}$$

where $\tilde{n}''_k$ is proportional to n''_k in Eq. (1.3) due to ASE noise. The sum over q_m can be written as

$$\begin{aligned} \sum_m q_m = j \sum_m \sum_n \Delta_{n,k} w_n p_d \left(-nT\right) + \\ \sum_{m,n,x_i \neq x_{[m]}} \tfrac{x_i}{x_{[m]}} j\Delta_{n,k} w_n p_d \left[(m-n)T - iT_s\right] \end{aligned} \tag{1.18}$$

The first term over m becomes multiplication by the number of taps of the moving average filter $2N_B + 1$. Drawing on our assumption that the block length is chosen such that phase noise $\phi_R(kT)$ is nearly constant over the moving average, we approximate the phase increments (relative to the n^{th} CD filter tap timing) to also remain unchanged, $\Delta_{n,m} \approx \Delta_{n,k}$, hence we use $\Delta_{n,k}$ in the summations.

The second term in Eq. (1.18) isolates the ISI contribution. We call this term $\tilde{n}'_k$ as it is proportional to the noise n'_k in Eq. (1.3). This summation is characterized by terms with random phase changes due to ISI that add incoherently. This zero mean term adds to the noise, which will be approximated as Gaussian and lumped with other noise at the output of the moving average filter. We now have

$$\sum_m q_m \approx \left(2N_B + 1\right) \sum_n j\Delta_{n,k} w_n p_d \left(-nT\right) \tag{1.19}$$

and the output of the moving average filter is

$$e^{j\phi_R(kT)}\left\{1+1+\sum_n j\Delta_{n,k}w_n p_d(-nT)\right\}+\tilde{n}'_k+\tilde{n}''_k \tag{1.20}$$

Taking the argument of the moving average filter output to get our phase estimate $\hat{\phi}_k$, we have

$$\hat{\phi}_k \approx \phi_R(kT)+\text{Arg}\left\{1+\sum_n j\Delta_{n,k}w_n p_d(-nT)\right\}+n'_k+n''_k \tag{1.21}$$

Consider the summation. The phase increment $\Delta_{n,k}$ is small for small index n since the phase holds constant over $2N_B$ samples. As the tap index n grows, so does the phase increment, however, the filter coefficients $\{w_n\}$ decay exponentially. Therefore the sum is small compared to one and we can use the approximation Arg $\{1+h\} \approx \text{I}m\{h\}$.

$$\hat{\phi}_k \approx \phi_R(kT)+\sum_n \Delta_{n,k}\text{Re}\left\{w_n p_d(-nT)\right\}+n'_k+n''_k \tag{1.22}$$

$$\theta_k = \sum_{m,n,i\neq[m]} \Delta_{n,k}\text{Re}\frac{x_i}{x_{[m]}}w_n p_d\left[(m-n)T-iT_s\right] \tag{1.23}$$

using arguments similar to those just presented.

Finally, this analysis assumed zero transmitter phase to simplify our derivation. As we have a linear system, and we have assumed the CD compensating filter exactly compensated the chromatic dispersion, we can use the superposition principle to include the transmitter phase noise and the ISI phase contribution to conclude

$$\hat{\phi}_k \approx \phi_R(kT)+\phi_T(kT)+\theta_k+\sum_n \Delta_{n,k}\text{Re}\left\{w_n p_d(-nT)\right\}+n'_k+n''_k \tag{1.24}$$

1.8 Appendix 2

The decision statistic y_k is the result of derotating the output of the CD compensating filter by $\hat{\phi}_k$. From Eq. (1.11), $y_k = r_k e^{-j\hat{\phi}_k}$ is

$$y_k = \sum_n w_n \sum_i x_i p_d\left[(k-n)T-iT_s\right]e^{j\phi_R((k-n)T)}e^{-j\hat{\phi}_k}+n_k \tag{1.25}$$

Using Eq. (1.24) with zero transmitter phase noise, $\phi_T=0$,

$$\begin{aligned} y_k \quad = \quad & \sum_n w_n \sum_i x_i p_d\left[(k-n)T-iT_s\right] \\ & e^{j\Delta_{n,k}}\cdot e^{-j\text{Re}\sum_l w_l p_d(-lT)\Delta_{l,k}}+\tilde{n}_k \end{aligned} \tag{1.26}$$

where we lump noise from $\hat{\phi}_k$ into a new additive noise term. We can write the exponentials in Eq. (1.26) as

$$\begin{aligned} & e^{j\Delta_{n,k}}\prod_l e^{-j\Delta_{l,k}\text{Re}\{w_l p_d(-lT)\}} \\ & \approx (1+j\Delta_{n,k})\prod_l\left(1-j\Delta_{l,k}\text{Re}\left\{w_l p_d(-lT)\right\}\right) \\ & \approx 1+j\Delta_{n,k}-\sum_l j\Delta_{l,k}\text{Re}\left\{w_l p_d(-lT)\right\} \end{aligned} \tag{1.27}$$

where we used the small angle approximation and only kept terms of first order in the product. Thus we have

$$\begin{aligned} y_k &= \left(1 - \sum_l j\Delta_{l,k}\mathrm{Re}\left\{w_l p_d(-lT)\right\}\right) \\ &\cdot \sum_n w_n \sum_i x_i p_d\left[(k-n)T - iT_s\right] \\ &+ \sum_n j\Delta_{n,k} w_n \sum_i x_i p_d\left[(k-n)T - iT_s\right] + \tilde{n}_k \\ &\approx x_{[k]} - x_{[k]}\sum_l j\Delta_{l,k}\mathrm{Re}\left\{w_l p_d(-lT)\right\} \\ &+ \sum_n j\Delta_{n,k} w_n \sum_i x_i p_d\left[(k-n)T - iT_s\right] + \tilde{n}_k \end{aligned} \tag{1.28}$$

using Eq. (1.12). Defining

$$s_{n,k} = j\sum_i x_i w_n p_d\left[(k-n)T - iT_s\right] - jx_{[k]}\mathrm{Re}\left\{w_n p_d(-nT)\right\} \tag{1.29}$$

we can write

$$y_k = x_{[k]} + \sum_n \Delta_{n,k} s_{n,k} + \tilde{n}_k \tag{1.30}$$

where $\tilde{n}_k$ is the cumulative noise in the estimation process with variance $\tilde{\sigma}^2$.

The conditional PDF $f(y_k|\mathbf{X})$ is a complex Gaussian PDF with mean vector $[\mathrm{Re}\{x_{[k]}\}, \mathrm{I}m\{x_{[k]}\}]^{tr}$ and a 2×2 covariance matrix C depending on $s_{n,k}$ and $\tilde{\sigma}^2$ and $\sigma^2 = 2\pi\Delta\nu_{LO}$. To find elements of this matrix, we consider real and imaginary parts of the right hand side of Eq. (1.7). The diagonal elements of C correspond to the variance of real and imaginary parts (each being a sum of iid Gaussian random variables, the variance is the sum of variances). The off-diagonal elements of C correspond to the covariances of the real and imaginary parts, which can be easily simplified using the fact that many terms in the covariance are products of independent random variables with zero mean. This matrix is the same as that in [9], however using Eq. (1.5) for $s_{n,k}$. We repeat the equations for the covariance matrix here for completeness.

$$C = \sigma^2 \begin{bmatrix} c_{11} + \tilde{\sigma}^2/\sigma^2 & c_{12} \\ c_{21} & c_{22} + \tilde{\sigma}^2/\sigma^2 \end{bmatrix} \tag{1.31}$$

where

$$c_{11} = \sum_{i=-N}^{-1}\left(\sum_{n=-N}^{i}\mathrm{Re}\{s_{n,k}\}\right)^2 + \sum_{i=1}^{N}\left(\sum_{n=i}^{N}\mathrm{Re}\{s_{n,k}\}\right)^2, \tag{1.32}$$

$$c_{22} = \sum_{i=-N}^{-1}\left(\sum_{n=-N}^{i}\mathrm{I}m\{s_{n,k}\}\right)^2 + \sum_{i=1}^{N}\left(\sum_{n=i}^{N}\mathrm{I}m\{s_{n,k}\}\right)^2, \tag{1.33}$$

$$\begin{aligned} c_{12} = c_{21} = &\sum_{i=-N}^{-1}\left(\sum_{n=-N}^{i}\mathrm{Re}\{s_{n,k}\}\sum_{n=-N}^{i}\mathrm{I}m\{s_{n,k}\}\right) + \\ &\sum_{i=1}^{N}\left(\sum_{n=i}^{N}\mathrm{Re}\{s_{n,k}\}\sum_{n=i}^{N}\mathrm{I}m\{s_{n,k}\}\right) \end{aligned} \tag{1.34}$$

Chapitre 2

Multi-level coded modulation in presence of phase noise

2.1 Introduction

While capacity can be pushed by leveraging constellation size and coding complexity independently, in practice there is always a compromise between the two. Coded modulation (CM) techniques promise more flexibility in efficient use of constellation size and coding strategy. A plethora of CM techniques exist in the literature [32, 8, 68, 33], however, application to the fiber-optic channel requires modifications as most existing techniques are developed for an AWGN or fading channel (suitable for wireline, wireless and satellite communications). Fiber-optic systems include chromatic dispersion, nonlinearity and relatively high levels of phase noise, often violating the AWGN assumption. In addition, due to the high speed of optical coherent systems today, CM complexity should be moderate, and implementation practical.

TCM is a classic example of CM studied for fiber-optic systems [40, 44, 42, 46, 43, 36, 38]. The channel is typically assumed to be AWGN, which may not be valid in practice. Even in short reach systems with negligible chromatic dispersion and nonlinearity, phase noise can substantially affect large constellations, like 16-QAM. The impact of phase noise on TCM for coherent polarization shift keying (POLSK) system was studied in [40]. However, the impact of phase noise on coded modulation systems for coherent systems is not studied thoroughly.

Block coded and BICM are other CM techniques studied well for optical communications [39, 69, 70, 71, 72]. Specifically, the use of low density parity check codes (LDPC) codes along with soft iterative decoding is studied extensively in [69, 70, 73]. Although BICM has a simple encoder, the decoding complexity can be quite high and may not be practical for very high baud rate. BICM performance strongly depends on the mapping of bits to symbols. The optimal mapping varies with system parameters like SNR, desired BER and number of iterations in the decoder [39], and may be difficult to find for constellations without strong symmetry.

Another CM technique known as MLCM has received less attention in optical communications. MLCM is a special case of generalized-LDPC (GLDPC) codes [74]. GLDPC codes are studied in [75] for optical coherent detection. However, no systematic mapping and rate optimization is used. Application of MLCM to a coherent system with nonlinear phase noise is studied in [11]. A heuristic set-partitioning is introduced in [11] and the code rates of RS component codes are optimized to find minimum BLER. While latency and error propagation can cause a problem in multi-stage decoders (MSDs) for MLCM, in [41] staircase component codes and independent decoding of each code can eliminate these concerns. For achieving high coding gain in [41], it is essential to use a Gray mapping of bits to symbols, which may not be feasible for an arbitrary constellation lacking strong symmetry, such as the phase noise optimized constellations in [10].

MLCM using RS component codes and hard decision MSD is studied in this chapter as a coding strategy for optical coherent systems. The choice of RS encoders and hard decision MSD enables a low complexity coding scheme. To this end, we concentrate in this chapter on RS codes with small code length and keep the total code rate high. This is in contrast with most of the GLDPC proposals for optical communication [74] which use component codes with large code length. We present a numerical method to 1) systemically find a set-partitioning for any constellation (even those lacking symmetry) and 2) optimize code rates for that set-partition. Our method exploits the conditional PDF of received symbols ; the PDF takes into account channel effects, while we manipulate that PDF to capture the impact of constellation geometry. Although our method is independent of the number of constellation points, we focus in this chapter on 16-QAM. There are a number of reasons for this choice : 1) 16-QAM provides a good compromise between spectral efficiency and performance [76], 2) our set-partitioning and rate optimization method is very fast for 16-QAM, while it may become too slow for constellations having more points, and 3) the latency incurred by MSD grows with constellation size.

This chapter is organized as follows. We first present our first contribution, which is the methods for 1) set-partitioning and 2) rate optimization, in section 2.2. BLER performance is investigated using semi-analytical techniques in section 2.3 for two scenarios : A) a system limited by amplifier induced nonlinear phase noise and B) a system limited by Wiener phase noise. Investigation of the first scenario is our second contribution where we consider a 16-QAM ring constellation and compare our set-partitioning with another presented in [11] ; optimized rate allocations are found for both set partitions. We show that our set-partitioning leads to a rate allocation that improves BLER. Investigation of the second scenario is our third contribution which is covered in sections 2.3 and 2.4. In these sections, we consider two constellations for 16-QAM : 1) a square constellation and 2) a phase-noise optimized constellation introduced in [10]. The BLER is greatly improved [77] using the constellation

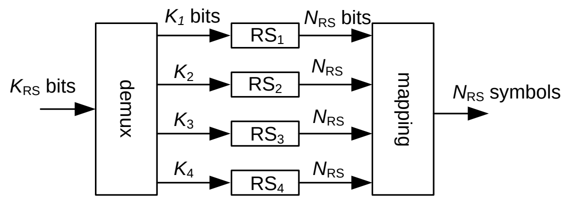


FIGURE 2.1 – Block diagram of MLCM encoder.

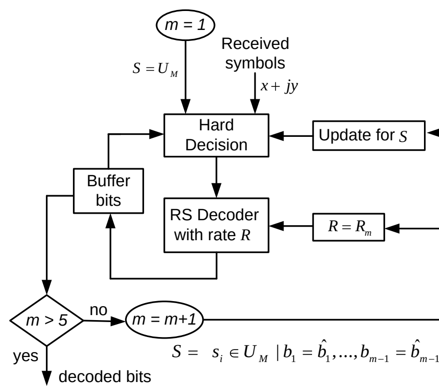


FIGURE 2.2 – Flowchart of MSD, m is the layer number, U_M is the set of all symbols, b_i is the ith bit of each symbol, $\hat{b}_i$ is the ith bit estimate (at ith layer).

from [10] with our set partition and rate optimization. We present Monte Carlo simulations of BER for the phase noise limited system of section 2.3. We implement both MLCM encoder and MSD. Our fourth and fifth contributions are also presented in section 2.4. We show that the set-partitioning and code rates that minimized BLER lead to the minimization of BER for ranges of SNR that are of practical interest. Low SNRs violate the assumption of zero error propagation in the BLER equations, and hence BER is suboptimal in these regions. We compare SNR advantage of the MLCM system with uniform rate system for both constellations and show that large SNR gains can be achieved at low BER, even if the set partition and rate allocation is suboptimal. The SNR advantage is examined for different levels of phase noise, and we find that our MLCM strategy leads to BER curves with the same improvement trends as symbol error rate (SER), validating an effective set-partitioning.

2.2 MLCM system design

The components of an MLCM system are first described, and then the method for designing the MLCM system is described in subsections.

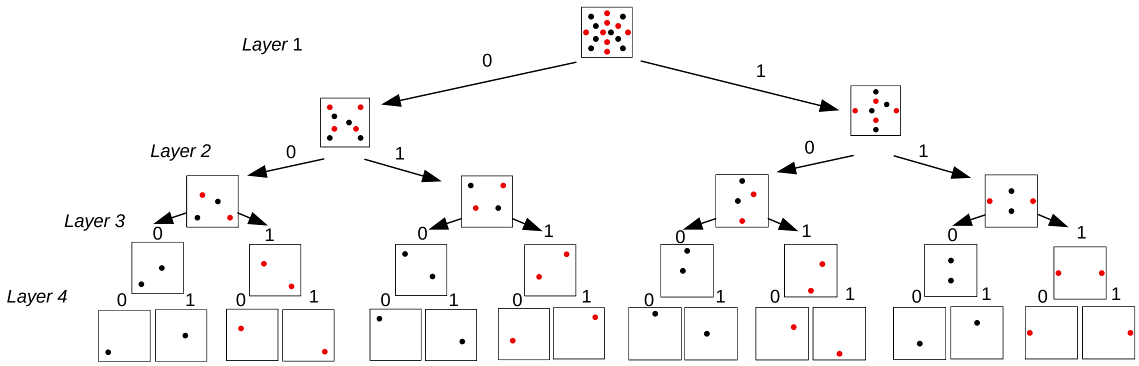


FIGURE 2.3 – Example of set partitioning (section 2B) for phase noise optimized constellation [10] at SNR = 20 dB and $\sigma_\theta^2 = 0.0185$ rad^2 ; red(gray)/black illustrate subsets feeding next layer.

2.2.1 MLCM encoder and MSD

In an MLCM encoder, a unique code rate R_m is found for each bit stream used to form symbols. For 16-QAM, the disparate code rates are implemented as illustrated in Fig. 2.1. At the encoder side, K_{RS} bits of information are demultiplexed into four blocks of (distinct) length K_m in the m^{th} encoder, and then encoded into four equal length blocks of N bits. In general, it is not necessary to have equal length output blocks of encoded bits, but we make this assumption for simplicity of the analysis. For a given total code rate $R = K_{RS}/4N_{RS}$, equal to the sum of component code rates $R_m = K_m/K_{RS}$, we optimize the vector of component code rates $[R_1 R_2 R_3 R_4]^t$.

A set-partitioning rule maps each four bit output symbol into a constellation point ; N 4-bit symbols per block are output by the encoder. According to the set-partitioning, the first bit (in the block of k_1 bits or layer 1) determines into which of two sets (each with 8 members) the symbol will fall. The second bit determines which one of the two subsets (each with 4 members) the symbol falls into, and so on. Each set-partition would require an optimized vector of component codes.

At the decoder side, the hard-decision MSD (HD-MSD) is composed of four RS decoders, one for each layer of set-partitioning. Decision regions are established so that the selected symbol has the smallest Euclidean distance from the received signal ; this leads to rectangular regions for square 16-QAM and more complex regions for the phase-optimized constellation. From the hard symbol decisions we extract all first position bits in the symbols to form a new sequence of layer-one bits. This sequence is input to the first decoder. The layer 1 decoder output is passed to layer 2. The decoded first bit determines the layer 2 subset of eight constellation points (see Fig. 2.2) that becomes the constellation used to make a new symbol decision. Once again the symbol with shortest Euclidean distance to the received signal is selected and now the sequence of second bits is sent to the layer-three decoder. This procedure is repeated for the third and fourth layers until the decoding is finished. In this way, possible errors in the encoded bits of an upper stage can be corrected up to the capacity of the upper RS code and a more reliable symbol decision can be made in the lower stage. A drawback of MSD is error propagation which means that decoding failure of a certain stage is propagated to the succeeding stages and increases the number of errors drastically. Although some methods are proposed to reduce impact of error propagation [78], these methods add complexity to the system.

It can be shown that using component codes whose rates are equal to the capacity of the layer are sufficient for the MSD to approach capacity [68]. The capacity of each layer equals the mutual information between the received symbol (a complex value) and the bit of a certain layer conditioned and averaged over bits in previous layers [68]. Although there are some techniques to reduce error propagation, they are either not very effective or costly [68]. Here our strategy to prevent excessive error propagation applies stronger codes at the top layers (first and second layers) and a set-partitioning devised for this strategy. In the set-partitioning and rate optimization method that we present in the next sections, we assign most of the allocated overhead of the FEC to the first and second layers.

2.2.2 Set-partitioning method

Set-partitioning determines the mapping of symbols to constellation points in the MLCM system. In order to approach capacity, it is not essential to use set-partitioning if there is no constraint on the component code lengths. Constraints are, however, inescapable in optical systems at high baud rate. Therefore, we consider RS codes with a fixed code length of 255 RS symbols. For this case of finite code length, set-partitioning promises the best performance [68].

Ungerboeck proposed maximizing Euclidian distance using a set-partitioning where the minimum distance between symbols in each set (i.e. intra-set minimum Euclidean distance) increases monotonically at each layer [8]. In an AWGN channel, maximizing Euclidean distance minimizes SER. We deal with a channel which is not AWGN, and therefore work with SER directly rather than Euclidean distance. Extension of the Ungerboeck set partitioning to irregular constellations is not straightforward. We adopt a heuristic approach that minimizes intra-set SER, thus guaranteeing that at each layer the conditional SER (SER provided there is no error in previous layers) decreases. For AWGN channel and square 16-QAM, our method yields the Ungerboeck set partitioning ; thus our extension to irregular geometries and non-AWGN channels is consistent with classic partitions.

When combined with MLCM, this leads to an optimal code rate vector whose elements are decreasing. This is advantageous as 1) it facilitates finding the optimal code vector (less prone to local minima) and 2) it leads to lower layers requiring weak codes (so weak that coding can be foregone to reduce complexity with little overall performance hit).

In order to calculate intra-set SER of candidate sets, we examine the conditional PDF $f(x, y|s_i)$ of received symbol $x + jy$ given that symbol s_i is transmitted ($i = 1, ..., M$). First we calculate the error probability p_{ij} that transmitted symbol s_i is received as symbol s_j ($i, j = 1, ..., M$ and $i \neq j$). These probabilities are found by numerically integrating $f(x, y|s_i)$ over the decision region for symbol s_j. The probabilities p_{ij} form a $M \times M$ matrix whose diagonal elements are zero ($p_{ii} = 0$).

We partition the M point constellation into two sets, $S(l)$ and its complement $S^c(l)$, each having $M/2$ elements. We denote the set of all possible partitions as F. For the l^{th} partition, the intra-set probability of symbol error for set $S(l)$ is

$$p_{SER}(l) = \sum_{x_i \in S(l)} \sum_{x_j \in S(l)} p_{ij} \tag{2.1}$$

found by summing the elements of the matrix $[p_{ij}]$ whose row and column indexes correspond to symbols in the set $S(l)$. We calculate a similiar error probability for set $S^c(l)$ and denote it as $P^c_{SER}(l)$. We define the optimal partition $S(l_{opt})$ to be that which minimizes the maximum of these two probabilities over all the members $S(l)$ of F, i.e. the minimax probability.

$$P_{SER}(l_{opt}) = \min_{S(l) \in F} \left\{ \max[P_{SER}(l), P^c_{SER}(l)] \right\} \tag{2.2}$$

In this way, we choose a pair of sets for the first layer whose maximum of intra-set SER is minimized. Recall that in an irregular constellation, intra-set SER will not be identical in the two subsets forming the partition. Performance will be dominated by the subset with the largest SER, hence this is the effective SER for the partition and it should be minimized. For subsequent layers, we start with each set in the preceding layer and apply the same methodology. That is, all partitions of that set are examined to find the minimax partition. We proceed until only sets with two elements remain (assuming M is a power of 2). As illustrated in Fig. 2.3 each splitting of a set into a partition is assigned to a bit, thus establishing the mapping of bits to constellation points (or equivalently symbols to constellation points). For AWGN channel and square 16-QAM, this method yields the Ungerboeck set partitioning.

In the set-partitioning algorithm described, finding the first pair of sets needs a search over M choose $M/2$ partitions. The number of partitions grows quickly with M, but for $M = 16$ the number of partitions (12870) is reasonable for a brute-force search. A typical result of set-partitioning for a phase noise optimized 16-QAM is shown in Fig. 2.3. A combination of SNR and phase noise determines the final set-partitioning.

2.2.3 Code rate optimization method

In this section, we describe a simple code rate optimization strategy. We propose the use of RS coding for each of the layers in the set partition, as illustrated in Fig. 2.1. We seek the vector of code rates $[R_1 R_2 R_3 R_4]^t$ that minimizes the BER, subject to a constraint on the total code rate (sum over component code rates). An analytical expression for BER is intractable, so we instead find the vector minimizing the block error rate BLER.

Let $p_b(m)$ be the probability of bit error for layer m, assuming no bit error was made in any previous layer. This probability can be estimated by Monte Carlo (MC) simulation, or calculated numerically from the final set partition and the matrix of p_{ij} (the probability of choosing symbol j when symbol i was transmitted). In this chapter, we use the latter approach and calculate $p_b(m)$ by following equation

$$p_b(m) = \frac{1}{M}\left(\sum_{x_i \in S_0(m)} \sum_{x_j \in S_1(m)} p_{ij} + \sum_{x_i \in S_1(m)} \sum_{x_j \in S_0(m)} p_{ij}\right) \tag{2.3}$$

where $S_0(m)$ and $S_1(m)$ are sets of symbols corresponding to bits 0 and 1 at layer m respectively.

The probability that a block of K bits, see Fig. 2.1, contains an error is the block error rate (BLER). The BLER for RS component codes can be approximated using the union bound to be [11].

$$\begin{aligned} BLER &\approx \sum_{m=1}^{4} \sum_{i=t_m+1}^{n} \binom{n}{i} p_s^i(m)(1-p_s(m))^{n-i} \\ &\approx \sum_{m=1}^{4} \sum_{i=t_m+1}^{n} \binom{n}{i} 8^i p_b^i(m)(1-8p_b(m))^{n-i} \end{aligned} \tag{2.4}$$

where n is the number of output symbols in an RS code with error correction capacity of $t_m = \lfloor n(1-R_m)/2 \rfloor$ and p_s is the RS symbol error rate and p_b its corresponding BER. As $p_b(m)$ assumes no error in previous layers, the BLER estimate is optimistic. Although (4) is approximate and deviates at low SNR from true BLER, optimization of code rates is usually performed in a region where BLER is small enough for very low post FEC BER. In these regions, (4) is accurate.

As our minimax set partition strategy yields layers with decreasing symbol error rates, we can assume that $R_1 < R_2 < R_3 < R_4$ will give best overall performance. This method is applicable to constellations of arbitrary size, however, we use a brute-force search to find the optimal code rate vector for 16-QAM which could become too slow for larger constellations. Candidate vectors are used to evaluate BLER in (2.4) ; we choose the set of code rates which minimizes the BLER. The code vector varies with system parameters such as SNR, phase noise or nonlinearity. The brute-force search can be improved by tweaking a few parameters to reduce the search space. The tweaking parameters must also be adjusted accordingly.

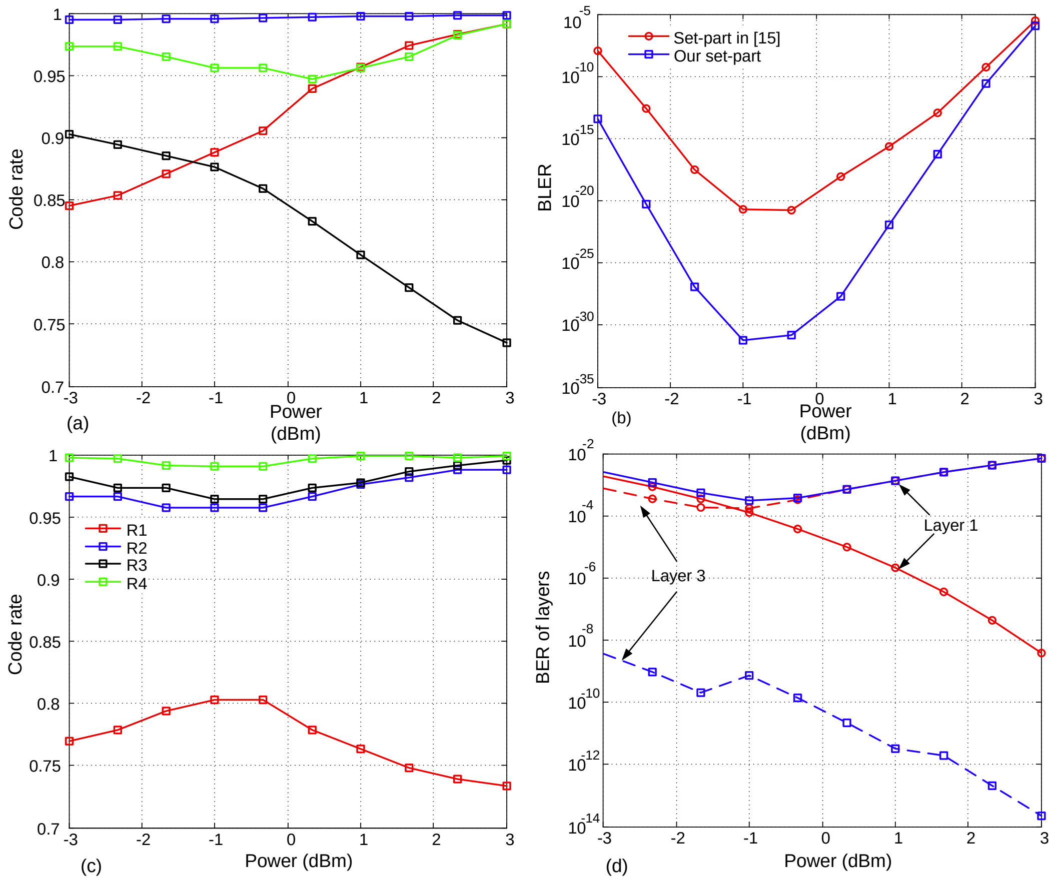


FIGURE 2.4 – (a) Optimized code rates vs. power for set-partition in [11] (b) BLER vs. power for our set-partition (squares) and set-partition in [11] (circles) (c) optimized code rates vs. power for our set-partition (d) uncoded BER vs. input power of layers 1 and 3 for two different set-partitions.

The complexity of the brute-force search can be reduced by compromising between computational complexity and accuracy of the optimal code rates. For example, a minimum code rate in the system can be set. The number of points searched will fall between the minimum code rate and one, and can be reduced by reducing the desired code rate resolution. In addition, other constraint optimization methods can be employed for large constellations [79].

2.3 Investigation of BLER improvement

In the following sections we consider three 16-QAM constellations. We first consider a ring constellation for a nonlinear phase noise (NLPN) system. Next, we examine the square constellation and a phase-noise optimized constellation with weak symmetry. For a given constellation, the set partition is critical to keeping BER low for a given symbol error rate performance. In this section we use BLER as an indicator of BER performance ; BLER can be found with semi-analytical techniques. In section 4 we will use Monte Carlo techniques to find BER performance and validate our assumption that optimizing BLER leads to optimization of BER for reasonable regions of signal to noise ratio.

An ad hoc set partition applicable to ring constellation in an NLPN channel was found in [11] ; we show in section 2.4 that our numerical technique for set partitioning leads to lower BLER. In [10] the SER performance of the phase-noise optimized constellation was shown to be superior to square 16-QAM, but the relationship between SER and BER for the optimized constellation was not examined. For a constellation with such weak symmetry, no Gray-coding can be applied. We show that combining set partitioning and rate optimization significantly improves BLER performance for the phase-noise optimized constellation.

2.3.1 Nonlinear phase noise limited system

In this section we apply our set-partitioning and rate optimization strategy to a coherent detection system affected by nonlinear phase noise (NLPN). A 16-QAM ring constellation is employed in [11] for NLPN, and a set-partitioning method is proposed where radial and angular set-partitioning are done separately ; an MLCM system with optimized RS component codes is used. A nonlinear post compensator for NLPN introduced in [80] is employed at the receiver.

The PDF of received symbols after NLPN compensation can be derived analytically if chromatic dispersion is ignored [80]. We use this PDF to calculate error probabilities p_{ij} and find set-partitioning and optimal code rates using the methods described in section 2. Our simulations used the system parameters from [11] : fiber length of 5000 km, baud rate of 42.7 Gbaud and total code rate of 0.929. We also recreated results from [11] using their published set-partition and a numerical search for the best rate allocation. Fig. 2.4a gives our recreation of the optimal code rates for the set-partition of [11], in close agreement with Fig. 6 of [11].

In Fig. 2.4b we present the optimized BLERs for the two set-partitions. The optimized BLER from the ad hoc radial/angular set-partition in [11] (red circles) is significantly worse than BLER using our minimax set-partition strategy (blue squares). MLCM is clearly more effective when combined with a methodical set-partition strategy exploiting knowledge of channel characteristics (probabilities p_{ij}). The optimal code rates for our set-partitioning are shown in Fig. 2.4c.

As expected, our set partition strategy leads to coding rates that decrease from lower (R_4) to upper (R_1) layers, as opposed to the variable nature of rates in Fig. 2.4a. The clumping of rates close to one for the first three layers in Fig. 2.4c for a wide range of input powers could be exploited to reduce complexity of the MLCM encoder, a side benefit of the minimax set partition strategy. In contrast, only the first layer (see Fig. 2.4a) has this property for the set-partition in [11]. In Fig. 2.4d we present $p_b(m)$, the BER of layer m ($m = 1$ and 3) assuming no bit error in previous layers. We see that the set partition of [11] does not have the property of decreasing $p_b(m)$ with increasing m.

For the NLPN channel we have assumed the use of a post-compensation algorithm. The set partitioning proposed offers improved performance, but at the cost of an increase in receiver complexity in order to implement the nonlinear post-compensator. The NLPN post- compensation algorithm requires an estimate of the symbol amplitude. This amplitude estimate is obtained in [11] from the detection algorithm where set-partitioning is naturally divided into two steps – one slicing for amplitude, and another slicing for phase. In our set partition the detection slicing is similar to a square 16-QAM constellation where both amplitude and phase must be taken together to select the most likely symbol.

In [11] the set-partition leads to a detection algorithm as follows : 1) amplitude selection/estimation, 2) calculation and application of NLPN post-compensation, and 3) phase selection. In the case of our set-partition, the detection would be : 1) selection of most likely symbol (with phase likely in error due to NLPN), 2) amplitude of selected symbol used to calculate and apply NLPN post-compensation, and 3) re-selection of most likely symbol (with NLPN compensated). Note that both in [11] and in our analysis of BLER, we assume that the estimation process provides perfect estimates of the received amplitude, so both approaches offer lower bounds for BLER. The bound in [11] is arguably tighter, as amplitude can be decoded before phase, but the enhanced performance of the new set partition could be approached by introducing an explicit amplitude estimation algorithm (increasing receiver latency and/or complexity).

2.3.2 Phase noise limited system

We consider in this section a coherent system with both AWGN and phase noise, which is ultimately limited by Wiener phase noise. This arises, for instance, for a short reach or metro area system where the impact of amplifier induced nonlinear phase noise is small and chromatic dispersion is compensated. We consider two 16-QAM constellations : 1) square and 2) a phase noise optimized constellation introduced in [10]. For the channel considered, the PDF of received sample $x + jy$ for a given transmitted symbol $a + jb$ is given by [81]

$$f(x,y|a,b) = \frac{1}{\sqrt{2\pi}\pi\sigma^2\sigma_\theta}\int_{-\pi}^{\pi} e^{-\left\{\frac{(x-a\cos\theta+b\sin\theta)^2}{\sigma^2}+\frac{(y-b\cos\theta-a\sin\theta)^2}{\sigma^2}+\frac{\theta^2}{2\sigma_\theta^2}\right\}} d\theta \tag{2.5}$$

where σ^2 and σ_θ^2 are the noise variance of ASE and phase error respectively. For a given average SNR over the complete constellation, SNR $= 1/\sigma^2$, and a, b are normalized to give unitary average signal power. Phase error variance is a function of the phase noise level (quantified by the bandwidth-linewidth product $\Delta\nu T_s$) and the phase tracking algorithm employed. Phase error variance must be found numerically using either simulated or experimental data. In this chapter we use a decision-directed, moving-average algorithm for carrier recovery and use numerical simulation to estimate σ_θ^2 for a given $\Delta\nu T_s$. We use (2.5) to calculate error probabilities p_{ij} for (2.4).

The SER of the optimized [10] and square 16-QAM constellations are compared in Fig. 2.5a, reproducing the result in [10] for $\Delta\nu T_s = 5 \times 10^{-4}$. The optimized constellation shows much better SER performance in the presence of large phase noise as compared to square 16-QAM. While Gray-coding can be applied to square 16-QAM, it is not compatible with the optimized constellation ; hence the relationship between SER and BER for the optimized constellation must be examined. We consider two approaches for the mapping of bits to 16-QAM symbols, and the assignment of forward error correction. In the first simplistic approach we use a single RS encoder and decoder. Gray coding for square 16-QAM and a Gray-inspired mapping (where we attempt minimum bit changes for symbols in close proximity) are used. This mapping and FEC would be appropriate for an AWGN channel. In the second approach an MLCM approach is considered where the exact error probabilities are calculated from (2.1. The probabilities are used to find the set-partition for the mapping of bits to symbols, as well as the optimal code rates for that set-partition per the methodology presented in section 2.2.2.

BLERs shown in Fig. 2.5b were calculated via (2.4) for both square and optimized 16-QAM constellations with either optimized (MLCM) or non-optimized (uniform rate) coding rates. For a fixed coding strategy (flat rate or MLCM), the phase-noise-optimized constellation (circles) performance equals or outperforms the square 16-QAM constellation (squares). The Gray-inspired mapping (dashed curves), which calls for a flat rate FEC, cannot exploit the probabilities of error that vary over bits in the QAM symbol, hence MLCM (the solid curves) always offers better performance. Clearly MLCM with an appropriate set partitioning can significantly enhance the use of a phase-noise-optimized constellation as seen in Fig. 2.5b. For the optimized constellation, constituent code rates minimizing BLER are plotted versus SNR in Fig. 2.5c for different layers.

Taking the non-Gaussian metric BLER, solid lines in Fig. 2.5b for $\Delta\nu T_s = 5 \times 10^{-4}$, we vary the level of phase noise $\Delta\nu T_s$ and compute the SNR advantage in using the phase-noise-optimized 16-QAM constellation instead of the square constellation. The advantage for several phase noise levels are presented in Fig. 2.5d versus BLER. For lower phase noise there is no gain in using the optimized constellation ; this is not surprising as the optimized constellation was developed assuming large residual phase offset. The phase error variance σ_θ is a function of both laser linewidth and the phase estimation algorithm. The optimized constellation offers robustness to this residual phase error, which can arise due to a wide linewidth laser, or feedback delay in a parallelized DSP architecture. The SNR advantage grows rapidly as phase error increases.

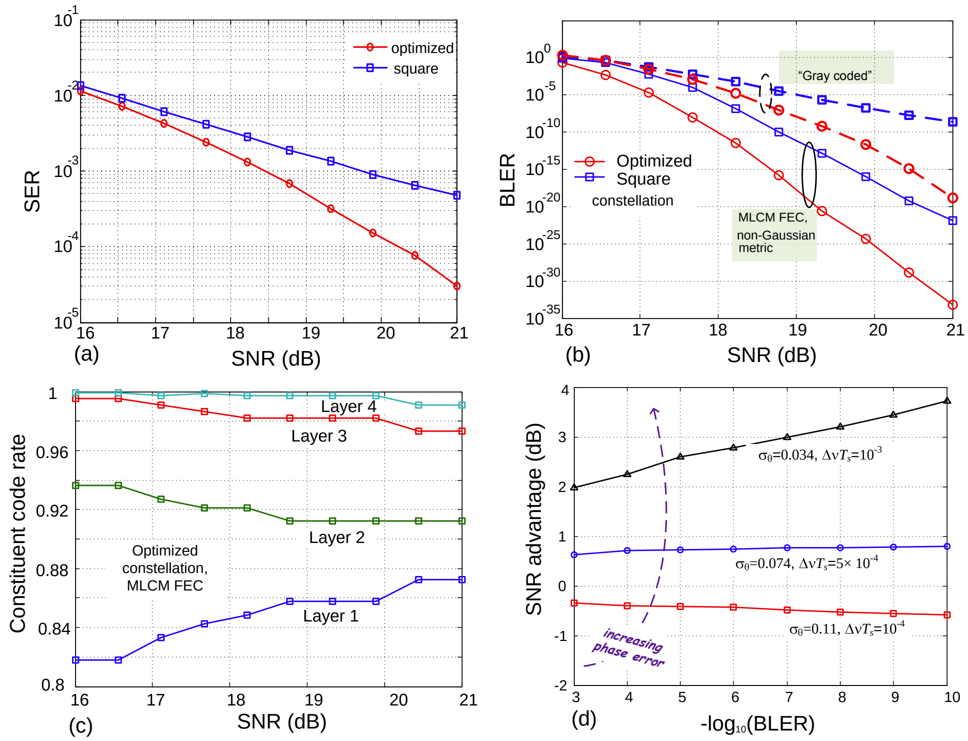


FIGURE 2.5 – (a) SER vs. SNR for two constellations, (b) BLER vs. SNR after coding comparing MLCM system (solid) with flat rate (dashed) system, (c) constituent code rates minimizing BLER for optimized 16-QAM constellation vs. SNR, (d) SNR advantage of optimized over square 16-QAM constellation for three phase error levels. Results in (a)-(c) for $\Delta\nu Ts = 5 \times 10^{-4}$.

2.4 Investigation of BER improvement

In this section, post FEC BER of the MLCM system is investigated using MC simulations. Our aim is to investigate the accuracy of BLER as a figure of merit for designing MLCM systems (based on the method presented in section 2.2.1) and verify BER improvement of phase noise optimized versus square 16-QAM constellation, and MLCM versus uniform rate coding.

BER simulations are performed in MATLAB^{TM} using RS encoders and HD-MSD. The MLCM encoder consists of four RS(n,k_i) encoders each having the same code length of $n = 255$ RS symbols. The number of input bits k_i for each RS encoder and mapping of bits to symbols are set based on the optimal code rates and the set-partitioning found using the method in section 2.2.2.

The BER versus SNR is shown in Fig. 2.6a for three different FEC systems applied to square 16-QAM. The solid and dotted curves represent an MLCM system where code rates of RS encoders at each point of SNR are optimized using the method in section 2.2.1. The difference between solid and dotted curves is in the implementation of the MSD ; in the solid curve, the error propagation in the MSD is disabled in the simulation by providing true encoded bits of each decoder to the succeeding stages for decision (see Fig. 2.2 flow chart). We can observe that although the impact of error propagation at low SNR is large, with increasing SNR the two curves converge. It should be noticed that the BLER used for rate optimization is an accurate measure of the MLCM system performance without error propagation (solid curve). Nonetheless, the BER curve without error propagation is an unrealistic indicator of system performance as the true encoded bits are unknown to the receiver. This BER curve is shown here simply as a reference, to identify regions where this curve and the realistic BER curve with error propagation converge, i.e., at high SNR. This convergence confirms the validity of the BLER approximation for regions having low post-FEC BER. The dashed curve in Fig. 2.6a shows BER of a uniform rate system having a single RS encoder and decoder with the same total code rate as MLCM system. The advantage of the MLCM system over uniform rate one at high SNR is visible in this figure. In Fig. 2.6b, SNR advantage of the MLCM system over uniform rate system is shown for square and optimized constellations versus BER. It can be observed from Fig. 2.6b that MLCM is superior to uniform rate system for both constellations at low BER which is of practical interest. It should be noticed that although SNR advantage of the optimized constellation in Fig. 2.6b is smaller than square one, its BER at high SNR is smaller than square constellation for a fixed coding schema (MLCM or uniform rate).

BER simulation of the MLCM system for optimized constellation is shown in Fig. 2.7a for different levels of phase noise. Again at each point, the set-partitioning and code rates are found using the method in section 2.2.1 based on BLER minimization. In Fig. 2.6b, SNR advantage of the MLCM system for optimized constellation over square 16-QAM is shown for three levels of phase noise. It can be observed that the curves follow the same trend as Fig. 2.5d for BLER ; for $\Delta\nu T_s = 10^{-4}$, there is no advantage in using optimized constellation, there is slight advantage at $\Delta\nu T_s = 5 \times 10^{-4}$ and this is increased at $\Delta\nu T_s = 10^{-3}$. This figure obtained for an MLCM system with error propagation, confirms that BLER can be used as a measure for BER minimization even in a realistic MLCM system assuming error propagation.

2.5 Discussion

In [82] independent low density parity check (LDPC) decoders are used, while we adopt a multi-stage decoder (MSD) and design the system based on this assumption. We find the set-partitioning and code rates in such a way to have decreasing code rates from top to bottom layers. In this way, the decoders and encoders do not have the same complexity as in [82]. The reduced complexity of the encoders and decoders in our scheme can compensate the complexity added by MSD.

We cannot compare our results with [82] due to differences in codes, decoders and overhead (overhead in [82] is $\sim 30\%$ and in our work $\sim 7\%$). More generally, an issue in implementing a soft decision SD-MLC to compare with our HD-MSD is finding an accurate equation for BLER for rate optimization. Equation (4) used for BLER optimization is a tight bound, at least at high SNR, for RS codes. However, the same equation may not be applicable to LDPC codes and more research is needed to find an accurate equation for BLER.

In both [73] and our scheme, the PDF of the received symbols in the presence of phase noise is exploited. This PDF is used in our scheme to find set-partition and code rates, and is used in [73] to find log-likelihood ratios. Based on Fig. 2 in [73] and our Fig. 2.7 we can compare the results. The LDPC code rate is 0.8 in [73] which is close to the rate 0.9373 in our thesis. In addition, both in [73] and Fig. 2.7, we have curves for the same amount of phase noise; $\Delta\nu Ts = 10^{-3}$ and the same number of constellation points; i.e. 16 points. In Fig. 2 of [73], the OSNR for BER of 10^{-6} is 12.4 dB for OSCD constellation and 13.2 dB for square 16-QAM. In Fig. 2.7a, the curve with triangle markers shows the BER for a 16-point optimized constellation with $\Delta\nu Ts = 10^{-3}$. At BER of 10^{-6}, the SNR is around 17.4 dB, or OSNR = 19.4 dB.[1] While we require around 5 dB more OSNR (comparing to OSCD in [73]), we should note that the stronger FEC with more overhead is used in [73]. In addition, we should note that systems and DSP may be different and optimal bit mapping is not used in [73]. The examination of soft decision MSD in the context of constellation optimization in the presence of phase noise is beyond the scope of this thesis and a question for future research, as is the relative performance of LDPC and RS codes.

2.6 Conclusion

A numerical method was proposed for designing an MLCM system with RS component codes. This method is based on known PDF of the channel and is applicable to both AWGN and non-AWGN channels. The channel PDF is exploited to find a set-partitioning optimizing a minimax criterion and a rate allocation minimizing BLER under the assumption of zero error propagation. The method is particularly useful in application of MLCM to coherent optical channels where large levels of phase noise could be present. We consider a 28 Gbaud coherent system with a coding scheme of total overhead 7%. The total linewidth (LO and Tx) is changed from 2.8 to 28 MHz.

Application of this method to three different constellations and two different channels was studied. We first considered a ring constellation in a coherent system impaired by AWGN and nonlinear phase noise and found optimal set-partitioning and code rates. We compared the system BLER results with the BLER of previous work which used an ad-hoc set-partitioning and showed large BLER improvement. We then considered square and a phase noise optimized

1. SNR is related to OSNR by $OSNR = (R_s/2)/B_{ref} * SNR$ [83]; $R_s = 28$ Gbaud being symbol rate and B_{ref} the reference bandwidth for the OSNR taken to be 12.5 GHz.

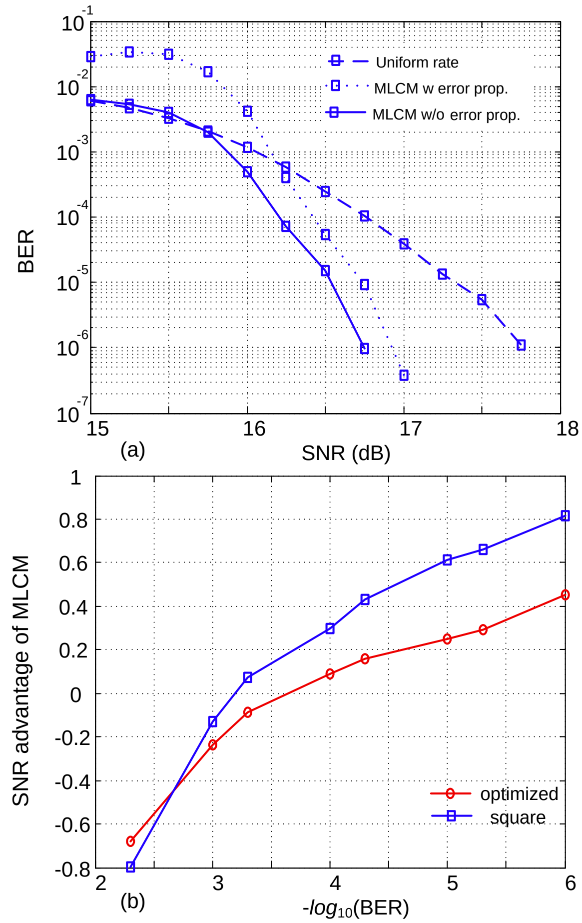


FIGURE 2.6 – (a) Post-FEC BER of square constellation for MLCM system without error propagation (solid), MLCM with error propagation (dotted) and uniform rate (dashed) (b) SNR advantage of the MLCM system with error propagation over uniform rate system versus post FEC BER for square (square markers) and optimized (circles) constellation. In all the curves $\Delta\nu Ts = 5 \times 10^{-4}$ and total rate is 0.9373.

16-QAM constellation in a coherent system impaired by AWGN and Wiener phase noise. We compared BLER of the optimized and square 16-QAM constellations in an MLCM system designed by our method.

Large improvement of BLER was observed for the optimized constellation in the regime of the optimality of the constellation. We also compared BLER of the MLCM and uniform rate systems (using RS code) and observed large SNR gains. In order to verify performance of the designed MLCM system in realistic implementation, we performed MC simulations using MLCM encoder and MSD in a coherent system. We studied BER with and without error propagation in the MSD. The BER performance of the MLCM system without error propagation is always better than that of uniform rate (with the same total rate) system.

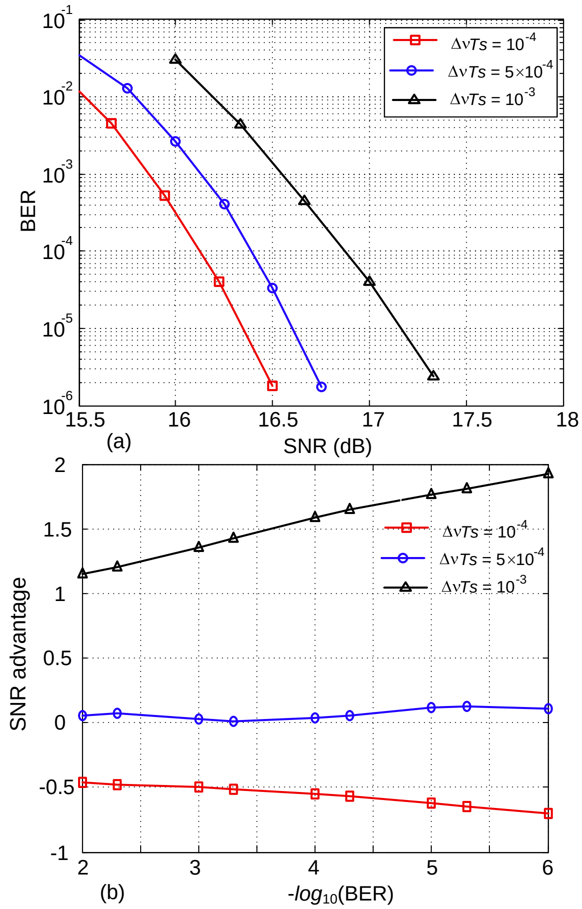


FIGURE 2.7 – (a) Post FEC BER of MLCM with error propagation for optimized constellation (b) SNR advantage of optimized versus square constellation when MLCM system with error propagation is used.

Although the BER performance of the MLCM system (with error propagation) at low SNR is worse than that of uniform rate system due to large error propagation in MSD, at low post FEC BER (where most of the practical systems operate) the MLCM system has an advantage. At BER of 10^{-6} and $\Delta\nu Ts = 5 * 10^{-3}$, the SNR advantage of the optimized and square 16-QAM over its uniform rate system is around 0.4 and 0.8 dB respectively. In our designed MLCM system, optimized constellation has negligible advantage over square 16-QAM at BER of 10^{-6} and $\Delta\nu Ts = 5 * 10^{-3}$. This advantage increases to around 2 dB at $\Delta\nu Ts = 10^{-3}$. In addition, the SNR advantage extracted from BER curves of the optimized and square 16-QAM for different levels of phase noise shows the same trend as the same curves for BLER. This confirms validity of the set-partitioning and code rates obtained assuming no error propagation.

Chapitre 3

Experimental verification of MLCM for 16-ary constellations

3.1 Introduction

Phase noise is a major impairment in transmission of 16-ary and higher order modulation formats used in optical coherent systems. High levels of phase noise can originate from linewidth of lasers, nonlinear effects or phase estimation algorithms. For instance, both parallelization of phase estimation algorithms with feedback and reducing complexity of feedforward algorithms increase phase noise. As observed in previous chapter, approaches based on MLCM are interesting as robustness toward other impairments can be achieved simultaneously [38, 84]. In addition, MLCM based on low complexity component codes like RS codes provides a good compromise on performance, overhead and complexity. MLCM has been studied by theory and simulation for optical systems [84, 41]. However, there is no experimental demonstration using MLCM for phase noise limited systems. Most experimental works consider a bit-interleaved coded modulation based on LDPC [85, 86, 87, 88].

In this chapter, we present an experimental demonstration of MLCM coding with RS component codes for 16-ary constellations. MLCM coding is designed using the approach in previous chapter. The MLCM coding is applied to two 16-ary constellations : a phase noise optimized 16-QAM proposed in [10] and standard square 16-QAM. Post FEC BER of the MLCM coding is calculated from offline processed data. By sweeping phase noise, we identify the phase noise regime where the optimized constellation outperforms square 16-QAM. In this chapter, phase noise is swept by an approach employed in DSP which effectively changes linewidth by symbol time product $\Delta\nu T_s$ by multiplying it with a factor M ; $M\Delta\nu T_s$. Experimental post FEC BER of uniform rate coding with a single RS encoder and decoder is also calculated and compared with MLCM coding. We show that there is up to 3 dB advantage in MLCM coding compared to uniform rate coding with identical overhead. It is also shown that

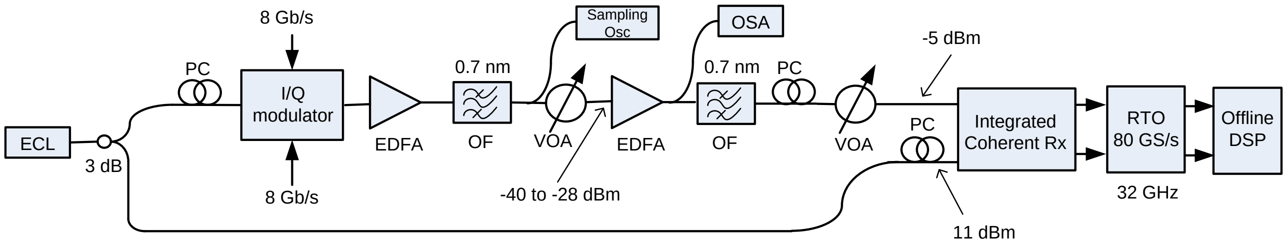


FIGURE 3.1 – Experimental setup, PC : polarization controller, OF : optical filter, EDFA : erbium dopped fiber amplifier, ECL : external cavity laser, RTO : real-time oscilloscope, VOA : variable optical attenuator, OSA : optical spectrum analyzer, DSP : digital signal processing.

the optimized constellation with MLCM coding, designed for a fixed $M\Delta\nu T_s$, has advantage over uniform rate coding and square 16-QAM with MLCM coding over a wide phase noise regime.

The MLCM coding in this chapter is designed by minimization of BLER using a formula that ignores error propagation in MSD. Nonetheless, experimental results show that the designed MLCM coding yields good post-FEC BER. We also examine the pre-FEC BER threshold of the MLCM coding. We find the threshold decreases with increasing phase noise. The chapter is organized as follows. We present in the next section parameters of MLCM coding used in the experiment. Then, the experimental setup is described and finally experimental results are presented.

3.2 Designing an MLCM system

The method in previous chapter is used to design the MLCM coding. We find the set-partitioning and optimal constituent code rates for a fixed SNR and phase error variance σ_θ. Phase error variance σ_θ depends on the product of linewidth $\Delta\nu$ by symbol time T_s and the carrier phase recovery algorithm. The method in previous chapter assumes that phase error is Gaussian distributed with zero mean and variance of σ_θ^2. The MLCM coding in this chapter is designed at OSNR of 15.5 dB and $\sigma_\theta = 0.11$ rad derived from $\Delta\nu T_s = 10^{-3}$ and DD carrier recovery. This design is fixed, despite our sweeping values of SNR and effective $\Delta\nu T_s$.

We design an MLCM code for each of the two constellations assuming total code rate of 0.9373. First we find the set-partitioning and then we minimize BLER. The optimal code rates for the optimized 16-QAM and square 16-QAM are [0.8586, 0.9116, 0.9800, 0.9989] and [0.7639, 0.9904, 0.9956, 0.9994] respectively. These correspond to the code rates of four RS codes each having 255 RS symbols at the output. The input symbols of RS encoders are [217, 231, 249, 253] and [193, 251, 253, 253] for the optimized 16-QAM and square 16-QAM respectively.

3.3 Experimental setup

We first describe generation of electrical signals for two constellations. Then the setup is described in detail and the DSP functions used in offline processing are explained. Finally, our method for sweeping phase noise in offline processing is discussed.

3.3.1 Generation of electrical signals

Transmitted bits are generated using a PRBS of length $2^{31}-1$. Transmitted bits are coded with a single RS encoder or four RS encoders in the uniform rate or MLCM coding respectively. The total number of encoded bits is $224,800$. We assign four bits to each symbol based on a mapping specific to each modulation and coding scheme. Gray coding is used for uniform rate coding while a mapping based on the designed set-partitioning is used in the MLCM coding. An arbitrary waveform generator (AWG) including an 8-bit resolution digital-to-analog converter (DAC) is used to generate 8 Gbaud symbols for either constellation. In-phase (I) and quadrature phase (Q) signals for each constellation are generated separately and then quantized to 64 levels before uploading to the DAC. Each data packet consists of 256×1024 points of the AWG field-programmable gate array (FPGA) where 240×1024 points are used for symbols and the rest for zeros serving as a guard time between different packets in a single capture of RTO data with 2 million points. Non-return to zero (NRZ) pulse shape with four points per symbol is used in the pulse generation. After uploading I and Q signals to the DAC, the quality of the generated RF signals are verified using a sampling oscilloscope.

3.3.2 The setup

The experimental setup is shown in Fig. 3.1. We use a homodyne coherent detection setup where an external cavity laser (ECL), with linewidth of 15 kHz, is used as both transmitter laser and LO. The optical signal is modulated by 8 Gbaud RF signals and then amplified by an EDFA. The second EDFA is used to load noise on the signal. Signal power entering the second EDFA is changed using a VOA from -40 to -28 dBm to sweep OSNR by steps of 1 dB. Signal spectrum after the second EDFA is monitored by an OSA with reference bandwidth of 0.1 nm and the OSNR is estimated. The second VOA is used to fix signal power entering the receiver at -5 dBm. The LO power is fixed at 11 dBm. Our setup is a single polarization coherent detection ; we maximize signal power by adjusting polarization controllers (PC). The signal and LO are mixed in an integrated coherent receiver and the electrical signal is sampled at 80 GS/s with a real-time scope having analog bandwidth of 32 GHz. The captured signal is then processed offline by DSP.

In the DSP unit, we first extract the data packets (maximum of six packets in each capture). Each packet is processed by dividing it into 40 frames (consisting of 1530 symbols) and processing each frame separately. The processing includes low pass filtering (LPF) with a super-Gaussian filter of bandwidth 0.7 $\times$ baud rate, timing recovery and down sampling to one sample per symbol (captured data contains 10 samples per symbol), MMSE filtering, phase noise estimation, symbol demodulation, processing training symbols, bit demodulation and decoding. The training symbols are used to update 20 taps of an MMSE filter and as updated reference symbols used in the decision units of phase estimation and symbol demodulation. DD carrier phase estimation is used in the DSP unit and the block length of the moving average is optimized in each frame for minimum SER. In the decoder, we use either 1) a single RS decoder for uniform rate coding, or 2) an MSD for MLCM coding. We consider the impact of error propagation in the MSD implementation. The post FEC BER is calculated using the known transmitted bits for each data packet. The maximum number of transmitted bits is around 12 million.

3.3.3 Sweeping phase noise

In the experimental setup, a narrow linewidth laser is used as transmit and LO laser. However, we are interested in exploring MLCM performance in a phase noise limited regime for which the MLCM coding is designed. In addition, the advantage of the optimized constellation over square 16-QAM will be observed at much larger linewidth. As sweeping of linewidth by employing different lasers is inconvenient as it requires lasers with a good distribution of linewidths, which may not be available. In this chapter, we change $2\Delta\nu T_s$ (factor of 2 accounting for the same linewidth of transmit and LO lasers) by applying a fixed phase noise estimate to M consecutive samples in the phase estimation unit. In this way the linewidth/symbol time product is multiplied by M; the effective product becomes $2M\Delta\nu T_s$. It can be shown that, assuming a Wiener process for phase noise, the statistics of phase error are the same as that of a linewidth M times larger, while the phase error distribution remains Gaussian.

3.4 Experimental results

In this section, SER and post-FEC BER of optimized and square constellations for MLCM system are compared. Then advantage of MLCM coding over uniform rate coding is verified for the phase noise optimized constellation.

3.4.1 MLCM performance for optimized and square 16-QAM

The optimized 16-QAM constellation is designed to minimize SER assuming a fixed phase rotation [10]. In practical systems the phase error is not fixed, but rather statistical in nature, and may be modeled as having a Gaussian distribution. The distribution is parameterized by the level of phase error (linewidth/time interval product), and influenced by the particular phase tracking algorithm used. We find a baseline for performance of square and optimized 16-QAM constellations by examining SER to understand to what extent the constellation designed for fixed offset is effective for random phase error. The use of MLCM will ensure that SER gains are efficiently transferred to BER as well.

The SER for two constellations is shown in Fig. 3.2 from both experiment (solid curves) and simulation (dashed curves). We note that as expected, for very low phase noise square 16-QAM outperforms optimized 16-QAM. For $2M\Delta\nu T_s > 10^{-3}$, however, the optimized 16-QAM has better SER. The phase noise level in the experimental results is varied by changing M in DSP. The linewidth is approximately $\Delta\nu = 15$ kHz for the ECL laser used. We sweep $2M\Delta\nu T_s$ from 3.7×10^{-6} to 0.003 by varying M. The experimental SER curves for $M = 1$ in Fig. 3.2 are in close agreement to the SER curves reported in [10] for optimized and square 16-QAM which is reported using an ECL laser. We observe $\sim$ 1 dB of penalty between simulation and experiment, which can be explained by filtering effects and quantization noise in the constellation coordinates.

For $2M\Delta\nu T_s = 1.5 \times 10^{-3}$ ($M = 400$), the Gaussian distributed phase error has variance σ_θ near 0.11, the value used to find the MLCM coding parameters. This coding strategy remains fixed, even as we examine other values for $2M\Delta\nu T_s$. The experimental post-FEC BER of MLCM coding for optimized and square 16-QAM constellations are shown in Fig. 3.3. For a fixed M, post-FEC BER is calculated after transmitting 12 million bits. In Fig. 3.3, a downward arrow indicates the post-FEC BER is less 10^{-6} for that and higher OSNR ; i.e., there were no bit errors.

To quantify the advantage of MLCM coding with an optimized constellation, we define OSNR_{req} as 1 dB plus the maximum OSNR for which post-FEC errors are detected. For example OSNR_{req} for $M = 400$ in Fig. 3.3 (a) is 16 dB. The advantage in OSNR_{req} for optimized over 16-QAM constellation is seen for $M = 400$ and greater and at $M = 600$ optimized constellation requires at least 2 dB less OSNR. The experimental OSNR advantage of the optimized 16-QAM over square 16-QAM is close to the SNR advantage reported in Fig. 2.7 (b) by simulation ; the baud rate in Fig. 2.7 (b) is different and the phase estimation method varies slightly from that used in this chapter.

Pre-FEC BER for the optimized constellation is shown in Fig. 3.4. The pre-FEC BER required to achieve zero error count post-FEC is decreased from 10^{-3} for $M = 200$ to 6×10^{-4} for $M = 600$. This is in contrast with uniform rate coding with a fixed pre-FEC BER threshold related to the correction capacity of the code. In the MLCM coding, the performance of the MSD decoder is highly dependent on the SER performance ; symbol errors cause error propagation in the MSD. As a result, the pre-FEC threshold becomes stricter with increasing phase noise.

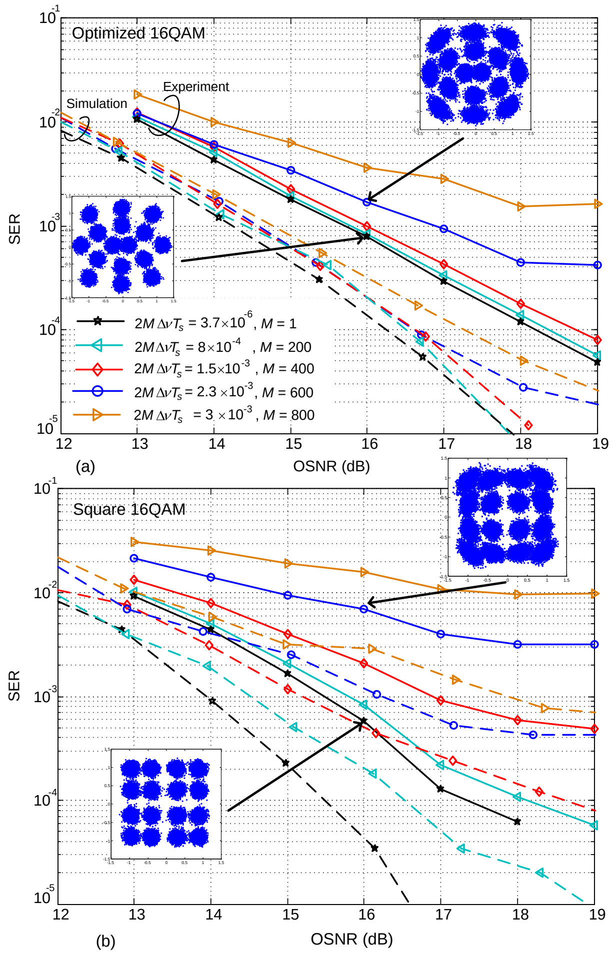


FIGURE 3.2 – SER of (a) optimized and (b) square 16-QAM constellation. Dashed curves are simulation results.

TABLE 3.1 – OSNR_{req} for MLCM and uniform rate coding

$2M\Delta\nu T_s$	8×10^{-4}	1.5×10^{-3}	2.3×10^{-3}	3×10^{-3}
M	200	400	600	800
MLCM	< 15	< 16	< 18	> 21
Uniform rate	< 16	< 17	> 21	> 21

3.4.2 MLCM advantage over uniform rate coding

In this section, performance of MLCM coding is compared with the uniform rate one assuming a total code rate of 0.9373. In Fig. 3.5, post-FEC BER curves of optimized 16-QAM are shown for MLCM and uniform rate coding. Despite the MLCM coding being designed for OSNR = 15.5 dB and $\Delta\nu T_s = 10^{-3}$, it performs better than uniform rate coding for a wide range of phase noise. For $\Delta\nu T_s = 2.3\times10^{-3}$ the SER in Fig. 3.2 reaches a floor, and the uniform rate coding fails to achieve error free transmission. The optimized constellation approaches a lower

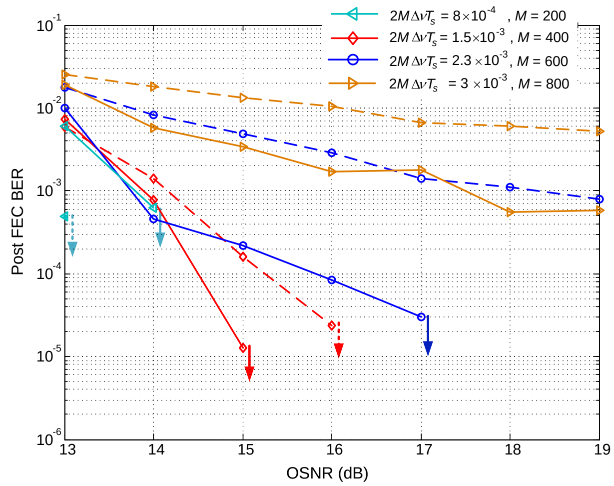


Figure 3.3 – Post-FEC BER of optimized (solid) and square (dashed) 16-QAM constellation for MLCM coding.

BER floor more slowly, and we can see in Fig. 3.5 that zero errors are observed by OSNR of 17 dB. In table 3.1, OSNR_{req} in dB of the MLCM and uniform rate coding for optimized constellation are compared for different phase noise values.

3.5 Conclusions

We presented experimental results for symbol error rate for MLCM coding for both phase noise optimized and square 16-QAM constellations. Post-FEC BER for MLCM coding for both constellations was calculated from experimental data. We show that the MLCM coding designed for the optimum performance of BLER (by optimizing code rates) performs well in terms of experimental post-FEC BER. Comparison of MLCM coding performance for optimized and square 16-QAM shows that the advantage in using optimized constellation can be greatly enhanced by employing MLCM coding. MLCM performance was also compared with that of uniform rate coding. We observed that MLCM coding is robust to various levels of phase noise, despite being designed for a fixed phase noise and OSNR.

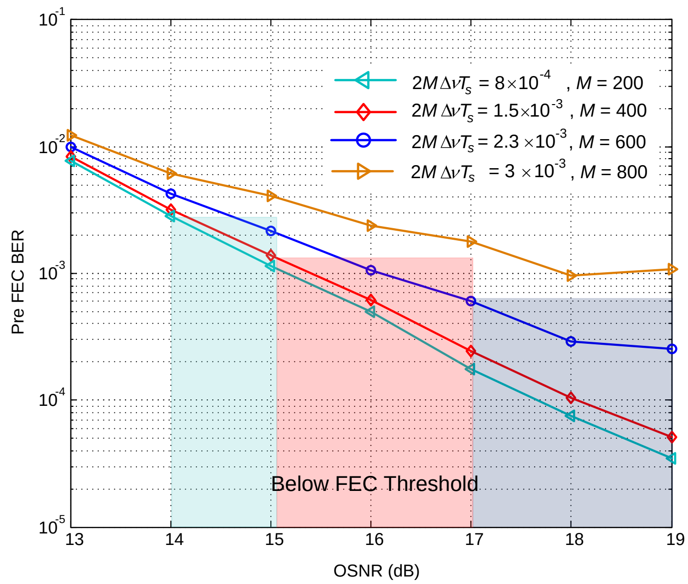


FIGURE 3.4 – Pre-FEC BER of the optimzied constellation with MLCM coding.

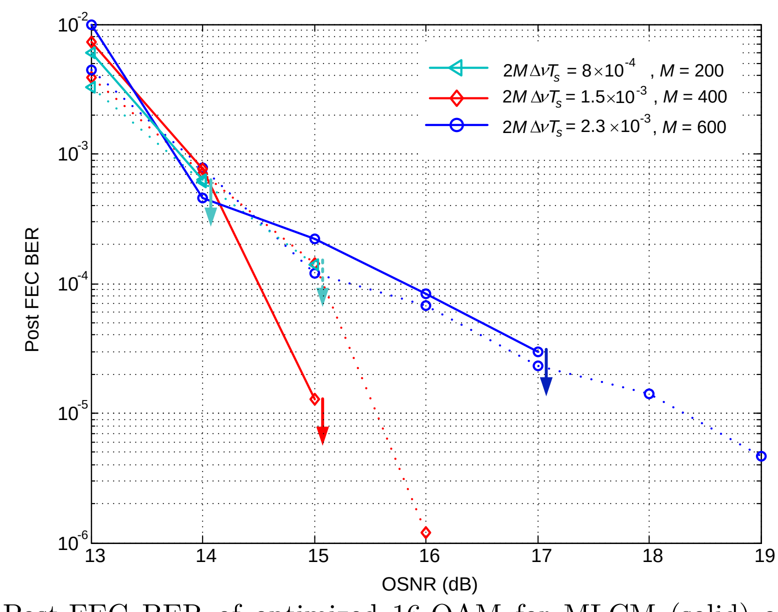


FIGURE 3.5 – Post-FEC BER of optimized 16-QAM for MLCM (solid) and uniform rate (dotted) coding.

Chapitre 4

Conclusions and Future Work

We explored two aspects of phase noise in optical coherent systems. First aspect is phase noise statistics in presence of CD where we modified the analysis in the literature. Second aspect is the application of MLCM to phase noise limited coherent systems where we proposed a novel method for designing of MLCM coding optimized for phase noise limited transmission.

Our first contribution was a more accurate formulation of phase noise and CD interaction. We showed that a previously published analysis must be modified due to an overly simple model for carrier phase recovery. We proposed a more accurate model for carrier phase recovery which corresponds to DD carrier phase recovery. We exploited this modified model to derive modified PDF of received symbols before decision. We also calculated BER of a DQPSK system semi-analytically using our modified PDF and showed that the result is well matched with the simulation for a wide range of system parameters. Our semi-analytical BER prediction is useful in finding maximum reach or maximum tolerable LO linewidth without resorting to simulations. Our semi-analytical BER prediction is also a lower bound of the BER from VV carrier phase recovery which is a more practical phase recovery method. In addition, our analytical PDF of received symbols can be used in designing soft decision FECs where accurate likelihood probabilities are required.

Our second contribution deals with the application of MLCM to phase noise limited coherent systems. We proposed a numerical method for designing MLCM coding ; the method can be applied to arbitrary constellations if PDF of received symbols before decision is known. We applied our MLCM design method to two phase noise limited coherent systems with known PDF ; 1) a system with nonlinear phase noise where we optimized BLER by designing optimal set-partitioning and code rates and showed that our design excels an already published MLCM design in terms of BLER 2) a system with AWGN and phase noise where we examined post FEC BER in addition to BLER and explored performance of a phase noise optimized and square 16-QAM constellation. We showed that superior SER performance of phase noise optimized constellation, over square 16-QAM, can be translated into post FEC BER and found the phase noise regime for the superior performance. Our post FEC BER results, obtained from MC simulation, showed that the BLER, used for optimization of our MLCM coding, is a reliable performance indicator and its minimizations translates into post FEC BER minimization. In addition, we performed experiments to compare MLCM and uniform rate coding for both phase noise optimized and square 16-QAM constellations. Experimental post FEC BER showed the same trend as simulated post FEC BER confirming validity of our method in practice. Although we designed MLCM coding for a specific combination of SNR and phase noise, our MLCM coding showed superior performance over uniform rate coding in experiments for a wide range of phase noise. This indicates robustness of our MLCM coding design which is important in practice.

In summary, we explored two aspects of phase noise in coherent systems which are unique to optical coherent systems and are not well studied in the literature of wireless communication. For the first contribution, electronic compensation of large CD has just recently attracted attention in optical communication community and problem of LO phase noise and CD interaction is a new problem. Our analysis of LO phase noise and CD interaction provides a more accurate formulation of this interaction. For the second contribution, application of MLCM to phase noise limited optical systems was not well studied before and we provided a systematic method to exploit MLCM coding in optical coherent systems. In addition, we explored for the first time accuracy of BLER optimization in MLCM design, rather than post FEC BER, through simulation and experiment.

Many aspects of our work can be extended or improved. For analysis of phase noise and CD interaction, we studied QPSK modulation in detail and 16-QAM briefly. However, extension of our approach to other modulation formats can be subject of future research. In addition, our analysis of phase noise and CD interaction ignores fiber nonlinearity, so our approach is applicable to short and medium reach transmission and may not be accurate in long-haul transmission. Developing a formulation by taking into account nonlinearity is another prospect of research in future. For application of our MLCM design, we demonstrated application of our method to 16-ary constellations. As we mentioned in chapter 3, our method becomes numerically complex when it is applied to higher order constellations. Improving this aspect of our method is an open subject for future research. In addition, we studied application of our method for systems where CD is fully compensated and there is no nonlinearity. Exploring application of our MLCM design to system with residual dispersion and nonlinearity is another interesting subject for future research.

Publication List

This is a list of my publications during my Ph.D. I have presented most of them in the previous chapters.

1. R. Farhoudi, A. Ghazisaeidi, and L. A. Rusch, “Analytical PDF of decision statistic for coherent MPSK with electronic dispersion equalization,” in Proc. of Optical fiber Conference (OFC) 2012, Los Angeles, CA..
2. R. Farhoudi, A. Ghazisaeidi, and L. A. Rusch, “Performance of carrier phase recovery for electronically dispersion compensated coherent systems," *Optics Express*, vol. 20, no. 24, pp. 26568-26582, Nov 2012.
3. R. Farhoudi and L. A. Rusch, “ Multi-level coded modulation for phase noise optimized constellations,” in Proc. of Optical fiber Conference (OFC) 2013, Anaheim, CA.
4. R. Farhoudi, and L. A. Rusch, “Multi-level coded modulation for 16-ary constellations in presence of phase noise," *Journal of Lightwave Technology*, vol. 32, no. 6, pp. 1159-1167, Nov 2013.
5. R. Farhoudi, A. T. Nguyen, and L. A. Rusch, “Experimental verification of multi-level coded modulation for 16-ary constellations," vol. 26, no. 17, pp. 1774-1777, *Photonics Technology Letters*, Jul 2014.

Bibliographie

[1] P. Winzer. Challenges and evolution of optical transport networks. In *Optical Communication (ECOC), 2010 36th European Conference and Exhibition on*, pages 1–3. IEEE.

[2] Cisco VNI Forecast. Cisco visual networking index : Global mobile data traffic forecast update 2012-2017. *Cisco Public Information, February*, 2013.

[3] E. Yamazaki, M. Tomizawa, and Y. Miyamoto. 100-Gb/s optical transport network and beyond employing digital signal processing. *Communications Magazine, IEEE*, 50(2) :s43–s49, 2012.

[4] E. Ip, A. Lau, D. Barros, and J. Kahn. Coherent detection in optical fiber systems. *Opt. Express*, 16(2) :753–791, 2008.

[5] G. Charlet. Coherent detection associated with digital signal processing for fiber optics communication. *Comptes Rendus Physique*, 9(9-10) :1012–1030, 2008.

[6] E. Ip and J. Kahn. Digital equalization of chromatic dispersion and polarization mode dispersion. *J. Lightwave Technol.*, 25(8) :2033–2043, 2007.

[7] M. Taylor. Phase estimation methods for optical coherent detection using digital signal processing. *J. Lightwave Technol.*, 27(7) :901–914, 2009.

[8] G. Ungerboeck. Channel coding with multilevel/phase signals. *IEEE Trans. Inf. Theory*, 28(1) :55–67, Jan. 1982.

[9] A. Lau, T. Shen, W. Shieh, and K. Ho. Equalization-enhanced phase noise for 100Gb/s transmission and beyond with coherent detection. *Opt. Express*, 18(16) :17239–17251, 2010.

[10] T. Pfau, X. Liu, and S. Chandrasekhar. Optimization of 16-ary quadrature amplitude modulation constellations for phase noise impaired channels. presented at the European Conference of Optical Commun. (ECOC), Conf., Geneva, Sep., 2011.

[11] L. Beygi, E. Agrell, P. Johannisson, and M. Karlsson. A novel multilevel coded modulation scheme for fiber optical channel with nonlinear phase noise. presented at the Global Telecommun. (GLOBECOM) Conf., Miami, FL, Dec., 2010.

[12] R. Saunders, M. Traverso, T. Schmidt, and C. Malouin. Economics of 100Gb/s Transport. In *National Fiber Optic Engineers Conference*. Optical Society of America, 2010.

[13] G. P. Agrawal. *Nonlinear Fiber Optics.* Academic, San Diego, 2001.

[14] A. Bononi, P. Serena, and N. Rossi. Nonlinear signal-noise interactions in dispersion-managed links with various modulation formats. *Optical Fiber Technology*, 16(2) :73–85, 2010.

[15] K. Ho. *Phase-modulated optical communication systems.* Springer Verlag, 2005.

[16] E. Ip and J. Kahn. Fiber impairment compensation using coherent detection and digital signal processing. *J. Lightwave Technol.*, 28(4) :502–519, 2010.

[17] S. Savory. Digital filters for coherent optical receivers. *Opt. Express*, 16(2) :804–817, 2008.

[18] K. Ho and J. Kahn. Electronic compensation technique to mitigate nonlinear phase noise. *J. Lightwave Technol.*, 22(3) :779–783, 2004.

[19] E. Ip and J. Kahn. Compensation of dispersion and nonlinear impairments using digital backpropagation. *J. Lightwave Technol.*, 26(20) :3416–3425, 2009.

[20] G. Goldfarb and G. Li. Efficient backward-propagation using waveletbased filtering for fiber backward-propagation. *Opt. Express*, 17(11) :8815–8821, 2009.

[21] G. Goldfarb and G. Li. Chromatic dispersion compensation using digital IIR filtering with coherent detection. *Photon. Technol. Lett., IEEE*, 19(13) :969–971, 2007.

[22] J. Proakis and M. Salehi. *Digital communications.* McGraw-hill New York, 2001.

[23] F. Derr. Coherent optical QPSK intradyne system : Concept and digital receiver realization. *J. Lightwave Technol.*, 10(9) :1290–1296, 2002.

[24] E. Ip and J. Kahn. Feedforward carrier recovery for coherent optical communications. *J. Lightwave Technol.*, 25(9) :2675–2692, 2007.

[25] H. Meyr, M. Moeneclaey, and S. Fechtel. *Digital communication receivers : synchronization, channel estimation, and signal processing.* John Wiley & Sons, Inc. New York, NY, USA, 1997.

[26] S. Zhang, P. Kam, J. Chen, and C. Yu. Decision-aided maximum likelihood detection in coherent optical phase-shift-keying system. *Opt. Express*, 17(2) :703–715, 2009.

[27] S. Zhang, P. Kam, C. Yu, and J. Chen. Decision-aided carrier phase estimation for coherent optical communications. *J. Lightwave Technol.*, 28(11) :1597–1607, 2010.

[28] A. Viterbi. Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission. *Information Theory, IEEE Transactions on*, 29(4) :543–551, 1983.

[29] D. Ly-Gagnon, S. Tsukamoto, K. Katoh, and K. Kikuchi. Coherent detection of optical quadrature phase-shift keying signals with carrier phase estimation. *J. Lightwave Technol.*, 24(1) :12–21, 2006.

[30] E. Vanin and G. Jacobsen. Analytical estimation of laser phase noise induced BER floor in coherent receiver with digital signal processing. *Opt. Express*, 18(5) :4246–4259, 2010.

[31] E. Zehavi. 8-psk trellis codes for a rayleigh channel. *Communications, IEEE Transactions on*, 40(5) :873–884, 1992.

[32] H. Imai and S. Hirakawa. A new multilevel coding method using error-correcting codes. *IEEE Trans. Inf. Theory*, 23(3) :371–377, May 1977.

[33] G. Caire, G. Taricco, and E. Biglieri. Bit-interleaved coded modulation. *IEEE Trans. Inf. Theory*, 44(3) :927–946, May 1998.

[34] E. Zehavi and G. Kaplan. Phase noise effects on m-ary psk trellis codes. *Communications, IEEE Transactions on*, 39(3) :373–379, 1991.

[35] T. Minowa, H. Ochiai, and H. Imai. Phase-noise effects on turbo trellis-coded over m-ary coherent channels. *Communications, IEEE Transactions on*, 52(8) :1333–1343, 2004.

[36] M. Magarini, R. Essiambre, B. Basch, A. Ashikhmin, G. Kramer, and A. de Lind van Wijngaarden. Concatenated coded modulation for optical communications systems. *IEEE Photon. Technol.*, 22(16) :1244–1246, June 2010.

[37] L. Schmalen and R. Dischler. Experimental evaluation of coded modulation for a coherent pdm system with high spectral efficiency. In *Optical Fiber Communication Conference*. Optical Society of America, 2012.

[38] H. Bülow and E. Masalkina. Coded modulation in optical communications. presented at the Optical Fiber Commun. (OFC), Conf., Los Angeles, CA, Mar., 2011.

[39] S. Lotz, W. Sauer-Greff, and R. Urbansky. Iterative demapping and decoding for bit-interleaved coded modulation in optical communication systems. presented at the Transparent Optical Networks (ICTON), Conf., Munich, Germany, June, 2010.

[40] S. Benedetto, G. Olmo, and P. Poggiolini. Trellis coded polarization shift keying modulation for digital optical communications. *IEEE Trans. Commun.*, 43(234) :1591–1602, Feb. 1995.

[41] B. Smith and F. Kschischang. A pragmatic coded modulation scheme for high-spectral-efficiency fiber-optic communications. *J. Lightwave Technol.*, 30(13) :2047–2053, Jul. 2012.

[42] P. Faraj, S. Schöllmann, J. Leibrich, and W. Rosenkranz. 8.4 dB net coding gain achieved with a serially concatenated coding scheme for differential quadrature phase shift keyed optical systems. presented at the European Conference of Optical Commun. (ECOC), Conf., Sep., 2005.

[43] T. Sakamoto, A. Chiba, I. Morohashi, and T. Kawanishi. Optical trellis-coded modulation with multi-parallel MZM. presented at the European Conference of Optical Commun. (ECOC), Conf., Vienna, Sep., 2009.

[44] H. Bülow, G. Thielecke, and F. Buchali. Optical trellis-coded modulation (oTCM). presented at the Optical Fiber Commun. (OFC), Conf., Los Angeles, CA, Feb., 2004. Paper. WM5.

[45] M. Kumar, H. Yoon, and N. Park. Performance evaluation of trellis code modulated odqpsk using the klse method. *Photonics Technology Letters, IEEE*, 19(16) :1245–1247, 2007.

[46] H. Zhao, E. Agrell, and M. Karlsson. Trellis-coded modulation in PSK and DPSK communications. presented at the European Conference of Optical Commun. (ECOC), Conf., Cannes, France, Sep., 2006.

[47] R. Saunders, M. Traverso, T. Schmidt, and C. Malouin. Economics of 100 Gb/s transport. In *Proc. of OFC*, Paper. NMB.2, 2010.

[48] T. Hoshida, H. Nakashima, T. Tanimura, S. Oda, Z. Tao, L. Liu, W. Yan, L. Li, and J. Rasmussen. Network innovations brought by digital coherent receivers. In *Proc. of OFC*, Paper. NMB.4, 2010.

[49] E. Ip and J. Kahn. Compensation of dispersion and nonlinear impairments using digital backpropagation. *J. Lightwave. Technol.*, 26(20) :3416–3425, 2008.

[50] V. Curri, P. Poggiolini, A. Carena, and F. Forghieri. Dispersion compensation and mitigation of nonlinear effects in 111-Gb/s WDM coherent PM-QPSK systems. *IEEE Photon. Technol. Lett.*, 20(17) :1473–1475, 2008.

[51] X. Chen, C. Kim, G. Li, and B. Zhou. Numerical study of lumped dispersion compensation for 40-Gb/s return-to-zero differential phase-shift keying transmission. *IEEE Photon. Technol. Lett.*, 19(8) :568–570, 2007.

[52] C. Xie. WDM coherent PDM-QPSK systems with and without inline optical dispersion compensation. *Opt. Express*, 17(6) :4815–4823, 2009.

[53] D. Crivelli, H. Carter, and M. Hueda. Adaptive digital equalization in the presence of chromatic dispersion, PMD, and phase noise in coherent fiber optic systems. In *Proc. of Globecom*, pages 2545–2551, 2004.

[54] C. Xie. Local oscillator phase noise induced penalties in optical coherent detection systems using electronic chromatic dispersion compensation. In *Proc. of OFC*, Paper. OMT.4, 2009.

[55] W. Shieh and K. Ho. Equalization-enhanced phase noise for coherent-detection systems using electronic digital signal processing. *Opt. Express*, 16(20) :15718–15727, 2008.

[56] I. Fatadin and S. Savory. Impact of phase to amplitude noise conversion in coherent optical systems with digital dispersion compensation. *Opt. Express*, 18(15) :16273–16278, 2010.

[57] K. Ho, A. Lau, and W. Shieh. Equalization-enhanced phase noise induced timing jitter. *Optics Letters*, 36(4) :585–587, 2011.

[58] S. Oda, C. Ohshima, T. Tanaka, T. Tanimura, H. Nakashima, N. Koizumi, T. Hoshida, H. Zhang, Z. Tao, and J.C. Rasmussen. Interplay between local oscillator phase noise and electrical chromatic dispersion compensation in digital coherent transmission system. In *Proc. of ECOC*, Paper. Mo.1.C.2, 2010.

[59] T. Xu, G. Jacobsen, S. Popov, J. Li, A. Friberg, and Y. Zhang. Analytical estimation of phase noise influence in coherent transmission system with digital dispersion equalization. *Opt. Express*, 19(8) :7756–7768, 2011.

[60] Q. Zhuge, C. Chen, and D. Plant. Dispersion-enhanced phase noise effects on reduced-guard-interval CO-OFDM transmission. *Opt. Express*, 19(5) :4472–4484, 2011.

[61] Tianhua Xu, Gunnar Jacobsen, Sergei Popov, Jie Li, Sergey Sergeyev, Ari T Friberg, and Yimo Zhang. Analytical BER performance in differential n-PSK coherent transmission system influenced by equalization enhanced phase noise. *Optics Communications*, 334 :222–227, 2015.

[62] Gunnar Jacobsen, Tianhua Xu, Sergei Popov, and Sergey Sergeyev. Phase noise influence in coherent optical DnPSK systems with DSP based dispersion compensation. *Journal of Optical Communications*, 35(1) :57–61, 2014.

[63] G. Colavolpe, T. Foggi, E. Forestieri, and M. Secondini. Impact of phase noise and compensation techniques in coherent optical systems. *J. Lightwave Technol.*, 29(18) :2790–2800, 2011.

[64] M. Secondini, G. Meloni, T. Foggi, G. Colavolpe, L. Poti, and E. Forestieri. Phase noise cancellation in coherent optical receivers by digital coherence enhancement. In *Proc. of ECOC*, Paper. P4.17, 2010.

[65] G. Colavolpe, T. Foggi, E. Forestieri, and M. Secondini. Phase noise sensitivity and compensation techniques in long-haul coherent optical links. In *Proc. of Globcom*, pages 1–6, 2010.

[66] Q. Zhuge, X. Xu, Z.A. El-Sahn, M. Mousa-Pasandi, M. Morsy-Osman, M. Chagnon, M. Qiu, and D.V. Plant. Experimental investigation of the equalization-enhanced phase noise in long haul 56 Gbaud DP-QPSK systems. *Opt. Express*, 20(13) :13841–13846, 2012.

[67] M. Magarini, A. Spalvieri, F. Vacondio, M. Bertolini, M. Pepe, and G. Gavioli. Empirical modeling and simulation of phase noise in long-haul coherent optical transmission systems. *Opt. Express*, 19(23) :22455–22461, 2011.

[68] U. Wachsmann, R. Fischer, and J. Huber. Multilevel codes : Theoretical concepts and practical design rules. *IEEE Trans. Inf. Theory*, 45(5) :1361–1391, Jul. 1999.

[69] I. Djordjevic, M. Arabaci, and L. Minkov. Next generation FEC for high-capacity communication in optical transport networks. *J. Lightwave Technol.*, 27(16) :3518–3530, Aug. 2009.

[70] H. Batshon, I. Djordjevic, and T. Schmidt. Ultra high speed optical transmission using subcarrier-multiplexed four-dimensional ldpc-coded modulation. *Optics Express*, 18(19) :20546–20551, Sep. 2010.

[71] Lotfollah Beygi, Erik Agrell, Joseph M Kahn, and Magnus Karlsson. Coded modulation for fiber-optic networks. *IEEE signal processing magazine*, 2014.

[72] Alex Alvarado and Erik Agrell. Four-dimensional coded modulation with bit-wise decoders for future optical communications. *arXiv*, 2014.

[73] I. Djordjevic and T. Wang. On the LDPC-Coded Modulation for Ultra-High-Speed Optical Transport in the Presence of Phase Noise. presented at the Optical Fiber Commun. (OFC), Conf., Anaheim, CA, Mar., 2013.

[74] N. Miladinovic and M. Fossorier. Generalized LDPC codes with Reed-Solomon and BCH codes as component codes for binary channels. presented at the Global Telecommun. (GLOBECOM) Conf., St. Louis, Mo, Nov., 2005.

[75] I. Djordjevic, L. Xu, and T. Wang. Multidimensional turbo product and generalized LDPC codes with component RS codes suitable for use in beyond 100 Gb/s optical

transmission. presented at the Telecommun. in Modern Satellite, Cable, and Broadcasting Services (TELSIKS) Conf., Nis, Serbia, Oct., 2009.

[76] P. Winzer. Beyond 100G ethernet. *IEEE Commun. Magazine*, 48(7) :26–30, Jul. 2010.

[77] R. Farhoudi and L. Rusch. Multi-level coded modulation for phase noise optimized constellations. presented at the Optical Fiber Commun. (OFC), Conf., Anaheim, CA, Mar., 2013.

[78] T. Woerz and J. Hagenauer. Multistage coding and decoding for a M-PSK system. presented at the Global Telecommun. (GLOBECOM) Conf., San Diego, CA, Dec., 1990.

[79] L. Beygi, E. Agrell, and M. Karlsson. On the dimensionality of multilevel coded modulation in the high snr regime. *IEEE Commun. Lett.*, 14(11) :1056–1058, Nov. 2010.

[80] A. Lau and J. Kahn. Signal design and detection in presence of nonlinear phase noise. *J. Lightwave Technol.*, 25(10) :3008–3016, Oct. 2007.

[81] P. Kam, K. Teo, Seow, Y. Some, and T. Tjhung. Approximate results for the bit error probability of binary phase shift keying with noisy phase reference. *IEEE Trans. Commun.*, 41(7) :1020–1022, Jul. 1993.

[82] I. Djordjevic and B. Vasic. Multilevel coding in M-ary DPSK/differential QAM high-speed optical transmission with direct detection. *J. Lightwave Technol.*, 24(1) :420–428, Jan. 2006.

[83] R. Essiambre, G. Foschini, P. Winzer, G. Kramer, and E. Burrows. The capacity of fiber-optic communication systems. presented at the Optical Fiber Commun. (OFC), Conf., San Diego, CA, Feb., 2008.

[84] R. Farhoudi and L. Rusch. Multi-level coded modulation for 16-ary constellations in presence of phase noise. *Journal of Lightwave Technol.*, 32(6) :1159–1167, 2014.

[85] J. Fu, M. Arabaci, I. Djordjevic, Y. Zhang, L. Xu, and T. Wang. First experimental demonstration of nonbinary ldpc-coded modulation suitable for high-speed optical communications. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2011 and the National Fiber Optic Engineers Conference*, pages 1–3. IEEE, 2011.

[86] H. Zhang, H. Batshon, D. Foursa, M. Mazurczyk, J. Cai, C. Davidson, A. Pilipetskii, G. Mohs, and N. Bergano. 30.58 tb/s transmission over 7,230 km using pdm half 4d-16qam coded modulation with 6.1 b/s/hz spectral efficiency. In *Optical Fiber Communication Conference*. Optical Society of America, 2013.

[87] J. Renaudier, O. Bertran-Pardo, A. Ghazisaeidi, P. Tran, H. Mardoyan, P. Brindel, A. Voicila, G. Charlet, and S. Bigo. Experimental transmission of nyquist pulse shaped 4-d coded modulation using dual polarization 16qam set-partitioning schemes at 28 gbaud. In *Optical Fiber Communication Conference*. Optical Society of America, 2013.

[88] M. Arabaci, I. Djordjevic, L. Xu, and T. Wang. Four-dimensional nonbinary ldpc-coded modulation schemes for ultra-high-speed optical fiber communication. *Photonics Technology Letters, IEEE*, 23(18) :1280–1282, 2011.