COHERENT PHASE-MODULATED OPTICAL FIBER COMMUNICATIONS WITH LINEAR AND NONLINEAR PHASE NOISE

ZHANG SHAOLIANG

NATIONAL UNIVERSITY OF SINGAPORE

2011

COHERENT PHASE-MODULATED OPTICAL FIBER COMMUNICATIONS WITH LINEAR AND NONLINEAR PHASE NOISE

ZHANG SHAOLIANG

(*B. Eng.(Hons.), Beijing University of Posts and Telecommunications, China*)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2011

Acknowledgement

Contents

CONTENTS

Summary

The exponentially-increasing growth of high-speed and high-capacity Internet traffics sees that the spectral efficiency (SE) becomes more and more important in the development of the backbone optical networks. To efficiently utilize the limited spectrum of the optical fibers, coherent detection has revived to support advanced modulation formats in optical transmission systems. Besides, the full information of the received electric field can be preserved in coherent receivers, thus enabling digital signal processing (DSP) algorithms to compensate for the fiber transmission impairments.

This thesis studies the DSP techniques in such following areas: phase estimation (PE) algorithms for laser phase noise and fiber nonlinear phase noise; frequency offset estimator (FOE) to tackle the frequency offset between transmitter and local oscillator (LO) lasers; and electrical compensation for the fiber nonlinearity.

Among these impairments, laser phase noise plays a significant role in affecting the performance of coherent receivers. For example, a good PE is capable of allowing for a laser with large linewidth, thus reducing the system cost. Although quite a few DSP-based PE algorithms have been proposed in the literatures, they require either nonlinear computations (Mth-power operation and phase unwrapping) or the statistics of the system noises (phase noise and additive noise). Nonlinear operations are likely to increase the power consumption of coherent receivers while the statistics of such information may be not known to the receiver especially in reconfigurable optical switching systems. In view of the disadvantages, a computationally-linear decision aided (DA)

maximum likelihood (ML) PE was introduced to eliminate the nonlinear computations while keeping or even improving the laser linewidth tolerance. We have conducted in-depth analysis on the performance of DA ML in different modulation formats, and observed that optimal memory length is related to the variances of the phase noise and additive noise. The parallel and serial implementations of the DA ML PE were also investigated to adapt itself to the high-speed optical receivers. Moreover, a coherent polarization-division-multiplexing (PDM) quadrature phase-shift-keying (QPSK) experiment was carried out to successfully demonstrate the DA ML PE which shows to achieve the same performance as the conventional V&V Mth-power method yet requires less computational loads.

However, the DA ML is subjected to the block length effect (BLE) because of a trade-off to average out the additive noise and phase noise. In order to address the BLE, a first-order filter was introduced to the DA ML algorithm, thus adaptively adjusting the filter gain based on the characteristics of the received signals. A Monte Carlo (M-C) simulation indicates that the adaptive DA algorithm has a powerful self-adaptation capability to acquire the optimal filter gain, resulting in optimal performance in all the signal-to-noise ratio (SNR) regions for constant-amplitude PSK formats. The adaptive DA algorithm was extended into the M-quadrature amplitude modulation (QAM) formats, where it was found that it suffers from the constellation penalty. Analysis was presented elaborately to show that the DA ML with the optimal memory length has a better performance than the adaptive DA at low and moderate SNRs. A long-haul coherent PDM-QPSK experiment was demonstrated that the adaptive DA algorithm can outperform the DA ML PE in the presence of nonlinear phase noise.

Finally, two novel DSP algorithms were proposed to address the phase noises originating from the frequency offset and fiber nonlinearity, respectively. A Gardner-timing-based FOE was experimentally demonstrated to achieve a wide estimation range from -5 GHz to 6 GHz. As for a joint self-phase-modulation compensation (SPMC)

scheme for fiber nonlinearity compensation, it is found that the joint-SPMC has a wider dynamic input power range compared to other conventional methods.

List of Figures

List of Tables

List of Abbreviations

ADC	Analog-to-digital converters
ASE	Amplified Spontaneous Emission
ASK	Amplitude Shift-Keying
AWG	Arbitrary Waveform Generator
AWGN	Additive White Gaussian Noise
B2B	Back-to-Back
BPSK	Binary Phase-Shift-Keying
BLE	Block Length Effect
CCI	Complex-Controlled Inverter
CD	Chromatic Dispersion
CDC	Chromatic Dispersion Compensation
CMA	Constant Modulus Algorithm
CMOS	Complementary Metal-Oxide-Semiconductor
DA	Decision Aided
DAC	Digital-to-Analog Converters
DCF	Dispersion-Compensating Fiber
DD	Decision-Directed
DE	Differential Encoding
DOF	Degrees of Freedom

Abbreviations

DPSK	Differential Phase-Shift-Keying
DQPSK	Differential Quadrature Phase-Shift-Keying
DSP	Digital Signal Processing
ECL	External Cavity Lasers
EDFA	Erbium-Doped Fiber Amplifier
FEC	Forward Error Control
FFT	Fast-Fourier Transform
FIR	Finite Impulse Response
FOE	Frequency Offset Estimator
GVD	Group Velocity Dispersion
IF	Intermediate Frequency
IIR	Infinite Impulse Response
IM/DD	Intensity-Modulation with Direct Detection
ISI	Intersymbol Interference
LDPC	Low-Density Parity Check
LMS	Least-Mean Square
LO	Local Oscillator
LSB	Least Significant Bit
MC	Monte Carlo
ML	Maximum Likelihood
MPE	Maximal Phase Error
MSB	Most Significant Bit
MZM	Mach-Zehnder Modulator
MZDI	Mach-Zehnder Delay Interferometers
NDE	Non-Differential-Encoded
NLMS	Normalized Least-Mean Square

Abbreviations

NRZ	Non-Return-to-Zero
NLSE	Nonlinear Schrödinger Equation
ODI	Optical Delay Interferometer
OFDM	Orthogonal Frequency Division Multiplexing
OOK	On-Off Keying
OSNR	Optical Signal-to-Noise Ratio
OTF	Optical Tunable Filter
PDM	Polarization-Division-Multiplexing
PE	Phase Estimation
PLL	Phase-Locked Loop
PMD	Polarization-Mode Dispersion
PolMux	Polarization-Multiplexing
PPG	Pulse Pattern Generators
PSD	Power Spectrum Density
PSK	Phase-Shift-Keying
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase-Shift-Keying
RP	Reference Phasor
RZ	Return-to-Zero
SE	Spectral Efficiency
SMF	Single-Mode Fiber
SNR	Signal-to-Noise Ratio
SPM	Self-Phase Modulation
SPMC	Self-Phase-Modulation Compensation
WDM	Wavelength-Division-Multiplexing
XPM	Cross-Phase Modulation

Chapter 1

Introduction

Optical communication refers to use optical fiber as a medium to transmit optical signal from the source to the destination. Since the 1990s the optical fiber has been widely installed for metropolitan and trans-ocean communications from 45 Mbit/s to now 100 Gbit/s. Depending on the detection methods, optical communication systems can be divided into two broad categories: direct detection and coherent detection. For direct detection system, the data is mapped into the intensity of the optical signals to simplify the system design and reduce the cost. Despite that the phase information of the optical signal can be modulated, the receiver complexity and cost will increase accordingly. To meet the bandwidth demand of the exponentially increasing internet traffic in the present and future, the capability of direct detection systems is extremely limited for upgrading the data rate from $40{\times}109$ Gbit/s up to $100{\times}109$ Gbit/s or even higher.

Recent advances in high-speed analog-to-digital converters (ADCs) have revived the coherent detection systems. When the received signals are sampled into digital waveforms using ADCs, its full information, such as phase, amplitude and polarization, can be preserved in the receiver such that higher-order modulation formats can be used to increase the transmission capacity. Besides, the bulky optical components are replaced by small and compact digital signal processing (DSP) processors to com-

pensate for the fiber transmission impairments. This DSP functions make the digital coherent detection universal and independent of the modulation formats.

In the following section, we will review the development of optical fiber communication, which actually sees the growth of "Global Village", where information moves faster and faster from one place to another!

1.1 Rebirth of Coherent Optical Communication

Coherent optical communication attracted considerable attentions in the 1980s, since it can approach the theoretical limit of the receiver sensitivity, thus extending the transmission reach of optical communication systems [1, 2]. In those days, experiments phase-locked the phase of the local oscillator (LO) laser to that of the incoming optical signals using an optical phase-locked loop (PLL). The scheme is called homodyne detection, which demodulates the optical signal directly to the baseband, because the LO laser has the same frequency as the signal optical carrier frequency [1]. However, it is complicated and unstable to use optical PLLs operating at optical frequency in practice [3, 4]. As a result, heterodyne detection was introduced to simplify the receiver design and relax the feedback delay of optical PLLs. The optical signal is first downconverted into an intermediate frequency (IF), and then an electrical PLL is used to track the phase of the IF signals at microwave frequency [5, 6]. It can be found that most works focused on the simple modulation formats, such as binary phase-shift-keying (BPSK), differential phase-shift-keying (DPSK) and amplitude-shift keying, due to the wide laser linewidth and stringent requirement on the loop delay of PLLs [7]. As claimed in [4], delays greater than a few tens of nanoseconds will result in loop instability in a 10 Gb/s transmission.

To develop a stable yet low-cost optical system, the interest in coherent detection cooled down and was shifted back to intensity modulation/direct detection (IM/DD)

scheme by resorting to erbium-doped fiber amplifiers (EDFAs) and wavelength-division-multiplexing (WDM) technology in the 1990's. The receiver sensitivity can be improved by using an optical amplifier as a preamplifier while the transmission distance can be up to thousands of kilometers by using cascaded EDFAs [8]. The capacity of lightwave communication systems had been increased to the order of Terabits per second with the aid of WDM technologies in S-band (1460 to 1530 nm), C-band (1530 to 1565 nm) and L-band (1565 to 1625 nm) ranges [9, 10].

As Internet networks continued to boom, their backbone optical systems saw the demanding requirement of spectral efficiency (SE) in the limited optical spectrum [11]. This contributed to the extensive researches on improving SE by using high-order modulation formats. Around the year 2000, differentially encoded phase-shift-keying (PSK) format was firstly demonstrated in experiments in conjunction with self-homodyne detection [8]. The self-homodyne receiver consists of Mach-Zehnder delay interferometers (MZDIs), which convert the phase difference between the current optical signal and its one-symbol delayed version into optical intensity, and a pair of balanced photodiodes [12]. The scheme removes an LO laser that is always present in a typical coherent receiver and consequently relaxes the laser linewidth tolerance. Moreover, SE can be raised in principle from 1 bit/s/Hz for IM/DD and DPSK to $\log_2 M$ bits/s/Hz in M-DPSK modulation ($M > 2$) [13]. Nonetheless, the potentials for further SE improvement have become limited within achievable receiver sensitivity by using the higher-order M-DPSK formats ($M > 4$), and the respective self-homodyne receiver has become quite complicated to implement [14]. In addition, the MZDIs remove the phase information of the current received M-DPSK signals (Note that MZDIs only detects the phase change between two adjacent symbols). Note that those phase and amplitude information of optical signals can be further used to compensate for the transmission impairments, such as chromatic dispersion (CD) and polarization-mode dispersion (PMD) [15].

1.1 Rebirth of Coherent Optical Communication

Recent advances in high-speed ADCs [16–18] have prompted extensive researches on coherent optical communication once again. Nowadays, the sampling speed of commercial built-in ADCs can be as high as 50 GSamples/s [19], or even up to 80 GSamples/s for two channels [20]. Rather than improving the receiver sensitivity only, the primary motivations nowadays aim to further improve SE and system capacity by using advanced modulation formats and schemes, such as M-PSK and 16-quadrature amplitude modulation (16QAM) in conjunction with single-carrier or orthogonal frequency division multiplexing (OFDM) techniques, and to compensate for the channel distortions through DSP algorithms [21–26]. Generally, homodyne detection is preferred to downconvert the optical signal to or near the baseband instead of using the heterodyne scheme, which requires quite a large bandwidth for subsequent electrical components [27]. Compared to the early coherent receivers, there are two dramatic differences in current ones. First, high-speed ADCs employed in current phase-diversity coherent receivers sample the photocurrents, corresponding to the received optical signals, at the Nyquist rate or above so as to retain full information of the electric field, which is lost in the self-homodyne scheme. Some papers have referred to it as a digital coherent receiver [28]. Since the amplitude and phase information of the received optical signals are preserved, both of them can be modulated simultaneously to increase SE, and can be further utilized for compensation of linear and even nonlinear channel impairments [15]. Second, bulky optical components are replaced by small and compact DSP processors to compensate for the fiber transmission impairments. This DSP functions make the digital coherent receiver universal and independent of the modulation formats [28]. In the future, customers may only set some parameters through the interface of their DSP-enabled coherent receivers to retrieve data from any fiber port. We did a literature survey about the SE and capacity achieved so far, as depicted in Figure 1.1. It can be observed that most of those experiments achieving the record system capacity and SE were conducted using coherent detection with advanced modulation

formats as well as polarization division multiplexing (PDM) techniques. The largest capacity till the year 2010 is 69.1 Tb/s (SE is 9 bits/s/Hz) by using PDM-16QAM format, which is done by NTT Lab, Japan [29]. As can be seen from Figure 1.1, the high capacity is driven by coherent detection and multiplexing technologies, including time-, wavelength, polarization and even mode multiplexing. All these multiplexing approaches are independent, thus enabling any combination depending on the system requirements.
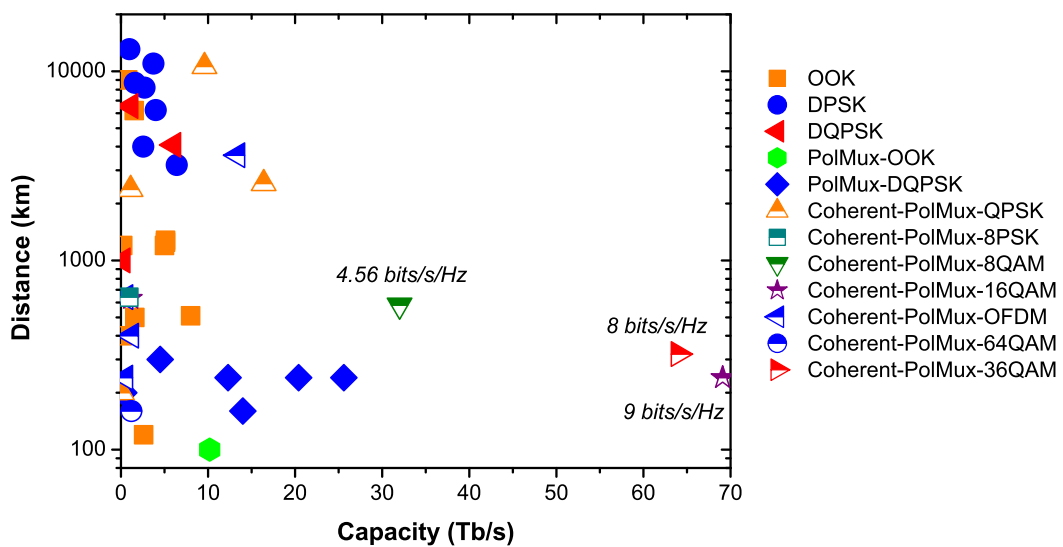
Figure 1.1: The review on capacity and SE reported in the experiments till the year 2010.

1.2 Literature Review

One of the challenges in coherent optical systems is to recover the carrier phase, which is perturbed, for example, by laser phase noise. An optical PLL is one solution to track the carrier phase with respect to the LO carrier in the early days of coherent optical communications. However, optical PLLs operating at optical wavelengths in combination with distributed feedback lasers are quite difficult to implement because of the large product of laser linewidth and loop delay [3]. It has been shown that delays greater than a few tens of nanoseconds would lead to loop instability even at a 10 Gbit/s

transmission [4].

Coherent detection can retain the information of the received electrical field. As a result, its linear impairments, such as CD, first-order PMD, can be fully compensated in principle. Instead of using bulky and complicated optical components, the DSP algorithms now play an important role in compensating for the fiber transmission impairments in digital coherent receivers [30]. With the aid of high-speed ADCs in a digital coherent receiver, the carrier PE can be done in high-speed DSP units rather than using optical PLLs for carrier phase tracking, allowing for a free-running LO laser. Recent experiments have demonstrated that DSP-based PE techniques are very effective to recover carrier phase. Based on a nonlinear transformation of received M-PSK signals, the commonly-used Viterbi & Viterbi (V&V) Mth-power scheme is capable of accurately tracking the unknown carrier phase [31]. Nevertheless, it relies heavily on nonlinear computations, such as rectangular-to-polar/polar-to-rectangular transformations. Despite that M-th power operation can be simply performed block by block [22], nonlinear functions like M-th power and phase unwrapping are still imperative such that risking increasing power consumption and memory requirements. This Mth-power scheme is further extended to non-constant-amplitude M-QAM formats by using a subgroup of symbols with phase modulation $\pi/4 + n\pi/2$ (n=0, 1, 2, 3) [32]. Since only a fraction of received signals are used to estimate the phase reference in a QAM system, its system performance is severely degraded, thus making the algorithm sensitive to laser phase noise [33]. A modified Mth-power scheme in [34,35] suggests utilizing all the symbols by sub-partitioning those symbols belonging to the middle ring of 16QAM constellation into two QPSK groups. However, all the Mth-power schemes require nonlinear operations, such as $\arctan(\cdot)$ function, subgrouping of symbols, and phase unwrapping. To address the nonlinear computations, we will first introduce a computationally-linear decision-aided (DA) maximum likelihood (ML) phase estimation (PE) into coherent optical communication systems to

eliminate the nonlinear operations while keeping or even improving the laser linewidth tolerance.

Akin to the Mth-power algorithm, DA ML is also subjected to block length effect (BLE) because of the trade-off between averaging over additive noise and phase noises [36]. It is suggested that the optimal memory length of DA ML can be calculated provided that the statistics of the additive noise and phase noise are available [37]. However, the same prerequisite can be applied to other PE algorithms, such as Wiener filter [4,38] and Kalman filter [39]. Note that such requirements may be impractical especially in a reconfigurable optical system. Therefore, a normalized least-mean square (NLMS) PE is provided to handle this problem with a fixed step size. The optimum step size needs to time-consumingly find out through either simulations or experiments, thus making it not suitable for online processing. In this thesis, we propose to send a sequence of training data to the receiver for acquiring the channel characteristics, and a recursive algorithm is introduced to adjust the filter gain in an adaptive version of the DA algorithm. The filter gain can enable the adaptive DA algorithm to operate at the optimal or suboptimal state even without the knowledge of the system noises.

Besides, the frequency offset between the transmitter and LO lasers, and the fiber nonlinearity can also lead to phase noise. This frequency offset can be as large as $\pm$ 5 GHz due to aging, heating and fabrication of lasers [40]. It is worth pointing out that PE algorithms generally require that frequency offset between the transmitter and LO lasers should be no more than 10 percent of symbol rate [41], which translates to be $\sim$1 GHz for a 10 GBaud coherent optical system. Hence, an additional DSP-based FOE is imperative to make sure PE algorithms work properly [42]. The conventional frequency offset estimators (FOE) usually build upon the nonlinear Mth-power operations [43–45] or decision feedback [46] to remove phase modulation. Nonlinear Mth-power operations would restrict such FOEs to a smaller estimation range compared to decision-feedback algorithms [46]. Meanwhile, decision-feedback FOEs suffers from

the feedback delays, thus limiting its practical application especially in high-speed optical systems. In view of those potential disadvantages, we will propose an ultra-wide feed-forward FOE without using nonlinear Mth-power operations.

On the other hand, fiber self-phase modulation (SPM) effect limits the performance of long-haul phase-modulated transmission systems through nonlinear phase noise [47,48]. A simple phase-rotation scheme depending on the received signal power has been proposed to mitigate fiber SPM at the receiver end [49]. However, the method is only effective in a dispersionless link or a specific dispersion map with small local dispersion. In an optical transmission system with strong dispersion, the interaction between fiber Kerr effect and CD causes this phase-rotation scheme to fail [47]. As a result, SPM pre- and post-compensation techniques have been individually proposed to reduce fiber nonlinearity effect [50–54]. The basic idea is to solve the inverse non-linear Schrödinger equation (NLSE) to either restore the waveform of received signals at the receiver side, or estimate the transmitted signals. It is observed that unknown system noises would lead to divergence in the solutions to the inverse NLSE [47]. We will propose a joint-SPM compensation (SPMC) at the transmitter and receiver side simultaneously, such that mitigating this divergence problem.

Of significance is that all DSP algorithms are able to co-exist with forward-looking all-optical signal regeneration approach, such as phase sensitive amplification, which would render one of the most cost-effective ways to improve system performance. In other words, with the aid of all-optical processing techniques, DSP algorithm can further enhance the system performance. Considering the fact that real-time coherent receivers are limited by the availability of high-speed ADCs, all-optical processing technologies become even more practical at beyond 100Gb/s systems.

1.3   Contribution of the Thesis

First of all, we successfully introduce the DA ML algorithm to recover the carrier phase in coherent optical M-PSK/QAM systems. Its phase error variance is analytically derived in both M-PSK/ QAM formats to explain the BLE of DA ML algorithm, thus enabling optimal operation of the DA ML algorithm. In addition, the serial and parallel implementations of DA ML algorithm are presented for the first time to operate it in real-time.

To deal with the BLE of DA ML and prerequisite on the statistics of system noises in other adaptive PEs (Wiener and Kalman filter), we introduce the adaptive DA algorithm and extend it into M-QAM formats. Its phase error variance is analytically obtained, and is subjected to the effect of constellation penalty. Moreover, we compare three adaptive PEs (NLMS, Kalman and the adaptive DA algorithm) in a comprehensive analysis, and find out that they share the same optimal performance but with difference in their approaches to determine stepsize. Experiments are also carried out to demonstrate DA ML and adaptive DA algorithms in a polarization-multiplexing (PolMux) quadrature phase-shift-keying (QPSK) system.

A novel FOE based on the Gardner timing recovery algorithm is proposed and experimentally demonstrated to track the frequency offset that keeps rotating the carrier phase. Its estimation range is shown from $-0.5R_s$ to $0.6R_s$ according to experimental results (R_s refers to the system symbol rate), which is the widest reported in the literature to the best of our knowledge.

Last not the least, we have proposed a joint pre- and post-compensation scheme to mitigate the fiber nonlinearity effect on the coherent receiver. The dynamic input power, leading to BER less than 10^{-3}, has been increased by 0.4 dB and 1.2 dB compared to individual pre- and post-compensation algorithm.

1.4   Thesis Outline

The remainder of this thesis is organized as follows.

In Chapter 2, we review the coherent optical communication system with the methods to generate different modulation formats in the optical domain by using Mach-Zehnder modulators (MZM). At the receiver side, a series of DSP algorithms are presented in detail to explain the principles how to compensate for the CD, de-multiplex polarization crosstalk, and recover carrier phase.

In Chapter 3, a DA ML PE is introduced to estimate the carrier phase perturbed by the laser phase noises in different modulation formats, where its performance is first investigated using Monte Carlo (MC) simulations. In addition, the implementation of the DA ML algorithm is discussed in the scenario of serial and parallel receiver structure.

In Chapter 4, to cope with BLE of DA ML, an adaptive version of the DA ML algorithm is introduced to recover the carrier phase in M-PSK formats, and we further develop the algorithm into M-QAM formats. The performance of the adaptive DA algorithm is experimentally demonstrated in a long-haul single-channel PolMux QPSK system.

In Chapter 5, a novel Gardner-timing-recovery-based FOE and a joint pre- and post-SPM compensation are proposed, respectively, to compensate for frequency offset and fiber nonlinear phase noise.

Finally, conclusion and future work are presented in Chapter 6.

Chapter 2

Fundamental Theory of Coherent Optical Systems

In this chapter, an overview of the coherent optical communication systems is presented in detail, including the transmitter design, transmission links, receiver structure as well as the DSP algorithms.

The modulation techniques in coherent optical communication systems can be further classified into single-carrier and multi-carrier or OFDM systems. They basically have the same optical modulators (such as I/Q MZM), transmission fibers and coherent receivers. The difference lies in their DSP methods to modulate the transmitted data and demodulate the received signals in the DSP modules, as depicted in Figure 2.1. As for the single-carrier system, DSP algorithms are generally not necessary at the transmitter side, whereas OFDM transmitter requires digital-to-analog converters (DAC) to perform OFDM modulation. To some extent, the maximum transmission speed is limited by the speed of DAC in OFDM systems [55]. In this thesis, we focus on single-carrier system in which optical transmitter is simpler than the one in OFDM systems.
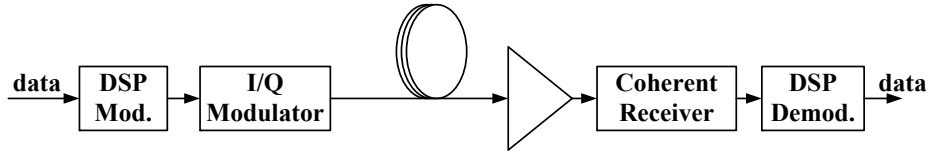
Figure 2.1: Coherent optical system using an I/Q modulator and DSP algorithms

2.1 Advanced Optical Modulation Formats

Advanced modulation formats in optical communication are attributed to the development of high-speed electro-optic modulators. Of importance is the development of MZM working by the principle of *interference*, controlled by modulating the optical *phase*. It is an essential component of I/Q modulators that are widely used in coherent optical communication systems nowadays [56].

2.1.1 Channel Capacity of Multi-Level Signals

The SE provided by the simple on-off keying (OOK) modulation cannot exceed 1 bit/s/Hz without using PDM technique [22]. Hence, multi-level signalling is a superior alternative to increase the SE of the optical transmission systems [28]. Besides, the slower symbol rate in multi-level modulation formats can reduce the impact of CD and PMD, compared to an OOK system with the equivalent bit rate.

The digital baseband representation of a phase- and/or amplitude-modulated signal can be expressed as [27]

$$X(k) = A_s(k) \exp\left(j\phi_s(k)\right), \tag{2.1}$$

where $A_s(k)$ and $\phi_s(k)$, respectively, denote the amplitude and the phase of the signal $X(k)$, $j = \sqrt{-1}$, and k stands for the symbol located at time period $[kT, (k+1)T)$. It is worth noting that various methods are available to generate multi-level signals [57–61]. In M-PSK modulation formats, the phase is divided into M even levels for representing M different symbols, each encoding $\log_2 M$ bits [62]. On the other hand,

2.1 Advanced Optical Modulation Formats

in M-QAM formats, two degrees of freedom (DOF), the amplitude and the phase, in Eq.(2.1) are used for denoting M different symbols. Although there are other variations to use M different frequencies or only M different levels of $A_s(k)$ [57,63], those methods usually do not provide better SE than M-PSK/QAM signals at low signal-to-noise ratio (SNR) [27]. To illustrate the superior SE using multi-level signalling and advantages of M-PSK/QAM formats, we compute the channel capacity according to the methods in [13].

Noise

N

X $Y = X+N$

Input Signal   Output Signal
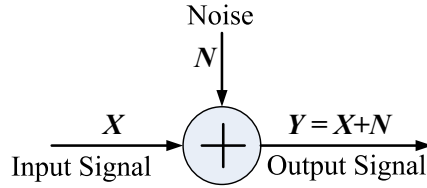
Figure 2.2: The schematic representation of the complex AWGN channel.

For a complex additive, white, Gaussian noise (AWGN) channel (see Figure 2.2) with discrete inputs, the channel capacity in bits per symbol is shown to be [64]

$$C = \max_{E[|X|^2] \leq P} I(X;Y), \tag{2.2}$$

subjected to the input power constraint $E[|X|^2] \leq P$. Here, $I(X;Y)$ is the mutual information of the channel input X and output Y, and $E[\cdot]$ denotes statistical expectation. Note that Eq.(2.2) is maximized only when all the symbols are equiprobable in the considered modulation formats [65]. The details about the calculation of the AWGN channel capacity are provided in Appendix A in the case of discrete constellations.

Figure 2.3 shows the capacity C in bits/symbol for several commonly-used modulation formats. The Shannon limit is given by $C = \log_2(1 + \gamma_s)$ [64], where γ_s refers to SNR per symbol. The SNR per bit (γ_b) is defined as

$$\gamma_b = \frac{\gamma_s}{\log_2 M}, \tag{2.3}$$

for a constellation that encodes $\log_2 M$ bits per symbol. It should be reminded that the optical SNR (OSNR) are usually used to evaluate the system performance. The
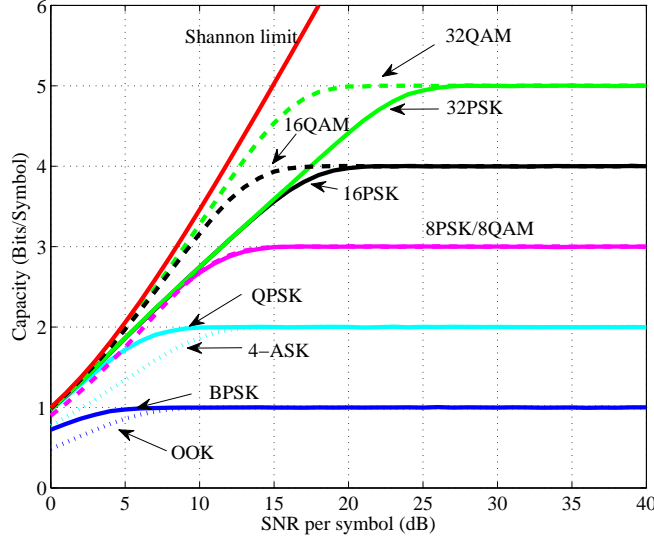
Figure 2.3: The capacity as a function of SNR per symbol for several modulation formats.

relationship between OSNR and SNR_b is shown to be [13]

$$\text{OSNR} = \frac{R_b}{2B_\text{ref}}\gamma_b, \tag{2.4}$$

where R_b is the information bit rate, and B_ref is the reference optical bandwidth, which is commonly chosen to be 0.1 nm resolution, corresponding to 12.5 GHz bandwidth of optical spectrum analyzers at 1550 nm carrier wavelength.

We highlight some observations from Figure 2.3 as follows.

- At high SNR, all formats with an alphabet size M saturate to $\log_2 M$ bits per symbol, the maximum capacity of each respective constellation;

- Higher-order modulation formats are preferred to approach the Shannon limit;

- At low SNRs, OOK and 4 amplitude shift-keying (ASK) formats are outperformed by BPSK and QPSK, respectively, due to the smaller Euclidean distances in such intensity-modulated formats;

- The constellations using both DOFs (phase and amplitude) approach their capacity limit much faster than those using only one DOF, for instance, M-QAM formats versus M-PSK ($M \geq 16$).

In this thesis, we hence focus on both M-PSK and M-QAM signals capable of achieving better channel capacity at low SNRs but vulnerable to phase noise. It is worth noting that the PDM technique can double the channel capacity in principle without requiring more optical spectrum [13]. As further shown in Figure 1.1, the technique is widely deployed in all historical transmission experiments.

2.1.2 The Principle of Mach-Zehnder Modulator
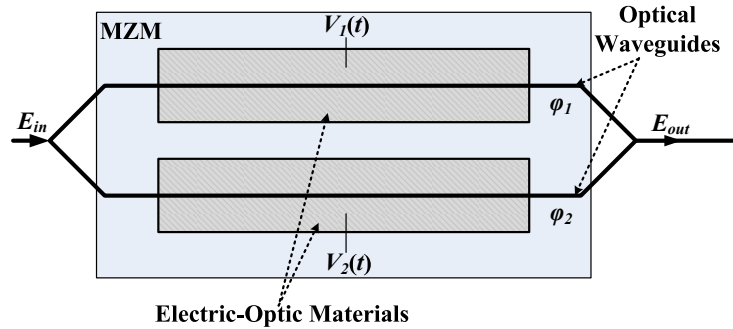
Figure 2.4: The structure of an MZM.

The structure of an MZM is shown in Figure 2.4. A coupler splits the incoming light E_{in} into two equal replicas, which pass through two identical optical waveguides controlled by the drive signals $V_1(t)$ and $V_2(t)$. Each replica experiences a phase shift ϕ_1 or ϕ_2 due to the electro-optic effect of the waveguide materials [66]. Generally, the phase shift, depending linearly on the drive signals, can be represented by

$$\phi_{1/2} = \frac{1}{2} \frac{V_{1/2}}{V_\pi} \pi, \tag{2.5}$$

where V_π is defined as the required voltage switching the output light intensity from its maximum to its minimum value [67]. Note that the factor 1/2 accounts for the

combined phase modulation for both two arms [27]. The two fields interfere at the output coupler to generate the output as given by

$$
\begin{aligned}
E_{out} &= \frac{E_{in}}{2} \left[\exp\left(j\phi_1(t)\right) + \exp\left(j\phi_2(t)\right)\right] \\
&= E_{in} \cos\left(\frac{V_1(t) - V_2(t)}{4V_\pi} \cdot \pi\right) \cdot \exp\left(j\frac{V_1(t) + V_2(t)}{4V_\pi} \cdot \pi\right).
\end{aligned}
\tag{2.6}
$$

Therefore, the output power of the MZM can be written as

$$
|E_{out}|^2 = \frac{1}{2}|E_{in}|^2 \left[1 + \cos\left(\frac{V_1(t) - V_2(t)}{2V_\pi}\pi\right)\right].
\tag{2.7}
$$

It can be observed that the output power is related to the drive voltage difference $\Delta V(t) = V_1(t) - V_2(t)$. According to the characteristics of the $\cos(\cdot)$ function, the intensity of the output optical signals exhibits 2π periodicity.
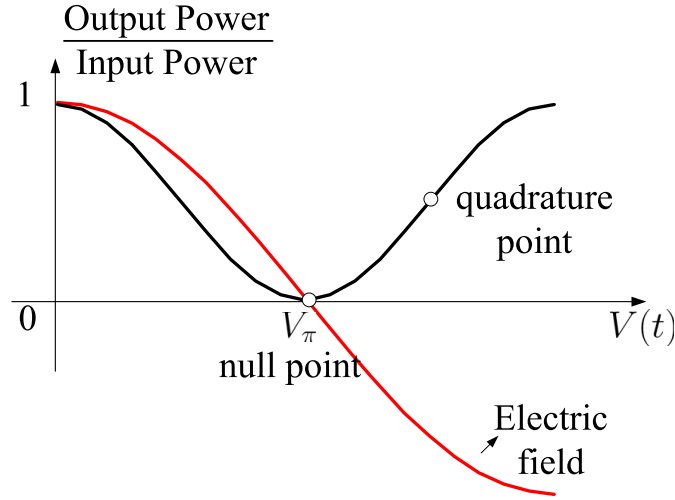
Figure 2.5: The power transfer function of an MZM.

As also shown from Eq.(2.6), there is a phase modulation $(\phi_1(t) + \phi_2(t))/2$ accompanying the intensity modulation. The ratio of phase to intensity modulation is called chirp parameter [68], which can be used to compensate for fiber dispersion [69]. In addition, this characteristic enables a single MZM to be used as an I/Q modulator to independently modulate the phase and amplitude by carefully adjusting $V_1(t) + V_2(t)$ and $V_1(t) - V_2(t)$ [70]. In digital optical communication, chirp is usually not desired

for the sake of increasing the transmission reach [56]. As a result, MZMs are typically operated at the push-pull condition, $V(t) = V_1(t) = -V_2(t)$, thus yielding into

$$E_{out} = E_{in} \cos\left(\frac{\pi V(t)}{2V_\pi}\right). \tag{2.8}$$

Its sine-shape transfer function is depicted in Figure 2.5. The null point refers to the V_π voltage to turn off the optical output of the MZM. Of importance in the phase-modulated systems is the electric field with respect to the electrical drive signal $V(t)$. The sign of the modulated optical field is capable of reflecting the change in drive signal $V(t)$, leading to the application of phase modulation. In contrast, the power transmission as a function of the drive signal is the key principle in intensity-modulated systems, in which the bias is normally set to the quadrature point, as indicated in Figure 2.5.

2.1.3 Generation of M-PSK/QAM

This section discusses a typical method using an I/Q modulator to generate M-PSK/QAM formats.
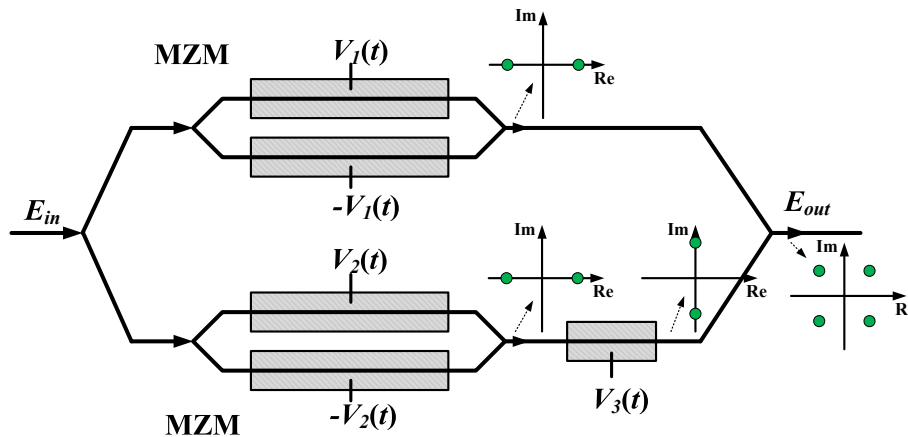
Figure 2.6: The structure of an I/Q modulator and the constellation of QPSK.

An integrated optical I/Q modulator has a nested MZM structure capable of converting the full complex plane of the optical field [67], as shown in Figure 2.6. The two

2.1 Advanced Optical Modulation Formats

sub-MZMs, operating at push-pull mode to achieve chirp-free signal, modulate their respective lightwave independently according to the drive signal $V_1(t)$ and $V_2(t)$. The third modulator driving by a DC $V_3(t)$ introduces a constant $90°$ phase shift between the two signals in upper and lower branch, thus putting them in quadrature to each other. Recalling the transfer function Eq.(2.8) of a single MZM, the output of the I/Q modulator can be written as

$$E_{out} = \frac{E_{in}}{\sqrt{2}} \left[\cos\left(\frac{\pi V_1(t)}{2V_\pi}\right) + j \cdot \cos\left(\frac{\pi V_2(t)}{2V_\pi}\right) \right]. \tag{2.9}$$

This transfer function indicates that an I/Q modulator can directly generate M-PSK/QAM formats when both the driving signals $V_1(t)$ and $V_2(t)$ are M-level. As an example illustrated in Figure 2.6, QPSK ($\pm 1, \pm j$) can be generated using two binary drive signals, 0 and $2V_\pi$. The multi-level electrical signals can be obtained using arbitrary waveform generator (AWG), in which the bit resolution and the speed of the DAC limits the maximum modulation speed. Since binary drive signals can promise higher-speed modulation formats and lower cost, serial MZMs implementation is another alternative to achieve higher-order M-PSK/QAM signals [28]. It is worth pointing out that the drive signals of $V_1(t)$ and $V_2(t)$ are not necessary evenly spacing for evenly spacing signal points due to the nonlinear transfer characteristic of Eq.(2.9) [71]. In addition, the bias should be properly applied to the I/Q modulator to remove the DC of the input drive signals. For instance, when the biases of the two sub-MZMs are set at null point (V_π) in a QPSK I/Q MZM, the drive signals $V_1(t)$ and $V_2(t)$ need to swing from $-V_\pi$ to V_π. Note that differential encoding of the drive signal $V_1(t)$ and $V_2(t)$ is usually necessary to avoid phase rotation because of the phase ambiguity in the systems, such as phase noise [72].
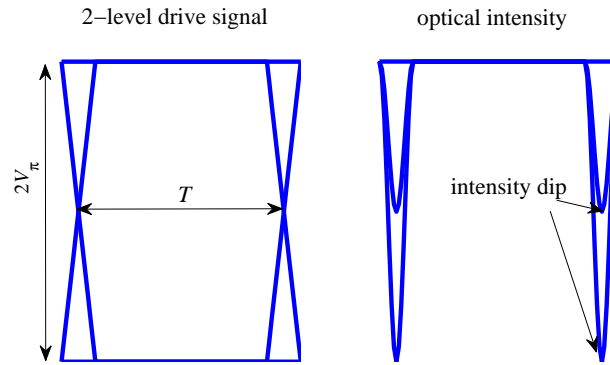
Figure 2.7: The eye diagram of the two-level drive signal and the generated output intensity in an I/Q modulator.

2.1.4 Pulse Carver

Figure 2.7 displays the eye diagram of the two-level electrical drive signal and the corresponding output intensity when using an I/Q modulator to generate non-return-to-zero (NRZ) QPSK signals. The drive signal has a peak-to-peak voltage $2V_\pi$ to obtain the 180° phase shift in each arm. Intensity dips occur between some consecutive symbols due to the transition between two levels [12]. A return-to-zero (RZ) pulse carver is usually applied to cut out the intensity dips in order to reduce its impact on the information-bearing optical phase [73]. Although RZ signals take up a broader optical spectrum than NRZ, it can have superior fiber nonlinearity tolerance over NRZ [74, 75]. The broader RZ signal spectrum, resulting in a 'soliton-like' properties [76], is shown to generally favor nonlinear transmission especially at high 10 Gb/s data rate and above [56].

An RZ pulse carver can be implemented using an MZM driven by a sinusoidal signal. Different duty cycle RZ pulses are obtained by changing the bias of an MZM modulator, the frequency and the amplitude of the sinusoidal signal [77], as depicted in Figure 2.8.

- **33% RZ**: driving an MZM (bias at $2V_\pi$) with a sinusoidal signal having peak-to-
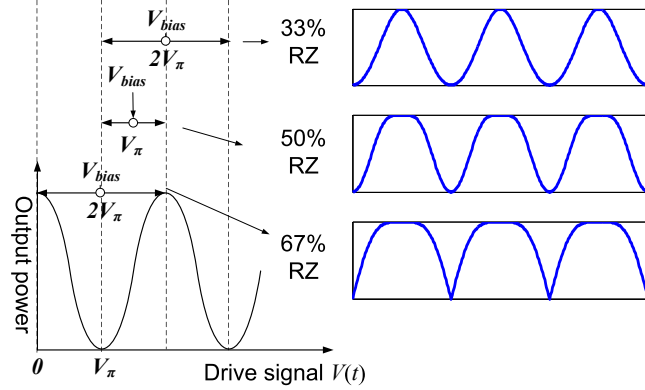
Figure 2.8: The illustration of generating 33%-, 50%- and 67%-duty cycle RZ pulse shape.

peak amplitude $2V_\pi$ at half the symbol rate;

- **50% RZ**: sinusoidally driving an MZM at the symbol rate with peak-to-peak amplitude V_π and bias at $3V_\pi/2$;

- **67% RZ**: sinusoidally driving an MZM at half the symbol rate with peak-to-peak amplitude $2V_\pi$ and bias at V_π. This format is also called carrier-suppressed RZ.

2.2 Transmission Links

The optical field E from the transmitter propagates through fiber, which imposes power attenuation, CD, PMD and fiber nonlinearity to the optical signals. When the other effects, such as Raman and Brillouin scattering, are negligible, the signal propagation in fiber can be described by the NLSE [78]:

$$\begin{aligned} \frac{\partial E}{\partial z} &= \left(\hat{D} + \hat{N} \right) E \\ \hat{D} &= -j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2} - \frac{\alpha}{2} \\ \hat{N} &= j\gamma|E|^2 \end{aligned} \tag{2.10}$$

where $\hat{D}$ is a differential operator accounting for group velocity dispersion (GVD) β_2 and attenuation α, and $\hat{N}$ is a nonlinear operator governing the effect of fiber nonlinearity γ. Note that higher-order dispersion, PMD and PDM are not taken into account here for simplification. The propagation equations governing evolution of the two polarization components along a fiber are characterized by two coupled-mode NLSEs [78]. Fully understanding the impact of different parameters on the signal propagation in optical fibers is helpful to design techniques to compensate for the transmission impairments in either optical or electrical domain.

2.2.1   Linear Fiber Impairments

We first study how the linear distortions, such as GVD and attenuation, affect the optical field E in the absence of the fiber nonlinearity ($\gamma = 0$). To solve the linear partial differential NLSE, Eq.(2.10) is transformed into frequency domain, thus yielding into [79]

$$\frac{\partial \mathbb{E}(z,\omega)}{\partial z} = \left(j\frac{\beta_2}{2}\omega^2 - \frac{\alpha}{2} \right) \mathbb{E}(z,\omega), \tag{2.11}$$

where $\mathbb{E}(z,\omega)$ is the Fourier transform of the optical electric field $E(z,t)$ located at distance z at the observation time t, and ω is the angular frequency. Note that the characteristic of Fourier transform $\frac{\partial^2}{\partial t^2} \xrightarrow{\mathbb{F}} (j\omega)^2$ is applied in deriving Eq.(2.11) [80]. The solution to the ordinary differential equation is easily to show as

$$\mathbb{E}(z,\omega) = \mathbb{E}(0,\omega) \exp\left(j\frac{\beta_2}{2}\omega^2 z - \frac{\alpha z}{2} \right). \tag{2.12}$$

The optical field at any distance z is obtained by converting the solution in Eq.(2.12) back into time domain through inverse Fourier transform:

$$E(z,t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} E(0,\omega) \exp\left(j\frac{\beta_2}{2}\omega^2 z - \frac{\alpha z}{2} - j\omega t \right) d\omega. \tag{2.13}$$

- **Fiber Loss**

Eq.(2.12) indicates that fiber loss α reduces the signal power. The power of the optical signal at distance z is given by $|E(z,t)|^2 = |E(0,t)|^2 \exp(-\alpha z)$. It is customary to express α in units of dB/km by using the relation [1]

$$\alpha(\text{dB/km}) = -\frac{10}{z} \log_{10} \left(\frac{|E(z,t)|^2}{|E(0,t)|^2} \right) \approx 4.343\alpha, \qquad (2.14)$$

and it is called the fiber-loss parameter. Fiber loss leads to exponential decay of the signal power, thus playing a critical role in the optical fiber communication. Attenuation is caused by absorption, scattering, and bending losses in an optical fiber. The sophisticated fabrication of the fiber has reduced the fiber loss from 20 dB/km in the early days to 0.2 dB/km in 1979 [81] and even commercial $\sim$0.17 dB/km at 1.55 μm nowadays [82]! On the other hand, the advent of optical amplifiers in the 1990s enables that transmission distances to easily exceed several thousands kilometers by compensating for accumulated losses periodically [83].

- **Fiber Dispersion**

Although GVD only changes the phase of each spectral component, it is capable of broadening the optical pulse, thus resulting in inter-symbol interference (ISI) [1]. This can be illustrated by considering the case of a Gaussian pulse traveling along a dispersive fiber. The incident Gaussian pulse has the form

$$E(0,t) = \exp \left(-\frac{t^2}{2T_0^2} \right), \qquad (2.15)$$

where T_0 is the half-width (at $1/e$-intensity point). By using Eq.(2.12) and carrying out the inverse Fourier transform like in Eq.(2.13), the electric field at distance z along the fiber is characterized by

$$E(z,t) = -\frac{T_0}{\sqrt{(T_0^2 - j\beta_2 z)}} \exp \left(-\frac{t^2}{2(T_0^2 - j\beta_2 z)} \right). \qquad (2.16)$$

Note that the Fourier transform of a Gaussian function is still a Gaussian, i.e., $\exp\left(-at^2\right)$ $\overset{\mathbb{F}}{\to} \sqrt{\pi/a}\exp\left[-\pi^2\omega^2/(4a)\right]$. The dispersion parameter D, expressed in units of ps/(km $\cdot$ nm), is commonly used to characterize the dispersive feature in the fiber, and is related to GVD β_2 through [78]

$$D = -\frac{2\pi c}{\lambda^2}\beta_2, \tag{2.17}$$

where c and λ are the light velocity in the vacuum and optical carrier wavelength, respectively.
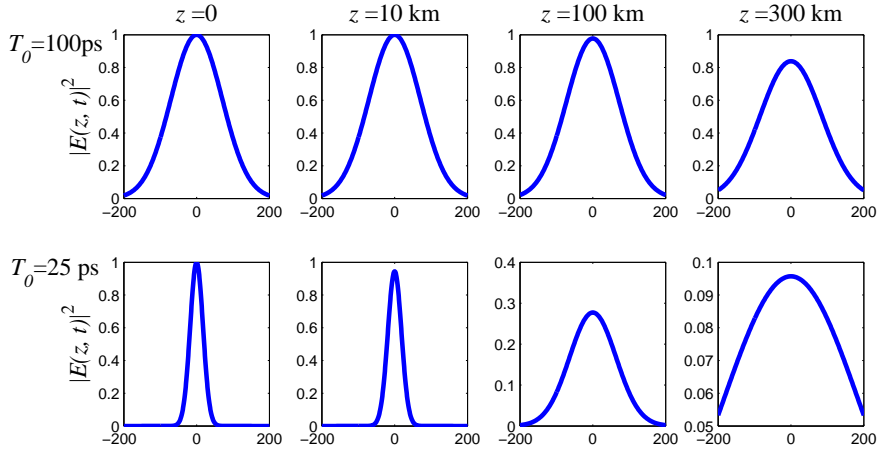
Figure 2.9: Dispersion-induced broadening of a Gaussian pulse in a fiber at different distances (z=0, 10 km, 100 km and 300 km) and pulse width (T_0=100 ps and 25 ps). λ=1.55 μm and D= 17 ps/(km·nm). (The unit of x-axis is ps.)

As illustrated in Figure 2.9, the extent of broadening depends on the initial pulse width and transmission distance. Shorter pulses, corresponding to higher symbol rate, experience more serious dispersion-induced broadening effects than wide pulses. As a result, the broadened pulses spread into the neighboring pulses to cause serious ISI, resulting in much shorter optical reach. When optical amplifiers are used to restore the attenuated signal power in long-haul transmission systems, the accumulated dispersion in the optical pulse becomes of serious concern. One solution is to reduce the symbol rate of the optical systems by means of employing higher-order modulation formats,

which have smaller symbol rate to achieve the same bit rate compared to OOK formats [84]. Further, considering that dispersion is a linear effect, dispersion management is a superior solution to keep the average GVD of the entire fiber link quite low or even zero [85, 86]. For a dispersion map consisting of two fiber segments, as shown in Figure 2.10 (a), Eq.(2.13) now becomes

$$E(L, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} E(0, \omega) \exp\left(\frac{j}{2}\omega^2(\beta_{21}L_1 + \beta_{22}L_2) - j\omega t\right) d\omega, \qquad (2.18)$$

where $L = L_1 + L_2$ is the dispersion-map period, β_{2i} is the GVD parameter of the fiber segment with length L_i. It can be observed that the original optical field can be restored, i.e., $E(L, t) = E(0, t)$, when the following condition is satisfied:

$$D_1 L_1 + D_2 L_2 = 0. \qquad (2.19)$$

Here, the GVDs β_{2i} are replaced by the commonly-used dispersion parameter D_i through Eq.(2.17). The prerequisite shows that the second fiber should have an opposite sign of the dispersion D_1. The second fiber with a negative dispersion, capable of fully canceling out the dispersion accumulated in the first span, is called dispersion-compensating fiber (DCF). In practical systems, a DCF module is added at the site of the optical amplifiers to compensate for the accumulated dispersion while the fiber losses are addressed by the amplifier. However, the nonlinear effect in DCFs are considerably enhanced due to the small effective mode area when the fiber input power is still high [87]. Transmission link consisting of single-mode fiber (SMF) without DCF (see Figure 2.10 (b)) becomes realistic only when the dispersion compensation is carried out at the transmitter or receiver side. The advance in high-speed DAC or ADC makes such compensation techniques feasible in current optical systems [88].

- **Polarization-Mode Dispersion**

SMF actually support two perpendicular polarizations of the original transmitted signal (fundamental mode). The two perpendicular polarization components of the
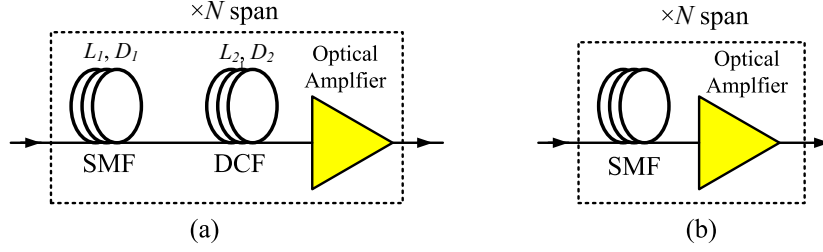
Figure 2.10: Transmission link: (a) dispersion-managed fiber; (b) uncompensated dispersion link.

signal light may travel at different velocities due to the imperfect cylindrical symmetry in fibers [89]. Therefore, the two components arrive at the end of the fiber at different times, leading to pulse broadening. This phenomenon is called PMD. PMD vector $\vec{\tau}$ can be expanded into Taylor series, given by

$$\vec{\tau} \approx \vec{\tau}(\omega_0) + \vec{\tau}^{(1)}(\omega_0)\Delta\omega, \tag{2.20}$$

where $\vec{\tau}^{(1)}(\omega_0) = d\vec{\tau}/d\omega|_{\omega=\omega_0}$. The first (0th-order in frequency) term of this expansion corresponds to the most widely used first-order PMD approximation [90]. Except the broadening of the pulse, the random variation of the polarization states imposes a performance degradation in coherent receiver. It is worth pointing out that the polarization state of light propagating in fibers would change randomly along the fiber during propagation. It is shown that PMD is a statistically random quantity [91].

The second-order PMD approximation adds the second term of the Taylor's expansion (first-order frequency derivative) [90]. It indicates that the PMD vector $\vec{\tau}$ can change either its direction or its magnitude. The change in the magnitude leads to a linear change in the group delay across the signal spectrum, i.e., CD. Pulses launched into the two PSPs (which are parallel and anti-parallel to the PMD vector) $\vec{\tau}(\omega_0)$ and $-\vec{\tau}(\omega_0)$ see different amounts of CD, with the difference being proportional to $d\tau/d\omega$ [90].

In a PDM optical system, PMD will cause cross-talk between the signals in the two polarization states in the presence of PMD and polarization dependent loss [15].

Hence, some extra processing procedures are necessary to separate the two polarization states.

2.2.2 Fiber Nonlinearity

The origin of fiber nonlinearity is that the refractive index of silica fiber is power dependent [78], given by

$$\gamma = \frac{2\pi\bar{n}_2}{A_{\text{eff}}\lambda}, \tag{2.21}$$

where $\bar{n}_2$ is the nonlinear index coefficient, and A_{eff} is the effective mode area. In the absence of GVD β_2, the general NLSE Eq.(2.10) can be rewritten as

$$\frac{\partial E}{\partial z} = j\gamma|E|^2 E \cdot \exp(-\alpha z), \tag{2.22}$$

Here, the exponential factor $\exp(-\alpha z)$ accounts for the fiber loss. The general solution to Eq.(2.22) has the form of $E = V\exp(j\Phi_{\text{NL}}(z))$, where V is the signal amplitude. Substituting the general solution to Eq.(2.22), we can have

$$\frac{\partial V}{\partial z} = 0; \qquad \frac{\partial \Phi_{\text{NL}}(z)}{\partial z} = \gamma V^2 \exp(-\alpha z). \tag{2.23}$$

It means that the signal amplitude V does not vary along the fiber length. The optical electric field at the end of the fiber length L becomes

$$E(L, t) = E(0, t)\exp(j\Phi_{\text{NL}}(z)), \tag{2.24}$$

where the nonlinear phase noise can be integrated analytically as

$$\Phi_{\text{NL}}(L) = \int_0^L \gamma|E(0, t)|^2 \exp(-\alpha z)dz = \gamma L_{\text{eff}}|E(0, t)|^2, \tag{2.25}$$

with effective length L_{eff} defined as

$$L_{\text{eff}} = \frac{1 - \exp(-\alpha L)}{\alpha}. \tag{2.26}$$

Eqs.(2.24)-(2.25) show that optical pulses experience intensity-dependent phase modulation, called SPM [92]. This SPM phenomenon will limit the maximum signal power

launched into optical fibers, which is in contrary to the conventional wireless channels. The temporally varying phase modulation also associates with the pulse shape, thus resulting in spectra broadening [78]. It is worth pointing out that the SPM effect occurs at the initial effective length L_{eff} of a fiber because fiber nonlinearity has a strong impact at high power [47]. Eventually, the attenuated signal power mitigates the effect after the effective length.

In view of this simple characteristic between SPM and signal input power, the overall accumulated nonlinear phase noise after passing through N amplifiers can be approximated by [49]

$$
\begin{aligned}
\phi_{\text{NL}}(t) &= \sum_{k=1}^{N} \phi_{\text{NL}}(k, t) \\
&= \gamma L_{\text{eff}} \sum_{k=1}^{N} |E(k, t)|^2 \\
&= \gamma L_{\text{eff}} \{ |E(0, t) + n_1(t)|^2 + |E(0, t) + n_1(t) + n_2(t)|^2 + \cdots \\
&\quad + |E(0, t) + n_1(t) + \cdots + n_N(t)|^2 \}
\end{aligned}
\tag{2.27}
$$

where $n_k(t), k = 1, \ldots, N$ are those AWGNs introduced by kth amplifier. Henceforth, the received signal can be written as

$$
E_r(t) = E(N, t) \cdot \exp\left[j\phi_{\text{NL}}(t) \right]
\tag{2.28}
$$

Note that nonlinear phase noise is suggested being χ^2 distributed due to square addition of the Gaussian ASE noises $\{n_1, n_2, \ldots, n_k, \ldots\}$, which is verified by experiment data in [93]. This nonlinear optical channel model has been well adopted to investigate performances of different modulation formats [27, 72, 94], though it would overestimate the penalty of fiber nonlinearity without taking into account the interaction between CD and fiber nonlinearity [48].

The power-dependent refractive index can also lead to other nonlinear effects, such as cross-phase modulation (XPM) [95], four-wave mixing [78], especially in WDM systems. All these nonlinear effects would interact with other linear fiber effects to

severely limit the performance of long-haul lightwave systems [77]. Of importance is that the SPM/XPM-induced phase shift would cause a severe performance degradation in phase-modulated systems.

2.2.3   Split-Step Fourier Method

We have separately investigated the impact of fiber linearity and nonlinearity. In fact, fiber dispersion and nonlinearity would interact with each other along the length of the fiber, though a general analytical solution is not available for the NLSE Eq.(2.10) [78]. Consequently, numerical approaches are necessary to emulate the optical pulse evolution in the fiber for evaluating various fiber impairments. One approach that has been used extensively to solve the pulse-propagation problem is the split-step Fourier method [96] when both fiber dispersion and nonlinearity are taken into account.
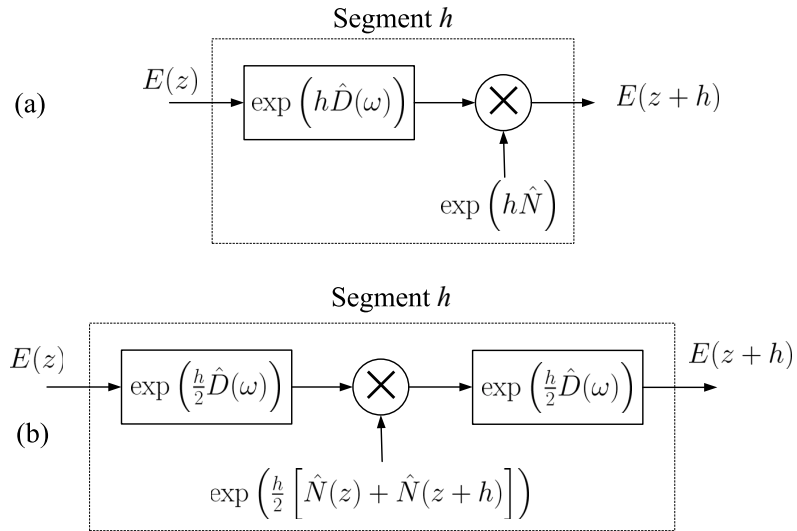
Figure 2.11: The schematic illustration of: (a) split-step Fourier method; (b) symmetrized split-step Fourier method.

The split-step Fourier method obtains an approximate solution by assuming that the dispersion and the nonlinearity factors can be treated separately in the propagation

of the optical field over a small distance h. The nonlinear section is first performed in the time domain, as illustrated in Figure 2.11 (a). Through Fast Fourier transform (FFT), the signal is then transformed into frequency domain where each spectral component is multiplied by a phase shift due to the fiber GVD, as described in Eq.(2.12). Note that $\hat{D}(\omega) = j\frac{\beta_2}{2}\omega^2$. To further improve the accuracy of the split-step Fourier method, fiber nonlinearity is evaluated in the middle of the segment h instead of carrying out the step at the beginning of the segment (see Figure 2.11 (b)) [97]. The mathematical expression of the modified method can be represented by

$$E(z+h,t) \approx \exp\left(\frac{h}{2}\hat{D}\right) \exp\left(\int_z^{z+h} \hat{N}(z')dz'\right) \exp\left(\frac{h}{2}\hat{D}\right) E(z,t). \qquad (2.29)$$

This scheme is called symmetrized split-step Fourier method because of the symmetric form of the exponential operators in Eq.(2.29), in which the integral is usually approximated by the trapezoidal rule [98]:

$$\int_z^{z+h} \hat{N}(z')dz' \approx \frac{h}{2}\left[\hat{N}(z) + \hat{N}(z+h)\right]. \qquad (2.30)$$

Note that higher-order dispersion is not included here. However, $\hat{N}(z+h)$ is unknown at the middle of the segment located at $z+h/2$. An iterative procedure is introduced to replace the term $\hat{N}(z+h)$ by $\hat{N}(z)$ at first. Eq.(2.29) is then used to estimate $E(z+h,t)$ which in turn is applied to calculate the new $\hat{N}(z+h)$. The procedure is repeated several times until certain tolerance is satisfied, such as the maximum phase shift and the loss of signal power during the segment. More efforts have been put to further improve the computational efficiency and enhance the simulation accuracy [99]. It is suggested that the fiber can be modeled as a concatenation of nonlinear and linear sections, in which the propagation of the optical field is governed by the prescription of Eq.(2.29) [78].

2.3   Coherent Receiver

Compared to direction detection, coherent detection can preserve full information of the optical electric field, such as the amplitude, phase and polarization, which provides modulation in more DOF for multi-level signalling [15] such that approaching Shannon capacity limit, as depicted in Figure 1.1. The application of high-speed ADCs in coherent receivers has triggered extensive research in DSP algorithms to compensate for the fiber transmission impairments rather than using costly and complicated optical components [24, 100–103].
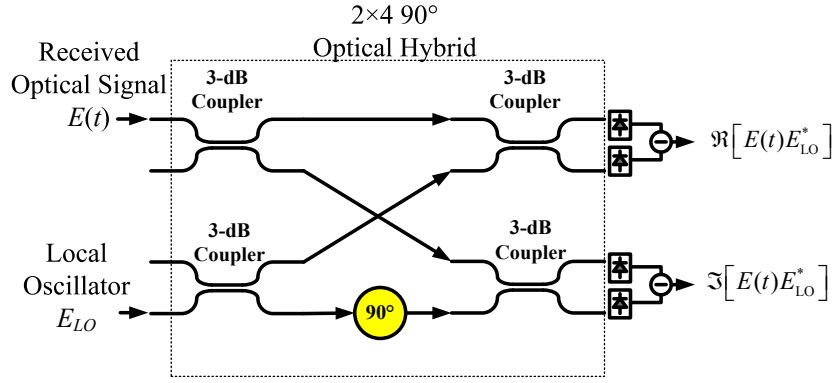
Figure 2.12: A schematic setup of a single-polarization coherent receiver.

The structure of a single-polarization coherent receiver with balanced photodetector is shown in Figure 2.12. The received optical field $E(t)$ is combined with an LO laser in a 2×4 $90°$ optical hybrid, whose output signals are detected by two balanced photodetectors corresponding to the in-phase (I) and quadrature-phase (Q) signals [104]. The optical hybrid is composed of four 3-dB couplers and an additional $90°$ phase shifter. The transfer matrix of the optical hybrid is given as [104]

$$S = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & j \\ 1 & -j \end{bmatrix}. \tag{2.31}$$

After the square-law photodetection, the photocurrents from the upper balanced detectors can be written as [94]

$$
\begin{aligned}
I &= \frac{1}{4}R\,|E(t)+E_{\mathrm{LO}}|^2 - \frac{1}{4}R\,|E(t)-E_{\mathrm{LO}}|^2 + i_{\mathrm{sh}}(t) \\
&= R\cdot\mathbf{Re}\left[E(t)E_{\mathrm{LO}}^*\right] + i_{\mathrm{sh1}}(t)
\end{aligned}
\tag{2.32}
$$

where R is the photodetector responsitivity, $i_{\mathrm{sh1}}(t)$ represents the shot noise, $\mathbf{Re}[x]$ and * denote the real part of x and conjugation operation, respectively. Here

$$
E_{\mathrm{LO}} = \sqrt{P_{\mathrm{LO}}}\exp\left(j\omega_{\mathrm{LO}}t + j\theta_{\mathrm{LO}}(t)\right),
\tag{2.33}
$$

where P_{LO}, ω_{LO} and $\theta_{\mathrm{LO}}(t)$, respectively, are the power, the carrier angular frequency and the phase noise of the LO signal. Applying the same derivation as in Eq.(2.32), the lower branch signal is derived as

$$
\begin{aligned}
Q &= \frac{1}{4}R\,|E(t)+jE_{\mathrm{LO}}|^2 - \frac{1}{4}R\,|E(t)-jE_{\mathrm{LO}}|^2 + i_{\mathrm{sh}}(t) \\
&= R\cdot\mathbf{Im}\left[E(t)E_{\mathrm{LO}}^*\right] + i_{\mathrm{sh2}}(t),
\end{aligned}
\tag{2.34}
$$

where $\mathbf{Im}[x]$ denotes the imaginary part of x and $i_{\mathrm{sh2}}(t)$ represents the shot noise. As can be seen from Eqs.(2.32) and (2.34), the outputs of the two pairs of balanced photodetectors correspond to the in-phase and quadrature of the received optical electric field. When sampling the signals at the Nyquist rate using ADCs, the digitalized signals can be further processed by either real-time or offline DSP algorithms to compensate for the transmission impairments [62]. This receiver structure is also universal for any multi-level modulation formats [28]. Note that a dual-polarization coherent receiver can be realized by splitting the polarization of the incident optical field into two orthogonal states, and each of them beats with the respective LO counterpart which has the identical polarization state as the signal's [41].

The impact of optical noises and the shot noises on the performance of a coherent system is worthy of investigation. Due to the fiber loss, the transmitted signal are

repeatedly amplified by optical amplifiers, which introduce ASE noises $N(t)$ into the transmitted signal $E_s(t)$. Hence, the received signal becomes $E(t) = E_s(t) + N(t)$. The transmitted signal is assumed to have the form of

$$E_s(t) = \sqrt{P_s} \exp\left(j\omega_c t + j\phi_s(t) + j\theta(t)\right), \qquad (2.35)$$

where P_s and $\phi_s(t)$ are the power and the signal phase, $\theta(t)$ is the phase noise associated with the source laser, and ω_c is the carrier angular frequency of the signal. The spectral density of the received ASE noise $N(t)$ with N_A-cascaded optical amplifiers is given by [1]

$$S_{\text{ASE}} = N_A n_{\text{sp}}(G - 1)h\nu, \qquad (2.36)$$

where n_{sp} and G are the spontaneous emission factor and gain of the optical amplifiers, and $h\nu$ is a photon energy. In terms of practical applications, of interest is homodyne detection ($\omega_c = \omega_{\text{LO}}$) because the bandwidth of the electrical low-pass filter required for heterodyne detection ($\omega_c \neq \omega_{\text{LO}}$) is R_s whereas one half of this bandwidth, i.e., $R_s/2$, is enough for the homodyne case [105]. Here, R_s is the symbol rate of a system. By combining the in-phase and quadrature signals together, we can derive the baseband complex signals as

$$
\begin{aligned}
r(t) &= R \cdot \mathbf{Re}\left[E(t)E_{\text{LO}}^*\right] + i_{\text{sh1}}(t) + j \cdot \left(R \cdot \mathbf{Im}\left[E(t)E_{\text{LO}}^*\right] + i_{\text{sh2}}(t)\right) \\
&= R\sqrt{P_s P_{\text{LO}}} \exp\left[j\left(\phi_s(t) + \theta_s(t) - \theta_{\text{LO}}(t)\right)\right] + i_{\text{LO-ASE}}(t) + i_{\text{sh}}(t),(2.37)
\end{aligned}
$$

where the LO-ASE beat noise $i_{\text{LO-ASE}}(t) = R \cdot \mathbf{Re}\left[N(t)E_{\text{LO}}^*\right] + j \cdot R \cdot \mathbf{Im}\left[N(t)E_{\text{LO}}^*\right]$, whose two-sided power spectrum density (PSD) is found to be

$$S_{\text{LO-ASE}} = R^2 P_{\text{LO}} S_{\text{ASE}}. \qquad (2.38)$$

The shot noise $i_{\text{sh}}(t)$ equals to $i_{\text{sh1}}(t) + j \cdot i_{\text{sh2}}(t)$ having a two-sided PSD S_{sh} of $eR(P_{\text{LO}} + P_s)$, where e is the electron charge. The signal power divided by the total noise power in the same band, $R_s/2$ in homodyne receivers, gives the SNR per

symbol (γ_s) at the detector, as given by [94]

$$\gamma_s = \frac{R^2 P_{\mathrm{LO}} P_s}{2\left(R^2 P_{\mathrm{LO}} S_{\mathrm{ASE}} + eR(P_{\mathrm{LO}} + P_s)\right) \cdot R_s/2}. \tag{2.39}$$
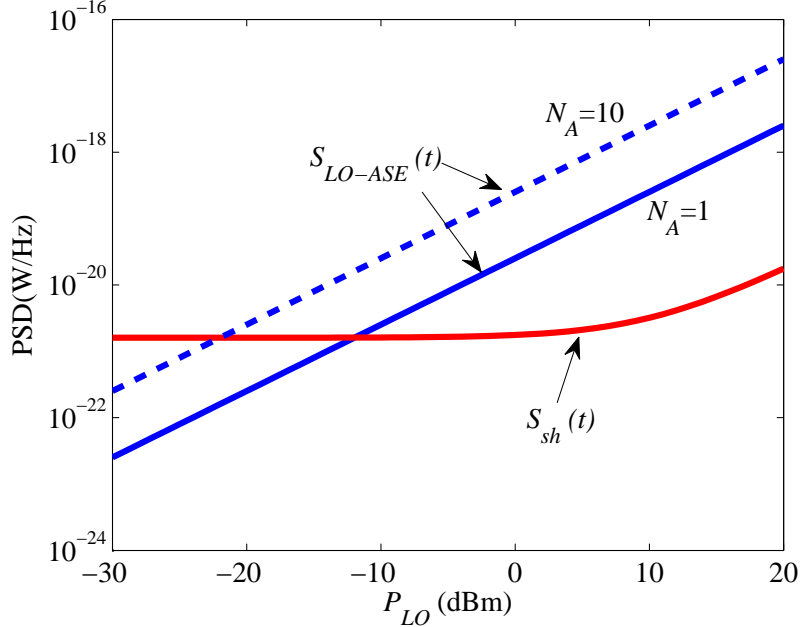
Figure 2.13: The PSD of different noises (LO-ASE beat noise and shot noise) as a function of LO power in a coherent receiver at P_s=10 dBm.

As illustrated in Figure 2.13, the shot noise is dominant over LO-ASE beat noise only when the power of LO laser is very small. In practice, the LO power is always kept at least above 0 dBm to improve the receiver sensitivity. Besides, the use of cascaded optical amplifiers will enhance the LO-ASE noises. Consequently, the LO-ASE beat noise is the major source of additive noises in coherent receivers, thus yielding the ASE-limited SNR per symbol as [27]

$$\mathrm{SNR}_{\mathrm{sym}} = \frac{N_s}{N_A n_{\mathrm{sp}}}, \tag{2.40}$$

because $P_s = G R_s N_s h\nu$ when assuming that the last optical amplifier acts as a preamplifier with gain G. Here, N_s is the average number of photons per symbol.

Other than the amplitude distortion, the received signals are also rotated by the phase noise $\theta_s(t) - \theta_{\mathrm{LO}}(t)$ from the transmitter and LO lasers. The power spectrum of

laser linewidth $\Delta\nu$ has a Lorentzian line-shape, inducing a Gaussian-distributed phase deviation with mean zero and variance [106]

$$\sigma_p^2 = 2\pi(2\Delta\nu)T_s \tag{2.41}$$

in a symbol interval $T_s = 1/R_s$. Note that $2\Delta\nu$ accounts for the total 3-dB linewidth for both lasers. The laser phase noise keeps rotating the received signals, thus requiring phase tracking in coherent receiver.

2.4   DSP Algorithms in Coherent Receivers

As reviewed in Section 2.2, the transmitted signal after fiber transmission will be deteriorated by the CD, PMD and fiber nonlinearity such that receivers are unable to recover the original transmitted data if no specific measures are employed in the systems. A digital coherent receiver samples the in-phase and quadrature signals into digital waveforms at Nyquist rate or above (see Section 2.3), consequently applying DSP algorithms to fully recover the transmitted data [80]. Several benefits are seen to using DSP-based algorithms in such coherent receivers: adaptive filters can provide better performance and design versatility while a universal receiver structure can be used for reception of different modulation formats by adjusting only the parameters of digital filters [107].

The DSP algorithms in digital coherent receivers are divided into several steps [24], which may consist of clock recovery, re-sampling, CD compensation (CDC), polarization-demultiplexing, carrier phase recovery including frequency and phase estimation, symbol decision, differential decoding and error detector, as depicted in the architecture of Figure 2.14. If nonlinear impairments such as SPM or XPM are present, additional fiber nonlinearity compensation is necessary to improve the system performance [47], which will be discussed in detail in Chapter 5.
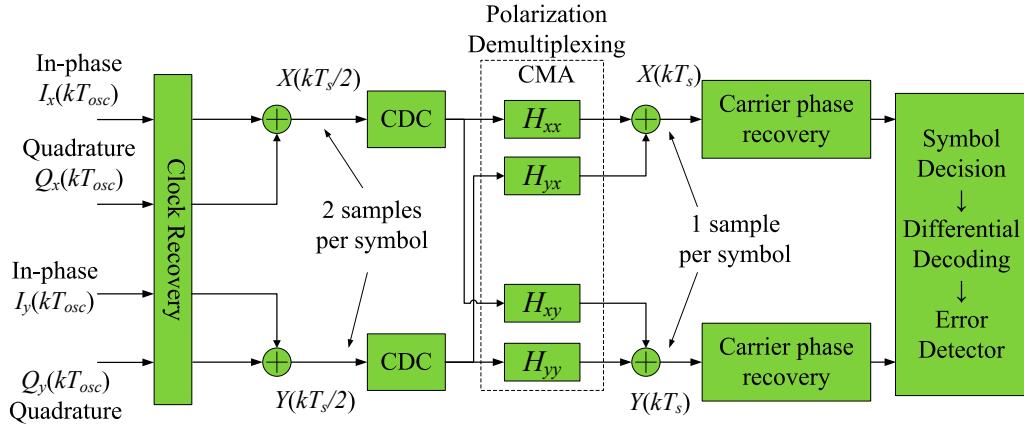
Figure 2.14: The architecture of the DSP algorithms in a polarization-multiplexing digital coherent receiver.

2.4.1   Clock Recovery and IQ Imbalance

The four streams of digitalized data from ADCs, corresponding to the in-phase and quadrature signals in two polarization states, are first re-sampled to 2 samples per symbol [21], because the sample rate $1/T_{\text{osc}}$ of a real-time oscilloscope may be not exactly two times the symbol rate of an optical system. In a clock recovery module, digital timing recovery algorithms are applied to correct the timing offset between the clocks of transmitter and receiver, which can be performed in a real-time oscilloscope. Commonly-used algorithms, such as the square timing recovery [108] and the Gardner algorithm [109], are able to estimate the timing error, which is then used to re-sample the data via an interpolator to obtain exactly 2 samples per symbol.

It is also worth mentioning that the normalization and IQ-imbalance compensation may be required when dealing with real experiment data [110, 111] because of the imperfections in optical front ends, where the amplitude of hybrid outputs may be not equal or the phase shifter in the hybrid (see Figure 2.12) may not be exactly 90° [107].

2.4.2 Channel Equalization

As noted in Section 2.2, fiber can be regarded as a linear system with only attenuation, CD and PMD in the absence of fiber nonlinearity. In principle, all the linear impairments can be fully compensated by using digital filters [62]. Since CD is a time-invariant distortion and the residual CD may exceed several thousands of ps/nm at the receiver end, it is desirable to use nonadaptive filter with fixed taps for the compensation [112]. Although infinite impulse response (IIR) filter requires less taps than finite impulse response (FIR) filters, IIR filters tend to be unstable especially when the residual CD is large [113]. This drawback limits its application in practical systems. On the other hand, the PMD effect may be modeled by the Jones matrix, which varies rapidly due to the fiber birefringence [89]. The compensation scheme of PMD should be adaptive to continuously adjust the tap coefficients [30]. As a result, the channel equalization in coherent receivers is split into two sub-steps: fixed-tap FIR filters for CDC and a time-domain adaptive butterfly-like filter for PMD compensation.

Recalling the fiber dispersion effect on the transmitted signal in Section 2.2, the fiber can be regarded as an all-pass filter only incurring a phase shift to different spectral components. It suggests that an inverse function of the all-pass filter suffices to fully compensate for the CD of a fiber link [114]. This approach can be realized in either time-domain or frequency-domain [112].

The time-domain CDC in the DSP unit is a fractionally-spaced $(T_s/2)$ FIR equalizer [115], as plotted in Figure 2.15. The tap coefficients are computed from the inverse Fourier transform of the fiber transfer function (see Eq.(2.12)) [116], as given by:

$$H(f) = \exp\left(-j\pi\frac{D\lambda^2 f^2}{c}z\right) \tag{2.42}$$

while changing the sign of D to the opposite. Note that the number of filter taps depends on the amount of the accumulated dispersion in the fiber link of length z. To approach the desired phase response $\mathbb{F}^{-1}\{H(f)\}$, the tap coefficients C_k is computed
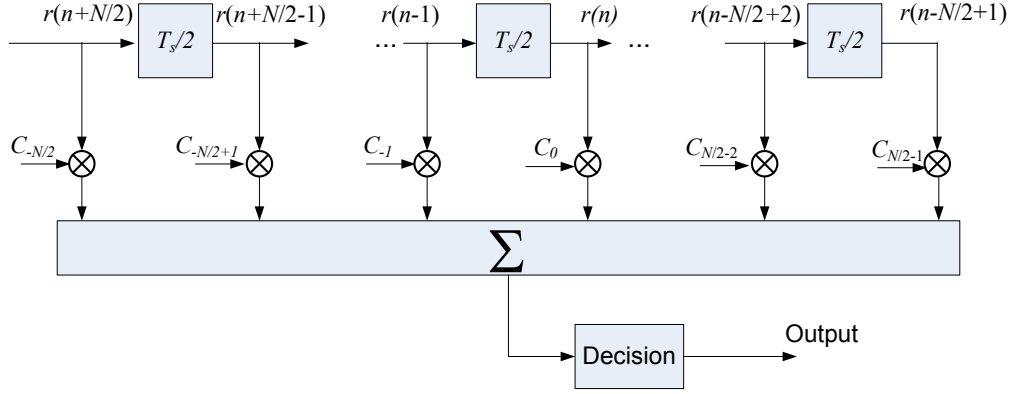
Figure 2.15: A fractionally-spaced ($T_s/2$) FIR filter with N taps.

through sampling the inverse Fourier transform of the transfer function Eq.(2.42), i.e. [116],

$$C_k = \frac{kT_s}{2}h(t = \frac{kT_s}{2}), \qquad (2.43)$$

where

$$
\begin{aligned}
h(t) &= \int_{-\infty}^{\infty} H(f)\exp{(j2\pi ft)}df \\
&= \Delta f \cdot \sum_{i=-\tilde{N}/2}^{\tilde{N}/2-1} H(f = i\Delta f)\exp{(j2\pi(i\Delta f)t)}. \qquad (2.44)
\end{aligned}
$$

Here, the integration is approximated by adding up those discrete components from $-1/T_s$ to $1/T_s$ with a stepsize Δf. If the accumulated dispersion Dz and symbol rate R_s are known, the FIR tap coefficients, as illustrated in Figure 2.16, are determined through Eqs.(2.43) and (2.44). In theory, CD can be completely compensated for if the number of taps of the FIR equalizer is sufficiently large [30]. It is found that a FIR with tap number $N = 4\pi|\beta_2|LR_s^2$ is sufficient to fully compensate for CD [117].

As for the frequency-domain equalizer, the re-sampled signals $X(kT_s/2)$ and $Y(kT_s/2)$ are transformed into frequency domain through FFT. The frequency components are then multiplied by the fiber transfer function $H(f)$ to cancel out the CD impact, and are converted back into time domain for subsequent signal processions. Generally, frequency-domain equalization is much faster than time-domain equalizer due
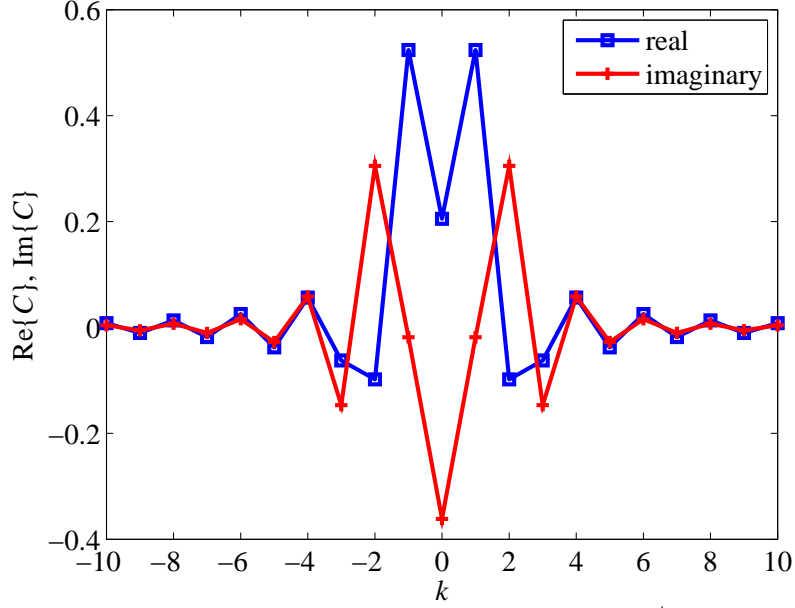
Figure 2.16: The tap coefficients of a fractionally-spaced ($T_s/2$) FIR filter with N=21, accumulated dispersion Dz= 1700 ps/nm, and symbol rate R_s=10 GBaud.

to the FFT operation, especially when the number of FIR taps becomes large [103,112]. It is worth noting that fiber nonlinearity compensation is necessary if there exists strong fiber nonlinearity effect. In this case, the fiber dispersion will be simultaneously compensated for through back-propagation of the received electric field [47, 118], which will be addressed elaborately in Chapter 5.

After the nonadaptive FIR filtering for CDC, the subsequent butterfly-like FIR filter aims to remove the cross-talk between two transmitted signals in these two orthogonal polarization states, because of the PMD effect [30]. In a PolMux system, the optical channel can be modeled as a 2×2 multi-input multi-output system [41], as given by

$$\begin{aligned}
X_{\text{out}}(k) &= \mathbf{H}_{xx}^T(k)\mathbf{X}_{\text{in}}(k) + \mathbf{H}_{yx}^T(k)\mathbf{Y}_{\text{in}}(k) \\
Y_{\text{out}}(k) &= \mathbf{H}_{xy}(k)^T\mathbf{X}_{\text{in}}(k) + \mathbf{H}_{yy}^T(k)\mathbf{Y}_{\text{in}}(k),
\end{aligned} \tag{2.45}$$

where $\mathbf{H}_{xx}(k)$, $\mathbf{H}_{xy}(k)$, $\mathbf{H}_{yx}(k)$ and $\mathbf{H}_{yy}(k)$ are $N \times 1$ tap coefficient vectors of the

butterfly FIR filter at $t = kT_s$, $X_{in}(k)$ and $Y_{in}(k)$ are $N \times 1$ input vectors at $t = kT_s$, and $(\cdot)^T$ denotes matrix transpose. Of importance is how to de-multiplex the two signals from the two polarization states. The blind constant modulus algorithm (CMA) is a popular method for constant-modulus modulation formats [119], such as M-PSK, where the magnitude errors of $\varepsilon_x = 1 - |X_{out}(k)|^2$ and $\varepsilon_y = 1 - |Y_{out}(k)|^2$ are examined to adjust the coefficient vectors through

$$\mathbf{H}_{xx}(k+1) = \mathbf{H}_{xx}(k) + \mu\varepsilon_x X_{out}(k)\mathbf{X}_{in}(k) \qquad (2.46)$$

$$\mathbf{H}_{yx}(k+1) = \mathbf{H}_{yx}(k) + \mu\varepsilon_x X_{out}(k)\mathbf{Y}_{in}(k) \qquad (2.47)$$

$$\mathbf{H}_{xy}(k+1) = \mathbf{H}_{xy}(k) + \mu\varepsilon_y Y_{out}(k)\mathbf{X}_{in}(k) \qquad (2.48)$$

$$\mathbf{H}_{yy}(k+1) = \mathbf{H}_{yy}(k) + \mu\varepsilon_y Y_{out}(k)\mathbf{Y}_{in}(k) \qquad (2.49)$$

where μ is the step size of the CMA algorithm. Note that the tap vectors are updated once every symbol. The decision-directed (DD) least-mean square (LMS) algorithm can be introduced to improve the system performance once the CMA algorithm has reached the steady state [23]. As for the non-constant-modulus QAM formats, this CMA criterion is usually applied at the first stage to acquire a pre-convergence [120]. A modified CMA such as multi-modulus CMA, and DD-LMS [121], are carried out to fully de-multiplex these two signals in conjunction with carrier phase recovery.

It is worth of pointing out that the higher-order dispersion, like dispersion slope, can still be compensated once the dispersion slope and channel wavelength are known to the receiver. Further, quite a few DSP algorithms [122] are already proposed to compensate for XPM from inter/intra channels and references therein.

2.4.3 Carrier Phase Recovery

When ISI of the received signals is removed by the channel equalization, the remaining distortions include only frequency offset, laser phase noise and additive noises.

2.4 DSP Algorithms in Coherent Receivers

The signal now becomes the same as Eq.(2.37), and the sampled waveforms are then described as [4]

$$
\begin{aligned}
r(k) &= I(kT_s) + j \cdot Q(kT_s) \\
&= A_0 \exp\left(j\phi_s(k) + j\theta(k) + jk\Delta f_{\text{offset}}\right) + n(k),
\end{aligned}
\tag{2.50}
$$

where $A_0 = R\sqrt{P_s P_{\text{LO}}}$ and Δf_{offset} is the frequency offset between transmitter and LO lasers. The additive noise is a complex AWGN, originating from LO-ASE beat noise and shot noises, with a circular Gaussian distribution of mean zero and variance N_0. The carrier phase recovery consists of FOE and PE, because both two factors would lead to carrier phase rotation.

Frequency mismatch always exists between two different lasers due to fabrication, heating and aging, thus resulting in a continuous phase increment in each symbol [46]. Conventional FOE algorithms can be implemented in either time-domain or frequency domain. Here, we only review a commonly-used frequency-domain FOE [43], described as

$$
\Delta \hat{f}_{\text{offset}} = \frac{R_s}{M} \arg \max_{|\bar{f}_0| < 1/2} \left| \frac{1}{N} \sum_{i=0}^{N-1} [r(i)]^M \exp\left(-j2\pi \bar{f}_0 i\right) \right|,
\tag{2.51}
$$

Here, $\bar{f}_0$ is the normalized frequency offset and N is the number of symbols used in the FOE. The M-th power operation is to eliminate phase modulation in the received M-PSK signals [31]. Note that the algorithm is also applicable in M-QAM formats but suffers from inaccuracy issue due to their phase modulations are not exactly $n\pi/M$ [44]. The accuracy of this FOE in M-QAM formats can be improved by incorporating more symbols in Eq.(2.51).

In addition to frequency offset, laser phase noise is a Wiener process associated with the 3-dB laser linewidth and system symbol rate, as shown in Figure 2.17a, and thus keeps rotating the phase of transmitted signals. In digital coherent receivers, DSP-based PE allows for a free-running LO laser rather than using an optical PLL for carrier

(a)

(b)
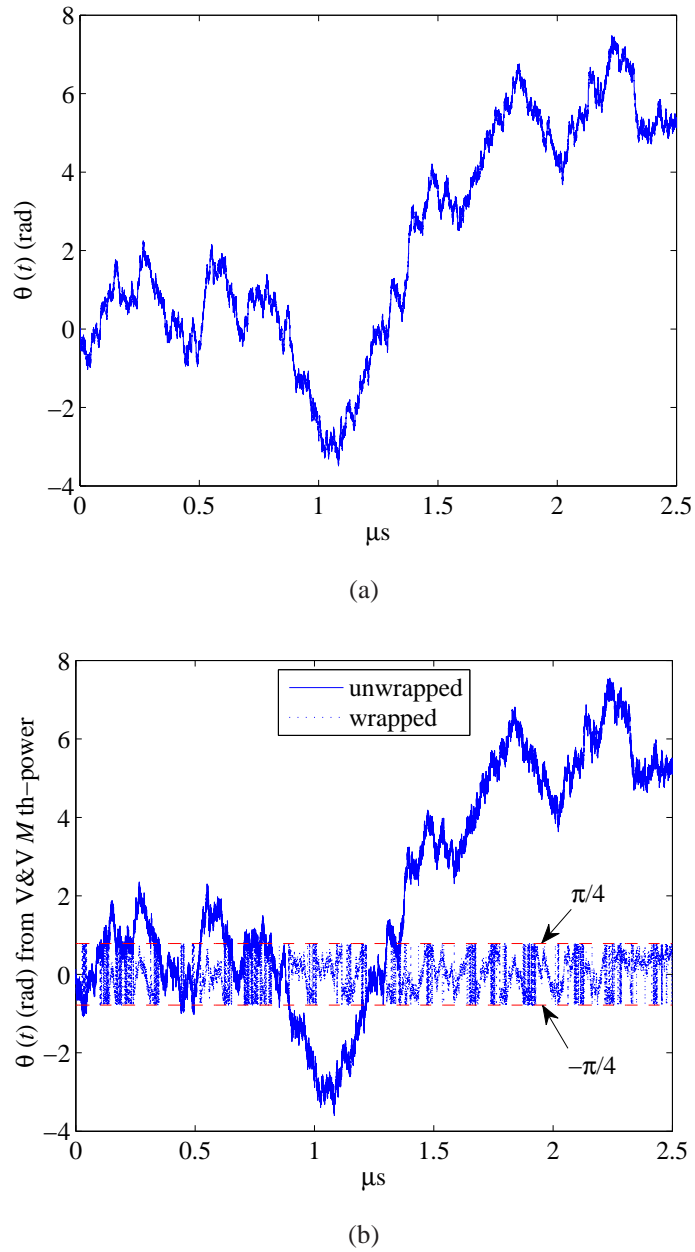
Figure 2.17: (a) Simulated Wiener laser phase noise at $2\Delta\nu T_s = 1\times10^{-4}$; (b) The esti-mated phase reference $\hat{\theta}(k)$ using the V&V Mth-power algorithm in a QPSK system

phase tracking [4]. To retrieve the phase modulation $\phi_s(k)$, the phase reference $\hat{\theta}(k)$ needs to be first estimated through PE algorithms. A conventional nonlinear V&V Mth-power scheme is proposed in [31] to raise the received M-PSK signals to the

2.4 DSP Algorithms in Coherent Receivers

Mth-power for estimating the carrier phase.

In the V&V Mth-power algorithm, a nonlinear transformation is performed to extract the unknown carrier phase from received M-PSK signals:

$$r'(k) \equiv \mathfrak{F}\left[|r(k)|\right] \exp\left[jM\angle r(k)\right], \qquad (2.52)$$

where $\mathfrak{F}[\cdot]$ denotes nonlinear transformation of the received signal amplitude. It has been observed that $\mathfrak{F}[|r(k)|] = 1$ or $|r(k)|^2$ achieves a better performance than the scenario of $\mathfrak{F}[|r(k)|] = |r(k)|^4$ for QPSK format. The nonlinear function of $\mathfrak{F}[|r(k)|] = |r(k)|^M$ is used throughout this thesis since we observe that there is trivial difference in terms of laser linewidth tolerance between them at high SNRs.

Considering the scenario of M-PSK signals, the received signal $r(k)$ is raised to the Mth-power to remove the phase modulation $\phi_s(k)$, since

$$r_M(k) = A_0^M \exp\left[j(M\phi_s(k) + M\theta(k))\right] = A_0^M \exp\left[jM\theta(k)\right]. \qquad (2.53)$$

Here, for illustration purpose,the amplitude noise term $n(k)$ is neglected and the nonlinear transformation is assumed to be $\mathfrak{F}[|r(k)|] = |r(k)|^M$. Averaging over $2L+1$ samples from $(k-L)T_s$ to $(k+L)T_s$, the estimated phase reference at $t = kT$ in the V&V Mth-power algorithm can be given as

$$\hat{\theta}(k) = \arctan\left[\sum_{l=(k-L)T_s}^{(k+L)T_s} r^M(l)\right]/M \qquad (2.54)$$

The argument of the received signal $r(k)$ is then subtracted by the estimated reference $\hat{\theta}(k)$ before symbol decoding. In the decoder, the phase difference $\arg(r(k)) - \hat{\theta}(k)$ is discriminated among M states in the M-PSK constellation. It can be observed that the phase reference $\hat{\theta}(k)$ is obtained from dividing the sum of $2L+1$-symbol long M-th power of $r(k)$. It indicates that the phase reference $\hat{\theta}(k)$ will lie between $-\pi/M$ and π/M due to the $\arctan(\cdot)$ function. Therefore, whenever the trajectory of the laser phase exceeds the range between $\pm\pi/M$, an instantaneous phase jump of $\pm 2\pi/M$

would occur, leading to irreducible symbol errors [123], as illustrated in Figure 2.17b. In order to avoid such phase jump, the phase estimate $\hat{\theta}(k)$ must be compared to the previous $\hat{\theta}(k-1)$. If the difference in the phase estimate between two adjacent symbols is within $(-\pi/M, \pi/M)$, the estimated phase reference $\hat{\theta}(k)$ is valid, otherwise it must be added or subtracted by multiple $2\pi/M$. The phase unwrapping process for M-PSK using V&V Mth-power algorithm is summarized as [22]:

$$\hat{\theta}(k) = \begin{cases} \hat{\theta}(k) + \frac{2\pi}{M}, & \text{if } \hat{\theta}(k) - \hat{\theta}(k-1) > \frac{\pi}{M} \\ \hat{\theta}(k) - \frac{2\pi}{M}, & \text{if } \hat{\theta}(k) - \hat{\theta}(k-1) < -\frac{\pi}{M} \\ \hat{\theta}(k), & \text{else} \end{cases} \quad (2.55)$$

Figure 2.17b demonstrates that the phase unwrapping is capable of accurately tracking the actual time-varying phase noise. However, phase unwrapping can cause cycle slips, which is a highly nonlinear phenomenon [4, 72]. Although the V&V Mth-power scheme is implemented in the constant-amplitude M-PSK formats, it has been extended to 16-QAM format by using a subgroup of symbols with phase modulation $\pi/4 + n\pi/2(n = 0, 1, 2, 3)$ [32].

2.4.4 Symbol Detector

After recovering the carrier phase, the symbol decision is made before carrying out differential decoding. It is important that the data is differentially encoded to avoid phase ambiguity or cycle slips [4, 33, 72]. Supposing that $a(k)$ is the information symbol, differentially encoded M-PSK symbol $c(k)$ can be obtained from $c(k) = c(k-1)a(k)$. On the other hand, differentially encoded QAM symbols are generated by means of the quadrant differential encoding rule [124]. The information symbol $a(k)$ can be represented by $a(k) = \rho(k)b(k)$, where $\rho(k)$ belongs to the first quadrant and $b(n) \in \pm 1, \pm j$ [125]. Differentially encoded QAM symbol $c(k)$ is given by $c(k) = \rho(k)d(k)$,

where $d(k) = b(k)d(k-1)$. Note that $E[|a(k)|^2]=1$ for both M-PSK/QAM formats. At the receiver side, to recover the original information symbol $a(k)$, differential decoding is imperative to reverse the differential encoding process. The information symbol $a(k)$ can be differentially decoded as $a(k) = c(k)c^*(k-1)$in M-PSK formats, and $a(k) = \rho(k)d(k)/d(k-1)$ in M-QAM systems. Error detector is performed to compare the recovered data with original transmitted data by counting bit errors. If forward error control (FEC) coding is used to achieve highly reliable communication in an optical transmission system [126], a uncoded BER level between $10^{-3} \sim 10^{-5}$ is a reasonable reference to compare the performances of coherent receivers.

2.5   Conclusion

In this chapter, we have studied the system principles of coherent optical communication, including transmitters, fiber transmission and coherent receivers. Advanced modulation formats are desirable to approach the Shannon limit in conjunction with coherent detection and PDM techniques. The development of MZM and IQ modulators facilitates optical communication systems to employ higher-order modulation formats rather than only using OOK/ASK to further improve SE.

During the propagation of the transmitted signal along optical fibers, not only do linear impairments, such as CD, PMD and fiber loss, deteriorate the signal, but the nonlinear SPM and XPM can become detrimental to the performance of long-haul optical communication systems at high input powers. Since coherent detection can retain the information of optical electric field, DSP algorithms are widely applied to compensate for the fiber impairments instead of using bulky and complicated optical components. The application of DSP CMOS chips render design versatility and universal structure to digital coherent receivers, which would be highly expected portable receivers in the near future. Finally, a series of DSP algorithms are reviewed in detail for the sake of

2.5 Conclusion

understanding the basic principles of digital coherent receivers.

Chapter 3

Decision-Aided Maximum Likelihood Phase Estimation

To allow a free-running LO laser for beating with the received optical signals in coherent receives, DSP-based PEs are widely introduced to recover the carrier phase [127]. Meanwhile, PE algorithms can relax the laser linewidth requirements compared to PLLs-based coherent receivers which are sensitive to the time delay in the feedback loop [4]. Furthermore, the laser linewidth poses a critical limit to the performance of high-order modulation formats, which are nowadays often employed in coherent optical communications for the sake of improving SE and system capacity [29, 120, 128]. Therefore, a large linewidth tolerance is desirable for a good PE. Although various PE algorithms, such as Mth-power [31], block Mth-power [22, 32], NLMS PE [129], Wiener filter-based PE [4, 38], Kalman filter-based PE [39], have been proposed to recover the carrier phase, they require either nonlinear computations (Mth-power operation and phase unwrapping) or the statistics of the system noises (phase noise and additive noise). Nonlinear computations may increase the power consumption of the complementary metal-oxide-semiconductor (CMOS) processors in coherent receivers while the statistics of such information cannot always be available, especially in re-

configurable optical switching systems.

In this chapter, a computationally-efficient, DA ML PE will be derived to achieve a comparable or even better performance than the Mth-power scheme. The performance of DA ML PE will be analytically conducted in different modulation formats. Further, a real-time structure of the DA ML receiver will be proposed to estimate carrier phase reference online in optical coherent detection.

3.1 The Principle of DA ML Phase Estimation

In order to simplify notations in the following formula, Eq.(2.50) can be rewritten as [37]

$$r(k) = m(k) \exp\left(j\theta(k)\right) + n(k) \tag{3.1}$$

where $m(k) = A_0(k) \exp\left(j\phi_s(k)\right)$ is the signal constellation point C_i. It is assumed that frequency offset and ISI are fully compensated through DSP algorithms as reviewed in Section 2.4 and all symbols are equiprobable. In a carrier-suppressed modulation format, it is possible to arrange constellation points such that $C_i = -C_{-i}$ [130]. Transmitted symbols are then chosen from ensemble of these constellation points, expressed as

$$m(k) \in \left\{C_i : C_i = -C_{i+M/2}, i = 1, 2, \cdots, M/2\right\}, \tag{3.2}$$

where M, a power of 2, is the total number of signal points in the constellation. The SNR per symbol (γ_s) is re-written as

$$\gamma_s = \frac{E\left[|m(k)|^2\right]}{N_0}. \tag{3.3}$$

The ML phase reference estimate $\hat{\theta}$ at time $t = kT$ is computed over the immediate past L received signals, i.e., $r(l)$, $k - L \leq l \leq k - 1$, where L is called memory length. The likelihood function $\Lambda(\theta, k)$ is given by joint PDF $p(r(k - L), \cdots, r(k - 1)|\theta)$. It is assumed that symbol timing is known, and θ is taken to be time invariant at

least over an interval longer than LT. Note that if there is no ISI affecting the samples at $t = kT$, $r(l)$ and $r(j)$ are independent because of the independence of $n(l)$ and $n(j)$ when $i \neq j$. The removal of ISI can be realized through channel equalization before carrying out PE algorithms. The PDF $p(r(l)|\theta)$ is obtained by writing it as the sum over all M of $p(r(l)|\theta, m(l) = C_i)P(m(l) = C_i)$. The conditional PDF of the received signal $r(k)$ is given by [131]

$$p(r(k)|\,\theta, m(k) = C_i) = \frac{1}{\pi N_0} \exp\left(-\frac{|r(k) - C_i e^{j\,\theta}|^2}{N_0}\right). \qquad (3.4)$$

Through this arrangement $C_i = -C_{i+M/2}$ of the constellation points, the log-likelihood function $L(\theta, k) = \ln \Lambda(\theta, k)$ can now be expressed as

$$L(\theta, k) = \sum_{l=k-L}^{k-1} \ln \left[\sum_{i=1}^{M/2} \exp\left(-S_i\right) \cosh q_i(l, \theta)\right] + c \qquad (3.5)$$

Here, $S_i = |C_i|^2/N_0$, $q_i(l, \theta) = (2/N_0)\mathbf{Re}[r(l)C_i^* e^{-j\theta}]$, and c is a constant independent of θ (see Appendix B for details). It is easy to derive the following equation for the ML estimate $\hat{\theta}$ from the likelihood equation $\partial L(\theta, k)/\partial\,\theta = 0$ at $\theta = \hat{\theta}(k)$ [130],

$$\begin{aligned}
\cos\hat{\theta}(k) &\sum_{l=k-L}^{k-1} \cdot \frac{\sum_{i=1}^{M/2} \exp\left(-S_i\right)\sinh q_i(l, \hat{\theta}(k))\mathbf{Im}\left[r(l)C_i^*\right]}{\sum_{i=1}^{M/2} \exp\left(-S_i\right)\cosh q_i(l, \hat{\theta}(k))} \\
&= \sin\hat{\theta}(k) \sum_{l=k-L}^{k-1} \cdot \frac{\sum_{i=1}^{M/2} \exp\left(-S_i\right)\sinh q_i(l, \hat{\theta}(k))\mathbf{Re}\left[r(l)C_i^*\right]}{\sum_{i=1}^{M/2} \exp\left(-S_i\right)\cosh q_i(l, \hat{\theta}(k))}
\end{aligned} \qquad (3.6)$$

Equation (3.6) is highly nonlinear and is hard to get an explicit solution for $\hat{\theta}(k)$. Supposing that we have obtained the ML phase estimate $\hat{\theta}(k)$, the receiver uses it in the detection of the kth data symbol $m(k)$ as if it were the true value θ. The receiver declares $m(k) = C_d$ according to decision statistics [62]

$$q_d(k) = \max_i \mathbf{Re}\left[r(k)C_i^* \exp\left(-j\hat{\theta}(k)\right)\right], i = \pm 1, \cdots, \pm M/2, \qquad (3.7)$$

in M-PSK formats, whereas in M-QAM formats the decision statistic q_d becomes [37]

$$q_d(k) = \max_i \mathbf{Re}\left[r(k)C_i^* \exp\left(-j\hat{\theta}(k)\right) - \frac{1}{2}|C_i|^2\right], i = \pm 1, \cdots, \pm M/2. \qquad (3.8)$$

3.1 The Principle of DA ML Phase Estimation

Here, the receiver decision is denoted by $\hat{m}(k)$ at time $t = kT$. An explicit result for $\hat{\theta}(k)$ allows the decision statistics (3.7) to be easily computed, thus allowing a completely digital receiver implementation.

Approximations are imperative to make Eq.(3.6) implementable. High SNR is the case of most interest in practice. When considering the high SNR limit, decision-feedback provides a nearly optimum implementation of the ML estimator (3.6) [132]. For instance, in BPSK format ($M = 2$), Eq.(3.6) reduces to

$$\hat{\theta}(k) = \arctan\left[\frac{\sum_{l=k-L}^{k-1} \mathbf{Im}[r(l)\hat{m}^*(l)]}{\sum_{l=k-L}^{k-1} \mathbf{Re}[r(l)\hat{m}^*(l)]}\right]. \tag{3.9}$$

where C_i^* at time $t = lT$ in Eq.(3.6) is replaced by the receiver decision $\hat{m}(l)$ in decision-aided PE. In the case of signal constellations with $M > 2$, for each fixed l in Eq.(3.6), each of the summations over L is dominated in magnitude by the term whose index corresponds to the decision. For instance, $|\sum_{i=1}^{M/2} \exp(-S_i) \sinh q_i (l, \hat{\theta}(k)) \mathbf{Im}[r(l)C_i^*]| \approx |\exp(-S_d) \sinh q_d(l, \hat{\theta}(k)) \mathbf{Im}[r(l)C_d^*]|$ if the decision is either $\hat{m}(l|k) = C_d$ or $\hat{m}(l|k) = -C_d$. Thus, Eq.(3.6) reduces again to Eq.(3.9) when $\tanh(\cdot)$ is replaced by $\text{sgn}(\cdot)$ and $\hat{m}(l|k)$ by $\hat{m}(l)$. The decision-feedback approximation is more accurate for BPSK ($M = 2$) than M-PSK with $M > 2$. For $M > 2$, the accuracy of the approximation depends on the "discriminability" between the signal points C_L's, which improves with higher values of SNR S_L's [130].

In order to totally eliminate nonlinear operation in this PE algorithm, we introduce a complex vector V as $V \triangleq [\mathbf{Re}\ V, \mathbf{Im}\ V] = [\hat{c}(k), \hat{s}(k)]$ such that Eq.(3.9) can be rewritten as [94, 130]

$$\hat{\theta}(k) = \arctan\left[\frac{\hat{s}(k)}{\hat{c}(k)}\right] \tag{3.10}$$

Therefore, from the comparison between Eq.(3.9) and Eq.(3.10), we can define

$$V(k) \triangleq U^{-1}(k) \sum_{l=k-L}^{k-1} r(l)\hat{m}^*(l), \tag{3.11}$$

where $U(k) = \sum_{l=k-L}^{k-1} |m(l)|^2$, a factor used to normalize the reference phasor (RP)

49

3.1 The Principle of DA ML Phase Estimation

$V(k)$. The quantities $[\hat{c}(k), \hat{s}(k)]$, respectively, are the in-phase and quadrature outputs of the receiver. In the form Eq.(3.10), it is clear that the π-radian ambiguity of $\arctan(\cdot)$ function is resolved by the signs of $\hat{c}(k)$ and $\hat{s}(k)$ if an initial training data is sent to start up the receiver. Figure 3.1 summarizes the receiver structure and the signal processing operations involved in DA ML algorithm for either M-PSK or M-QAM signals.

RP $V(k)$ Estimator

$r(k)$

$$V(k) = U(k) \sum_{l=k-L}^{k-1} r(l)\hat{m}^*(l)$$

$V(k)$ $\qquad\qquad\qquad\qquad$ $\hat{m}(k)$

Symbol Detector

Set $\quad \hat{m}(k) = C_d \quad$ if

$$C_d = \begin{cases} \max_i \mathbf{Re}[r(k)V^*(k)C_i^*], & \text{for } M\text{-ary PSK} \\ \max_i \mathbf{Re}[r(k)V^*(k)C_i^* - |C_i|^2/2], & \text{for } M\text{-QAM} \end{cases}$$
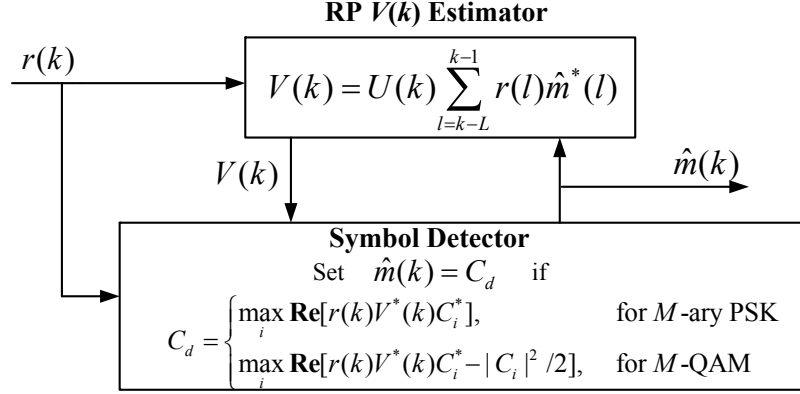
Figure 3.1: The structure of the DA ML receiver for M-ary PSK/QAM systems.

Compared to the Mth-power scheme, the contributions of introducing the RP $V(k)$ in a phasor form are significant when implementing PE algorithms. In the Mth-power scheme, nonlinear operations are necessary, such as $\arctan(\cdot)$ operation and the Mth power operation of the received signals [33]. Besides, due to the fact that $\arctan(\cdot)$ operation limits estimated phase reference $\hat{\theta}(k)$ to an interval $(-\pi/M, \pi/M)$, as already illustrated in Section 2.4.3, an additional step of phase unwrapping hence needs to be performed [4] to avoid cycle slips. In contrast, the DA ML PE only requires linear computations and there is no phase ambiguity. Note that the feedback in DA ML would be subjected to decision-feedback errors, thus possibly causing error propagation. The impact of decision-feedback errors is found to be negligible when BER is less than 10^{-3}. A detailed comparison between DA ML and the Mth-power schemes will be given in Section 3.2.

3.2 The Performance of DA ML in M-PSK and QAM

To investigate the performance of the derived DA ML PE, we carry out MC simulations in different M-PSK and 16-QAM formats. Differential encoding is employed at the transmitter for preventing error propagation due to past decision errors.
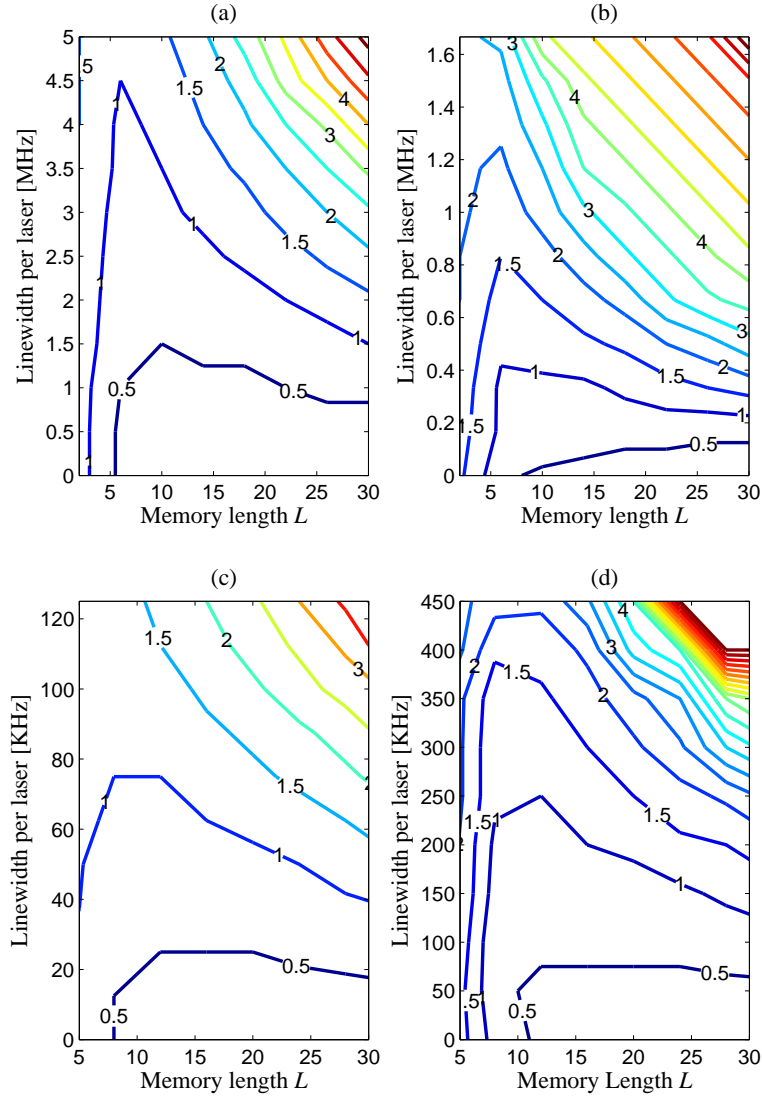
Figure 3.2: Receiver sensitivity penalties at BER=10^{-4} versus the linewidth per laser and memory length L for different modulation formats: (a) QPSK, (b) 8PSK, (c) 16P-SK and (d) 16-QAM. Bit rate @ 40Gb/s.

As a performance measure, the contour plots of the receiver sensitivity penalty

at BER=10^{-4} are depicted in Figure 3.2 for different modulation formats with respect to the linewidth per laser and different memory lengths L. The reference is the SNR per bit (γ_b) in ideal coherent detection of differentially encoded M-PSK/16QAM with perfect carrier PE. The reference γ_b at BER=10^{-4} are 8.8 dB, 12.1 dB, 16.6 dB and 12.5 dB for QPSK, 8PSK, 16PSK and 16-QAM, respectively. Note that a certain minimum memory length L is necessary to approach the ideal coherent performance even in the absence of phase noise. For a small laser linewidth, which means the phase noise varies slowly, a longer memory length L is preferred to average out the additive noise. As the laser linewidth increases, the optimal memory length L becomes smaller, because rapidly changing phase noise becomes less correlated over the long memory length, compared to the case of short L. Generally, a trade-off exists between averaging over the additive noise and over the phase noise. If a 1-dB γ_b penalty is acceptable, the optimal memory lengths are found to be $L = 6$ for PSK formats and L = 12 for 16QAM. From the above results, the faster the phase noise varies the shorter the memory length L to be used. This is called block length effect (BLE) phenomenon that exists in both DA ML and the Mth-power schemes [32]. One alternative is to use an adaptive PE technique to achieve the optimal performance by adaptively adjusting the phase estimator input gain [133], which will be addressed in Chapter 4.

It is also obviously observed that requirements for laser linewidth is more rigid if there are more phase states in the modulation format. The receiver sensitivity penalty at BER=10^{-4} is shown in Figure 3.3 versus the ratio of the linewidth per laser to the symbol rate ($\Delta\nu T_s$), for the four modulation formats with optimal memory length L. The limitations of the ratio of linewidth per laser to symbol rate leading to 1 dB-penalty are found to be 2.2×10^{-4}, 3.0×10^{-5}, 8.8×10^{-6} and 2.5×10^{-5} for QPSK, 8PSK, 16PSK and 16-QAM, respectively, which correspond to 4.4 MHz, 400 kHz, 88 kHz and 250 kHz linewidth for each laser at 40 Gbit/s. In order to show the performance improvement, we compare our results with the one using the V&V Mth-power scheme.
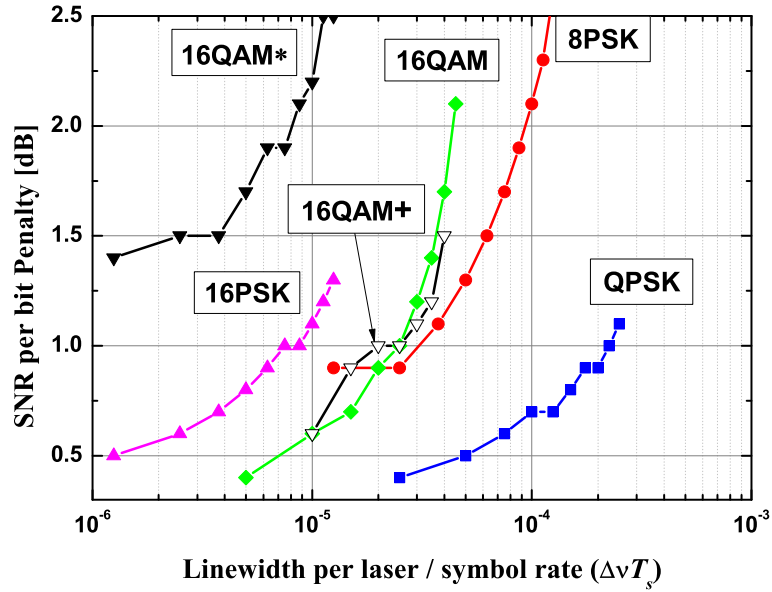
Figure 3.3: Receiver sensitivity penalty at BER=10^{-4} versus the ratio of the linewidth per laser to symbol rate $\Delta\nu T_s$ using DA ML. Note: 16QAM* denotes the penalty from the V&V Mth power scheme (optimal block length: 64); 16QAM+ represents the penalty from the modified Mth power scheme (optimal block length: 16).

The optimal length for M-PSK and 16-QAM formats are found to be 8 and 64 in [32]. As indicated in Table 3.1, the performance of DA ML consistently outperforms the conventional V&V Mth-power scheme. This performance improvement of DA ML over the V&V Mth-power becomes more evident as the number of phase states increases. The advantage of using DA ML is to avoid the contribution from the higher powers (≥ 2) of additive noise present in the Mth power scheme [94,130].

In addition, as indicated in Figure 3.2, the optimal memory length L for 16QAM is comparable to the PSK formats. This improvement is due to the fact that all symbols are used to estimate the RP $V(k)$ in DA ML method, whereas only subgroup symbols (Class I) with modulation phases of $\pi/4 + n\pi/2$ (n=0, 1, 2, 3) can be utilized to determine phase estimate in the V&V Mth-power scheme [32], as depicted in Figure 3.4. As a result, the requirement on laser linewidth for the V&V Mth-power becomes worse
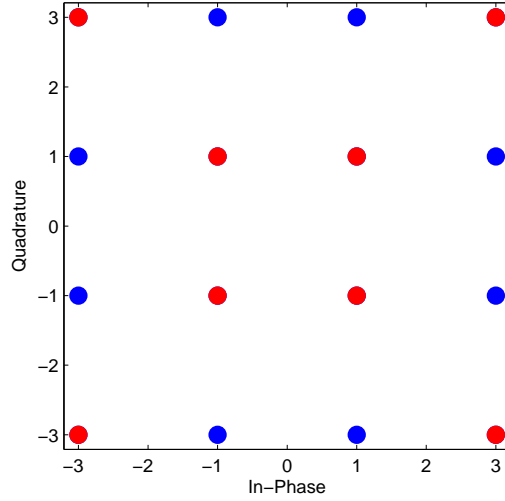
Figure 3.4: The Class I symbols (marked as red) in 16-QAM format are used in Mth-power PEs. The blue symbols located on the middle ring belongs to Class II.

Table 3.1: The maximum tolerance to linewidth per laser leading to a 1-dB γ_b penalty at BER=10^{-4} for DA ML and V&V Mth-power PEs (MC Simulations)

	QPSK	8PSK	16PSK	16-QAM
DA ML	4.4 MHz	400 kHz	88 kHz	250 kHz
V&V Mth-power	4.4 MHz	330 kHz	50 kHz	< 12.5 kHz

than our DA ML algorithm (Figure 3.3). The tolerance to laser linewidth using DA M-L is increased by more than 10 times in 16-QAM format compared to the one using V&V Mth-power, as listed in Table 3.1. Additionally, a modified Mth-power scheme has been proposed in [34, 35], wherein all symbols are exploited to estimate carrier phase in QAM systems. Figure 3.3 shows that its performance is comparable to DA ML algorithm (the optimum block length of the modified Mth-power is found to be 16 in our simulation). However, the complexity of the modified Mth-power scheme is relatively large through acquiring the phase information in those symbols belonging

to the middle ring of QAM constellation (Class II): i) Mth-power and phase unwrapping are necessary. It can be observed that the modified scheme needs at least two phase unwrapping processes, which are involved in the PE using Class I symbols and Class II symbols; ii) the symbol decision and rotation are required in the sub-group partition stage, which determines the sign of phase rotation $\pm \arctan 1/3$ when estimating carrier phase [33]. Instead, DA ML PE is computationally linear and simple to be implemented without involving any nonlinear computation and sub-step. Table 3.2 summarizes the complexity comparison between DA ML and V&V Mth-power PEs. In short, the proposed DA ML algorithm not only has a better performance than the conventional Mth-power PEs, but requires less computational complexity.

Table 3.2: The Complexity Comparison between DA ML and V&V Mth-power PEs with Memory Length L

	Complex Multiplication	Complex Addition	Phase Unwrapping	Argument Operation
DA ML	$2L^1$	$L-1$	No	No
V&V Mth-power	$(M-1)L$	$L-1$	Yes	Yes

3.3 Performance Evaluation of DA ML

3.3.1 Analysis of Phase Error

Phase error variance is usually used to evaluate the performance of PE algorithms. Not only can it provide an insightful investigation on the impacts of different parameters on the performance of PE algorithms, but it also indicates how to achieve the ultimate

---

[1]$2L$ denotes that $2L$ numbers of multiplications.

optimal performance [4, 134]. The phase error is defined as [134]

$$\Delta\theta(k) = \theta(k) - \hat{\theta}(k), \tag{3.12}$$

where $\hat{\theta}(k)$ refers to the phase reference obtained from PE algorithms. In DA ML PE, the phase reference $\hat{\theta}(k)$ is the argument of RP $V(k)$, i.e., $\hat{\theta}(k) = \arg V(k)$. For simplicity, we first assume that the receiver decision is correct ($\hat{m(l)} = m(l)$) in a constant-amplitude M-PSK system. Replacing $r(l)$ in Eq.(3.11) with Eq.(3.1), the phase reference $\hat{\theta}(k)$ is yielded into

$$\hat{\theta}(k) = \arg\left[\sum_{l=k-L}^{k-1} \exp\left(j\theta(l)\right) + n'(l)\right], \tag{3.13}$$

where $n'(l) \equiv n(l)/|m(l)|^2$ having a statistics with zero mean and variance $\sigma_{n'}^2 = 1/\gamma_s$. Note that real factors in Eq.(3.13) have been removed since they do not affect the argument value of a complex.

By pulling out the phase $\theta(k-1)$ out of the bracket in Eq.(3.13), the phase reference $\hat{\theta}(k)$ can be re-written as [37]:

$$\begin{aligned}
\hat{\theta}(k) &= \theta(k-1) + \arg\left[\sum_{l=k-L}^{k-1} \exp\left(j\theta(l) - j\theta(k-1)\right) + n'(l)\exp\left(j\theta(k-1)\right)\right] \\
&= \theta(k-1) + \arg\left[\sum_{l=k-L}^{k-1} \exp\left(-j\sum_{i=l+1}^{k-1}\nu(i)\right) + \right. \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \left. n'(l)\exp\left(j\theta(k-1)\right)\right].
\end{aligned} \tag{3.14}$$

Here, the discrete Wiener laser phase noise $\theta(k)$ is modeled by [4]

$$\theta(k) = \sum_{m=-\infty}^{k} \nu(m), \tag{3.15}$$

where $\nu(m)$'s are independent, identically distributed, Gaussian random variables with mean zero and variance σ_p^2. Since $|\sum_{i=l+1}^{k-1}\nu(i)| \ll 1$, the approximation $1 + jx \approx$

$\exp(jx)$ is applied, thus giving

$$
\begin{aligned}
\sum_{l=k-L}^{k-1} \exp\left(-j\sum_{i=l+1}^{k-1}\nu(i)\right) &\approx \sum_{l=k-L}^{k-1}\left(1-j\sum_{i=l+1}^{k-1}\nu(i)\right) \\
&= L - j\sum_{p=1}^{L-1}(L-p)\nu(k-p) \\
&\approx L\exp\left(-j\frac{1}{L}\sum_{p=1}^{L-1}(L-p)\nu(k-p)\right). \quad (3.16)
\end{aligned}
$$

Substituting the approximation Eq.(3.16) into Eq.(3.14), the phase reference $\hat{\theta}(k)$ is now simplified into

$$
\hat{\theta}(k) \approx \theta(k-1) - \frac{1}{L}\sum_{p=1}^{L-1}(L-p)\nu(k-p) + \arg\left[1+\frac{1}{L}n''(l)\right], \quad (3.17)
$$

where

$$
n''(l) \equiv n'(l)\cdot\exp\left(j\theta(k-1)-j\frac{1}{L}\sum_{p=1}^{L-1}(L-p)\nu(k-p)\right). \quad (3.18)
$$

Note that $n''(l)$ is statistically identical to $n'(l)$ because of circularly symmetric characteristics of $n'(l)$. The remaining argument in Eq.(3.16) can be further re-written into the form as follows by using this approximation that $\arg[1+x]\approx \mathbf{Im}[x]$ for $|x|\ll 1$:

$$
\hat{\theta}(k) \approx \theta(k-1) - \frac{1}{L}\sum_{p=1}^{L-1}(L-p)\nu(k-p) + \frac{1}{L}\mathbf{Im}[n''(l)]. \quad (3.19)
$$

Therefore, the phase error is derived as [37]

$$
\Delta\theta(k) \approx \frac{1}{L}\sum_{p=0}^{L-1}(L-p)\nu(k-p) - \frac{1}{L}\mathbf{Im}[n''(l)]. \quad (3.20)
$$

Due to the independence between additive noise and phase noise, Eq.(3.20) leads to a Gaussian approximation for $\Delta\theta(k)$ with moments [37]:

$$
E\left[\Delta\theta(k)\right] = 0, \quad (3.21)
$$

$$
E\left[\Delta\theta(k)^2\right] = \frac{2L^2+3L+1}{6L}\sigma_p^2 + \frac{1}{2L}\sigma_{n'}^2. \quad (3.22)
$$
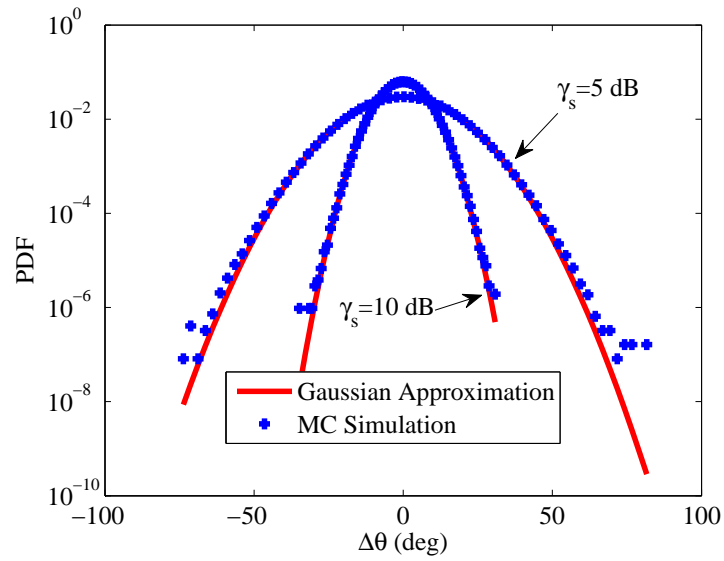
Figure 3.5: The PDF of the simulated phase error variance using DA ML PE in M-PSK formats(L=5).
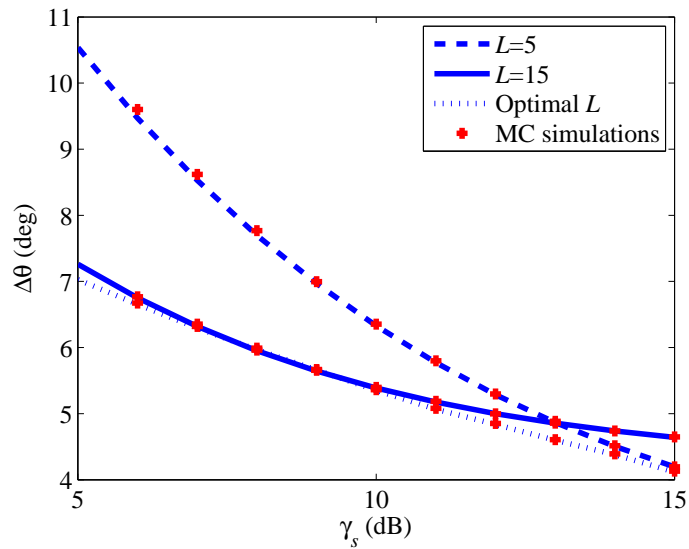
Figure 3.6: The simulated STD of the phase error variance using DA ML PE in M-PSK/QAM formats with(L=5) and optimal L.

MC simulations are performed to verify the derived phase error in M-PSK formats. Figure 3.5 proves that the mean of the phase error is zero. Besides, the PDF of the phase error in DA ML can be well approximated by a Gaussian distribution up to the probability of 10^{-6} even though γ_s is only 5 dB. The small difference at the tails is due to the limited samples (10^6) in MC simulations. Figure 3.6 shows that the standard deviation (STD) of the analytical phase error variance, where ideal decision feedback is assumed, i.e., $\hat{m}(k) = m(k)$. Excellent agreement can be observed between MC simulations and the analytical approximation at different memory length L. This indicates that we can rely on the analytical phase error to evaluate the performance of different M-PSK systems when using DA ML PE.

It is seen from Eq.(3.22) that the contribution of phase noise to the variance of phase error $\Delta\theta$ is enhanced as memory length L increases, while the impact of additive noise $\sigma_{n'}^2$ is reduced. The expression for the variance of phase error $\Delta\theta$ explains the tradeoff that is discussed previously based on those MC simulation results. As illustrated in Figure 3.6, the STD of phase error at L=15 is much smaller than the one using DA ML with L=5 at low and moderate SNR levels, whereas short memory length becomes better to track phase noise at high SNR in which phase noise is dominant over additive noise, as explained in Section 3.2.

From Eq.(3.22), an optimal L can be obtained to give the minimum variance of phase error $\Delta\theta$ and, hence, the smallest BER:

$$L_{\text{opt}} = \left\lfloor \frac{1}{4}\sqrt{1 + 24\frac{\sigma_{n'}^2}{\sigma_p^2}} - \frac{3}{4} \right\rfloor, \tag{3.23}$$

where $\lfloor x \rfloor$ denotes the largest integer less than or equal to x. It is seen from Figure 3.6 that DA ML with optimal memory length L_{opt} always delivers a minimal phase error variance, thus leading to the optimal BER performance.

In addition, though the analysis here is developed from constant-modulus M-PSK formats, these conclusions are also applicable in M-QAM formats due to the average

over L symbols [135]. MC simulations in Figure 3.7 verify that the analytical phase error variance Eq.(3.22) is sufficient to characterize the performance of DA ML in any order of QAM formats.
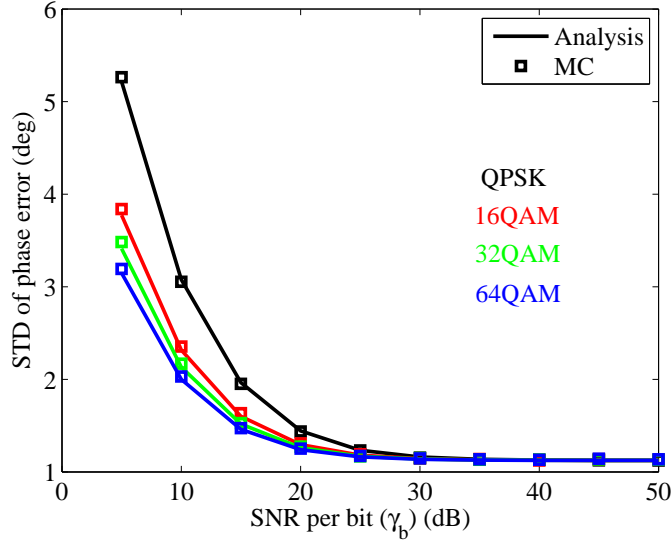
Figure 3.7: The STD of the phase error of DA ML PE ($L = 10$) obtained from analysis (line) and MC simulations (marker '+'). $\sigma_p^2 = 1 \times 10^{-4}$ rad^2.

3.3.2 Impact of Decision Errors on DA ML

In the above simulations, the decision feedback is assumed to be error free. In fact, this is not valid any more since there exist bit errors when receivers normally operate at a low or moderate SNR to reduce power consumption and avoid fiber nonlinearity. As illustrated in Figure 3.8, if RP $V(k)$ deviates from the correct one by $\pi/2$ rad or more, due to incorrect decision feedback from previous decision errors, those subsequent received signals $r(k)$'s would be constantly rotated by $\pi/2$ or more, resulting in cycle slips and error propagation in coherent detection of QPSK signals [72]. Although feedforward Mth-power schemes are not affected by decision errors, their phase ambiguity still causes catastrophic error propagation [4]. Such problem can be avoided

by using differential encoding (DE) [124] as the phase difference between two consecutive symbols is independent of a constant phase offset of RP $V(k)$, though there is a DE-induced performance penalty. For example, the BER of a DE-BPSK increases by a factor of 2 at high SNR compared to a non-DE BPSK format [62]. This performance penalty can be even larger if fiber nonlinearity is present. Alternatively, periodic pilots can be inserted to data sequence for the sake of correcting the possibly deviated RP and thus phase offset [72]. Pilot is very efficient and does not take up much extra power when pilots and data length are carefully selected.
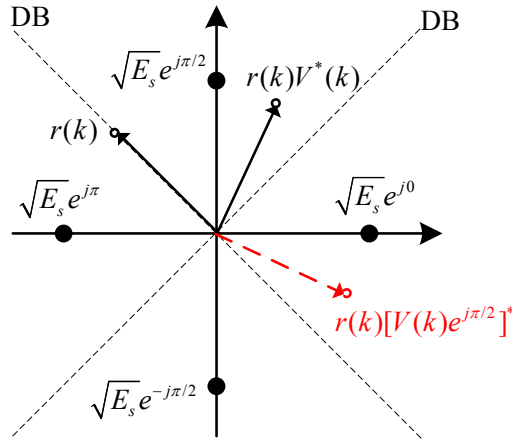
Figure 3.8: Effect of a RP $V(k)$ error on the symbol decision in a QPSK constellation. DB: decision boundary; $r(k)$: received signal; E_s: symbol energy.

The impact of decision error is further studied using MC simulations, as depicted in Figure 3.9. Frequently occurring decision errors[1] would keep rotating RP $V(k)$'s from their correct ones, thus causing phase jumping and large STD of phase error in DA ML. If neither DE nor pilots are employed in the system, error propagation will occur in non-differential-encoded (NDE) systems [4]. As explained above, because DE-QPSK is independent of constant phase offset of RP $V(k)$, DE is able to avert

---

[1]In real optical system, the acceptable BER is about 4.5×10^{-3} with the current commercial FEC decoder whose output error less than 1×10^{-15} [136].
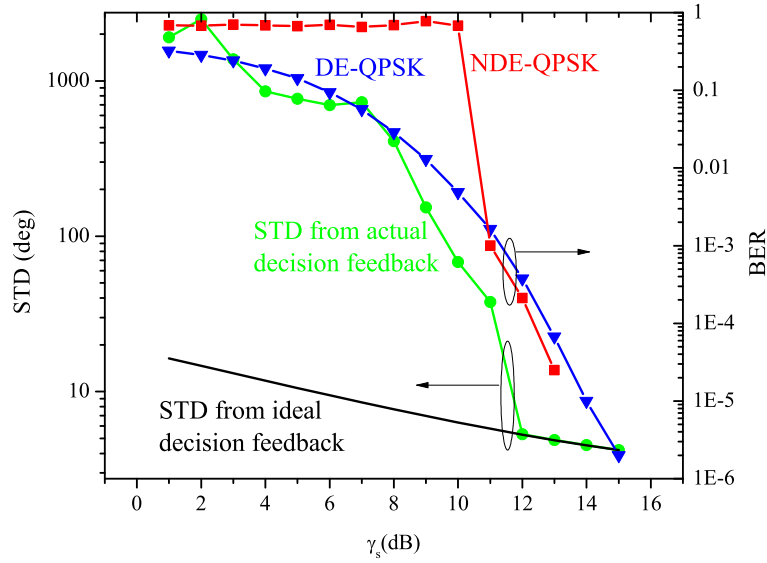
Figure 3.9: The impact of decision error on the STD of phase error using DA ML in a QPSK constellation (left axis); BER performance comparison between DE and NDE-QPSK (right axis).

catastrophic error propagation (see Figure 3.9). It is worth noting that the BER of DE-QPSK is slightly worse than NDE-QPSK when BER is lower than 10^{-3}, where infrequently decision errors are not critical any more. DE is usually applied in coherent M-PSK/QAM simulations and experiments to prevent error propagation due to its simplicity [24]. In our thesis it is still called coherent M-PSK/QAM instead of DE M-PSK/QAM for short notation unless specified clearly in certain context.

3.3.3 Analytical Performance of DA ML in Non-DE M-PSK/QAM

In view of the good agreement between MC simulations and analytical results, we can now resort to analytical tools which are more efficient and powerful compared to extensive MC simulations to evaluate the performance of DA ML.

In the presence of a phase error $\Delta\theta$, the BER $P_b(e)$ of M-PSK/QAM can be

numerically evaluated as [62]

$$P_b(e) = \int_{-\pi}^{\pi} P_b(e|\Delta\theta)p(\Delta\theta)d\Delta\theta, \tag{3.24}$$

where $P_b(e|\Delta\theta)$ stands for the BER of different M-PSK/QAM signals conditioned on a fixed value of the phase error $\Delta\theta$, and $p(\Delta\theta)$ is the PDF of the phase error $\Delta\theta$, as given in Eqs.(3.21) and (3.22) for DA ML PE.

The well-known expression $P_b(e|\Delta\theta)$ of BPSK, QPSK and 8PSK can be found in many literatures [137–139], which are listed as:

$$P_b(e|\Delta\theta) = \frac{1}{2}\text{erfc}\left(\sqrt{\gamma_b}\cos\Delta\theta\right) \tag{3.25}$$

for BPSK, where $\text{erfc}(x) = 2/\sqrt{\pi} \cdot \int_x^{+\infty} \exp\left(-t^2\right)dt$;

$$P_b(e|\Delta\theta) = \frac{1}{4}\left[\text{erfc}\left(\sqrt{\gamma_s}\sin\left(\frac{\pi}{4}-\Delta\theta\right)\right) + \text{erfc}\left(\sqrt{\gamma_s}\sin\left(\frac{\pi}{4}+\Delta\theta\right)\right)\right] \tag{3.26}$$

for QPSK; and

$$\begin{aligned} P_b(e|\Delta\theta) &+ P_b(e|-\Delta\theta) \\ &= \frac{1}{3}\text{erfc}\left[\sqrt{\gamma_s}\sin\left(\frac{\pi}{8}-\Delta\theta\right)\right] + \frac{1}{3}\text{erfc}\left[\sqrt{\gamma_s}\sin\left(\frac{\pi}{8}+\Delta\theta\right)\right] \\ &+ \frac{1}{6}\text{erfc}\left[\sqrt{\gamma_s}\cos\left(\frac{\pi}{8}-\Delta\theta\right)\right] \cdot \text{erfc}\left[\sqrt{\gamma_s}\sin\left(\Delta\theta-\frac{\pi}{8}\right)\right] \\ &+ \frac{1}{6}\text{erfc}\left[\sqrt{\gamma_s}\cos\left(\frac{\pi}{8}+\Delta\theta\right)\right] \cdot \text{erfc}\left[-\sqrt{\gamma_s}\sin\left(\Delta\theta+\frac{\pi}{8}\right)\right] \end{aligned} \tag{3.27}$$

for 8PSK. Note that the integration range in Eq.(3.24) for 8PSK is reduced to $[0,\pi)$ since a pair of $P_b(e|\Delta\theta)$ and $P_b(e|-\Delta\theta)$ are used in Eq.(3.27).

Recent researches in optical communication have focused on other high-order modulation formats, such as 16-PSK and 16-QAM [59, 120], capable of further increasing SE. It is worthy of investigating the tolerance of such systems to phase noise. Here, we have derived the $P_b(e|\Delta\theta)$ expressions for 16-PSK and 16-QAM in the presence of the phase error $\Delta\theta$ [135], as described elaborately in Appendix C.
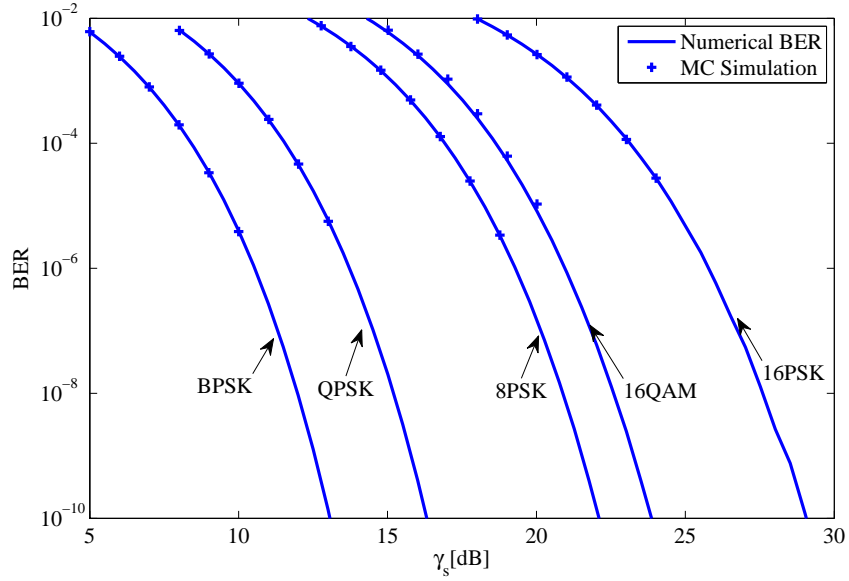
Figure 3.10: The BER performance comparison of DA ML PE between M-ary PSK and 16QAM formats through numerical integration and MC simulations with the optimum memory length L. $\sigma_p^2 = 1 \times 10^{-4}$ rad^2.

To justify the validity of the analytical BER evaluation, the BER performances of DA ML in different modulation formats are examined using MC simulations when $\sigma_p^2 = 1 \times 10^{-4}$ rad^2, as demonstrated in Figure 3.10. Note that the optimum memory length L of DA ML is selected by minimizing the phase error variance in either M-PSK (M=2, 4, 8, and 16) or 16-QAM format. MC simulation results conform well to the numerical integration of Eq.(3.24) when the statistics of the phase error $\Delta\theta$ in DA ML PE is applied. The excellent agreement indicates that the analytical performance can be utilized to investigate the performance of DA ML in different modulation formats, such as laser linewidth tolerance. Note that the analysis here does not take into account the error propagation in the non-DE format. It is because we are interested in the BER level less than 10^{-3}, where the error propagation is not of big concern for decision-feedback algorithms.

3.4   Implementation of DA ML Algorithm

The implementation of DA ML algorithm can be realized using either serial [94] or parallel [140] structure. Serial structure is simple but requires high-speed processing CMOS units, which are difficult to be fabricated based on the states-of-arts technologies. As a result, parallel processing is preferred in high-speed receivers to avoid the bottleneck of CMOS processor speed, and thus reduce the cost of receivers [141]. In this section, we will simplify the serial processing for BPSK and QPSK formats. The performance DA ML algorithm is then investigated in a parallel structure.

3.4.1   Simplified Serial Structure

As for BPSK, the decision statistic Eq.(3.7) can be reduced to

$$q_d(k) = \text{sgn}\left(\mathbf{Re}\left[r(k)V^*(k)\right]\right), \tag{3.28}$$

due to $C_i^* \in \pm A_0$, which can be easily implemented using a slicer with the real part of $r(k)V^*(k)$ as input. Here $\text{sgn}(\cdot)$ is a sign function. In the case of QPSK format, the decision statistic is equivalently expressed as

$$
\begin{aligned}
q_d(k) &= \arg\max \mathbf{Re}\left[2\exp\left(j\pi/4\right)r(k)V^*(k)C_i^*\exp\left(-j\pi/4\right)\right] \\
&= \arg\max \mathbf{Re}\left[\mu(k)m'^*(k)\right]. \tag{3.29}
\end{aligned}
$$

Here, $\mu(k) \equiv 2\exp\left(j\pi/4\right)r(k)V^*(k)$, and $m'(k) \equiv C_i^*\exp\left(-j\pi/4\right) \in (\pm 1 \pm j)A_0/\sqrt{2}$ are rotated symbols in a $45°$-tilted QPSK constellation. The scale factor of 2 in Eq.(3.29) does not affect the decision rule, because the decision statistics take index i only when the real part is the maximum. The scaling will guarantee a simple implementation of DA ML algorithm, as illustrated in the following steps. It can be observed that the decision in Eq.(3.29) over the tilted constellation reduces to decisions on the

signs of the real and imaginary parts of $\mu(k)$ [142]:

$$\hat{m}'(k) = \frac{\text{sgn}(\mathbf{Re}[\mu(k)]) + j \cdot \text{sgn}(\mathbf{Im}[\mu(k)])}{\sqrt{2}}. \qquad (3.30)$$

The decision on variables $\mathbf{Re}[\mu(k)]$ and $\mathbf{Im}[\mu(k)]$ can be implemented using two sign-decision devices or two slicers. Although $m'(k)$ is a complex rotated signal which is different from the previous $m(k)$, the gray bit mapping of the QPSK signals is still preserved after symbol decision:

$$\sqrt{2}\hat{m}'(k) = \pm 1 \pm j \in \sqrt{2}\exp{(j\pi/4)}\{1, j, -1, j\} \longleftrightarrow \{00, 10, 11, 01\}. \qquad (3.31)$$

As for M-PSK/QAM ($M > 4$), the decision rule cannot be easily simplified into taking the sign of the decision statistics. An alternative is to use high-speed ADCs to conduct the decision statistics in Eq.(3.7) or Eq.(3.8) in DSP processors, which also make the real-time implementation of M-PSK/QAM($M > 4$) more difficult compared to BPSK and QPSK formats.
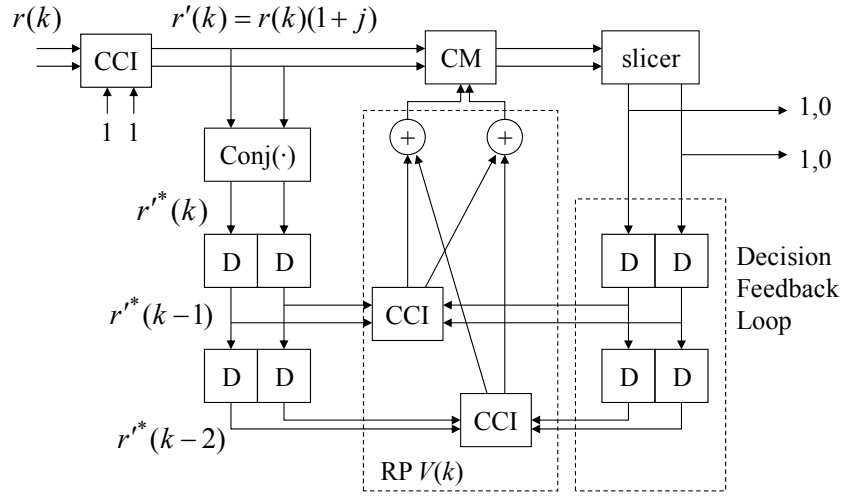
Figure 3.11: A simplified serial structure of DA ML algorithm for QPSK format (L=2). D: Time delay. Note that the input complex signal is formed by its real and imaginary parts.

Here, we use the analog devices to implement the simplified serial DA ML structure for BPSK/QPSK modulated signals. By considering the QPSK simplified decision

statistics in Eq.(3.29), decision variable $\mu(k)$ should be fed into the slicer to decode the received signals, so a four-quadrant complex multiplexer (CM) is required to perform multiplications of the form [94]:

$$
\mu(k) = \left(\mathbf{Re}\left[2\exp\left(j\pi/4\right)r(k)\right] + j \cdot \mathbf{Im}\left[2\exp\left(j\pi/4\right)r(k)\right]\right)
$$
$$
\cdot\left(\mathbf{Re}\left[V^*(k)\right] + j \cdot \mathbf{Im}\left[V^*(k)\right]\right), \tag{3.32}
$$

where the term $2\exp\left(j\pi/4\right)r(k) = (1+j)\cdot(\mathbf{Re}[r(k)]+j\mathbf{Im}[r(k)])$ can be realized by using complex-controlled inverters (CCIs), whose control bit inputs are fixed to $\{1,1\}$. The transfer function of a CCI is given by

$$
\begin{aligned}
y_{\text{out}}^{\text{re}}(t) &= (2\bar{b}_0 - 1)x_{\text{in}}^{\text{re}} - (2\bar{b}_1 - 1)x_{\text{in}}^{\text{im}}, \\
y_{\text{out}}^{\text{im}}(t) &= (2\bar{b}_1 - 1)x_{\text{in}}^{\text{re}} - (2\bar{b}_0 - 1)x_{\text{in}}^{\text{im}},
\end{aligned} \tag{3.33}
$$

where $\{x_{\text{in}}^{\text{re}}, x_{\text{in}}^{\text{im}}\}$ are the real and imaginary parts of input signals to the CCI and $\{y_{\text{out}}^{\text{re}}, y_{\text{out}}^{\text{im}}\}$ are the counterparts of its output signals, $\{\bar{b}_0, \bar{b}_1\}$ are the control bits (binary inverted) to realize the multiplication between $\pm 1 \pm j$ and $x_{\text{in}}^{\text{re}} + jx_{\text{in}}^{\text{im}}$ [142]. Since CCI is a specialized complex multiplier with $\pm 1 \pm j$, its implementation is considerably simpler than that of a general four-quadrant CM. Furthermore, the decision-feedback RP $V(k)$ is generated by the decision-feedback loop, which consists of $2L$ time delay (T) of the slicer output. Since

$$
\begin{aligned}
\sqrt{2}V^*(k) &= \left[\sum_{l=k-L}^{k-1} \sqrt{2}r(l)\hat{m}'^*(l)\right]^* = \sum_{l=k-L}^{k-1} \sqrt{2}r^*(l)\hat{m}'(l)\exp\left(-j\pi/4\right) \\
&= \sum_{l=k-L}^{k-1} [r'(l)]^*\hat{m}'(l), \tag{3.34}
\end{aligned}
$$

phase reference $V^*(k)$ can be obtained by two adders and L CCIs. $V^*(k)$ times the rotated received signal $r'(k) = \sqrt{2}r(k)\exp\left(j\pi/4\right)$ to generate the decision variable $\mu(k)$, which is the input signal to the slicer for decoding the received signal $r(k)$. From the previous explanations, the scale factor of 2 is split into two factors of $\sqrt{2}$

so that $r'(k)$ can be generated through a simple CCI not a four-quadrant CM, in order to reduce the complexity. As for BPSK signals, the difference is that the received $r(k)$ does not require a 45° rotation and the slicer can decode the received signals based on the real part of $r(k)V^*(k)$. A simplified structure of DA ML algorithms for QPSK format with memory length $L = 2$ is plotted in Figure 3.11. Note that this QPSK DA ML serial structure with memory length L comprises $2L + 1$ CCIs, $4L$ time delays, two adders, one CM and one slicer. A similar structure for differentially encoded M-PSK signals is shown in [142] in conjunction with several pairs of optical delay interferometers (ODIs) with respective time delays and phase shifters. However, such a receiver structure with a window length D needs D-1 pairs of complex ODIs, and it requires data be differentially encoded. In our receiver structure, there is no requirement for data to be differentially encoded, thus delivering a better performance.

3.4.2 Parallel Structure

The current speed of a CMOS processor is limited to less than a few GHz. As a result, a parallel structure is necessary in 10 Gbit/s and higher optical systems to split ultra-high-speed data stream into several lower-speed branches, in which CMOS processors can implement DSP algorithms online [143]. Current commercial FPGA chips are able to deliver up to 2.8 Tb/s serial bandwidth [144]. It is desirable to have no feedback in the PE scheme for real-time operation [141]. In this subsection, we will propose a parallel DA ML PE for real-time processing, and will examine its performance in different M-PSK formats.

The serial DA ML structure suggested in Section 3.4.1 can only operate at symbol rate, thus making it difficult to implement by using current commercial CMOS processors. Here, we use the same RP from previous frame to adapt our DA ML to a parallel implementation, as shown in Figure 3.12. Since the whole DA ML PE algorithm can-

3.4 Implementation of DA ML Algorithm
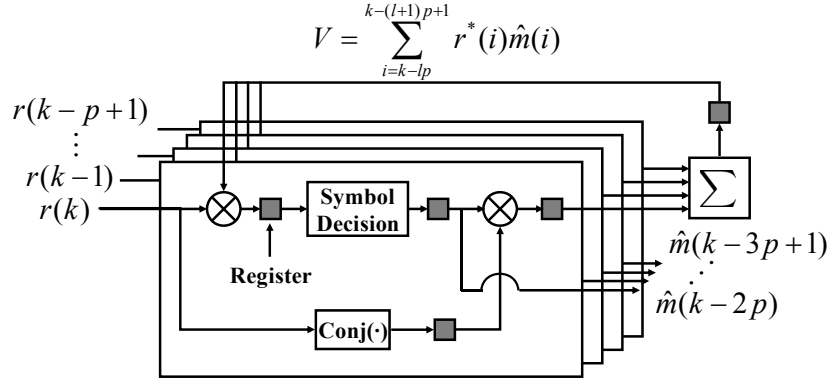
Figure 3.12: A schematic of a parallelized DA ML receiver with pipeline stages $l=4$. Conj$(\cdot)$ stands for a complex conjugation, and $\sum$ denotes summation of all its inputs.

not be realized completely within one clock cycle, those intermediate results need to be stored in registers (represented by black square boxes in Figure 3.12). This process is called pipelining [141]. The high-speed data stream is de-multiplexed into p branches, whose signals are rotated using the same RP from the previous p symbols [140], i.e., $V = \sum_{i=k-lp}^{k-(l+1)p+1} r(i)\hat{m}^*(i)$, when taking into account l pipeline stages. Note that the received signals $r(k)$'s and V are first conjugated to save conjugation operations in the parallel structure, as illustrated in Figure 3.12. A pipelining stage of $l = 4$ is found to be sufficient for a parallel DA ML PE [140].

It can be observed that this common RP V is used for the entire p symbols in each frame. The performance will be degraded due to this delayed phase tracking, compared to the symbol-rate one in Section 3.4.1. The phase error $\Delta\theta$ of the parallel DA ML is analytically investigated on the center symbol, i.e., $r(k-(p-1)/2)$, which thus accounts for an average performance over the entire p branches. The phase error in the parallel DA ML is defined as

$$\Delta\theta \equiv \theta(k-(p-1)/2) - \arg V. \qquad (3.35)$$

Following the same approach as described in Section 3.3, the argument of the RP V in

the parallel DA ML can be written as

$$\hat{\theta} \equiv \arg V = \arg \left[\sum_{i=k-lp}^{k-(l+1)p+1} \{\exp\left[j\theta(i)\right] + n'(i)\} \right], \qquad (3.36)$$

which is further simplified into [140]

$$\arg V \approx \theta(k - lp) - \frac{1}{p}\sum_{i=1}^{p-1}(p-i)\nu(k - lp - i + 1) + \frac{1}{p}\sum_{i=k-lp}^{k-(l+1)p+1} \mathrm{Im}[n'(i)]. \quad (3.37)$$

Consequently, the phase error variance of the parallel DA ML is easily derived as

$$E\left[\Delta\theta^2\right] = \frac{(6l-1)p^2+1}{6p}\sigma_p^2 + \frac{1}{2p}\sigma_{n'}^2. \qquad (3.38)$$

To justify our analysis results, both the mean and the variance of the phase error are measured using MC simulations in M-PSK formats. In our simulations, the variance and mean of phase errors are obtained by averaging over those entire p branches. The assumption on the zero-mean phase error is found to be still reasonable even at a large parallelism degree (p =20) and the number of pipelines (l =20), and linewidth as high as 20 MHz per laser at 10 GSymbols/s for illustration. The parallelism degree (p =10) and the number of pipelines (l =4) are selected at 10 GSymbols/s. As shown in Figure 3.13, the average phase error variance from MC simulations conforms well to the analytical results based on the center symbol despite an increasing laser linewidth.

The feedback delay of the parallel DA ML can be defined as $\Delta T = (l-1) \cdot p + (p+1)/2$ [141]. It can be observed from Figure 3.14 that there exists an optimum delay ΔT for each pair of σ_p^2 and $\sigma_{n'}^2$, and the number of pipeline stages (l), thus minimizing the variance in Eq.(3.35). The results point out that a certain degree of parallelism is actually beneficial to the performance of a parallel DA ML. Besides, it is found from Eq.(3.35) and Figure 3.14 that the number of pipeline stages (l) affects the phase error more seriously than parallelism degree (p), because the phase error is linearly proportional to the number of pipeline stages.
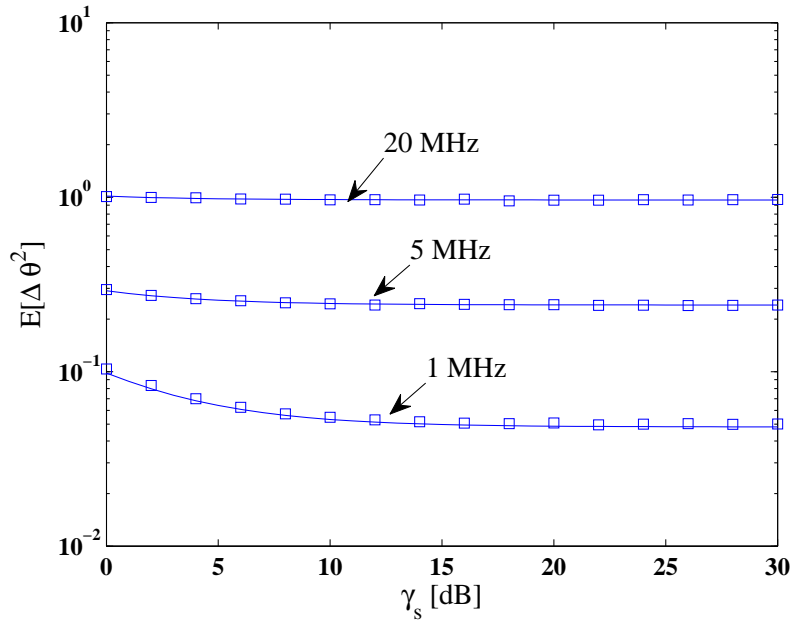
Figure 3.13: Comparison of the phase error variance from analysis (line) and MC simulations (marker '□') in a 10 GSymbols/s M-PSK format at different linewidths of each laser ($p = 10, l = 4$).
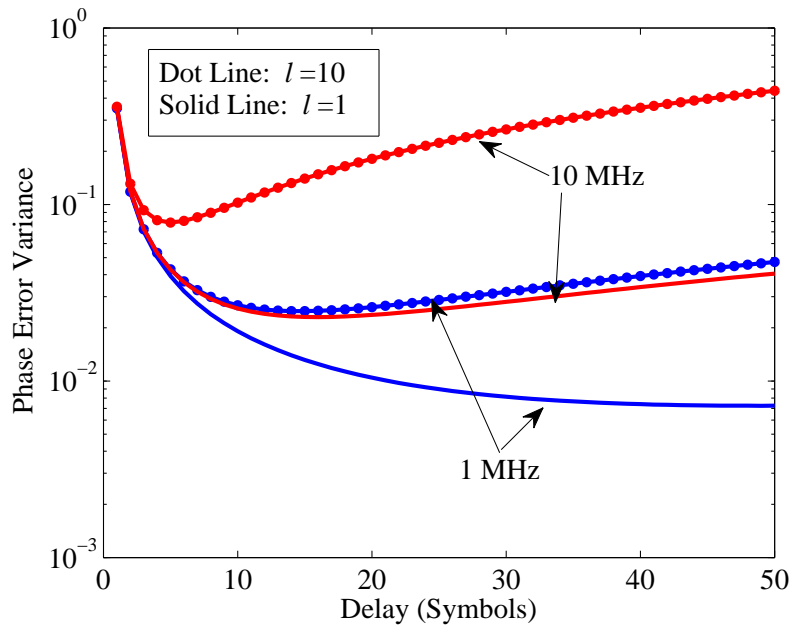
Figure 3.14: The phase error variance as a function of feedback delay ΔT in a parallel DA ML ($p = 10$) with different laser linewidths and pipeline stages at 10 GSymbols/s and $\gamma_s = 10$ dB.
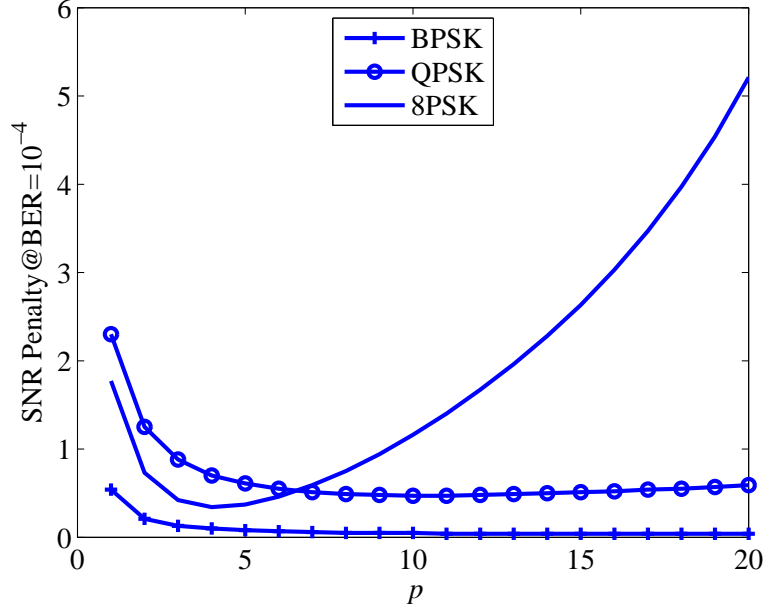
Figure 3.15: The SNR per bit (γ_b) penalty at BER=10^{-4} versus different parallelism degrees (p) in a 40 Gbit/s M-PSK system with 100 KHz linewidth of each ECL (l =5).

In addition, we performed BER analysis in 40 Gbit/s BPSK, QPSK and 8PSK systems by using the approach illustrated in Section 3.3. The linewidth of each laser is assumed to be 100 KHz for emulating external cavity lasers (ECL), and l=5 as suggested in [141] for realistic reason. The γ_b penalty at BER=10^{-4} as a function of parallelism degree (p) is plotted in Figure 3.15. The SNR per bit (γ_b) penalty is with respect to ideal coherent detection of each M-PSK with perfect PE, given by

$$\gamma_b = \frac{1}{\log_2 M} \left[\frac{\mathrm{erfc}^{-1}(\log_2 M \cdot \mathrm{BER})}{\sin{(\pi/M)}} \right]^2 \tag{3.39}$$

for M-PSK with $M > 4$. Note that the well-known BER of BPSK and QPSK in the absence of phase error is given by $\gamma_b = [\mathrm{erfc}^{-1}(2 \cdot \mathrm{BER})]^2$. It is seen from Figure 3.15 that the performance degradations of a parallel DA ML in BPSK and QPSK are limited within 0.2 dB and 0.5 dB, respectively. However, the penalty rises sharply in 8PSK and even high-order formats. One reason is that higher-order M-PSK is more sensitive to phase noise, thus imposing a rigid feedback delay of DA ML algorithm. In

general, the parallel DA ML is feasible for coherent real-time receiver provided that laser linewidth, the degree of parallelism (p) and the number of pipelining stages (l) are carefully selected.

3.5   Filtering Effect

Although significant researches have been carried out to analyze PEs in coherent detection [4,37,127], they did not take into account filtering effect as suggested in [145]. In communication theory, the matched filter would offer high SNR at the sampling point [62], thus ensuring better receiver sensitivity. In this section, the filtering process of DA ML algorithm is studied by assuming a matched filter at the receiver side. Therefore, the filtered M-PSK signal can be represented by [138]

$$r(k) = \sqrt{E_s} \exp\left[j\phi_s(k)\right] \cdot \int_{kT}^{(k+1)T} \frac{\exp\left[j\varphi(t)\right]}{T} dt + n(t), \qquad (3.40)$$

where $\varphi(t)$ is the laser phase noise that is a Wiener process with $E[\varphi(t)] = 0$ and $E[\varphi(t_1)\varphi(t_2)] = \sigma_0^2 \min(t_1, t_2)$, and $\sigma_0^2 = 2\pi(2\Delta\nu)$. Here, the matched filter is assumed to have a rectangular pulse shape.

Replacing $r(l)$ in the RP $V(k)$ given in Eq.(3.11) with Eq.(3.40), it yields into

$$V(k) = \frac{1}{L\sqrt{E_s}} \sum_{l=k-L}^{k-1} z(l) + \exp\left[-j\hat{\phi}_s(l)\right]n(l), \qquad (3.41)$$

where

$$z(l) \equiv \sqrt{E_s} \exp\left[j\varphi(lT)\right] \cdot \int_{lT}^{(l+1)T} \frac{\exp\left[j\varphi(t) - j\varphi(lT)\right]}{T} dt. \qquad (3.42)$$

To facilitate the analysis here, the phase-noise exponent commutation theory in [145] is applied to give the subsequent approximation:

$$
\begin{aligned}
z(l) &= \sqrt{E_s} \exp\left[j\theta(l)\right] \cdot \exp\left[j \int_{lT}^{(l+1)T} \frac{\varphi(t) - \theta(l)}{T} dt\right] \\
&= \sqrt{E_s} \exp\left[j\theta(k)\right] \cdot \exp\left[-j \sum_{i=l+1}^{k} \nu(i) + j \int_{lT}^{(l+1)T} \frac{\varphi(t) - \theta(l)}{T} dt\right]. \quad (3.43)
\end{aligned}
$$

3.5 Filtering Effect

Here, $\theta(l) \equiv \varphi(lT) = \theta(l-1) + \nu(l)$, and $\theta(l) = \theta(k) - \sum_{i=l+1}^{k} \nu(i)$. The additive noise terms are not taken into account temporally for simplicity. In fact, the additive noise term can be simply added into the final equation as described in [37], and will be included later. Using the similar approach as illustrated in Section 3.3, the argument of the RP $V(k)$ in the presence of the matched filter can be expressed as

$$
\begin{aligned}
\hat{\theta}(k) \equiv \arg V(k) &= \theta(k) + \arg \left[\sum_{l=k-L}^{k-1} \exp \left[-j \sum_{i=l+1}^{k} \nu(i) + j \int_{lT}^{(l+1)T} \frac{\varphi(t) - \theta(l)}{T} dt \right] \right] \\
&\approx \theta(k) + \arg \left[\sum_{l=k-L}^{k-1} \left[1 - j \sum_{i=l+1}^{k} \nu(i) + j \int_{lT}^{(l+1)T} \frac{\varphi(t) - \theta(l)}{T} dt \right] \right] \\
&\approx \theta(k) + \frac{1}{L} \sum_{l=k-L}^{k-1} \left[\int_{lT}^{(l+1)T} \frac{\varphi(t) - \theta(l)}{T} dt - \sum_{i=l+1}^{k} \nu(i) \right].
\end{aligned}
\tag{3.44}
$$

Thus, the statistic of the phase error variance becomes

$$
E\left[(\Delta\theta(k))^2 \right] = E \left[\frac{1}{L^2} \left(\sum_{l=k-L}^{k-1} \left[\int_{lT}^{(l+1)T} \frac{\varphi(t) - \theta(l)}{T} dt - \sum_{i=l+1}^{k} \nu(i) \right] \right)^2 \right].
\tag{3.45}
$$

The details to derive the final expression of the phase error variance can be found in Appendix D, where the phase error variance is shown as

$$
E\left[(\Delta\theta(k))^2 \right] = \frac{L}{3}\sigma_p^2 + \frac{1}{2L\gamma_s}.
\tag{3.46}
$$

Recalling that the phase error variance without filtering is $\frac{2L^2+3L+1}{6L}\sigma_p^2 + \frac{1}{2L}\sigma_{n'}^2$, as given in Eq.(3.22). It is found that the phase error variance with filtering is smaller than the one without filtering (refer to Figure 3.16). This may be due to the fact that filtering makes the phase noise vary slower. The result shows that filtering is helpful to mitigate the impact of phase noise on coherent receivers. In addition, a longer memory length L is able to reduce the phase error difference between the DA ML with and without matched filtering but in the sacrifice of enhancing absolute phase error variance due to the BLE phenomenon.
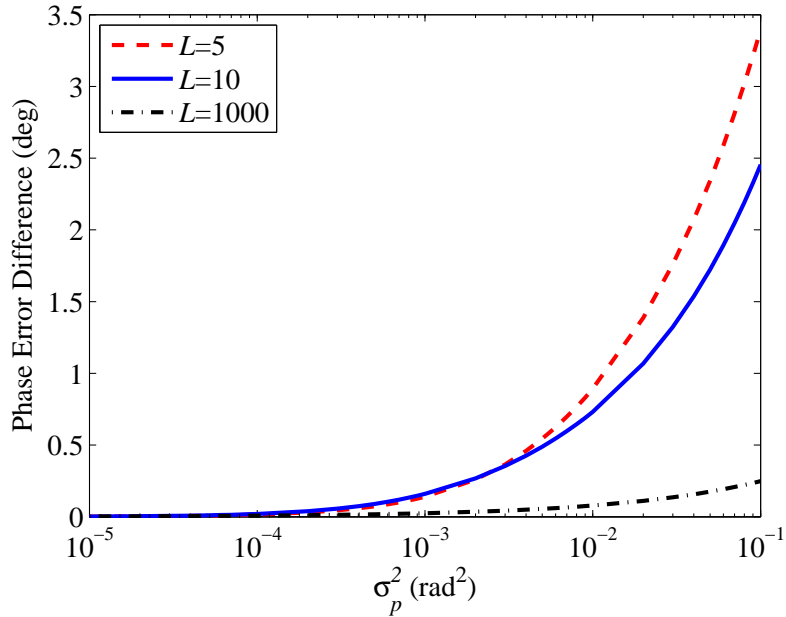
Figure 3.16: The difference between the STD of phase error with and without a matched filtering at different memory length L ($\gamma_s = 10$ dB).

3.6 Experiment

An experiment was conducted to verify the performance of the introduced DA M-L algorithm. An 8-channel 42.8-Gbit/s WDM coherent PolMux-QPSK system setup is shown in Figure 3.17 [146]. Note that 7% overhead accounts for the FEC overhead in 40 Gbit/s data transmission, resulting in total 42.8 Gbit/s per channel. A 42.8-Gbit/s PolMux-QPSK transmitter modulated eight 100 GHz-spaced tunable E-CLs with linewidth about 100 kHz. The PolMux-QPSK transmitter consisted of an IQ-modulator, an additional MZM as a pulse carver, and a polarization-multiplexing u-nit. The IQ-modulator was driven by 10.7 Gbit/s 2^{15}-1 pseudo random binary sequence data, while the MZM carved the optical pulse into 50%-RZ-QPSK signals. Note that the two polarization states of 21.4 Gbit/s data signals were de-correlated by introducing different bit delays with respect to each other in a polarization-multiplexing unit. The transmission link comprised 10 spans of 80 km standard SMF (average
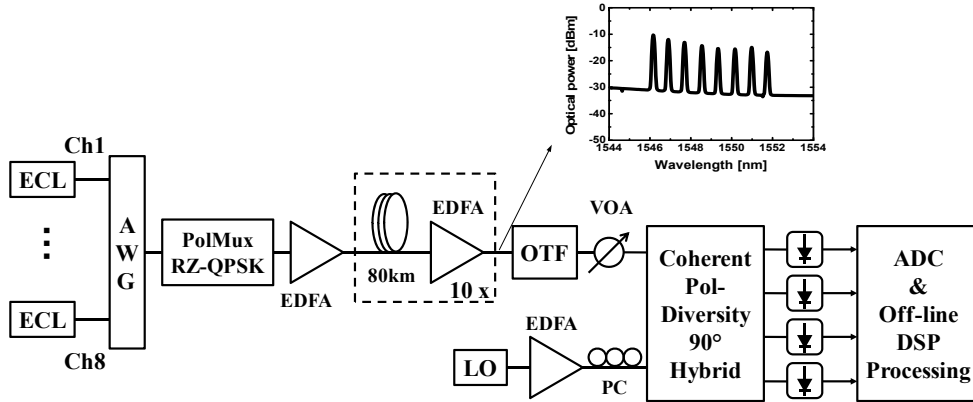
Figure 3.17: Experimental setup for an 8-channel WDM 42.8 Gbit/s coherent PolMux-QPSK system. AWG: arrayed waveguide grating; PC: polarization controller; VOA: variable optical attenuator.

span loss around 17.3 dB) and EDFA-only optical amplification. No dispersion compensation fiber was used in this experiment. The total power (after the boost EDFA) launching into the transmission fiber was kept at 9 dBm. At the receiver end, the measured 42.8 Gbit/s signal was selected by one optical tunable filter (OTF) with 0.5 nm bandwidth. After down-converting the selected optical signal by a tunable ECL LO (about 100 KHz linewidth) in a polarization-diversity coherent receiver, the outputs from four single-ended photodetectors were sampled by using a 4-channel digital storage scope with 50 GSamples/s sample rate. The captured data (1 million symbols) was then post-processed using a desktop PC. A series of standard DSP algorithms as illustrated in Section 2.4 of Chapter 2 were carried out involving CDC, polarization de-multiplexing, FOE and so on. Note that the unequal optical spectrum (inset of Figure 3.17) after transmission was due to the EDFA amplifier unequal gain profile.

Figure 3.18 (a) shows the back-to-back (B2B) BER performance of DA ML and V&V Mth-power as a function of OSNR in 0.1 nm for center channel 5 (1549.5 nm). According to the SNR penalty contour in Figure 3.2, the optimal memory length L should be greater than 5 to mitigate the impact of the major distortion from receiv-
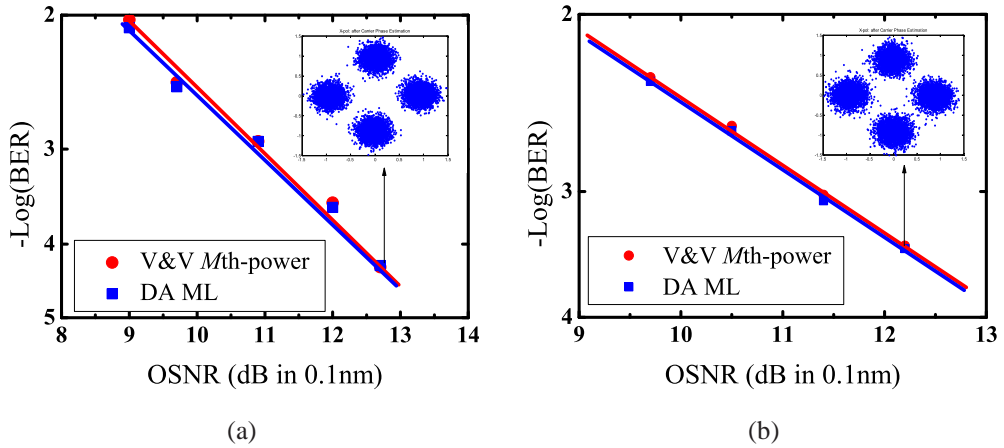
Figure 3.18: Experimental results: (a) measured B2B BER performance; (b) measured BER performance after 800 km SSMF transmission

er additive noise. In our experiment, L was selected to be 21 when considering the small laser linewidth of ECLs. As can be seen, the computationally-linear DA M-L algorithm is able to achieve a comparable performance as the conventional V&V Mth-power at much lower computational complexity. The inset of Figure 3.18 depicts the recovered constellation of PolMux-QPSK signals using DA ML. After 800km standard SMF transmission, DA ML is still effective to recover the carrier phase by observing Fig 3.18 (b). The results of other channels are omitted because of the similar performances. Although a larger laser linewidth tolerance is observed in Section 3.2 using DA ML compared with V&V Mth-power, the laser around a few MHz was not available in the lab such that the performance of laser linewidth tolerance cannot be experimentally proved. Due to the limited resources, no QAM experiments were carried out to further demonstrate the advantages of DA ML over Mth-power algorithm. In addition, the data processing was performed in offline, thus making the comparison of time consumption not meaningful from this point of view. This comparison of complexity has already proved in the Table 3.2 through rigid mathematics derivation.

In this experiment, linear phase noise is of particular interest to investigate. With

high input power, nonlinear phase noise would become dominant in the system, which will be examined in the next chapter. Moreover, the number of channels would not affect the performance of phase estimation much if their channel spacing is 50 GHz or above and the input channel power is low to keep fiber impairments within the linear region. However, when the input power goes up, the XPM would introduce additional phase noise to the testing channel. It would be desirable to apply additional DSP algorithms, like XPM compensation [122], to address those nonlinear phase noise. Those nonlinear compensation algorithms will be further elaborated in Chapter 5.

3.7 Conclusion

In this chapter, we have analytically introduced a computationally-linear DA ML PE algorithm to recover the carrier phase instead of the conventional V&V Mth-power, which suffers from nonlinear computational complexity. Its performance was investigated in both analysis and MC simulations, and was found to have a comparable or even larger laser linewidth than the V&V Mth-power algorithm. Only limited symbols does the V&V Mth-power utilize to estimate the carrier phase in M-QAM formats. As a result, the laser linewidth tolerance is much smaller than DA ML, in which all symbols are used. In addition, the memory length L plays an critical role in averaging over additive noise and over phase noise, thus affecting the performance of DA ML. This is called block length effect: the optimal memory length L becomes smaller as laser linewidth increases, because rapidly changing phase noise becomes less correlated over the long memory length compared to the case of short L, and vice versa. Subsequently, the analysis was carried out to evaluate the phase error of DA ML in different modulation formats, and pointed out that the optimal memory length of DA ML is related to the ratio of laser phase noise and additive noise. With the knowledge of the phase error variance of DA ML algorithm, we can analytically evaluate the per-

formance of DA ML in different modulation formats. Generally, the analytical BERs give the lower bound of the DA ML performance at high BER levels ($> 10^{-3}$) due to the fact that error propagation of non-DE modulation formats occurs from the frequent decision errors.

Moreover, the issues of serial and parallel implementation of DA ML algorithm were extensively studied. A simplified serial structure for BPSK and QPSK was presented to reduce the complexity. As a result of the bottleneck of electrical CMOS processors, it is desirable to implement algorithm in a parallel structure especially in high-speed coherent optical receivers. The analysis shows that parallel DA ML is quite feasible in BPSK/QPSK formats whereas in higher-order formats the laser linewidth, the degree of parallelism (p) and the number of pipelining stages (l) should be carefully selected for on-line processing. The performance of DA ML PE in the presence of a matched filter suggests that filtering is helpful to mitigate the impact of phase noise on coherent receivers, which may be attributed to the fact that filtering makes phase noise vary slower. Finally, an experiment was carried out to demonstrate DA ML algorithm in an 8-channel 42.8-Gbit/s WDM coherent PolMux-QPSK system.

Chapter 4

Adaptive Decision-Aided Phase Estimation

The proposed DA ML PE in Chapter 3 can outperform the conventional V&V Mth-power method; however, it suffers from the impact of BLE like Mth-power algorithms [36]. Although the optimal memory length L_{opt} can be determined when the statistical information of laser phase noise and additive noise becomes known to receivers [37], as illustrated in Section 3.3, these parameters may not be available especially in reconfigurable optical systems. In consequence, it is desirable that the phase estimator be adaptive to estimate the carrier phase without the aid of such information, and without BLE. We will introduce an adaptive DA PE in this chapter to meet the aforementioned motivations.

4.1 The Principle of Adaptive Decision-Aided Phase Estimation

The adaptive DA PE algorithm has been proposed for M-PSK signals in the work [133]. Here, we extend that work to include both constant-amplitude and non-constant-amplitude signals, e.g., QAM. A first-order DA adaptive filter is introduced to assign weights α and $1-\alpha$ to the previous RP $V(k)$ and the current RP, i.e., $r(k)\hat{m}^*(k)|\hat{m}(k)|^{-2}$, respectively, so that the RP for demodulating the $(k+1)$th symbol can be represented as

$$
\begin{aligned}
V(k+1) &= \alpha V(k) + (1-\alpha)\frac{r(k)\hat{m}^*(k)}{|\hat{m}(k)|^2} \\
&= \alpha V(k) + (1-\alpha)\frac{r(k)}{\hat{m}(k)}.
\end{aligned}
\tag{4.1}
$$

The filter gain α is chosen at each time k based on the observations $\{r(l), 0 \leq l \leq k\}$ to minimize the conditional risk function $R(k)$ [133], i.e., the conditional expectation of the cost, where

$$
R(k) = E\left[\sum_{l=1}^{k} |r(l) - V(l)\hat{m}(l)|^2 \,\middle|\, \{r(l)\}_{l=1}^{k}\right].
\tag{4.2}
$$

Minimizing $R(k)$ for each sample of observations $\{r(l), 0 \leq l \leq k\}$ results in minimizing the average risk $E[R(k)]$. The optimal value, $\hat{\alpha}(k)$, of α at each time k, is obtained by solving $dR(k)/d\alpha = 0$, given as [37]:

$$
\hat{\alpha}(k) = A(k)/B(k), \text{ for } k \geq 1,
\tag{4.3}
$$

where

$$
A(k) = A(k-1) + |\hat{m}(k)|^2 \cdot \left[|g(l-1)|^2 - \mathbf{Re}\left[V(l-1)g^*(l-1) - g(l)\cdot\right.\right.
$$
$$
\left.\left.[V^*(l-1) - g^*(l-1)]\right]\right];
\tag{4.4}
$$

$$
B(k) = B(k-1) + |\hat{m}(k)|^2 \cdot |V(l-1) - g(l-1)|^2.
\tag{4.5}
$$

4.1 The Principle of Adaptive Decision-Aided Phase Estimation

Here, we have $g(k) \equiv r(k)\hat{m}^{-1}(k)$. The detailed derivation of $\hat{\alpha}(k)$ is given in Appendix E. Since $|\hat{m}(k)|$ is constant in M-PSK modulation formats, it can be removed simultaneously from the numerator and denominator of Eq.(4.3), yielding the same adaptive gain derived in [133] for M-PSK signals. Of importance is that the numerator and denominator in Eq.(4.3) are formed recursively, thus requiring little memory and low computational complexity. In operating the adaptive filter Eq.(4.1), the initial condition $\hat{\alpha}(0) = 0$ is selected to give the maximum gain of one to the first received signal $r(0)\hat{m}^{-1}(0)$ (This is done by setting $A(0) = B(0) = 1$ as the initial conditions in Eq.(4.3).). The subsequent weights $\hat{\alpha}(k)$ are then adaptively adjusted based on the received signals $r(k)$ for $k \geq 1$. An initial preamble of K known symbols (training data) is used to aid in acquiring a steady-state value of $\hat{\alpha}(k)$, enabling better tracking of the phasor $\exp(j\theta(k))$.
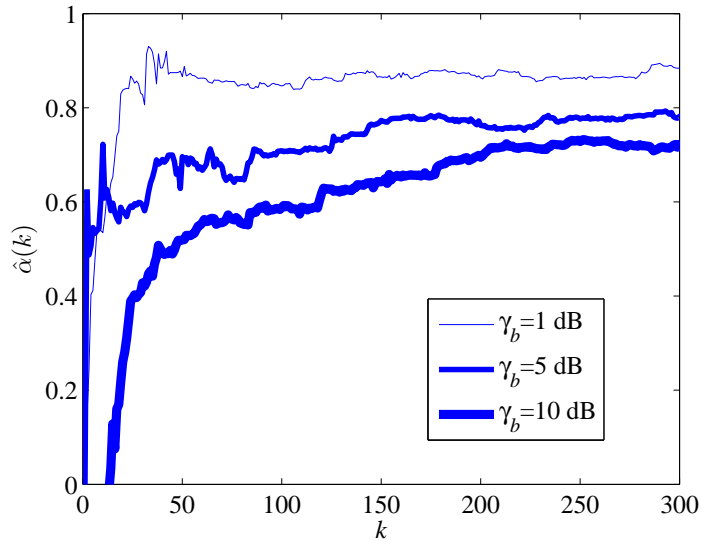
Figure 4.1: The trajectory of the filter gain $\hat{\alpha}(k)$ with ideal decision feedback $\hat{m}(k) = m(k)$ in a simulated QPSK system ($\sigma_p^2 = 1 \times 10^{-3}$ rad^2).

The trajectory of the first-order filter gain $\hat{\alpha}(k)$ versus k for different SNR γ_b is shown in Figure 4.1. Simulation shows that starting from $\hat{\alpha}(0) = 0$, $\hat{\alpha}(k)$ increases to a steady-state value between 0 and 1, as k increases. Thus, the gain $1 - \hat{\alpha}(k)$ on the

4.1 The Principle of Adaptive Decision-Aided Phase Estimation

input $r(k)\hat{m}^*(k)|\hat{m}(k)|^{-2}$ decreases from one at $k = 0$ to a steady-state value less than one. Figure 4.1 clearly shows that the higher the SNR, the smaller the value $\hat{\alpha}(k)$, because $r(k)\hat{m}^*(k)|\hat{m}(k)|^{-2}$ gives much more information on the RP $V(k + 1)$ than the previous one at high SNR [37]. As SNR goes to infinity, the steady-state $\hat{\alpha}(k)$ goes to zero, and the receiver becomes a differential detector. In this limit each signal sample $r(k)$ is almost noiseless, and it becomes unnecessary to do noise smoothing by taking the weighted average of the past samples [133]. It can be observed that the time-varying gain $\hat{\alpha}(k)$ is similar to the adaptive NLMS one-tap filter, though our adaptive DA receiver requires no preset parameters like the step size in the NLMS scheme. The generalized adaptive DA receiver not only inherits the merits of the DA ML receiver: no phase unwrapping and nonlinear operations, but also operates without the statistical knowledge of phase noise and additive noise. These characteristics make the adaptive receiver very suitable in a reconfigurable network.
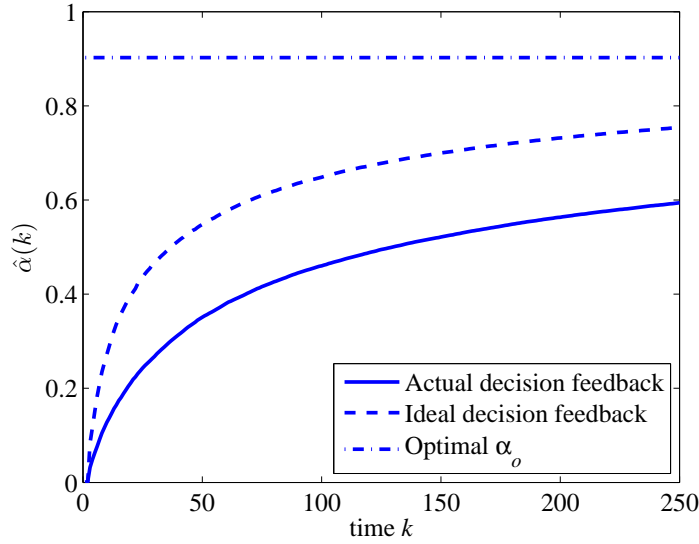
Figure 4.2: The adaptation process of $\hat{\alpha}(k)$ from actual decision-feedback, ideal decision-feedback ($\hat{m}(k) = m(k)$), and optimal $\hat{\alpha}_o$ Eq.(4.12), in a 16QAM system at γ_s=15 dB. $2\Delta\nu T = 5 \times 10^{-5}$. $\hat{\alpha}(k)$ is obtained by averaging over 500 runs.

The impact of decision errors on the performance of the adaptive DA algorith-

m is illustrated in Figure 4.2. After an initial 50-symbol training period, the adaptive DA algorithm switches to a decision-directed mode. As can be seen from Figure 4.2, the decision errors make the adaptation process slower than the ideal decision-feedback case, where the decision $\hat{m}(k)$ is identical to the transmitted one, i.e., $\hat{m}(k) = m(k)$. It can be observed that the risk function $R(k)$, ignoring the additive noise term for high SNR, becomes $\sum_{l=1}^{k} |1 - \exp(j\epsilon(l))|^2 \approx \sum_{l=1}^{k} |\epsilon(l)|^2$, by using the approximation $1 + jx \approx \exp(jx)$ for $|x| \ll 1$, in the case of ideal decision feedback. Note that the phase tracking error $\epsilon(l)$ in the adaptive DA algorithm is defined as $\epsilon(l) = \theta(l) - \hat{\theta}(l)$. As a result, the adaptive filter tracks the carrier phase by minimizing phase errors. However, decision errors in the actual decision feedback mode degrades the tracking performance by modifying the above risk function into $\sum_{l=1}^{k} |m(l)/\hat{m}(l) \exp(j\Delta\phi_s(l)) - \exp(j\epsilon(l))|^2$, leading to a longer tracking process to reach the steady-state value of $\hat{\alpha}(k)$ than that in the ideal decision feedback case [37]. The phase difference $\Delta\phi_s(l)$ is the signal phase error due to the symbol decision error, i.e., $\Delta\phi_s(l) = \phi_s(l) - \hat{\phi}_s(l)$.

4.2 Performance Investigation

The performance of the adaptive DA algorithm is evaluated using analysis and M-C simulations in M-PSK/QAM formats. In addition, the performance of DA ML is plotted for comparison.

4.2.1 MC Simulations

Figure 4.3 shows the BER performance in a QPSK system for DA ML with the optimum memory length ($L = 6$) for M-PSK formats, as suggested in the Section 3.2 of Chapter 2 and the adaptive DA algorithm. Note that one of the advantages of our
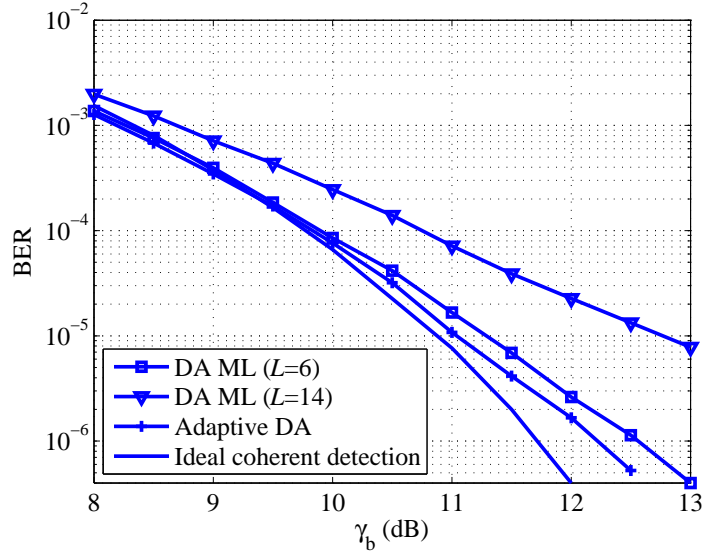
Figure 4.3: Performance comparison in a QPSK system among the three different PE algorithms: DA ML ($L = 6$, 14) and the adaptive DA algorithm. Ideal coherent detection and differential detection are also indicated for comparison. Linewidth per laser: 5 MHz @40 Gb/s.

adaptive DA algorithm is its self-adaptation capability without specifying any preset parameters. When the linewidth of each laser is 5 MHz, the adaptive DA algorithm has a significant performance improvement over the optimal DA ML. It is noted that the selection of the memory length L in DA ML is pivotal to determine its BER performance. For instance, DA ML with $L = 14$ performs very worse due to the fast-varying phase noise. However, the adaptive DA algorithm does not suffer from the BLE as in DA ML algorithm. Only 50 training data is necessary for the receiver to acquire channel characteristics. In addition, as illustrated in Figure 4.3, the adaptive DA algorithm can approach the performance of ideal coherent detection even if the laser linewidth is as large as 5 MHz per laser!

The performance comparisons in 16-QAM systems are illustrated in Figure 4.4, where the linewidth is assumed to be 250 KHz for each laser and the optimal L is 12 for DA ML. As can be seen, DA ML with optimal L can even outperform the adaptive DA
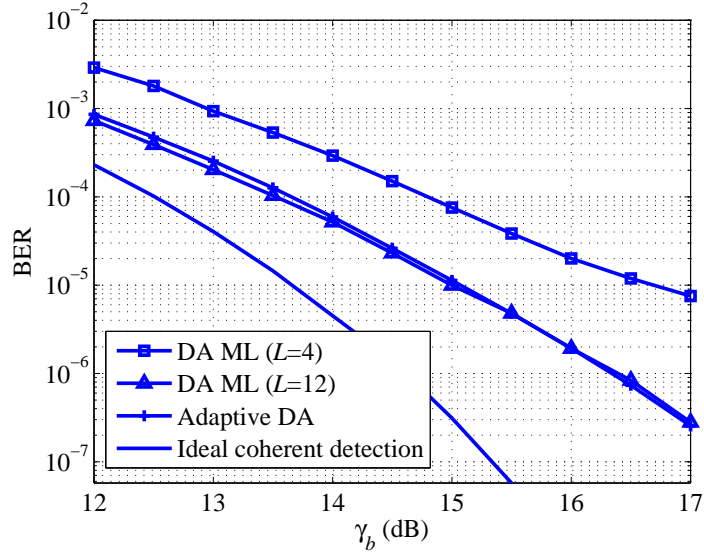
Figure 4.4: Performance comparison in a 16-QAM system between the DA ML ($L =$ 4, 12) and the adaptive DA algorithm. Linewidth per laser: 250KHz @40 Gbit/s.

receiver in 16-QAM system, though the performance improvement is not consistent. It can be observed that the adaptive DA receiver gradually outperforms DA ML ($L=$ 12) when BER is less than 10^{-7}. This observation indicates that DA ML cannot always be optimal due to its BLE, as explained in Section 3.2. Despite the small performance degradation ($<$ 0.1 dB compared to DA ML with optimal $L =$ 12) for γ_b less than 16.5 dB, the adaptive DA algorithm can avoid the exhaustive search for the optimal memory length L in DA ML algorithm. This feature renders the adaptive DA receiver more suitable than DA ML in a reconfigurable optical system. It is worth pointing out that DA ML with optimal memory length L in M-QAM formats is capable of outperforming the adaptive DA algorithm at low and moderate SNR.

The maximum tolerance to the laser linewidth leading to a 1-dB γ_b penalty at BER = 10^{-4} is listed in Table 4.1 for each modulation format. As can be seen, the adaptive DA algorithm can improve the laser linewidth tolerance in M-PSK systems. Also, we note from our simulation results that the adaptive receiver is consistently the best in M-PSK modulation formats, whether with a small laser linewidth or a large

Table 4.1: Linewidth per laser tolerance at 1-dB γ_b penalty for DA ML and the adaptive DA algorithms at BER=10^{-4} and 40 Gbit/s.

	QPSK	8PSK	16PSK	16-QAM
DA ML	4.4 MHz	400 KHz	88 KHz	250 KHz
Adaptive DA	5 MHz	547 KHz	95 KHz	200 KHz

one. This shows the adaptive DA algorithm can effectively operate in a time-varying and reconfigurable system without having to specify the memory length or search for the optimal one. Note that the laser linewidth tolerance of the adaptive DA algorithm is outperformed by DA ML with optimum memory length L. This issue will be addressed in the next section. Generally speaking, the adaptive DA algorithm is a very good candidate for PE algorithms in coherent receivers due to its self-acquisition and superior performance especially in M-PSK formats. On the other hand, the ultimate performance of the adaptive DA algorithm is still the best, which can be only illustrated through subsequent analytical works.

4.2.2   Phase Tracking Performance

As we have noted, the BER performance of the adaptive DA algorithm is not always better than DA ML in M-QAM formats [37]. Besides, MC simulations cannot explore the ultimate performance of the adaptive DA algorithm because it is time-consuming and even impossible to evaluate BER less than 10^{-9}[1] using conventional MC simulation techniques [134]. Here, we resort to analysis tools to investigate the reason behind the performance degradation when using the adaptive DA algorithm in M-QAM formats.

---

[1]Sometimes it is suggested to run the BER to below this level to check whether there is any error floor in the algorithm.

In the steady state, the RP $V(k)$ can be well approximated by $\exp\left(j\hat{\theta}(k)\right)$. Thus, at high SNR where $\hat{m}(k) = m(k)$, Eq.(4.1) can be re-written as [133]

$$\exp\left[j\hat{\theta}(k+1)\right] = \alpha \exp\left[j\hat{\theta}(k)\right] + (1-\alpha)\left[\exp\left[j\theta(k)\right] + n'(k)\right], \qquad (4.6)$$

where $n'(k) \equiv n(k)/m(k)$ with mean zero and variance $E[|n'(k)|^2] = N_0 E[|1/m(k)|^2]$. Taking into account the Wiener process assumed for the laser phase noise in Eq.(3.15) in Section 3.3 of Chapter 3, and dividing both sides of Eq.(4.6) by $\exp\left[j\theta(k)\right]$, the phase error Eq.(4.6) can be further simplified into

$$\exp\left[j\nu(k+1) - \epsilon(k+1)\right] = \alpha_k \exp\left[-j\epsilon(k)\right] + (1-\alpha_k)\left[1 + n'(k)\exp\left[-j\theta(k+1)\right]\right].$$
$$(4.7)$$

Note that the phase estimate $\hat{\theta}(k)$ consists of two parts: the actual laser phase noise $\theta(k)$ and the phase tracking error $\epsilon(k)$, i.e., $\hat{\theta}(k) = \theta(k) - \epsilon(k)$. Using the approximation $1 + jx \approx \exp\left(jx\right)$ for $|x| \ll 1$, Eq.(4.7) can be linearly expanded into the form

$$\epsilon(k+1) \approx \nu(k+1) + \alpha_k \epsilon(k) - (1-\alpha_k)\mathbf{Im}\left[n'(k)\exp\left[-j\theta(k+1)\right]\right]. \qquad (4.8)$$

It is easy to show that $E[\epsilon(k+1)] = \alpha_k E[\epsilon(k)]$ by taking expectation of both sides of Eq.(4.8). Thus we have $E[\epsilon(k)] = 0$ because $\alpha_k \neq 0$. Further, squaring both sides of Eq.(4.8), the variance of the phase error $\epsilon(k)$ can be shown as

$$\sigma_\epsilon^2(k+1) = \alpha_k^2 \sigma_\epsilon^2(k) + \sigma_p^2 + \frac{1}{2}(1-\alpha_k)^2 E\left[|n'(k)|^2\right]. \qquad (4.9)$$

The steady-state variance σ_ϵ^2 is obtained through letting $\sigma_\epsilon^2 = \sigma_\epsilon^2(k+1) = \sigma_\epsilon^2(k)$ in Eq.(4.9) when $k \to \infty$, thus yielding [37]

$$\sigma_\epsilon^2 = \frac{\sigma_p^2}{1-\alpha_k^2} + \frac{\eta}{2\gamma_s} \cdot \frac{1-\alpha_k}{1+\alpha_k}, \qquad (4.10)$$

where

$$\eta \equiv N_0 E[|1/m(k)|^2]$$

$$= \begin{cases} 1, & M\text{-PSK}; \\ 1.8889, 2.2283, \text{ and } 2.6854, & M\text{-QAM}(M = 16, 32 \text{ and } 64); \end{cases} \tag{4.11}$$

is called constellation penalty [4]. It can be observed that Eq.(4.10) reduces to the same as in [133] for constant-amplitude M-PSK formats, where the constellation penalty is 1.

The optimal steady-state α, denoted by α_o, that minimizes the variance σ_ϵ^2 is easily derived as Eq.(4.12) by solving $\partial \sigma_\epsilon^2 / \partial \alpha = 0$:

$$\alpha_o = 1 + \xi - \xi \sqrt{2\xi^{-1} + 1}, \tag{4.12}$$

where $\xi \equiv \gamma_s \sigma_p^2 / \eta$. Substituting the optimal α_o into Eq.(4.10), the minimum steady-state variance $\sigma_{\epsilon,\text{min}}^2$ is given by

$$\sigma_{\epsilon,\text{min}}^2 = \frac{\sigma_p^2}{2\xi \left[\left(\sqrt{2\xi^{-1} + 1} - 1 \right) (1 + \xi) - 1 \right]} + \frac{\sigma_p^2}{2\xi \sqrt{2\xi^{-1} + 1}}. \tag{4.13}$$

To justify the minimum steady-state variance derived in Eq.(4.13), we plot the sample variance from simulations as a function of the SNR per symbol (γ_s) in Figure 4.5. As can be seen, at high SNR, the simulated phase error variances approach their respective steady-state values for M-PSK and 16-QAM. In contrast, at a lower SNR, the variances of the phase error from simulations deviate from the theoretical $\sigma_{\epsilon,\text{min}}^2$, as a result of those frequently occurring decision errors. It is worth mentioning that the minimum variance $\sigma_{\epsilon,\text{min}}^2$ of 16QAM is larger than that of M-PSK due to the constellation penalty found in Eq.(4.11). In addition, the adaptive DA algorithm for 16-QAM is much more sensitive to decision errors compared to the M-PSK scenario, as illustrated in Figure 4.5 within the range of $\gamma_s < 15$ dB. This can be explained as follows. The risk function in Eq.(4.2) for constant-amplitude M-PSK

formats can be rewritten as $\sum_{l=1}^{k} |\exp(j\Delta\phi_s(l)) - \exp(j\epsilon(l))|^2$, after neglecting additive noises. However, in the case of M-QAM formats, this risk function is expressed as $\sum_{l=1}^{k} |m(l)/\hat{m}(l) \exp(j\Delta\phi_s(l)) - \exp(j\epsilon(l))|^2$. Comparing these two forms of the risk function, we see that both the magnitude and the phase of the decision symbol would affect the risk function in QAM systems, thus degrading the phase tracking performance more severely at low SNR region.
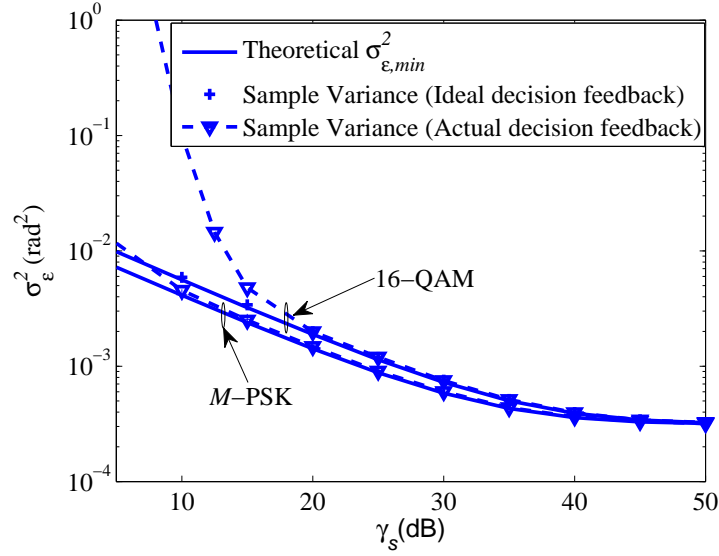
Figure 4.5: The theoretical minimum variance $\sigma_{\epsilon,\min}^2$, and the sample variance from simulation (ideal and actual decision feedback) versus γ_s in M-PSK and 16-QAM formats. $\sigma_p^2 = 3.41 \times 10^{-4}$ rad^2.

4.2.3 Performance Comparison

In order to address the issue that the adaptive DA algorithm is outperformed by DA ML at low and moderate SNR in M-QAM formats, the performance of the adaptive DA algorithm is first compared with DA ML in different modulation formats. On the other hand, there are some other adaptive PEs, such as NLMS and Kalman filter PE, to estimate carrier phase. Their performances are also analytically investigated to

show that they share the same performance as the adaptive DA algorithm, though the difference lies in their approaches to determine the filter stepsize [147].

4.2.3.1   Adaptive DA versus DA ML Algorithm
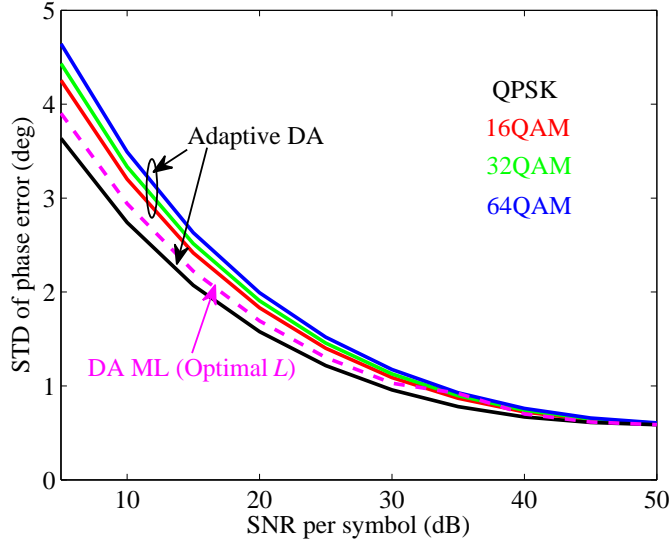
Figure 4.6: The comparison of STD between DA ML and adaptive DA PE in M-QAM formats obtained from analysis. $\sigma_p^2 = 1 \times 10^{-4}$ rad^2.

Figure 4.6 compares the STD of DA ML with optimal memory length L and the adaptive DA PE in M-QAM formats (M=4, 16, 32 and 64). The STD of DA ML algorithm remains the same in higher-order QAM formats since its phase error variance depends only on the SNR per symbol and memory length L. In contrast, due to the constellation penalty η in the adaptive DA algorithm, the adaptive DA algorithm is outperformed by DA ML PE at low SNR region in non-constant modulus QAM formats while it is expected to achieve a better performance in M-PSK formats. Furthermore, the adaptive DA algorithm has the same performance as DA ML with the optimal memory length at high SNR regardless of modulation formats. In a real optical system with the aid of FEC, uncoded BER around 10^{-3} is sufficient to achieve error-free transmission. Additionally, it is desirable to operate optical systems at low SNR region

to reduce fiber nonlinearity and channel crosstalk effects. Hence, of importance are the performances of PEs at low and moderate SNR region. In other words, the adaptive DA PE is more suitable in M-PSK format in view of the optimal performance, though its self-adaptation characteristics renders it a superior candidate in reconfigurable systems [148].

4.2.3.2   Adaptive DA versus Other Adaptive PEs

Several adaptive PEs have been recently proposed in coherent optical communication systems, such as Kalman filter and NLMS. Here, we take a quick analysis of the phase error in both Kalman and NLMS PE to compare their performances.

- **Kalman Filter**

A canonical Kalman filter with stepsize ρ can be further simplified into the form [149]

$$\hat{\theta}(k+1) = \hat{\theta}(k) + (1-\rho)\arg\left[r(k)\exp\left(-j\hat{\theta}(k)\right)\hat{m}^*(k)\right]. \qquad (4.14)$$

With the assumption that decision feedback is error-free, the phase error variance of Kalman filter at the steady state is shown as [147]

$$\sigma_\epsilon^2 = \frac{\sigma_p^2 + (1-\rho)^2\sigma_{n'}^2/2}{1-\rho^2}, \qquad (4.15)$$

by using the same analytical approaches as in Section 4.2.2. Here, the stepsize $\rho = (1+\xi/2) - \sqrt{(1+\xi/2)^2 - 1}$.

- **NLMS Filter**

The principle of PE using NLMS algorithm is given by [129]

$$c(k+1) = c(k) + \frac{\mu}{|r(k)|^2}e^*(k)r(k); \qquad (4.16)$$

$$e(k) = \hat{m}(k) - c^*(k)r(k), \qquad (4.17)$$
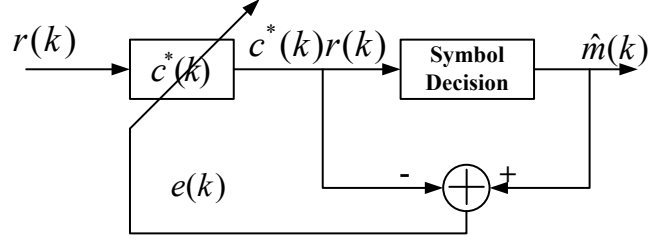
Figure 4.7: The structure of NLMS PE.

where $c(k)$ is the coefficient of an one-tap filter, and μ is the stepsize of the NLMS algorithm, which is manually set in advance between 0 and 1. Figure 4.7 depicts the entire structure of the NLMS filter. Here, we demonstrate how to derive its analytical phase error variance for the first time. Substituting $r(k)$ (see Eq.(3.1)) and $e(k)$ into Eq.(4.16), it becomes

$$c(k+1) = (1-\mu)c(k) + \frac{\mu}{|r(k)|^2}\hat{m}^*(k)r(k)$$

$$\approx (1-\mu)c(k) + \mu \exp\left[j\theta(k)\right] + \frac{\mu}{m(k)}n(k). \tag{4.18}$$

when $|r(k)| \approx m(k)$ is assumed. In the steady state, it is reasonable to assume that the coefficient $c(k)$ of the one-tap filters tracks the time-varying laser phase noise $\theta(k)$, which means $c(k) = \exp\left(j\hat{\theta}(k)\right)$ [150]. Applying the similar method that the phase estimate is given by $\hat{\theta}(k) = \theta(k) - \epsilon(k)$ as demonstrated in Section 4.2.2, it is easy to show that the phase error variance of the NLMS filter is expressed as

$$\sigma_\epsilon^2 = \frac{\sigma_p^2 + \mu^2\sigma_{n'}^2/2}{1 - (1-\mu)^2}. \tag{4.19}$$

The theoretically derived phase error variance of Kalman filter is verified using MC simulations in 4-, 16- and 32-QAM formats. As shown in Figure 4.8, the STD of the phase error in the Kalman filter through MC simulation agrees well with the analytical STD from Eq.(4.15), especially at high SNRs. The small deviation at low
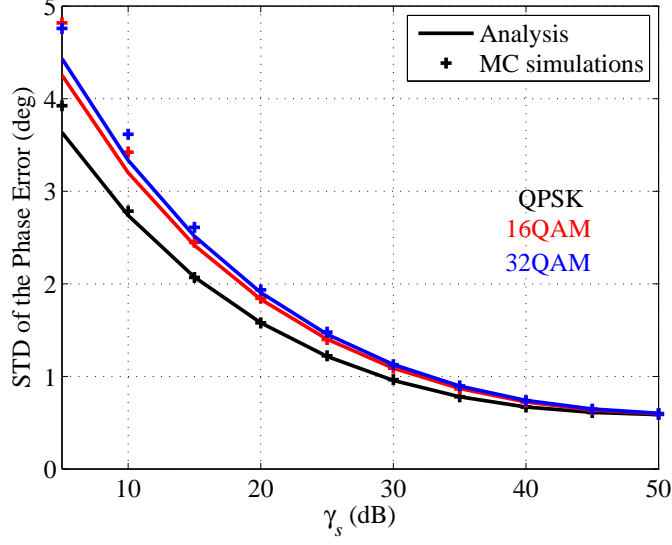
Figure 4.8: The STD of the phase error of Kalman PE in 4-QAM (black), 16-QAM (red) and 32-QAM (blue) formats obtained from analysis (line) and MC simulations (marker '+'). $\sigma_p^2 = 1 \times 10^{-4}$ rad^2.

SNRs is due to the fact that the approximations that are made in the derivation of the phase error are not very accurate at low SNRs. In addition, the NLMS PE is evaluated in 32-QAM format at $\sigma_p^2 = 10^{-3}$ rad^2, as illustrated in Figure 4.9 (a). At low SNRs, a smaller stepsize μ is preferred to average out more additive noise; whereas a large stepsize μ performs better to track the fast-varying carrier phase at high SNRs. To minimize the phase error variance, the optimal stepsize μ_{opt}, by solving the equation $\partial \sigma_\epsilon^2 / \partial \mu = 0$, is found to be

$$\mu_{\text{opt}} = \xi \left(\sqrt{1 + 2\xi^{-1}} - 1 \right). \tag{4.20}$$

Note that $\xi \equiv \gamma_s \sigma_p^2 / \eta$. Figure 4.9 (a) shows that the STD of the phase error from the optimal stepsize is consistently minimal. Besides, Figure 4.9 (b) indicates that error floor may be avoided by selecting a suitable stepsize.

So far, we have obtained the phase error variances for these three adaptive PEs considered in this chapter. It would be interesting to take a deeper look at their final
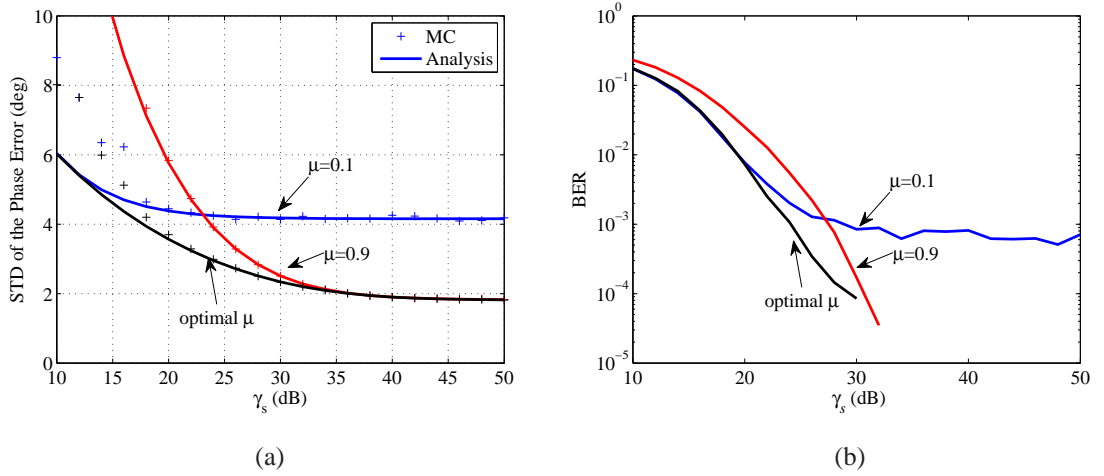
Figure 4.9: The STD of phase error of NLMS PE with different stepsizes $\mu = 0.1, 0.9$ and optimal μ in a 32-QAM format from analysis (line) and MC simulations (marker '+'). $\sigma_p^2 = 1 \times 10^{-3}\text{rad}^2$.

expressions Eqs.(4.10), (4.15) and (4.19): they have the same form with a minor difference in the selection of the stepsize ($\alpha = \rho = 1 - \mu$). This relationship can also be justified by comparing their optimal stepsizes. In other words, these three adaptive PEs should have the same performance under the same circumstances (the same system noises and stepsize). The difference lies in the processes by which they acquire the optimal stepsize.

The stepsize in NLMS is normally fixed to estimate carrier phase [129] but with possible performance degradation, as already illustrated in Figure 4.9. As a result, it is necessary to put some effort to find the optimal stepsize at the SNR of interest and different laser linewidths. As for Kalman PE, it requires the exact statistics of the system noises to compute state transition matrix [39, 151]. Therefore, the optimal stepsize can be computed in advance such that Kalman PE operates at the optimal state. However, the optimal stepsizes in the other two methods are also available if all the information is known to the receiver, according to our analysis here.

In contrast, the stepsize α in the adaptive DA algorithm is dynamically obtained by using a sequence of training data to acquire the knowledge of a channel. Simulations show that it can quickly approach the optimal stepsize [37]. In addition, the algorithm's complexity during the adaptation process is reduced because of recursive computations [133]. On the other hand, like NLMS, a fixed stepsize is expected to work in the adaptive DA algorithm, though at the expense of possible performance loss [147]. In short, it is more suitable in a reconfigurable optical system than the other two methods.

4.3   Experiments of Long-Haul Transmission

As explained in Section 3.6 of Chapter 3, the B2B experiments are not presented because of the small laser linewidth, where they all have the same performances. Instead, a long-haul transmission experiment was carried out to demonstrate the performance improvement of the adaptive DA algorithm.
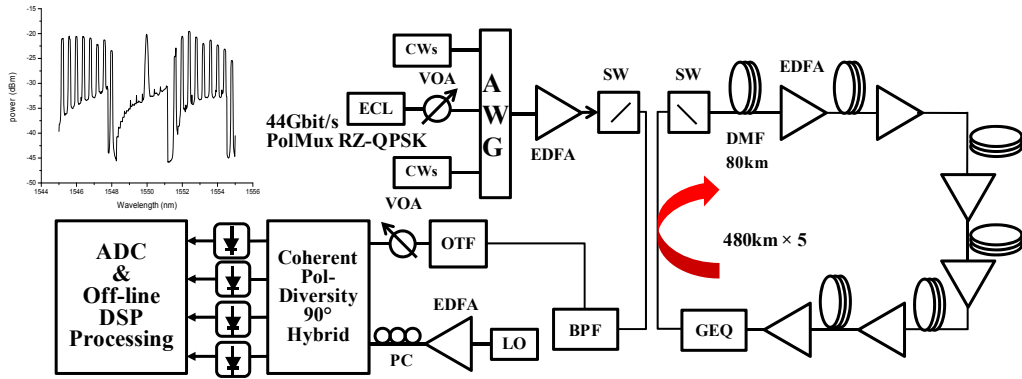
Figure 4.10: The experimental setup for a long-haul PolMux-QPSK transmission with a coherent receiver.

The system setup is illustrated in Figure 4.10. A typical PolMux RZ-QPSK transmitter [26], centering at wavelength 1550.0 nm modulated by 44 Gbit/s bit rate, was coupled into the transmission loop. The other channels were occupied by CW lasers for noise suppression in EDFAs. The fiber loop consisted of 6-span ~80 km dispersion

managed fiber, resulting in a total 800 ps/nm residual chromatic dispersion after each loop. The measured signal was transmitted over 5 loops, resulting in a total transmission length of 2,400 km. The spectra after 5-loop transmission are depicted in the inset of Figure 4.10. The received signal was filtered through a 1nm optical tunable filter, and was then fed into a coherent receiver. The received signal power was kept at -5 dBm throughout all the measurements.

To realize the training transmission, two sequences of known bits were mapped into a QPSK constellation, and differentially decoded offline to resolve the possible phase ambiguity in received signals. The training bits for inphase and quadrature branches were uploaded into two synchronized pulse pattern generators (PPGs). The conventional DSP processing was implemented before PE steps. To find the beginning of the training data, cross correlations between the received bits and training in both inphase and quadrature branches were calculated, as depicted in Figure 4.11. The purpose of using both correlations was to monitor the possible run-off of the synchronization of two PPGs. The length of the training bits was selected to 17,152, which was originally served to decode low-density parity check (LDPC) codes.
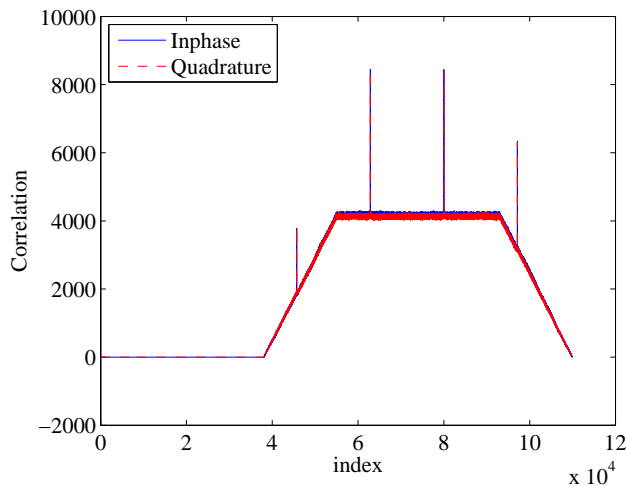
Figure 4.11: The illustration of the correlation between the recovered bits in each branch and the training bits after 2,400 km transmission. $P_t = -8$ dBm.

4.3 Experiments of Long-Haul Transmission

The first 50 symbols of the LDPC training were used for the acquisition process of the adaptive DA algorithm, and the remaining training data were assumed to be payload data for calculating BER. The conventional V&V Mth-power and a commonly-used block M-th power PEs [22] were also plotted for comparison. There were 20 sets of data having at least 2 groups of training data each set, thus resulting in around 1.2 million total bits for calculating BER, as shown in Figure 4.12.
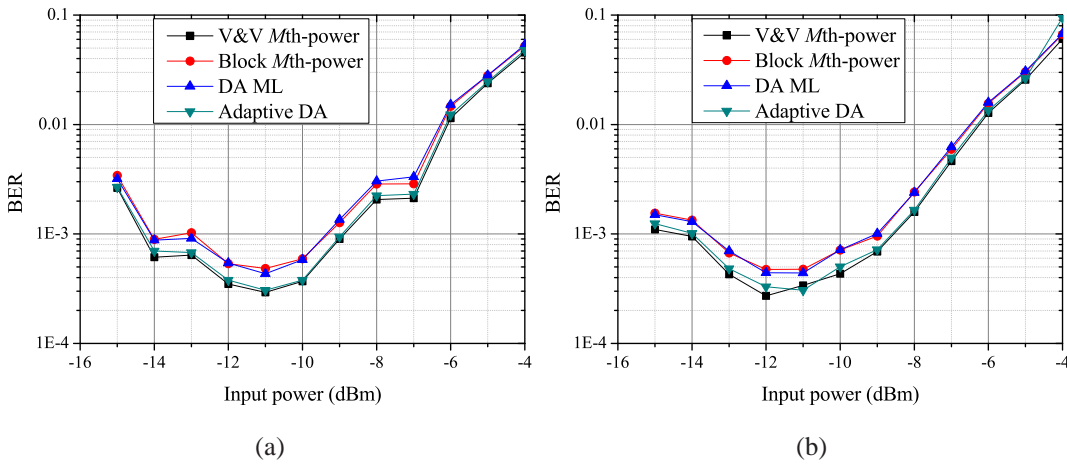
Figure 4.12: The experimental BER versus input power for different PEs: V&V Mth-power (N=11), block M-th power (N=11), DA ML ($L = 11$), and the adaptive DA algorithm. (a) X-polarization; (b) Y-polarization.

In the presence of fiber nonlinearity, the V&V Mth-power performs better than the DA ML algorithm whereas both of them have a comparable performance in the B2B case when only small laser phase noise is dominant. It may be due to the fact that the V&V Mth-power is capable of fully utilizing the pre- and post-symbols to exploit carrier phase information. In contrast, the DA ML algorithm only relies on the past symbols to predict the current phase information. As can be observed, the adaptive DA algorithm outperforms DA ML even in the presence of fiber nonlinearity because it can adaptively adjust the filter gain according to the received signals. The commonly-used block M-th power algorithm, which has less computational load than

the V&V Mth-power, has the same performance as DA ML PE. The common phase estimate $\hat{\theta}$ for all the N symbols in the entire block degrades the performance of block M-th power algorithm.

4.4 Conclusion

As demonstrated in Chapter 3, the DA ML algorithm exhibits BLE phenomenon which requires extensive MC simulations to find out the optimal memory length. Alternatively, the optimal memory length L_{opt} can be determined if the statistics of the channel noises are known to the receiver, though these parameters may not be available especially in reconfigurable optical systems. As a result, an adaptive DA PE in this chapter was developed in different M-PSK/QAM formats to eliminate BLE and even improve system performance as well.

MC simulations indicate that the adaptive DA algorithm outperforms DA ML PE in M-PSK formats, thus increasing laser linewidth tolerance. However, its performance in M-QAM formats suffers from the constellation penalty due to the non-constant-modulus characteristics, which was investigated through a comprehensive analysis. Despite the fact that DA ML and the adaptive DA algorithms have the same performance at high SNRs, DA ML is independent of modulation formats, thus rendering DA ML algorithm a better choice in M-QAM formats at low and moderate SNRs.

Compared with Kalman filter and NLMS algorithm, we find that the adaptive DA shares the same performance as these two adaptive algorithms except that their approaches to determine the filter stepsize. The analytical results of Kalman filter and NLMS PE algorithms were presented for the first time to examine their performances in different modulation formats. The optimal stepsize in Kalman filter can be computed in advance once the exact statistics of the system noises are available; however, this

prerequisite is also applied to the other two adaptive methods and even the DA ML algorithm. As for NLMS PE algorithm, its stepsize is usually fixed to estimate carrier phase in sacrifice of some possible performance degradation. In contrast, the stepsize α in the adaptive DA algorithm is dynamically obtained by using a sequence of training data to acquire the knowledge of a channel. Simulations show that it can quickly approach the optimal stepsize, though training data should be sent to the receiver to acquire the channel information.

Finally, a 2,400 km-long 44 Gbit/s coherent PolMux-RZ-QPSK experiment was conducted to examine the performance of the adaptive DA algorithm in the presence of fiber nonlinearity. The adaptive DA algorithm has the same performance as the V&V Mth-power performs, whose performances are better than the DA ML and the commonly-used block M-th power algorithm.

Chapter 5

Study of DSP Algorithms for Large Frequency Offset and Fiber Nonlinearity

The aim of DSP-based PE algorithms is to recover carrier phase noise, thus allowing using a free-running LO laser. In addition to laser phase noise, frequency offset between transmitter and LO lasers also leads to linear increment of carrier phases between two consecutive symbols, which is another form of phase noise[1]. This frequency offset can be as large as $\pm$ 5 GHz due to aging, heating and fabrication of lasers [40]. It is worth pointing out that PE algorithms generally require that frequency offset between the transmitter and LO lasers should be no more than 10 percent of symbol rate [41], which translates to be $\sim$1 GHz for a 10 GBaud coherent optical system. Hence, an additional DSP-based FOE is imperative to make sure PE algorithms work properly [42]. On the other hand, fiber nonlinearity would induce nonlinear phase noise in the received signals through the interaction between fiber Kerr effect and ASE

---

[1]The constant frequency offset will add a phase offset between consecutive two symbols, which can be regarded as a kind of phase noise.

noises of optical amplifiers [152]. Due to the access to the electric field through co-herent receivers, we propose two novel DSP algorithms to address these two critical factors, frequency offset and fiber nonlinearity, in a coherent optical system.

5.1 Dual-Stage FOE based on Gardner Timing Recovery Algorithm

A proposed FOE is able to estimate frequency offset up to 8 GHz in a 43 Gbit/s Pol-Mux QPSK system [153]; however, a feedback structure makes it difficult to implement in parallel processing. A feed-forward FOE is desirable to avoid performance degradation when implementing with parallel processing [141] in a real-time coherent receiver. Mth-power operation is usually applied to remove data modulation in those feed-forward FOEs. As a result, the maximal estimation range for M-PSK formats is limited to [-$R_s/2M$, $R_s/2M$] [43], where R_s refers to the system symbol rate. Taking a 40 Gbit/s PolMux-QPSK system as an example, the maximum estimation range is only [-1.25 GHz, 1.25 GHz] in an FOE that utilizes Mth-power to remove data modulation [45, 46, 154]. In this section, a novel, dual-stage, cascaded FOE scheme consisting of a coarse FOE and a fine FOE will be presented. In a 42.8 Gbit/s PolMux-RZ-QPSK system, the estimation range of the proposed dual-stage FOE will be experimentally shown to have 4 times that which can be achieved using Mth-power-based FOE [42]. The estimation range of the dual-stage FOE can be up to almost $\pm 0.9 R_s$ in simulation, and $\pm 0.5 R_s$ in experiment. The achieved results have, to our best knowledge, the largest range among FOEs reported so far in the literatures.
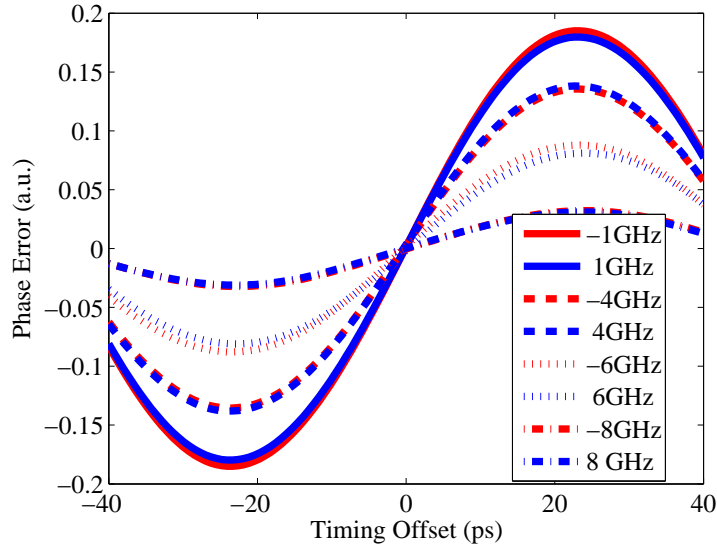
Figure 5.1: Simulation results on S-curves (phase error output of Gardner algorithm) versus timing offset at different frequency offsets in a coherent 42.8 Gbit/s PolMux-RZ-QPSK system (OSNR=10 dB).

5.1.1 The Principle of Coarse FOE

Timing recovery using, for instance, Gardner algorithm [109], is usually necessary to correct timing phase error between the transmitter and receiver clocks in coherent receivers [24, 112, 155]. The simple Gardner algorithm can be applied to yield a phase error output when only two samples per symbol are available. In a coherent PolMux M-PSK system with Nyquist sampling rate (2 samples per symbol), Gardner algorithm can be represented by [109]

$$U_t(2k) = I(2k-1)[I(2k) - I(2k-2)] + Q(2k-1)[Q(2k) - Q(2k-2)], \quad (5.1)$$

where U is a phase error, and I and Q are the samples of in-phase and quadrature-phase, respectively, for each polarization state. Note that an S curve shows the relationship between phase error U_t and input signal timing error t [155]. By averaging U_t over a long sequence of samples, it gives a mean phase error corresponding to the current sampling offset t. However, the performance of Gardner algorithm is subjected to
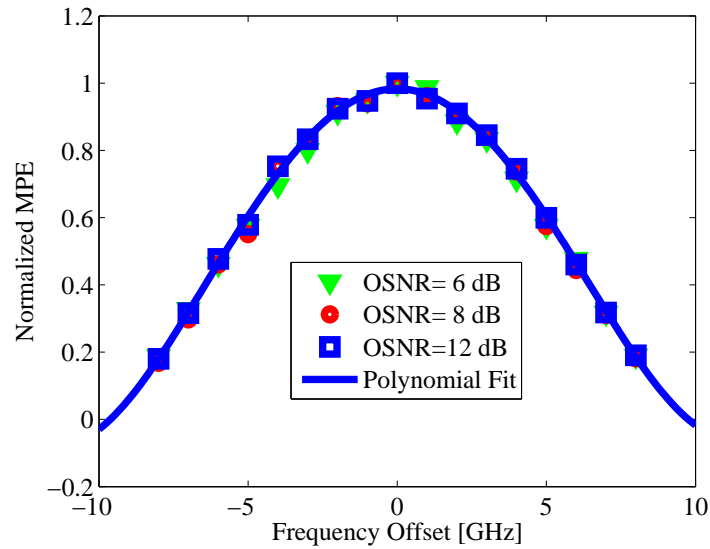
Figure 5.2: Normalized MPE versus frequency offsets under different OSNRs.
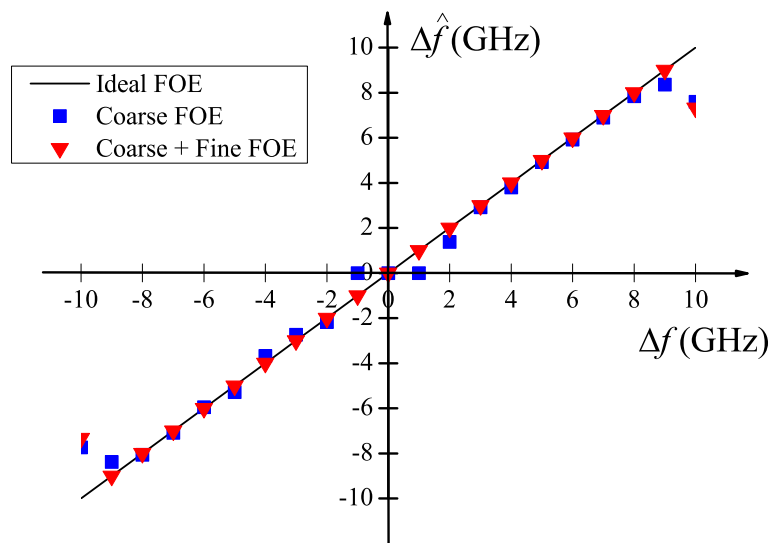
Figure 5.3: Estimated frequency offset ($\Delta \hat{f}_c$) by coarse FOE and dual-stage cascaded FOE at OSNR = 9.5 dB.

5.1 Dual-Stage FOE based on Gardner Timing Recovery Algorithm

frequency offset [155, 156]. To investigate the impact of frequency offset on the phase error output of Gardner algorithm, we first perform simulations to emulate a 42.8 G-bit/s coherent PolMux RZ-QPSK system. The linewidth of both transmitter and LO lasers is set to be 100 KHz for emulating ECLs. The outputs of balanced detectors are passed through a 5th-order Bessel electrical low-pass filter with bandwidths at 75% of symbol rate. The effective number of bits for ADCs is 8 in simulation.

A series of S-curves, representing the phase error output of Gardner algorithm from Eq.(5.1), is shown in Figure 5.1 when sweeping the sampling timing offset t and tuning frequency offset (Δf) in the system. It can be observed that this S-curve becomes flat as frequency offset increases, which means the maximal phase error (MPE) outputs of the Gardner algorithm drop. Thus, a non-zero frequency offset makes Gardner timing recovery algorithm less sensitive to the sampling offset in coherent receivers, conforming to the results in [155]. On the other hand, we may take advantage of this underlying relationship between frequency offset and the MPE of Gardner algorithm as a measure to estimate frequency offset in a system. Figure 5.2 shows that normalized MPEs under different OSNR levels exhibit almost the same trend within ± 9 GHz. We use a 4th-order polynomial to well emulate the relationship between the normalized MPE outputs and frequency offsets ($|\Delta f| \leq 9$ GHz). It should be noticed that the absolute value of MPE varies at different OSNRs so that normalization is required with respect to those frequency-offset-free MPEs, which are stored firstly in a look-up table. Therefore, we can rely on this fixed relationship between the normalized MPE and frequency offset to measure frequency offset in a system. The proposed FOE based on the normalized MPE is called a coarse FOE, since it can only offer a coarse estimation of frequency offset ($\Delta \hat{f}_c$), as illustrated in Figure 5.3. The coarse FOE cannot accurately estimate frequency offset around ± 1 GHz, because the MPE of Gardner algorithm is less affected by such a small amount of frequency offset compared to the case when $|\Delta \hat{f}| \geq 2$ GHz, which is further justified by the smaller MPE curve slope around zero

frequency offset in Figure 5.2.

5.1.2   Implementation of the Dual-Stage FOE

It can be observed that the emulated polynomial is an even function (see Figure 5.2). In other words, MPE alone cannot identify the sign of frequency offset. To resolve this sign ambiguity, we propose a sign identifier that monitors the maximum in the discrete frequency domain of the received signals:

$$\Delta \hat{f_c} = \begin{cases} \Delta \hat{f_c}, & 1 < i_{\max} \leq N/2; \\ -\Delta \hat{f_c}, & \text{otherwise} \end{cases} \quad (5.2)$$

where $i_{\max}$ is the index corresponding to the maximum of $\text{FFT}(x)$, $x(k) = I(k) + j \cdot Q(k)$ is the received samples in either X or Y polarization, $\text{FFT}(\cdot)$ is FFT operation with size of N. The principle of the sign identifier is to monitor the maximum of the signal spectrum which is shifted by the frequency offset existing in a system. Based on the characteristics of FFT operation and Fourier transform, if $\Delta f > 0$, it means the maximum of the received signals' spectrum will move to the right-hand side of zero frequency; otherwise, it will move to the left-hand side of zero frequency [80]. As displayed in Figure 5.3, the proposed sign identifier successfully addresses the sign ambiguity in the proposed coarse FOE.

In addition, through scrutinizing the coarse frequency offset estimate of the coarse FOE, i.e., $\Delta \hat{f_c}$, its estimation error is limited to be within ± 1 GHz in our simulated system. To tackle the residual frequency offset, a traditional FOE is introduced in cascade with the coarse FOE, as shown in Figure 5.4. Since the traditional FOE involving Mth-power can only estimate Δf within $\pm R_s/2M$, it is called a fine FOE [42]. Here, we choose an FFT-based FOE where data modulation is removed by Mth-power, as explained in Section 2.4.3 of Chapter 2. A possible structure of a dual-stage cascaded FOE is proposed in Figure 5.4, where the polynomial between frequency offset and
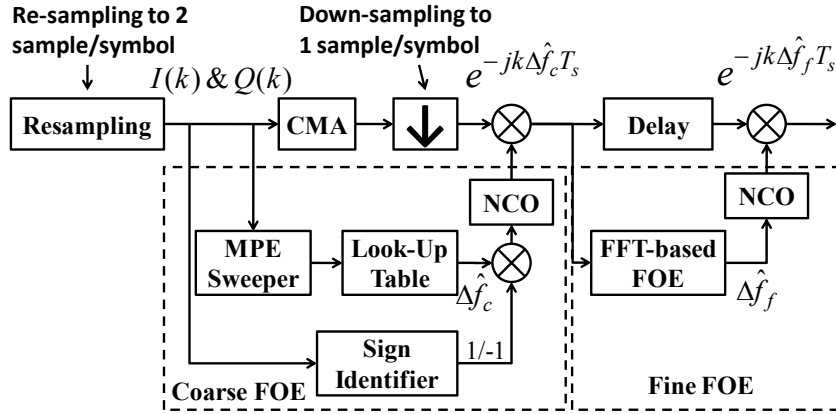
Figure 5.4: A proposed structure of the dual-stage cascaded FOE. CMA: constant modulus algorithm; NCO: numerically-controlled oscillator.

normalized MPE can be realized by using a look-up table. Through sweeping sampling offset t from 0 to a symbol duration T_s, MPE obtained in Gardner algorithm Eq.(5.1) is normalized to estimate a coarse frequency offset ($\Delta \hat{f}_c$) while the residual frequency offset can be recovered through the FFT-based FOE ($\Delta \hat{f}_f$). The sign ambiguity of the coarse FOE is resolved by the proposed sign identifier. To implement the proposed FOE, as shown in Figure 5.4, the coarse FOE can work independently in parallel with CMA, which is applied to de-multiplex two orthogonal channels. After down-sampling the outputs of CMA to one sample per symbol, the output sample is firstly de-rotated by the coarse frequency offset $\Delta \hat{f}_c$. The de-rotated signals will be then used in the fine FOE to estimate the fine frequency offset $\Delta \hat{f}_f$, thus further compensating for the residual frequency offset in those de-rotated samples. Of importance is that the inputs to the fine FOE should be roughly compensated for through the coarse frequency offset such that two numerically-controlled oscillators (NCO) are necessary. In addition, in the presence of both PMD and CD effects, FIR filters can be introduced to firstly compensate for CD before the dual-stage FOE and CMA are applied [30]. Figure 5.3 plots the estimated $\Delta \hat{f}$ by single coarse FOE and dual-stage cascaded FOE

at OSNR = 9.5 dB. The results show that the proposed dual-stage FOE is capable of accurately estimating a frequency offset up to ± 9 GHz, around $\pm 0.9 R_s$, the largest frequency offset estimation range reported so far n simulations.

5.1.3   Experiment

Experiments were carried out to demonstrate this dual-stage cascaded FOE. The configuration of a 42.8-Gbit/s coherent PolMux RZ-QPSK system was the same as the one given in Figure 3.17 of Chapter 3.
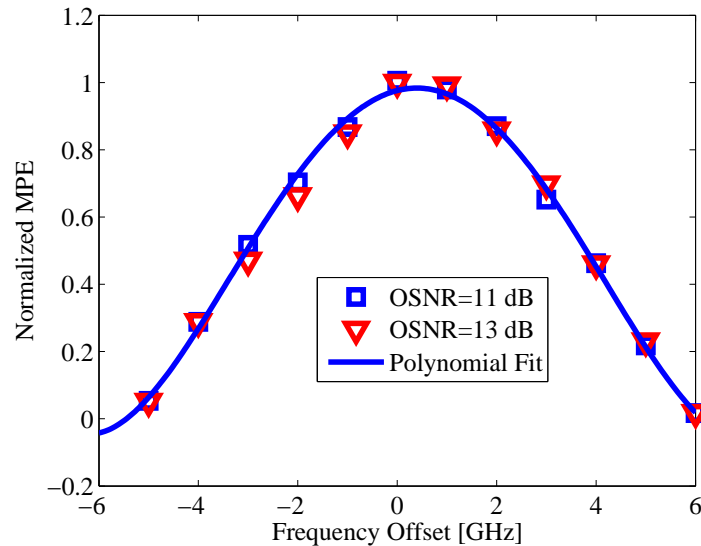
Figure 5.5: The experimentally measured normalized MPE versus frequency offsets at OSNR= 11 dB and 13 dB.

Those normalized MPEs were measured under different frequency offsets (1 GHz steps) at different OSNRs in 0.1 nm. It was found that a 4th-order polynomial can also be applied to fit the trend in our experimental results and the estimation range of cascaded FOE reduces to almost [-5 GHz, 6 GHz] for our experimental configuration, as plotted in Figure 5.5. With the knowledge of the normalized MPE as a function of frequency offset, we can set up a look-up table[2] for the dual-stage cascaded FOE

---

[2]The size of look-up table in our experiments was only 1 kb.
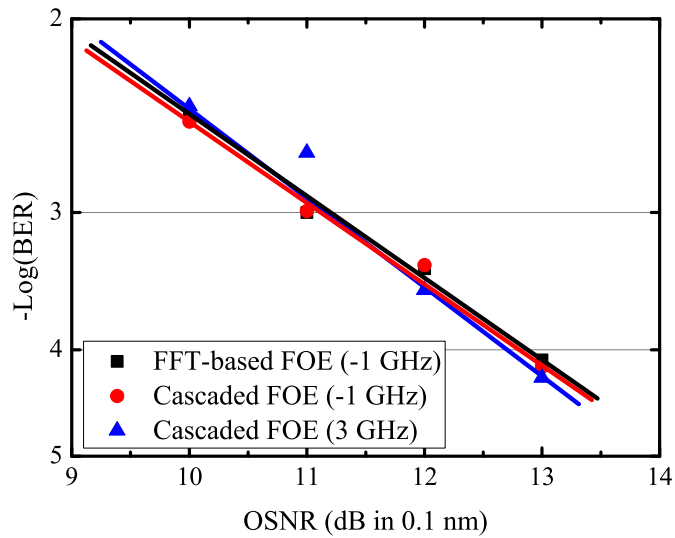
Figure 5.6: Measured BER performance of cascaded and conventional FOE at frequency offset of -1 GHz and 3 GHz. .
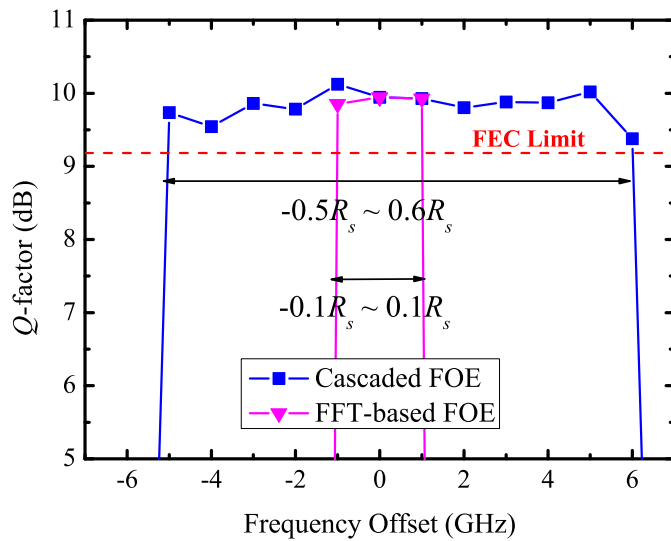
Figure 5.7: A comparison of the measured Q-factor between proposed dual-stage cascaded FOE and single FFT-based FOE. The enhanced FEC limit is 2×10^{-3}. (The signal OSNR is set at 12 dB)

for further data processing. We also plot the performance of single FFT-based FOE for comparison. Figure 5.6 shows the BER performance of dual-stage or single FFT-based FOE used in the DSP processing of those captured data at different frequency offsets (-1 GHz and 3 GHz). Note that the BER of single FFT-based FOE at 3GHz frequency offset is omitted since it does not work at a frequency offset exceeding its theoretical limit ($\pm$ 1.34 GHz for 10.7 GBaud). It can be seen that our proposed dual-stage cascaded FOE consistently and accurately recovers the original data even though frequency offset is 3 GHz without having any performance degradation. Furthermore, Figure 5.7 illustrates the Q-factor against different frequency offsets. The Q-factor is directly calculated from BER through

$$Q = \sqrt{2}\text{erfc}^{-1}(2 \times \text{BER}), \tag{5.3}$$

where $\text{erfc}^{-1}(\cdot)$ stands for the inverse function of $\text{erfc}(\cdot)$ function. The Q-factor corresponding to the enhanced FEC limit (2×10^{-3}) is also plotted for reference. These experimental results justify that this dual-stage FOE is capable of estimating a frequency offset up to the range of around $[-0.5R_s, 0.6R_s]$, which is almost 4 times the theoretical limit of single FFT-based FOE using Mth-power. This excellent improvement is attributed to the fact that the coarse FOE based on MPE of the Gardner algorithm can well track frequency offset variations. System complexity is determined by counting the number of gates to implement this dual-stage FOE, which requires $T_s/\Delta T \cdot N_s$ multipliers and $T_s/\Delta T \times 1.5N_s$ adders (ΔT is the step of the scanned timing offset, and the calculated phase error is averaged over N_s symbols).

5.2 Joint SPM Compensation

PolMux QPSK with coherent detection is regarded as a promising technology for upgrading optical networks up to 100 Gb/s per channel. Since coherent detection can

retain the full electric field information of received optical signals, those linear distortions accumulated in the transmission links can be totally compensated for by using DSP techniques [21, 30, 115]. This leads to intensive investigation on pre- and post-electrical CD compensation via linear equalizers [157]. On the other hand, fiber SPM effect limits the performance of such linear equalizers in long-haul transmission systems [47]. Therfore, SPM pre- and post-compensation techniques have been individually proposed to reduce fiber nonlinearity effect [50–54]. As for pre-SPM compensation (pre-SPMC), transmitted signals are pre-distorted by calculating the inverse NLSE, such that trying to restore the waveform of received signals at the receiver side. Post-SPM compensation (post-SPMC), well-known as back-propagation [47, 118], is carried out at the received side by solving the inverse NLSE to estimate the transmitted signals. When solving NLSE with many sections, noises will cause the solution of NLSE to diverge from the actual input signals. In this section, we propose a SPMC at both the transmitter and receiver sides, such that mitigating this divergence problem [158]. Our results show that the proposed joint-SPMC can increase dynamic power range compared to individual pre-SPMC and post-SPMC scheme. Due to the limited resource, we were unable to perform the experiments to verify our proposed joint-SPMC algorithm.

5.2.1 Principle of Joint SPM Compensation

In a PolMux fiber transmission system, two orthogonal polarization components E_x and E_y are governed by a coupled NLSE [78]:

$$\frac{\partial E_{x(y)}}{\partial z} + j\frac{\beta_2}{2}\frac{\partial^2 E_{x(y)}}{\partial t^2} + \frac{\alpha}{2}E_{x(y)} = j\gamma\left[|E_{x(y)}|^2 + \frac{2}{3}|E_{y(x)}|^2\right] \qquad (5.4)$$

Recalling that split-step method is applied to simulate a pulse propagation over fibers, in order to compensate for fiber nonlinearity and dispersion, it is feasible to use the same method to solve an inverse NLSE at either transmitter or receiver side [50–54].

5.2 Joint SPM Compensation

A proposed joint-SPMC here is to compensate for fiber nonlinearity simultaneously at transmitter and receiver sides. The desired PolMux-QPSK signals $E_{x(y)}(t, L)$, that have no CD and PMD, can be located at any point along the transmission link. In other words, a pre-SPMC at the transmitter side is aiming to address fiber dispersion and nonlinearity in the first N_{tx} spans. Meanwhile, fiber nonlinearity and dispersion in the remaining N_{rx} spans are compensated for by using a post-SPMC at the receiver side [158]. For example, if the desired signal $E_{x(y)}(t, L)$ is at the transmitter side ($L = 0$), and the whole fiber nonlinearity and dispersion are compensated in the receiver, it is called back-propagation. The parameters N_{tx} and N_{rx} need to be optimized in order to achieve a better performance. The advantage of this proposed joint-SPMC is to reduce the divergence when solving NLSE without knowing exact noises at those locations of optical amplifiers. A coarse-step computation of inverse NLSE is applied in the simulation for fast and simple calculation [47].

The pre-equalized optical waveform generated by this joint-SPMC at the transmitter (see Figure 5.8 (a)) is computed as follows: the target optical signals $E_{x(y)}(t, N_{tx})$ after N_{tx}-span transmission are used to compute the signals at the end of $N_{tx} - 1$ span based on a coarse-step split-step method, as given by [158]

$$E_{x(y)}(t, N_{tx} - 1) = \mathbb{F}^{-1} \left\{ \mathbb{F} \left[E_{x(y)}(t, N_{tx}) \cdot \exp\left(-jD(f)\right) \right] \right\} \tag{5.5}$$

$$E_{x(y)}(t, N_{tx} - 1) = E_{x(y)}(t, N_{tx} - 1) \cdot$$
$$\exp\left[-j\xi_{tx}\gamma L_{\text{eff}} \left(|E_{x(y)}(t, N_{tx} - 1)|^2 + \frac{2}{3}|E_{y(x)}(t, N_{tx} - 1)|^2 \right) \right] \tag{5.6}$$

Here, $\mathbb{F}$ and $\mathbb{F}^{-1}$ represent the Fourier and inverse Fourier transformations, $D(f)$ is a dispersion function related to the considered fiber span, L_{eff} is the effective fiber length in each span, and ξ_{tx} is a scaling factor ranging from 0 to 1. This procedure will be repeatedly calculated until the pre-compensated signals are generated at $L = 0$. Further, the driver signals of an IQ modulator are generated by inverting its transfer

function to obtain the exact electric field signals at the output of the IQ modulator [71], as given by

$$I_{x(y)} = \arccos\left[\frac{\mathbf{Re}\left[E_{x(y)}(t,0)\right]}{\sqrt{P_t}}\right] \cdot \frac{2V_\pi}{\pi}; \tag{5.7}$$

$$Q_{x(y)} = \arccos\left[\frac{\mathbf{Im}\left[E_{x(y)}(t,0)\right]}{\sqrt{P_t}}\right] \cdot \frac{2V_\pi}{\pi}. \tag{5.8}$$

Note that the driving voltages for the inverse equation of the IQ-modulator in VPI are only half of the voltages given in Eq.(5.7) because V_π there is defined as the voltages leading to π-phase change.

The other joint-SPMC in the receiver (Figure 5.8 (b)) is carried out using the same method as post-SPMC in Eqs.(5.5) and (5.6). Note that ξ_{tx} is replaced by ξ_{rx} to optimize the performance. The signals at the receiver end have been sampled in coherent receivers, and are then back-propagated to estimate the signals at the end of N_{tx} spans.

5.2.2 Simulation

The proposed joint-SPMC scheme was evaluated in a single-channel 112 Gb/s PolMux-QPSK system, where high-speed systems are more vulnerable to fiber nonlinearity [13]. The transmitted sequences for those two polarization states were first fed into a joint-SPMC to pre-compensate for SPM and CD in the first N_{tx}-span transmission link, as illustrated in Figure 5.8 (c). The transmission link consisted of 20 × 100 km SMF with dispersion coefficient 17 ps/nm/km, fiber loss 0.2 dB/km, and fiber nonlinearity 1.31 W^{-1}km^{-1}. An optical amplifier fully compensates for fiber loss in each span. Its noise figure was 5 dB. Laser linewidths for transmitter and LO lasers were 100 kHz each, and ADC and DAC were assumed to be ideal. At the receiver end, the other joint-SPMC for the remaining N_{rx} span was conducted before polarization de-multiplexing, carrier phase recovery and data decoding. Note that $N_{tx} + N_{rx} = 20$

(a)

(b)

(c)
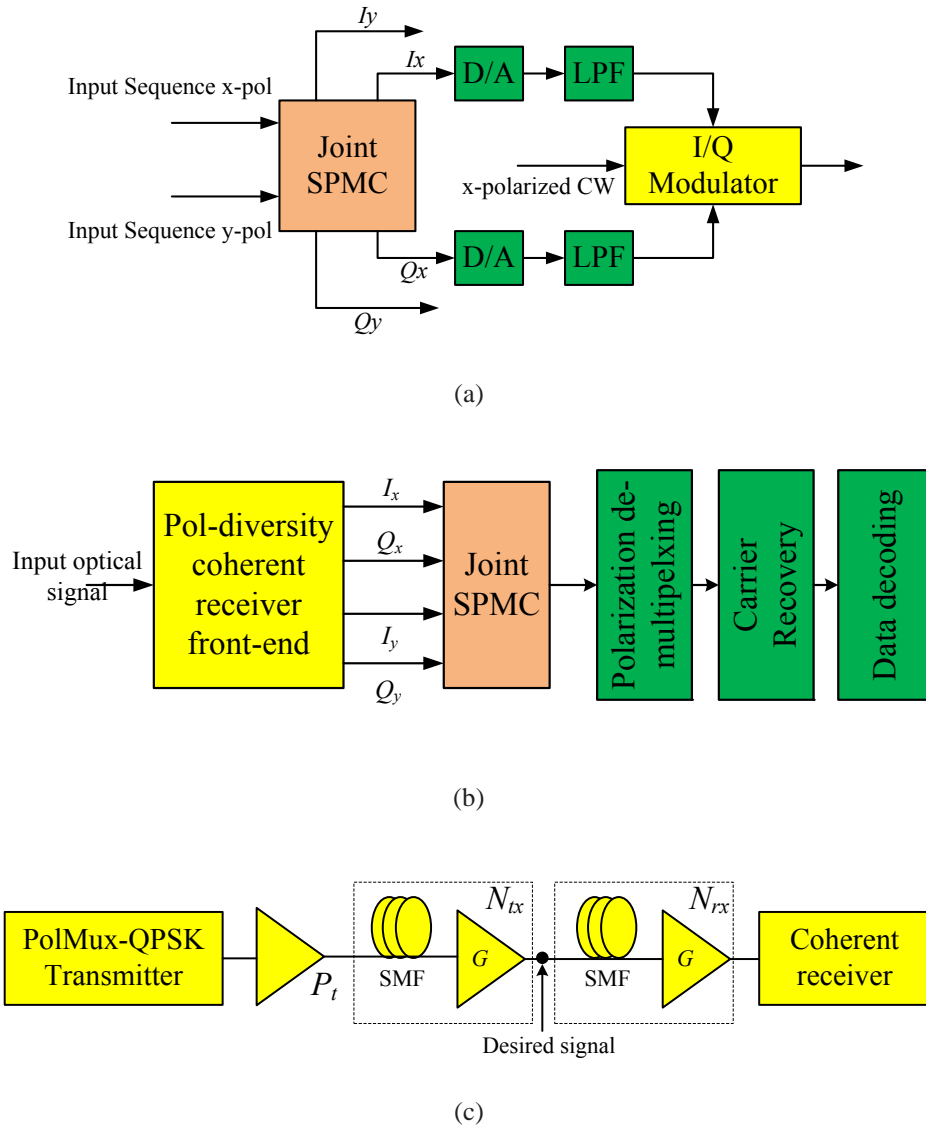
Figure 5.8: (a) The transmitter structure of the joint-SPMC; (b) the receiver structure of the joint-SPMC; (c) simulated system setup where desired signals are located in the middle of the link instead of at the end of the link.
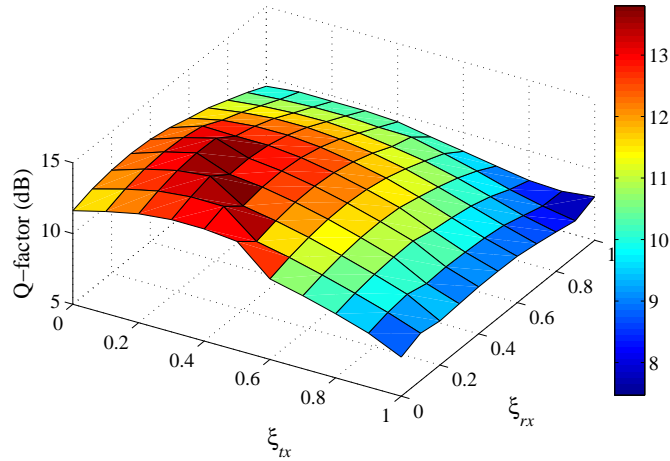
Figure 5.9: Q-factor as a function of ξ_{tx} and ξ_{rx} using the joint SPMC (N_{tx}=10, N_{rx}=10, P_t=5 dBm).

in simulation.
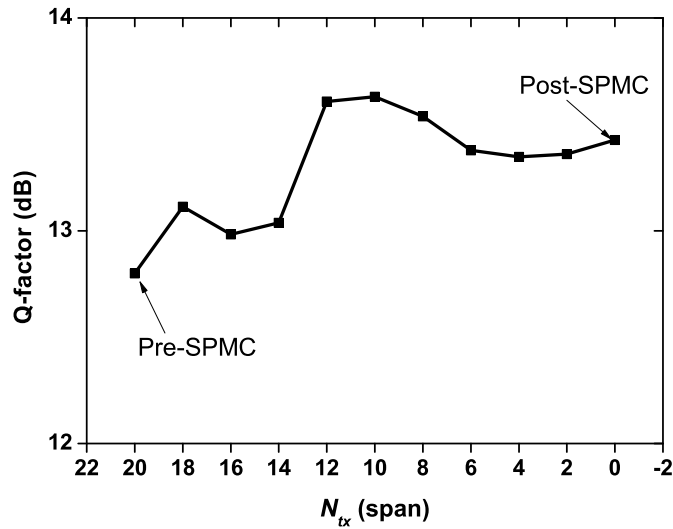
Figure 5.10: Q-factor versus N_{tx} and N_{rx} ($N_{tx} + N_{rx} = 20$, P_t=5 dBm.)

To determine the optimal scaling factor ξ_{tx} and ξ_{rx}, we scanned their values from 0 to 1 by setting $N_{tx} = N_{rx} = 10$ spans in the simulated system. The Q-factor was obtained by using the same approach as reported in [159]. As shown in Figure 5.9, the Q-factor was improved by 2.7 dB at $\xi_{tx} = 0.3$ and $\xi_{rx} = 0.5$, compared to no SPM compensation scenario. The scaling factors were fixed to their optimal values in all the

subsequent simulations.

(a)   (b)   (c)
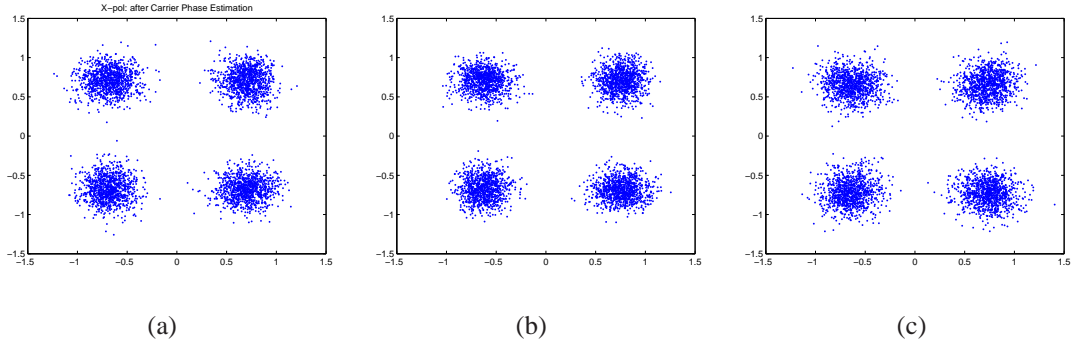
Figure 5.11: Constellation map of the recovered signals in a 20×100km transmission system. (a) Pre-SPMC (N_{tx}=20, $N_{rx} = 0$); (b) Joint SPMC($N_{tx} = 10$, $N_{rx} = 10$); (c) Post-SPMC ($N_{tx} = 0$, $N_{rx} = 20$). P_t=5 dBm.

The parameters N_{tx} and N_{rx} were also swept to find the optimal solution to the joint-SPMC. As displayed in Figure 5.10, the optimal span for pre-SPMC and post-SPMC was suggested to be around half of the total length. Note that these simulation were also found to be quite similar at other input powers. Figure 5.11 shows recovered constellation maps by using different SPMC schemes. It can be observed that the proposed joint SPM can outperform the pre- and post-SPMC by 1.1 dB and 0.2 dB in Q-factor improvement, respectively. This may be due to the fact that the joint-SPMC is able to mitigate the divergence problem when solving the inverse NLSE through splitting it into two parts: pre- and post-SPMC. Figure 5.12 illustrates the Q-factor performance of these three SPMC schemes. The input power dynamic range is defined as the launch power allowing BER less than FEC limit (10^{-3}) [160]. The dynamic range of the proposed joint-SPMC can be as large as 10.2 dB. Compared to pre-SPMC (9.8 dB) and post-SPMC scheme (9 dB), the joint-SPMC scheme increased the dynamic input power range by 0.4 dB and 1.2 dB, respectively. In addition, pre-SPMC scheme cannot perform as good as joint-SPMC and post-SPMC since its inverse NLSE
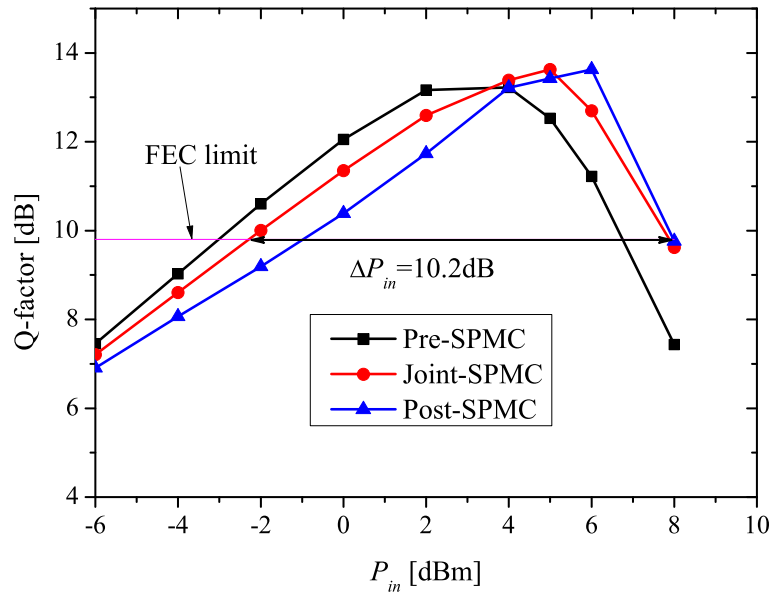
Figure 5.12: Performance comparison among the three SPMC schemes: pre-SPMC (N_{tx}=20), joint-SPMC (N_{tx} = 10, N_{rx}=10) and post-SPMC (N_{rx} =20).

computation does not take into account system noises added by optical amplifiers. Although the improvement in dynamic input power range is less than 1 dB, we believe that this improvement will be even larger when more steps are performed in each span instead of one step each span during carrying out SPMC. Besides, the analog transmitters has been gradually replaced by digital transmitter, thus the pre-SPMC can be easily implemented through look-up tables.

5.3 Conclusion

Frequency offset and fiber nonlinearity would cause carrier phaser noise, which are detrimental to system's performance as well. It is desirable to apply independent DSP modules to address these problems separately. An FOE based on Gardner timing recovery algorithm was proposed to deal with frequency offset. A underlying relationship between normalized MPE and frequency offset was utilized to obtain a coarse estimate of frequency offset. The estimation range was found to be from 9 GHz to -9

GHz in a simulated 40 Gbit/s PolMux-QPSK system. The residual frequency offset was further addressed by a cascaded fine FOE based on M-power and FFT operations. This dual-stage cascaded FOE was experimentally demonstrated in a 42.8 Gbit/s coherent PolMux-RZ-QPSK system. The estimation range was measured to be [-5 GHz, 6 GHz], where the estimation range of conventional FOE using Mth-power is limited within [-1.34 GHz, 1.34 GHz].

As for fiber nonlinearity, SPM in a single channel was simulated in a 112 Gbit/s PolMux-QPSK system to investigate its impact. Back-propagation algorithms are mostly investigated at either transmitter or receiver side, though it is susceptible to divergence when solving inverse NLSE. In our proposed joint-SPMC approach, we split NLSE into two parts: fiber nonlinearity in the first N_{tx}-span was pre-compensated at the transmitter side while the remaining fiber nonlinearity in the N_{rx}-span was post-compensated in coherent receivers. It was found that this joint-SPMC scheme can enhance the tolerance of fiber nonlinearity in the simulated system.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

In this thesis, we have conducted a systematic study on DA ML and adaptive DA PE algorithms in M-PSK/QAM formats. Besides, the issues of frequency offset and nonlinear phase noise were addressed by using those proposed DSP algorithms.

The performances of DA ML were investigated in different modulation formats. It shows that this computationally-linear DA ML PE algorithm performs much better than the conventional V&V Mth-power methods in M-QAM formats. However, like Mth-power algorithm, it suffers from BLE due to the average over additive noises and laser phase noise. The phase error variance of DA ML PE was analytically derived to illustrate this trade-off in selecting the optimal memory length L. In addition, the optimal memory length can be determined if the statistics of phase noise and additive noise are known to the receivers. On the other hand, the feedback in DA ML PE actually limits its application in a real-time coherent receiver, in which a parallel structure is desirable to reduce receiver cost and complexity. In view of those required features, a parallel structure of DA ML algorithm was proposed to make itself operate online, where it has a reasonable performance in both BPSK and QPSK formats. Finally, an

experiment was demonstrated that DA ML algorithm has the same performance as the V&V Mth-power method but requires less computational efforts.

To avoid BLE in DA ML algorithm, a first-order filter, called the adaptive DA algorithm, was introduced to adaptively adjust this filter gain, thus mitigating or even eliminating BLE phenomenon. The adaptive DA algorithm was extended to M-QAM formats, where it was found that its performance was degraded by constellation penalty. Compared to DA ML algorithm, the adaptive DA algorithm is more suitable in constant-modulus PSK formats. It is also observed that the self-adaptation capability makes the adaptive DA algorithm even more attractive in M-QAM formats without spending amounts of time in finding the optimal memory length. Further, we examined another two adaptive PE algorithms, NLMS and the Kalman filter, through a similar analysis as deriving the phase error variance of the adaptive DA algorithm. Of interest is that these three adaptive DA algorithms share the same performances but differentiates from each other in the approach of determining their stepsizes. The adaptive DA algorithm was then experimentally demonstrated in a long-haul PolMux coherent QPSK system. It was found that the adaptive DA performed similarly as the V&V Mth-power method in the presence of fiber nonlinearity while DA ML and block Mth-power algorithms had the same performance.

Asides from laser phase noise, frequency offset and fiber nonlinearity are also those factors causing carrier phase noise, thus requiring additional DSP methods to compensate before applying PE algorithms. Based on Gardner timing algorithm, we proposed a dual-stage cascaded FOE whose estimation range is as large as [-9 GHz, 9 GHz] in a simulated 40 Gbit/s PolMux-QPSK system. In a 42.8 Gbit/s coherent PolMux-QPSK system, the experimentally measured estimation range reduced to [-5 GHz, 6 GHz] because of the imperfect experiments setup. As for nonlinear phase noise compensation, the proposed joint-SPMC was verified in simulation to increase the dynamic power range at BER= 10^{-3} by 0.4 dB and 1.2 dB, compared to separate

pre- and post-SPMC scheme.

6.2 Future Work

The following areas, from my point of view, may become the future hot topics in coherent optical communication systems.

6.2.1 Estimation of the SNR and Phase Noise Variance

As demonstrated in Chapter 3 and 4, the optimal operation of PE algorithms is related to the statistics of additive noise and phase noise, though these knowledge may be not available at the receiver side. It is desirable to estimate those parameters through those received signals, thus adjusting the memory length L of DA ML and Mth-power methods, or computing the optimal stepsize in the adaptive DA algorithm. Besides, the estimated SNR and phase noise variance are helpful for designing the Wiener filter, which is shown to be an optimal minimum mean-square-error filter capable of approaching the theoretical performance [4, 151].

6.2.2 Joint Equalization and Phase Estimation

Nowadays, researches have pursued a standard to achieve higher data rate (112 G-bit/s [24, 120], 400 Gbit/s [161] or even 1 Tbit/s [55]) to meet the increasingly demand on traffic bandwidth. The bottleneck is the speed of electronic devices that are far less than data rate [141]. Alternatively, higher-order modulation formats have become candidates to support such ultra-high-speed transmissions, such as QAM formats [162, 163]. In a PolMux M-QAM format, polarization de-multiplexing butterfly filters using DD-LMS algorithm and PE need to be conducted simultaneously due to the non-constant-modulus of QAM signals [120]. In M-PSK format, CMA algorithm

is sufficient to fully de-multiplex the signals of the two polarization states. Instead, DD-LMS should be applied in QAM formats to fully equalize the cross-talk between these two polarization states. Nonetheless, DD-LMS algorithm is sensitive to carrier phase of the equalized signals. Of the importance is the issue that how to incorporate PE modules into DD-LMS algorithm in order to obtain a better performance [128].

6.2.3 Mode Multiplexing

Recently, the successful demonstration of higher-order mode to increase fiber nonlinearity tolerance has triggered intense research on few-mode multiplexing/demultiplexing technologies. How to separate those different modes at the transmitter side is a challenge issue, because the inter-mode delay is too large thus requiring a huge amount of taps to de-multiplex those modes. Therefore, a reasonable DSP algorithm is necessary to keep the de-multiplexing filters within certain limit of complexity for practical implementation.

6.2.4 Coding

The advance of high-speed ADC and DSP processors enables more and more signal processing techniques to be employed in coherent optical communication systems. Specifically, coding techniques provide a great advantage in the tolerance of fiber nonlinearity and channel capacity [126, 164]. Conventionally, only 7% overhead is used for FEC coding which may limit its performance improvement. It has been suggested increasing the overhead of FEC coding up to 20% for further performance improvement [165]. In addition, more advanced coding techniques, such as LDPC and Turbo codes, are still very powerful methods for increasing FEC limit from 10^{-3} to even higher [166].

Appendix A

Channel Capacity

In a complex AWGN channel (see Figure 2.2), the channel input X and output Y are complex random variables with

$$Y = X + N, \tag{A.1}$$

where N is the complex AWGN with a circular Gaussian distribution of mean zero and variance σ_n^2. Due to the independence between the channel input X and noise N, the conditional PDF of the channel output can be written as

$$p_{Y|X}(y|x) = \frac{1}{\pi\sigma_n^2} \exp\left[-\frac{|y-x|^2}{\sigma_n^2}\right]. \tag{A.2}$$

The mutual information of the channel input X and output Y is given by [65]

$$
\begin{aligned}
I(X;Y) &= \int\int p_{XY}(x,y)\log_2\left(\frac{p_{Y|X}(y|x)}{p_Y(y)}\right)dxdy \\
&= H(Y) - H(Y|X) \\
&= H(Y) - H(X+N|X) \\
&= H(Y) - H(N) \\
&= H(Y) - \log_2\left(\pi e\sigma_n^2\right),
\end{aligned}
\tag{A.3}
$$

where $H(Y)$ and $H(Y|X)$, respectively, represent the entropy of the random variable Y and the average conditional entropy. The channel capacity in bits per symbol is

A. Channel Capacity

defined as [64]

$$C = \max_{E[\|X\|^2] \leq P} I(X; Y), \tag{A.4}$$

subjected to the input power constraint $E[\|X\|^2] \leq P$.

Research works suggest that it is suffices to consider digitized to approach the ultimate capacity limits of noisy channels by using Shannon's information theory [13]. As can be observed from Eq.(A.3), the PDF $p_Y(y)$ needs to be derived in order to compute the channel capacity. It can be shown that [131]

$$p_Y(y) = \sum_{x_i \in X}^{i=0,\cdots,M-1} P_X(x_i) p_{Y|X}(y|x_i), \tag{A.5}$$

where $P_X(x_i)$ is the probability that the random variable X takes the value x_i, and M refers to the alphabet size in the given modulation format. Therefore, the capacity is a maximization of $H(Y)$ over $P_X(x_i)$, $i = 0, \cdots, M-1$. In fact, the optimal probability distribution turns out to be $P_X(x_i) = 1/M$ for all i [65], thus yielding

$$p_Y(y) = \frac{1}{M\pi\sigma_n^2} \sum_{x_i \in X}^{i=0,\cdots,M-1} \exp\left[-\frac{|y-x_i|^2}{\sigma_n^2}\right]. \tag{A.6}$$

The entropy $H(Y)$ can be evaluated through

$$H(Y) = E\left[-\log_2\left(\frac{1}{M\pi\sigma_n^2} \sum_{x_i \in X}^{i=0,\cdots,M-1} \exp\left(-\frac{|y-x_i|^2}{\sigma_n^2}\right)\right)\right]. \tag{A.7}$$

To numerically compute the entropy $H(Y)$ in Eq.(A.7), we follow the similar methods described in [13]. One million symbols are uniformly chosen over the considered modulation formats with alphabet size M, and are then added by complex Gaussian noises to emulate the AWGN channel. Finally, we average over the received signals to obtain the ultimate channel capacity from Eq.(A.3).

Appendix B

Derivation of the log-likelihood function $L(\theta, k) = \ln \Lambda(\theta, k)$

The likelihood function $\Lambda(\theta, k)$ is given by

$$\Lambda(\theta, k) = p(r(k - L), \cdots, r(k - 1)|\theta) \tag{B.1}$$

As explained in 3.1 of Chapter 3, if $i \neq j$, $r(l)$ and $r(j)$ are independent due to the independence of $n(l)$ and $n(j)$. So (B.1) can be rewritten as products of these PDFs [131]:

$$\Lambda(\theta, k) = \prod_{l=k-M}^{k-1} p(r(l)|\theta) \tag{B.2}$$

By using the law of total probability and PDF of $r(k)$ (3.4), the likelihood function is expressed as

$$
\begin{aligned}
\Lambda(\theta, k) &= \prod_{l=k-M}^{k-1} \sum_{i=-M/2}^{M/2} p(r(l)|\theta, m_i = C_i) P(C_i) \\
&= \prod_{l=k-M}^{k-1} \sum_{i=-M/2}^{M/2} \frac{1}{M\pi N_0} \exp\left[-\frac{|r(l) - C_i e^{j\,\theta}|^2}{N_0}\right] \\
&= \prod_{l=k-M}^{k-1} \sum_{i=-M/2}^{M/2} \frac{1}{M\pi N_0} \exp\left[-\frac{|r_l|^2 + |C_i|^2}{N_0}\right] \exp\left[\frac{2\mathbf{Re}[r(l)C_i^* e^{-j\theta}]}{N_0}\right]
\end{aligned}
$$

B. Derivation of the log-likelihood function $L(\theta, k) = \ln \Lambda(\theta, k)$

$$= \prod_{l=k-M}^{k-1} \frac{1}{M\pi N_0} \exp\left[-\frac{|r_l|^2}{N_0}\right] \sum_{i=-M/2}^{M/2} \exp\left[-\frac{|C_i|^2}{N_0}\right] \exp\left[\frac{2\mathbf{Re}[r(l)C_i^* e^{-j\theta}]}{N_0}\right]$$

(B.3)

Here, $P(C_i) = 1/M$ because all symbols are equiprobable. In addition, $C_i = -C_{-i}$, so we can combine those C_i and C_{-i} terms in (B.3) into $\cosh(\cdot)$[1].

$$\Lambda(\theta, k) = \prod_{l=k-M}^{k-1} \frac{1}{M\pi N_0} \exp\left[-\frac{|r_l|^2}{N_0}\right] \sum_{i=1}^{M/2} \exp\left[-\frac{|C_i|^2}{N_0}\right] \cosh\left[\frac{2\mathbf{Re}[r(l)C_i^* e^{-j\theta}]}{N_0}\right]$$

(B.4)

Therefore, the log-likelihood function $L(\theta, k) = \ln \Lambda(\theta, k)$ can now be written as

$$
\begin{aligned}
L(\theta, k) &= \sum_{l=k-M}^{k-1} \ln\left[\sum_{i=1}^{M/2} \exp\left[-\frac{|C_i|^2}{N_0}\right] \cosh\left[\frac{2\mathbf{Re}[r(l)C_i^* e^{-j\theta}]}{N_0}\right]\right] \\
&+ \sum_{l=k-M}^{k-1} \ln\left[\frac{1}{M\pi N_0} \exp\left[-\frac{|r_l|^2}{N_0}\right]\right]
\end{aligned}
$$

(B.5)

Since the received signals $\{r(l), k - L \leq l \leq k - 1\}$ are observed quantities and independent of θ and i, the second term on the right-hand side of (B.5) can be treated as a constant c. To make the equation shorter, let $S_i = |C_i|^2/N_0$, $q_i(l, \theta) = (2/N_0)\mathbf{Re}[r(l)C_i^* e^{-j\theta}]$. Now, the log-likelihood function $L(\theta, k)$ reduces to

$$L(\theta, k) = \sum_{l=k-L}^{k-1} \ln\left[\sum_{i=1}^{M/2} \exp(-S_i) \cosh q_i(l, \theta)\right] + c$$

(B.6)

---

[1] $\cosh(x) = \frac{e^x + e^{-x}}{2}$

126

Appendix C

BER of 16-PSK/QAM in the Presence of Phase Error

C.1   BER in 16PSK
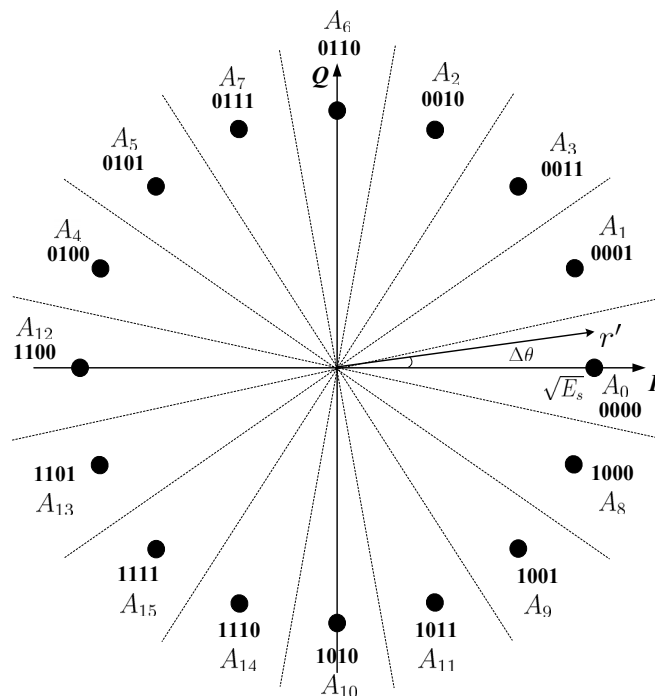
Figure C.1: Constellation map and gray bit mapping of 16PSK.

C.1 BER in 16PSK

Although BERs conditioned on different transmitted symbols in 16PSK are not the same, the trivial difference especially at high SNR [167] can guarantee that our analysis here is sufficiently accurate for investigating its BER. Constellation map and bit mapping of 16PSK are given in Fig. C.1, where Gray coding is used. Supposing that "0000" is transmitted, the BER conditioned on a phase error $\Delta\theta$ can be written as

$$
\begin{aligned}
P_b(e|\Delta\theta) = \frac{1}{4}\Big[&P(A_1|\Delta\theta) + 2P(A_3|\Delta\theta) + P(A_2|\Delta\theta) + 2P(A_6|\Delta\theta) + 3P(A_7|\Delta\theta) \\
&+ 2P(A_5|\Delta\theta) + P(A_4|\Delta\theta) + 2P(A_{12}|\Delta\theta) + 3P(A_{13}|\Delta\theta) \\
&+ 4P(A_{15}|\Delta\theta) + 3P(A_{14}|\Delta\theta) + 2P(A_{10}|\Delta\theta) + 3P(A_{11}|\Delta\theta) \\
&+ 2P(A_9|\Delta\theta) + P(A_8|\Delta\theta)\Big],
\end{aligned}
\tag{C.1}
$$

where each A_i denotes a decision region, and $P(A_i|\Delta\theta)$ is the probability of the received signal vector falling in the decision region A_i. The order $\{A1, A3, A2, A6\cdots\}$ is arranged in the anti-clockwise direction as shown in Fig. C.1 for easier understanding and derivation rather than using the order $\{A1, A2, A3\cdots\}$. Observing that $U_1 = A_1 \bigcup A_3 \bigcup A_2 \bigcup A_6 \bigcup A_7 \bigcup A_5 \bigcup A_4 \bigcup A_{12}$ and $U_2 = A_{12} \bigcup A_{13} \bigcup A_{15} \bigcup A_{14} \bigcup A_{10} \bigcup A_{11} \bigcup A_9 \bigcup A_8$, are half-planes, and thus Eq.(C.1) can be rewritten as

$$
\begin{aligned}
P_b(e|\Delta\theta) = \frac{1}{4}\Big[&P(U_1|\Delta\theta) + P(U_2|\Delta\theta) + P(A_3|\Delta\theta) + P(A_6|\Delta\theta) + 2P(A_7|\Delta\theta) \\
&+ P(A_5|\Delta\theta) + 2P(A_{13}|\Delta\theta) + 3P(A_{15}|\Delta\theta) + 2P(A_{14}|\Delta\theta) \\
&+ P(A_{10}|\Delta\theta) + 2P(A_{11}|\Delta\theta) + P(A_9|\Delta\theta)\Big].
\end{aligned}
\tag{C.2}
$$

Similar to the method of deriving the BER of 8PSK as in [139], Eq.(C.2) can be simplified into

$$
\begin{aligned}
P_b(e|\Delta\theta) + P_b(e|-\Delta\theta) = \frac{1}{2}\Big[&P(U_1|\Delta\theta) + P(U_2|\Delta\theta) + P(U_3|\Delta\theta) + P(U_4|\Delta\theta) \\
&+ P(A_3 + A_2|\Delta\theta) + P(A_7 + A_5|\Delta\theta) \\
&+ P(A_{15} + A_{14}|\Delta\theta) + P(A_{11} + A_9|\Delta\theta)\Big],
\end{aligned}
\tag{C.3}
$$

C.1 BER in 16PSK

where both $U_3 = A_6 \bigcup A_7 \bigcup A_5 \bigcup A_4$ and $U_4 = A_{13} \bigcup A_{15} \bigcup A_{14} \bigcup A_{10}$ are quadrants.

It is easy to show that

$$P(U_1|\Delta\theta) = \frac{1}{2}\text{erfc}\left[\sqrt{\gamma_s}\sin\left(\frac{\pi}{16} - \Delta\theta\right)\right], \tag{C.4a}$$

$$P(U_2|\Delta\theta) = \frac{1}{2}\text{erfc}\left[\sqrt{\gamma_s}\sin\left(\frac{\pi}{16} + \Delta\theta\right)\right], \tag{C.4b}$$

$$P(U_3|\Delta\theta) = \frac{1}{4}\text{erfc}\left[-\sqrt{\gamma_s}\sin\left(\frac{\pi}{16} + \Delta\theta\right)\right] \cdot \text{erfc}\left[\sqrt{\gamma_s}\cos\left(\frac{\pi}{16} + \Delta\theta\right)\right], \tag{C.4c}$$

$$P(U_4|\Delta\theta) = \frac{1}{4}\text{erfc}\left[-\sqrt{\gamma_s}\sin\left(\frac{\pi}{16} - \Delta\theta\right)\right] \cdot \text{erfc}\left[\sqrt{\gamma_s}\cos\left(\frac{\pi}{16} - \Delta\theta\right)\right]. \tag{C.4d}$$

Now we derive those remaining terms in Eq.(C.3). $P(A_3 + A_2|\Delta\theta)$ can be expressed as

$$
\begin{aligned}
P(A_3 + A_2|\Delta\theta) &= P\left(x' \geq 0, y_0 \leq y' \leq y_0 + x'\right) \\
&= \int_0^{+\infty} \frac{1}{\sqrt{\pi N_0}} \exp\left[-\frac{(x'-x_0)^2}{N_0}\right] dx' \cdot \int_{y_0}^{y_0+x'} \frac{1}{\sqrt{\pi N_0}} \exp\left[-\frac{y'^2}{N_0}\right] dy' \\
&= \frac{1}{2}\int_0^{+\infty}\left[\text{erfc}\left(\frac{y_0}{\sqrt{N_0}}\right) - \text{erfc}\left(\frac{y_0+x'}{\sqrt{N_0}}\right)\right] f(x'-x_0)dx', \tag{C.5}
\end{aligned}
$$

where $x_0 = \sqrt{E_s}\cos(3\pi/16 - \Delta\theta)$, $y_0 = \sqrt{E_s}\sin(3\pi/16 - \Delta\theta)$, and $f(x) = \frac{1}{\sqrt{\pi N_0}}$ $\cdot \exp(-t^2/N_0)$. Similar to the derivation in $P(A_3 + A_2|\Delta\theta)$, the other terms can also be obtained as follows:

$$P(A_9 + A_{11}|\Delta\theta) = \frac{1}{2}\int_0^{+\infty}\left[\text{erfc}\left(\frac{y_1}{\sqrt{N_0}}\right) - \text{erfc}\left(\frac{y_1+x'}{\sqrt{N_0}}\right)\right] \cdot f(x'-x_1)dx', \tag{C.6}$$

$$P(A_{14} + A_{15}|\Delta\theta) = \frac{1}{2}\int_0^{+\infty}\left[\text{erfc}\left(\frac{y_2}{\sqrt{N_0}}\right) - \text{erfc}\left(\frac{y_2+x'}{\sqrt{N_0}}\right)\right] \cdot f(x'+x_2)dx', \tag{C.7}$$

$$P(A_5 + A_7|\Delta\theta) = \frac{1}{2}\int_0^{+\infty}\left[\text{erfc}\left(\frac{y_3}{\sqrt{N_0}}\right) - \text{erfc}\left(\frac{y_3+x'}{\sqrt{N_0}}\right)\right] \cdot f(x'+x_3)dx', \tag{C.8}$$

where $x_1 = \sqrt{E_s}\cos(3\pi/16 + \Delta\theta)$, $y_1 = \sqrt{E_s}\sin(3\pi/16 + \Delta\theta)$, $x_2 = \sqrt{E_s}\sin(\pi/16$ $\Delta\theta)$, $y_2 = \sqrt{E_s}\cos(\pi/16 + \Delta\theta)$, $x_3 = \sqrt{E_s}\sin(\pi/16 - \Delta\theta)$, and $y_3 = \sqrt{E_s}\cos(\pi/16$ $-\Delta\theta)$. Therefore, the BER of 16PSK conditioned on the phase error $\Delta\theta$ is given from Eqs.(C.3) to (C.8). The final expression of $P_b(e|\Delta\theta) + P_b(e|-\Delta\theta)$ for 16PSK is omitted to save space.

C.2 BER in 16QAM

The signal constellation and bit mapping of 16QAM are shown in Fig. C.2, where each signal point is represented by a 4-bit symbol (i_1, q_1, i_2, q_2). Data bit stream is split into inphase (I) and quadrature (Q) bit streams. The I and Q components together are mapped to complex symbols using Gray coding. Bits i_1 and q_1 are the most significant bits (MSB) of I and Q, and bits i_2 and q_2 are their least significant bits (LSB) [168]. The average symbol energy is $\bar{E}_s = 10d^2$. Note that the two bit values of each I and Q are equally likely.
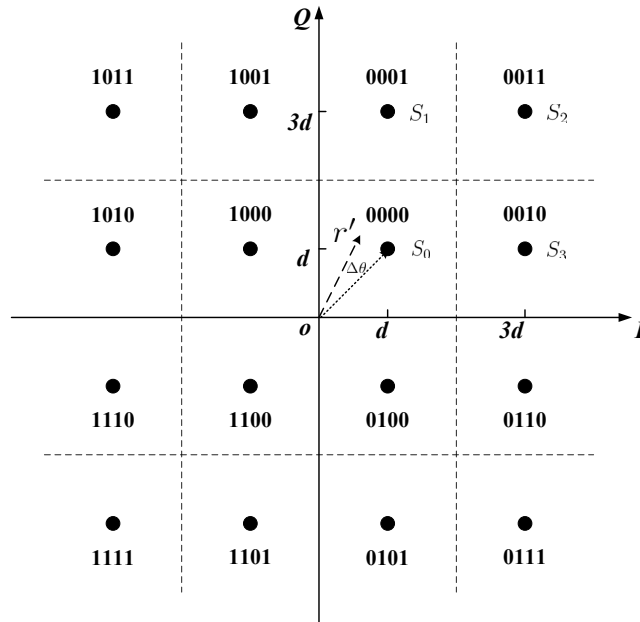
Figure C.2: Constellation map and bit mapping of 16QAM.

Due to the symmetry of signal constellation, as well as the symmetry between MSB and LSB, the average BER conditioned on a phase error $\Delta\theta$ can be represented by [168]

$$P_b(e|\Delta\theta) = \frac{1}{2} \left[P_{MSB}(e|\Delta\theta) + P_{LSB}(e|\Delta\theta) \right], \tag{C.9}$$

which accounts for the BER of either I or Q component. It can be observed that the probability of MSB bit error conditioned on $i_1 = 0$ is the same as that conditioned on

$i_1 = 1$. Without loss of generality, we use the four signal points located in the first quadrant, i.e., S_0, S_1, S_2, and S_3. It is easy to show that

$$P_{MSB}(e|S_0, \Delta\theta) = \frac{1}{2}\text{erfc}\left(\frac{\sqrt{2}d \cdot \cos(\pi/4 + \Delta\theta)}{\sqrt{N_0}}\right), \tag{C.10a}$$

$$P_{MSB}(e|S_1, \Delta\theta) = \frac{1}{2}\text{erfc}\left(\frac{\sqrt{10}d \cdot \cos(\arctan(3) + \Delta\theta)}{\sqrt{N_0}}\right), \tag{C.10b}$$

$$P_{MSB}(e|S_2, \Delta\theta) = \frac{1}{2}\text{erfc}\left(\frac{\sqrt{18}d \cdot \cos(\pi/4 + \Delta\theta)}{\sqrt{N_0}}\right), \tag{C.10c}$$

$$P_{MSB}(e|S_3, \Delta\theta) = \frac{1}{2}\text{erfc}\left(\frac{\sqrt{10}d \cdot \cos(\arctan(1/3) + \Delta\theta)}{\sqrt{N_0}}\right). \tag{C.10d}$$

Thus, $P_{MSB}(e|\Delta\theta)$ can be computed from $P_{MSB}(e|\Delta\theta) = [P_{MSB}(e|S_0, \Delta\theta) + P_{MSB}(e|S_1, \Delta\theta) + P_{MSB}(e|S_2, \Delta\theta) + P_{MSB}(e|S_3, \Delta\theta)]/4$. Now the LSB bit error can be dealt with using a similar approach as the MSB bit error, though $P_{LSB=0}(e, |\Delta\theta) \neq P_{LSB=1}(e|\Delta\theta)$. It can be shown that

$$P_{LSB=0}(e|S_0, \Delta\theta) = \frac{1}{2}\text{erfc}\left(\frac{2d + \sqrt{2}d \cdot \cos(\pi/4 + \Delta\theta)}{\sqrt{N_0}}\right)$$
$$+ \frac{1}{2}\text{erfc}\left(\frac{2d - \sqrt{2}d \cdot \cos(\pi/4 + \Delta\theta)}{\sqrt{N_0}}\right), \tag{C.11a}$$

$$P_{LSB=0}(e|S_1, \Delta\theta) = \frac{1}{2}\text{erfc}\left(\frac{2d + \sqrt{10}d \cdot \cos(\arctan(3) + \Delta\theta)}{\sqrt{N_0}}\right)$$
$$+ \frac{1}{2}\text{erfc}\left(\frac{2d - \sqrt{10}d \cdot \cos(\arctan(3) + \Delta\theta)}{\sqrt{N_0}}\right), \tag{C.11b}$$

$$P_{LSB=1}(e|S_3, \Delta\theta) = \frac{1}{2}\text{erfc}\left(\frac{\sqrt{18}d \cdot \cos(\pi/4 + \Delta\theta) - 2d}{\sqrt{N_0}}\right)$$
$$- \frac{1}{2}\text{erfc}\left(\frac{\sqrt{18}d \cdot \cos(\pi/4 + \Delta\theta) + 2d}{\sqrt{N_0}}\right), \tag{C.11c}$$

$$P_{LSB=1}(e|S_2, \Delta\theta) = \frac{1}{2}\text{erfc}\left(\frac{\sqrt{10}d \cdot \cos(\arctan(1/3) + \Delta\theta) - 2d}{\sqrt{N_0}}\right)$$
$$- \frac{1}{2}\text{erfc}\left(\frac{\sqrt{10}d \cdot \cos(\arctan(1/3) + \Delta\theta) + 2d}{\sqrt{N_0}}\right). \tag{C.11d}$$

C.2 BER in 16QAM

$P_{LSB}(e|\Delta\theta)$ can be represented by $P_{LSB}(e|\Delta\theta) = [P_{LSB=0}(e|S_0, \Delta\theta) + P_{LSB=0}(e|S_1,$
$\Delta\theta) + P_{LSB=1}(e|S_2, \Delta\theta) + P_{LSB=1}(e|S_3, \Delta\theta)]/4$.

Combining Eq.(C.10a) through Eq.(C.11d), the BER conditioned on the phase error $\Delta\theta$ in 16QAM is expressed as

$$P_b(e|\Delta\theta) =$$

$$\frac{1}{16}\left[\operatorname{erfc}\left(\frac{\sqrt{2}d \cdot \cos\left(\pi/4 + \Delta\theta\right)}{\sqrt{N_0}} \right) + \operatorname{erfc}\left(\frac{\sqrt{10}d \cdot \cos\left(\arctan\left(3\right) + \Delta\theta\right)}{\sqrt{N_0}} \right) \right.$$

$$+ \operatorname{erfc}\left(\frac{\sqrt{18}d \cdot \cos\left(\pi/4 + \Delta\theta\right)}{\sqrt{N_0}} \right) + \operatorname{erfc}\left(\frac{\sqrt{10}d \cdot \cos\left(\arctan\left(1/3\right) + \Delta\theta\right)}{\sqrt{N_0}} \right)$$

$$+ \operatorname{erfc}\left(\frac{2d + \sqrt{2}d \cdot \cos\left(\pi/4 + \Delta\theta\right)}{\sqrt{N_0}} \right) + \operatorname{erfc}\left(\frac{2d - \sqrt{2}d \cdot \cos\left(\pi/4 + \Delta\theta\right)}{\sqrt{N_0}} \right)$$

$$+ \operatorname{erfc}\left(\frac{2d + \sqrt{10}d \cdot \cos\left(\arctan\left(3\right) + \Delta\theta\right)}{\sqrt{N_0}} \right)$$

$$+ \operatorname{erfc}\left(\frac{2d - \sqrt{10}d \cdot \cos\left(\arctan\left(3\right) + \Delta\theta\right)}{\sqrt{N_0}} \right)$$

$$+ \operatorname{erfc}\left(\frac{\sqrt{18}d \cdot \cos\left(\pi/4 + \Delta\theta\right) - 2d}{\sqrt{N_0}} \right) - \operatorname{erfc}\left(\frac{\sqrt{18}d \cdot \cos\left(\pi/4 + \Delta\theta\right) + 2d}{\sqrt{N_0}} \right)$$

$$+ \operatorname{erfc}\left(\frac{\sqrt{10}d \cdot \cos\left(\arctan\left(1/3\right) + \Delta\theta\right) - 2d}{\sqrt{N_0}} \right)$$

$$\left. - \operatorname{erfc}\left(\frac{\sqrt{10}d \cdot \cos\left(\arctan\left(1/3\right) + \Delta\theta\right) + 2d}{\sqrt{N_0}} \right) \right]. \tag{C.12}$$

Appendix D

Phase Error Variance of DA ML PE with a Matched Filter

When a matched filter is used at the receiver side, the phase error variance Eq.(3.45) can be expanded into

$$
E\left[(\Delta\theta(k))^2\right] = \frac{1}{L^2}\left[E\left[\left(\int_{lT}^{(l+1)T}\frac{\varphi(t)-\theta(l)}{T}dt - \sum_{i=l+1}^{k}\nu(i)\right)^2\right] + \right.
$$
$$
\sum_{l_1=k-L}^{k-1}\sum_{\substack{l_2=k-L \\ l_2\neq l_1}}^{k-1}E\left[\left(\int_{l_1T}^{(l_1+1)T}\frac{\varphi(t)-\theta(l_1)}{T}dt - \sum_{i=l_1+1}^{k}\nu(i)\right)\right.
$$
$$
\left.\left.\cdot\left(\int_{l_2T}^{(l_2+1)T}\frac{\varphi(t)-\theta(l_2)}{T}dt - \sum_{i=l_2+1}^{k}\nu(i)\right)\right]\right].
\tag{D.1}
$$

Applying the important characteristics of Wiener process $\varphi(t)$, $E[\varphi(t)] = 0$ and $E[\varphi(t_1)\varphi(t_2)] = \sigma_0^2\min(t_1,t_2)$ [131], the 1st term in Eq.(D.1) can be shown as:

$$
E\left[\left(\int_{lT}^{(l+1)T}\frac{\varphi(t)-\theta(l)}{T}dt - \sum_{i=l+1}^{k}\nu(i)\right)^2\right]
$$
$$
= E\left[\int_{lT}^{(l+1)T}\int_{lT}^{(l+1)T}\frac{(\varphi(t_1)-\theta(l))(\varphi(t_2)-\theta(l))}{T^2}dt_1dt_2\right.
$$
$$
\left. - 2\sum_{i=l+1}^{k}\nu(i)\int_{lT}^{(l+1)T}\frac{\varphi(t)-\theta(l)}{T}dt + \left(\sum_{i=l+1}^{k}\nu(i)\right)^2\right]
$$

133

D. Phase Error Variance of DA ML PE with a Matched Filter

$$
= \frac{\sigma_p^2}{3} - 2E\left[\nu(l+1)\int_{lT}^{(l+1)T}\frac{\varphi(t)-\theta(l)}{T}dt\right] + \sum_{i=l+1}^{k}\sigma_p^2
$$

$$
= (k - l - \frac{2}{3})\sigma_p^2. \tag{D.2}
$$

The 2nd term in Eq.(D.1) can be further divided into the following form through some basic mathematical operations:

$$
E\left[\left(\int_{l_1T}^{(l_1+1)T}\frac{\varphi(t)-\theta(l_1)}{T}dt - \sum_{i=l_1+1}^{k}\nu(i)\right)\cdot\left(\int_{l_2T}^{(l_2+1)T}\frac{\varphi(t)-\theta(l_2)}{T}dt - \sum_{i=l_2+1}^{k}\nu(i)\right)\right]
$$

$$
= E\Bigg[\int_{l_1T}^{(l_1+1)T}\int_{l_2T}^{(l_2+1)T}\frac{(\varphi(t_1)-\theta(l_1))(\varphi(t_2)-\theta(l_2))}{T^2}dt_1dt_2
$$

$$
- \sum_{i=l_1+1}^{k}\nu(i)\int_{l_2T}^{(l_2+1)T}\frac{\varphi(t)-\theta(l_2)}{T}dt - \sum_{i=l_2+1}^{k}\nu(i)\int_{l_1T}^{(l_1+1)T}\frac{\varphi(t)-\theta(l_1)}{T}dt
$$

$$
+ \sum_{i_1=l_1+1}^{k}\sum_{i_2=l_2+1}^{k}\nu(i_1)\nu(i_2)\Bigg]. \tag{D.3}
$$

As a result, now it is easy to demonstrate that

$$
E\left[\int_{l_1T}^{(l_1+1)T}\int_{l_2T}^{(l_2+1)T}\frac{(\varphi(t_1)-\theta(l_1))(\varphi(t_2)-\theta(l_2))}{T^2}dt_1dt_2\right]
$$

$$
= \int_{l_1T}^{(l_1+1)T}\int_{l_2T}^{(l_2+1)T}\frac{\sigma_0^2}{T^2}\Big[\min(t_1,t_2) - \min(t_1,l_2T) - \min(l_1T,t_2)
$$

$$
- \min(l_1T,l_2T)\Big]dt_1dt_2
$$

$$
= 0. \tag{D.4}
$$

The 2nd expectation in Eq.(D.3) is evaluated as

$$
E\left[-\sum_{i=l_1+1}^{k}\nu(i)\int_{l_2T}^{(l_2+1)T}\frac{\varphi(t)-\theta(l_2)}{T}dt\right]
$$

$$
= E\left[-\sum_{i=l_1+1}^{k}(\varphi(iT)-\varphi((i-1)T))\int_{l_2T}^{(l_2+1)T}\frac{\varphi(t)-\theta(l_2)}{T}dt\right]
$$

$$
= \begin{cases} E\left[-[\varphi((l_2+1)T)-\varphi(l_2T)]\int_{l_2T}^{(l_2+1)T}\frac{\varphi(t)-\theta(l_2)}{T}dt\right], & l_1 < l_2; \\ 0, & l_1 > l_2 \end{cases}
$$

134

D. Phase Error Variance of DA ML PE with a Matched Filter

$$
= \begin{cases} -\sigma_p^2/2, & l_1 < l_2; \\ \\ 0, & l_1 > l_2 \end{cases} \tag{D.5}
$$

Similarly,

$$
E\left[-\sum_{i=l_2+1}^{k} \nu(i) \int_{l_1 T}^{(l_1+1)T} \frac{\varphi(t) - \theta(l_1)}{T} dt \right] = \begin{cases} 0, & l_1 < l_2; \\ \\ -\sigma_p^2/2, & l_1 > l_2 \end{cases} \tag{D.6}
$$

The expectation of the last term in Eq.(D.4) is obviously shown as

$$
E\left[\sum_{i_1=l_1+1}^{k} \sum_{i_2=l_2+1}^{k} \nu(i_1)\nu(i_2) \right] = (k - \max(l_1, l_2))\sigma_p^2 \tag{D.7}
$$

By summing up the expectations from Eqs.(D.4) to (D.7) over the double summation in Eq.(D.1) , together with the first expectation of Eq.(D.1), we can derive the phase error variance of DA ML with a matched filter in the absence of additive noise, given by

$$
E\left[(\Delta\theta(k))^2 \right] = \frac{1}{L^2} \left[\frac{L(L+1)}{2} - \frac{2L}{3} - \frac{L(L-1)}{2} + \frac{L(L^2-1)}{3} \right] \sigma_p^2 = \frac{L}{3}\sigma_p^2. \tag{D.8}
$$

Note that phase noise $\varphi(t)$ and additive noise $n(k)$ are independent. In consequence, the final phase error variance can be derived as

$$
E\left[(\Delta\theta(k))^2 \right] = \frac{L}{3}\sigma_p^2 + \frac{1}{2L\gamma_s}. \tag{D.9}
$$

Appendix E

Derivation $\hat{\alpha}(k)$ from Risk Function $R(k)$

In the risk function $R(k)$, given the observations $\{r(l), 1 \leq l \leq k\}$, we can remove the expectation operation in Eq.(4.2) in Chapter 4. By substituting $V(l)$ in $R(k)$ with Eq.(4.1), the risk function now can be rewritten as

$$
\begin{aligned}
R(k) &= E\left[\sum_{l=1}^{k} |r(l) - V(l)\hat{m}(l)|^2 \,\Big|\, \{r(l)\}_{l=1}^{k}\right] \\
&= \sum_{l=1}^{k} \left| r(l) - \left[\alpha V(l-1) + (1-\alpha)\frac{r(l-1)}{\hat{m}(l-1)}\right]\hat{m}(l)\right|^2 \\
&= \sum_{l=1}^{k} \left| r(l) - \alpha V(l-1)\hat{m}(l) - (1-\alpha)\frac{r(l-1)}{\hat{m}(l-1)}\hat{m}(l)\right|^2 \\
&= \sum_{l=1}^{k} \left\{ |r(l)|^2 + \alpha^2|V(l-1)\hat{m}(l)|^2 + (1-\alpha)^2\left|\frac{r(l-1)}{\hat{m}(l-1)}\hat{m}(l)\right|^2 \right. \\
&\qquad - 2\alpha\mathbf{Re}\left[r(l)V^*(l-1)\hat{m}^*(l)\right] - 2(1-\alpha)\mathbf{Re}\left[r(l)\left[\frac{r(l-1)}{\hat{m}(l-1)}\hat{m}(l)\right]^*\right] \\
&\qquad \left. + 2\alpha(1-\alpha)|\hat{m}(l)|^2\mathbf{Re}\left[V(l-1)\left[\frac{r(l-1)}{\hat{m}(l-1)}\right]^*\right]\right\}
\end{aligned}
$$

(E.1)

E. Derivation $\hat{\alpha}(k)$ from Risk Function $R(k)$

Through taking the first derivative of $R(k)$ with respect to α, i.e., $dR(k)/d\alpha = 0$, we can get

$$
\begin{aligned}
\frac{dR(k)}{d\alpha} = \sum_{l=1}^{k} & \left\{ 2\alpha |V(l-1)\hat{m}(l)|^2 - 2(1-\alpha)\left|\frac{r(l-1)\hat{m}(l)}{\hat{m}(l-1)}\right|^2 \right. \\
& - 2\mathbf{Re}\left[r(l)V^*(l-1)\hat{m}^*(l)\right] + 2\mathbf{Re}\left[r(l)r^*(l-1)\hat{m}^*(l)/\hat{m}*(l-1)\right] \\
& \left. + (2-4\alpha)|\hat{m}(l)|^2\mathbf{Re}\left[\frac{V(l-1)r^*(l-1)}{\hat{m}*(l-1)}\right] \right\} \\
& = 0.
\end{aligned}
\tag{E.2}
$$

The optimal $\hat{\alpha}(k)$ can be obtained after some simple mathematics, as given by:

$$
\hat{\alpha}(k) = \frac{A(k)}{B(k)},
\tag{E.3}
$$

where

$$
A(k) = \sum_{l=1}^{k} |\hat{m}(l)|^2 \cdot \left\{ \left|\frac{r(l-1)}{\hat{m}(l-1)}\right|^2 + \mathbf{Re}\left[\frac{r(l)}{m(l)}V^*(l-1) - \frac{r^*(l-1)}{\hat{m}^*(l-1)}\right.\right.
$$
$$
\left.\left.\left[\frac{r(l)}{\hat{m}(l)} + V(l-1)\right]\right]\right\}
\tag{E.4}
$$

$$
B(k) = \sum_{l=1}^{k} |\hat{m}(l)|^2 \cdot \left[|V(l-1)|^2 - 2\mathbf{Re}\left[V(l-1)\frac{r^*(l-1)}{\hat{m}^*(l-1)}\right] + \left|\frac{r(l-1)}{\hat{m}(l-1)}\right|^2\right].
$$
$$
\tag{E.5}
$$

It is seen that the numerator and denominator are formed recursively, therefore $A(k)$ and $B(k)$ can be re-written as shown in Eqs.(4.4) and (4.5) in Chapter 4.

Bibliography

[1] G. P. Agrawal, *Fiber-Optic Communication Systems*. New York: Wiley-Interscience, 2002.

[2] L. G. Kazovsky, G. Kalogerakis, and W.-T. Shaw, "Homodyne phase-shift-keying systems: Past challenges and future opportunities," *J. Lightw. Technol.*, vol. 24, no. 12, pp. 4876–4884, 2006.

[3] R. Noé, "PLL-free synchronous QPSK polarization multiplex/diversity receiver concept with digital I&Q baseband processing," *IEEE Photon. Technol. Lett.*, vol. 17, no. 4, pp. 887–889, 2005.

[4] E. Ip and J. M. Kahn, "Feedforward carrier recovery for coherent optical communications," *J. Lightw. Technol.*, vol. 25, no. 9, pp. 2675–2692, 2007.

[5] L. G. Kazovsky, "Performance analysis and laser linewidth requirements for optical PSK heterodyne communications systems," *J. Lightw. Technol.*, vol. 4, no. 4, pp. 415–425, 1986.

[6] C. Wree, D. Becker, D. Mohr, and A. Joshi, "Coherent receivers for phase-shift keyed transmission," in *Proc. OFC/NFOEC*, 2007, paper OMP6.

[7] S. Norimatsu and K. Iwashita, "Damping factor influence on linewidth requirements for optical PSK coherent detection systems," *J. Lightw. Technol.*, vol. 11, no. 7, pp. 1266–1233, 1993.

[8] C. Xu, X. Liu, and X. Wei, "Differential phase-shift keying for high spectral efficiency optical transmissions," *IEEE J. Sel. Topics in Quantum Electron.*, vol. 10, no. 2, pp. 281–293, 2004.

[9] D. G. Foursa, C. R. Davidsou, M. Nwsov, M. A. Mills, L. Xu, J. X. Cai, A. N. Pilipetskii, Y. Cai, C. Breverman, R. R. Cordell, T. J. Carvelli, P. C. Corbett, H. D. Kidorf, and N. S. Bergano, "2.56 Tb/s (256×10 Gb/s) transmission over 11,000 km using hybrid raman/EDFAs with 80 nm of continuous bandwidth," in *Proc. OFC/NFOEC*, 2002, paper FC3-1.

[10] R. Freund, L. Molle, F. Raub, C. Caspar, M. Karkri, and C. Weber, "Triple-(S/C/L)-band WDM transmission using erbium-doped fibre amplifiers," in *Proc. ECOC*, 2005, paper Mo.4.2.3.

[11] C. Lam, "Google optical network," in *Proc. OFC/NFOEC*, 2010, paper NWA3.

[12] A. H. Gnauck and P. J. Winzer, "Optical phase-shift-keyed transmission," *J. Lightw. Technol.*, vol. 23, no. 1, pp. 115–130, 2005.

[13] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, 2010.

[14] Y. Cai, "Coherent detection in long-haul transmission systems," in *Proc. OFC/NFOEC*, 2008, paper OTuM1.

[15] E. Ip, A. P. T. Lau, D. J. F. Barros, and J. M. Kahn, "Coherent detection in optical fiber systems," *Opt. Express*, vol. 16, no. 2, pp. 753–791, 2008.

[16] H. Jiang and R. Saunders, "Advances in SiGe ICs for 40G signal equalization," in *Proc. OFC/NFOEC*, 2006, paper OTuE1.

[17] P. Schvan, J. Bach, C. Fait, P. Flemke, R. Gibbins, Y. Greshishchev, N. Ben-Hamida, D. Pollex, J. Sitch, S.-C. Wang, and J. Wolczanski, "A 24GS/s 6b ADC in 90nm CMOS," in *Proc. IEEE Int. Solid-State Circuits Conf.*, 2008, pp. 544–634.

[18] I. Dedic, "56 GS/s ADC: Enabling 100GbE," in *Proc. OFC/NFOEC*, 2010, paper OThT6.

[19] "DPO/DSA/MSO70000 digital & mixed signal oscilloscope series." [Online]. Available: http://www.tek.com/products/oscilloscopes/dpo70000_dsa70000/

[20] "Wavemaster 8 zi oscilloscopes." [Online]. Available: http://www.lecroy.com/Oscilloscope/OscilloscopeSeries.aspx?mseries=45&capid=102&mid=504

[21] M. G. Taylor, "Coherent detection method using DSP for demodulation of signal and subsequent equalization of propagation impairments," *IEEE Photon. Technol. Lett.*, vol. 16, no. 2, pp. 674–676, 2004.

[22] D.-S. Ly-Gagnon, S. Tsukamoto, K. Katoh, and K. Kikuchi, "Coherent detection of optical quadrature phase-shift keying signals with carrier phase estimation," *J. Lightw. Technol.*, vol. 24, no. 1, pp. 12–21, 2006.

[23] S. J. Savory, G. Gavioli, R. I. Killey, and P. Bayvel, "Transmission of 42.8 Gbit/s polarization multiplexed NRZ-QPSK over 6400 km of standard fiber with no optical dispersion compensation," in *Proc. OFC/NFOEC*, 2007, paper OTuA1.

[24] C. Fludger, T. Duthel, D. van den Borne, C. Schulien, E.-D. Schmidt, T.Wuth, J. Geyer, E. DeMan, K. Giok-Djan, and H. deWaardt, "Coherent equalization and POLMUX-RZ-DQPSK for robust 100-GE transmission," *J. Lightw. Technol.*, vol. 26, no. 1, pp. 64–72, 2008.

[25] H. Sun, K.-T.Wu, and K. Roberts, "Real-time measurements of a 40 Gb/s coherent system," *Opt. Express*, vol. 16, no. 2, pp. 873–879, 2008.

[26] J. Yu, M.-F. Huang, S. Zhang, L. Xu, Y. Inada, T. Ogata, and Y. Aoki, "Transmission of 42.8-Gb/s polarization multiplexed RZ-QPSK DWDM signals over 3900 km with 12.5-Ghz channel spacing and coherent detection," in *Proc. OFC/NFOEC*, 2010, paper OTuD4.

[27] K.-P. Ho, *Phase-Modulated Optical Communication Systems*. New York: Springer, 2005.

[28] X. Zhou, J. Yu, D. Qian, T. Wang, G. Zhang, and P. D. Magill, "High-spectral-efficiency 114-Gb/s transmission using PolMux-RZ-8PSK modulation format and single-ended digital coherent detection technique," *J. Lightw. Technol.*, vol. 27, no. 3, pp. 146–152, 2009.

[29] A. Sano, H. Masuda, T. Kobayashi, M. Fujiwara, K. Horikoshi, E. Yoshida, Y. Miyamoto, M. Matsui, M. Mizoguchi, H. Yamazaki, Y. Sakamaki, and H. Ishii, "69.1-Tb/s (432×171-Gb/s) C- and Extended L-Band transmission over 240 km using PDM-16-QAM modulation and digital coherent detection," in *Proc. OFC/NFOEC*, 2010, paper PDPB7.

[30] S. J. Savory, "Digital filters for coherent optical receivers," *Opt. Express*, vol. 16, no. 2, pp. 804–817, 2008.

[31] A. J. Viterbi and A. N. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 4, pp. 543–551, 1983.

[32] M. Seimetz, "Laser linewidth limitations for optical systems with high-order modulation employing feed-forward digital carrier phase estimation," in *Proc. OFC/NFOEC*, 2008, paper OTuM2.

[33] S. Zhang, P. Y. Kam, C. Yu, and J. Chen, "Laser linewidth tolerance of decision-aided maximum likelihood phase estimation in coherent optical M-ary PSK and QAM systems," *IEEE Photon. Technol. Lett.*, vol. 21, no. 15, pp. 1075–1077, 2009.

[34] F. Rice, M. Rice, and B. Cowley, "A new algorithm for 16QAM carrier phase estimation using QPSK partitioning," *Digital Signal Processing*, vol. 12, no. 1, pp. 77–86, 2002.

[35] H. Louchet, K. Kuzmin, and A.Richter, "Improved DSP algorithms for coherent 16-QAM transmission," in *Proc. ECOC*, 2008, paper Tu.1.E.6.

[36] S. Zhang, P. Y. Kam, and C. Yu, "Block length effect of decision-aided maximum likelihood phase estimation in coherent optical communication systems," in *Proc. CLEO*, 2009, paper CMZ3.

[37] S. Zhang, P. Y. Kam, C. Yu, and J. Chen, "Decision-aided carrier phase estimation for coherent optical communications," *J. Lightw. Technol.*, vol. 28, no. 11, pp. 1597–1607, 2010.

[38] M. G. Taylor, "Accurate digital phase estimation for coherent detection using a parallel digital processor," in *Proc. ECOC*, 2005, paper Tu 4.2.6, pp. 263–264.

[39] L. M. Pessoa, H. M. Salgado, and I. Darwazeh, "Joint mitigation of optical impairments and phase estimation in coherent optical systems," in *Proc. IEEE LEOS Summer Topical Meetings*, 2008, pp. 169–170.

[40] H. Nakashima, T. Tanimura, T. Hoshida, S. Oda, J. C. Rasmussen, L. Li, Z. Tao, Y. Ishii, K. Shiota, K. Sugitani, and H. Adachi, "Novel wide-range frequency offset compensator demonstrated with real-time digital coherent receiver," in *Proc. ECOC*, 2008, paper Mo.3.D.4.

[41] Y. Han and G. Li, "Coherent optical communication using polarization multiple-input-multiple-output," *Opt. Express*, vol. 13, no. 19, pp. 7527–7534, 2005.

[42] S. Zhang, L. Xu, J. Yu, M.-F. Huang, P. Y. Kam, C. Yu, and T. Wang, "Dual-stage cascaded frequency offset estimation for digital coherent receivers," *IEEE Photon. Technol. Lett.*, vol. 22, no. 6, pp. 401–403, 2010.

[43] Y. Wang, E. Serpedin, and P. Ciblat, "Non-data aided feedforward estimation of psk-modulated carrier frequency offset," in *Proc. ICC*, 2002.

[44] M. Selmi, Y. Jaouen, and P. Ciblat, "Accurate digital frequency offset estimator for coherent PolMux QAM transmission systems," in *Proc. ECOC*, 2009, paper P3.08.

[45] A. Leven, N. Kaneda, U.-V. Koc, and Y.-K. Chen, "Frequency estimation in intradyne reception," *IEEE Photon. Technol. Lett.*, vol. 19, no. 6, pp. 366–369, 2007.

[46] L. Li, Z. Tao, S. Oda, T. Hoshida, and J. C. Rasmussen, "Wide-range, accurate and simple digital frequency offset compensator for optical coherent receivers," in *Proc. OFC/NFOEC*, 2008, paper OWT4.

[47] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *J. Lightw. Technol.*, vol. 26, no. 20, pp. 3416–3425, 2008.

[48] A. Demir, "Nonlinear phase noise in optical-fiber-communication systems," *J. Lightw. Technol.*, vol. 25, no. 8, pp. 2002–2032, 2007.

[49] K.-P. Ho and J. M. Kahn, "Electronic compensation technique to mitigate non-linear phase noise," *J. Lightw. Technol.*, vol. 22, no. 3, pp. 779–783, 2004.

[50] K. Roberts, C. Li, L. Strawczynski, M. O'Sullivan, and I. Hardcastle, "Electronic precompensation of optical nonlinearity," *IEEE Photon. Technol. Lett.*, vol. 18, no. 2, pp. 403–405, 2006.

[51] X. Liu and D. A. Fishman, "A fast and reliable algorithm for electronic pree-qualization of SPM and chromatic dispersion," in *Proc. OFC/NFOEC*, 2006, paper OThD4.

[52] G. Charlet, M. Salsi, P. Tran, M. Bertolini, H. Mardoyan, J. Renaudier, O. Bertran-Pardo, and S. Bigo, "72×100 Gb/s transmission over transoceanic distance, using large effective area fiber, hybrid Raman-Erbium amplification and coherent detection," in *Proc. OFC/NFOEC*, 2009, paper PDPB6.

[53] S. Oda, T. Tanimura, T. Hoshida, H. Nakashima, Z. Tao, and J. C. Rasmussen, "112 Gb/s DP-QPSK transmission using a novel nonlinear compensator in digital coherent receiver," in *Proc. OFC/NFOEC*, 2009.

[54] D. S. Millar, S. Makovejs, V. Mikhailov, R. I. Killey, P. Bayvel, and S. J. Savory, "Experimental comparison of nonlinear compensation in long-haul PDM-QPSK transmission at 42.7 and 85.4 Gb/s," in *Proc. ECOC*, 2009, paper Th9.4.4.

[55] Y. Ma, Q. Yang, Y. Tang, S. Chen, and W. Shieh, "1-Tb/s single-channel coherent optical OFDM transmission with orthogonal-band multiplexing and sub-wavelength bandwidth access," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 308–315, 2010.

[56] P. J. Winzer and R.-J. Essiambre, "Advanced optical modulation formats," in *Optical Fiber Telecommunications*, V ed., ser. Systems and Networks, I. P. Kaminow, T. Li, and A. E. Willner, Eds.   New York: Academic Press, 2008, vol. B.

[57] S. Walklin and J. Conradi, "Multilevel signaling for increasing the reach of 10-Gb/s lightwave systems," *J. Lightw. Technol.*, vol. 17, no. 11, pp. 2235–2248, 1999.

[58] R. A. Griffin and A. C. Carter, "Optical differential quadrature phase shift key (oDQPSK) for high-capacity optical transmission," in *Proc. OFC/NFOEC*, 2002, paper WX6.

[59] M. Seimetz, M. Noelle, and E. Patzak, "Optical systems with high-order dpsk and star qam modulation based on interferometric direct detection," *J. Lightw. Technol.*, vol. 25, no. 6, pp. 1515–1530, 2007.

[60] N. Kikuchi, K. Mandai, K. Sekine, and S. Sasaki, "Incoherent 32-level optical multilevel signaling technologies," *J. Lightw. Technol.*, vol. 26, no. 1, pp. 150–157, 2008.

[61] S. Makovejs, D. S. Millar, V. Mikhailov, G. Gavioli, R. I. Killey, S. J. Savory, and P. Bayvel, "Novel method of generating QAM-16 signals at 21.3 Gbaud and transmission over 480 km," *IEEE Photon. Technol. Lett.*, vol. 22, no. 1, pp. 36–38, 2010.

[62] J. G. Proakis, *Digital Communications*, 4th ed.   New York: McGraw-Hill, 2000.

[63] M. Ohm and J. Speidel, "Quaternary optical ask-dpsk and receivers with direct detection," *IEEE Photon. Technol. Lett.*, vol. 15, no. 1, pp. 159–161, 2003.

BIBLIOGRAPHY

[64] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, and 623–656, 1948.

[65] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.

[66] F. Heismann, S. K. Korotky, and J. J. Veselka, "Lithium niobate integrated optics: Selected contemporary devices and system applications," in *Optical Fiber Telecommunications*, III ed., I. P. Kaminow and T. L. Koch, Eds. New York: Academic Press, 1997.

[67] T. Kawanishi, S. Sakamoto, and M. Izutsu, "High-speed control of lightwave amplitude, phase, and frequency by use of electrooptic effect," *IEEE J. Sel. Top. Quantum Electron.*, vol. 13, no. 1, pp. 79–91, 2007.

[68] F. Koyama and K. Oga, "Frequency chirping in external modulators," *J. Lightw. Technol.*, vol. 6, no. 1, pp. 87–93, 1988.

[69] A. H. Gnauck, S. K. Korotky, J. J. Veselka, J. N. C. T. Kemmerer, W. J. Minford, and D. T. Moser, "Dispersion penalty reduction using an optical modulator with adjustable chirp," *J. Lightw. Technol.*, vol. 3, no. 10, pp. 916–918, 1991.

[70] K.-P. Ho and H.-W. Cuei, "Generation of arbitrary quadrature signals using one dual-drive modulator," *J. Lightw. Technol.*, vol. 23, no. 2, pp. 764–770, 2005.

[71] C. Weber, J. K. Fischer, C.-A. Bunge, and K. Petermann, "Electronic precompensation of intrachannel nonlinearities at 40 Gb/s," *IEEE Photon. Technol. Lett.*, vol. 18, no. 16, pp. 1759–1761, 2006.

[72] S. Zhang, X. Li, P. Y. Kam, C. Yu, and J. Chen, "Pilot-assisted, decision-aided, maximum likelihood phase estimation in coherent optical phase-modulated sys-

tems with nonlinear phase noise," *IEEE Photon. Technol. Lett.*, vol. 22, no. 6, pp. 380–382, 2010.

[73] P. J. Winzer, D. Stahl, and H. Kim, "Impact of pulse carver chirp on RZ-DPSK receiver performance," in *Proc. ECOC*, 2003, paper We3.5.6.

[74] M. I. Hayee and A. E. Willner, "NRZ versus RZ in 10 - 40Gb/s dispersion-managed WDM transmission systems," *IEEE Photon. Technol. Lett.*, vol. 11, no. 8, pp. 991–993, 1999.

[75] C. Caspar, H. M. Foisel, A. Gladisch, N. Hanik, F. Kppers, R. Ludwig, A. Mattheus, W. Pieper, B. Strebel, and H. G. Weber, "RZ versus NRZ modulation format for dispersion compensated SMF-based 10 Gb/s transmission with more than 100-km amplifier spacing," *IEEE Photon. Technol. Lett.*, vol. 11, no. 4, pp. 481–483, 1999.

[76] E. Ip and J. M. Kahn, "Power spectra of return-to-zero optical signals," *J. Lightw. Technol.*, vol. 24, no. 3, pp. 1610–1618, 2006.

[77] P. J. Winzer and R.-J. Essiambre, "Advanced modulation formats for high-capacity optical transport networks," *J. Lightw. Technol.*, vol. 24, no. 12, pp. 4711–4728, 2006.

[78] G. P. Agrawal, *Nonlinear Fiber Optics*, 3rd ed. San Diego: Academic Press, 2001.

[79] L. D. Coelho, L. Molle, D. Gross, N. Hanik, R. Freund, C. Caspar, E.-D. Schmidt, and B. Spinnler, "Modeling nonlinear phase noise in differentially phase-modulated optical communication systems," *Opt. Express*, vol. 17, no. 5, pp. 3226–3241, 2009.

[80] A. V. Oppenheim, A. S. Willsky, and H. Nawab, *Signals and Systems*, 2nd ed. Upper Saddle River, N.J.: Prentice Hall, 1997.

[81] T. Miya, Y. Terunuma, T. Hosaka, and T. Miyashita, "Ultimate low-loss single-mode fibre at 1.55 μm," *Electron. Lett.*, vol. 15, no. 4, pp. 106–108, 1979.

[82] "Corning® SMF-28® ULL optical fiber." [Online]. Available: http://www.corning.com/opticalfiber/products/SMF-28_ULL_fiber.aspx

[83] A. Bjarklev, *Optical Fiber Amplifiers: Design and System Applications*. Boston: Artech House, 1993.

[84] G. Charlet, "Progress in optical modulation formats for high-bit rate wdm transmissions," *IEEE J. Sel. Top. Quantum Electron.*, vol. 12, no. 4, pp. 469–483, 2006.

[85] C. Lin, H. Kogelnik, and L. G. Cohen, "Optical-pulse equalization of low-dispersion transmission in single-mode fibers in the 1.3-1.7-μm spectral region," *Opt. Lett.*, vol. 5, no. 11, pp. 476–478, 1980.

[86] V. K. Mezentsev, S. K. Turitsyn, and N. J. Doran, "System optimisation of 80 Gb/s single channel transmission over 1000 km of standard fibre," *Electron. Lett.*, vol. 36, no. 23, pp. 1949–1950, 2000.

[87] L. Grüner-Nielsen, M. Wandel, P. Kristensen, C. Jørgensen, L. V. Jørgensen, B. Edvold, B. Pálsdóttir, and D. Jakobsen, "Dispersion-compensating fibers," *J. Lightw. Technol.*, vol. 23, no. 11, pp. 3566–3579, 2005.

[88] G. Li, "Recent advances in coherent optical communication," *Adv. Optics Photonics*, vol. 1, no. 2, pp. 279–307, 2009.

[89] C. Yu, "Polarization mode dispersion (PMD) monitoring," in *Optical Performance Monitoring Techniques for Next Generation Photonic Networks*, C. K. Chan, Ed.   Academic Press.

[90] M. Brodsky, N. J. Frigo, and M. Tur, "Polarization mode dispersion," in *Optical Fiber Telecommunications*, V ed., ser. Components and Subsystems, I. P. Kaminow, T. Li, and A. E. Willner, Eds.   New York: Academic Press, 2008, vol. A.

[91] C. D. Poole, "Statistical treatment of polarization dispersion in single-mode fiber," *Opt. Lett.*, vol. 13, no. 8, pp. 687–689, 1988.

[92] J. P. Gordon and L. F. Mollenauer, "Phase noise in photonic communications systems using linear amplifiers," *Opt. Lett.*, vol. 15, no. 5, pp. 1351–1353, 1990.

[93] H. Kim and A. H. Gnauck, "Experimental investigation of the performance limitation of DPSK systems due to nonlinear phase noise," *IEEE Photon. Technol. Lett.*, vol. 15, no. 2, pp. 320–322, 2003.

[94] S. Zhang, P. Y. Kam, J. Chen, and C. Yu, "Decision-aided maximum likelihood detection in coherent optical phase-shift-keying system," *Opt. Express*, vol. 17, no. 2, pp. 703–715, 2009.

[95] T.-K. Chiang, N. Kagi, M. E. Marhic, and L. G. Kazovsky, "Cross-phase modulation in fiber links with multiple optical amplifiers and dispersion compensators," *J. Lightw. Technol.*, vol. 14, no. 3, pp. 249–260, 1996.

[96] R. H. Hardin and F. D. Tappert, "Applications of the split-step Fourier method to the numerical solution of nonlinear and variable coefficient wave equations," *SIAM Rev.*, vol. 15, p. 423, 1973.

[97] A. Fleck, J. R. Morris, and M. D. Feit, "Time dependent propagation of high energy laser beams through the atmosphere," *Appl. Phys.*, vol. 10, pp. 129–160, 1976.

[98] A. M. Lax, J. H. Batteh, and G. P. Agrawal, "Channeling of intense electromagnetic beams," *J. Appl. Phys.*, vol. 52, pp. 109–125, 1981.

[99] X. Liu and B. Lee, "A fast method for nonlinear schrödinger equation," *IEEE Photon. Technol. Lett.*, vol. 15, no. 11, pp. 1549–1551, 2003.

[100] Y. Sun, M. E. Ali, K. Balemarthy, R. L. Lingle Jr, S. E. Ralph, and B. E. Lemoff, "10 Gb/s transmission over 300m OM3 fiber from 990-1080nm with electronic dispersion compensation," in *Proc. OFC/NFOEC*, 2006, paper OTuE2.

[101] B. Franz, F. Buchali, D. Roesener, and H. Büelow, "Adaptation techniques for electronic equalizers for the mitigation of time-variant distortions in 43 Gbit/s optical transmission systems," in *Proc. OFC/NFOEC*, 2007, paper OMG1.

[102] A. Leven, N. Kaneda, and Y. Chen, "A real-time CMA-based 10 Gb/s polarization demultiplexing coherent receiver implemented in an FPGA," in *Proc. OFC/NFOEC*, 2008, paper OTuO2.

[103] R. Kudo, T. Kobayashi, K. Ishihara, Y. Takatori, A. Sano, E. Yamada, H. Masuda, Y. Miyamoto, and M. Mizoguchi, "Two-stage overlap frequency domain equalization for long-haul optical systems," in *Proc. OFC/NFOEC*, 2009, paper OMT3.

[104] M. Seimetz and C.-M. Weinert, "Options, feasibility, and availability of 2×4 90° hybrids for coherent optical systems," *J. Lightw. Technol.*, vol. 24, no. 3, pp. 1317–1322, 2006.

[105] K. Kikuchi and S. Tsukamoto, "Evaluation of sensitivity of the digital coherent receiver," *J. Lightw. Technol.*, vol. 26, no. 13, pp. 1817–1822, 2008.

[106] G. Nicholson, "Probability of error for optical heterodyne DPSK system with quantum phase noise," *Electron. Lett.*, vol. 20, no. 24, pp. 1005–1007, 1984.

[107] S. J. Savory, "Digital coherent optical receivers: Algorithms and subsystems," *IEEE J. Sel. Top. Quantum Electron.*, vol. 16, no. 5, pp. 1164–1179, 2010.

[108] M. Oerder and H. Meyr, "Digital filter and square timing recovery," *IEEE Trans. Commun.*, vol. 36, no. 5, pp. 605–612, 1988.

[109] F. M. Gardner, "A BPSK/QPSK timing-error detector for sampled receivers," *IEEE Trans. Commun.*, vol. 34, no. 5, pp. 423–429, 1986.

[110] I. Fatadin, S. J. Savory, and D. Ives, "Compensation of quadrature imbalance in an optical QPSK coherent receiver," *IEEE Photon. Technol. Lett.*, vol. 20, no. 20, pp. 1733–1735, 2008.

[111] S. H. Chang, H. S. Chung, and K. Kim, "Impact of quadrature imbalance in optical coherent QPSK receiver," *IEEE Photon. Technol. Lett.*, vol. 21, no. 11, pp. 709–711, 2009.

[112] M. Kuschnerov, F. Hauske, K. Piyawanno, B. Spinnler, M. S. Alfiad, A. Napoli, and B. Lankl, "Dsp for coherent single-carrier receivers," *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3614–3622, 2009.

[113] G. Goldfarb and G. Li, "Chromatic dispersion compensation using digital IIR filtering with coherent detection," *IEEE Photon. Technol. Lett.*, vol. 19, no. 13, pp. 969–971, 2007.

[114] S. U. H. Qureshi, "Adaptive equalization," *Proceedings of the IEEE*, vol. 73, no. 9, pp. 1349–1387, 1985.

[115] S. Tsukamoto, K. Katoh, and K. Kikuchi, "Unrepeated transmission of 20-Gb/s optical quadrature phase-shift-keying signal over 200-km standard single-mode fiber based on digital processing of homodyne-detected signal for group-velocity dispersion compensation," *IEEE Photon. Technol. Lett.*, vol. 18, no. 9, pp. 1016–1018, 2006.

[116] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing : Principles, Algorithms, and Applications*, 3rd ed.   Upper Saddle River, N.J.: Prentice-Hall, 1996.

[117] E. Ip and J. M. Kahn, "Digital equalization of chromatic dispersion and polarization mode dispersion," *J. Lightw. Technol.*, vol. 25, no. 8, pp. 2033–2043, 2007.

[118] X. Li, X. Chen, G. Goldfarb, E. Mateo, I. Kim, F. Yaman, and G. Li, "Electronic post-compensation of WDM transmission impairments using coherent detection and digital signal processing," *Opt. Express*, vol. 16, no. 2, pp. 881–888, 2008.

[119] D. Godard, "Self-recovering equalization and carrier tracking in twodimensional data communication systems," *IEEE Trans. Commun.*, vol. 28, no. 11, pp. 1867–1875, 1980.

[120] P. Winzer, A. Gnauck, C. Doerr, M. Magarini, and L. Buhl, "Spectrally efficient long-haul optical networking using 112-Gb/s polarization multiplexed 16-QAM," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 547–556, 2010.

[121] I. Fatadin, D. Ives, and S. J. Savory, "Blind equalization and carrier phase recovery in a 16-QAM optical coherent system," *J. Lightw. Technol.*, vol. 27, no. 15, pp. 3042–3049, 2009.

[122] L. Du and A. Lowery, "Practical XPM compensation method for coherent optical OFDM systems," *IEEE Photon. Technol. Lett.*, vol. 22, no. 5, pp. 320–322, 2010.

[123] S. Hoffmann, S. Bhandare, T. Pfau, O. Adamczyk, C. Wördehoff, R. Peveling, M. Porrmann, and R. Noé, "Frequency and phase estimation for coherent QPSK transmission with unlocked DFB lasers," *IEEE Photon. Technol. Lett.*, vol. 20, no. 18, pp. 1569–1571, 2008.

[124] W. J. Weber, "Differential encoding for multiple amplitude and phase shift keying systems," *IEEE Trans. Commun.*, vol. 26, no. 3, pp. 385–391, 1978.

[125] G. Colavolpe and R. Raheli, "Noncoherent sequence detection," *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1376–1385, 1999.

[126] I. B. Djordjevic, M. Arabaci, and L. L. Minkov, "Next generation FEC for high-capacity communication in optical transport networks," *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3518–3530, 2009.

[127] M. G. Taylor, "Phase estimation methods for optical coherent detection using digital signal processing," *J. Lightw. Technol.*, vol. 27, no. 7, pp. 901–914, 2009.

[128] X. Zhou, J. Yu, M. Huang, Y. Shao, T. Wang, L. Nelson, P. Magill, M. Birk, P. I. Borel, D. W. Peckham, and R. Lingle, "64-Tb/s (640×107-Gb/s) PDM-36QAM transmission over 320km using both pre- and post-transmission digital equalization," in *Proc. OFC/NFOEC*, 2010, paper PDPB9.

[129] Y. Mori, C. Zhang, K. Igarashi, K. Katoh, and K. Kikuchi, "Unrepeated 200-km transmission of 40-Gbit/s 16-QAM signals using digital coherent receiver," *Opt. Express*, vol. 17, no. 3, pp. 1435–1441, 2009.

BIBLIOGRAPHY

[130] P. Y. Kam, "Maximum-likelihood carrier phase recovery for linear suppressed-carrier digital data modulations," *IEEE Trans. Commun.*, vol. COM-34, pp. 522–527, 1986.

[131] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. Dubuque, Iowa: McGraw-Hill, 2002.

[132] D. D. Falconer and J. Salz, "Optimal reception of digital data over the Gaussian channel with unknown delay and phase jitter," *IEEE Trans. Inf. Theory*, vol. 23, no. 1, pp. 117–126, 1977.

[133] P. Y. Kam, K. H. Chua, and X. Yu, "Adaptive symbol-by-symbol reception of MPSK on the gaussian channel with unknown carrier phase characteristics," *IEEE Trans. Commun.*, vol. 46, pp. 1275–1279, 1998.

[134] G. Goldfarb and G. Li, "BER estimation of QPSK homodyne detection with carrier phase estimation using digital signal processing," *Opt. Express*, vol. 14, no. 18, pp. 8043–8053, 2006.

[135] S. Zhang, P. Y. Kam, J. Chen, and C. Yu, "Bit-error rate performance of coherent optical M-ary PSK and 16-QAM using decision-aided maximum likelihood phase estimation," *Opt. Express*, vol. 18, no. 12, pp. 12 088–12 103, 2010.

[136] M. Scholten, T. Coe, and J. Dillard, "Continuously-interleaved BCH (CIBCH) FEC delivers best in class NECG for 40G and 100G metro applications," in *Proc. OFC/NFOEC*, 2010, paper NTuB3.

[137] P. Y. Kam, S. K. Teo, Y. K. Some, and T. T. Tjhung, "Approximate results for the bit error probability of binary PSK with noisy phase reference," *IEEE Trans. Commun.*, vol. 41, no. 7, pp. 1020–1022, 1993.

[138] Y. K. Some and P. Y. Kam, "Bit error probability of QPSK with noisy phase reference," *IEE Proceedings - Communications*, vol. 142, no. 5, pp. 292–296, 1995.

[139] M. P. Fitz and R. J. M. Cramer, "A performance analysis of a digital PLL-based MPSK demodulator," *IEEE Trans. Commun.*, vol. 43, no. 2,3,4, pp. 1192–1201, 1995.

[140] S. Zhang, C. Yu, P. Y. Kam, and J. Chen, "Parallel implementation of decision-aided maximum likelihood phase estimation in coherent M-ary phase-shifted keying systems," *IEEE Photon. Technol. Lett.*, vol. 21, no. 19, pp. 1471–1473, 2009.

[141] T. Pfau, S. Hoffmann, and R. Noé, "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for M-QAM constellations," *J. Lightw. Technol.*, vol. 27, no. 8, pp. 989–999, 2009.

[142] M. Nazarathy, X. Liu, L. Christen, Y. K. Lize, and A. Willner, "Self-coherent multisymbol detection of optical differential phase-shift-keying," *J. Lightw. Technol.*, vol. 26, no. 13, pp. 1921–1934, 2008.

[143] T. Pfau, S. Hoffmann, R. Peveling, S. Bhandare, S. K. Ibrahim, O. Adamczyk, M. Porrmann, R. Noé, and Y. Achiam, "First real-time data recovery for synchronous QPSK transmission with standard DFB lasers," *IEEE Photon. Technol. Lett.*, vol. 18, no. 9, pp. 1907–1909, 2006.

[144] "Virtex-7 FPGA family." [Online]. Available: http://www.xilinx.com/products/silicon-devices/fpga/virtex-7/index.htm

[145] Y. Atzmon and M. Nazarathy, "Laser phase noise in coherent and differential optical transmission revisited in the polar domain," *J. Lightw. Technol.*, vol. 27, no. 1, pp. 19–29, 2009.

[146] S. Zhang, L. Xu, J. Yu, P. Y. Kam, C. Yu, and T. Wang, "Experimental demonstration of decision-aided maximum likelihood phase estimation in 8-channel 42.8-Gbit/s DWDM coherent PolMux-QPSK system," in *Proc. OFC/NFOEC*, 2010, paper OMK1.

[147] S. Zhang, L. Xu, P. Y. Kam, C. Yu, J. Chen, and T. Wang, "A performance investigation of adaptive phase estimations in coherent optical communications," *IEEE Photon. Technol. Lett.*, vol. 23, no. 8, pp. 462–464, 2011.

[148] S. Zhang, C. Yu, P. Y. Kam, and J. Chen, "Performance comparison between decision-aided maximum likelihood and adaptive decision-aided phase estimation," in *Proc. ICOCN*, 2010, pp. 253–257.

[149] L. Pessoa, H. Salgado, and I. Darwazeh, "Performance evaluation of phase estimation algorithms in equalized coherent optical systems," *IEEE Photon. Technol. Lett.*, vol. 21, no. 17, pp. 1181–1183, 2009.

[150] S. Zhang, C. Yu, P. Y. Kam, and J. Chen, "Optimizing the performance of normalized least-mean square phase estimation for digital coherent receivers," in *Proc. OECC*, 2010, pp. 286–287.

[151] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, N. J.: Prentice-Hall, 2002.

[152] A. P. T. Lau and J. M. Kahn, "Design of inline amplifier gains and spacings to minimize the phase noise in optical transmission systems," *J. Lightw. Technol.*, vol. 24, no. 3, pp. 1334–1341, 2006.

[153] K. Piyawanno, M. Kuschnerov, B. Spinnler, and B. Lankl, "Fast and accurate automatic frequency control for coherent receivers," in *Proc. ECOC*, 2009, paper 7.3.1.

[154] S. Zhang, P. Y. Kam, C. Yu, and J. Chen, "Frequency offset estimation using a Kalman filter in coherent optical phase-shift keying systems," in *Proc. CLEO*, 2010, paper CThDD4.

[155] T. Tanimura, T. Hoshida, S. Oda, H. Nakashima, M. Yuki, Z. Tao, L. Liu, and J. C. Rasmussen, "Digital clock recovery algorithm for optical coherent receivers operating independent of laser frequency offset," in *Proc. ECOC*, 2008, paper Mo.3.D.2.

[156] S. Zhang, L. Xu, J. Yu, M.-F. Huang, P. Y. Kam, C. Yu, and T. Wang, "Novel ultra wide-range frequency offset estimation for digital coherent optical receiver," in *Proc. OFC/NFOEC*, 2010, paper OWV3.

[157] H. Bülow, F. Buchali, and A. Klekamp, "Electronic dispersion compensation," *J. Lightw. Technol.*, vol. 26, no. 1, pp. 158–167, 2008.

[158] S. Zhang, L. Xu, P. Y. Kam, C. Yu, and T. Wang, "Performance investigation of the joint-SPM compensation in a long-haul coherent dual-polarization QPSK system," in *Proc. ECOC*, 2010, paper P3.15.

[159] L. B. Du and A. J. Lowery, "Improved single channel backpropagation for intra-channel fiber nonlinearity compensation in long-haul optical communication systems," *Opt. Express*, vol. 18, no. 16, pp. 17 075–17 088, 2010.

[160] S. J. Savory, G. Gavioli, E. Torrengo, and P. Poggiolini, "Impact of interchannel nonlinearities on a split-step intrachannel nonlinear equalizer," *IEEE Photon. Technol. Lett.*, vol. 22, no. 10, pp. 673–675, 2010.

[161] X. Liu, S. Chandrasekhar, B. Zhu, P. J. Winzer, A. H. Gnauck, and D. W. Peckham, "Transmission of a 448-Gb/s reduced-guard-interval CO-OFDM signal with a 60-GHz optical bandwidth over 2000 km of ULAF and five 80-GHz-Grid ROADMs," in *Proc. OFC/NFOEC*, 2010, paper PDPC2.

[162] M. Nakazawa, "Optical quadrature amplitude modulation (QAM) with coherent detection up to 128 states," in *Proc. OFC/NFOEC*, 2009, paper OThG1.

[163] J. Yu, X.Zhou, Y.-K. Huang, S. Gupta, M.-F. Huang, T. Wang, and P. Magill, "112.8-Gb/s PM-RZ-64QAM optical signal generation and transmission on a 12.5GHz WDM grid," in *Proc. OFC/NFOEC*, 2010, paper OThM1.

[164] T. Mizuochi, "Recent progress in forward error correction and its interplay with transmission impairments," *IEEE J. Sel. Top. Quantum Electron.*, vol. 12, no. 4, pp. 544–554, 2006.

[165] F. R. Kschischang and B. P. Smith, "Forward error correction (FEC) in optical communication," in *Proc. CLEO*, 2010, paper CThDD1.

[166] K. Onohara, T. Sugihara, Y. Konishi, Y. Miyata, T. Inoue, S. Kametani, K. Sugihara, K. Kubo, H. Yoshida, and T. Mizuochi, "Soft-decision-based forward error correction for 100 Gb/s transport systems," *IEEE J. Sel. Top. Quantum Electron.*, vol. 16, no. 5, pp. 1258–1267, 2010.

[167] J. Lassing, E. Ström, E. Agrell, and T. Ottosson, "Computation of the exact bit-error rate of coherent M-ary PSK with gray code bit mapping," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1758–1760, 2003.

[168] H. Fu and P. Y. Kam, "A simple bit error probability analysis for square QAM in rayleigh fading with channel estimation," *IEEE Trans. Commun.*, vol. 57, no. 8, pp. 2193–2197, 2009.

List of Publications

Journal Papers

1. *Shaoliang Zhang*, Lei Xu, Pooi Yuen Kam, Changyuan Yu, Jian Chen, and Ting Wang, "A performance investigation of adaptive phase estimations in coherent optical communications," *IEEE Photon. Technol. Lett.*, vol. 23, no. 8, pp. 462–464, 2011.

2. *Shaoliang Zhang*, Pooi Yuen Kam, Jian Chen, and Changyuan Yu, "Bit-error rate performance of coherent optical M-ary PSK and 16-QAM using decision-aided maximum likelihood phase estimation," *Opt. Express*, vol. 18, no. 12, pp. 12 088–12 103, 2010.

3. *Shaoliang Zhang*, Pooi Yuen Kam, Changyuan Yu, and Jian Chen, "Decision-aided carrier phase estimation for coherent optical communications," *J. Lightw. Technol.*, vol. 28, no. 11, pp. 1597–1607, 2010.

4. *Shaoliang Zhang*, Lei Xu, Jianjun Yu, Ming-Fang Huang, Pooi Yuen Kam, Chang yuan Yu, and Ting Wang, "Dual-stage cascaded frequency offset estimation for digital coherent receivers," *IEEE Photon. Technol. Lett.*, vol. 22, no. 6, pp. 401–403, 2010.

5. *Shaoliang Zhang*, Xiaojing Li, Pooi Yuen Kam, Changyuan Yu, and Jian Chen, "Pilot-assisted, decision-aided, maximum likelihood phase estimation in coher-

ent optical phase-modulated systems with nonlinear phase noise," *IEEE Photon. Technol. Lett.*, vol. 22, no. 6, pp. 380–382, 2010.

6. *Shaoliang Zhang*, Changyuan Yu, Pooi Yuen Kam, and Jian Chen, "Parallel implementation of decision-aided maximum likelihood phase estimation in coherent M-ary phase-shifted keying systems," *IEEE Photon. Technol. Lett.*, vol. 21, no. 19, pp. 1471–1473, 2009.

7. *Shaoliang Zhang*, Pooi Yuen Kam, Changyuan Yu, and Jian Chen, "Laser linewidth tolerance of decision-aided maximum likelihood phase estimation in coherent optical M-ary PSK and QAM systems," *IEEE Photon. Technol. Lett.*, vol. 21, no. 15, pp. 1075–1077, 2009.

8. *Shaoliang Zhang*, Pooi Yuen Kam, Jian Chen, and Changyuan Yu, "Decision-aided maximum likelihood detection in coherent optical phase-shift-keying system," *Opt. Express*, vol. 17, no. 2, pp. 703–715, 2009.

Conference Papers

1. *Shaoliang Zhang*, Lei Xu, Pooi Yuen Kam, Changyuan Yu, and Ting Wang, "Study on the performance of decision-aided maximum likelihood phase estimation with a forgetting factor," *Proc. OFC/NFOEC*, Paper JWA022, 2011.

2. Zixiong Wang, Wen-De Zhong, *Shaoliang Zhang*, Changyuan Yu, and Yicheng Ding, "Performance evaluation of OOK free-space optical transmission system with coherent detection and dynamic decision threshold," in *Proc. IEEE Photonics Global Conference (IPGC)*, Paper 3-1F-4, 2010.

3. Changyuan Yu, Pooi Yuen Kam, *Shaoliang Zhang*, and Jian Chen, "Phase estimation in coherent communication systems with semiconductor laser noises," in

Proc. Photonics Asia 2010, Paper 7844-19, (invited paper).

4. *Shaoliang Zhang*, Changyuan Yu, Pooi Yuen Kam, and Jian Chen, "Performance comparison between decision-aided maximum likelihood and adaptive decision-aided phase estimation," in *Proc. ICOCN*, pp. 253-257, 2010.

5. *Shaoliang Zhang*, Lei Xu, Pooi Yuen Kam, Changyuan Yu, and Ting Wang, "Performance investigation of the joint-SPM compensation in a long-haul coherent dual-polarization QPSK system," in *Proc. ECOC*, 2010, paper P3.15.

6. Changyuan Yu, Pooi Yuen Kam, *Shaoliang Zhang*, and Jian Chen, "Decision-aided maximum likelihood phase estimation in coherent communication systems," in *Proc. OECC* 2010, pp. 764–765 (invited paper).

7. *Shaoliang Zhang*, Changyuan Yu, Pooi Yuen Kam, and Jian Chen, "Optimizing the performance of normalized least-mean square phase estimation for digital coherent receivers," in *Proc. OECC* 2010, pp. 286–287.

8. Hongyu Zhang, *Shaoliang Zhang*, Pooi Yuen Kam, Changyuan Yu, and Jian Chen,"Optimized phase error tolerance of 16-star quadrature amplitude modulation in coherent optical communication systems," in *Proc. OECC* 2010, pp. 592–593.

9. *Shaoliang Zhang*, Pooi Yuen Kam, Changyuan Yu, and Jian Chen, "Frequency offset estimation using a Kalman filter in coherent optical phase-shift keying systems," in *Proc. CLEO*, 2010, paper CThDD4.

10. *Shaoliang Zhang*, Lei Xu, Jianjun Yu, Pooi Yuen Kam, Changyuan Yu, and Ting Wang, "Experimental demonstration of decision-aided maximum likelihood phase estimation in 8-channel 42.8-Gbit/s DWDM coherent PolMux-QPSK system," in *Proc. OFC/NFOEC*, 2010, paper OMK1.

11. *Shaoliang Zhang*, Lei Xu, Jianjun Yu, Ming-Fang Huang, Pooi Yuen Kam, Changyuan Yu, and Ting Wang, "Novel ultra wide-range frequency offset estimation for digital coherent optical receiver," in *Proc. OFC/NFOEC*, 2010, paper OWV3.

12. Jianjun Yu, Ming-Fang Huang, *Shaoliang Zhang*, Lei Xu, Yoshihisa Inada, Takaaki Ogata, and Yasuhiro Aoki, "Transmission of 42.8-Gb/s polarization multiplexed RZ-QPSK DWDM signals over 3900 km with 12.5-Ghz channel spacing and coherent detection," in *Proc. OFC/NFOEC*, 2010, paper OTuD4.

13. *Shaoliang Zhang*, Jian Chen, Changyuan Yu, Weifeng Rong, and Pooi Yuen Kam, "ADC bandwidth optimization for coherent optical detection in phase-modulated systems," in *Proc. Asia-Pacific Optical Communications (APOC)* 2009, Paper FC3.

14. *Shaoliang Zhang*, Pooi Yuen Kam, and Changyuan Yu, "Block length effect of decision-aided maximum likelihood phase estimation in coherent optical communication systems," in *Proc. Conference on Lasers and Electro-Optics (CLEO)* 2009, Paper CMZ3.

15. Xiaojing Li, *Shaoliang Zhang*, Changyuan Yu, and Pooi Yuen Kam, "Pilot decision aided maximum likelihood phase estimation in coherent optical QPSK and 8PSK systems with nonlinear phase noise," in *Proc. Conference on Lasers and Electro-Optics (CLEO)* 2009, Paper CMZ4.

16. *Shaoliang Zhang*, Pooi Yuen Kam, Jian Chen, and Changyuan Yu, "A Comparison of Phase Estimation in Coherent Optical PSK System," in *Proc. IEEE Photonics Global Conference (IPGC)* 2008, Paper C3-4A-03.

17. *Shaoliang Zhang*, Pooi Yuen Kam, Jian Chen, and Changyuan Yu, "Adaptive Decision-Aided Maximum Likelihood Phase Estimation in Coherent Optical

DQPSK System," in *Proc. OptoElectronics and Communications Conference (OECC)* 2008, Paper TuA-4, pp. 1-2.

18. *Shaoliang Zhang*, Pooi Yuen Kam, Jian Chen, and Changyuan Yu, "Receiver sensitivity improvement using decision-aided maximum likelihood phase estimation in coherent optical DQPSK system," in *Proc. Conference on Lasers and Electro-Optics (CLEO)* 2008, paper CThJJ2.