CHANNEL CHARACTERIZATION AND SIGNAL TRANSMISSION FOR DIGITAL SIGNAL PROCESSING BASED OPTICAL COMMUNICATION SYSTEMS

SUI QI

Ph.D

The Hong Kong Polytechnic University

2015

The Hong Kong Polytechnic University

Department of Electrical Engineering

CHANNEL CHARACTERIZATION AND SIGNAL TRANSMISSION FOR DIGITAL SIGNAL PROCESSING BASED OPTICAL COMMUNICATION SYSTEMS

SUI QI

A thesis submitted in partial fulfillment of the requirements for

The Degree of Doctor of Philosophy

September 2014

Certificate of Originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

…….. (Signed)

Sui Qi …….. (Name of student)

Abstract

The presence of various degradations in optical transmission systems raises the demand for channel characterization. In the aid of optical performance monitoring (OPM) techniques or channel estimation algorithms, monitoring or estimating the impairments such as optical signal-to-noise ratio (OSNR), chromatic dispersion (CD), polarization mode dispersion (PMD) can be achieved. The OPM can facilitate link supervision, impairment-aware routing and channel compensation. With theoretical analysis, three problems in channel characterization are solved for different generations of optical transmission system.

For non-coherent system, OSNR monitoring technique using polarization diversity is studied. The relation among input polarization states, electrical bandwidth and monitored OSNR results are theoretically investigated. It is proposed to use a polarization scrambler (PS) to search for the particular polarization states. Based on the analysis, a simplified OSNR monitoring technique is derived, which advantages in low cost and insensitivity of CD and first order PMD. The proposed technique can achieve modulation format transparent OSNR monitoring at the intermediate nodes for non-coherent systems and enable reliable network operation, supervision and routing.

For coherent system, a fast and robust CD estimation technique is proposed by examining a peak in the auto-correlation of the received signal power waveform. Theoretical analysis shows that the location of the peak is indicative of the accumulated CD. Due to its search-free nature, the required number of symbols is less than 1/10 of

other searching based CD estimation techniques. The reduction in required symbols can reduce the delay and extra payload in building up a connection. It can also help to reduce the complexity and power consuming of the receiver DSP.

Finally the channel of quasi-single-mode (QSM) is analytically characterized, in which the single-mode fiber (SMF) is replaced by few-mode fiber (FMF) to suppress the nonlinear effect and allows higher signal power and OSNR, with the tradeoff of additional loss by higher mode power stripped off at the end of each link and the presence of modal coupling induced multi-path interference (MPI). In order to characterize the aggregate performance improvement, 6 liner-polarized (LP) mode model is built to analyze the effect of modal coupling strength, differential modal delay (DMD) and differential modal loss (DML). It is found that for weakly coupling and long-haul transmission, the in- and cross-polarization channels can be modeled as Rayleigh and Gaussian channels. And the MPI effect can be adaptively compensated by the decision-directed least-mean-square (DD-LMS) algorithm for practical implementations. In the experiments for 256 Gb/s PM-16-QAM system, the transmission distance for QSM reaches 2600 km with 501-tap DD-LMS, almost twice as that with SMF. After theoretically and experimentally investigating the performance improvement of QSM over traditional SMF, QSM is shown to be a good candidate in future optical communication systems. The full compatibility with SMF system makes it easy in upgrading the current links and deploying new links. The marginal incremental cost coming from MPI compensation DSP brings about almost double

transmission distance as a reward. All of these attributes grant QSM a promising technique in the long-haul transmissions.

Publications Arising From the Thesis

Journal articles

[1] **Qi Sui**, Alan Pak Tao Lau and Chao Lu, "OSNR monitoring in the presence of first-order PMD using polarization diversity and DSP," *IEEE Journal of Lightwave Technology*, vol. 28, no. 15, pp. 2105-2114, Aug. 2010.

[2] **Qi Sui**, Alan Pak Tao Lau and Chao Lu, "Fast and robust blind chromatic dispersion estimation using auto-correlation of signal power waveform for digital coherent systems," *IEEE Journal of Lightwave Technology*, vol. 31, no. 2, pp. 306–312, Jan. 2013.

[3] Alan Pak Tao Lau, Yuliang Gao, **Qi Sui**, Dawei Wang, Qunbi Zhuge, Mohamed H. Morsy-Osman, Mathieu Chagnon, Xian Xu, Chao Lu and David V. Plant, "Advanced DSP Techniques Enabling High Spectral Efficiency and Flexible Transmissions," *IEEE Signal Processing Magazine*, vol. 31, no. 2, pp. 82-92, Mar. 2014.

[4] **Qi Sui**, HongYu Zhang, John D. Downie, William A. Wood, Jason Hurley, Snigdharaj Mishra, Alan Pak Tao Lau, Chao Lu, Hwa-Yaw Tam and P. K. A. Wai, "Long-haul Quasi-Single-Mode Transmissions using Few-Mode Fiber in presence of Multi-Path Interference," *Manuscript submitted for publication*.

Conference articles

[5] **Q. Sui**, C. Lu and A. P. T. Lau, "In-Band OSNR Monitoring by Polarization Diversity and Electronic Signal Processing," *The 14th OptoElectronics and Communications Conference (OECC)*, paper FW5, 2009.

[6] A. P. T. Lau, **Q. Sui**, T. S. R. Shen, F. N. Khan and C. Lu, "Optical Performance Monitoring Using Statistical Signal Processing," *International Conference on Advanced Infocomm. Technology (ICAIT)*, Haikou, Hainan Island, China, Jul. 2010.

[7] A. P. T. Lau, **Q. Sui**, T. S. R. Shen, F. N. Khan and C. Lu, "Optical Performance Monitoring Using Statistical Signal Processing," *International Conference on Optical Communications Networks (ICOCN)*, Nanjing, China, Oct. 2010.

[8] C. Lu, A. P. T. Lau, F. N. Khan, **Q. Sui**, J. Zhao, Z. Li, H. Y. Tam and P. K. A. Wai, "Optical performance monitoring techniques for high capacity optical networks," in *7th International Symposium on Communication Systems Networks and Digital Signal Processing (CSNDSP)*, pp. 678 – 681, July 2010.

[9] A. P. T. Lau, **Q. Sui**, T. S. R. Shen, F. N. Khan and C. Lu, "Optical Performance Monitoring for Next Generaton Optical Networks: Progress and Challenges," *International Conference on Advanced Infocomm. Technology (ICAIT)*, Wuhan, China, Jul. 2011.

[10] **Q. Sui**, A. P. T. Lau and C. Lu, "Fast and robust chromatic dispersion estimation using auto-correlation of signal power waveform for DSP based-coherent systems," in *Proc. Opt. Fiber Commun. Conf. (OFC/NFOEC) 2012*, paper OW4G.3.

[11] C. Lu, A. P. T. Lau, **Q. Sui**, F. Khan and Z. Dong, "Optical Performance Monitoring: from direct detection to coherent optical communication systems," in *Commun. and Photonics Conference (ACP)*, 2012, Paper AF3G.3.

[12] A. P. T. Lau, Y. Gao, **Q. Sui**, D. Wang, Q. Zhuge, M. M. Osman, M. Chagnon, X. Xu, C. Lu and D. V. Plant, "Beyond 100 Gb/s: Advanced DSP Techniques Enabling High Spectral Efficiency and Flexible Optical Communications," in *International Conference on Optical Commun. and Networks (ICOCN) 2013*, paper SC1In9.

[13] **Q. Sui**, H. Zhang, J. D. Downie, W. A. Wood, J. Hurley, S. Mishra, A. P. T. Lau, C. Lu, H. Y. Tam and P. K. A. Wai, "256 Gb/s PM-16-QAM Quasi-Single-Mode Transmission over 2600 km using Few-Mode Fiber with Multi-Path Interference Compensation," in *Proc. Opt. Fiber Commun. Conf. (OFC/NFOEC) 2014*, paper M3C.5.

[14] A. P. T. Lau, **Q. Sui**, H. Y. Tam, C. Lu and P. K. A. Wai, J. D. Downie, W. A. Wood, J. Hurley and S. Mishra, "Long-haul Quasi-Single-Mode Transmission using Few-Mode Fiber with Multi-Path Interference Compensation" in *Proc. of International Conference on Optical Internet 2014*, paper FB3–1.

[15] A. P. T. Lau, Y. Gao, **Q. Sui**, D. Wang and C. Lu, "Advanced DSP for High Spectral Efficiency and Flexible Optical Communications," in *Signal Processing in Photonic Commun. 2014*, paper SM2D.1.

[16] A. P. T. Lau , Y. Gao, **Q. Sui**, D. Wang, C. Lu, "Advanced DSP Techniques Enabling Flexible Transmissions and Elastic Optical Networks," in *International Photonics and OptoElectronics Meetings (POEM) 2014*, paper OTh3B.3.

Acknowledgments

I am deeply indebted to my supervisor Dr. Alan Pak Tao Lau, for his guidance, encouragement, support, patients and sharing of experiences.

Specially thanks Prof. LU Chao, from whom I learnt the attitude of research and the way of thinking.

Thanks to all the other members of Photonics Research Centre, for their help and insightful discussion.

Thanks to my parents who give me support all the way along.

Table of Contents

List of Abbreviations

ADC	Analogue-digital converter
ASE	Amplified spontaneous emission
BER	Bit error rate
CD	Chromatic dispersion
CPE	Carrier phase estimation
DCF	Dispersion compensating fiber
DD-LMS	Decision directed-least mean square
DGD	Differential group delay
DMD	Differential modal delay
DML	Differential modal loss
DOF	Degree of freedom
DOP	Degree of polarization
DPSK	Differential phase shift keying
DQPSK	Differential quadrature phase shift keying
DSP	Digital signal processing
ECL	External cavity laser
EDFA	Erbium doped fiber amplifier
FMF	Few-mode fiber
FTN	Faster-than Nyquist
FWM	Four-wave mixing

GVD	Group velocity dispersion
i.i.d.	independent identically distributed
ICI	Inter-carrier-interference
ISI	Inter-symbol-interference
LP	Linear polarized
MIMO	Multi-input multi-output
MMSE	Minimum mean square error
MPI	Multi-path interference
NRZ	Non-return-to-zero
OFDM	Orthogonal frequency-division multiplexing
OOK	On-off-keying
OPM	Optical performance monitoring
OSNR	Optical signal-to-noise ratio
PBS	Polarization beam splitter
PDF	Probability density function
PDL	Polarization dependent loss
PM	Polarization multiplexing
PMD	Polarization mode dispersion
PS	Polarization scrambler
PSD	Power spectral density
PSP	Principal states of polarization

QAM	Quadrature amplitude modulation
QPSK	Quadrature phase-shift keying
QSM	Quasi-single mode
ROADM	Reconfigurable optical add/drop multiplexer
RZ	Return-to zero
SD-FEC	soft decision-forward error correction
SDM	Space division multiplexing
SE	Spectral efficiency
SMF	Single mode fiber
SOP	State of polarization
SPM	Self-phase modulation
WDM	Wavelength division multiplexed
XPM	Cross-phase modulation

Chapter 1 Introduction

Fueled by the ever-increasing bandwidth requirement all over the world, extensive efforts have been made in pushing the optical transmission speed to their limit. Over the past few years, optical communications have experienced a transition from non-coherent system where information is encoded on the amplitude, or energy, of the carrier signal to coherent system where information is encoded in both the amplitude and phase of the carrier. In addition, receiver signal processing has long been simple bit decision algorithms for traditional non-coherent systems and complex statistical signal processing is only relevant for copper-wire or wireless communications. Advances in digital signal processing (DSP) hardware over the past decade together with coherent detection have paved the way for non-trivial signal processing algorithms to be finally applied in optical communication systems. The field of optical communications has undergone a period of intensive research in applying various signal processing techniques from wireless communications and optimization/custom designs for fiber-specific features such as high symbol rates and fiber nonlinearity effects.

1.1. Evolution of Optical Communication Systems

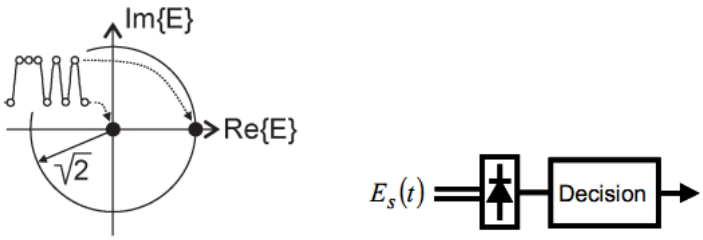
1.1.1. Non-Coherent System

On-off-keying (OOK) is a popular modulation format in optical communication systems with data rates 10 Gbit/s or below. The information is encoded in the power

waveform of signal which can be received by a simple photo diode, as shown in Fig. 1.1. The output photo-current can be written as

$$I_{OOK}(t) = R\left|E_s(t)\right|^2 . \tag{1.1}$$

where R denotes the responsivity of the photo-diode, $E_s(t)$ denotes the received electrical field.



(a) (b)

Fig. 1.1 (a) Constellation diagram and (b) detection schemes for OOK system.

There are 2 major limitations for non-coherent systems [1]. First, the detection for OOK are based on pulse energy and prohibit the use of more than one degree of freedom (DOF) in single mode fiber (SMF), which actually supports 4 DOFs, with two quadratures in the two orthogonal polarization states. Second, the loss of phase information prevents the appropriate equalization of the inline impairments. e.g. the chromatic dispersion (CD) has to be compensated for by the dispersion compensating fiber (DCF) within each span and consequently induce a power or optical signal-to-noise (OSNR) penalty.

1.1.2. Differentially-Coherent System

In differential-coherent system such as differential phase-shift keying (DPSK) and differential quadrature phase-shift keying (DQPSK), the information is encoded in the

phase shift between the neighboring symbols. As shown in Fig. 1.2, the output photo-currents for DPSK and DQPSK are given by

$$I_{DPSK}(t) = R \cdot \mathrm{Re}\left\{E_s(t)E_s^*(t-T_s)\right\} \tag{1.2}$$

and

$$\begin{cases} I_{DQPSK,i}(t) = \dfrac{1}{2}R \cdot \mathrm{Re}\left\{E_s(t)E_s^*(t-T_s)\right\} \\ I_{DQPSK,1}(t) = \dfrac{1}{2}R \cdot \mathrm{Im}\left\{E_s(t)E_s^*(t-T_s)\right\} \end{cases} \tag{1.3}$$

where T_s is the symbol period.

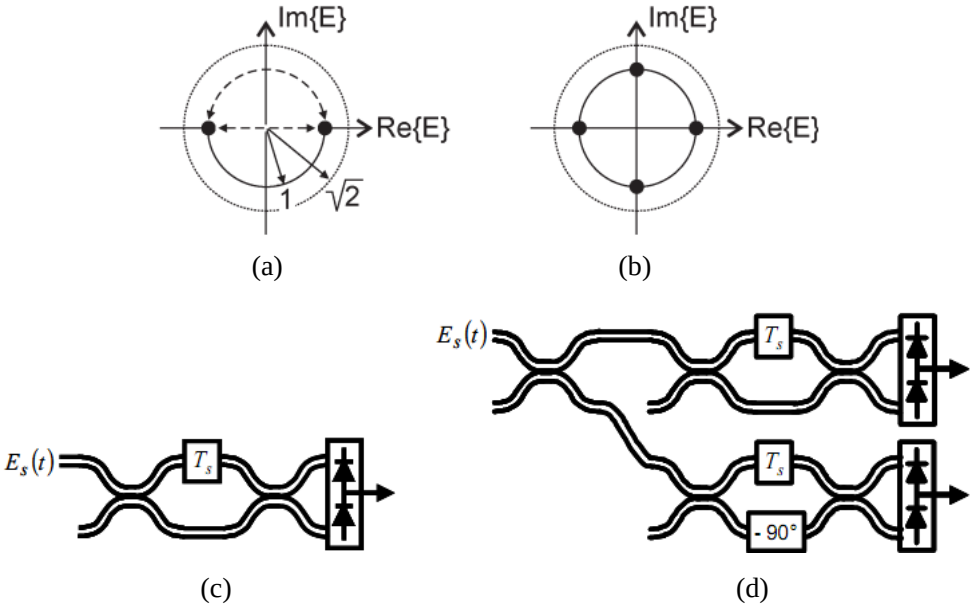


Fig. 1.2. Constellation diagrams of (a) DPSK, (b) DQPSK and detection schemes for (c) DPSK, (b) DQPSK systems.

Compared with non-coherent systems, DQPSK allows information encoded in 2 DOFs and thus has a higher spectral efficiency (SE). Although CD and polarization mode dispersion (PMD) equalization is still infeasible in differential-coherent system, the sensitivity and tolerance for CD, PMD and narrow-filtering are better than that of

non-coherent system [2].

1.1.3. Coherent System

The structure of coherent receiver is as shown in Fig. 1.3. The output electrical signal is given by

$$\begin{cases} I_{x(y)i}(t) \propto \left|E_{sx(y)}(t)+E_{LO}(t)\right|^2-\left|E_{sx(y)}(t)-E_{LO}(t)\right|^2=2\,\mathrm{Re}\left\{E_{sx(y)}(t)E_{LO}(t)\right\} \\ I_{x(y)q}(t) \propto \left|E_{sx(y)}(t)+jE_{LO}(t)\right|^2-\left|E_{sx(y)}(t)-jE_{LO}(t)\right|^2=2\,\mathrm{Im}\left\{E_{sx(y)}(t)E_{LO}(t)\right\} \end{cases} \tag{1.4}$$

where $E_{sx(y)}(t)$ denotes the *x*(*y*)-polarization tributary of $\boldsymbol{E}_s(t)$. With coherent receiver, we are able to extract all the information in the 4 DOFs, hence enables signal with high-order modulation formats such as quadrature phase-shift keying (QPSK) and higher order quadrature amplitude modulation (QAM) transmitted simultaneously in two orthogonal polarization states as a 2×2 multi-input multi-output (MIMO) system to achieve a much higher SE than that of non-coherent and differential-coherent system. E.g. *M*-QAM allows $\log_2 M$ bits simultaneously transmitted in a symbol period, whereas polarization multiplexed (PM) signaling further doubles this rate. A channel capacity comparison is made from single polarization DPSK to PM-16-QAM at 10 G baud rate, as shown in Fig. 1.4

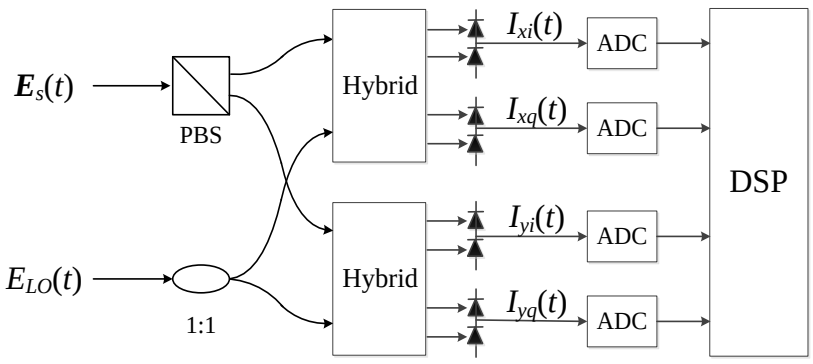


Fig. 1.3. Coherent receiver structure. PBS: polarization beam splitter. ADC: analogue-digital converter. DSP: digital signal processing.

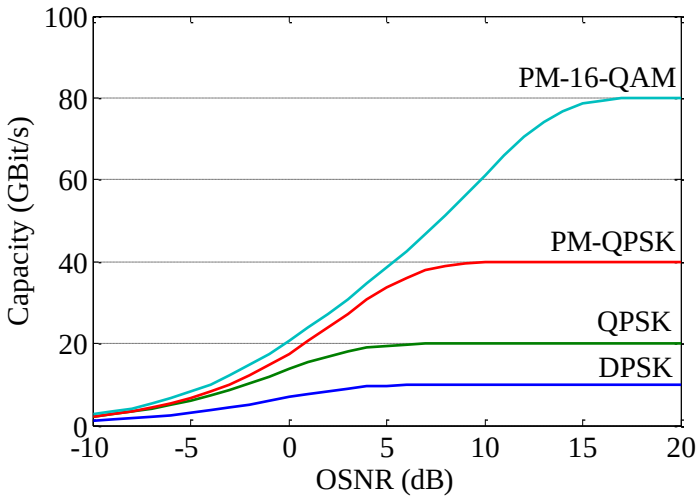


Fig. 1.4. Channel capacity comparison for different modulation formats at a symbol rate of 10 G baud.

After coherent detection of the signal, with analogue-to-digital converters (ADCs) operating above Nyquist sampling rate, arbitrarily complex DSP algorithms are allowed in principle to compensate various transmission impairments and transceiver impairments to recover the transmitted signal [3]-[9]. A basic set of DSP algorithms for a standard receiver DSP unit is shown in Fig. 1.5. After beating with the LO, the received optical signal is converted into electrical domain and digitized to 2 samples per symbol. Then the DSP compensates for the hybrid imperfection and normalizes the received signal power. The static channel impairments such as CD and inline filters are compensated before the dynamic equalization, which handles the time-varying polarization rotation and PMD effects. Then the signal is down-sampled to 1 sample per symbol with the timing recovery algorithm ensuring each sample locates at the

center of each symbol. The following frequency estimation compensates for the frequency difference between the transmitter laser and LO, which can also be seen as a steady and fast phase variation. And the residual temporary and slow-varying phase change is estimated and compensated afterwards. Finally symbol decision is made and decoded with FEC.

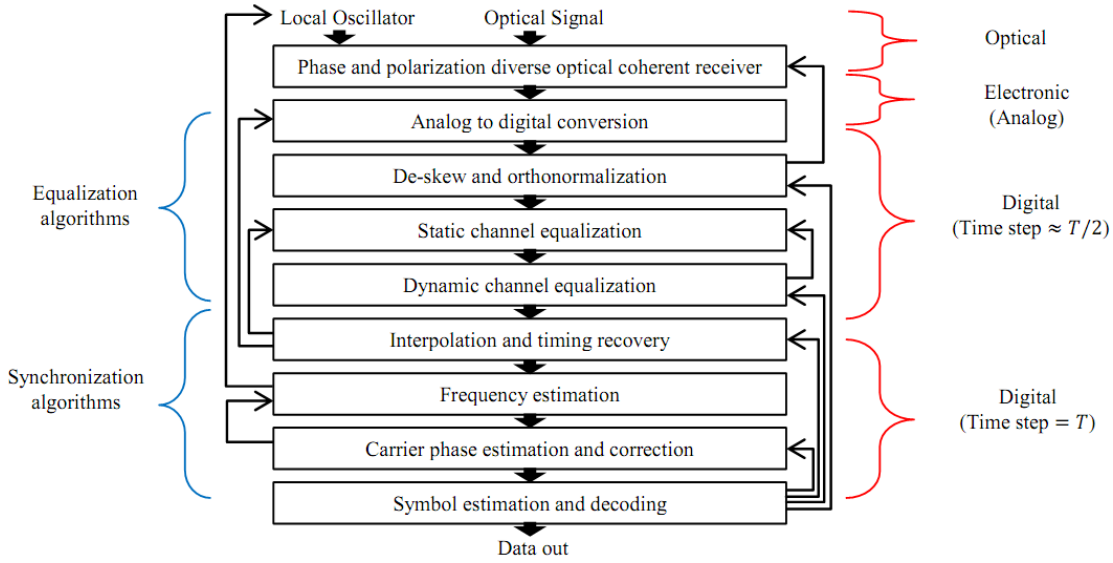


Fig. 1.5. Basic DSP structure in coherent receiver that compensates various transmission impairments and transceiver imperfections [5].

1.2. Degradations in the Transmission Channel

1.2.1. Amplified Spontaneous Emission Noise

The noise in long-haul optical communication systems mostly comes from amplified spontaneous emission (ASE) of the inline amplifier, such as erbium doped fiber amplifier (EDFA) and Raman amplifier. For a system with N_A amplifiers operating with the same gain G, the total ASE noise power spectral density (PSD) is given by

$$S_{ASE} = 2n_{sp}h\nu_0 N_A(G-1) \tag{1.5}$$

where n_{sp} is the spontaneous-emission factor and $h\nu_0$ is the photon energy [10]. The factor takes into account the unpolarized nature of ASE. According to Shannon limit, the OSNR fundamentally limits the channel capacity. In a specific transmission link, the OSNR budget determines various link specifications such as signal power, noise figure of the amplifiers, transmission length, baud rate, modulation format, receiver sensitivity and pulse shaping. And all of these choices are carefully chosen to ensure that the bit error ratio (BER) is well controlled under a threshold for reliable data detection.

1.2.2. Chromatic Dispersion

Chromatic Dispersion refers to the phenomenon of dependence of the medium response on the optical frequency when an electromagnetic wave interacts with the bound electrons of a dielectric. The CD manifests itself through the refractive index $n(\omega)$ varying with frequency ω [11]. Mathematically, CD can be expressed by expanding the mode-propagation constant β in a Taylor series around the center frequency ω_0:

$$\beta(\omega) = n(\omega)\frac{\omega}{c} = \beta_0 + \beta_1(\omega-\omega_0) + \frac{1}{2}\beta_2(\omega-\omega_0)^2 + \cdots, \tag{1.6}$$

where c is the speed of light and β_m is the mth derivative of β. β_1 and β_2 are related to $n(\omega)$ and its derivatives through

$$\beta_1 = \frac{1}{v_g} = \frac{n_g}{c} = \frac{1}{c}\left(n + \omega\frac{dn}{d\omega}\right), \tag{1.7}$$

$$\beta_2 = \frac{1}{c}\left(2\frac{dn}{d\omega} + \omega\frac{d^2n}{d\omega^2}\right), \tag{1.8}$$

where n_g and v_g are the group index and group velocity, respectively, as shown in Fig. 1.6. β_1 relates with the group velocity which defines the propagation speed of envelop. While β_2 represents group-velocity dispersion (GVD) and causes pulse broadening. The dispersion parameter D is often used in practice. It is defined as $d\beta_1 / d\lambda$ and related to β_2 and n as

$$D = \frac{d\beta_1}{d\lambda} = -\frac{2\pi c}{\lambda^2}\beta_2 = -\frac{\lambda}{c}\cdot\frac{d^2n}{d\lambda^2}. \tag{1.9}$$

And the GVD induced degradation can described by an all-pass optical filter

$$H_{GVD}(z,\omega) = \exp\left(i\beta_2\omega^2 z/2\right), \tag{1.10}$$

where z is the transmission length.

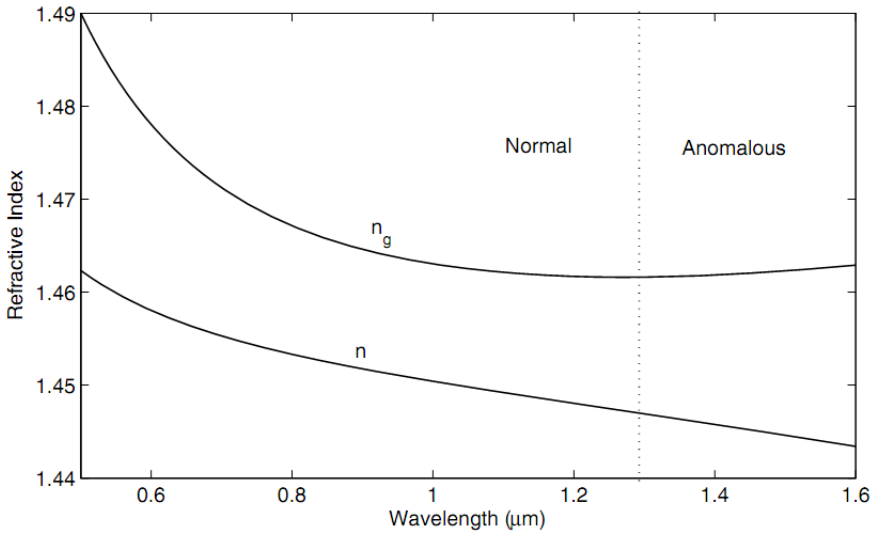


Fig. 1.6. Variation of refractive index n and group index n_g with wavelength for fused silica [11].

For a single carrier system, CD may results in the pulse broadening and thus introduces inter-symbol-interference (ISI). To avoid such detrimental effects, CD is

usually compensated by either DCF in non-coherent systems or electronic equalization in coherent systems.

1.2.3. Polarization-Mode Dispersion

Polarization-mode dispersion arises from the random birefringence of the fiber, which stems from the geometrical deformation or random imperfections and asymmetries of the fiber. PMD is often defined in Stokes space, where for a polarized input polarization vector $\boldsymbol{S}_{in}$, the out polarization vector $\boldsymbol{S}_{out}$ is related with PMD vector $\boldsymbol{\Omega}$ in a frequency dependent manner:

$$\frac{d}{d\omega}\boldsymbol{S}_{out} = \boldsymbol{\Omega} \times \boldsymbol{S}_{out}. \tag{1.11}$$

And $\boldsymbol{\Omega}$ is a function of ω as well. Nth order PMD is characterized by the Nth derivative of the vector with respect to ω. And its 1st order component, differential group delay (DGD), which models the relative delay between the two principal states of polarization (PSP), can be expressed in Jones space as

$$\boldsymbol{H}_{PMD} = \boldsymbol{R}_2 \begin{bmatrix} e^{-i\omega\Delta\tau/2} & 0 \\ 0 & e^{i\omega\Delta\tau/2} \end{bmatrix} \boldsymbol{R}_1, \tag{1.12}$$

where $\Delta\tau$ is the DGD, $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ denote unitary polarization rotation matrices satisfying

$$\boldsymbol{R}_k = \begin{bmatrix} \cos\theta_k e^{-i\varphi_k/2} & \sin\theta_k e^{i\phi_k/2} \\ -\sin\theta_k e^{-i\phi_k/2} & \cos\theta_k e^{i\varphi_k/2} \end{bmatrix}. \tag{1.13}$$

Unlike the static nature of CD, PMD is a stochastic dispersive impairment. For a specific link configuration, the DGD follows the Maxwellian distribution [12]

$$f_{\Delta\tau}(x) \sim \frac{32x^2}{\overline{\Delta\tau}^3 \pi^2} \exp(-\frac{4x^2}{\overline{\Delta\tau}^2 \pi}) \tag{1.14}$$

where $\overline{\Delta\tau}$ denotes the mean DGD. Thus PMD may cause fatal error with very small probability.

1.2.4. Polarization Dependent Loss

The existence of polarization dependent loss (PDL) results in the launched signal suffers a power loss in selective input state of polarization (SOP). The mathematical expression of PDL is given by

$$\boldsymbol{H}_{PDL} = \boldsymbol{R}^{\dagger} \begin{bmatrix} 1 & 0 \\ 0 & e^{-PDL} \end{bmatrix} \boldsymbol{R} \tag{1.15}$$

where '†' denotes conjugated transpose and $\boldsymbol{R}$ follows the form of (1.13). PDL is usually caused by the aging or breakdown of the inline devices and components. For single polarization system, the PDL is equivalent with an excess of insertion loss, whereas for polarization multiplexing (PM) system, it may results in uneven OSNR in the two polarization channels, and the consequent BER outage.

1.2.5. Kerr Nonlinearity

Kerr effect is the change of refractive index with the electric field, manifesting itself as self-phase modulation (SPM), cross-phase modulation (XPM) and four-wave mixing (FWM). In presence of Kerr nonlinearity, the propagation of electrical field in SMF with center frequency ω_0 is governed by

$$\frac{\partial E}{\partial z}+\beta_1\frac{\partial E}{\partial t}+\frac{i\beta_2}{2}\cdot\frac{\partial^2 E}{\partial t^2}+\frac{\alpha}{2}E=i\gamma(\omega_0)\left|E\right|^2 E\,, \tag{1.16}$$

which is known as the scalar form of nonlinear Schrödinger equation. Unlike CD and PMD, the existence of the nonlinearity makes the channel transfer function un-invertible and the uncompensated nonlinear effect bounds the aggregate channel capacity together with ASE noise. In such a nonlinear channel, the capacity is bounded by the nonlinear Shannon capacity, which is achieved with an optimum launched signal power [13]-[15].

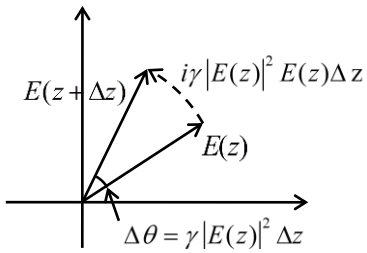


Fig. 1.7 Kerr Nonlinearity induced Phase shift.

1.2.6. Modal Dispersion in Few-Mode Fibers

The cross-section of fiber is shown in Fig. 1.8, where n_{co} and n_{cl} stands for the refractive index of core and cladding, respectively and *a* denotes the radius of the core. Then the generalized frequency *V* and generalized guide index *b* are given by

$$V=ka\sqrt{n_{co}^2-n_{cl}^2} \tag{1.17}$$

$$b=\frac{N^2-n_{cl}^2}{n_{co}^2-n_{cl}^2} \tag{1.18}$$

where *k* is the vacuum vector and *N* is the effective guide index. The excited linear polarized (LP) mode are defined by both *b* and *V*, as shown in Fig. 1.9. When *V* is

smaller than 2.408, only the fundamental LP mode (LP_{01}) can be excited, which is the SMF case.

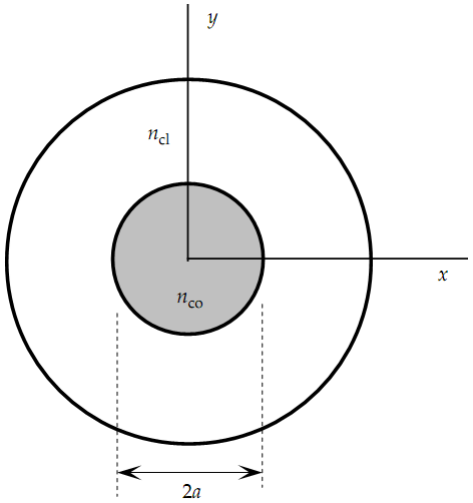


Fig. 1.8. Cross section of fiber.

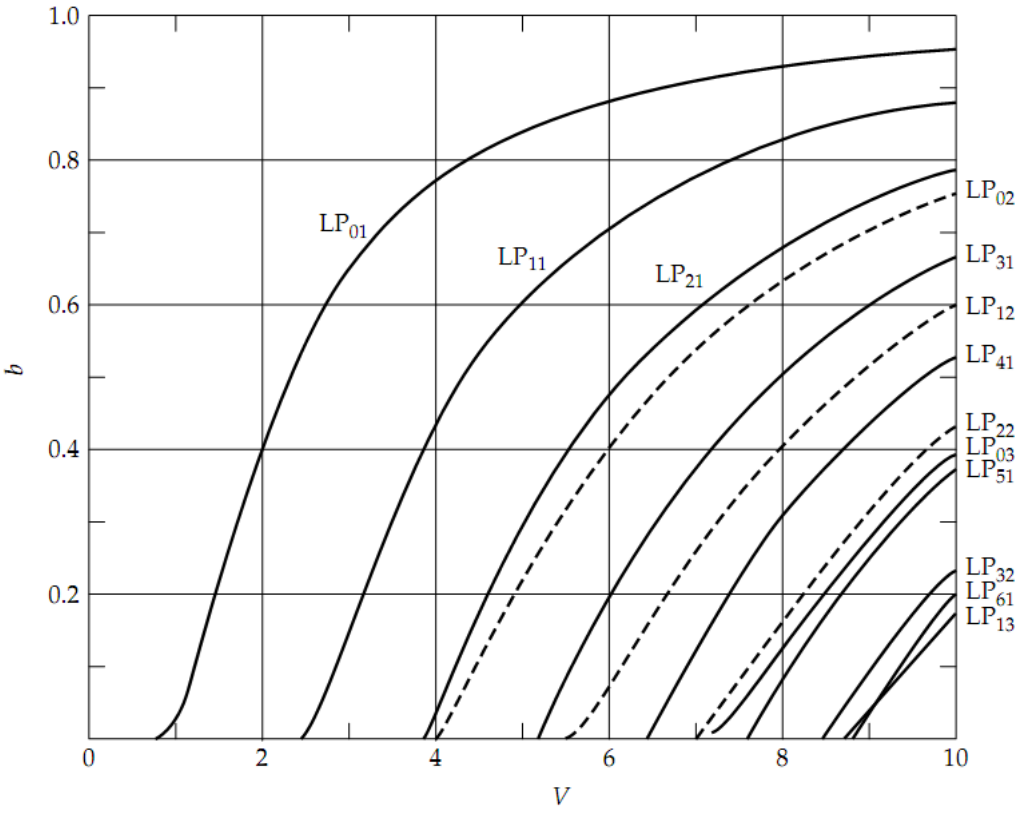


Fig. 1.9. *b*-*V* characteristics of weakly guiding fibers [16].

One way to mitigate the nonlinear effect is enlarging the core area so as to increase

the effective area of the fundamental mode. However, constrained by (1.17), the increase of a and hence V permits higher order LP mode excited in the fiber as in Fig. 1.9. E.g. when $V \approx 3$, the fiber would support one LP_{01} and two LP_{11} mode and hence is named as few-mode fiber (FMF). Taking the advantage of weaker nonlinear effect, traditional SMF can be replaced by FMF with the signal launched in only LP_{01} mode to accommodate higher signal power and thus increasing the OSNR. Since all the higher order modes are not excited when launching the signal power, such optical transmission system is called quasi-single mode (QSM) transmission. However, due to the nature of fiber, modal coupling is inevitable in fibers supporting more than one LP mode and the pulse broadening effect caused thereby is named as inter-modal dispersion. This is characterized by the modal coupling strength, differential modal delay (DMD) and differential modal loss (DML). In principle, its transfer function can be expressed as a invertible 2×2 matrix and hence can be compensated by 2×2 equalizers in MIMO processing as shown in Fig. 1.10.

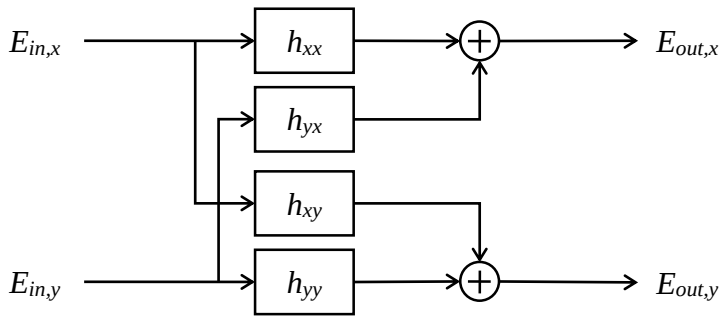


Fig. 1.10. 2×2 equalizer to mitigate inter-modal dispersion in a polarization-multiplexed QSM transmission system.

It should be noted that the FMF suffers a higher bending loss compared with SMF and bending FMF also introduces more modal coupling and undesired inter-modal

dispersion. However, QSM take the most advantages in long-haul transmission scenario, where the bending effect occurs much less frequently than that in metropolitan area network.

1.3. Objectives of Channel Characterization

1.3.1. Non-Coherent Systems

In the legacy non-coherent systems such as OOK and differential DPSK with direct detection, the objective of channel characterization is to monitor the impairments such as OSNR, CD and PMD at intermediate node in order to achieve: 1) system diagnosis, fault detection and localization; 2) impairment-aware routing; and 3) adaptive resource allocation with signal wavelength, signal power, channel bandwidth, tunable compensation, data coding and path determination to maximize the network efficiency. The monitoring techniques are desired to be fast, accurate, robust, cost-effective, transparent to modulation format and date rate, and able to isolate different impairments.

1.3.2. Coherent Systems

In coherent systems with DSP, linear impairments such as CD and PMD are no longer needed to be monitored since they can be electronically compensated. Nonetheless, they need to be correctly estimated before they can be compensated part of the objective of channel characterization becomes estimating and compensating the impairments in the coherent receiver, where the CD estimation becomes vital in the receiver DSP since the

accumulated CD can be in the order of 100,000 ps/nm while the accuracy of estimation are required to be within around ±200 ps/nm. This is particularly difficult as CD compensation is typically in the first step of the receiver DSP and hence the information of SOP, frequency offset and timing phase are all unavailable in CD estimation. In addition, in order to reduce the latency in initializing connection, the number of symbols consumed in CD estimation should be as few as possible. This results in a problem of its own as typical DSP techniques for CD estimation will take an unacceptably long time and it might not even converge to the right value due to the long tail of the CD impulse response.

1.3.3. Quasi-Single Mode Transmissions

QSM transmission is a promising candidate for next generation long-haul optical communications, in which the SMF is replaced by FMF with single-mode operation in the lowest LP mode to mitigate the Kerr nonlinearity through the larger effective area and thus achieving a higher operable signal power and OSNR. In comparison with space division multiplexing (SDM), QSM benefits in full compatibility with current SMF systems by retaining all the single-mode transponders and inline devices whereas upgrading SMF to SDM systems requires multimode EDFA, ROADM and other components which are far from mature. However, mode coupling effects and differential mode delay results in excess loss from higher order mode stripping and multi-path interference (MPI). Estimating the MPI transmission channel will be vital to

its appropriate compensation so as to fully realize the benefits of QSM transmissions

1.4. Dissertation Outline

Chapter 1 presents the background, motivation and organization of the dissertation. The channel impairments are introduced. The objectives of channel characterization is discussed for non-coherent, coherent and QSM systems, respectively.

In chapter 2, OSNR monitoring using polarization diversity and DSP in the presence of first-order PMD is analytically studied for non-coherent systems. The conditions on the input signal SOP, PSP and DGD of the fiber, polarization beam splitter (PBS) axes and receiver electrical low-pass filter bandwidth such that the monitored OSNR is least affected by PMD is derived. A new OSNR monitoring technique proposed by using a PS before the PBS to achieve such conditions. In addition, two special polarization states for which the proposed technique can operate independent of first-order PMD are identified. The theoretical predictions are verified through simulations and experiments. The analytical insights developed represent a more general and unified framework of which various OSNR monitoring techniques based on polarization diversity proposed in the literature can be understood as special cases.

In chapter 3, a CD estimation technique by using the auto-correlation of signal power waveform is proposed for DSP-based coherent systems. For single-carrier systems with Gaussian pulse shapes, it is analytically shown that the auto-correlation of signal power waveform contains a peak whose location is indicative of the

accumulated CD and is insensitive to ASE noise, laser frequency offset, laser phase noise and first-order PMD. As a linear scan of a preset range of CD is not needed, the number of symbols of the proposed technique is dramatically reduced. Simulation results show that to estimate the link CD for 28 Gbaud PM-QPSK and PM-16-QAM systems, 8192 symbols are sufficient for non-return-to-zero (NRZ)/return-to zero (RZ) pulse shapes and 32768 symbols are sufficient for root-raised-cosine pulse shapes with the roll-off factor as low as 0.1 with estimation error below 175 ps/nm and 400 ps/nm, respectively.

In chapter 4, the analytical model is built for QSM system using 6-spatial-polarization mode FMF for signal transmission. It is shown that in the weakly coupling regime, the QSM channel approaches a Gaussian random process in frequency. MPI compensation filters are derived and OSNR penalties due to MPI and signal power loss from higher-order modes are characterized in terms of coupling strength, DMD and DML. The experiment is carried out for 256 Gb/s PM-16-QAM QSM transmissions over a record distance of 2600 km with 100-km span using decision directed-least mean square (DD-LMS) algorithm for MPI compensation.

At the end the conclusion is made and the possible research directions in the future are discussed.

Chapter 2 OSNR Monitoring Using Polarization Diversity

2.1. Introduction

OSNR is an important parameter that determines the performance of an optical communication system. Its effective monitoring is very important in ensuring the proper operation of an optical network [17]. Linearly interpolating the ASE noise in optical amplifiers for wavelength division multiplexed (WDM) systems has been used for OSNR estimation [18]. However, with the increase of spectral efficiency, the gap between neighboring channels may no longer exist to measure the ASE spectrum. And the introduction of ROADM in WDM optical networks makes the in-band and out-band ASE with different power spectral density. Hence the effectiveness of the out-band OSNR technique is significantly limited. Polarization nulling method [19]-[22] has been introduced to overcome this problem. It measures in-band OSNR based on the fact that signal is polarized while the ASE noise is unpolarized. It is simple, insensitive to CD and bit rate independent. However, the scheme will underestimate the value of OSNR in the presence of PMD. PMD causes the change of SOP in the signal spectrum and leads to signal depolarization. This makes it impossible to separate signal and noise using a simple polarizer as in polarization-nulling scheme. With the increase of data transmission rate in optical communication systems, the effect of PMD is especially important due to the increased signal bandwidth. It is essential to study the OSNR

monitoring schemes that are insensitive to PMD. Degree of polarization (DOP) based scheme has been proposed [23]. By measuring the SOP of different spectral component, DOP variation caused by PMD can be found and this enables the effect of PMD in OSNR measurement be removed. The main drawback of this scheme is the requirement of a polarimeter, which significantly increases the system complexity. Other proposed schemes include orthogonal-polarization heterodyne OSNR monitoring [24] and off-center narrow band filtering based OSNR monitoring techniques [25]. Both are based on the assumption that signal can be considered to be polarized if the optical filter bandwidth is narrow enough. However, the measurement is decided by the bandwidth of the optical filter used and precise spacing control of the center wavelengths of the two optical filters is also necessary.

In 2005, an OSNR monitoring scheme using polarization diversity was proposed [26]-[27]. It separates the signal to be monitored into two polarization paths using a PBS followed by a signal processing unit to calculate OSNR. It claims that the PMD effect can be neglected as long as the product of electrical bandwidth and DGD is much smaller than 1. And it is demonstrated with 70 MHz bandwidth 60 ps DGD. However, no in depth analysis was provided on the relations between DGD, low-pass filter bandwidth, input signal polarization and the PSP of the fiber such that OSNR can be accurately monitored. In this chapter, we present a detailed analytical study on OSNR monitoring using polarization diversity and DSP with a PS at the receiver [28]-[33]. In absence of PMD, we analytically derive OSNR estimates obtained from the first- and

second-order statistics of the photo-detector outputs. With first-order PMD, we derive conditions on the input signal polarization states, DGD and PSP of the fiber, PBS axes and electrical low-pass filter bandwidth for OOK, BPSK and/or DPSK systems such that OSNR monitoring is least sensitive to the effect of PMD. In addition, two specific conditions such that OSNR monitoring is truly insensitive to first-order PMD will be derived and a corresponding simplified experimental setup is proposed. The theoretical insights developed can be viewed as a more general formulation for understanding in-band OSNR monitoring using polarization diversity in the presence of first-order PMD. Simulations are conducted to verify the theoretical predictions.

2.2. System Model

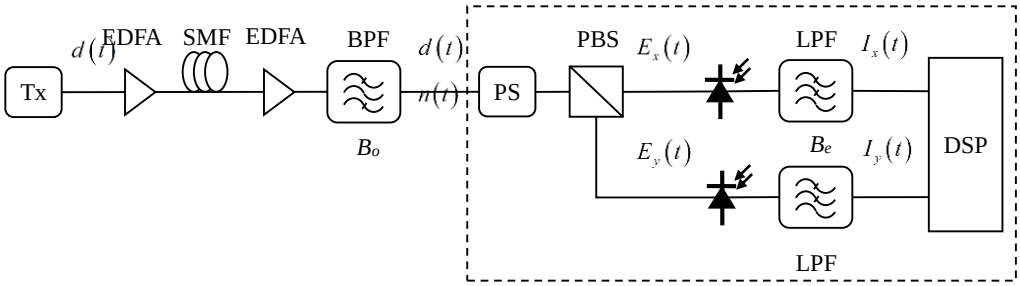


Fig. 2.1. OSNR monitoring using polarization diversity and DSP. Tx: Transmitter; BPF: band-pass filter; PS: polarization scrambler; PBS: polarization beam splitter; LPF: low-pass filter; DSP: digital signal processing unit.

Consider an optical communication system in which ASE noise is introduced by inline optical amplifiers located along the transmission link as shown in Fig. 2.1. Assume that the effect of CD is negligible or has been compensated by inline devices such as dispersion compensating fibers. Fiber nonlinearity and multi-channel effects are also

neglected for simplicity. At the receiver, an optical band-pass filter with bandwidth B_o is used to filter out the signal. Here we assume that B_o is large enough so that the signal passes through the filter without distortion. Let $d(t)$ be the baseband representation of the transmitted signal and the signal after the optical band-pass filter. Note that for modulation formats such as OOK, DPSK or BPSK, $d(t)$ is real while it is complex for QPSK or higher-order QAM systems. A PS and a PBS are used to split the received signal into two orthogonal polarizations. Without loss of generality, one can assume that the PBS output polarizations are aligned with the x and y axis. Let $n_x(t)$ and $n_y(t)$ be the collective ASE noise after the optical band-pass filter in the x and y polarization which are independent zero-mean complex circularly symmetric Gaussian noise process. If we denote $\tilde{D}(\omega)$, $\tilde{N}_x(\omega)$ and $\tilde{N}_y(\omega)$ as the Fourier transform of $d(t)$, $n_x(t)$ and $n_y(t)$ respectively, the Fourier transform of the PBS output $E_x(t)$ and $E_y(t)$ can be expressed as

$$\begin{aligned}\begin{bmatrix}\tilde{E}_x(\omega)\\ \tilde{E}_y(\omega)\end{bmatrix} &= \begin{bmatrix}\alpha_2 & \beta_2\\ -\beta_2^* & \alpha_2^*\end{bmatrix}\begin{bmatrix}e^{j\omega\Delta\tau/2} & 0\\ 0 & e^{-j\omega\Delta\tau/2}\end{bmatrix}\begin{bmatrix}\alpha_1 & \beta_1\\ -\beta_1^* & \alpha_1^*\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix}\tilde{D}(\omega)+\begin{bmatrix}\tilde{N}_x(\omega)\\ \tilde{N}_y(\omega)\end{bmatrix}\\ &= \begin{bmatrix}\alpha_1\alpha_2\tilde{D}(\omega)e^{j\omega\Delta\tau/2}-\beta_1^*\beta_2\tilde{D}(\omega)e^{-j\omega\Delta\tau/2}\\ -\alpha_1\beta_2^*\tilde{D}(\omega)e^{j\omega\Delta\tau/2}-\beta_1^*\alpha_2^*\tilde{D}(\omega)e^{-j\omega\Delta\tau/2}\end{bmatrix}+\begin{bmatrix}\tilde{N}_x(\omega)\\ \tilde{N}_y(\omega)\end{bmatrix}\end{aligned} \quad (2.1)$$

where '*' denotes complex conjugate, $\Delta\tau$ denotes the DGD of the link, $\begin{bmatrix}\alpha_1 & \beta_1\\ -\beta_1^* & \alpha_1^*\end{bmatrix}$ models the projection of signal onto the PSP of the fiber and $\begin{bmatrix}\alpha_2 & \beta_2\\ -\beta_2^* & \alpha_2^*\end{bmatrix}$ models polarization de-multiplexing collectively by the PS and PBS. In this formulation, $|\alpha_1|^2+|\beta_1|^2=|\alpha_2|^2+|\beta_2|^2=1$ and the PS varies the received signal over all polarization

states through α_2 and β_2. The time-domain waveform of the PBS output can be written as

$$\begin{bmatrix} E_x(t) \\ E_y(t) \end{bmatrix} = \begin{bmatrix} \alpha_1\alpha_2 d\left(t+\Delta\tau/2\right) - \beta_1^*\beta_2 d\left(t-\Delta\tau/2\right) + n_x(t) \\ -\alpha_1\beta_2^* d\left(t+\Delta\tau/2\right) - \beta_1^*\alpha_2^* d\left(t-\Delta\tau/2\right) + n_y(t) \end{bmatrix}. \tag{2.2}$$

The PBS is then followed by the photo-detectors and electrical low-pass filters with bandwidth B_e and impulse response $h_e(t)$ and the output $I_x(t)$ and $I_y(t)$ are then sampled and processed in a DSP unit. In the following sections, we will study the use of $I_x(t)$ and $I_y(t)$ for OSNR monitoring under different degrees of polarization effects. Various proposed OSNR monitoring techniques based on polarization diversity will be presented under a unified framework and relationships among them will be analytically illustrated.

2.3. OSNR Monitoring in Absence of PMD

In the absence of PMD (i.e. when $\Delta\tau = 0$), the PBS output will be given by

$$\begin{bmatrix} E_x(t) \\ E_y(t) \end{bmatrix} = \begin{bmatrix} \left(\alpha_1\alpha_2 - \beta_1^*\beta_2\right) d(t) + n_x(t) \\ \left(-\alpha_1\beta_2^* - \beta_1^*\alpha_2^*\right) d(t) + n_y(t) \end{bmatrix} \tag{2.3}$$

and the corresponding photo-detector output is

$$\begin{bmatrix} |E_x(t)|^2 \\ |E_y(t)|^2 \end{bmatrix} = \begin{bmatrix} \underbrace{\left|\alpha_1\alpha_2 - \beta_1^*\beta_2\right|^2 |d(t)|^2}_{E_{x,s}(t)} + \underbrace{2\,\mathrm{Re}\left\{\left(\alpha_1\alpha_2 - \beta_1^*\beta_2\right) d(t) n_x^*(t)\right\}}_{E_{x,sn}(t)} + |n_x(t)|^2 \\ \underbrace{\left|-\alpha_1\beta_2^* - \beta_1^*\alpha_2^*\right|^2 |d(t)|^2}_{E_{y,s}(t)} + \underbrace{2\,\mathrm{Re}\left\{\left(-\alpha_1\beta_2^* - \beta_1^*\alpha_2^*\right) d(t) n_y^*(t)\right\}}_{E_{y,sn}(t)} + |n_y(t)|^2 \end{bmatrix}, \tag{2.4}$$

assuming the responsivity of the photo-detector to be 1 for simplicity purpose. As the value of B_e does not affect the mean value of the photo current and the signal-noise

beating terms $E_{x,sn}(t)$ and $E_{y,sn}(t)$ are known to have zero mean, the time average of $I_x(t)$ and $I_y(t)$ can be expressed as

$$\begin{aligned}\langle I_x(t)\rangle &= \left|\alpha_1\alpha_2 - \beta_1^*\beta_2\right|^2 \mathbf{E}\left[\left|d(t)\right|^2\right] + \mathbf{E}\left[\left|n_x(t)\right|^2\right] \\ &= r\mathbf{E}\left[\left|d(t)\right|^2\right] + \mathbf{E}\left[\left|n_x(t)\right|^2\right] \\ &= rP_s + P_n/2\end{aligned} \tag{2.5}$$

and

$$\begin{aligned}\langle I_y(t)\rangle &= \left|-\alpha_1\beta_2^* - \beta_1^*\beta_2^*\right|^2 \mathbf{E}\left[\left|d(t)\right|^2\right] + \mathbf{E}\left[\left|n_y(t)\right|^2\right] \\ &= (1-r)\mathbf{E}\left[\left|d(t)\right|^2\right] + \mathbf{E}\left[\left|n_y(t)\right|^2\right] \\ &= (1-r)P_s + P_n/2\end{aligned} \tag{2.6}$$

where $r = \left|\alpha_1\alpha_2 - \beta_1^*\beta_2\right|^2$ can be regarded as the signal power splitting ratio by the PBS, $\mathbf{E}[\cdot]$ denotes expectation, P_s and P_n are the signal and noise power after the band-pass filter, respectively. Note that B_o is implicitly included in the noise power P_n. As $d(t)$ and $n(t)$ are Ergodic processes, the time average operator $\langle\cdot\rangle$ and the ensemble average operator $\mathbf{E}[\cdot]$ will be used interchangeably throughout the rest of the chapter. Now, when r = 1 or 0, the system setup reduce to that of polarization nulling [19] and one can monitor OSNR by simply solving for P_s and P_n from (2.5) and (2.6). In general, however, polarization control may not be available at the monitoring unit such that r does not equal to 0 or 1. Despite of the fact that ASE noise will always be split equally as they are unpolarized. In this case, additional relationships between r, P_s and P_n will be required for OSNR monitoring. In particular, one can consider the second order moment $\left\langle\left[I_x(t) - \frac{r}{1-r}I_y(t)\right]^2\right\rangle = \left\langle\left[I_x(t) - \mu I_y(t)\right]^2\right\rangle$.

Let the in-band ASE noise be $\begin{bmatrix} n_x(t) \\ n_y(t) \end{bmatrix} = \begin{bmatrix} n_{xi}(t) + jn_{xq}(t) \\ n_{yi}(t) + jn_{yq}(t) \end{bmatrix}$ where $n_{xi}(t)$, $n_{xq}(t)$, $n_{yi}(t)$ and $n_{yq}(t)$ are independent identically distributed (i.i.d.) zero-mean Gaussian random process with PSD

$$S_n(f) = \begin{cases} \dfrac{P_n}{4B_o} & -\dfrac{B_o}{2} \le f \le \dfrac{B_o}{2} \\ 0 & \text{otherwise} \end{cases} \tag{2.7}$$

and corresponding autocorrelation function $R_n(\tau) = \dfrac{P_n \operatorname{sinc}(B_o\tau)}{4B_o}$. Let $h_e(t)$ be the impulse response of the electrical low-pass filter. After cancelling the signal component in $|E_x(t)|^2$ and $|E_y(t)|^2$ from, we obtain

$$\begin{aligned} &\mathbf{E}\left[\left(I_x(t) - \mu I_y(t)\right)^2\right] \\ &= \mathbf{E}\left[\left\{\left(E_{x,sn}(t) - \mu E_{y,sn}(t) + |n_x(t)|^2 - \mu|n_y(t)|^2\right) \otimes h_e(t)\right\}^2\right] \\ &= \mathbf{E}\left[\left(E_{x,sn}(t) \otimes h_e(t)\right)^2\right] + \mu^2 \mathbf{E}\left[\left(E_{y,sn}(t) \otimes h_e(t)\right)^2\right] \\ &\quad + \left(1+\mu^2\right)\mathbf{E}\left[\left(|n_x(t)|^2 \otimes h_e(t)\right)^2\right] - 2\mu\left(\mathbf{E}\left[|n_x(t)|^2\right]\right)^2 \otimes h_e(t) \\ &= \int_{-B_e}^{B_e} S_{x,sn}(f)df + \mu^2 \int_{-B_e}^{B_e} S_{y,sn}(f)df + \left(1+\mu^2\right)\int_{-B_e}^{B_e} S_{|n_x|^2}(f)df - \frac{1}{2}\mu P_n^2, \end{aligned} \tag{2.8}$$

where $S_{x,sn}(f)$, $S_{y,sn}(f)$ and $S_{|n_x|^2}(f)$ are the power spectral densities of $E_{x,sn}(t)$, $E_{y,sn}(t)$ and $|n_x(t)|^2$ respectively and ' $\otimes$ ' denotes convolution. As $\mathbf{E}\left[n_{xi}^4(t)\right] = \mathbf{E}\left[n_{xq}^4(t)\right] = \mathbf{E}\left[n_{yi}^4(t)\right] = \mathbf{E}\left[n_{yq}^4(t)\right] = 3\left(P_n/4\right)^2 = 3P_n^2/16$, the auto-correlation function of $|n_x(t)|^2$ is given by

$$
\begin{aligned}
R_{|n_x^2|}(\tau) &= \mathbf{E}\left[|n_x(t)|^2 \, |n_x(t-\tau)|^2 \right] \\
&= \mathbf{E}\left[n_{xi}^2(t) n_{xi}^2(t-\tau) \right] + \mathbf{E}\left[n_{xi}^2(t) n_{xq}^2(t-\tau) \right] \\
&\quad + \mathbf{E}\left[n_{xq}^2(t) n_{xi}^2(t-\tau) \right] + \mathbf{E}\left[n_{xq}^2(t) n_{xq}^2(t-\tau) \right] \\
&= 2\mathbf{E}\left[n_{xi}^2(t) n_{xi}^2(t-\tau) \right] + P_n^2 / 8 \\
&\overset{(a)}{=} 2\left[\frac{P_n^2}{16} + 2 \cdot \frac{P_n^2}{16} \cdot R_n^2(\tau) \right] + \frac{P_n^2}{8} \\
&= P_n^2 \left(1 + R_n^2(\tau) \right) / 4
\end{aligned}
\tag{2.9}
$$

where the equality in (a) follows from the Isserlis' theorem [34]. Therefore, $S_{|n_x|^2}(f)$ is given by

$$
S_{|n_x|^2}(f) = \frac{P_n^2}{4} \left(\delta(f) + S_n(f) \otimes S_n(f) \right) \tag{2.10}
$$

where $\delta(f)$ is the dirac-delta function and $S_{|n_x|^2}(f)$ is graphically shown in Fig. 2.2. With electrical low-pass filtering, one can easily show that

$$
\int_{-B_e}^{B_e} S_{|n_x|^2}(f) df = \frac{1}{4} \left(1 + \frac{2B_e}{B_o} - \frac{B_e^2}{B_o^2} \right) P_n^2 . \tag{2.11}
$$

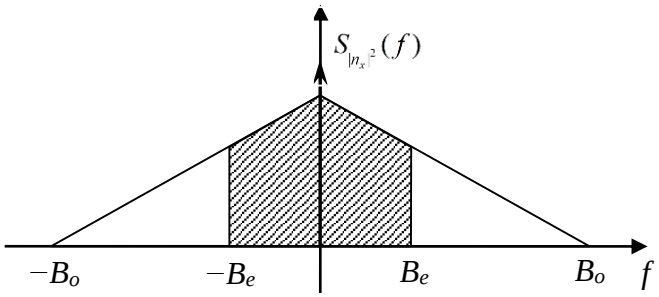


Fig. 2.2. PSD of the noise-noise beating term $|n_x(t)|^2$.

For the PSD of the signal-noise beating terms, let $R_d(\tau)$ and $S_d(f)$ be the autocorrelation and PSD of the signal $d(t)$ respectively. Now, consider the autocorrelation function

$$
\begin{aligned}
R_{x,sn}(\tau) &= \mathbf{E}\left[E_{x,sn}(\tau) E_{x,sn}^{*}(t-\tau) \right] \\
&= 4\mathbf{E}\Big[\operatorname{Re}\left\{ \left(\alpha_1 \alpha_2 - \beta_1^{*} \beta_2 \right) d(t) n_x^{*}(t) \right\} \\
&\quad \cdot \operatorname{Re}\left\{ \left(\alpha_1 \alpha_2 - \beta_1^{*} \beta_2 \right) d(t-\tau) n_x^{*}(t-\tau) \right\} \Big] \\
&= 2r\mathbf{E}\left[d(t) d^{*}(t-\tau) \right] \mathbf{E}\left[n_x(t) n_x^{*}(t-\tau) \right] \\
&= 4r R_d(\tau) R_n(\tau)
\end{aligned}
\tag{2.12}
$$

and hence $S_{x,sn}(f) = 4rS_d(f) \otimes S_n(f)$. When the electrical bandwidth B_e is small, we can approximate the integral

$$
\int_{-B_e}^{B_e} S_{x,sn}(f) df \approx \int_{-B_e}^{B_e} S_{x,sn}(0) df = \frac{2B_e}{B_o} r P_s P_n \tag{2.13}
$$

and simulation results show that the approximation holds for $0 < B_e < B_o / 2$. For $B_e > B_o / 2$, we can approximate $\int_{-B_e}^{B_e} S_{x,sn}(f) df \approx r P_s P_n$. Similarly, one can show that

$$
\int_{-B_e}^{B_e} S_{y,sn}(f) df = \frac{2B_e}{B_o} (1-r) P_s P_n \tag{2.14}
$$

when $0 < B_e < B_o / 2$. Finally, substituting (2.11), (2.13) and (2.14) into (2.8), we obtain

$$
\mathbf{E}\left[\left(I_x(t) - \mu I_y(t) \right)^2 \right] = \frac{2B_e}{B_o} \mu P_s P_n + \frac{1}{4} \left(1 + \frac{2B_e}{B_o} - \frac{B_e^2}{B_o^2} \right) \left(1 + \mu^2 \right) P_n^2 - \frac{1}{2} \mu P_n^2 . \tag{2.15}
$$

With further algebraic manipulations, (2.5), (2.6) and (2.15) results in

$$
\left\langle I_x(t) \right\rangle + \left\langle I_y(t) \right\rangle = P_s + P_n \tag{2.16}
$$

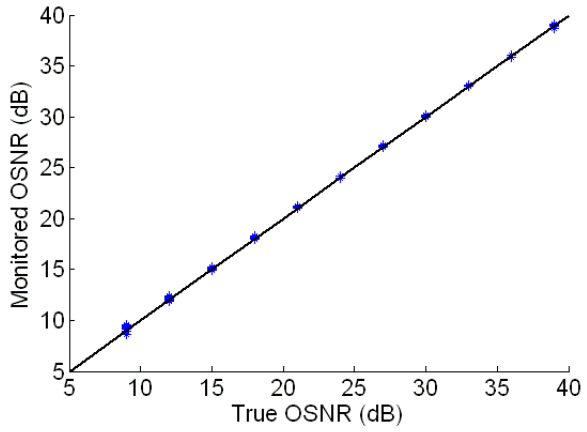
and

$$\left\langle I_x^2(t)\right\rangle - 2\frac{\left\langle I_x(t)\right\rangle - P_n/2}{\left\langle I_y(t)\right\rangle - P_n/2}\left\langle I_x(t)I_y(t)\right\rangle + \left(\frac{\left\langle I_x(t)\right\rangle - P_n/2}{\left\langle I_y(t)\right\rangle - P_n/2}\right)^2\left\langle I_y^2(t)\right\rangle$$
$$= \frac{2B_e}{B_o}\frac{\left\langle I_x(t)\right\rangle - P_n/2}{\left\langle I_y(t)\right\rangle - P_n/2}P_n\left[\left\langle I_x(t)\right\rangle + \left\langle I_y(t)\right\rangle - P_n\right] \tag{2.17}$$
$$+\frac{1}{4}\left[\left(1+\frac{2B_e}{B_o}-\frac{B_e^2}{B_o^2}\right)\left(1+\left(\frac{\left\langle I_x(t)\right\rangle - P_n/2}{\left\langle I_y(t)\right\rangle - P_n/2}\right)^2\right) - 2\frac{\left\langle I_x(t)\right\rangle - P_n/2}{\left\langle I_y(t)\right\rangle - P_n/2}\right]P_n^2.$$

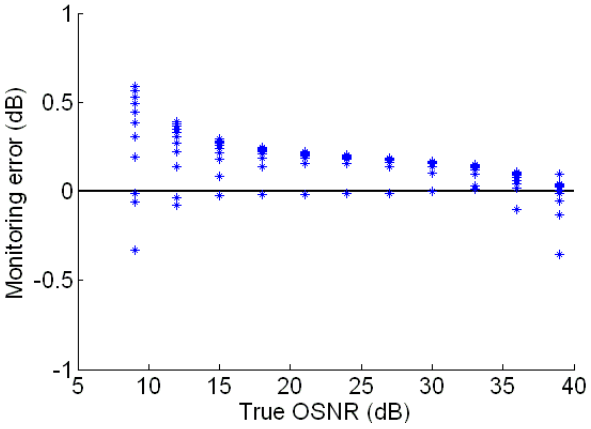
Therefore, with $\left\langle I_x(t)\right\rangle$, $\left\langle I_y(t)\right\rangle$, $\left\langle I_x^2(t)\right\rangle$, $\left\langle I_y^2(t)\right\rangle$ and $\left\langle I_x(t)I_y(t)\right\rangle$ obtained from the DSP unit, P_s and P_n can be calculated using (2.16) and (2.17) for any *r* and hence OSNR can be monitored without the need for polarization control.

We investigated the proposed technique for 40 Gb/s NRZ-OOK and RZ-QPSK systems without PMD and polarization control through simulations. The average power of the received signal is 0 dBm and the PS scrambles the received signal polarization so that power splitting ratios from 0 to 0.5 are covered. For each PS configuration, 4096 samples with a sampling rate of 2.5 GS/s of $I_x(t)$ and $I_y(t)$ are used to obtain their first- and second-order moments. The bandwidth of the optical band-pass filter is 80 GHz and the bandwidth of the electrical low-pass filter is 1.25 GHz. The sampling rate is set to be twice the value of B_e for all the simulations in this chapter. Moreover, note that if the sampling rate is equal or higher than Nyquist sampling rate, either synchronous or asynchronous sampling can be performed. Otherwise, asynchronous sampling is necessary. In practice, sampling rate for asynchronous systems can be far lower than the symbol rate or the low-pass filter bandwidths which will guarantee statistical independence of the noise samples. Fig. 2.3 shows the monitored OSNR

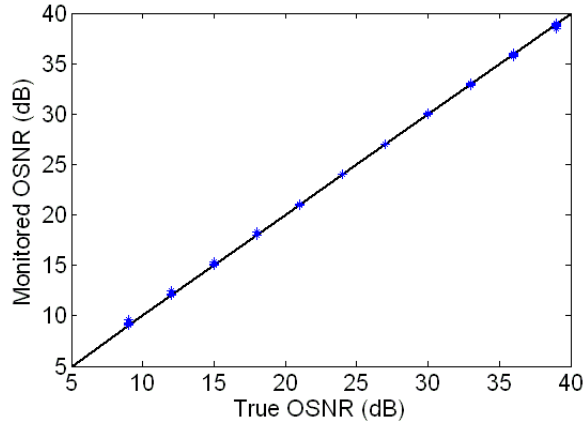
(measured in 0.1 nm bandwidth) and the corresponding monitoring errors. For various values of *r*, a monitoring range of 9 dB to 39 dB can be obtained with a mean error of 0.5 dB for both modulation formats. According to the simulation results, we find that the monitoring error is large when $r = 0$ or 0.05.



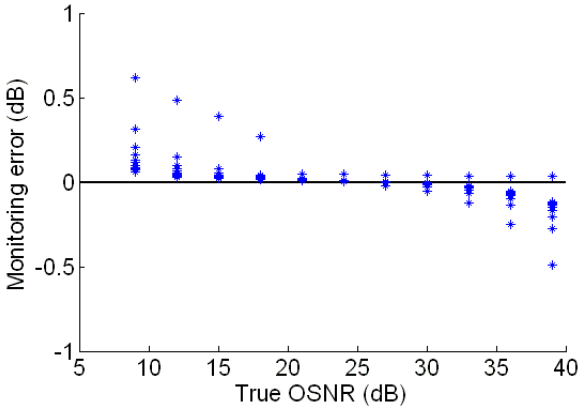
(a)



(b)



(c)



(d)

Fig. 2.3. Monitored OSNR vs True OSNR for (a) 40 Gbit/s NRZ-OOK and (c) RZ-QPSK systems without PMD and polarization control. The data points for each OSNR value correspond to power splitting ratios ranging from 0 to 0.5 with 0.05 step width. Monitoring errors for NRZ-OOK and RZ-QPSK systems are shown in (b) and (d) respectively. With lower power splitting ratio, the errors tend to be larger due to the stochastic nature of noise.

2.4. OSNR Monitoring in the Presence of First-Order PMD

In the presence of PMD, and when the signal polarization is not aligned with the PSP of the fiber, the PBS output is given by (2.2) and the photo detector output is given by

$$\begin{bmatrix} |E_x(t)|^2 \\ |E_y(t)|^2 \end{bmatrix} = \begin{bmatrix} \underbrace{\left|\alpha_1\alpha_2 d\left(t+\frac{\Delta\tau}{2}\right)\right|^2 + \left|\beta_1\beta_2 d\left(t-\frac{\Delta\tau}{2}\right)\right|^2 - 2\,\mathrm{Re}\left\{\alpha_1\alpha_2\beta_1\beta_2^* d\left(t+\frac{\Delta\tau}{2}\right) d^*\left(t-\frac{\Delta\tau}{2}\right)\right\}}_{E_{x,s}(t)} \\ \underbrace{+2\,\mathrm{Re}\left\{\left(\alpha_1\alpha_2 d\left(t+\frac{\Delta\tau}{2}\right)+\beta_1^*\beta_2 d\left(t-\frac{\Delta\tau}{2}\right)\right) n_x^*(t)\right\} + |n_x(t)|^2}_{E_{x,sn}(t)} \\ \underbrace{\left|\alpha_1\beta_2 d\left(t+\frac{\Delta\tau}{2}\right)\right|^2 + \left|\beta_1\alpha_2 d\left(t-\frac{\Delta\tau}{2}\right)\right|^2 + 2\,\mathrm{Re}\left\{\alpha_1\alpha_2\beta_1\beta_2^* d\left(t+\frac{\Delta\tau}{2}\right) d^*\left(t-\frac{\Delta\tau}{2}\right)\right\}}_{E_{y,s}(t)} \\ \underbrace{+2\,\mathrm{Re}\left\{\left(-\alpha_1\beta_2^* d\left(t+\frac{\Delta\tau}{2}\right)-\beta_1^*\alpha_2^* d\left(t-\frac{\Delta\tau}{2}\right)\right) n_y^*(t)\right\} + |n_y(t)|^2}_{E_{y,sn}(t)} \end{bmatrix}. \quad (2.18)$$

And we can see that as a result of PMD, $E_{x,s}(t)$ and $E_{y,s}(t)$ comprises of two components originating from the two PSP and one component from their beating. At this point, we note that for modulation formats such as OOK, BPSK and DPSK, the signal $d(t)$ is real and

$$\begin{aligned} &2\,\mathrm{Re}\left\{\alpha_1\alpha_2\beta_1\beta_2^* d\left(t+\Delta\tau/2\right)d^*\left(t-\Delta\tau/2\right)\right\} \\ &=2\,\mathrm{Re}\left\{\alpha_1\alpha_2\beta_1\beta_2^*\right\} d\left(t+\Delta\tau/2\right)d^*\left(t-\Delta\tau/2\right). \end{aligned} \tag{2.19}$$

Furthermore, for polarization states such that the phase of $\alpha_1\alpha_2\beta_1\beta_2^*$

$$\arg\left(\alpha_1\alpha_2\beta_1\beta_2^*\right)=\frac{\pm\pi}{2}, \tag{2.20}$$

(2.19) will be identically zero. In Stokes space representation, if $\Delta\vec{\tau}$ and $\hat{s}_d\left(\omega_0\right)$ denotes the PMD vector and the Stokes vector of the signal at center frequency ω_0 respectively, (2.20) is equivalent to the condition when the PS rotates $\Delta\vec{\tau}$ and $\hat{s}_d\left(\omega_0\right)$ on the Poincaré sphere so that the Stokes vector (±1,0,0) falls on the plane defined by $\Delta\vec{\tau}$ and $\Delta\vec{\tau}\times\hat{s}_d\left(\omega_0\right)$. Alternatively, for a given $\Delta\vec{\tau}$ and $\hat{s}_d\left(\omega_0\right)$, we can think of (2.20) as the condition when the PS rotates the Stokes vector of the PBS output polarizations $\pm\hat{p}$ onto the plane defined by $\Delta\vec{\tau}$ and $\Delta\vec{\tau}\times\hat{s}_d\left(\omega_0\right)$. In this case, $I_x(t)$ and $I_y(t)$ are then given by

$$\begin{bmatrix} I_x(t) \\ I_y(t) \end{bmatrix}=\begin{bmatrix} \dfrac{1}{2\pi}\displaystyle\int_{-2\pi B_e}^{2\pi B_e}\left(\left|\alpha_1\alpha_2\right|^2 e^{j\omega\Delta\tau/2}+\left|\beta_1\beta_2\right|^2 e^{-j\omega\Delta\tau/2}\right)\tilde{D}_2(\omega)e^{j\omega t}d\omega+\left|n_x(t)\right|^2\otimes h_e(t) \\ +\dfrac{1}{2\pi}\displaystyle\int_{-2\pi B_e}^{2\pi B_e}2\,\mathrm{Re}\left\{\left(\alpha_1\alpha_2\tilde{D}(\omega)e^{j\omega\Delta\tau/2}-\beta_1^*\beta_2\tilde{D}(\omega)e^{-j\omega\Delta\tau/2}\right)\otimes\tilde{N}_x^*(-\omega)\right\}e^{j\omega t}d\omega \\ \dfrac{1}{2\pi}\displaystyle\int_{-2\pi B_e}^{2\pi B_e}\left(\left|\alpha_1\beta_2\right|^2 e^{j\omega\Delta\tau/2}+\left|\beta_1\alpha_2\right|^2 e^{-j\omega\Delta\tau/2}\right)\tilde{D}_2(\omega)e^{j\omega t}d\omega+\left|n_y(t)\right|^2\otimes h_e(t) \\ +\dfrac{1}{2\pi}\displaystyle\int_{-2\pi B_e}^{2\pi B_e}2\,\mathrm{Re}\left\{\left(-\alpha_1\beta_2^*\tilde{D}(\omega)e^{j\omega\Delta\tau/2}-\beta_1^*\alpha_2^*\tilde{D}(\omega)e^{-j\omega\Delta\tau/2}\right)\otimes\tilde{N}_y^*(-\omega)\right\}e^{j\omega t}d\omega \end{bmatrix} \tag{2.21}$$

where $\tilde{D}_2\left(\omega\right)$ is the Fourier transform of $\left|d(t)\right|^2$. Now, note that when $\Delta\tau B_e\ll 1$,

$e^{\pm j\omega\Delta\tau/2} \approx 1$ for $-2\pi B_e < \omega < 2\pi B_e$ and $I_x(t)$ and $I_y(t)$ are reduced to

$$\begin{bmatrix} I_x(t) \\ I_y(t) \end{bmatrix} = \begin{bmatrix} \left(|\alpha_1\alpha_2|^2 + |\beta_1\beta_2|^2\right)|d(t)|^2 \otimes h_e(t) + |n_x(t)|^2 \otimes h_e(t) \\ +2\,\mathrm{Re}\left\{\left(\alpha_1\alpha_2 - \beta_1^*\beta_2\right)d(t) \otimes n_x^*(t)\right\} \otimes h_e(t) \\ \left(|\alpha_1\beta_2|^2 + |\beta_1\alpha_2|^2\right)|d(t)|^2 \otimes h_e(t) + |n_y(t)|^2 \otimes h_e(t) \\ +2\,\mathrm{Re}\left\{\left(-\alpha_1\beta_2^* - \beta_1^*\alpha_2^*\right)d(t) \otimes n_y^*(t)\right\} \otimes h_e(t) \end{bmatrix} \tag{2.22}$$

and their time average

$$\begin{bmatrix} \langle I_x(t)\rangle \\ \langle I_y(t)\rangle \end{bmatrix} = \begin{bmatrix} \left(|\alpha_1\alpha_2|^2 + |\beta_1\beta_2|^2\right)P_s + P_n/2 \\ \left(|\alpha_1\beta_2|^2 + |\beta_1\alpha_2|^2\right)P_s + P_n/2 \end{bmatrix} = \begin{bmatrix} rP_s + P_n/2 \\ (1-r)P_s + P_n/2 \end{bmatrix}. \tag{2.23}$$

In addition, one can show that with (2.22), (2.15) will also hold and therefore (2.16) and (2.17) can be used to monitor OSNR. Note that (2.22) and (2.23) is true only when (2.20) and the condition $\Delta\tau B_e \ll 1$ hold. Otherwise, (2.16) and (2.17) is not applicable for the system concerned and in that case, it can be shown that solving (2.16) and (2.17) will always produce under-estimates of the true OSNR with the proof shown below.

We note that (2.16) is always true regardless of whether PMD exists in the system or not. When PMD is absent, one can substitute $\mu = \dfrac{\langle I_x(t)\rangle - P_n/2}{\langle I_y(t)\rangle - P_n/2}$ into (2.15) to solve for P_n. More generally, one can define a function

$$f(z)=\left\langle\left(I_x(t)-\frac{\langle I_x(t)\rangle-z/2}{\langle I_y(t)\rangle-z/2}I_y\right)^2\right\rangle-\frac{1}{4}\left(1-\frac{\langle I_x(t)\rangle-z/2}{\langle I_y(t)\rangle-z/2}\right)^2 z$$
$$-\frac{2B_e}{B_o}\cdot\frac{\langle I_x(t)\rangle-z/2}{\langle I_y(t)\rangle-z/2}\cdot z\left(\langle I_x(t)\rangle+\langle I_y(t)\rangle-z\right) \tag{2.24}$$
$$-\frac{2B_e}{B_o}\cdot\frac{1}{4}\left[1+\left(\frac{\langle I_x(t)\rangle-z/2}{\langle I_y(t)\rangle-z/2}\right)^2\right]z^2$$

assuming $B_e/B_o \ll 1$ such that the solution to $f(z)=0$ is $\tilde{z}=P_n$. In this case, the OSNR can be estimated by

$$\text{OSNR}=\frac{\langle I_x(t)\rangle+\langle I_y(t)\rangle-\tilde{z}}{\tilde{z}}. \tag{2.25}$$

We will prove that in presence of PMD, the corresponding solution to $f(z)=0$ will be $\tilde{z}\geq P_n$ with equality only for certain polarization states. Therefore, the noise power will always be over-estimated and by (2.25), OSNR will always be under-estimated except for certain polarization states.

In general, the photo-detector output is given by

$$\begin{bmatrix}|E_x(t)|^2\\ |E_y(t)|^2\end{bmatrix}=\begin{bmatrix}E_{x,s}(t)+E_{x,sn}(t)+|n_x(t)|^2\\ E_{y,s}(t)+E_{y,sn}(t)+|n_y(t)|^2\end{bmatrix} \tag{2.26}$$

where $E_{x,s}(t)$, $E_{y,s}(t)$, $E_{x,sn}(t)$ and $E_{y,sn}(t)$ are defined in (2.4) when PMD is absent and (2.18) when PMD is present. Let

$$I_x(t)=I_{x,s}(t)+I_{x,sn}(t)+I_{x,n}(t) \tag{2.27}$$

and

$$I_y(t)=I_{y,s}(t)+I_{y,sn}(t)+I_{y,n}(t) \tag{2.28}$$

where $I_{x(y),s}(t)$, $I_{x(y),sn}(t)$ and $I_{x(y),n}(t)$ are the low-pass filtered output of

$E_{x(y),s}(t)$, $E_{x(y),sn}(t)$ and $|n_{x(y)}(t)|^2$, respectively. When PMD is present, one can still denote

$$\langle I_x(t)\rangle = \mathbf{E}\left[E_{x,s}(t)\right] + \mathbf{E}\left[\left|n_x(t)\right|^2\right] = rP_s + P_n/2 \tag{2.29}$$

and

$$\langle I_y(t)\rangle = \mathbf{E}\left[E_{y,s}(t)\right] + \mathbf{E}\left[\left|n_y(t)\right|^2\right] = (1-r)P_s + P_n/2 \tag{2.30}$$

since for given P_s and P_n, one can always find a numerical value of *r* such that (2.29) and (2.30) hold even when PMD is present. Therefore, $\langle I_{x,s}(t)\rangle = rP_s$, $\langle I_{y,s}(t)\rangle = (1-r)P_s$, $\langle I_{x,sn}(t)\rangle = \langle I_{y,sn}(t)\rangle = 0$, $\langle I^2_{x,sn}(t)\rangle = 2B_e rP_sP_n/B_o$ and $\langle I^2_{y,sn}(t)\rangle = 2B_e(1-r)P_sP/B_o$ regardless of the present of PMD. In this case, as shown in the appendix,

$$\begin{aligned}
&\left\langle\left(I_x(t) - \frac{\langle I_x(t)\rangle - z/2}{\langle I_y(t)\rangle - z/2} I_y(t)\right)^2\right\rangle \\
&= \mathbf{Var}\left[I_{x,s}(t) - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} I_{y,s}(t)\right] \\
&\quad + \frac{2B_e}{B_o}\left[r + \left(\frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2}\right)^2 (1-r)\right] P_sP_n \\
&\quad + \frac{2B_e}{B_o}\cdot\frac{1}{4}\left[1 + \left(\frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2}\right)^2\right] P_n^2 + \left[\frac{(1-2r)P_s z/2}{(1-r)P_s + (P_n - z)/2}\right]^2
\end{aligned} \tag{2.31}$$

and

$$\begin{aligned}
f(z) = &\underbrace{\mathbf{Var}\left[I_{x,s}(t) - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} I_{y,s}(t)\right]}_{f_1(z)} \\
&+ \underbrace{\frac{2B_e}{B_o}\left\{\left[rP_s + (P_n - z)/2\right]^2\left[2 - \frac{(1-r)^2P_s^2}{\left[(1-r)P_s + (P_n - z)/2\right]^2}\right] - r^2P_s^2\right\}}_{f_2(z)}
\end{aligned} \tag{2.32}$$

We note that without PMD, $I_{x,s}(t) = \frac{r}{1-r} I_{y,s}(t) = \mu I_{y,s}(t)$ and $f_1(z) = f_2(z) = 0$

when $z = P_n$ and hence the solution to $f(z) = 0$ is an unbiased estimate of the noise power. In addition, one can show that $f_2(z) > 0$ when $0 \le z < P_n$ and $f_2(z) < 0$ when $P_n \le z < P_n + 2rP_s$. However, $I_{x,s}(t)$ and $I_{y,s}(t)$ can no longer be related by a simple scaling factor when PMD is present. Consequently, $f_1(z) > 0$ for all *z*. In addition, as $f_2(z)$ does not change with PMD for a given *r*, $f(z) = 0$ will only be true for some $\tilde{z} \ge P_n$ as illustrated in Fig. 2.4. Therefore, the noise power will always be over-estimated and hence the monitored OSNR will always be smaller than the true value. And hence one can solve for (2.16) and (2.17) to obtain an OSNR estimate for each polarization state scrambled by the PS and the maximal value among the estimates can be used to monitor OSNR.

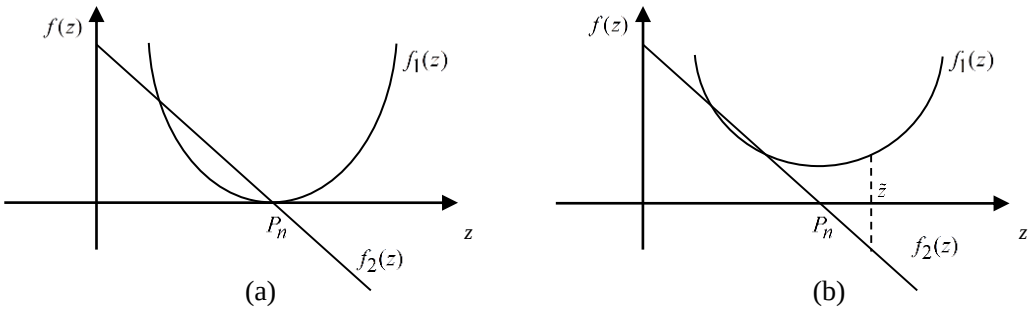
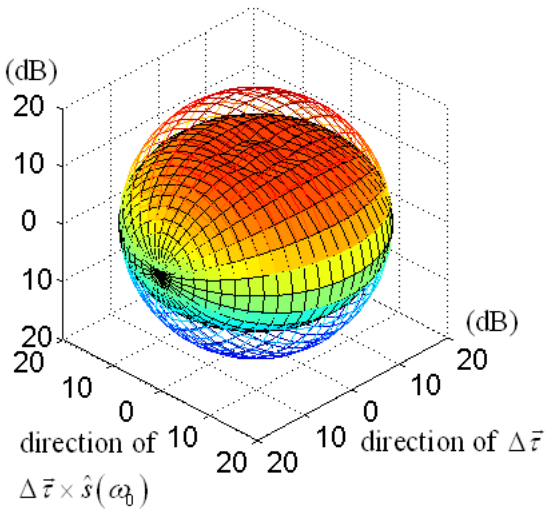


Fig. 2.4. A graphical illustration of the solution to $f(z) = f_1(z) + f_2(z) = 0$ (a) without PMD. (b) with first-order PMD. In presence of first-order PMD, the solution $\tilde{z}$ is always larger than the true noise power P_n.

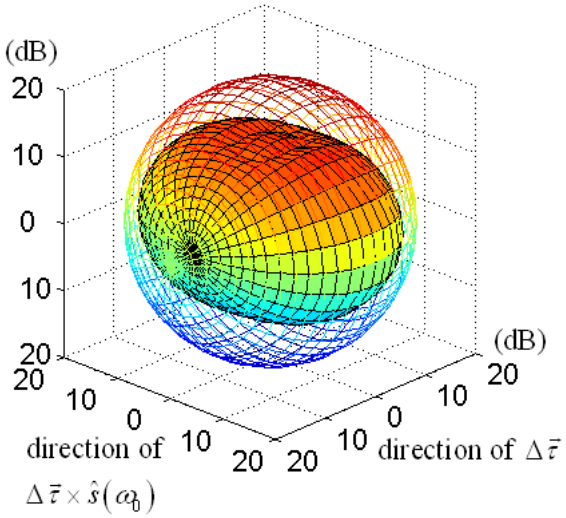
Then we carry out simulations on 40 Gbit/s NRZ-OOK systems for three sets of DGD and electrical low-pass filter bandwidths: ($\Delta\tau$, B_e) = (12.5 ps, 312.5 MHz), ($\Delta\tau$, B_e) = (12.5 ps, 10 GHz) and ($\Delta\tau$, B_e) = (400 ps, 312.5 MHz). The true OSNR is 20 dB and the optical filter bandwidth is 80 GHz. We studied the case when the signal power is split equally onto the PSP (hence $\Delta\vec{\tau}$ is perpendicular to $\hat{s}_d(\omega_0)$) in order to

maximize the effect of PMD. Fig. 2.5 shows the monitoring results as a function of $\hat{p}$ in Stokes space for various $\Delta\tau$ and B_e. In this representation, the direction from the origin represents the direction of $\hat{p}$ and the distance from origin is the monitored OSNR value in dB. Accurate OSNR monitoring for all PS configurations correspond to a constant magnitude for all directions and is graphically shown as a hollow sphere for reference. From the figures, we can see that all the monitored OSNR values are equal to or smaller than the true OSNR. Accurate OSNR estimates can be obtained when $\hat{p}$ lies on the plane defined by $\Delta\vec{\tau}$ and $\Delta\vec{\tau}\times\hat{s}_d(\omega_0)$ while it is under-estimated otherwise. Fig. 2.5(d) shows the monitoring results for $\hat{p}$ lying in the $\Delta\vec{\tau}$ and $\Delta\vec{\tau}\times\hat{s}_d(\omega_0)$ plane. Note that when ($\Delta\tau$, B_e) = (12.5 ps, 10 GHz) and ($\Delta\tau$, B_e) = (400 ps, 312.5 MHz) with $\Delta\tau B_e$ = 0.125 for both cases, the monitored OSNR is similar and thus verify that the performance of the proposed technique depends on the product $\Delta\tau B_e$. Unfortunately, when $\Delta\tau B_e$ become large, OSNR will be under-estimated even when $\hat{p}$ lies on that plane. In addition, the monitored OSNR is affected most when $\hat{p}$ aligns with $\pm\Delta\vec{\tau}$. This is further illustrated in Fig. 2.6(a) and (b) for various OSNR values. The dependence of the monitoring performance on $\Delta\tau$ and B_e was also reported in [27] and such system setup correspond to a specific $\hat{p}$ in the $\Delta\vec{\tau}$ and $\Delta\vec{\tau}\times\hat{s}_d(\omega_0)$ plane. Finally, we note that (2.19) will also be identically zero when either α_1, α_2, β_1 or β_2 is zero. When α_1 or β_1 is zero, the signal polarization aligns with the PSP of the fiber and hence the transmission is not affected by PMD. α_2 or β_2 is zero correspond to the case when the PBS output polarization aligns with the PSP i.e. $\hat{p}$

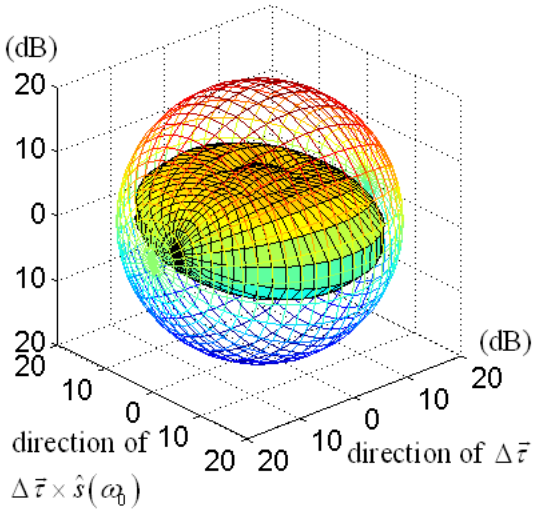
aligns with $\pm\Delta\vec{\tau}$.



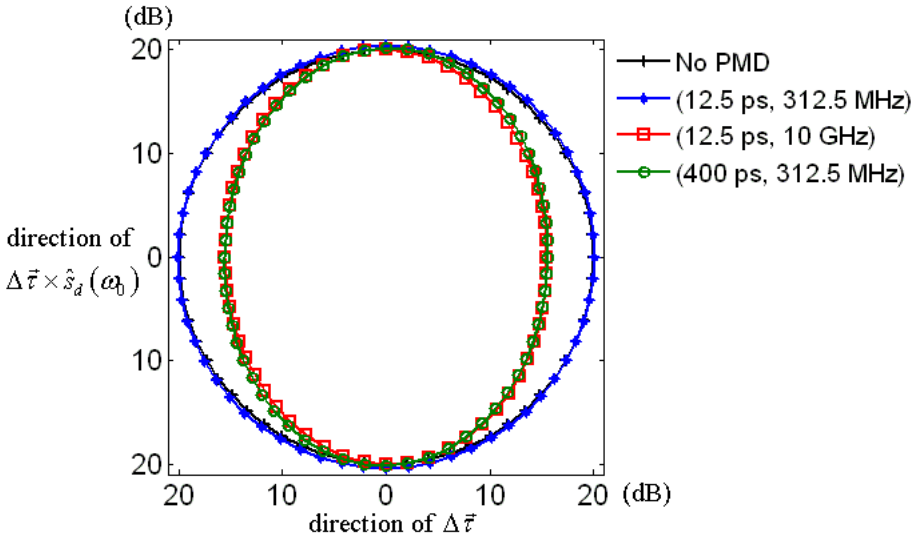
(a)



(b)

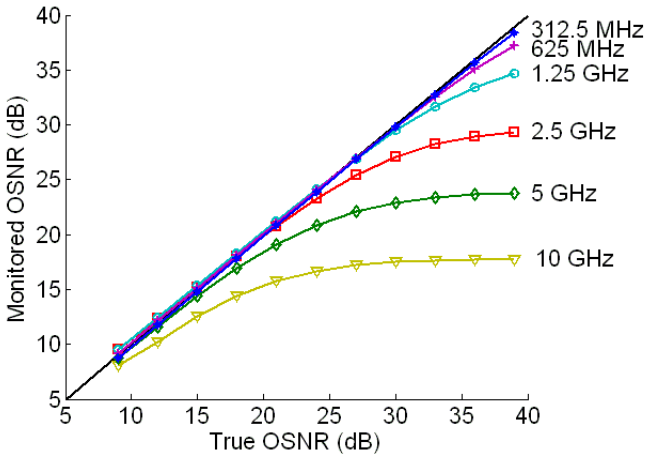


(c)

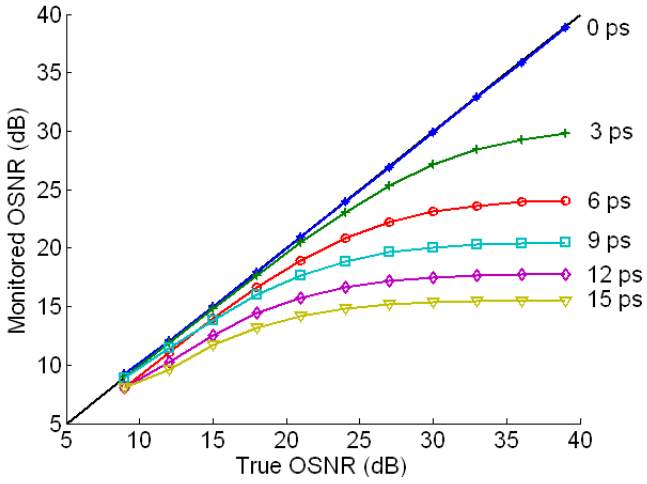


(d)

Fig. 2.5. Monitored OSNR vs. true OSNR for a 40 Gb/s NRZ-OOK system in presence of first-order PMD for (a) ($\Delta\tau$, B_e) = (12.5 ps, 312.5 MHz), (b) ($\Delta\tau$, B_e) = (12.5 ps, 10 GHz) and (c) ($\Delta\tau$, B_e) = (400 ps, 312.5 MHz). The direction and magnitude of any point on the surface from the origin correspond to the direction of $\hat{p}$ and monitored OSNR respectively. The hollow sphere represents accurate OSNR estimates for all $\hat{p}$. (d) Monitored OSNR vs. true OSNR for $\hat{p}$ lying on the plane defined by $\Delta\vec{\tau}$ and $\Delta\vec{\tau}\times\hat{s}_d(\omega_0)$ for various $\Delta\tau$ and B_e.



(a)



(b)

Fig. 2.6. Monitored OSNR vs true OSNR for a 40 Gb/s NRZ-OOK system when $\hat{p}$ aligns with $\pm\Delta\vec{\tau}$ for (a) $\Delta\tau$ = 12.5 ps and various B_e and (b) B_e = 10 GHz and various $\Delta\tau$.

2.5. True First-Order PMD Independent OSNR Monitoring and Simplified Implementation Strategies

As shown in Fig. 2.5(d), the OSNR will be under-estimated even if (2.20) holds when $\Delta\tau$ and/or B_e is not small. Therefore, the proposed OSNR monitoring technique is not truly PMD independent in general. However, note that there are two particular points in Fig. 2.5(d) whose corresponding OSNR estimates do not decrease with the increase of $\Delta\tau B_e$. This can be explained as follows. When (2.20) holds and also when

$$\left|\alpha_2\right|^2 = \left|\beta_2\right|^2 = \frac{1}{2} \tag{2.33}$$

or equivalently, when the PS rotates $\hat{p}$ to align with $\pm\Delta\vec{\tau}\times\hat{s}_d(\omega_0)$ in the Poincaré Sphere, (2.21) becomes

$$\begin{bmatrix} I_x(t) \\ I_y(t) \end{bmatrix} = \begin{bmatrix} \left(\alpha_1\alpha_2 d\left(t+\frac{\Delta\tau}{2}\right) - \beta_1^*\beta_2 d\left(t-\frac{\Delta\tau}{2}\right)\right) \otimes h_e(t) + \left|n_x(t)\right|^2 \otimes h_e(t) \\ + \frac{1}{4\pi}\int_{-2\pi B_e}^{2\pi B_e} \left(\left|\alpha_1\right|^2 e^{j\omega\Delta\tau/2} + \left|\beta_1\right|^2 e^{-j\omega\Delta\tau/2}\right) D_2(\omega) e^{j\omega t} d\omega \\ \left(-\alpha_1\beta_2^* d\left(t+\frac{\Delta\tau}{2}\right) - \beta_1^*\alpha_2^* d\left(t-\frac{\Delta\tau}{2}\right)\right) \otimes h_e(t) + \left|n_y(t)\right|^2 \otimes h_e(t) \\ + \frac{1}{4\pi}\int_{-2\pi B_e}^{2\pi B_e} \left(\left|\alpha_1\right|^2 e^{j\omega\Delta\tau/2} + \left|\beta_1\right|^2 e^{-j\omega\Delta\tau/2}\right) D_2(\omega) e^{j\omega t} d\omega \end{bmatrix}. \quad (2.34)$$

As the signal component of $I_x(t)$ and $I_y(t)$ is identical regardless of $\Delta\tau$ and B_e, one can show from (2.34) that

$$\begin{bmatrix} \langle I_x(t) \rangle \\ \langle I_y(t) \rangle \end{bmatrix} = \frac{1}{2}\begin{bmatrix} P_s + P_n \\ P_s + P_n \end{bmatrix} \quad (2.35)$$

and

$$\left\langle \left[I_x(t) - I_y(t) \right]^2 \right\rangle = \frac{2B_e}{B_o} P_n \left(\langle I_x(t) \rangle + \langle I_y(t) \rangle - P_n \right) + \left(\frac{B_e}{B_o} - \frac{B_e^2}{2B_o^2} \right) P_n^2 \quad (2.36)$$

which correspond to (2.5), (2.6) and (2.15) with r = 1/2. Therefore, when the PS scrambles the signal polarization state such that (2.20) and (2.23) hold, (2.16) and (2.17) can be solved to obtain accurate OSNR estimates which are truly independent of PMD (and also the bandwidth of the electrical filter).

Fig. 2.7 shows the monitoring results using 4096 samples for all 24 combinations of B_e = 312.5 MHz, 625 MHz, 1.25 GHz, 2.5 GHz, 5 GHz, 10 GHz and $\Delta\tau$ = 12.5 ps, 25 ps, 50 ps, 100 ps when $\hat{p}$ is aligned with $\Delta\vec{\tau} \times \hat{s}_d(\omega_0)$. For various $\Delta\tau$ and B_e value, the mean monitoring error is within 0.5 dB. Compared with Fig. 2.6, it is obvious that the monitored OSNR is insensitive to $\Delta\tau$ and B_e and hence the proposed technique is truly independent of first-order PMD when $\hat{p}$ is aligned with $\Delta\vec{\tau} \times \hat{s}_d(\omega_0)$.

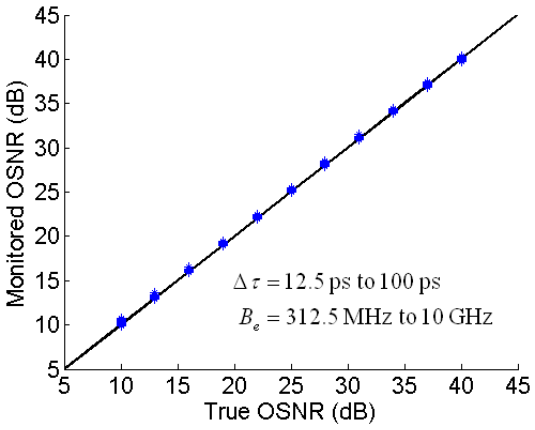


Fig. 2.7. Monitored OSNR vs true OSNR for a 40 Gb/s NRZ-OOK system when $\hat{p}$ aligns with $\Delta\vec{\tau}\times\hat{s}_d(\omega_0)$ for various $\Delta\tau B_e$.

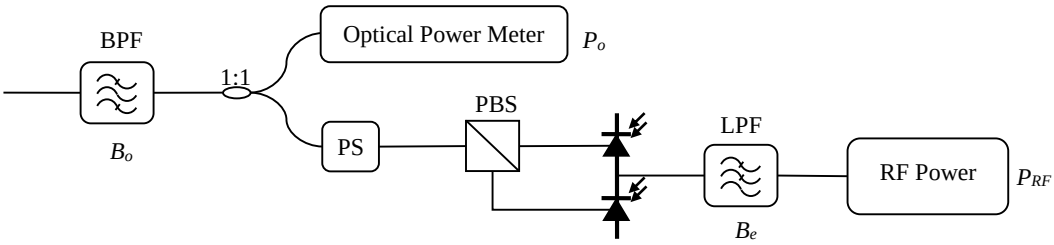


Fig. 2.8. Simplified OSNR monitoring technique using polarization diversity with a polarization scrambler, a polarization bean splitter, optical power meter, and a balanced detecter. BPF: band-pass filter; PS: polarization scrambler; PBS: polarization beam splitter; LPF: low-pass filter.

Finally, the fact that r = 1/2 in (18) and (19) also enable simplified experiment setups to realize OSNR monitoring truly independent of first-order PMD. Instead of a full DSP unit that samples $I_x(t)$ and $I_y(t)$, a balanced detector followed by an RF power meter and an optical power meter prior to the PS can be used as shown in Fig. 2.8. In this case, when (2.20) and (2.23) holds, the optical power P_o and the RF power P_{RF} is given by

$$P_o = P_s + P_n \tag{2.37}$$

$$P_{RF} = \frac{2B_e}{B_o} P_s P_n + \left(\frac{B_e}{B_o} - \frac{B_e^2}{2B_o^2} \right) P_n^2 \tag{2.38}$$

and (2.37) and (2.38) can be solved to obtain OSNR estimate given by

$$OSNR = \left(\frac{B_e}{B_o} \frac{P_o^2}{P_{RF}} - 1 \right) + \sqrt{\left(\frac{P_o^2}{P_{RF}} \right)^2 \left(\frac{B_e}{B_o} \right)^2 - \left[\frac{B_e}{B_o} + \frac{1}{2} \left(\frac{B_e}{B_o} \right)^2 \right] \frac{P_o^2}{P_{RF}}} \,. \tag{2.39}$$

It should be noted that for the proposed technique, like the polarization nulling method, there are only 2 operable polarization states, which may potentially affect the monitored OSNR results since the PS may not reach the exact 2 points in the Poincaré sphere. However, this effect can be mitigated by extending the PS searching time to increase the probability in locating the two polarization states, with a fast PS operating in tens or hundreds of kHz. It should be noted that the time required for more accurate OSNR estimation should grow logarithmically with searching time i.e. a reasonable amount of searching time should give almost optimal OSNR monitoring results already.

2.6. Experimental Results

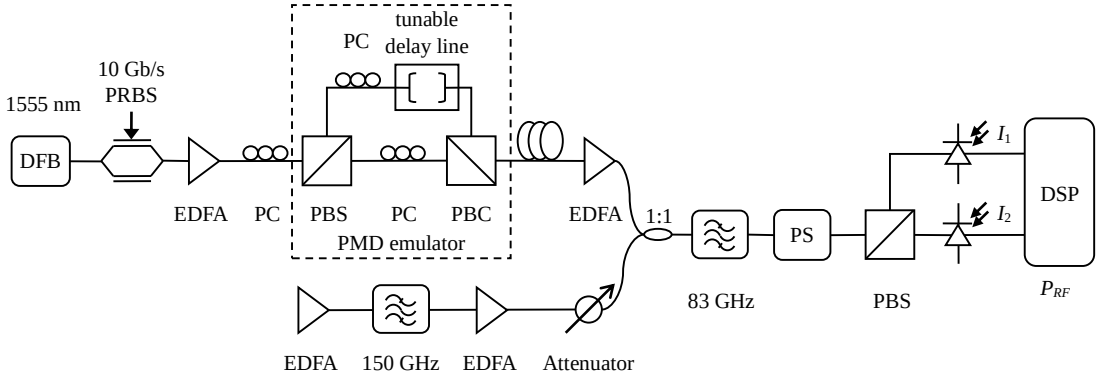


Fig. 2.9. Experimental setup for OSNR monitoring using polarization diversity and polarization scrambling. PC: Polarization controller. PS: polarization scrambler.

The experimental setup is shown in Fig. 2.9. 10 Gb/s NRZ-OOK or NRZ-DPSK signals are generated by externally modulating a DFB laser. The signal is amplified and passes through a first-order PMD emulator with 0, 50, 100 ps DGD, and the signal polarization state is 45 ° aligned with the two emulated PSPs to enhance the effect of PMD. By introducing a SSMF with a length of 0, 50 and 100 km, 0, 850 and 1700 ps/nm CD can be realized. Then the signal is amplified again to compensate for the loss before it is combined with the ASE noise and filtered with 83 GHz bandwidth optical filter. The signal power entering the PS is kept to be −18 dBm. OSNR (0.1 nm) ranging from 8.3 to 23.3 dB in step of 2.5 dB is realized by tuning the attenuator after the ASE source. The PS used is Miniature Polarization Scrambler Card from General Photonics and operates at a scrambling speed of 2 kHz. After the PBS, the signal is detected by two high-gain and low-noise DC coupled photo receivers NewFocus 1811 with 125 MHz bandwidth instead of a balanced detector. The detected signal is then sampled using a PicoScope 5203 at a rate of 62.5MS/s. The sampled data is divided into groups corresponding to different polarization states. Each group contains 80 μs worth of samples which is equivalent to 5000 samples for each polarization state. The RF power P_{RF} which is equal to the variance of $I_1 - I_2$ is then calculated for 2500 groups of data, from which the minimum P_{RF} is chosen for OSNR estimation.

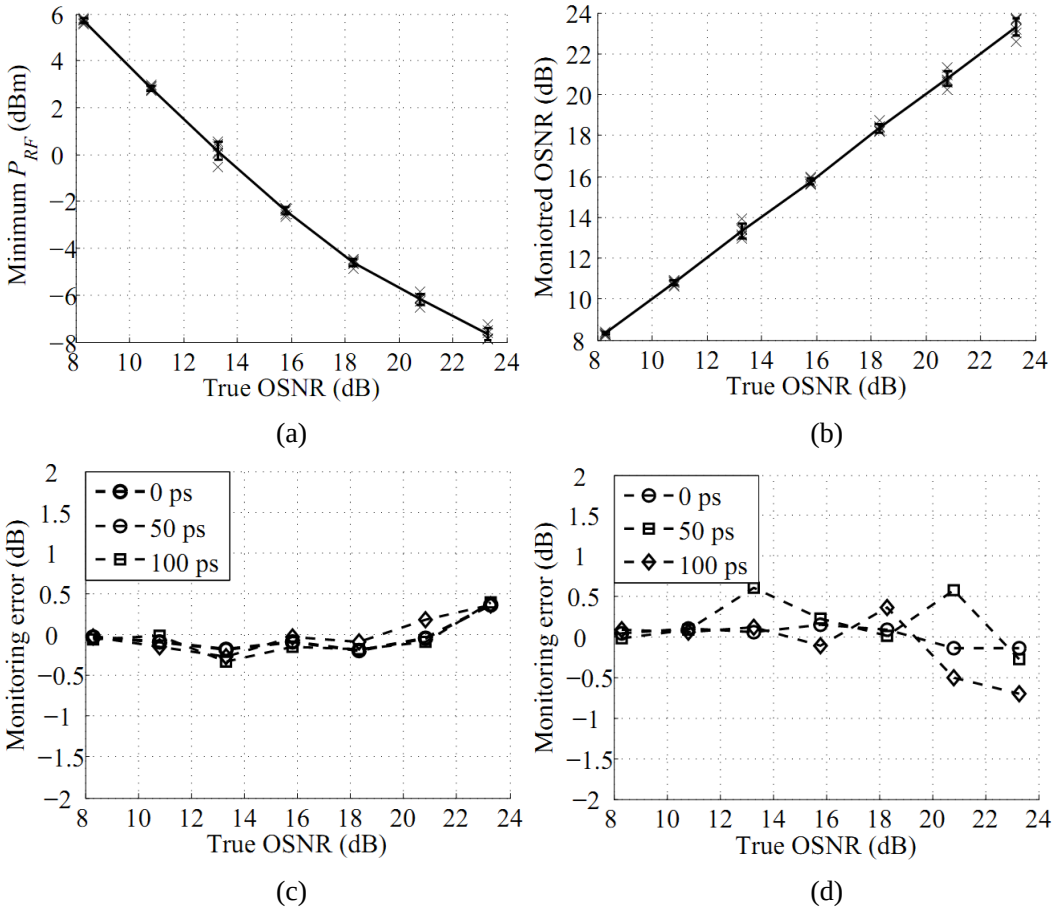
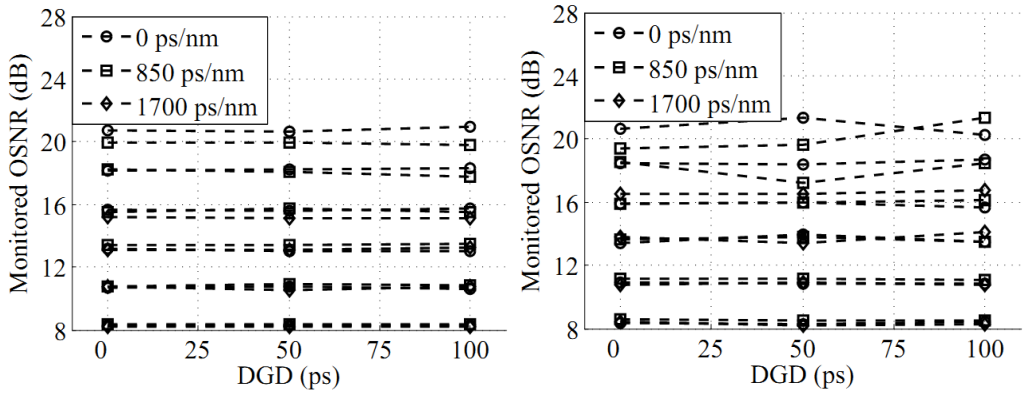


Fig. 2.10. (a) Minimum P_{RF} and (b) corresponding monitored OSNR vs. true OSNR (0.1 nm) for 10 Gb/s NRZ-OOK and NRZ-DPSK systems with OSNR ranging from 8.3 to 23.3 dB, and DGD ranging from 0 to 100 ps. The monitoring errors for NRZ-OOK and NRZ-DPSK systems are shown in (c) and (d) respectively.

The overall results of P_{RF} vs. OSNR are as shown in Fig. 2.10(a). It can be seen that when OSNR increases above 18 dB, thermal noise and shot noise begins to affect P_{RF}. The calibrating results are as shown in Fig. 2.10(b). The maxmim monitored OSNR error is within 0.7 dB for OSNR ranging from 8.3 to 23.3 dB. Fig. 2.11 shows the monitored OSNR vs. DGD for both NRZ-OOK and NRZ-DPSK systems. It can be seen that the results are insensitive to PMD of up to 100 ps, in agreement with the theoretical predictions. Yet it should be noted that at OSNR range higher than 23 dB,

the noise beating power would be too small so that the shot noise and thermal noise would strongly affect P_{RF} and cause monitoring error.



(a) (b)

Fig. 2.11. Monitored OSNR vs. DGD for (a) 10 Gb/s NRZ-OOK and (b) 10 Gb/s NRZ-DPSK systems. The monitoring results exhibit no PMD and weak CD dependence.

In the previous derivations, one of the conditions for PMD-independent OSNR monitoring is that the signal needs to be purely real and hence it may not perform well in the presence of CD. Nevertheless, we experimentally investigate the CD tolerance of the proposed technique by introducing 50 and 100 km of fiber in the experimental setup and the results are shown in Fig. 2.11. It can be seen that in basis of 1 dB monitoring error, the tolerance of CD is around 1700 ps/nm, illustrating that the proposed technique is applicable to practical deployed networks with typical values of residual dispersion. This may due to the fact that although the signal will be complex in the presence of CD, the majority of the signal power still remains in the in-phase component even after a considerable amount of CD.

2.7. Summary

In this chapter, we studied OSNR monitoring using polarization diversity and DSP analytically and through simulations for non-coherent systems. In absence of PMD, the scheme is shown to achieve an OSNR monitoring range of 9-39 dB for 40 Gb/s NRZ-OOK and RZ-QPSK systems with less than 1dB error. With first-order PMD, we developed conditions for the input signal polarization and the PSP of the fiber such that the proposed technique is insensitive to PMD if the product of DGD and electrical low-pass filter bandwidth of the optical detector is sufficiently small. For a 12.5 ps DGD and an electrical low-pass filter bandwidth of 312.5 MHz, a mean monitoring error of less than 0.5 dB is obtained from simulations. Furthermore, we derived conditions for first-order PMD independent OSNR monitoring and proposed a corresponding simple experimental setup based on a PS, PBS and balanced detectors. Simulation results show that a mean monitoring error of less than 0.5 dB can be realized regardless of the DGD value and the bandwidth of the electrical low-pass filter. And the experiment is demonstrated for an OSNR monitoring range of 8 to 23 dB. The observed PMD independence is in accordance with theoretical predictions.

Chapter 3 CD Estimation Using Auto-Correlation of Signal Power Waveform

3.1. Introduction

Coherent receiver is capable of compensating linear impairments such as CD and PMD, which can be modeled as all-pass filters. Hence inline optical CD compensation devices such as DCF are no longer necessary nor desired in future coherent systems. With a digital coherent receiver, CD can be estimated by inserting training symbols or specific bit sequences [35]-[37]. On the other hand, blind non-data aided CD estimation is often required/preferred [38] for equalizer initialization especially for links in dynamic optical networks without significant changes to the transmitter. Traditional CD estimation techniques rely on compensating the exact CD and have to scan through a range of preset CD values [38]-[44]. Hence to estimate large CD values, the number of symbols required is typically in the order of 100,000. Although there is no strict limitations on the time and symbols required in CD estimation, reducing the excess payloads can always benefit the link efficiency and power consumption, especially for long-haul transmission with large CD values. Recently, Malouin, et al. proposed an interesting CD estimation technique by examining the clock tone power without scanning [45]-[46]. Although the speed is fast, such clock tone will be considerably suppressed and the CD estimation technique may potentially be affected by PMD, especially when the DGD equals to half of the symbol period.

In this chapter, we propose the CD estimation technique for digital coherent receivers based on examining the auto-correlation of signal power waveform [47]-[49][48]. Due to CD-induced inter-symbol interference, the auto-correlation of signal power waveform exhibits a peak whose location is analytically shown to be indicative of the accumulated CD and hence can be employed for CD estimation. In addition, one can analytically show that the estimation technique is insensitive to amplifier noise, laser frequency offset, laser phase noise, first-order PMD and is applicable to single carrier PSK and QAM systems. Simulation results for 112 Gb/s PDM-QPSK as well as 224 Gb/s PDM-16-QAM systems show that the standard deviation of the CD estimate using only 8192 symbols is as low as 25 ps/nm for RZ and NRZ pulse shapes even in the presence of fiber nonlinearity. For systems with root-raised- cosine pulse shapes, we propose to use a band-pass digital filter to regenerate the correlation peak. And the required number of symbols is increased to 32768 for the roll-off factor down to 0.1.

3.2. Theoretical Foundations

Consider a single polarization coherent system in which the transmitted baseband signal can be expressed as

$$E_t(t) = \sum_n s_n p(t - nT), \tag{3.1}$$

where $p(t)$ denotes the pulse shape, T is the symbol period and s_n denotes i.i.d. PSK or QAM information symbols with

$$\mathbf{E}[s_n] = 0, \quad \mathbf{E}[s_n s_m^*] = \begin{cases} 1 & n = m \\ 0 & n \neq m \end{cases}. \tag{3.2}$$

Impairments other than CD such as ASE noise, multi-channel crosstalk, fiber nonlinearity and polarization effects are neglected in the theoretical analysis and will be discussed afterwards. To aid our theoretical derivations, we will assume a Gaussian pulse shape i.e.

$$p(t) = \exp\left(-\frac{t^2}{2T_0^2}\right), \tag{3.3}$$

where T_0 is the Gaussian pulse width. It should be noted that the insights drawn from the analysis can be applied to other practical pulse shapes such as NRZ and RZ through subsequent simulation results. For a link with length z and GVD coefficient β_2, the accumulated CD is characterized by the impulse response and corresponding transfer function

$$h(t) = \frac{1}{\sqrt{2\pi i\beta_2 z}}\exp\left(\frac{t^2}{2i\beta_2 z}\right) \xrightarrow{\mathcal{F}} H(f) = \exp(2i\pi^2\beta_2 z f^2), \tag{3.4}$$

where '$\mathcal{F}$' denotes Fourier transform. In this case, the pulse at the receiver is broadened to become

$$q(t) = \frac{T_0}{\sqrt{T_0^2 - i\beta_2 z}}\exp\left[-\frac{t^2}{2(T_0^2 - i\beta_2 z)}\right] \tag{3.5}$$

and the received baseband signal can be expressed as

$$E_r(t) = \sum\nolimits_n s_n q(t - nT)\cdot e^{j(2\pi\Delta f t + \psi(t))}, \tag{3.6}$$

where Δf denotes the frequency offset between transmitter laser and local oscillator, $\psi(t)$ represents the combined transmitter and local oscillator phase noise. If we directly detect the signal or calculate the optical power in a digital coherent receiver, we obtain the intensity as

$$I(t)=\left|E_r(t)\right|^2=\underbrace{\sum\nolimits_n\left|s_n q(t-nT)\right|^2}_{I_0(t)}+\underbrace{\sum\nolimits_{k\neq m}s_k s_m^* q(t-kT)q^*(t-mT)}_{I_1(t)}, \tag{3.7}$$

where $I_1(t)$ represents the CD-induced overlapping among neighboring pulses, $I_0(t)$ represents the sum of the power of the individual pulses. The average auto-correlation and power spectrum of $|s_n|^2$ are given by [50]

$$\bar{R}_{|s_n|^2}(k)=\begin{cases}\mathbf{E}\left[|s_n|^4\right] & k=0\\ \left(\mathbf{E}\left[|s_n|^2\right]\right)^2 & k\neq 0\end{cases} \tag{3.8}$$

and

$$\begin{aligned}\bar{S}_{|s_n|^2}(f)&=\sum\nolimits_k \bar{R}_{|s_n|^2}(k)e^{-2\pi ikfT}\\ &=\mathbf{Var}\left[|s_n|^2\right]+\left(\mathbf{E}\left[|s_n|^2\right]\right)^2\sum\nolimits_k e^{-2\pi ikfT}\\ &=\mathbf{Var}\left[|s_n|^2\right]+\frac{1}{T}\sum\nolimits_k \delta\left(f-\frac{k}{T}\right).\end{aligned} \tag{3.9}$$

And the average power spectrum of $I_0(t)$ is given by

$$\begin{aligned}\bar{S}_{I_0}(f)&=\frac{1}{T}\left|\mathcal{F}\left\{|q(t)|^2\right\}\right|^2\cdot\bar{S}_{|s_n|^2}(f)\\ &=\frac{\pi T_0^2}{T}\exp\left[-\frac{2\pi^2 f^2(T_0^4+\beta_2^2 z^2)}{T_0^2}\right]\cdot\left(\mathbf{Var}\left[|s_n|^2\right]+\frac{1}{T}\sum_k\delta\left(f-\frac{k}{T}\right)\right).\end{aligned} \tag{3.10}$$

Since the width of Gaussian envelop $\exp\left[-2\pi^2 f^2(T_0^4+\beta_2^2 z^2)/T_0^2\right]$ approaches 0 when $\beta_2 z$ increases to infinity. When the CD is large enough, $\bar{S}_{I_0}(f)$ approaches a delta function

$$\lim_{\beta_2 z\to\infty}\bar{S}_{I_0}(f)=\delta(f)\cdot\pi T_0^2/T^2 \tag{3.11}$$

and hence $I_0(t)$ approaches a constant

$$\lim_{\beta_2 z\to\infty} I_0(t)=\sqrt{\pi}T_0/T\,. \tag{3.12}$$

For a 28 G baud system, $I_0(t)$ can be well approximated by a DC component after 10

km SMF. And in the following analysis, we assume the CD is large enough for the approximation. Since

$$\mathbf{E}\left[I_1(t)\right]=\sum_{k\neq m}\mathbf{E}\left[s_k s_m^*\right]q(t-kT)q^*(t-mT)=0\,, \tag{3.13}$$

the cross-correlation between $\sqrt{\pi}T_0/T$ and I_1 is 0. Hence the auto-correlation of $I(t)$ is given by

$$R_I(\tau)=R_{I_0}(\tau)+R_{I_1}(\tau)\approx \pi T_0^2/T^2+R_{I_1}(\tau)\,. \tag{3.14}$$

Furthermore, since $I(t)$ is a cyclo-stationary process, the average auto-correlation is given by [51]

$$\bar{R}_{I_1}(\tau)=\frac{1}{T}\int_{-T/2}^{T/2}\mathbf{E}\left[I_1(t)I_1^*(t-\tau)\right]dt\,. \tag{3.15}$$

Since

$$E[s_k s_m^* s_r^* s_u]=\begin{cases}1 & k=r\ and\ m=u\\ 0 & otherwise\end{cases} \tag{3.16}$$

inserting (3.7) and (3.16) into (3.15) yields

$$\begin{aligned}
\bar{R}_{I_1}(\tau)&=\frac{1}{T}\int_{-T/2}^{T/2}\sum_{k\neq m,r\neq u}E[s_k s_m^* s_r^* s_u]q(t-kT)q^*(t-mT)q^*(t-rT-\tau)q(t-uT-\tau)dt\\
&=\frac{1}{T}\int_{-T/2}^{T/2}\sum_{k\neq m}q(t-kT)q^*(t-mT)q^*(t-kT-\tau)q(t-mT-\tau)dt\\
&=\frac{1}{T}\int_{-T/2}^{T/2}\sum_{v\neq 0}\sum_{k}q(t-kT)q^*(t-kT+vT)q^*(t-kT-\tau)q(t-kT+vT-\tau)dt\\
&=\frac{1}{T}\sum_{v\neq 0}\int_{-\infty}^{\infty}q(t)q^*(t+vT)q^*(t-\tau)q(t+vT-\tau)dt\\
&=\frac{T_0^4}{T(T_0^4+\beta_2^2z^2)}\sum_{v\neq 0}\exp\left[-\frac{(v^2T^2+\tau^2)T_0^2-2i\tau vT\beta_2 z}{2(T_0^4+\beta_2^2z^2)}\right]\int_{-\infty}^{\infty}\exp\left[-\frac{(2t+vT-\tau)^2T_0^2}{2(T_0^4+\beta_2^2z^2)}\right]dt\\
&=\frac{T_0^3}{T}\sqrt{\frac{\pi}{2(T_0^4+\beta_2^2z^2)}}\exp\left[-\frac{\tau^2T_0^2}{2(T_0^4+\beta_2^2z^2)}\right]\sum_{v\neq 0}\exp\left[-\frac{v^2T^2T_0^2-2i\tau vT\beta_2 z}{2(T_0^4+\beta_2^2z^2)}\right].
\end{aligned} \tag{3.17}$$

Then the term $\sum_{v\neq 0}\exp\left[-\frac{v^2T^2T_0^2-2i\tau vT\beta_2 z}{2(T_0^4+\beta_2^2z^2)}\right]$ can be further expressed as

$$
\begin{aligned}
&\sum_{v\neq 0}\exp\left[-\frac{v^2T^2T_0^2-2i\tau vT\beta_2 z}{2(T_0^4+\beta_2^2z^2)}\right]\\
&\xrightarrow{\mathcal{F}}\sum_{v\neq 0}\exp\left[-\frac{v^2T^2T_0^2}{2(T_0^4+\beta_2^2z^2)}\right]\delta\left[f-\frac{vT\beta_2 z}{2\pi(T_0^4+\beta_2^2z^2)}\right]\\
&\overset{(a)}{=}\exp\left[-\frac{2\pi^2f^2T_0^2(T_0^4+\beta_2^2z^2)}{\beta_2^2z^2}\right]\sum_{v\neq 0}\delta\left[f-\frac{vT\beta_2 z}{2\pi(T_0^4+\beta_2^2z^2)}\right]\\
&=\exp\left[-\frac{2\pi^2f^2T_0^2(T_0^4+\beta_2^2z^2)}{\beta_2^2z^2}\right]\sum_{v}\delta\left[f-\frac{vT\beta_2 z}{2\pi(T_0^4+\beta_2^2z^2)}\right]-\delta(f)\\
&\xrightarrow{\mathcal{F}^{-1}}\left\{\sqrt{\frac{\beta_2^2z^2}{2\pi^2T_0^2(T_0^4+\beta_2^2z^2)}}\exp\left[-\frac{\tau^2\beta_2^2z^2}{2T_0^2(T_0^4+\beta_2^2z^2)}\right]\right\}\\
&\quad\otimes\left\{\frac{2\pi(T_0^4+\beta_2^2z^2)}{T\beta_2 z}\sum_k\delta\left[\tau+\frac{2k\pi(T_0^4+\beta_2^2z^2)}{T\beta_2 z}\right]\right\}-1\\
&=\frac{\sqrt{2\pi(T_0^4+\beta_2^2z^2)}}{TT_0}\exp\left[-\frac{\tau^2\beta_2^2z^2}{2T_0^2(T_0^4+\beta_2^2z^2)}\right]\otimes\sum_k\delta\left[\tau+\frac{2k\pi(T_0^4+\beta_2^2z^2)}{T\beta_2 z}\right]-1
\end{aligned}
\tag{3.18}
$$

where the equality (a) uses the properties of Dirac delta function. Substituting (3.18) into (3.17), $\bar{R}_{I_1}(\tau)$ is given by

$$
\begin{aligned}
\bar{R}_{I_1}(\tau)=&\frac{\pi T_0^2}{T^2}\exp\left[-\frac{\tau^2T_0^2}{2(T_0^4+\beta_2^2z^2)}\right]\\
&\cdot\left\{\exp\left[-\frac{\tau^2\beta_2^2z^2}{2T_0^2(T_0^4+\beta_2^2z^2)}\right]\otimes\sum_k\delta\left[\tau+\frac{2k\pi(T_0^4+\beta_2^2z^2)}{T\beta_2 z}\right]-\frac{TT_0}{\sqrt{2\pi(T_0^4+\beta_2^2z^2)}}\right\}.
\end{aligned}
\tag{3.19}
$$

When $\beta_2 z\gg T_0^2$, let

$$
R_{env}(\tau)=\frac{\pi T_0^2}{T^2}\exp\left[-\frac{\tau^2T_0^2}{2(T_0^4+\beta_2^2z^2)}\right]\approx\frac{\pi T_0^2}{T^2}\exp\left(-\frac{\tau^2T_0^2}{2\beta_2^2z^2}\right) \tag{3.20}
$$

denotes the Gaussian envelop of $\bar{R}_{I_1}(\tau)$ and

$$
\tau_0=-\frac{2\pi(T_0^4+\beta_2^2z^2)}{T\beta_2 z}\approx-\frac{2\pi\beta_2 z}{T}. \tag{3.21}
$$

It should be noted that since $\beta_2<0$ for standard SMF, the negative sign in (3.21) ensures $\tau_0>0$. Then (3.19) becomes

$$\bar{R}_{I_1}(\tau) \approx R_{env}(\tau) \cdot \left[p(\tau) \otimes \sum_k \delta(\tau - k\tau_0) \right] = R_{env}(\tau) \cdot \sum_k p(\tau - k\tau_0) \qquad (3.22)$$

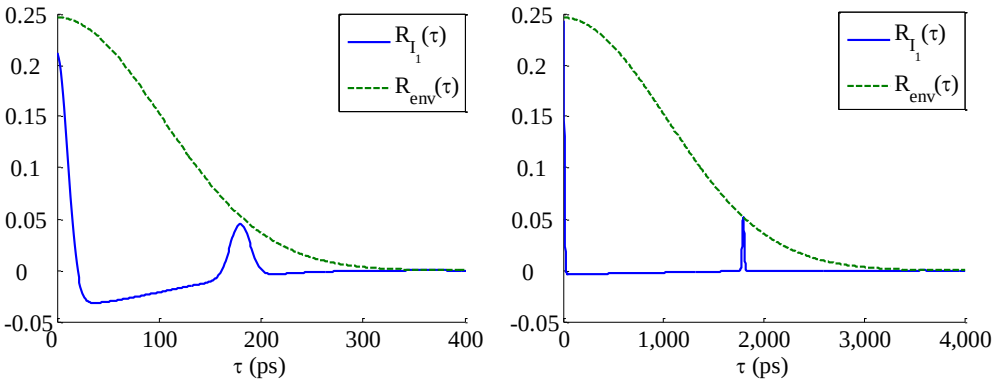
which comprises of a periodic Gaussian pulse train with shape $p(\tau)$ as defined in (3.3) and enclosed in a wide Gaussian envelop $R_{env}(\tau)$ as shown in Fig. 3.1. Both the period of the Gaussian pulse train τ_0 and the width of $R_{env}(\tau)$ increase proportionally with $\beta_2 z$. Hence the height of the Gaussian pulse at $\tau = k\tau_0$ is given by

$$\bar{R}_{I_1}(k\tau_0) \approx R_{env}(k\tau_0) \cdot p(0) \approx \frac{\pi T_0^2}{T^2} \exp\left(-\frac{2k^2\pi^2 T_0^2}{T^2} \right) \qquad (3.23)$$

which is independent of $\beta_2 z$ but exponentially decays with k^2. In addition, $R_{env}(\tau)$ will essentially suppress all the pulses of $\sum_k p(\tau - k\tau_0)$ except for $|k| \le 1$ as shown in Fig. 3.1. Therefore, τ_0, the location of the first pulse in $\bar{R}_{I_1}(\tau)$ as well as $\bar{R}_I(\tau)$ can be employed for CD estimation through

$$CD = -\frac{2\pi\beta_2 zc}{\lambda^2} \approx \frac{\tau_0 Tc}{\lambda^2} \qquad (3.24)$$

where λ is the carrier wavelength and c is the speed of light.



(a) (b)

Fig. 3.1. The function $\bar{R}_{I_1}(\tau)$ (solid line) and $R_{env}(\tau)$ (dashed line) obtained from (3.19) and (3.20) for 28 G baud single polarization system with T_0 = 10 ps in the presence of (a) 800 and (b) 8000 ps/nm CD.

In (3.6), we did not take into consideration of the ASE noise, which can be modeled as a band-limited Gaussian random process with bandwidth larger than 1/*T*, hence its auto-correlation will essentially be zero for $\tau \geq T$. Such correlation property will also hold for the signal-noise beating and noise-noise beating component in the optical intensity. In the following sections, simulation results show that the correlation peak will not be affect even when OSNR is 10 dB or below. Therefore, $\bar{R}_I(\tau)$ will essentially be unaffected by ASE noise for $\tau > T$. Therefore, τ_0 and hence the CD estimate will not be affected as long as $\tau > T$, which translates into 10 km SMF for 28 G baud system.

The discussions above provide the theoretical analysis of the proposed CD estimation technique using the auto-correlation of signal power waveform for single polarization system and analytically demonstrate its robustness against ASE noise, laser frequency offset and laser phase noise. For a PDM system, the optical intensity at the receiver is given by

$$I(t) = I_{x0}(t) + I_{y0}(t) + I_{x1}(t) + I_{y1}(t) \approx 2\sqrt{\pi}T_0 / T + I_{x1}(t) + I_{y1}(t) \qquad (3.25)$$

where the '*x*' and '*y*' in the subscripts identify the contributions from the two polarizations. As the symbols in the two polarizations are uncorrelated, $I_{x1}(t)$ and $I_{y1}(t)$ are uncorrelated as well. Since from (3.13) we also have $\mathbf{E}\left[I_{x1}(t)\right] = \mathbf{E}\left[I_{y1}(t)\right] = 0$, the cross-correlation between any two of $2\sqrt{\pi}T_0 / T$, $I_{x1}(t)$ and $I_{y1}(t)$ is 0. Therefore, the auto- correlation of signal power waveform for a PDM system is given by

$$R_I(\tau) = 4\pi T_0^2 / T^2 + R_{I_{x1}}(\tau) + R_{I_{y1}}(\tau) \tag{3.26}$$

and the peak at τ_0 in $R_{x1}(\tau)$ and $R_{y1}(\tau)$ can be used for CD estimation as well.

3.3. Immunity to First-order PMD

The baseband signal for a PDM system can be expressed as

$$\boldsymbol{E}_t(t) = \sum_n \begin{bmatrix} s_{x,n} \\ s_{y,n} \end{bmatrix} \cdot p(t - nT) \tag{3.27}$$

where $s_{x,n}$ and $s_{y,n}$ denote the i.i.d. information symbols for the two polarizations satisfying (3.2). In the presence of CD and first-order PMD, the received signal in the frequency domain can be expressed as

$$\boldsymbol{E}_r(f) = \boldsymbol{R}_2 \begin{bmatrix} e^{-2\pi i f \Delta\tau/2} & 0 \\ 0 & e^{2\pi i f \Delta\tau/2} \end{bmatrix} \boldsymbol{R}_1 H(f) \boldsymbol{E}_t(f) \tag{3.28}$$

where $\Delta\tau$ is the DGD, $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ denote unitary polarization rotation matrices satisfying

$$\boldsymbol{R}_k = \begin{bmatrix} \cos\theta_k e^{-i\varphi_k/2} & \sin\theta_k e^{i\varphi_k/2} \\ -\sin\theta_k e^{-i\varphi_k/2} & \cos\theta_k e^{i\varphi_k/2} \end{bmatrix}. \tag{3.29}$$

Let $\boldsymbol{E}_{x'}(t)$ and $\boldsymbol{E}_{y'}(t)$ denotes the signal on the two PSPs and

$$\begin{bmatrix} s'_{x,n} \\ s'_{y,n} \end{bmatrix} = \begin{bmatrix} \cos\theta_1 e^{-i\varphi_1/2} s_{x,n} & \sin\theta_1 e^{i\varphi_1/2} s_{y,n} \\ -\sin\theta_1 e^{-i\varphi_1/2} s_{x,n} & \cos\theta_1 e^{i\varphi_1/2} s_{y,n} \end{bmatrix} \tag{3.30}$$

denotes the equivalent symbol projected on the two PSP, then the time-domain received signal is given by

$$\boldsymbol{E}_r(t) = \boldsymbol{R}_2 \begin{bmatrix} E_{x'}(t) \\ E_{y'}(t) \end{bmatrix} = \boldsymbol{R}_2 \sum_n \begin{bmatrix} s'_{x,n} q(t - nT - \Delta\tau/2) \\ s'_{y,n} q(t - nT + \Delta\tau/2) \end{bmatrix}. \tag{3.31}$$

Since it can be shown that

$$\mathbf{E}[s'_{x,n}] = \mathbf{E}[s'_{y,n}] = \mathbf{E}[s'_{x,n}s'_{y,n}] = \mathbf{E}[s'_{x,n}s'^{*}_{y,n}] = 0 \tag{3.32}$$

and

$$E\left[|\, s'_{x,n}\,|^2\right] = E\left[|\, s'_{y,n}\,|^2\right] = 1 \tag{3.33}$$

hence $s'_{x,n}$ and $s'_{y,n}$ are uncorrelated. Then (3.32) can be seen as a new PDM signal with symbols $\left[s'_{x,n} \quad s'_{y,n}\right]^T$ and a mismatch between the two polarizations. Then the overall intensity is given by

$$\begin{aligned} I(t) &= |\, E_{x'}(t)\,|^2 + |\, E_{y'}(t)\,|^2 \\ &= I_{x'}(t) + I_{x'}(t) \\ &\approx \frac{2\sqrt{\pi}T_0}{T} + I_{x'1}(t) + I_{y'1}(t) \end{aligned} \tag{3.34}$$

as in (3.25). Since $s'_{x,n}$ and $s'_{y,n}$ are uncorrelated, $I_{x'1}(t)$ and $I_{y'1}(t)$ are uncorrelated as well. Hence

$$R_I(\tau) = 4\pi T_0^2 / T^2 + R_{I_{x'1}}(\tau) + R_{I_{y'1}}(\tau). \tag{3.35}$$

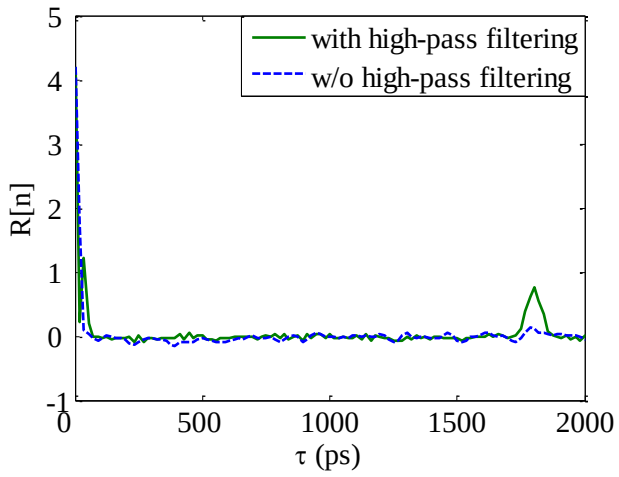
Since the new signal has the same power, symbol rate, pulse shape and accumulated CD with the actual signal, both $R_{I_{x'1}}(\tau)$ and $R_{I_{y'1}}(\tau)$ satisfy (3.19) and comprise a peak at τ_0. Therefore, we conclude that the peak at τ_0 is independent of $\boldsymbol{R}_1$ and $\Delta\tau$, meaning that the estimated CD is independent of first order PMD.

Yet it should be noted that the first order PMD independence is only valid for PDM system. For signal polarization system, the signal on the two PSPs is always correlated and hence $\bar{R}_I(\tau)$ will change with PMD.

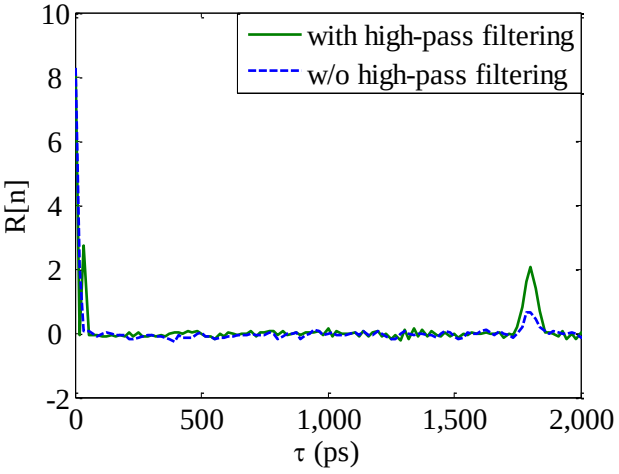
3.4. Simulation Results for NRZ and RZ Systems

Using VPI transmission maker, we carry out three simulations for 28 G baud NRZ, 50%

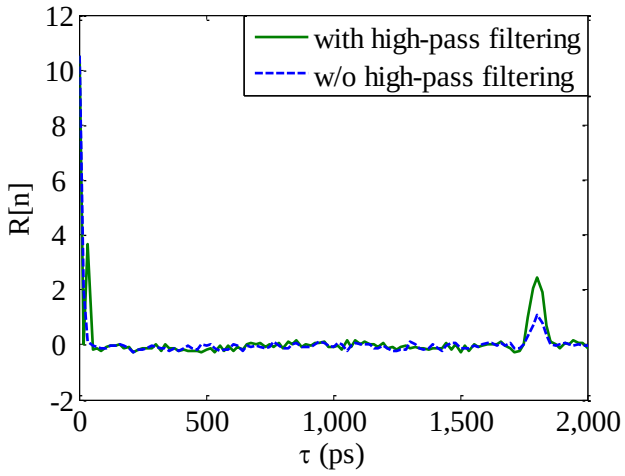
RZ and 66% RZ PDM coherent systems. For all the simulations, 8192 symbols are generated from 2^{30}–1 PRBS sequences and the signal is used to modulate a CW laser with 1550 nm wavelength and 1 MHz linewidth. At the receiver, the optical signal is passed through a 50 GHz -bandwidth 3rd-order Gaussian optical filter, followed by a 90 degree hybrid and balanced detection. The local oscillator has a random frequency offset of –2 GHz ~ 2 GHz and a linewidth of 1 MHz. The output electrical signal after the balanced detection is electrically filtered with 28 GHz bandwidth and sampled at 56 GSa/s. The auto-correlation of signal power waveform is calculated for CD estimation.



(a)



(b)



(c)

Fig. 3.2. $R[n]$ (DC blocked) for 224 Gb/s PDM-16-QAM systems in the presence of 8,000 ps/nm CD with (a) NRZ, (b) 50% RZ and (c) 66% RZ pulse shape. The amplitude $R[n]$ is normalized with the optical power (after the filter).

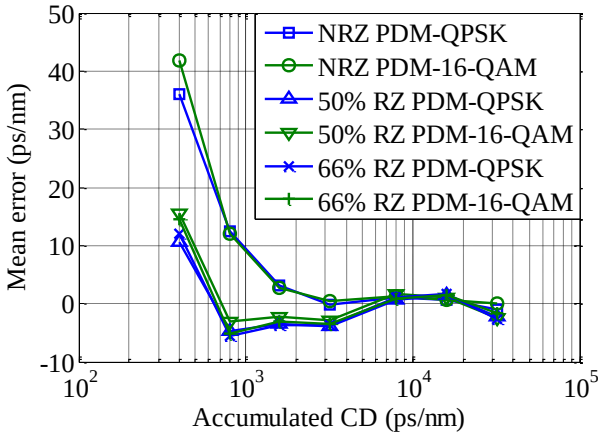
The first simulation is carried out for PDM-16-QAM systems with 8000 ps/nm accumulated CD. Let $\boldsymbol{E}_r[n]$ denotes the vector of sampled waveforms in both polarizations after balanced detection, the discrete time version of $R(\tau)$ is given by

$$R[n] = IFFT\left\{\left|FFT\left\{\boldsymbol{E}_r^{\dagger}[n]\boldsymbol{E}_r[n]\right\}\right|^2\right\} \tag{3.36}$$

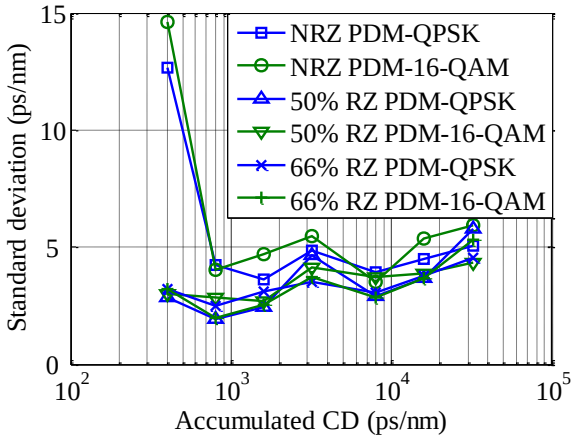
which is as shown in Fig. 3.2 with dashed line. For RZ pulse shape, it is easy to observe the pulse located at τ_0 in $R(\tau)$ as in Fig. 3.2(b) and (c). However, in the case of NRZ pulse shape, the pulse located at τ_0 in $R(\tau)$ is too 'weak' for CD estimation. To cope with the problem, we found that the peak in $R(\tau)$ can be boosted with high-pass filtering for all the pulse shapes. In particular, a delay-and-subtract operation on the sampled electrical waveform will serve as a high-pass filter and the corresponding auto-correlation of signal power waveform is given by

$$R[n] = IFFT\left\{\left|FFT\left\{\left(\boldsymbol{E}_r[n]-\boldsymbol{E}_r[n+1]\right)^{\dagger}\left(\boldsymbol{E}_r[n]-\boldsymbol{E}_r[n+1]\right)\right\}\right|^2\right\} \qquad (3.37)$$

which is as shown in Fig. 3.2 with solid lines.

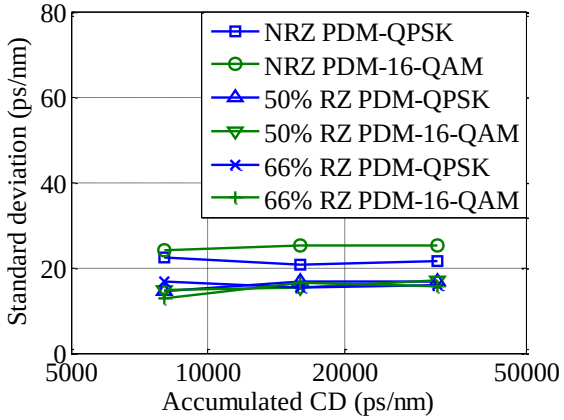


(a)

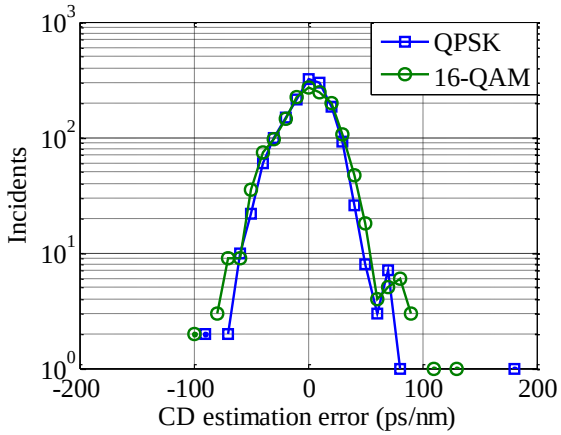


(b)

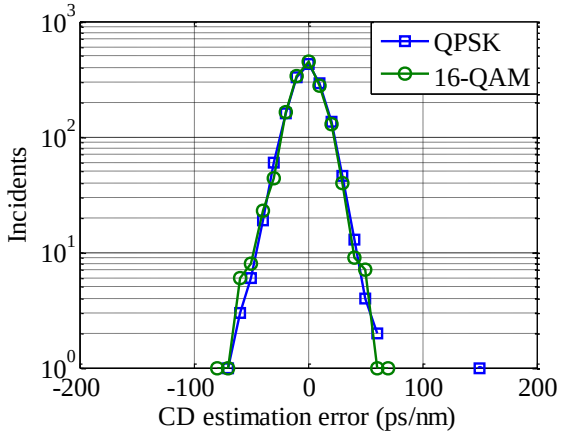
Fig. 3.3. (a) Mean error and (b) standard deviation obtained from 50 independent trials for 28 G baud PDM-QPSK and PDM-16-QAM systems with NRZ, 50% RZ and 66% RZ pulses in the presence of 400 to 32,000 ps/nm accumulated CD.


NRZ PDM-QPSK
NRZ PDM-16-QAM
50% RZ PDM-QPSK
50% RZ PDM-16-QAM
66% RZ PDM-QPSK
66% RZ PDM-16-QAM
Standard deviation (ps/nm)
Accumulated CD (ps/nm)
0
20
40
60
80
5000
10000
20000
50000

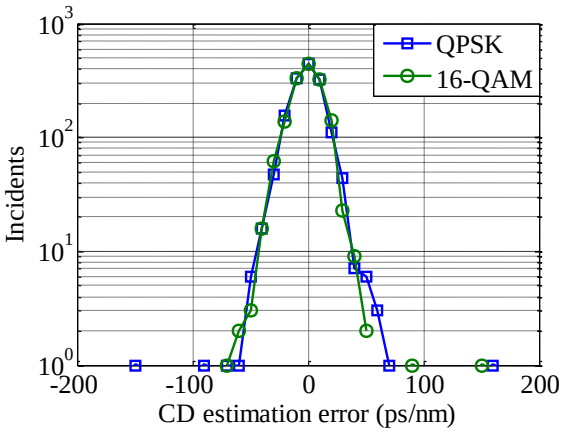
(a)


QPSK
16-QAM
Incidents
CD estimation error (ps/nm)
-200
-100
0
100
200

(b)


QPSK
16-QAM
Incidents
CD estimation error (ps/nm)
-200
-100
0
100
200

(c)



(d)

Fig. 3.4. CD estimation results for 10-channel NRZ, 50% RZ and 66% RZ 28 G baud PDM-QPSK and PDM-16-QAM systems with 10 dB OSNR, 25 ps mean DGD, –2 ~ 2 GHz frequency offset and 0 dBm signal power per channel. (a) Standard deviation of the estimation results. Error incidents of (b) NRZ (c) 50% RZ and (d) 66% RZ pulse shape.

In the second simulation, we investigate the performance of the proposed technique for PDM-QPSK and PDM-16-QAMsystems in the presence of 400 ~ 32,000 ps/nm CD. The mean error and standard deviation obtained from 50 independent trials are calculated and shown in Fig. 3.3. From Fig. 3.3 (a), we can see that the CD is over estimated at lower CD value region especially for NRZ pulse shapes. This is probably due to the fact that T_0 is neglected in (3.24). When CD increases above 800 ps/nm, the technique demonstrates excellent accuracy and precision for all the pulse shapes considered.

In the third simulation, the performances of our proposed CD estimation technique are investigated in a 10-channel DWDM coherent system with PDM-QPSK or PDM-16-QAM signals. The optical power for each channel is 0 dBm and the channel spacing is 50 GHz. Then the optical signal is sent through 500 to 2,000 km transmission links

with a span length of 83km and each EDFA fully compensates the loss of one span of fiber. The attenuation, dispersion and nonlinear coefficient of the fiber are 0.2 dB/km, 16 ps/nm ŀkm and 1.32/W-km, respectively. For all the cases, the OSNR (0.1 nm) at the receiver is 10 dB and the mean DGD is 25 ps. For each combination of modulation formats, pulse shapes and fiber lengths, 500 independent trials are performed to obtain the estimation statistics. The standard deviation of the estimation results are shown in Fig. 3.4(a) and the maximum standard deviation is 25 ps/nm. The distributions of the estimation error from all the simulation trials are shown in Fig. 3.4(b-d). For all the 9000 independent trials, 99.9% of the estimation errors are less than 100 ps/nm and the maximum error is 175 ps/nm.

3.5. Simulation Results for Systems with Root-Raised-Cosine Pulse Shapes

In order to achieve higher spectral efficiency, the spectrum of the transmitted signal is often shaped to be root-raised-cosine function as given by

$$P(f)=\begin{cases} 1 & |f|\le\dfrac{1-\beta}{2T} \\ \left[\dfrac{1}{2}+\dfrac{1}{2}\cos\left(\dfrac{\pi T}{\beta}\left(|f|-\dfrac{1-\beta}{2T}\right)\right)\right]^{1/2} & \dfrac{1-\beta}{2T}\le|f|\le\dfrac{1+\beta}{2T} \\ 0 & \textit{otherwise} \end{cases} \tag{3.38}$$

where β is the roll-off factor that is desired to be as small as possible. For such systems, even employing the two-taps filter as mentioned in the last section, the peak at τ_0 in the auto-correlation of the power waveform is still too small to be observed, especially

when $\beta < 0.5$. To regenerate the peak, we propose to filter the four tributaries of signal (in-phase and quadrature components in the two polarizations) with a band-pass filter given by

$$P(f)=\begin{cases}\left[\frac{1}{2}+\frac{1}{2}\cos\left(\frac{\pi T}{\beta}\left(\frac{1+\beta}{2T}-|f|\right)\right)\right]^{1/2} & \frac{1-\beta}{2T}\le|f|\le\frac{1+\beta}{2T}\\ 0 & otherwise\end{cases}. \tag{3.39}$$

The spectra of the transmitted signal, band-pass filter and the signal after filtering are as shown in Fig. 3.5. The aim of using such filter is to suppress the low-frequency component of the signal and we note that this approach is equally applicable to NRZ and RZ systems and digital filters with spectral shape similar to (3.39) will serve the purpose just as well.

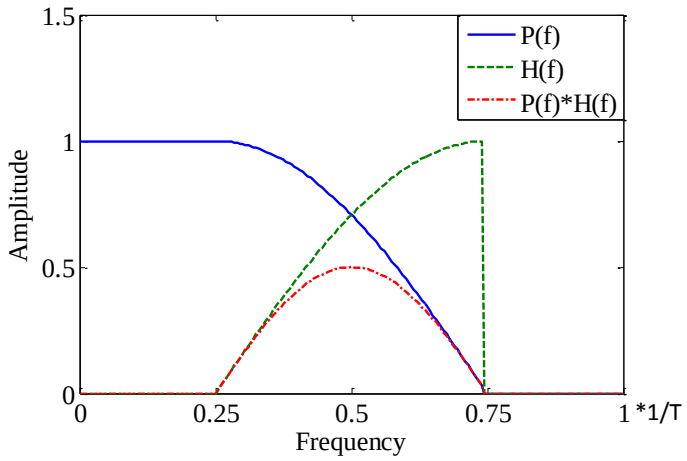


Fig. 3.5. Spectrum of the root-raised-cosine pulse $P(f)$ with $\beta = 0.5$, the corresponding filter $H(f)$ and the pulse spectrum after filtering.

By realizing (3.39) with a 64-tap digital filter, Fig. 3.6 shows the auto-correlation of the signal power waveform for 112 Gb/s PDM-QPSK system using root-raised-cosine pulse shapes with β ranging from 0.05 to 0.5 in the presence of 8000 ps/nm CD. It should be noted that a longer number of taps is desired to achieve (3.39), especially

when β is small. As can be seen in the figure, when β decreases, both the peaks at $\tau = 0$ and $\tau = \tau_0$ become broader. In order to resolve the two peaks for CD estimation, the lower bound of τ_0 and hence the CD value becomes larger for smaller β. Note that this problem can be easily solved by intentionally adding some CD using the digital filter before calculating $R[n]$. In addition, the peak at $\tau = \tau_0$ becomes broader, suggesting that the precision of the estimated τ_0 would become worse.

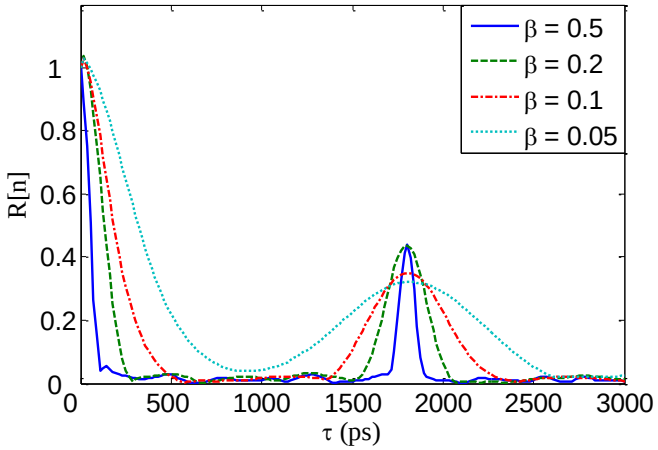
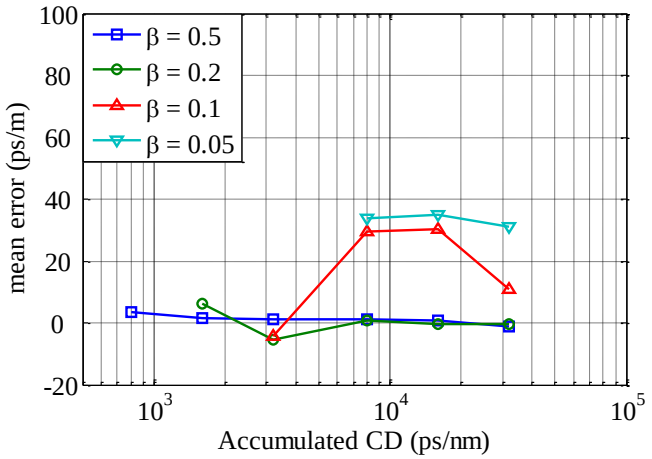
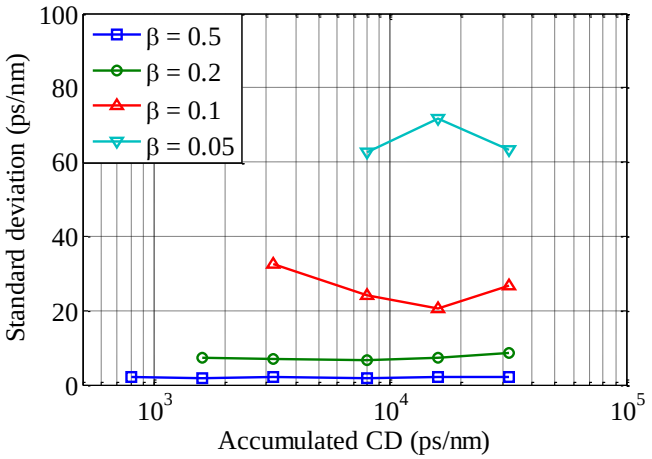


Fig. 3.6. $R[n]$ (DC blocked) for 112 Gb/s PDM-QPSK system using root-raised-cosine pulse shape with β ranging from 0.05 to 0.5 in the presence of 8000 ps/nm CD. The received signals in two quadratures and polarizations first passed through a 64-tap digital filter with transfer functions described in (3.39) before the calculation of $R[n]$.

For PDM-QPSK systems with β ranging from 0.05 to 0.5, we investigated the CD estimation performance with the proposed band-pass filter for 800, 1600, 3200 and 8000 ps/nm of CD. With 50 independent trials for each combination of β and CD value, the mean error and standard deviation of the CD estimation are shown in Fig. 3.7. An increase of standard deviation is observed for lower β.



(a)

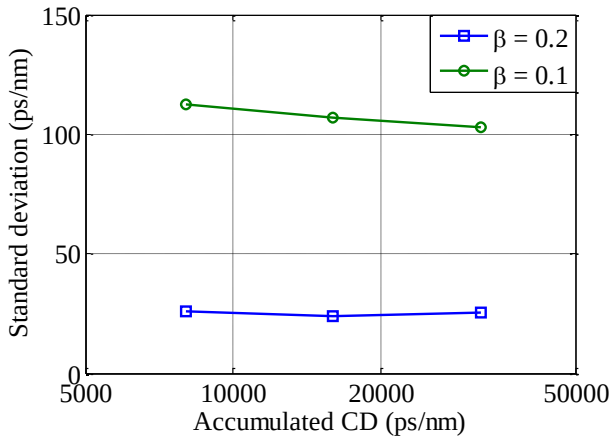


(b)

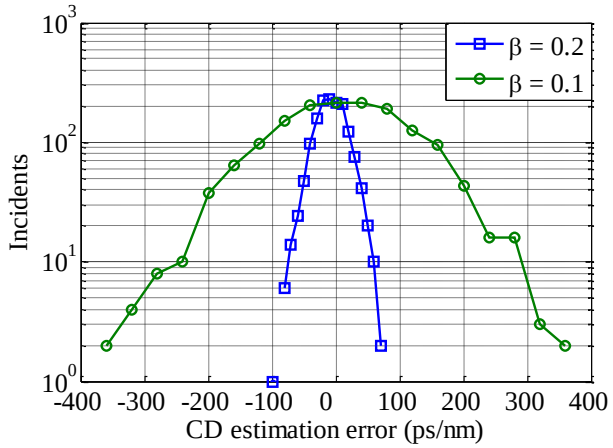
Fig. 3.7. (a) Mean error and (b) standard deviation obtained from 50 independent trials for 28G baud PDM-QPSK systems using root-raised-cosine spectrum with β ranging from 0.05 to 0.5 in the presence of 800 to 32000 ps/nm CD.

Finally, in order to investigate the robustness to other impairments for systems with root-raised-cosine pulse shapes, simulations are carried out for 10-channel DWDM PDM-QPSK systems transmitted over 500 to 2000 km of fiber. The optical power per channel is 0 dBm and the channel spacing is 50 GHz. The attenuation, dispersion and nonlinear coefficient of the fiber are 0.2 dB/km, 16 ps/nm ŀkm and 1.32/W-km, respectively. For all the cases, the OSNR (0.1 nm) at the receiver is 13 dB

and the mean DGD is 25 ps. For each combination of β and transmission distance, 500 independent trials are performed and the CD is estimated from 32768 symbols. The standard deviation and the distributions of the estimation error are shown in Fig. 3.8, both of which significantly increase when β decreased from 0.2 to 0.1



(a)



(b)

Fig. 3.8. (a) Standard deviation and (b) distributions of the CD estimation error for 10-channel DWDM 28G baud PDM-QPSK with root-raised-cosine spectral shaping. The OSNR, mean DGD and signal power per channel are 10 dB, 25 ps mean and 10 dBm, respectively.

3.6. Summary

In this chapter, we discuss a simple blind CD estimation technique using the auto-correlation of signal power waveform for digital coherent receivers. The technique demonstrates high estimation accuracies, large dynamic range, short estimation time and robustness against other transmission impairments. It can also potentially be implemented as a CD monitor with simple direct detection for intermediate nodes of long haul optical communication systems. The only drawback is the performance degradation with very narrow band filtering. The experimental verification of the proposed technique is shown in [52]. With less than 10,000 samples, CD can be accurately estimated up to 50,000 ps/nm with a maximum error of 186 ps/nm in the presence of 0~80 ps first order PMD.

Chapter 4 Quasi-Single-Mode Transmissions

4.1. Introduction

Fiber nonlinearity is a fundamental performance limiting factor in long-haul transmissions. Nonlinearity compensation algorithms such as digital back-propagation [53] are new tools to combat fiber nonlinearities. But they use split-step method to solve the inverse of the nonlinear Schrödinger equation which is too complex and only brings moderate transmission improvements at best. This partly motivates the current wave of research in mode division multiplexing transmissions [54]-[56] as alternatives to increase transmission capacities. Yet modal-multiplexing with FMF or multi-core fiber is highly impractical at present as they require few-mode or multi-core amplifiers, multiplexers, de-multiplexers in addition to highly complex DSP. In this regard, Yaman et al. [57]-[58] proposed to use the FMF as a large effective-area single-mode fiber and launch signals in the fundamental LP_{01} mode while keeping other single-mode components such as EDFA and transceivers in the link, otherwise known as QSM transmissions. The large-effective-area nature of the LP_{01} mode considerably extended the transmission distance. However, this may be specific and overly-optimistic scenario. In general, mode coupling between the LP_{01} and LP_{11} modes are expected to be present in practical QSM systems which will induce MPI and potentially limit transmission performance. In other significant QSM transmission work, 16×100 Gb/s PM-QPSK signals were successfully transmitted over 635 km of standard 50 μm core diameter

multimode fiber, but the reach was ultimately limited by MPI effects [59]. Bai et al. [60] briefly studied DSP techniques to mitigate MPI but long-haul QSM transmission experiments with MPI compensation has yet to be demonstrated.

In this chapter, we develop the theoretical foundations for polarization-multiplexed QSM transmission systems in the presence of mode coupling and differential mode loss between different mode groups in the FMF. We show that in the weak coupling regime, the multi-span QSM channel is a Gaussian random process in frequency in the weak coupling regime. MPI compensation filters are derived and mode-coupling-induced performance penalties are characterized. In particular, we show that MPI impairments can be effectively compensated by appropriate DSP techniques while signal loss in higher-order modes at the end of each span plays a significant role in the overall performance gain of QSM transmission systems.

The experiment is carried out for 256 Gb/s PM-16-QAM QSM transmission with 100 km per span and EDFA amplification [61]-[63]. The MPI is compensated by an additional DD-LMS algorithm after carrier phase estimation (CPE) in a standard DSP platform. The DD-LMS increases the optimal signal launch power by around 3 dB and the transmission distance for 256 Gb/s PM-16-QAM signals with 100 km per span can be extended up to 2600 km, comparable with the performance with Corning ultra low loss and effective area fiber EX2000 [64].

4.2. QSM Channel Model

Consider a span of few-mode fiber of length L modeled as a concatenation of K sections of length ΔL between which mode coupling may occur. For simplicity, we will initially only study the coupling between the fundamental mode and a higher-order mode with DMD $\Delta\tau$ and DML $\Delta\alpha$ as shown in Fig. 4.1. The input $E_{in}(\omega)$ is the transmitted signal launched into the fundamental mode with unit power and $E_{out}(\omega)$ is the signal output at the fundamental mode at the fiber end. Because of mode-coupling along the fiber, some signals will be present at the higher-order mode at the fiber end and will be stripped off as a consequence of butt-coupling with single-mode EDFAs or coherent receiver.

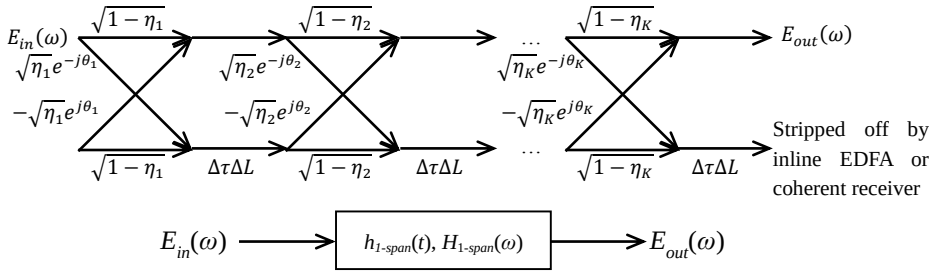


Fig. 4.1. The 2-mode coupling model for one span of QSM transmission.

In this case, the output signal is given by

$$E_{out}(\omega) = [1 \quad 0]\prod_{k=1}^{K}\left(\boldsymbol{G}(\omega)\boldsymbol{R}_k\right)\cdot\begin{bmatrix} E_{in}(\omega) \\ 0 \end{bmatrix} \tag{4.1}$$

where

$$\boldsymbol{G}(\omega) = \begin{bmatrix} 1 & 0 \\ 0 & e^{-(j\omega\Delta\tau+\Delta\alpha/2)\Delta L} \end{bmatrix} \tag{4.2}$$

takes into account of the effects of DMD and DML. The matrix

$$\boldsymbol{R}_k = \begin{bmatrix} \sqrt{1-\eta_k} & \sqrt{\eta_k}\, e^{-j\theta_k} \\ -\sqrt{\eta_k}\, e^{j\theta_k} & \sqrt{1-\eta_k} \end{bmatrix} \tag{4.3}$$

models random mode coupling at the k^{th} section through i.i.d. uniform random variables η_k and θ_k with distributions U[0, $2\kappa\Delta L$] and U[0, 2π] respectively. The coupling coefficient κ can be viewed as a measure of coupling strength. We note that there are other 2-mode models proposed in the literature but they are fundamentally very similar to ours and give identical statistical predictions. Furthermore, the dynamics of statistical mode coupling in linear fibers have been extensively studied in the '90s in which a continuous-fiber model is adopted [65]-[66]. We hereby choose to use the discrete concatenating-fiber model to better align with theoretical analysis in recent spatial-division multiplexing literature and we again emphasize that the conclusions drawn will be identical regardless of the fiber model chosen.

The transfer function of one span of QSM system is then given by

$$H_{1-span}(\omega) = [1 \quad 0] \prod_{k=1}^{K} \left(\boldsymbol{G}(\omega) \boldsymbol{R}_k \right) \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{4.4}$$

and let $h_{1-span}(t)$ denote the corresponding impulse response. We are interested in the statistics of $h_{1-span}(t)$ as it characterizes the amount of MPI and the length of the MPI compensation filter. In particular, starting from the PSD relation

$$\mathbf{E}\left[\left| h_{1-span}(t) \right|^2 \right] \xleftrightarrow{\mathcal{F}} \mathbf{E}\left[H^*_{1-span}(\omega') H_{1-span}(\omega' + \omega) \right], \tag{4.5}$$

the auto-correlation function of $H_{1-span}(\omega)$ in frequency i.e.

$$\begin{aligned} &\mathbf{E}\left[H^*_{1-span}(\omega') H_{1-span}(\omega' + \omega) \right] \\ &= \mathbf{E}\left[[1 \quad 0] \prod_{k=0}^{K-1} \left(\boldsymbol{R}^*_{K-k} \boldsymbol{G}^*(\omega') \right) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \prod_{k=1}^{K} \left(\boldsymbol{G}(\omega' + \omega) \boldsymbol{R}_k \right) \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right] \end{aligned} \tag{4.6}$$

Since R_m is statistically independent of R_n for $m \neq n$, we have

$$\begin{aligned}&\mathbf{E}[H^*_{1-span}(\omega')H_{1-span}(\omega'+\omega)]\\&=\mathbf{E}\left[\begin{bmatrix}1 & 0\end{bmatrix}\prod_{k=0}^{K-2}\boldsymbol{R}^*_{K-k}\boldsymbol{G}^*(\omega')\cdot\mathbf{E}\left[\boldsymbol{R}_1^*\boldsymbol{G}^*(\omega')\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\boldsymbol{G}(\omega'+\omega)\boldsymbol{R}_1\right]\cdot\prod_{k=2}^{K}\boldsymbol{G}(\omega'+\omega)\boldsymbol{R}_k\begin{bmatrix}1\\0\end{bmatrix}\right]\end{aligned} \tag{4.7}$$

in which

$$\mathbf{E}\left[\boldsymbol{R}_1^*\boldsymbol{G}^*(\omega')\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\boldsymbol{G}(\omega'+\omega)\boldsymbol{R}_1\right]=\mathbf{E}\left[\begin{bmatrix}1-\eta_1 & \sqrt{\eta_1(1-\eta_1)}e^{-j\theta_1}\\\sqrt{\eta_1(1-\eta_1)}e^{j\theta_1} & \eta_1\end{bmatrix}\right]=\begin{bmatrix}1-\kappa\Delta L & 0\\0 & \kappa\Delta L\end{bmatrix}. \tag{4.8}$$

Next, if we let $\xi = e^{-(j\omega\Delta\tau+\Delta\alpha)\Delta L}$

$$\begin{aligned}&\mathbf{E}\left[\boldsymbol{R}_2^*\boldsymbol{G}^*(\omega')\boldsymbol{R}_1^*\boldsymbol{G}^*(\omega')\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\boldsymbol{G}(\omega'+\omega)\boldsymbol{R}_1\boldsymbol{G}(\omega'+\omega)\boldsymbol{R}_2\right]\\&=\mathbf{E}\left[\boldsymbol{R}_2^*\boldsymbol{G}^*(\omega')\begin{bmatrix}1-\kappa\Delta L & 0\\0 & \kappa\Delta L\end{bmatrix}\boldsymbol{G}(\omega'+\omega)\boldsymbol{R}_2\right]\\&=\begin{bmatrix}(1-\kappa\Delta L)^2+\kappa^2\Delta L^2\xi & 0\\0 & (1-\kappa\Delta L)\kappa\Delta L\left(\xi+\xi^2\right)\end{bmatrix}\end{aligned} \tag{4.9}$$

and

$$\begin{aligned}&\mathbf{E}\left[\boldsymbol{R}_3^*\boldsymbol{G}^*(\omega')\boldsymbol{R}_2^*\boldsymbol{G}^*(\omega')\boldsymbol{R}_1^*\boldsymbol{G}^*(\omega')\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\boldsymbol{G}(\omega'+\omega)\boldsymbol{R}_1\boldsymbol{G}(\omega'+\omega)\boldsymbol{R}_2\boldsymbol{G}(\omega'+\omega)\boldsymbol{R}_3\right]\\&=\begin{bmatrix}(1-\kappa\Delta L)^3+(1-\kappa\Delta L)\kappa^2\Delta L^2(2\xi+\xi^2) & 0\\0 & (1-\kappa\Delta L)\kappa\Delta L\left(\xi+\xi^2+\xi^3\right)\end{bmatrix}\end{aligned} \tag{4.10}$$

and the process can be continued recursively. For a fiber with *K* sections, we have

$$\begin{aligned}&\mathbf{E}\left[\prod_{k=0}^{K-1}\left(\boldsymbol{R}^*_{K-k}\boldsymbol{G}^*(\omega')\right)\begin{bmatrix}1\\0\end{bmatrix}\begin{bmatrix}1 & 0\end{bmatrix}\prod_{k=1}^{K}\left(\boldsymbol{G}(\omega'+\omega)\boldsymbol{R}_k\right)\right]\\&\approx\begin{bmatrix}(1-\kappa\Delta L)^K+(1-\kappa\Delta L)^{K-2}\kappa^2\Delta L^2\cdot\sum_{p=1}^{K}(K-p)\xi^p & 0\\0 & (1-\kappa\Delta L)\kappa\Delta L\cdot\sum_{p=1}^{K}\xi^p\end{bmatrix}\end{aligned} \tag{4.11}$$

in the weak coupling regime and neglect $\kappa^p\Delta L^p$ for $p \geq 3$. In this case,

$$\begin{aligned}&\mathbf{E}[H_{1-span}^{*}(\omega')H_{1-span}(\omega'+\omega)]\\&=(1-\kappa\Delta L)^{K}+(1-\kappa\Delta L)^{K-2}\kappa^{2}\Delta L^{2}\cdot\sum_{p=1}^{K}(K-p)\xi^{p}.\end{aligned}\tag{4.12}$$

And its Fourier transform is given by

$$\mathbf{E}\left[\left|h_{1-span}(t)\right|^{2}\right]=(1-\kappa\Delta L)^{K}\left(\delta(t)+\left(\frac{\kappa\Delta L}{1-\kappa\Delta L}\right)^{2}\sum_{p=1}^{K}(K-p)e^{p\Delta\alpha\Delta L}\delta(t-p\Delta\tau\Delta L)\right).\tag{4.13}$$

Finally, in the limit $K\to\infty$, we obtain

$$\lim_{K\to\infty}\mathbf{E}\left[\left|h_{1-span}(t)\right|^{2}\right]=\begin{cases}e^{-\kappa L}\delta(t)+e^{-\kappa L}\cdot\underbrace{\frac{\kappa^{2}L}{\Delta\tau}e^{-\frac{\Delta\alpha t}{\Delta\tau}}\left(1-\frac{t}{\Delta\tau L}\right)}_{\mathbf{E}\left[\left|h_{MPI,1-span}(t)\right|^{2}\right]}, & 0\le t\le \Delta\tau L\\ 0 & \text{otherwise}\end{cases},\tag{4.14}$$

where $\mathbf{E}\left[\left|h_{MPI,1-span}(t)\right|^{2}\right]$ is the MPI component of $h_{1-span}(t)$. For a multi-span QSM system with per-span amplification, let $\otimes$ denotes convolution and $h_{MPI,n}(t)$ be the MPI component of the n^{th} span of fiber. In this case, the overall impulse response $h(t)$ is given by

$$\begin{aligned}h(t)&=\left(\delta(t)+h_{MPI,1}(t)\right)\otimes\left(\delta(t)+h_{MPI,2}(t)\right)\otimes\ldots\otimes\left(\delta(t)+h_{MPI,N}(t)\right)\\&\approx\left(\delta(t)+\sum_{n=1}^{N}h_{MPI,n}(t)\right)\end{aligned}\tag{4.15}$$

in the weak coupling regime where $h_{MPI,n}(t)$ is small. Now, since $h_{MPI,n}(t)$ are statistically independent for different spans,

$$\begin{aligned}\mathbf{E}\left[\left|h(t)\right|^{2}\right]&\approx\delta(t)+\sum_{n=1}^{N}\mathbf{E}\left[\left|h_{MPI,n}(t)\right|^{2}\right]\\&=\begin{cases}\delta(t)+\underbrace{\frac{N\kappa^{2}L}{\Delta\tau}e^{-\frac{\Delta\alpha t}{\Delta\tau}}\left(1-\frac{t}{\Delta\tau L}\right)}_{\mathbf{E}\left[\left|h_{MPI}(t)\right|^{2}\right]} & 0\le t\le \Delta\tau L\\ 0 & \text{otherwise}\end{cases}\end{aligned}\tag{4.16}$$

and the corresponding MPI power is

$$P_{MPI} = \int_0^{\Delta\tau L} \mathbf{E}\left[\left|h_{MPI}(t)\right|^2\right] dt = N\kappa^2 \cdot \frac{\Delta\alpha L + e^{-\Delta\alpha L} - 1}{\Delta\alpha^2}. \tag{4.17}$$

Note that the MPI power is independent of DMD.

The 2-mode coupling model can be readily generalized to 6 spatial-polarization modes for realistic polarization-multiplexed QSM transmissions. In this case, the input-output relation is given by

$$\begin{aligned} \begin{bmatrix} E_{out,x}(\omega) \\ E_{out,y}(\omega) \end{bmatrix} &= \boldsymbol{H}(\omega)\boldsymbol{E}_{in}(\omega) + \boldsymbol{V}(\omega) \\ &= \begin{bmatrix} H_{xx}(\omega) & H_{xy}(\omega) \\ H_{yx}(\omega) & H_{xx}(\omega) \end{bmatrix} \begin{bmatrix} E_{in,x}(\omega) \\ E_{in,y}(\omega) \end{bmatrix} + \begin{bmatrix} V_x(\omega) \\ V_y(\omega) \end{bmatrix} \end{aligned} \tag{4.18}$$

where $V_{x(y)}(\omega)$ is the Fourier Transform of the collective amplifier noise of the link, which is modeled as a circularly symmetric white Gaussian random process. Signals in the two fundamental modes (and any two modes of the 4 higher-order modes) are assumed to couple freely between each other i.e. the corresponding η_k will be uniformly distributed between 0 and 1. For coupling between any fundamental mode and any higher-order mode, the coupling coefficient is changed from κ to $\kappa/4$. The derivations for the expected impulse response powers for polarization-multiplexed QSM systems are largely similar and will be omitted here. The results are

$$\mathbf{E}\left[\left|h_{xx}(t)\right|^2\right] = \mathbf{E}\left[\left|h_{yy}(t)\right|^2\right] = \delta(t) + \frac{1}{4}\mathbf{E}\left[\left|h_{MPI}(t)\right|^2\right] \tag{4.19}$$

and

$$\mathbf{E}\left[\left|h_{xy}(t)\right|^2\right] = \mathbf{E}\left[\left|h_{yx}(t)\right|^2\right] = \frac{1}{4}\mathbf{E}\left[\left|h_{MPI}(t)\right|^2\right]. \tag{4.20}$$

In addition, we are also interested in the probability density function (PDF) of the QSM channel $H(\omega)$. In the weak coupling regime, we consider the channel output

consists of signals that have stayed in the fundamental mode along transmission as well as accumulated MPI from signals that have 'coupled up' to higher-order mode and 'coupled back down' to fundamental mode only once during transmission. There are $K(K-1)/2$ such configurations in a span of K sections and hence $H_{1-span}(\omega)$ can be expressed as

$$\begin{aligned} H_{1-span}(\omega) &= \prod_{k=1}^{K}\sqrt{1-\eta_k} + \prod_{k=1}^{K}\sqrt{1-\eta_k}\cdot\sum_{m=1}^{K-1}\sum_{l=1}^{K-m}\sqrt{\frac{\eta_l\eta_{m+l}}{(1-\eta_l)(1-\eta_{m+l})}}e^{-m\Delta\alpha\Delta L+j(\theta_{m+l}-\theta_l+m\Delta\tau\Delta L\omega)} \\ &= \prod_{k=1}^{K}\sqrt{1-\eta_k} + \prod_{k=1}^{K}\sqrt{1-\eta_k}\cdot\sum_{m=1}^{K-1}\sum_{l=1}^{K-m}\sqrt{\frac{\eta_l\eta_{m+l}}{(1-\eta_l)(1-\eta_{m+l})}}e^{-m\Delta\alpha\Delta L+j\sum_{q=l}^{m+l-1}x_q}, \end{aligned} \tag{4.21}$$

where m denotes the number of consecutive sections the signal 'stays' in the higher-order mode and $x_q = \theta_{q+1} - \theta_q + \Delta\tau\Delta L\omega$. Note that the x_q's are i.i.d. random variables with distribution U[0, 2π] and is in fact independent of ω due to the modulo 2π nature of phase. Consequently, the distribution of $H_{1-span}(\omega)$ is also independent of ω. As $K\to\infty$, $\prod_{k=1}^{K}\sqrt{1-\eta_k} \to e^{-\kappa L/2}$ and hence $\mathbf{E}\left[H_{1-span}(\omega)\right] \to e^{-\kappa L/2}$ and

$$\mathbf{Var}\left(H_{1-span}(\omega)\right) = \int \mathbf{E}\left[\left|h_{MPI,1-span}(t)\right|^2\right]dt = \frac{1}{N}P_{MPI}. \tag{4.22}$$

For an N-span QSM transmission system, denote

$$Y_n = \sum_{m=1}^{K-1}\sum_{l=1}^{K-m}\sqrt{\frac{\eta_l\eta_{m+l}}{(1-\eta_l)(1-\eta_{m+l})}}e^{-m\Delta\alpha\Delta L+j\sum_{q=l}^{m+l-1}x_q} \tag{4.23}$$

as the MPI terms incurred in the n^{th} span. In the weak coupling regime, the transfer function with per-span amplification gain $e^{\kappa L}$ becomes

$$H(\omega) = \prod_{n=1}^{N}(1+Y_n) \approx \left(1+\sum_{n=1}^{N}Y_n\right) \xrightarrow{N \text{ large}} \mathcal{N}\left(1, \frac{1}{2}P_{MPI}\boldsymbol{I}\right) \tag{4.24}$$

by the central limit theorem as N gets large. Together with the derived $E\left[\left|h(t)\right|^2\right]$ in the paper, one can show that $H(\omega)$ is a stationary complex Gaussian random process in

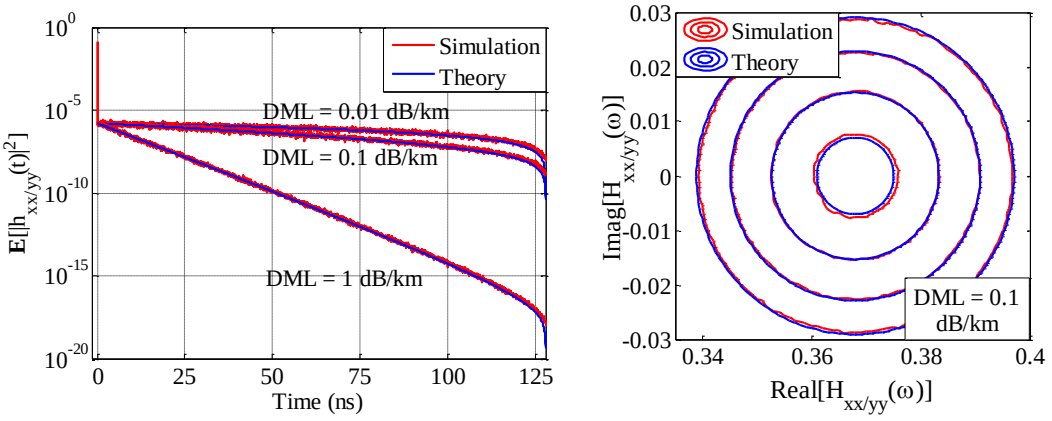
ω. Generalizations to the QSM transfer function matrix $H(\omega)$ for polarization-multiplexed transmissions in 6-spatial-polarization mode FMF systems are straight forward and are given by

$$H_{xx(yy)}(\omega) \sim \mathcal{N}\left(1, \frac{1}{8} P_{MPI} \boldsymbol{I}\right) \tag{4.25}$$

and

$$H_{xy(yx)}(\omega) \sim \mathcal{N}\left(0, \frac{1}{8} P_{MPI} \boldsymbol{I}\right) \tag{4.26}$$

where $\boldsymbol{I}$ is the identity matrix. Altogether, in the weak coupling regime, the transfer function of a N-span QSM system $H(\omega)$ is a stationary Gaussian random process in frequency with auto-correlation function $\mathbf{E}\left[|h_{xx(yy)}(t)|^2\right]$ and $\mathbf{E}\left[|h_{xy(yx)}(t)|^2\right]$



(a) (b)

Fig. 4.2. Theoretical and simulation results for (a) Expected impulse response power $\mathbf{E}\left[|h_{xx(yy)}(t)|^2\right]$ and (b) PDF of transfer function $H_{xx(yy)}(\omega)$ for a 32 Gbaud polarization-multiplexed QSM system. The link consists of 20 spans of 100-km fiber with κL = 0.1, DMD $\Delta\tau$ = 1.28 ns/km and DML $\Delta\alpha$ = 0.1 dB/km.

Simulations are conducted to verify the theoretical predictions and results shown in Fig. 4.2, which indicate good agreements between them. Further simulations reveal

that $\mathbf{E}\left[|h_{xx(yy)}(t)|^2\right]$ is insensitive to typical DGD between the two fundamental modes and thus PMD effects. This should be expected as typical FMFs have DMDs much larger than the DGD between the fundamental modes.

4.3. Transmission Penalties and DSP Mitigation Strategies

Mode coupling effects and DMD induce MPI and ISI in QSM transmissions. In principle, the ISI or MPI length should scale with increasing DMD and decreasing DML. The MPI length dictates the number of taps and hence the computational complexity of the MPI compensation filter in a digital coherent receiver. If we define τ_{MPI} as the length of time that contains 99% of the MPI power, we have

$$\int_0^{\tau_{MPI}} \left|h_{MPI}(t)\right|^2 dt = \int_0^{\tau_{MPI}} \left(1-\frac{t}{\Delta\tau L}\right)\cdot e^{-\frac{\Delta\alpha t}{\Delta\tau}} dt = 0.99\int_0^{\Delta\tau L}\left(1-\frac{t}{\Delta\tau L}\right)\cdot e^{-\frac{\Delta\alpha t}{\Delta\tau}} dt \quad (4.27)$$

and the solution is given by

$$\begin{aligned}\tau_{MPI} &= \frac{\Delta\tau}{\Delta\alpha}\left(\Delta\alpha L - W\left(\left(0.01\left(\Delta\alpha L-1\right)-0.99e^{-\Delta\alpha L}\right)e^{\Delta\alpha L-1}\right)-1\right)\\ &\xrightarrow{\Delta\alpha L \text{ large}} \frac{\Delta\tau}{\Delta\alpha}\cdot\ln 100\end{aligned} \quad (4.28)$$

where $W(\cdot)$ is the Lambert W function [69]. Fig. 4.3 shows the MPI length τ_{MPI} for various DMD and DML values for a 32Gbaud QSM transmission system.

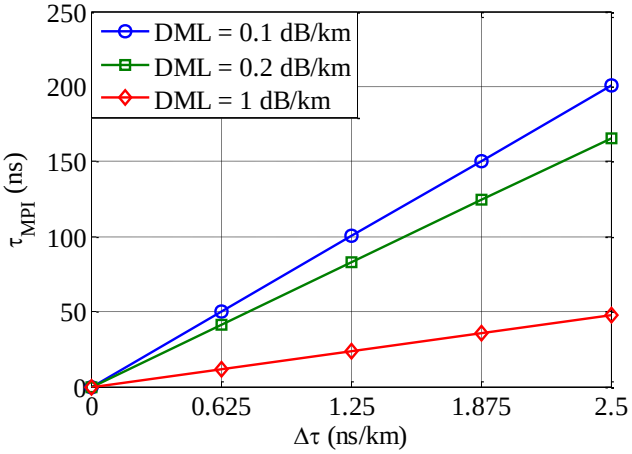


Fig. 4.3. The MPI length τ_{MPI} for various DMD $\Delta\tau$ and DML $\Delta\alpha$ for a 32 Gbaud polarization-multiplexed QSM transmission system.

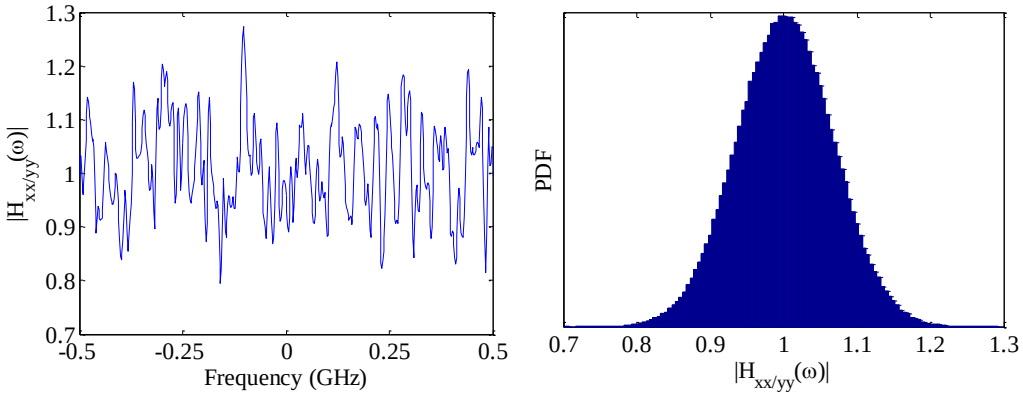


Fig. 4.4. (a) A sample realization of $H_{xx/yy}(\omega)$ and (b) PDF of the amplitude $|H_{xx/yy}(\omega)|$ for a 20-span QSM system with 100-km span length, $\kappa L = 0.1$, $\Delta\alpha = 1$ dB/km and $\Delta\tau = 1.28$ ns/km.

In the frequency domain, MPI manifests itself as frequency-selective fading, very much like multi-path fading in wireless communications. A sample realization of $H_{xx(yy)}(\omega)$ and the PDF of its amplitude is shown in Fig. 4.4(a). Minimum mean square error (MMSE) filter $W(\omega)$ can be used to compensate MPI and is given by

$$\boldsymbol{W}(\omega) = \left(\boldsymbol{H}^{\dagger}(\omega)\boldsymbol{H}(\omega)S_S(\omega) + N_0\boldsymbol{I}\right)^{-1}\boldsymbol{H}(\omega)S_S(\omega), \tag{4.29}$$

where $S_S(\omega)$ is the PSD of the input signal and the '†' in the superscript denotes

conjugate transpose. N_0 is the PSD of the amplifier noise $V_{x(y)}(\omega)$ and the filter output is given by

$$\begin{aligned}\begin{bmatrix} R_{out,x}(\omega) \\ R_{out,\mathrm{y}}(\omega) \end{bmatrix} &= \boldsymbol{W}(\omega)\boldsymbol{H}(\omega)\boldsymbol{E}_{in}(\omega) + \boldsymbol{W}(\omega)\boldsymbol{V}(\omega) \\ &= \boldsymbol{U}(\omega)\boldsymbol{E}_{in}(\omega) + \boldsymbol{W}(\omega)\boldsymbol{V}(\omega).\end{aligned} \tag{4.30}$$

Here we assume the power of $\boldsymbol{H}(\omega)$ is normalized and the transmitted signal has Nyquist pulse shaping. Note that the '(ω)' is omitted from (4.31) to (4.37) for simplicity. The channel transfer function can be expressed as

$$\boldsymbol{H} = \begin{bmatrix} H_{xx} & H_{yx} \\ H_{xy} & H_{yy} \end{bmatrix}. \tag{4.31}$$

The MMSE filter is given by

$$\boldsymbol{W} = \left(\boldsymbol{H}^{\dagger}\boldsymbol{H} + SNR^{-1}\boldsymbol{I}\right)^{-1}\boldsymbol{H}^{\dagger} = D\left(\det(\boldsymbol{H}^{*})\begin{bmatrix} H_{yy} & -H_{yx} \\ -H_{xy} & H_{xx} \end{bmatrix} + SNR^{-1}\boldsymbol{H}^{\dagger}\right) \tag{4.32}$$

where

$$D = \left(\det(\boldsymbol{H}^{\dagger}\boldsymbol{H}) + SNR^{-1}\left(\left|H_{xx}\right|^2 + \left|H_{xy}\right|^2 + \left|H_{yx}\right|^2 + \left|H_{yy}\right|^2\right)\right)^{-1}. \tag{4.33}$$

The signal after equalization is

$$\begin{aligned}\boldsymbol{W}\boldsymbol{E}_{out} &= D\left(\det(\boldsymbol{H}^{\dagger}\boldsymbol{H})\boldsymbol{I} + SNR^{-1}\boldsymbol{H}^{\dagger}\boldsymbol{H}\right)\boldsymbol{E}_{in} + \boldsymbol{W}\boldsymbol{N} \\ &= D\left(\begin{bmatrix} \det(\boldsymbol{H}^{\dagger}\boldsymbol{H}) + SNR^{-1}\left(\left|H_{xx}\right|^2 + \left|H_{xy}\right|^2\right) & SNR^{-1}\left(H_{xx}^{*}H_{yx} + H_{xy}^{*}H_{yy}\right) \\ SNR^{-1}\left(H_{yx}^{*}H_{xx} + H_{yy}^{*}H_{xy}\right) & \det(\boldsymbol{H}^{\dagger}\boldsymbol{H}) + SNR^{-1}\left(\left|H_{yy}\right|^2 + \left|H_{yx}\right|^2\right) \end{bmatrix}\boldsymbol{E}_{in} + \boldsymbol{W}\boldsymbol{N}\right).\end{aligned} \tag{4.34}$$

The effective signal, interference and noise power spectral densities in x polarization are given by

$$P'_{Sx} = \left\{\mathbf{E}\left[D\left(\det(\boldsymbol{H}^{\dagger}\boldsymbol{H}) + SNR^{-1}\left(\left|H_{xx}\right|^2 + \left|H_{xy}\right|^2\right)\right)\right]\right\}^2 \tag{4.35}$$

$$P'_{Ix} = \mathrm{var}\left[D\left(\det(\boldsymbol{H}^{\dagger}\boldsymbol{H}) + SNR^{-1}\left(\left|H_{xx}\right|^2 + \left|H_{xy}\right|^2\right)\right)\right] + \mathbf{E}\left[\left|SNR^{-1}D(H_{xx}^{*}H_{yx} + H_{xy}^{*}H_{yy})\right|^2\right] \tag{4.36}$$

$$P'_{Nx} = SNR^{-1}\left\{\mathbf{E}\left[\left|D\left(H_{yy}\det(\boldsymbol{H}^{\dagger}) + SNR^{-1}H^{*}_{xx}\right)\right|^{2}\right] + \mathbf{E}\left[\left|D\left(-H_{yx}\det(\boldsymbol{H}^{\dagger}) + SNR^{-1}H^{*}_{xy}\right)\right|^{2}\right]\right\} \quad (4.37)$$

where the expectation and variance are taken in frequency domain. And similar in y polarization.

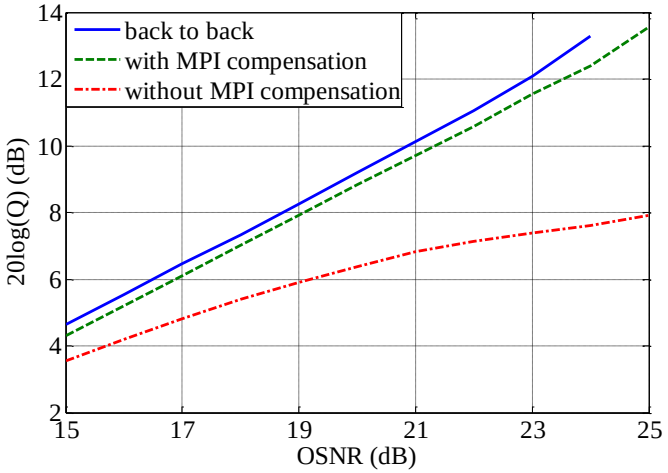


Fig. 4.5. *Q* vs. OSNR for 32 Gbaud-PM-16-QAM transmissions with and without MPI compensation for a 20-span QSM system with a random realization of the transfer function $\boldsymbol{H}(\omega)$. The span length is 100 km, coupling strength κL = 0.1, DMD $\Delta\tau$ = 1.28 ns/km and DML $\Delta\alpha$ = 0.1 dB/km.

Fig. 4.5 shows the *Q* vs. OSNR for 32 Gbaud-PM-16-QAM transmissions with and without MPI compensation for a 20-span QSM system with a random realization of the transfer function $\boldsymbol{H}(\omega)$. The *Q* value is derived from $Q = \sqrt{2}erfc^{-1}(2\cdot\mathrm{BER})$. It is clear that MPI significantly degrades system performance if left uncompensated. Fortunately, MMSE filtering can almost eliminate the effects of MPI, but not completely. This is because the minimum mean-squared error-criterion introduces noise enhancement, residual ISI and hence a performance penalty. Furthermore, the penalty is in general random due to the random fading nature of $H_{xx(yy)}(\omega)$. If we also consider the residual ISI as an additional source of noise, this MPI-induced penalty Λ_{MPI} will

be given by (assuming unit input signal PSD)

$$\Lambda_{MPI} = \frac{P'_{Ix} + P'_{Iy} + P'_{Nx} + P'_{Ny}}{2N_0\left(P'_{Sx} + P'_{Sy}\right)}. \tag{4.38}$$

In addition to Λ_{MPI}, another mode-coupling-induced performance penalty Λ_{HOM} stems from the loss of signals in higher-order modes at the end of each span. This penalty can be approximated by the power remaining in the fundamental mode at the end of each span, i.e.

$$\Lambda_{HOM} \approx \lim_{K\to\infty} \prod_{k=1}^{K} \left(1-\eta_k\right) = e^{-\kappa L}. \tag{4.39}$$

Compared with standard fiber systems, the QSM system has a larger effective area, smaller nonlinearity coefficient which in turn result in an OSNR gain

$$G_{OSNR} = A_{eff,FMF} / A_{eff,Std.fiber}. \tag{4.40}$$

Assuming the fundamental mode loss coefficient are the same for standard fiber systems and QSM systems and taking into account MPI effects and higher-order mode stripping, the overall performance gain $G_{overall}$ of QSM relative to standard single-mode fiber transmission systems is then given by

$$G_{overall} = G_{OSNR}\,(\text{dB}) - \Lambda_{MPI}\,(\text{dB}) - \Lambda_{HOM}\,(\text{dB}). \tag{4.41}$$

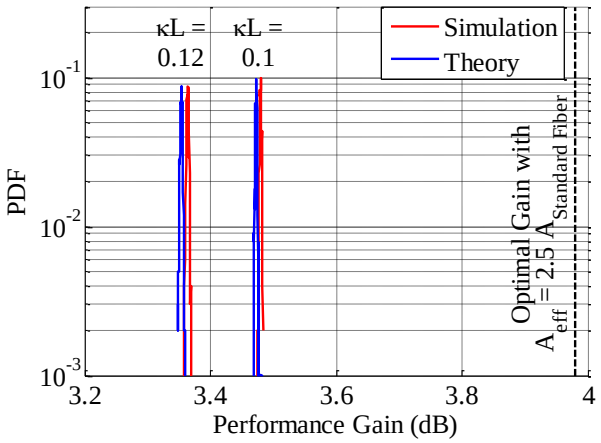


Fig. 4.6. PDF of the overall performance gain $G_{overall}$ of a 20-span QSM transmission system in presence of mode-coupling-induced penalties Λ_{MPI} and Λ_{HOM} and an effective area $A_{eff,QSM}$ 2.5 times that of a standard single-mode fiber. The span length is 100 km, DMD $\Delta\tau$ = 1.28 ns/km and the DML $\Delta\alpha$ = 0.5 dB/km.

Fig. 4.6 shows the pdf of overall performance gain $G_{overall}$ of a 20-span QSM transmission system in presence of mode-coupling-induced penalties Λ_{MPI} and Λ_{HOM} and an effective area $A_{eff,FMF}$ 2.5 times that of a standard single mode fiber for different coupling strengths. Fiber nonlinearity effects are not included in the simulation as we want to highlight the effects of MPI and higher-order mode stripping loss and nonlinearity should induce the same amount of OSNR penalty for both QSM and standard fiber systems at their respective optimal signal launched powers. It can be seen that, when MPI is appropriately compensated, the higher-order mode loss Λ_{HOM} plays a larger role than residual MPI-induced penalties in limiting the overall performance gain. One can define an appropriate threshold, say 1 dB from optimal performance, to statistically characterize acceptable level of performance and outage probabilities of QSM transmission systems.

4.4. Experiments

The experiments are also conducted to investigate the performance of QSM transmissions with MPI compensation in practical settings. The FMF used for QSM transmission in our experiments is a step-index silica core fiber with core radius around 9 μm and supports 6 polarization modes. The nominal effective area for the fundamental

mode is 220 μm^2. The DMD is measured to be around 1.12 ns/km. The higher order modes suffers a higher loss compared with the fundamental mode. The average measured attenuation of two 100 km spans constructed is about 19.1 dB including splices and connectors. At the beginning and end of each span, the FMF is connected with standard single mode fiber jumpers through a bridge fiber to ensure compatibility with single-mode components such as EDFAs and transmitters/receivers. The total loss on the input side including connector loss, splice from single-mode to bridge fiber, and splice from bridge fiber to FMF is about 0.6 dB. At the output of each span, the same level of connection and splicing loss is expected, in addition to a loss rising from the mode stripping. The average fiber loss is approximately in the range of approximately 17.5 dB and the rest is due to splices and connectors. Since the fibers are drawn from two different blanks, measurements of the accumulated MPI or crosstalk level of each of the two 100 km spans shows significantly different behavior with estimated levels of about −24 dB for one span and −30 dB or less for the second span.

The two 100-km spans of FMF are placed in a re-circulating loop configuration for transmission measurements. The experimental setup is shown in Fig. 4.7. Ten lasers spaced by 50 GHz are modulated together in the 16-QAM format at a symbol rate of 32 Gbaud, and then polarization-multiplexed to produce 256 Gb/s PM-16-QAM channels. The channel under measurement is encoded on an external cavity laser (ECL) while the other 9 channels are encoded on DFB lasers. The channels are passed through a short piece of positive dispersion fiber to de-correlate adjacent channels by a few

symbols before being launched in the loop. EDFAs are used to amplify the signals at the output of each FMF span.

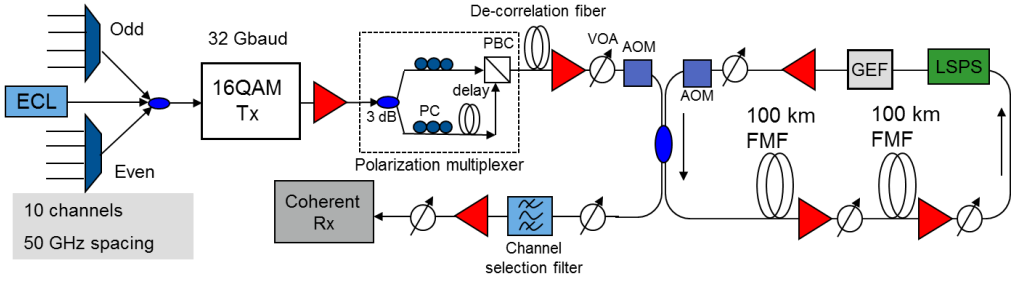


Fig. 4.7. Experimental setup for QSM transmission and MPI compensation. GEF: gain equalization filter; LSPS: loop synchronous polarization scrambler.

The receiver DSP algorithms consist of orthogonalization and normalization followed by frequency-domain CD compensation. The received data is then re-sampled to 2 samples per symbol for timing phase recovery [70]. CMA and CMMA on a 21-tap 2×2 MIMO filters are used as a two-stage algorithm for polarization de-multiplexing and PMD compensation. The signals are then down-sampled to 1 sample per symbol followed by frequency offset estimation and CPE [71]-[72]. A joint-polarization DD-LMS filters with maximum 501 taps is inserted after CPE for MPI compensation. This is followed by symbol decisions and BER calculation.

The working principle of DD-LMS can be described as follows. At time slot n, let $\boldsymbol{u}_{x,n}$ and $\boldsymbol{u}_{y,n}$ denotes the input signal vector in the two polarization, $\boldsymbol{h}_{xx,n}$, $\boldsymbol{h}_{xy,n}$, $\boldsymbol{h}_{yx,n}$, $\boldsymbol{h}_{yy,n}$ denote the butterfly filter taps as in Fig. 1.10, $v_{x,n}$ and $v_{y,n}$ denote the output signal, the DD-LMS equalization can be expressed as

$$\begin{cases} v_{x,n} = \boldsymbol{h}_{xx,n}^{T}\boldsymbol{u}_{x,n} + \boldsymbol{h}_{yx,n}^{T}\boldsymbol{u}_{y,n} \\ v_{y,n} = \boldsymbol{h}_{xy,n}^{T}\boldsymbol{u}_{x,n} + \boldsymbol{h}_{yy,n}^{T}\boldsymbol{u}_{y,n} \end{cases}. \tag{4.42}$$

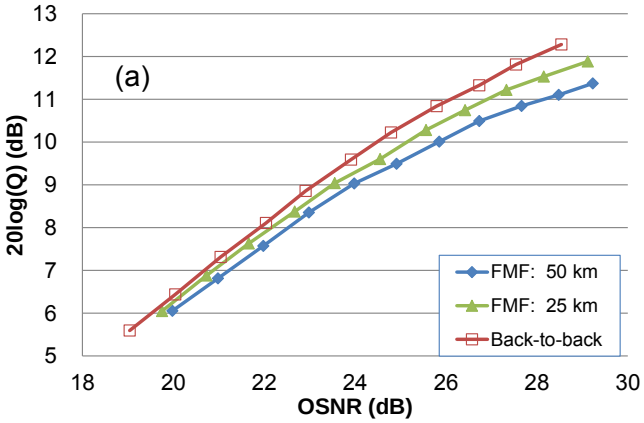
And at time slot $n + 1$, the filter taps is updated according to the decisions at time slot

n,

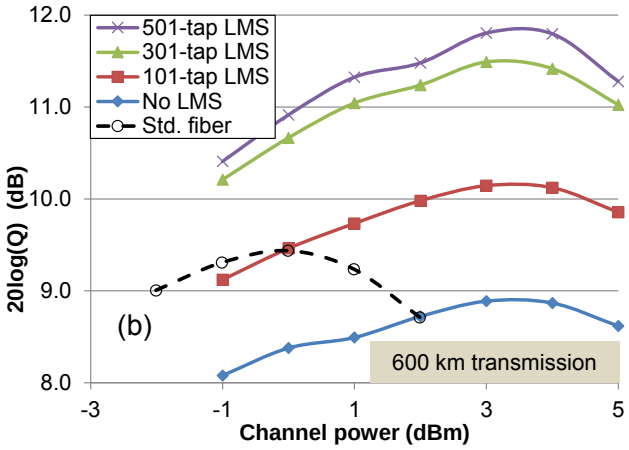
$$\boldsymbol{h}_{ab,n+1} = \boldsymbol{h}_{ab,n} + \mu(\hat{v}_{b,n} - v_{b,n}) \cdot \boldsymbol{u}_{a,n}^{*}, \tag{4.43}$$

where μ is the update step, a and b in the subscript can be either x or y. It should be noted that, in principle, such DD-LMS is not necessary in single-mode fiber systems where MPI is absent.

Fig. 4.8(a) shows the BER vs. OSNR for linear transmission for one of the FMF spans cut into different lengths without using DD-LMS for MPI compensation. These results are obtained for straight-line transmission through the two fiber lengths shown with ASE noise-loading before the receiver. The presence and distributed nature of MPI is evident from the fact that performance degrades with increasing FMF length. We then conducte QSM transmission over 600 km, or 3 loops, and vary the channel launched power to determine the optimal channel power. The measured Q values are shown in Fig. 4.8(b) along with results over a comparable 600-km standard single-mode fiber system. The optimal launched power for the QSM system is about 3-4 dB higher than that of the standard fiber system by virtue of the larger effective area. Without DD-LMS, the FMF system has a maximum Q value about 0.5 dB lower than the standard SMF system at this distance because of significant MPI. However, a 101-tap joint-polarization DD-LMS significantly improves QSM performance and a 501-tap DD-LMS enables QSM to outperform the standard fiber by more than 2 dB in Q. The convergence time for the DD-LMS is in the order of 50,000 symbols in most of our data sets.



(a)



(b)

Fig. 4.8. (a) *Q* vs. OSNR for segments of span no. 1 with different lengths in linear transmission. (b) *Q* vs. channel power of central channel for standard single-mode fiber and QSM transmission over 600 km using DD-LMS for MPI compensation.

It should be noted that the 501-tap DD-LMS does not fully compensate the MPI and further improvement, albeit marginal, can be obtained with increasing number of DD-LMS taps. However, as the delay between x and y polarization in our experimental setup is around 280 symbols, extending the number of taps beyond 560 taps results in double counting the signal and results in an under-estimate of BER. This is clearly evident in Fig. 4.9 for a single-span 50-km QSM system. Therefore, we choose the maximum number of taps to be 501 in our studies to ensure that our results are free

from signal double counting.

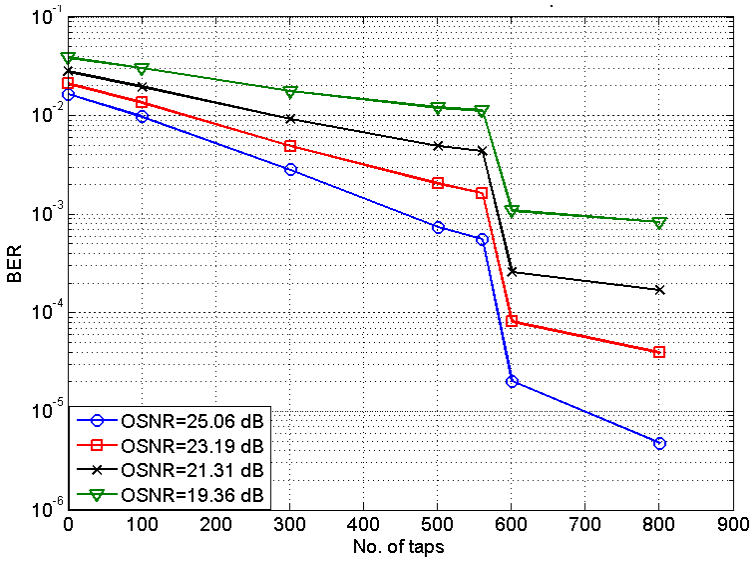


Fig. 4.9. BER vs. number of taps of the DD-LMS filter used to compensate MPI for a single-span 50-km QSM link in which the relative symbol delay between the x- and y-polarization is 280 symbols. When the number of taps exceed 560, the DD-LMS double counts the signal and result in an abrupt BER reduction.

With the channel power set at the optimum for both FMF and standard single-mode fiber systems, we then measured the *Q* of the central channel as a function of distance. The results are shown in Fig. 4.10. Without DD-LMS, QSM transmission over the FMF suffers from severe performance degradation due to MPI for all transmission distances. The detrimental effects of MPI offset the benefits of reduced nonlinearity so that the QSM transmission is inferior to standard fiber transmission. However, a joint-polarization 101-tap DD-LMS increases the *Q* considerably so that it becomes slightly better than that of the standard fiber system. The *Q* continues to improve with increasing number of taps until around 501 taps when only marginal improvement is obtained. This may imply that most of the MPI has been compensated with 301-taps DD-LMS.

Furthermore, the transmission distance for a soft decision-forward error correction (SD-FEC) threshold of 2E-2 BER (or *Q* of 6.25 dB) has been extended from 1200 km to 2600 km with the help of DD-LMS. It is worth noting that at 2600 km, the calculated total accumulated MPI for the QSM transmission would be greater than −12 dB. In comparison, the standard single-mode fiber system has a maximum reach of about 1500 km and the same strength DD-LMS brings negligible performance improvement, thus verifying that the DD-LMS is in fact compensating MPI in QSM systems.

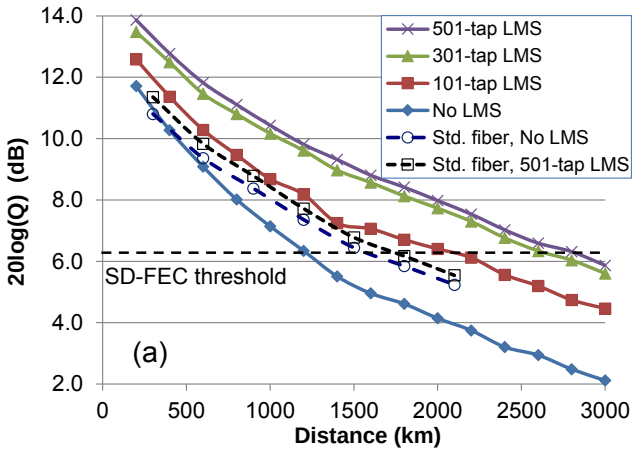
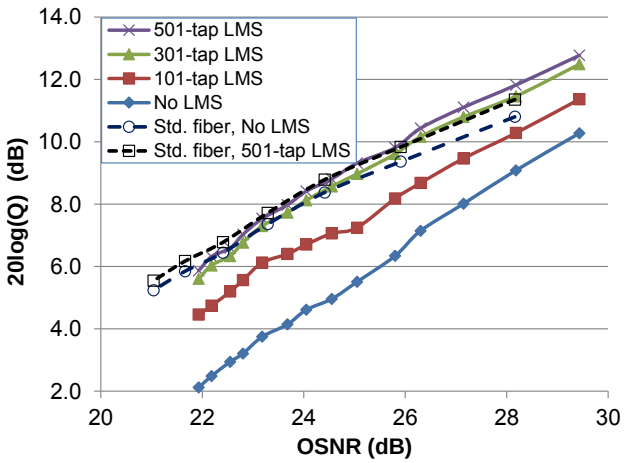


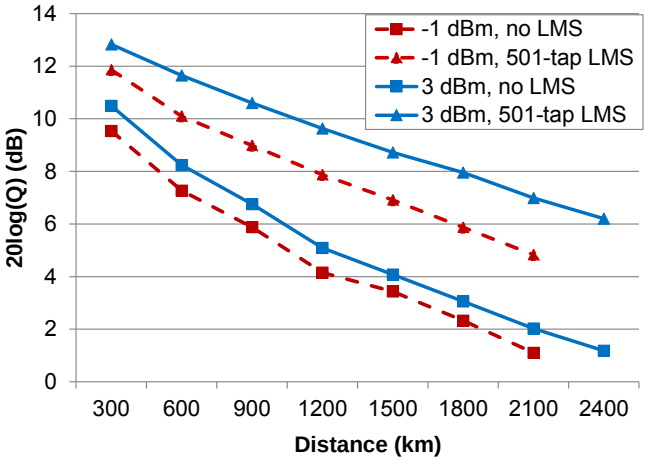
Fig. 4.10. *Q* vs. distance at optimal launch power using DD-LMS for MPI compensation.

To further quantify the effectiveness of DD-LMS in MPI compensation, we show the *Q* value of the central channel as a function of measured OSNR at the receiver in Fig. 4.11(a) for both QSM and standard fiber systems at their respective optimal signal launched powers. Note that as the optimal signal launched power is higher for QSM than standard fiber systems due to its reduced fiber nonlinearity, same OSNR actually means a higher signal launched power for QSM system with proportionately more noise, which in turn means the QSM link is longer compared to the corresponding standard

fiber system with the same OSNR. The close agreements between QSM with 301- and 501-taps DD-LMS and the standard fiber suggests that the MPI in the FMF is nearly fully compensated. The improvement of the DD-LMS filter for standard fiber systems is minimal, indicating that the transceiver circuitry and components are more or less optimized and there are negligible residual impairments and the DD-LMS are solely compensating MPI effects. Furthermore, Fig. 4.11(b) shows the *Q* factor vs. distance at −1 and 3 dBm signal launched power and the DD-LMS provides considerable performance improvements in both cases, illustrating that the DD-LMS is not compensating fiber nonlinearity.



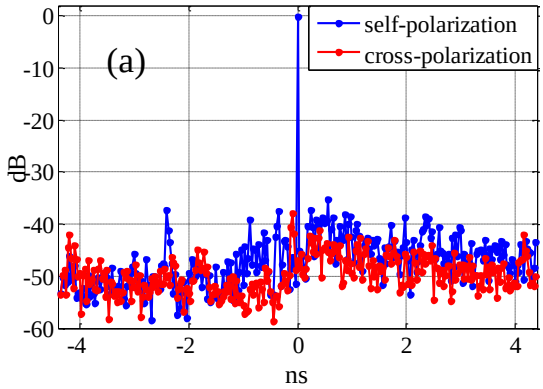
(a)



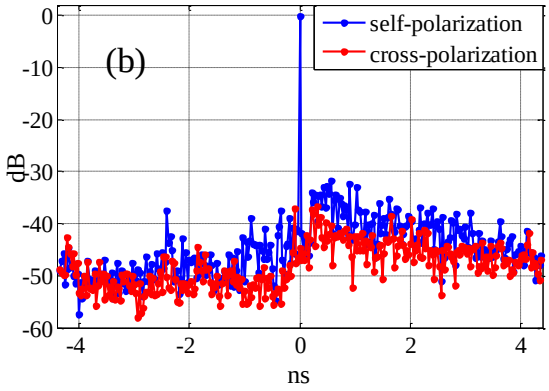
(b)

Fig. 4.11. (a) *Q* vs. OSNR at optimal signal launched power with DD-LMS for QSM and standard fiber systems and (b) *Q* vs. distance at −1 and 3 dBm signal launched power with and without DD-LMS filter.

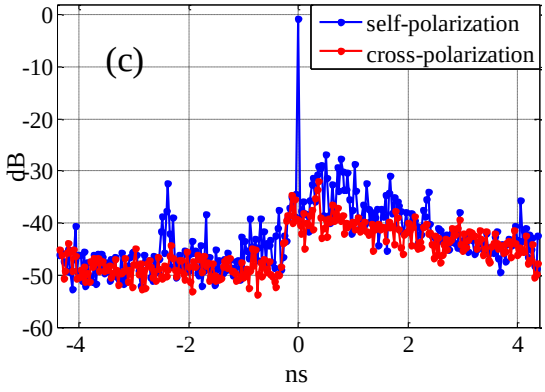
The DD-LMS taps after convergence allow us to recover/estimate the impulse response of MPI impairment. Alternatively, applying the method described in Ryf et al.[73], the estimated magnitude squared of the self-polarization impulse responses $|h_{xx}(t)|^2$ and $|h_{yy}(t)|^2$ and cross-polarization impulse responses $|h_{xy}(t)|^2$ and $|h_{yx}(t)|^2$ are depicted in Fig. 4.12 for various transmission distances. The impulse responses shown are obtained by averaging estimates over 5 independent data sets. It is clear that the MPI in QSM systems are distributed in nature as its strength grows with fiber length. Also, self-polarization MPI is more dominant than cross-polarization MPI. It should be noted that the unusual spike at the $t < 0$ region is merely an artifact of our 16-QAM generation method based on cascading a QPSK signal with its delayed and scaled version. Such configuration has no effect on the overall 16-QAM signal processing and BER performance



(a)



(b)



(c)

Fig. 4.12. Estimated magnitude squared of self-polarization and cross polarization impulse responses for (a) 600, (b) 1200 and (c) 2400 km of QSM transmission in the presence of MPI.

4.5. Summary

In this chapter, we analytically investigated polarization-multiplexed long-haul QSM transmission systems using 6-spatial-polarization-mode FMF parameterized by the differential mode delay and differential mode loss between the fundamental mode group and higher-order mode group. In the weak coupling regime, a multi-span QSM channel can be modeled as a Gaussian random process in frequency. Mode coupling in FMF

result in MPI and additional signal loss at each span of a QSM system. We derived appropriate filters for MPI compensation and showed that they induce minimal noise enhancement and performance penalty. In addition, we experimentally demonstrated 256 Gb/s PM-16-QAM QSM transmission over a record distance of 2600 km for 100-km spans by using a DD-LMS filter to compensate MPI. We showed that QSM transmissions can be a practical and viable option for transmission reach and capacities upgrade for future long-haul systems.

Chapter 5 Conclusions

The dissertation focused on channel characterization and signal transmission in various types of long-haul optical transmission systems from non-coherent system with single polarization transmission, coherent transmission in SMF and future QSM systems. The monitoring/estimation/compensation techniques of the impairments such as OSNR, CD and modal coupling caused MPI are investigated in detail.

For single-polarization system with direct detection, OSNR monitoring using polarization diversity and DSP is discussed in chapter 2. Analytical derivations are presented in exploring the relation among the PBS angle (polarization state of PBS), electrical bandwidth and monitoring results in the presence of first order PMD. It is shown that the polarization state of PBS can be classified in three groups. The first group comprises of a pair of orthogonal polarization states in which the monitored OSNR is accurate independent of receiver electrical bandwidth. The second group of polarization states lies on an equator in the Poincaré sphere except the two points in the first group. The ONSR can be correctly monitored in the second group given that the product of electrical bandwidth and DGD is small enough. And finally all the rest polarization states are in the third group where the OSNR is always under-estimate. Based on the above findings, we proposed to use a polarization scrambler to search for the first group of polarization states while the actual OSNR corresponds to the largest monitoring result. The proposed technique is verified in both simulations and experiments for OOK and DPSK systems with 1 dB error.

The third chapter is engaged in fast and PMD, phase-noise insensitive CD estimation in coherent systems. Theoretical analysis reveals that the auto-correlation of the received signal power contains a peak whose location is proportional to the accumulated CD, which enables a search-free CD estimation technique. Simulation results verify the proposed technique with a maximum estimation error of 175 ps/nm. Due to its search-free nature, the number of required symbols is only 8192, much less than that of the other search based techniques.

The fourth chapter discusses the channel characterization and signal transmissions in QSM systems. The FMF has larger effective area and allows higher launched power and thus increases OSNR. However, it brings about modal coupling and causes MPI and additional loss. In order to evaluate the aggregate performance improvement, we theoretically analyze the OSNR degradation in terms of the modal coupling strength, modal delay and modal loss with MMSE equalization in the coherent receiver compensating for MPI. Simulation results show that the in- and cross-polarization transfer functions can be fairly modeled as Rayleigh and Gaussian channel. For the first time, we experimentally demonstrate 256 Gb/s PM-16-QAM QSM transmission over a record distance of 2600 km, which doubles the length in the SMF system with the same transponder configuration. The MPI in the link is compensated by 501-taps one-sample-per-symbol DD-LMS after carrier phase recovery.

Chapter 6 Future Perspective

The current challenge in OSNR monitoring is to achieve it in polarization- multiplexed coherent systems with various advanced modulation formats. Moreover, researchers begin to investigate Nyquist or Faster-than Nyquist (FTN) systems in recent years in which the characterizations of the signal in both time and frequency domain are noise-like, especially in the presence of various inline impairments including huge amount of residual CD. The basis of OSNR monitoring is to utilize the characterizations of the signal to segregate or identify it from the noise, which would become unrealistic for the FTN scenario without any prior knowledge of the transmitted symbols or any modification of the transmitter. Hence in the future, data-aided OSNR monitoring technique may be needed for inline cost-effective OSNR monitoring.

For CD estimation in coherent receivers, the only drawback of the proposed technique is the performance degradation in the Nyquist pulse shaping system, especially with a low roll-off factor, where the correlation peak may vanish. Accordingly, we proposed to regenerate the correlation peak with a high-pass filter before calculating the auto-correlation in the expense of slightly more symbols needed and larger estimation error. Future research prospects may performance improvement in Nyquist systems and exploiting new techniques in FTN systems.

For QSM transmission, the nonlinearity in the link is left uncompensated in the experiments. It is known that nonlinear compensation such back propagation can bring about some performance improvement for SMF system. However its effectiveness in

the QSM systems is left to be investigated in the future. Moreover, it is know that the interplay between the phase noise of local oscillator and CD produces equalization enhanced phase noise in SMF system. And it would be reasonable to infer that such degradation would have bigger impact in QSM system due to the presence of modal coupling, which would be worth investigating in future. Also, higher-order modulation formats transmission, OFDM as well as super-channel QSM transmissions would be valuable to demonstrate in the future.

Appendix

$$
\begin{aligned}
&\left\langle \left(I_x(t) - \frac{\langle I_x(t)\rangle - z/2}{\langle I_y(t)\rangle - z/2} I_y(t) \right)^2 \right\rangle \\
&= \left\langle \left[I_{x,s}(t) - \frac{rP_s + (P_n - z)/2}{(1-r)(1-r)P_s + (P_n - z)/2} I_{y,s}(t) + I_{x,sn}(t) \right.\right. \\
&\quad \left.\left. - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} I_{y,sn}(t) + I_{x,n}(t) - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} I_{y,n}(t) \right]^2 \right\rangle \\
&= \mathbf{Var}\left[I_{x,s}(t) - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} I_{y,s}(t) + I_{x,sn}(t) \right. \\
&\quad \left. - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} I_{y,sn}(t) + I_{x,n}(t) - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} I_{y,n}(t) \right] \\
&\quad + \left[\langle I_{x,s}(t)\rangle - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} \langle I_{y,s}(t)\rangle + \langle I_{x,sn}(t)\rangle \right. \\
&\quad \left. - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} \langle I_{y,sn}(t)\rangle + \langle I_{x,n}(t)\rangle - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} \langle I_{y,n}(t)\rangle \right]^2 \\
&= \mathbf{Var}\left[I_{x,s}(t) - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} I_{y,s}(t) \right] + \left(\mathbf{Var}\left[I_{x,sn}(t) \right] \right. \\
&\quad \left. + \left[\frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} \right]^2 \mathbf{Var}\left[I_{y,sn}(t) \right] \right) + \left(\mathbf{Var}\left[I_{x,n}(t) \right] \right. \\
&\quad \left. + \left[\frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} \right]^2 \mathbf{Var}\left[I_{y,n}(t) \right] \right) + \left[\frac{(1-2r)P_s z/2}{(1-r)P_s + (P_n - z)/2} \right]^2 \\
&= \mathbf{Var}\left[I_{x,s}(t) - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} I_{y,s}(t) \right] \\
&\quad + \frac{2B_e}{B_o} \left[r + \left(\frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} \right)^2 (1-r) \right] P_s P_n \\
&\quad + \frac{2B_e}{B_o} \cdot \frac{1}{4} \left[1 + \left(\frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} \right)^2 \right] P_n^2 + \left[\frac{(1-2r)P_s z/2}{(1-r)P_s + (P_n - z)/2} \right]^2
\end{aligned}
$$

$$
\begin{aligned}
& f(z) \\
&= \mathbf{Var}\left[I_{x,s}(t) - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} I_{y,s}(t) \right] \\
&\quad + \frac{2B_e}{B_o}\left[r + \left(\frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} \right)^2 (1-r) \right] P_s P_n \\
&\quad + \frac{2B_e}{B_o} \cdot \frac{1}{4}\left[1 + \left(\frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} \right)^2 \right] P_n^2 + \left[\frac{(1-2r)P_s z/2}{(1-r)P_s + (P_n - z)/2} \right]^2 \\
&\quad - \frac{1}{4}\left[\frac{(1-2r)P_s}{(1-r)P_s + (P_n - z)/2} \right]^2 z^2 - \frac{2B_e}{B_o} \cdot \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} \cdot z(P_s + P_n - z) \\
&\quad - \frac{2B_e}{B_o} \cdot \frac{1}{4}\left[1 + \left(\frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} \right)^2 \right] z^2 \\
&= \mathbf{Var}\left[I_{x,s}(t) - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} I_{y,s}(t) \right] \\
&\quad + \frac{2B_e}{B_o}\left\{ \left[\frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} \right]^2 \left[(1-r)P_s P_n + \frac{1}{4}P_n^2 - \frac{1}{4}z^2 \right] \right. \\
&\quad \left. - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} \cdot z(P_s + P_n - z) + \left(rP_s P_n + \frac{1}{4}P_n^2 - \frac{1}{4}z^2 \right) \right\} \\
&= \mathbf{Var}\left[I_{x,s}(t) - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} I_{y,s}(t) \right] + \frac{2B_e}{B_o}\left\{ \left[\frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} \right]^2 \right. \\
&\quad \cdot \left[(1-r)P_s P_n + \frac{1}{4}P_n^2 - \frac{1}{4}z^2 \right] - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} \cdot z\left(P_s + P_n - z\right) \\
&\quad \left. + \left[rP_s + (P_n - z)/2 \right]^2 + z\left[rP_s + (P_n - z)/2 \right] - r^2 P_s^2 \right\} \\
&= \mathbf{Var}\left[I_{x,s}(t) - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} I_{y,s}(t) \right] + \frac{2B_e}{B_o}\left\{ \left[\frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} \right]^2 \right. \\
&\quad \cdot \left[(1-r)P_s P_n + \frac{1}{4}P_n^2 - \frac{1}{4}z^2 \right] - \frac{\left[rP_s + (P_n - z)/2 \right]^2}{(1-r)P_s + (P_n - z)/2} \cdot z \\
&\quad \left. + \left[rP_s + (P_n - z)/2 \right]^2 - r^2 P_s^2 \right\} \\
&= \mathbf{Var}\left[I_{x,s}(t) - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} I_{y,s}(t) \right] + \frac{2B_e}{B_o}\left\{ \left[\frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} \right]^2 \right. \\
&\quad \left. \cdot \left[\left((1-r)P_s + (P_n - z)/2 \right)^2 - (1-r)^2 P_s^2 \right] + \left[rP_s + (P_n - z)/2 \right]^2 - r^2 P_s^2 \right\} \\
&= \underbrace{\mathbf{Var}\left[I_{x,s}(t) - \frac{rP_s + (P_n - z)/2}{(1-r)P_s + (P_n - z)/2} I_{y,s}(t) \right]}_{f_1(z)} \\
&\quad + \underbrace{\frac{2B_e}{B_o}\left\{ \left[rP_s + (P_n - z)/2 \right]^2 \left[2 - \frac{(1-r)^2 P_s^2}{\left[(1-r)P_s + (P_n - z)/2 \right]^2} \right] - r^2 P_s^2 \right\}}_{f_2(z)}
\end{aligned}
$$

Reference

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

[2] 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, Dce. 2006.

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

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

[5] S. J. Savory, "Digital signal processing for coherent systems," in *Proc. Opt. Fiber Commun. Conf. (OFC/NFOEC) 2012*, Los Angeles, CA, Paper OTh3C.7.

[6] A. P. T. Lau, Y. Gao, Q. Sui, D. Wang, Q. Zhuge, M. H. M.-Osman, M. Chagnon, X. Xu, C. Lu and D. V. Plant, "Advanced DSP Techniques Enabling High Spectral Efficiency and Flexible Transmissions," *IEEE Signal Processing Magazine*, vol. 31, no. 2, pp. 82-92, Mar. 2014.

[7] A. P. T. Lau, Y. Gao, Q. Sui, D. Wang, Q. Zhuge, M. M. Osman, M. Chagnon, X. Xu, C. Lu and D. V. Plant, "Beyond 100 Gb/s: Advanced DSP Techniques Enabling High Spectral Efficiency and Flexible Optical Communications," in *International Conference on Optical Commun. and Networks (ICOCN) 2013*, paper SC1In9.

[8] A. P. T. Lau, Y. Gao, **Q. Sui**, D. Wang and C. Lu, "Advanced DSP for High Spectral Efficiency and Flexible Optical Communications," in *Signal Processing in Photonic Commun. 2014*, paper SM2D.1.

[9] A. P. T. Lau , Y. Gao, **Q. Sui**, D. Wang, C. Lu, “Advanced DSP Techniques Enabling Flexible Transmissions and Elastic Optical Networks,” in *International Photonics and OptoElectronics Meetings (POEM) 2014*, paper OTh3B.3.

[10] G. P. Agrawal, *Lightwave Technology: Telecommunication Systems*. New York: Wiley, 2005.

[11] G. P. Agrawal, *Nonlinear Fiber Optics*, 4th ed. San Diego, CA: Elsevier, 2006.

[12] A. Galtarossa and C. R. Menyuk, Eds., *Polarization Mode Dispersion*. New York: Springer, 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, Feb. 2010.

[14] A. D. Ellis, J. Zhao D. Cotter, “Approaching the non-linear Shannon limit,” *J. Lightw. Technol*., vol. 28, no. 4, pp. 423–433, Feb. 2010.

[15] A. Mecozzi and R. J. Essiambre, “Nonlinear shannon limit in pseudo-linear coherent systems,” *J. Lightw. Technol*., vol. 30, no. 12, pp. 2010–2024, 2012.

[16] C. L. Chen, *Foundations for guide-wave optics*. New Jersey: Wiley-interscience, 2007.

[17] D. C. Kilper, R. Bach, D. J. Blumenthal, D. Einstein, T. Landolsi, L. Ostar, M. Preiss and A. E.Willner, “Optical performance monitoring,” *J. Lightw. Technol*., vol. 22, no. 1, pp. 294–304, Jan. 2004.

[18] D. C. Kilper, S. Chandrasekhar, L. Buhl, A. Agarwal and D. Maywar, “Spectral monitoring of OSNR in high speed networks,” in *Europ. Conf. Optical Commun. (ECOC) 2002*, Paper 7.4.4.

[19] D. K. Jung, C. H. Kim, and Y. C. Chung, “OSNR monitoring technique using polarization-nulling method,” in *Proc. Opt. Fiber Commun. Conf. (OFC/NFOEC)*

2005, Baltimore, MD, Paper WK4.

[20] J. H. Lee and Y. C. Chung, "OSNR monitoring technique using polarization nulling method," *IEEE Photon. Technol. Lett*., vol. 13, no. 1, pp. 88–90, Jan. 2001.

[21] J. H. Lee, H. Y. Choi, S. K. Shin, and Y. C. Chung, "A review of the polarization-nulling technique for monitoring optical-signal-to-noise ratio in dynamic WDM networks," *J. Lightw. Technol.*, vol. 24, No. 11, pp. 4162–4171, Nov. 2006.

[22] J. H. Lee and Y. C. Chung, "Improved OSNR monitoring technique based on polarization-nulling method," *Electronics Letters*, vol. 37, No. 15, pp. 972–973, Jul. 2001.

[23] M. Skold, B.-E. Olsson, H. Sunnerud and M. Karlsson, "PMD-insensitive DOP-based OSNR monitoring by spectral SOP measurements," in *Proc. Opt. Fiber Commun. Conf. (OFC/NFOEC) 2005*, Anaheim, CA, Paper OThH3.

[24] C. Xie, D. C. Kilper, L. Moller and R. Ryf, "Orthogonal-polarization heterodyne OSNR monitoring insensitive to polarization-mode dispersion and nonlinear polarization scattering," *J. Lightw. Technol*., vol. 25, no. 1, pp. 177–183, Jan. 2007.

[25] M. H. Cheung, L. K. Chen and C. K. Chan, "PMD-insensitive OSNR monitoring based on polarization-nulling with off-center narrow-band filtering," *IEEE Photon. Technol. Lett*., vol. 16, no. 11, pp. 2562–2564, Nov. 2004.

[26] S. Dod and T. B. Anderson, "In-band optical signal-to-noise monitoring technique using polarization diversity," in *Proc. LEOS 2005*, pp. 700-701, 2005.

[27] T. B. Anderson, K. Clarke, S. D. Dods and M. Bakaul, "Robust, low cost, in-band optical signal to noise monitoring using polarization diversity," in *Proc. Opt. Fiber Commun. Conf. (OFC/NFOEC) 2007*, Anaheim, CA, Paper OMM3.

[28] Q. Sui, A. P. T. Lau and C. Lu, "OSNR monitoring in the presence of first-order PMD using polarization diversity and DSP," *IEEE Journal of Lightwave*

Technology, vol. 28, no. 15, pp. 2105-2114, Aug. 2010.

[29] Q. Sui, C. Lu and A. P. T. Lau, "In-Band OSNR Monitoring by Polarization Diversity and Electronic Signal Processing," in *Proc. OECC 2009*, Hong Kong, Paper FW5.

[30] A. P. T. Lau, Q. Sui, T. S. R. Shen, F. N. Khan and C. Lu, "Optical Performance Monitoring Using Statistical Signal Processing," *International Conference on Advanced Infocomm. Technology (ICAIT)*, Haikou, Hainan Island, China, Jul. 2010.

[31] A. P. T. Lau, Q. Sui, T. S. R. Shen, F. N. Khan and C. Lu, "Optical Performance Monitoring Using Statistical Signal Processing," *International Conference on Optical Communications Networks (ICOCN)*, Nanjing, China, Oct. 2010.

[32] C. Lu, A. P. T. Lau, F. N. Khan, Q. Sui, J. Zhao, Z. Li, H. Y. Tam and P. K. A. Wai, "Optical performance monitoring techniques for high capacity optical networks," in *7th International Symposium on Communication Systems Networks and Digital Signal Processing (CSNDSP)*, pp. 678 – 681, July 2010.

[33] A. P. T. Lau, Q. Sui, T. S. R. Shen, F. N. Khan and C. Lu, "Optical Performance Monitoring for Next Generaton Optical Networks: Progress and Challenges," *International Conference on Advanced Infocomm. Technology (ICAIT)*, Wuhan, China, Jul. 2011.

[34] G. Rangarajan and M. Ding, *Processes with long-range correlations: theory and applications*, New York: Springer, 2003.

[35] F. N. Hauske, M. Kuschnerov, B. Spinnler and B. Lankl, "Optical performance monitoring in digital coherent receivers," *J. Lightw. Technol*., vol. 27, no. 16, pp. 3623–3630, Aug. 2009.

[36] C. C. Do, C. Zhu, A.V. Tran, S. Chen, T. Anderson, D. Hewitt and E. Skafidas, "Chromatic Dispersion Estimation in 40 Gb/s Coherent Polarization-Multiplexed

Sinigle Carrier System using Complementary Golay Sequences," in *Proc. Opt. Fiber Commun. Conf. (OFC/NFOEC) 2012*, paper OW4G.1.

[37] F. Pittala, F. N. Hauske, Y. Ye, N. G. Gonzalez and I. T. Monroy, "Combined CD and DGD Monitoring Based on Data-Aided Channel Estimation," in *Proc. OSA Signal Processing in Photonics Communications (SPPCom) 2011*, paper SPTuC3.

[38] R. A. Soriano, F. N. Hauske, N. G. Gonzalez and I. T. Monroy, "Chromatic dispersion estimation in digital coherent receivers," *J. Lightw. Technol*., vol. 29, no. 11, pp. 1627–1637, Jun. 2011.

[39] M. Kuschnerov, F. N. Hauske, K. Piyawanno, B. Spinnler, A. Napoli and B. Lankl, "Adaptive chromatic dispersion equalization for non-dispersion managed coherent systems," in *Proc. Opt. Fiber Commun. Conf. (OFC/NFOEC) 2009*, paper OMT1.

[40] F. N. Hauske, Z. Zhang, C. Li, C. Xie, and Q. Xiong, "Precise, robust and least complexity CD estimation," in *Proc. Opt. Fiber Commun. Conf. (OFC/NFOEC) 2011*, San Diego, CA, paper JWA032.

[41] D. Wang, C. Lu, A. P. T. Lau and S. He, "Adaptive Chromatic Dispersion Compensation for Coherent Communication Systems Using Delay-Tap Sampling Technique," *IEEE Photon Technol. Lett*., vol. 23, no. 14, pp. 1016–1018, July. 2011.

[42] R. Borkowski, X. Zhang, D. Zibar, R. Younce and I. T. Monroy, "Experimental demonstration of adaptive digital monitoring and compensation of chromatic dispersion for coherent DP-QPSK receiver," *Opt. Exp.*, vol. 19, no. 26, pp. 728–735, Dec. 2011.

[43] S. M. Ranzini, E. C. Magalh ães, V. B. Ribeiro, V. V. Nascimento and J. C. R. F. Oliveira, "Accurate Blind Chromatic Dispersion Estimation in Long-haul 112Gbit/s PM-QPSK WDM Coherent Systems" in *Proc. OSA Signal Processing in Photonics Communications (SPPCom) 2012*, paper SpTh2B.3.

[44] H. Wymeersch and P. Johannisson, “Maximum-Likelihood-Based Blind Dispersion Estimation for Coherent Optical Communication,” *J. Lightw. Technol*., vol. 30, no. 18, pp. 2976–2982, Sept. 2012.

[45] C. Malouin, P. Thomas, B. Zhang, J. O'Neil, T. Schmidt, “Natural Expression of the Best-Match Search Godard Clock-Tone Algorithm for Blind Chromatic Dispersion Estimation in Digital Coherent Receivers” in *Proc. OSA Signal Processing in Photonics Communications (SPPCom)* 2012, paper SpTh2B.4.

[46] C. Malouin, M. Arabaci, P. Thomas, B. Zhang, T. Schmidt and R. Marcoccia, “Efficient, Non-Data-Aided Chromatic Dispersion Estimation via Generalized, FFT-Based Sweep,” in *Proc. Opt. Fiber Commun. Conf. (OFC/NFOEC) 2013*, Anaheim, CA, paper JW2A.45.

[47] Q. Sui, A. P. T. Lau and C. Lu, “Fast and robust blind chromatic dispersion estimation using auto-correlation of signal power waveform for digital coherent systems,” *IEEE Journal of Lightwave Technology*, vol. 31, no. 2, pp. 306–312, Jan. 2013.

[48] Q. Sui, A. P. T. Lau and C. Lu, “Fast and Robust Chromatic Dispersion Estimation Using Auto-Correlation of Signal Power Waveform for DSP based-Coherent Systems,” in *Proc. Opt. Fiber Commun. Conf. (OFC/NFOEC) 2012*, paper OW4G.3.

[49] C. Lu, A. P. T. Lau, **Q. Sui**, F. Khan and Z. Dong, “Optical Performance Monitoring: from direct detection to coherent optical communication systems,” in *Commun. and Photonics Conference (ACP)*, 2012, Paper AF3G.3.

[50] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, Inc., 2001.

[51] J. Zhao, A. P. T. Lau, K. K. Qureshi, Z. Li, C. Lu and H. Y. Tam, “Chromatic dispersion monitoring for DPSK systems using RF power spectrum,” *J. Lightw.*

Technol., vol. 27, no. 24, pp. 5704–5709, Dec. 2009.

[52] F. C. Pereira, V. N. Rozental, M. Camera, G. Bruno, D. A. A. Mello, "Experimental Analysis of the Power Auto-Correlation-Based Chromatic Dispersion Estimation Method," *IEEE Photonics Journal*., vol. 5, no. 4, Aug. 2013

[53] 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, Oct. 2008.

[54] S. Randel, "Space Division Multiplexed Transmission," in *Proc. Opt. Fiber Commun. Conf. (OFC/NFOEC) 2013*, paper OW4F.

[55] D. J. Richardson, J. M. Fini and L. E. Nelson, "Space-division multiplexing in optical fibres," *Nat. Photon.*, vol. 7, pp. 354–362, Apr. 2013.

[56] L. W. Luo, N. Ophir, C. P. Chen, L. H. Gabrielli, C. B. Poitras, K. Bergmen and M. Lipson, "WDM-compatible mode-division multiplexing on a silicon chip," *Nat. Commun*., vol. 5, num. 3069, Jan. 2014.

[57] F. Yaman, N. Bai, B. Zhu, T. Wang and G. Li, "Long distance transmission in few-mode fibers," *Opt. Exp.*, vol. 18, no. 12, pp. 13250–13257, Jun. 2010.

[58] F. Yaman, N. Bai, Y. K. Huang, M. F. Huang, B. Zhu, T. Wang and G. Li, "10 x 112Gb/s PDM-QPSK transmission over 5032 km in few-mode fibers," *Opt. Exp.*, vol. 18, no. 20, pp. 21342–21349, Sep. 2010.

[59] J. D. Downie, J. E. Hurley, D. V. Kuksenkov, C. M. Lynn, A. E. Korolev and V. N. Nazarov, "Transmission of 112 Gb/s PM-QPSK signals over up to 635 km of multimode optical fiber," *Opt. Exp*., vol. 19, no. 26, pp. B363–B369, Nov. 2011.

[60] N. Bai, C. Xia and G. Li, "Adaptive frequency-domain equalization for the transmission of the fundamental mode in a few-mode fiber," *Opt. Exp.*, vol. 20, no. 21, pp. 24010–24017, Oct. 2012.

[61] Q. Sui, H. Zhang, J. D. Downie, W. A. Wood, J. Hurley, S. Mishra, A. P. T. Lau, C. Lu, H-Y. Tam and P. K. A. Wai, "256 Gb/s PM-16-QAM Quasi-Single-Mode Transmission over 2600 km using Few-Mode Fiber with Multi-Path Interference Compensation," in *Proc. Opt. Fiber Commun. Conf. (OFC/NFOEC) 2014*, paper M3C.5.

[62] Q. Sui, H. Zhang, J. D. Downie, W. A. Wood, J. Hurley, S. Mishra, A. P. T. Lau, C. Lu, H-Y Tam and P. K. A. Wai, "Long-haul Quasi-Single-Mode Transmissions using Few-Mode Fiber in presence of Multi-Path Interference," *Manuscript submitted for publication*.

[63] A. P. T. Lau, Q. Sui, H. Y. Tam, C. Lu and P. K. A. Wai, J. D. Downie, W. A. Wood, J. Hurley and S. Mishra, "Long-haul Quasi-Single-Mode Transmission using Few-Mode Fiber with Multi-Path Interference Compensation" in *Proc. of International Conference on Optical Internet (COIN) 2014*, paper FB3–1.

[64] J. D. Downie, J. Hurley, D. Pikula, S. Ten and C. Towery, "Study of EDFA and Raman system transmission reach with 256 Gb/s PM-16QAM signals over three optical fibers with 100 km spans," *Opt. Exp.*, vol. 21, no. 14, pp. 17372–17378, July. 2013.

[65] F. Yaman, E. Mateo and T. Wang, "Impact of Modal Crosstalk and Multi-Path Interference on Few-Mode Fiber Transmission," in *Proc. Opt. Fiber Commun. Conf. (OFC/NFOEC) 2012*, paper OTu1D.2.

[66] J. Vuong, P. Ramantanis, A. Seck, D. Bendimerad and Y. Frignac, "Understanding discrete linear mode coupling in few-mode fiber transmission systems," *Proceedings of ECOC 2011*, Paper Tu.5.B.2.

[67] D. Marcuse, C. R. Manyuk and P. K. A. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightw. Technol.*, vol. 15, no. 9, pp. 1735–1746, Sep. 1997.

[68] P. Wai and C. R. Menyuk, “Polarization mode dispersion, decorrelation and diffusion in optical fibers with randomly varying birefringence,” *J. Lightw. Technol.*, vol. 14, no. 2, pp. 148–157, Feb. 1996.

[69] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey and D. E. Knuth, “On the LambertW function,” *Advances in Computational Mathematics*, vol. 5, no. 1, pp. 329-359, 1996.

[70] F. M. Gardner, “A BPSK/QPSK Timing-Error Detector for Sampled Receivers,” *IEEE Trans. Commun.*, vol. 34, no. 5, pp. 423–429, May. 1986.

[71] A. Leven, N. Kaneda, U.-V. Koc, Y.-K. Chen “Frequency Estimation in Intradyne Reception,” *IEEE Photon. Technol. Lett.*, vol. 19, no. 6, pp. 366–368, March. 2007.

[72] X. Zhou, “An improved feed-forward carrier recovery algorithm for coherent receivers with M-QAM modulation format,” *IEEE Photon. Technol. Lett.*, vol. 22, no. 14, pp. 1051–1053, July. 2010.

[73] R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, A. Sierra, S. Mumtaz, M. Esmaeelpour, E. C. Burrows, R-J. Essiambre, P. J. Winzer, D. W. Peckham, A. H. McCurdy and R. Lingle, “Mode-division multiplexing over 96km of few-mode fiber using coherent 6 × 6 MIMO processing,” *J. Lightw. Technol.*, vol. 30, no. 4, pp. 521–531, Feb. 2012.