Advanced Transmission Technologies for High-Capacity Optical Networks

by

Qian Hu

Submitted in total fulfilment of

The requirements of the degree of

Doctor of Philosophy

Department of Electrical and Electronic Engineering

The University of Melbourne

VIC 3010, Melbourne, Australia

August 2016

Printed on archival quality paper.

Abstract

Advanced Transmission Technologies for High-Capacity Optical Networks

by **Qian Hu**

In order to keep up with the exponential growth of Internet traffic, there emerges an urgent demand for high-capacity optical networks. The optical networks can be classified into two categories, the long-haul and the short-reach optical networks, in consideration of their different requirements in transmission distance, data rate, and construction cost. Great challenges exist in both long-haul and short-reach optical networks when upgrading the capacities.

Various multiplexing techniques have been proposed to explore the available modulation freedoms in optical fibers, leading to a constant capacity increase in long-haul optical networks over the past few decades. A significant capacity increase has been witnessed during the last decade with the revival of the coherent optical technique, which makes full utilization of the available degrees of modulation freedoms in single-mode fiber (SMF). With the development of the coherent technique, the capacity of the long-haul optical networks rapidly approaches the Shannon limit, which imposes an upper limit for the channel capacity in SMF. To support a sustainable capacity increase, the long-haul optical networks are faced with the challenge to achieve a capacity beyond the Shannon limit in SMF.

On the contrary, the traffic is relatively low in the short-reach optical networks, where the service is provided for fewer users. The system cost shared by each user is dramatically increased, making the cost a primary consideration. A cost-efficient solution can be readily obtained using direct modulation and direct detection (DM/DD) system. However, the data rate and transmission distance of such system are limited due to the signal distortion resulting from the nonlinear channel. To meet with the increasing traffic demand between the data centers, the short-reach optical networks are confronted with the challenge to improve the capacity while maintaining the system cost at an acceptable level.

In this thesis, we explore the promising techniques for high-capacity optical networks with various transmission distances, from short-reach to long-haul. For the long-haul optical networks, a channel capacity beyond the Shannon limit of SMF can be achieved with the help of few-mode fiber (FMF), where the signal is multiplexed in different spatial modes. The multiple-input multiple-output (MIMO) digital signal processing (DSP) is necessary to recover the signal after the random mode coupling. The channel matrix plays a critical role in the DSP, describing the linear impact from the FMF channel. By investigating into the properties of the FMF channel matrix, we obtain the essential knowledge to design a FMF transmission system with optimal performance and efficient DSP, providing a feasible way towards the ultra-high-speed optical transmission.

For short-reach optical networks, various advance modulation and detection schemes have been proposed to improve the channel capacity with low-cost transceivers. In this thesis, we will report the recent progress on high-performance short-reach optical networks, where a linear channel is obtained to reach a capacity beyond that of simple DM/DD system. We will introduce serveral advanced modulation formats, which can linearize the channel in direct detection (DD) systems using self-coherent approach. The emphasis will be placed on the Stoke Vector Direct Detection (SV-DD) scheme where superior electrical spectral efficiency can be achieved. After that, we will investigate into two advanced detection schemes for high-performance direct modulation (DM) system. By utilizing the phase modulation from the directly modulated laser (DML) due to the frequency chirp effect, the receiver sensitivity can be significantly improved, leading to a greater potential of high-capacity transmission.

This is to certify that

(i) the thesis comprises only my original work,

(ii) due acknowledgement has been made in the text to all other material used,

(iii) the thesis is less than 100,000 words in length, exclusive of table, maps, bibliographies, appendices and footnotes.

Signature_____

Date_____

Declaration

I hereby declare that this thesis comprises only my original work. No material in this thesis has been previously published and written by another person, except where due reference is made in the text of the thesis. I further declare that this thesis contains no material which has been submitted for a degree or diploma or other qualifications at any other university. Finally, I declare that the thesis is less than 100,000 words in length, exclusive of tables, figures, bibliographies, appendices and footnotes.

Acknowledgements

This thesis is a summary of my academic work during my Ph.D. candidature at the University of Melbourne, which cannot be finished without the helps and supports from so many incredible individuals.

First and foremost, I would like to give my sincerest gratitude to my supervisor, Prof. William Shieh, for his high-quality supervision, patient guidance, constant encouragement and generous support during my entire Ph.D. candidature. His profound thinking and technical insights have inspired me and helped me develop deep understanding on my research topic.

In addition, I would like to express my sincere gratitude to Dr. Fred Buchali at Bell Labs, Alcatel-Lucent Germany, for providing me an internship opportunity. It was an invaluable experience for me to learn from him and other experts at Bell Labs.

My sincere gratitude also goes to my Ph.D. advisory committee members, Prof. Marcus Nathan Brazil and Dr. Alan Lee, for their stimulating discussion and precious advice. Moreover, I am thankful to the University of Melbourne and Department of Electrical & Electronic Engineering for offering me the scholarship, as well as the outstanding research environment and facilities.

Also, I would like to thank my colleagues and friends who have helped me during the past: An Li, Xi Chen, Jiayuan He, Hamid Khodakarami, Yifei Wang, Di Che, Jian Fang, Liaquat Ali, Feng Yuan, Miao Sun, Ke Wang, Chen Zhu, and Meng Wang. I especially appreciate their assistance and encouragement during the time I experienced difficulties.

Last but not the least, I would like to express my greatest gratitude to my loving parents, Ping Hu and Na Li, for their unconditional love, sacrifice and endless support. They are the emotional anchor for me. I would not be the same without their unwavering support.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Innovation of optical fiber and optical communication

The invention of telegraph by Samuel Morse in 1838 started the era of modern information communication. Since then, the communication demand has been constantly increasing. Different forms of communication systems have been developed and installed all over the world to provide better communication service for the people. A serial of evolutions in transmission medium has been witnessed during the development of communication systems: from single wireline used in the earliest telegraphy, to twisted pair wires used in the early stage telephony, then to coaxial cables used in the multi-channel telephony. Each evolution in transmission medium has led to a significant improvement in the transmission capacity.

After the concept of optical transmission was proposed, the scientists had been searching for the best optical transmission medium for a long time [1]. Back in 1920s, un-cladded glass fiber was successful fabricated. Since then, the optical signal is transmitted via the fiber channel [2-4]. Later on, the fiber characteristics were significantly improved by adding a cladding layer surrounding the core. Since the cladding layer has lower refractive index compared to the core. the transmitted optical signal is well-constrained within the core area under the principle of so-called total internal reflection due to the index difference between the core and cladding layer. The concept of optical field was first introduced into the fiber design in 1950s [5-8]. After that, the optical field of fibers has been intensively investigated, providing the essential knowledge to design high-performance fibers.

The high transmission loss was the fundamental factor that limited the application of optics fiber in the early days. The low-loss silica fiber was first successfully fabricated in 1970 based an earlier idea from Kao, reducing the loss of the silica fiber to below 20 dB/km [9]. After that, a great progress had been made in manufacturing low-loss silica fiber. In 1979, the fiber fabrication technique was able to reduce the transmission loss as low as 0.2 dB/km in the 1.55 μm wavelength window [10,11], corresponding to the C- or L-band in the modern optical communication. Given the advantages of small attenuation over long distance and immunity to electromagnetic interference, optical fibers are widely used transmission medium in modern communication systems.

Together with the optical fiber, the inventions of laser sources, optical amplifiers and other optical modulation and detection components have led to the popularity of optical communications [12]. With the development of optical transmission technologies, the optical networks have witnessed a huge capacity increase during the past few decades. In the meantime, the traffic demand has also been greatly increased due to the prevalence of the internet connection all over the world. To fulfill the increasing traffic demand, there is great interest to explore the promising techniques for the next-generation high-capacity optical networks.

Significant amount of effort has been devoted to develop the advanced techniques, enabling the optical transmission with higher capacity [13]. To increase the channel capacity within a limited bandwidth, various advanced modulation and multiplexing schemes have been proposed to improve the spectrum efficiency of the optical systems. After decades of development, the data rate of the optical system has been increased from 1 Gb/s, to 10 Gb/s, then to 100 Gb/s. The superchannel has also been demonstrated recently for Tb/s transport [14,15]. With the popularity of coherent technique, along with the high-order modulation, polarization-division multiplexing (PDM), and wavelength-division multiplexing (WDM) techniques, transmission rate over 100 Gb/s has become commercially available. However, the optical transmission systems are always faced with the challenge to achieve a higher channel capacity in order to keep up with the increasing traffic demand. In this thesis, we investigate into the promising techniques and schemes to improve the capacity for future high-capacity optical networks.

1.2 Review of multiplexing techniques for high-capacity transmission

The spectrum efficiency of optical systems can be improved using various multiplexing schemes, such as optical code division multiple access (OCDMA) [16-19],optical time division multiplexing (OTDM) [20-23],wavelength division multiplexing (WDM) [24-28], as well as coherent optical transmission [29-31]. By exploring different dimensions of freedom for signal modulation, the multiplexing techniques provide a feasible pathway towards high-capacity optical transmission.

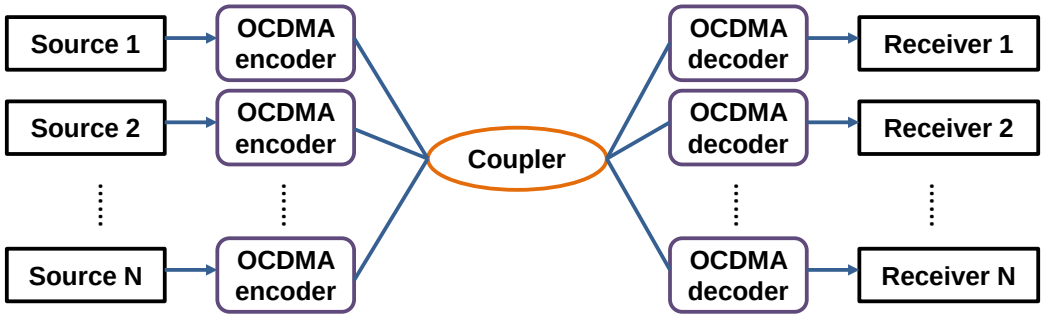

Fig. 1.1 Conceptual diagram of an OCDMA network.

For optical code division multiple access technique, also as optical code division multiplexing technique, all users in the network share the same channel, while a unique signature sequence is assigned to each user. The signature sequence is used to encode the signal at the transmitter and decode it at the receiver. In that way, the desired information can be selected from the information flow in the OCDMA networks. The structure of a typical OCDMA network consisting of *N* users is shown in Fig. 1.1 [32 , 33]. The desired signal can be ideally recovered using the corresponding signature sequence, if the signature sequences for different users are perfectly orthogonal to each other. However, it is hard to find a set of signature sequences, which are perfectly orthogonal to each other, in the real OCDMA networks. The imperfect orthogonality of the signature sequences leads to the cross talk between different users, degrading the system performance. The noise due to the cross talk increases with the number of users. Therefore, to have sufficient signal to noise ratio (SNR), the OCDMA network can only support a limited number of users. The capacity of the OCDMA network is significantly limited by imperfect orthogonality of the signature sequences.

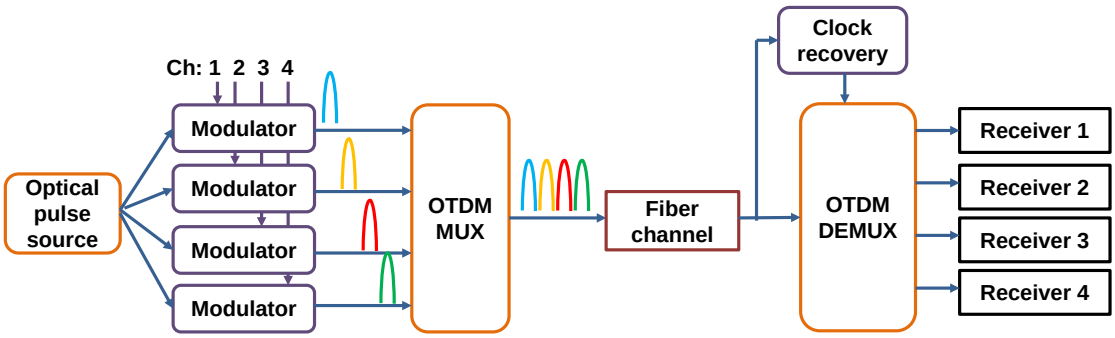

Fig. 1.2 Conceptual diagram of an OTDM transmission system.

Optical time division multiplexing technique provides an alternative approach to integrate multiple low-rate data streams into a high-rate one. Without multiplexing

technique, the achievable line rate of a channel is limited by the maximum data rate of its electrical components. To achieve a higher data rate, the OTDM technique divides one transmission period into several time slots, and each slot is loaded with data from different channels. By loading the data from *N* channels into *N* time slots, the multiplexed channel can reach a line rate *N* times higher [34-37]. Fig. 1.2 is the schematic depiction of a typical OTDM system, consisting of 4 tributary channels with bit rate of 10 Gb/s. At the transmitter, optical pulse train is generated by a mode-locked laser diode at repetition rate of 10 GHz with a narrow pulse width. The optical pulse train is then divided into 4 paths. In each path, the signal from one tributary channel is modulated onto the optical pulse train through a $LiNbO_3$ intensity modulator. The four modulated optical pulse trains are then delayed with different clock periods and interleaved by an optical timing multiplexer. After the time domain multiplexing, the aggregated data stream has a bit rate of 4x10 Gb/s. The OTDM receiver typically consists of three sub-systems, the optical gate, the timing extraction device, and the optoelectronic receiver, where the optical pulses for the four tributary channels are separated for subsequent detection and signal processing.

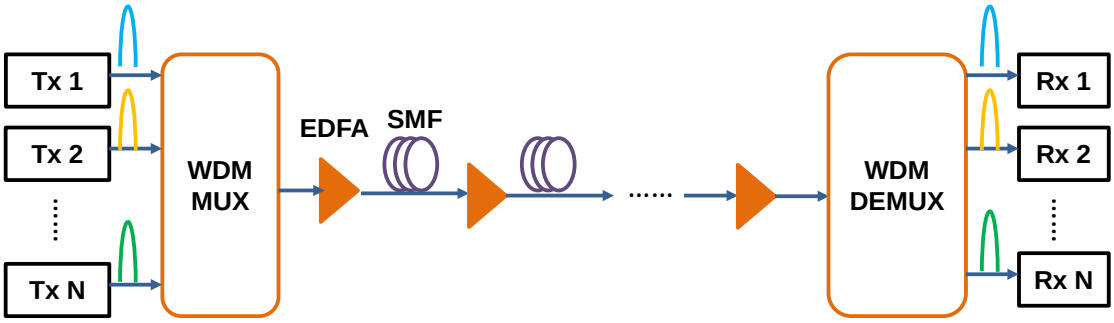

Fig. 1.3 Conceptual diagram of a WDM transmission system.

Wavelength division multiplexing technique increases the fiber capacity by transmitting parallel channels using different wavelengths. With the commercialization of the optical devices, such as arrayed wave-guide grating (AWG), dispersion-managed fibers and erbium-doped fiber amplifier (EDFA), WDM has become a preferred scheme to increase the channel capacity due to its simple implementation. The basic configuration of a WDM system is illustrated in Fig. 1.3. After signal modulation, all wavelength channels are combined with a WDM multiplexer at the transmitter. The parallel wavelength channels are separated with a WDM de-multiplexer at the receiver. With the development of the WDM technique, the number of wavelength channels supported in a WDM system has been greatly

increased. The state-of-art WDM optical system can support more than hundreds of wavelength channels simultaneously over a single fiber. In addition to the increased channel capacity, WDM optical networks also have the advantages of easy routing and switching based on wavelength scheme[38-41]. With the help of optical add/drop multiplexer (OADM) and optical cross connect (OXC), the information can be added, dropped and exchanged flexibly in WDM optical networks [42-46]. The capacity of the WDM system can be increased by adding more wavelength channels into the network. However, an upper limit is imposed by the available bandwidth of the optical components.

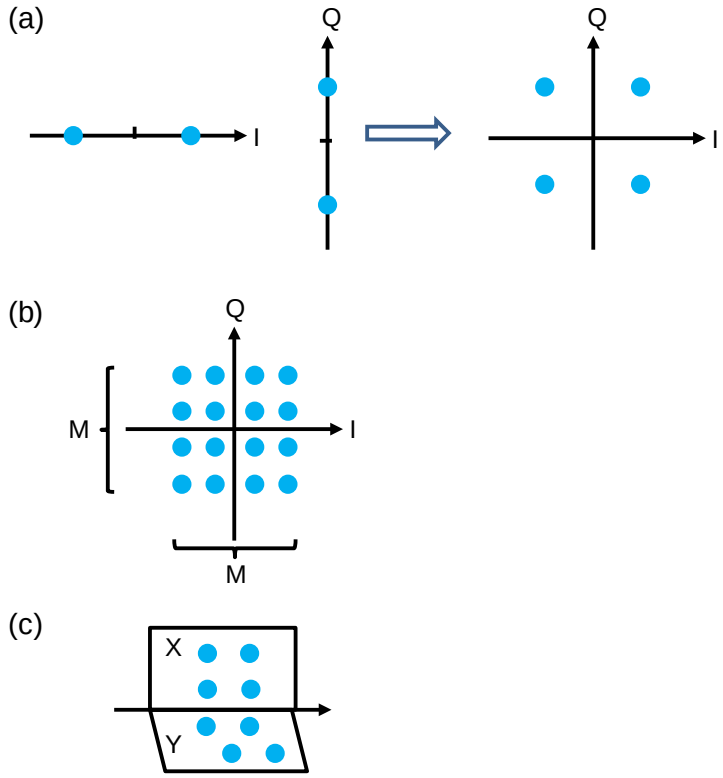

Fig. 1.4 Signal multiplexing in coherent optical transmission. (*a*) *signal multiplexing in real and imaginary parts.* (*b*) *higher-order modulation.* (*c*) *signal multiplexing in two polarizations.*

To keep up with the ever increasing traffic demand, it is of great interest to explore the available dimensions of freedoms for signal modulation and multiplexing. Coherent optical transmission can significantly improve the spectrum efficiency by multiplexing the signal into different dimensions of the optical field [47,48]. Firstly, the signal can be multiplexed in the real and imaginary parts of the optical field using the coherent technique. If the real and imaginary parts are modulated with the bipolar

signal, a constellation of four points is obtained in the optical field as shown in Fig. 1.4(a), doubling the capacity. The spectrum efficiency can be further improved using higher-order modulation, which increases the modulation level to *M* for both real and imaginary parts. By doing that, the constellation size is increased to *M*x*M*, as shown in Fig. 1.4(b). Moreover, the coherent technique is able to multiplex the signal using the two polarizations of the optical field. The capacity is doubled again when both two polarizations are loaded with signal as shown in Fig. 1.4(c). Beside of the increased spectrum efficiency, the coherent technique also brings the benefit of improved sensitivity. In a coherent transmission system, the RF signal is up-converted to the optical field at the transmitter, and the optical signal is converted back to the RF field at the receiver. With the linear mapping between RF and optical signal, the linear channel impairment can be easily compensated using the powerful digital signal processing (DSP) technique, leading to a significant improvement in system performance [49-52].

1.3 Improved spectral efficiency in short-reach optical networks

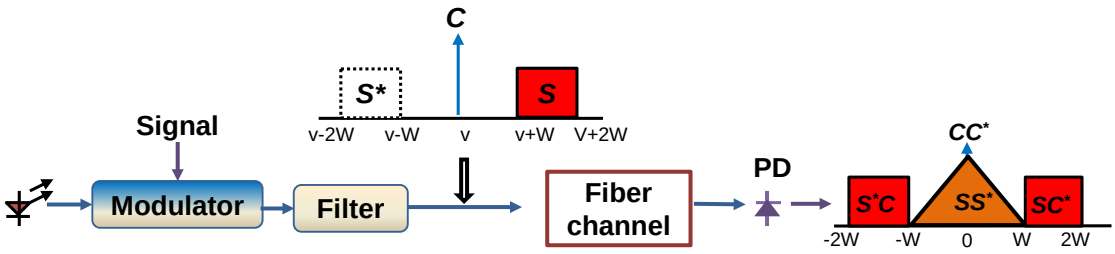

Fig. 1.5 Conceptual diagram of SSB transmission system.

Coherent optical system has been widely applied in the long-haul optical networks to extend the transmission distance and improve the channel capacity, due to its high performance and easy integration with DSP. Meanwhile, the short-reach optical networks are faced with an urgent demand to upgrade its capacity to support the increasing traffic between the data centers. Different from the long-haul transmission, the short-reach networks have to give the system cost a primary consideration. Therefore, the high-performance coherent transceivers are unsuitable for the applications in short-reach optical networks due to the high cost. Single-side band (SSB) modulation has been widely investigated for the high-performance short-reach interconnects, providing a favorable solution to reach a compromise between the system cost and performance [53-55]. The benefits are twofold: (1) SSB modulation

scheme uses a single-ended photodiode (PD) as the receiver, significantly reducing the system cost; (2) SSB modulation scheme adopts the self-coherent detection approach to achieve a linear channel, enabling a linear channel compensation at the receiver. Fig. 1.5 shows the typical structure of a SSB transmission system. The signal generated at the transmitter has a Hermitian symmetric spectrum. An optical filter is then used to filter out half side of the spectrum, leaving another half side of the spectrum. The optical carrier is transmitted together with the optical signal by adjusting the modulation bias of the modulator away from the null point. At the receiver, the signal and carrier beat with each other in a single-ended PD. The signal is recovered from the linear beating term, resulting in a linear channel.

1.4 Motivation

With the development of global internet, significant amount of effort has been made to upgrade the optical networks in order to keep up with the extraordinary increase of the traffic demand. Aiming at different transmission distance and data rate, the long-haul and short-reach optical networks are faced with different challenges during the capacity upgrade. For the long-haul transmission, various modulation and multiplexing schemes have been proposed to fully utilize the available modulation freedoms in single mode fiber. Since the fiber nonlinearity imposes a hard limit on the channel capacity of SMF [56], the long-haul transmission will be confronted with a traffic crunch in the near future, without inducing a new modulation dimension to sustain the capacity increase. Multi-core fiber (MCF) and multi-mode fibers (MMF) can potentially enable the simultaneous propagation of tens or hundreds of spatial modes, overcoming the capacity limitation in SMF [57- 62]. Multiple-input multiple-output (MIMO) DSP is used in MMF transmission to compensate the linear channel impairments, including the random coupling and group delay between the different spatial modes [63- 66]. To maintain the DSP complexity to a manageable level, few-mode fiber (FMF) is proposed [67], supporting the propagation of a small number of modes (e.g., 2 to 6). The performance of FMF transmission largely depends on our knowledge of the channel matrix, which represents the linear impact of the FMF channel and plays a critical role in the signal recovery. The channel matrix is not fixed, but changes gradually in both frequency and time domain during the transmission. A thorough study of its dynamic properties will lead to a more effective signal recovery, providing a feasible way towards the high-capacity transmission based on FMF.

For the short-reach optical networks, the cost becomes a primary consideration, shared by a smaller number of users. Therefore, the expensive coherent transceivers are cost prohibitive in short-reach networks, despite of the high performance. Direct modulation and direct detection (DM/DD) system provides a cost-effective alternative [68]. However, the nonlinear channel in such system leads to significant signal distortion, making the DM/DD system unsuitable for the high-speed transmission. To fulfil the increasing traffic demand, the short-reach network is faced with the challenge to improve its system performance while maintaining the system cost to an acceptable level. Considering the choices of the transceiver, the system cost can be reduced in two ways. We can either reduce the cost at receiver by using direct detection, or reduce the cost at transmitter by using directly modulated laser (DML). High system performance is ensured by a linear channel where channel compensation can be easily performed. To achieve a linear down-conversion of optical field using direct detection, an optical carrier is transmitted together with the signal. The linear down-conversion is given by the linear beating term between the signal and the carrier. Large amount of efforts have been devoted to improve the spectral efficiency of the so-called self-coherent detection scheme, providing promising potential for the high-capacity short-reach networks [54,55].

When DML is adopted to reduce the cost of the transmitter, coherent detection is necessary at the receiver to ensure a linear channel for the system. Though both intensity and phase information can be recovered at the coherent receiver, the phase information is usually discarded during the signal recovery, since the intensity modulation is more straightforward for the signal recovery [69,70]. However, the optical phase is also modulated with the signal due to the chirp effect in DML. By combining the phase and intensity modulation during the signal recovery, the receiver sensitivity can be greatly improved. Two novel schemes have been proposed to perform the signal recovery based on the combined phase and intensity information for the short-reach transmission with DML. Compared to the conventional scheme using the intensity-only signal recovery, the two schemes show greater potentials in achieving a higher data rate for short-reach networks.

1.5 Thesis outline

The content of the thesis is structured as follows:

Chapter 1 Introduction This chapter reviews the development of optical fiber communication, as well as various modulation and multiplexing techniques proposed for the high-capacity optical transmission. The motivation of this piece of work is also introduced in this chapter. The factors that limit the capacity in long-haul and short-reach optical systems are discussed in detail. This work is devoted to the promising solutions that can overcome the capacity limitations in long-haul and short-reach optical systems and support the capacity increase in future optical networks.

Chapter 2 Principle of Optical OFDM System In this chapter, the principles of coherent optical transmission are discussed, especially for the coherent optical OFDM (CO-OFDM) transmission. The fundamentals of OFDM are introduced, including its basic mathematical formulation, discrete Fourier transform implementation and cyclic prefix. The coherent optical OFDM techniques, including the system architecture, modulation and detection devices, polarization-division-multiplexing, and MIMO signal processing, are also introduced in detail.

Chapter 3 Space-Division-Multiplexing and Few-Mode-Fiber This chapter introduces the fundamentals of space-division-multiplexing (SDM), where SDM systems based few-mode-fiber (FMF) transmission is presented. The enabling techniques, including the FMF design, mode multiplexing/de-multiplexing components, and MIMO signal processing, are discussed. After that, a study is conducted for the property of FMF channel matrix, which plays a critical role in the receiver design and ultimate channel performance. The frequency dependence of the channel matrix is studied in terms of the autocorrelation function (ACF), showing the divergence of the fiber channel in frequency domain, while the dynamics of the channel matrix is measured with the help of Stokes-space analysis, showing the change of the fiber channel in time domain.

Chapter 4 Advanced Modulation Formats for Direct Detection System In this chapter, we study the advanced modulation formats which can improve the spectral efficiency of the direct detection (DD) system for the high-performance short-reach optical networks. We first provide a review of the current advanced modulation formats that can achieve superior electrical spectral efficiency and data rate beyond that of simple direct modulation direct detection (DM/DD) system. Among these formats, Stokes vector direct detection (SV-DD) achieves the highest electrical spectrum efficiency, presenting itself as a promising candidate for future short-reach

networks. Then, we expound several novel algorithms to improve the performance of the SV-DD system under the impairments of polarization mode dispersion (PMD) and polarization dependent loss (PDL).

Chapter 5 Complex Modulation and Detection with Directly Modulated Laser In this chapter, we provide thorough analysis of two recovery schemes, which utilize the combined phase and intensity information to improve the performance of the short-reach optical system using directly modulated laser (DML). The phase modulation mechanism in DML is analyzed, derived from the dynamic laser chirp equation. Based on that, the DSP of the two schemes is revealed, including the chirp characterization and signal recovery algorithms. The experimental demonstrations of the two schemes are also presented, showing the improvement achieved by the two recovery schemes.

Chapter 6 Conclusions In this chapter the main results of the thesis are reviewed and summarized.

1.6 Contributions

The contributions of this work in thesis are listed as follows:

Chapter 3 We have derived a canonical stochastic differential equation (SDE) for the channel matrix of FMF in the regime of strong coupling, and studied the frequency properties of FMF channel in terms of the autocorrelation function (ACF). We have also measured the fluctuation speed of FMF channel using the Stokes-space analysis. The studies provide an important guide for the receiver design in high-performance FMF transmission.

Chapter 4 We have proposed the Stokes vector detection scheme for high-performance direct detection system, enabling the polarization recovery in Stokes space using direct detection. We have also proposed advanced DSP algorithms to improve the performance of the SV-DD system under the impairments of PMD and PDL.

Chapter 5 We have introduced two signal recovery schemes to fully utilize the phase and intensity modulation in the DML system. The experimental demonstrations show great improvement in receiver sensitivity compared to the conventional DML system.

1.7 Publications related to this thesis

1. Q. Hu, D. Che, Y. Wang, F. Yuan, Q. Yang, and W. Shieh, “Complex modulation and detection with directly modulated laser,” Optics express 23 (25), 32809-32819 (2015).
2. Q. Hu, D. Che, Y. Wang, A. Li, J. Fang and W. Shieh, “Beyond amplitude-only detection for digital coherent system using directly modulated laser,” Optics Letter 40(12), 2762-2765 (2015).
3. Q. Hu, D. Che, Y. Wang and W. Shieh, “PMD induced impairment mitigation in Stokes vector direct detection systems,” Optical Fiber Communication Conference 2015, Paper Th1E.2.
4. Q. Hu, D. Che, Y. Wang and W. Shieh, “Advanced modulation formats for high-performance short-reach optical interconnects,” Optics Express 23 (3), 3245-3259 (2015).
5. Q. Hu, D. Che, and W. Shieh, “Mitigation of PMD induced nonlinear noise in Stokes vector direct detection system,” European Conference on Optical Communication (ECOC) 2015, Paper P.3.4.
6. Q. Hu, X. Chen, A. Li, and W. Shieh, “High-dimensional Stokes-space analysis for monitoring fast change of mode dispersion in few-mode fibers,” Optical Fiber Communication Conference 2014, Paper W3D.3.
7. Q. Hu, X. Chen, and W. Shieh, “Comparison of channel fluctuation between single-mode fibers and few-mode fibers using Stokes-space analysis,” OECC/ACOFT 2014, Paper MO2C-3.
8. Q. Hu, and W. Shieh, “Stochastic analysis of channel transmission matrix for optical fibers with gradually varying birefringence model,” Asia Communications and Photonics Conference 2013, Paper AW3G.2.
9. Q. Hu, and W. Shieh, “Autocorrelation function of channel matrix in few-mode fibers with strong mode coupling,” Optics Express 21(19), 22153-22165 (2013).
10. D Che, Q Hu, F Yuan, Q Yang, W Shieh, “Enabling Complex Modulation of Directly Modulated Signals Using Laser Frequency Chirp,” Photonics Technology Letters, IEEE 27 (22), 2407-2410 (2015)

11. D Che, Q Hu, W Shieh, “High-spectral-efficiency optical direct detection using the stokes vector receiver,” European Conference on Optical Communication (ECOC) 2015, 1-3
12. D Che, Q Hu, F Yuan, W Shieh, “Enabling complex modulation using the frequency chirp of directly modulated lasers,” European Conference on Optical Communication (ECOC) 2015, 1-3
13. D Che, A Li, Q Hu, X Chen, and W Shieh, ” Implementing Simplified Stokes Vector Receiver for Phase Diverse Direct Detection,” Optical Fiber Communication Conference 2015, Th1E. 4.
14. W Shieh, D Che, Q Hu, and A Li, “Linearization of Optical Channels with Stokes Vector Direct Detection,” Optical Fiber Communication Conference 2015, Th1E. 5.
15. D Che, A Li, X Chen, Q Hu, Y Wang, and W Shieh, “Stokes Vector Direct Detection for Linear Complex Optical Channels,” Journal of Lightwave Technology 33 (3), 678-684 (2015).
16. A Li, Y Wang, Q Hu, and W Shieh, “Few-mode fiber based optical sensors,” Optics express 23 (2), 1139-1150 (2015).
17. A Li, Y Wang, Q Hu, D Che, X Chen, and W Shieh, “Measurement of distributed mode coupling in a few-mode fiber using a reconfigurable Brillouin OTDR,” Optics letters 39 (22), 6418-6421 (2014).
18. Y Wang, A Li, X Chen, Q Hu, and W Shieh, “Characterization of in-line interferometric based temperature sensors in two-mode fibers,” Asia Communications and Photonics Conference 2014, Paper AW3D. 4.
19. D Che, X Chen, A Li, Q Hu, Y Wang, and W Shieh, “Optical direct detection for 100G short reach applications,” Asia Communications and Photonics Conference 2014, Paper AF1H. 4.
20. A Li, Q Hu, D Che, Y Wang, and W Shieh, “Measurement of distributed mode coupling in a few-mode fiber using a Brillouin optical time domain reflectometer,” European Conference on Optical Communication (ECOC) 2014, 1-3.
21. D Che, A Li, X Chen, Q Hu, Y Wang, and W Shieh, “Stokes vector direct detection for short-reach optical communication,” Optics letters 39 (11), 3110-3113 (2014).

22. A Li, Q Hu, X Chen, BY Kim, W Shieh, B Wang, H Huang, K Wang, and H Long, "Characterization of distributed modal birefringence in a few-mode fiber based on Brillouin dynamic grating," Optics letters 39 (11), 3153-3156 (2014).
23. D Che, A Li, X Chen, Q Hu, Y Wang, and W Shieh, "160-Gb/s stokes vector direct detection for short reach optical communication," Optical Fiber Communication Conference 2014, 1-3.
24. D Che, X Chen, J He, A Li, W Shieh, "102.4-Gb/s single-polarization direct-detection reception using signal carrier interleaved optical OFDM," Optical Fiber Communication Conference 2014, Paper Tu3G. 7.
25. X Chen, A Li, Q Hu, J He, D Che, Y Wang, and W Shieh, "Demonstration of direct detected optical OFDM signals via block-wise phase switching," Journal of Lightwave Technology 32 (4), 722-728 (2014).
26. Y Wang, Q Hu, X Chen, A Li, T Anderson, and W Shieh, "3-λ characterization of phase response for optical receivers," Optics letters 39 (3), 670-673 (2014).
27. Y Wang, Q Hu, X Chen, T Anderson, and W Shieh, "Using 3 wavelengths to measure photodiode phase response with laser phase noise robustness," Engineers Australia 2014, 1-3.
28. A Li, Q Hu, and W Shieh, "Flexible few-mode Brillouin optical time-domain analyzer based on spatial mode multiplexer," Engineers Australia 2014, 1-3.
29. A Li, Q Hu, and W Shieh, "Characterization of stimulated Brillouin scattering in a circular-core two-mode fiber using optical time-domain analysis," Optics express 21 (26), 31894-31906 (2013).
30. X Chen, D Che, A Li, J He, Y Wang, Q Hu, and W Shieh, "Training Symbol Assisted Signal-to-Signal Beat Noise Cancellation for Direct Detected Optical OFDM Systems," Asia Communications and Photonics Conference 2013, Paper AF1E. 2.
31. A Li, D Che, X Chen, Q Hu, Y Wang, and W Shieh, "Signal-to-Signal Beat Noise Cancellation for Direct Detection Optical OFDM System Based on Block-wise Signal Phase Switching," Asia Communications and Photonics Conference 2013, Paper AF1E. 6.

32. A Li, D Che, X Chen, Q Hu, Y Wang, and W Shieh, “48.8-Gb/s 16-QAM direct-detection optical OFDM based on block-wise signal phase switching,” European Conference on Optical Communication (ECOC) 2013, 1-3.
33. A Li, D Che, X Chen, Q Hu, Y Wang, and W Shieh, “61 Gbits/s direct-detection optical OFDM based on blockwise signal phase switching with signal-to-signal beat noise cancellation,” Optics letters 38 (14), 2614-2616 (2013).
34. X Chen, A Li, D Che, Q Hu, Y Wang, J He, and W Shieh, “Block-wise phase switching for double-sideband direct detected optical OFDM signals,” Optics express 21 (11), 13436-13441 (2013).
35. X Chen, J Ye, Y Xiao, A Li, J He, Q Hu, and W Shieh, “Equalization of two-mode fiber based MIMO signals with larger receiver sets,” Optics express 20 (26), B413-B418 (2013).

2 Coherent Optical OFDM for High-capacity Long-haul Transmission

2.1 Introduction

The fast growth of bandwidth-rich internet applications, such as online mobile applications and cloud computing, has led to a huge demand on the bandwidth of optical transports. To satisfy the ever increasing bandwidth demand, extensive studies have been conducted to increase the spectral efficiency in the state-of-the-art optical transmission systems. With the motivation to enhance the channel capacity within the limited bandwidth, various multiplexing techniques has been proposed, including the optical wavelength-division multiplexing (WDM) [24-28], time-domain multiplexing (OTDM) [20-23], as well as the coherent optical communication [29-31]. With the increase of the data rate and transmission reach, the optical signal is extremely sensitive to the chromatic dispersion (CD), polarization mode dispersion (PMD), reconfigurable optical add/drop multiplexer (ROADM) filtering effects, and imperfections of the electric-optics components. Coherent communication provides a preferable solution to overcome those challenges in long-haul transmission due to its easy compensation of the linear channel impairments with the help of powerful DSP. High-speed coherent optical system has been demonstrated with either single-carrier modulation (SCM) or multi-carrier modulation (MCM) scheme. The SCM scheme transmits all data with one carrier, while the MCM scheme reaches the same data rate by binding up a number of carriers modulated at lower data rate. Orthogonal frequency-division multiplexing (OFDM) is a MCM scheme which can utilize the bandwidth in an efficient way [71,72]. Orthogonal subcarriers are banded up tightly in the OFDM scheme without inducing any inter-carrier interference (ICI).

OFDM has already been adopted on broad range of applications, such as digital television, audio broadcasting, wireless networking and broadband internet access [73,74]. In fiber-optic communications, coherent optical OFDM (CO-OFDM) was proposed by combining the OFDM scheme with the coherent transmission [75-77]. While the coherent transmission provides a linear channel for the OFDM signal, the OFDM scheme enables a sophisticated way to improve the performance of the coherent system [78]. By inserting cyclic prefix (CP) and training symbols into the OFDM signal, CO-OFDM is able to estimate the channel, and compensate the inline

CD and PMD easily. Because of the high receiver sensitivity, high spectral efficiency and robustness against the linear impairments, CO-OFDM has attracted great interest from the optical communication community. Despite of the disadvantages of high peak-to-average-power ratio (PAPR) and reduced tolerance to phase noise, the advantages of the CO-OFDM are also obvious compared to the SCM scheme. CO-OFDM has tight spectrum, scalable bandwidth, and can be easily upgraded to higher order modulation. With the help of pilot symbols and pilot subcarriers, CO-OFDM enables the channel and phase estimations in a computationally efficient way. With these benefits, CO-OFDM provides a promising solution for the future high-speed optical transmission. Besides CO-OFDM, the OFDM signal also finds wide applications in direct detection systems, aiming at simpler receiver structure and lower system cost. In this chapter, we will first review the basic principle of OFDM signal and the associated DPS for coherent transmission. The application of OFDM signal in various direct-detection systems will be discussed in Chapter 4.

2.2 Coherent transmitter and receiver

Coherent transmitter and receiver are the key components to achieve a linear channel in the coherent optical system, providing linear conversions between the baseband and optical signal. Coherent transmitter up-converts the baseband signal to the optical field using a pair of Mach-Zehnder modulators (MZMs) with a 90^{O} phase shift, while coherent detection down-converts the optical signal back to the baseband using a pair of balanced photodiodes (PDs). With a linear channel, DSP can perform the best role to mitigate the channel impairments in the coherent systems.

2.2.1 I/Q modulator

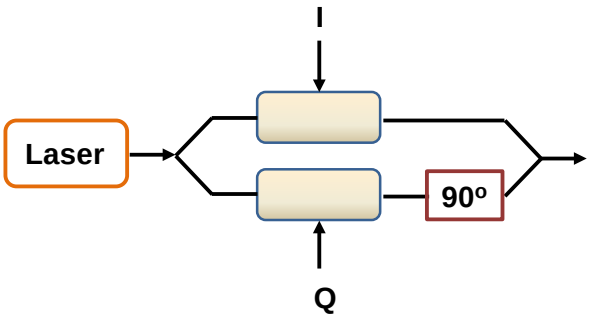

Fig. 2.1 Conceptual structure of an IQ modulator.

Optical I/Q modulator is used to up-convert the complex baseband signal into the optical field [79-81]. As show in Fig. 2.1, an I/Q modulator consists of two Mach-

Zehnder modulators (MZMs) and a 90° phase shift. The MZMs modulate the real and imaginary parts of the signal onto the optical carrier respectively, while a 90° phase shift is induced in one branch to ensure the orthogonality between the I and Q. Combining the optical signal from the two branches, a linear up-conversion from the baseband to the optical field is achieved.

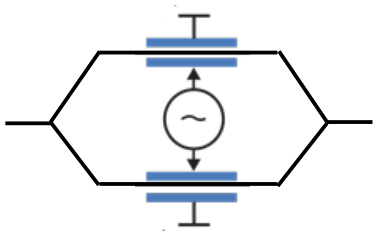

Fig. 2.2 Conceptual structure of a MZM.

MZM is typically fabricated on $LiNbO_3$ crystal, with a structure shown in Fig. 2.2. The input optical signal E_{in} is divided into two paths at the input, and a phase delay is induced in each path when the optical signal passes through the Kerr medium. The phase delay can be controlled by the voltage applied to the Kerr medium. When voltages, $V_1(t)$ and $V_2(t)$, are applied, the corresponding phase delays, $\phi_1(t)$ and $\phi_2(t)$, in the two paths are

$$\phi_1(t) = K \cdot V_1(t) = \phi_0 - \pi \frac{V_1}{V_\pi} \tag{2.1}$$

$$\phi_2(t) = K \cdot V_2(t) = \phi_0 - \pi \frac{V_2}{V_\pi}, \tag{2.2}$$

where K is the electro-optic factor, V_π is the voltage to induce a phase delay of π, and ϕ_0 is the phase delay when the applied voltage is 0. The optical signals from the two paths are then combined again, giving the output signal E_{out} as

$$\begin{aligned} E_{out} &= \frac{E_{in}}{2}\left(\exp(j\phi_1(t)) + \exp(j\phi_2(t))\right) \\ &= E_{in}\cos\left(\frac{\phi_1(t)-\phi_2(t)}{2}\right)\exp\left(j\frac{\phi_1(t)+\phi_2(t)}{2}\right). \end{aligned} \tag{2.3}$$

Typically, MZM is working at the “push-pull” mode, in other words, opposite voltages are applied for the two paths, $V_1(t) = -V_2(t)$. As a result, the phase term in Eq. (2.3) equals to zero, and the equation can be rewritten as

$$E_{out} = E_{in} \cos(K \cdot V_1(t)) . \quad (2.4)$$

Therefore, the optical phase maintains the same after the modulation in MZM, while the optical amplitude is modulated as a cosine function of the applied voltage. The transfer function of a MZM is plotted in Fig. 2.3. An approximately linear transfer can be found near the null-point. By working within the linear region, the real and imaginary parts of the baseband signal is up-converted to the optical field with a pair of MZMs. Before combining the two path, a 90^o phase shift is inserted to ensure the orthogonality between the real and imaginary parts.

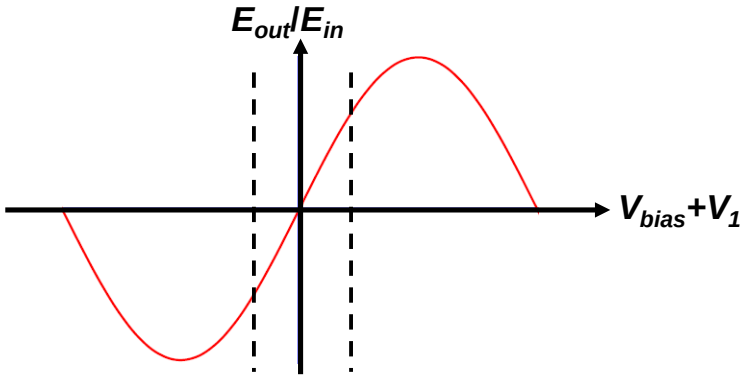

Fig. 2.3 Transfer function of MZM.

2.2.2 Coherent detection

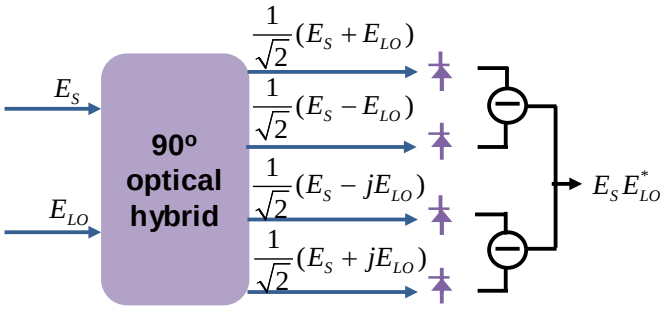

Fig. 2.4 Conceptual diagram of a coherent detection.

The coherent receiver can be used to detect the optical field, including its intensity, phase and polarization, with the help of local oscillator (LO). Base on where the LO frequency is placed, coherent receiver can be divided into two categories, the homodyne receiver and heterodyne receiver. The homodyne receiver recovers the real and imaginary parts of the signal in parallel, and therefore needs twice as many photodetectors (PDs) as in the heterodyne receiver. In contrast, the heterodyne receiver needs an image rejection filter to recover the real and imaginary parts, and

therefore requires twice the bandwidth as required in the homodyne receiver. The conceptual diagram of a homodyne receiver is shown in Fig. 2.4, which consists of a 6-port 90^o optical hybrid, and a pair of balanced PD [80,81]. The received optical signal is mixed with the LO in 90^o optical hybrid, generating a 180^o phase shift between the two branches for the balanced PD, and a 90^o phase shift between the *I* and *Q* components [82,83]. Denoting the input signal and LO as E_s and E_{LO}, the four outputs of the 90^o optical hybrid are $E_1 = \frac{1}{\sqrt{2}}\left[E_s + E_{LO}\right]$, $E_2 = \frac{1}{\sqrt{2}}\left[E_s - E_{LO}\right]$, $E_3 = \frac{1}{\sqrt{2}}\left[E_s - jE_{LO}\right]$ and $E_4 = \frac{1}{\sqrt{2}}\left[E_s + jE_{LO}\right]$. The PD takes square operation for the optical signal. Thus, the output photocurrents of PD1 and PD2, I_1 and I_2, are

$$I_1 = \frac{1}{2}\left\{\left|E_s\right|^2 + \left|E_{LO}\right|^2 + 2\,\mathrm{Re}\left\{E_s E_{LO}^*\right\}\right\} \tag{2.5}$$

$$I_2 = \frac{1}{2}\left\{\left|E_s\right|^2 + \left|E_{LO}\right|^2 - 2\,\mathrm{Re}\left\{E_s E_{LO}^*\right\}\right\}. \tag{2.6}$$

Considering the ASE noise of the optical signal and the relative intensity noise (RIN) of the LO source, we can write the intensity of the signal and LO, $\left|E_s\right|^2$ and $\left|E_{LO}\right|^2$, as

$$\left|E_s\right|^2 = \left|E_r\right|^2 + \left|n_o\right|^2 + 2\,\mathrm{Re}\left\{E_r n_o^*\right\} \tag{2.7}$$

$$\left|E_{LO}\right|^2 = I_{LO}\left(1 + I_{RIN}\left(t\right)\right). \tag{2.8}$$

where E_r is the received signal with on ASE noise, n_o is the ASE noise, I_{LO} is the average power of LO and I_{RIN} is the RIN of LO. The balanced PD subtracts the photocurrents obtained from the two PDs, and provides us

$$I_I\left(t\right) = I_1 - I_2 = 2\,\mathrm{Re}\left\{E_s E_{LO}^*\right\}. \tag{2.9}$$

The square terms in Eqs. (2.5) and (2.6) are cancelled during the subtraction operation, suppressing the impact from ASE noise and RIN. From the left term, the *I* component of the signal is recovered, giving the real part of the beating term. Similarly, the *Q* component is recovered by another pair of PDs, PD3 and PD4. Combining the *I* and *Q* components, the constructed complex signal is the replica of the optical field with frequency down-shifted by the LO.

Coherent detection represents the ultimate performance in receiver sensitivity, spectral efficiency. The advantages are twofold: (1) the linear down-conversion of the received optical signal, and (2) the suppressing of the common mode noise. Together with the *I*/*Q* modulator, the coherent detection, leads to a linear channel, where complex signal can be transmitted and processed easily using DSP.

2.3 Basics of OFDM

In the conventional MCM scheme, the input serial data stream is divided into N_{sc} parallel sub-streams, reducing the data rate by N_{sc} times for each sub-stream. The sub-streams are transmitted via carriers with different frequency. The combination of the subcarriers, $s(t)$, can be expressed as

$$s(t) = \sum_{k=1}^{N_{sc}} c_k e^{j2\pi f_k t}, \tag{2.10}$$

where c_k is the complex data loaded in the k-th (k=1,2,…,N_{sc}) subcarrier with frequency f_k. At the receiver, a filter or a correlator that matches the subcarrier waveform is used to detect the data in the corresponding subcarrier. Non-overlapped subcarriers, and large frequency spacing are adopted in the conventional MCM scheme to minimize the interference between the subcarriers, leading to the reduced spectral efficiency. Meanwhile, large number of oscillators and filters are needed at the receiver ends, increasing the system cost.

As a special form of MCM, OFDM can minimize the subcarrier spacing by using a set of overlapped subcarriers. The OFDM subcarriers are orthogonal to each other, satisfying the following condition [71-74]

$$f_k - f_l = n\frac{1}{T_s}. \tag{2.11}$$

The frequency of the subcarrier is spaced at n times of the inverse of the symbol period, T_s. By satisfying the orthogonality condition of Eq. (2.11), the subcarriers are free from the inter-carrier interference (ICI) in spite of strong spectral overlapping.

An OFDM frame includes a number of OFDM symbols, which can be expressed as [71-74]

$$s(t) = \sum_{i=-\infty}^{+\infty} \sum_{k=-Nsc/2+1}^{k=N_{sc}/2} c_{ki} \exp(j2\pi f_k (t - iT_s)) \Pi(t - iT_s), \tag{2.12}$$

$$f_k = \frac{k-1}{T_s}, \tag{2.13}$$

$$\Pi(t) = \begin{cases} 1, (0 < t \le T_s) \\ 0, (t \le 0, t > T_s) \end{cases}, \tag{2.14}$$

where c_{ki} is the *i*-th information symbol at the *k*-th subcarrier, $\Pi(t)$ is the pulse waveform of the symbol, and f_k is the frequency of the *k*-th subcarrier. The OFDM signal is modulated as demodulated based on the IDFT/DFT operation. When the OFDM symbol is sampled at the sampling rate of T_s/N_{sc}, the signal at the *m*-th sampling point of Eq. (2.12) is

$$s_m = \sum_{k=1}^{N_{sc}} c_k \exp(j2\pi f_k \frac{(m-1)T_s}{N_{sc}}). \tag{2.15}$$

Substituting Eq. (2.13) into Eq. (2.15), we have

$$s_m = \sum_{k=1}^{N_{sc}} c_k \exp(j2\pi \frac{(m-1)(k-1)}{N_{sc}}) = F^{-1}\{c_k\}. \tag{2.16}$$

where F^{-1} stands for the IDFT operation. The discrete value of the transmitted OFDM signal is given by N_{sc}-point IDFT of the frequency domain signal. Similarly, the frequency domain signal $c_k^{'}$ can be recovered at the receiver using N_{sc}-point DFT as

$$c_k^{'} = F\{r_m\}, \tag{2.17}$$

where and F stands for the DFT operation, and r_m is the received value sampled at time interval of T_s/N_{sc}.

The orthogonality of the subcarriers will be ruined by the inter-symbol interference (ISI) induced by the dispersion effects in the fiber channel, such as the chromatic dispersion (CD) and polarization mode dispersion (PMD). The ISI can be avoided by inserting cyclic prefix (CP) into the OFDM symbol [84]. As shown in Fig. 2.5, the CP is an identical copy of the last or the first portion of the data symbol. By appending the CP to the front or end of the symbol, the OFDM symbol is cyclically

extended. If the CP is longer the time delay induced by the channel dispersion, the ISI falls into the guard interval, leaving the data in the observation window clean. Thus, the orthogonality of the subcarriers is reserved. However, the CP induces redundant information into the OFDM symbol, reducing the spectral efficiency. For this reason, a shorter CP is preferable as long as the data in observation window is free from the ISI. The minimum guard interval for ISI-free transmission is

$$\frac{c}{f^2}\left|D_t\right| \cdot N_{SC} \cdot \Delta f + DGD_{\max} \le \Delta_G , \tag{2.18}$$

where f is the frequency of the optical carrier, c is the speed of light, D_t is the total accumulated chromatic dispersion in units of ps/km, N_{sc} is the number of subcarriers, Δf is the subcarrier spacing, and $DGD_{\max}$ is the maximum budget DGD.

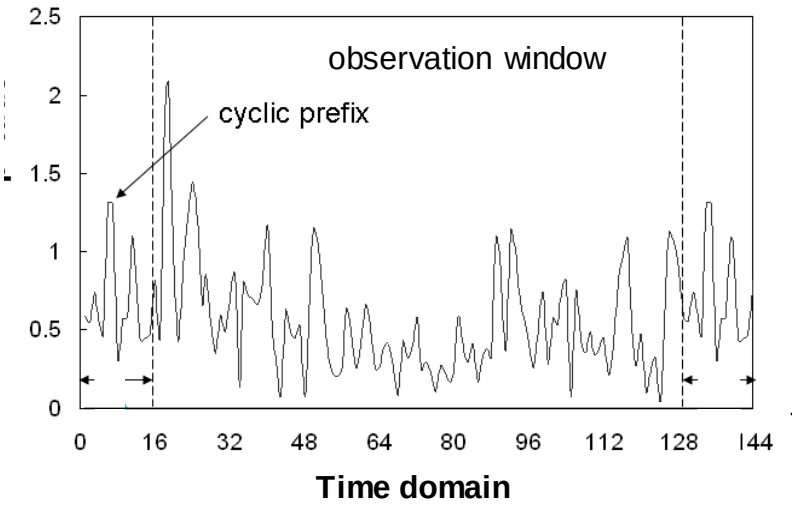

Fig. 2.5 OFDM symbol in time domain.

2.4 Signal processing for CO-OFDM

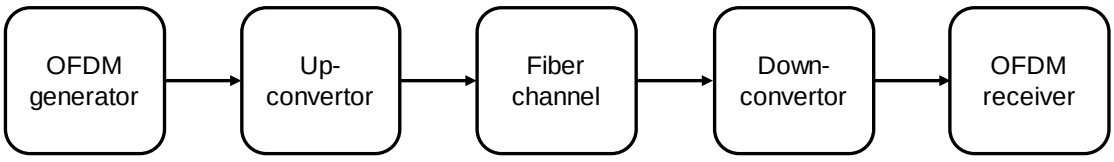

Fig. 2.6 Functional blocks of a CO-OFDM system.

As shown in Fig. 2.6, a CO-OFDM system is constructed by five functional blocks, including the RF OFDM generator, RF-to-optical (RTO) up-converter, optical channel, optical-to-RF (OTR) down-converter and RF OFDM receiver [75-78,85]. The RTO

up-convertor and OTR down-convertor have been discussed in Section 2.2, which achieve linear conversions between the baseband and optical signal. The RF OFDM generator and receiver will be discussed in the following sections, including the DSP needed at the generator and receiver.

2.4.1 Signal generation

To generate the baseband OFDM signal with RF OFDM generator, the data bits are first mapped to complex signals using modulation schemes such as QAM or PSK. The OFDM symbol consisting of N_{sc} subcarriers is then converted to time domain using IDFT expressed in Eq. (2.16). After that, the guard interval is inserted to the time domain OFDM symbol to prevent the ISI due to the channel dispersion. The digital signal is converted to the analogue one using digital-to-analogue converter (DAC), and the alias signal is filtered out by a low-pass filter (LPF). The obtained analogue signal is up-converted to the optical field using the IQ modulator described in section 2.2.1. The up-converted optical signal can be expressed as

$$E(t) = s_B(t) \cdot \exp\left(j\omega_{LD1} t + \phi_{LD1} \right), \tag{2.19}$$

where $s_B(t)$ is the baseband OFDM signal, ω_{LD1} and ϕ_{LD1} denote the angular frequency and the phase of the laser source.

2.4.2 Signal recovery

The linear effect of the fiber channel can be described by an impulse response $h(t)$. After the propagation in optical fiber medium, the optical signal becomes as

$$E(t) = \exp\left(j\omega_{LD1} t + \phi_{LD1} \right) s_B(t) \otimes h(t), \tag{2.20}$$

where '$\otimes$' stands for convolution.

After the down-conversion of the optical signal using coherent detection, the obtained RF signal can be expressed as

$$r(t) = r_0(t) \cdot \exp\left(j\omega_{off} t + \Delta\phi \right), \tag{2.21}$$

$$r_0(t) = s_B(t) \otimes h(t), \tag{2.22}$$

$$\omega_{off} = \omega_{LD1} - \omega_{LD2}, \quad \Delta\phi = \phi_{LD1} - \phi_{LD2}, \tag{2.23}$$

where $r_0(t)$ is the RF signal with the linear impact from the fiber channel, $\Delta\omega_{off}$ is the angular frequency, and $\Delta\phi$ is the phase offset between transmit and receive lasers. The OFDM signal is then sampled with an analogue-to-digital converter (ADC) at RF OFDM receiver. The sampled signal is processed with advanced DSP to remove the linear impact from fiber channel and recover the transmitted information.

DSP has become one of the most powerful tools in modern optical transmission system. With the improved speed and reliability of the integrated circuits, DSP has been widely applied in coherent transmission to ensure the high performance of the transmission system. In the meantime, CO-OFDM is able to provide the easiness for the DSP by using the pilot symbols and pilot subcarriers. The signal processing for CO-OFDM can be divided into three parts: (i) FFT window synchronization where OFDM symbol is properly delineated to avoid ISI, (ii) frequency synchronization, namely, frequency offset ω_{off} is estimated and compensated, and (iii) the subcarrier recovery, where the linear channel impact on each subcarrier is estimated and compensated [75-78]. The details are provided below.

- **FFT windows synchronization**

As shown in Fig. 2.5, an OFDM symbol consists of guard interval and observation window. The start point of the observation window has to be determined properly to have an improper DFT window and avoid the ISI. A popular window synchronization method has been proposed by Schmidl and Cox, which uses a pilot or preamble consisting of two identical segments as

$$S_m = S_{m+N_{sc}/2}, \qquad m \in [1, N_{sc}/2] \tag{2.24}$$

where S_m is the m-th digital signal with random value. The correlation function is calculated for the received pilot with shifted start point d, which can be expressed as

$$R_d = \sum_{m=1}^{N_{sc}/2} r_{m+d}^{*} r_{m+d+N_{sc}/2}, \tag{2.25}$$

where r_{m+d} is the received signal at point $m+d$. Since the second half of the received pilot symbol is identical to the first half except for a phase shift, the correlation function achieves its maximum value when $r_{d+1} = S_1$. Therefore, the correct time window is found.

- **Frequency synchronization**

The frequency offset in the received OFDM signal can break the orthogonality between the subcarriers, inducing ICI penalty. Frequency synchronization can be accomplished following two steps: frequency acquisition and frequency tracking. Frequency acquisition is used for coarse estimation of the frequency offset, which brings the receiver laser close to the center frequency, approximately within several times of the subcarrier frequency spacing, making the following signal processing easier. A convenient approach for frequency acquisition is to send CW and measure the offset tone directly. For more accurate identification of the frequency offset, frequency tracking is needed. The Schmidl and Cox approach used for window synchronization can also be used for accurate frequency tracking. We can rewrite the correlation function of Eq. (2.25) as

$$R_d = \sum_{m=1}^{N_{sc}/2} \left| r_{m+d} \right|^2 e^{j\pi f_{off}/\Delta f} . \tag{2.26}$$

Therefore, the frequency offset can be estimated by

$$f_{off} = \frac{\Delta f}{\pi} \arg(R_d) . \tag{2.27}$$

- **Subcarrier recovery**

After the successful DFT window synchronization and frequency synchronization, the sampled OFDM signal of Eq. (2.21) is converted to the frequency domain through the DFT operation. The rotation of the constellation c_{ki} at *k*-th subcarrier and *i*-th symbol is contributed by three factors: the channel dispersion, the DFT sampling timing offset, and the phase noise. The first two factors can be represented by the channel transfer function h_{ki}, and the third factor can be represented by the common phase error (CPE) ϕ_i. Therefore, the recovered OFDM signal r_{ki} after DFT becomes

$$r_{ki} = e^{\phi_i} h_{ki} c_{ki} + n_{ki} , \tag{2.28}$$

where n_{ki} is the noise. We assume that the channel transfer function h_{ki} is invariant within the OFDM block, and the laser phase noise ϕ_i varies based on the OFDM symbol.

Various approaches can be used for channel estimation. The following content describes a frequency domain pilot-assisted approach, where pilot symbols are added at the beginning of the OFDM symbols for the channel estimation. With the knowledge of the transmitted pilot symbols, the channel transfer function can be estimated as

$$h_k = \sum_{i=1}^{p} e^{-j\angle r_{ki}} \frac{r_{ki}}{c_{ki}}, \tag{2.29}$$

where c_{ki} and r_{ki} are the transmitted and received pilot subcarriers respectively, $\angle r_{ki}$ is the common of the k-th carrier in the i-th OFDM symbol, and p is the number of the pilot symbols.

The laser phase noise ϕ_i in i-th symbol can be estimated using the pilot subcarriers. Assuming N_p subcarriers in i-th symbol are used for the phase estimation, the maximum likelihood CPE can be estimated as

$$\phi_i = \arg\left(\sum_{k=1}^{N_p} \frac{c_{ki}^{'} h_k^* c_{ki}^*}{\delta_k^2} \right), \tag{2.30}$$

where δ_k is the standard deviation of the constellation spread.

Once h_k and ϕ_i are known, an estimated value of c_{ki} is given by zero-forcing method as

$$\hat{c}_{ki} = \frac{h_k^*}{|h_k|^2} e^{-j\phi_i} r_{ki}. \tag{2.31}$$

$\hat{c}_{ki}$ is then used for the symbol decision, which is subsequently mapped back to digital bits.

2.5 PDM supported transmission

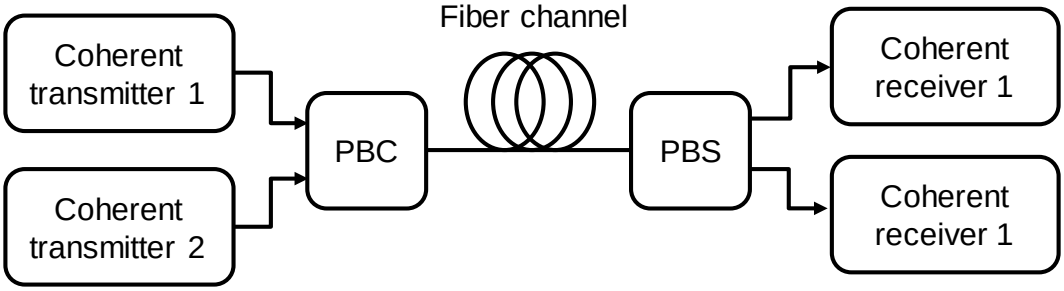

Fig. 2.7 conceptual diagram of PDM supported coherent transmission.

Polarization-division multiplexed (PDM) signal can be easily transmitted in a coherent system, where the linear channel leads to an efficient compensation of polarization related effects, such as polarization mode dispersion (PMD) and polarization dependent loss (PDL). By transmitting the signal in two orthogonal polarization modes, the capacity of the single-mode fiber (SMF) can be doubled, contributing to a higher spectral efficiency. Fig. 2.7 shows the conceptual diagram of the PDM supported coherent transmission [86-88]. Two coherent transmitter are used to generate the signals for two polarization modes. The generated signals are combined together using a polarization beam combiner (PBC) before entering into the fiber. At the receiver, the two polarization modes are separated again with a polarization beam splitter (PBS) and detected by the coherent receivers individually.

In presence of the polarization-dependent effects, such as PMD and PDL, the two polarization modes travel at different group velocity, and experience different loss during the fiber transmission. Due to the random cross talk between the two polarization modes, a fiber channel is usually described as a two-input-two-output (TITO) model for the polarization-division-multiplexed transmission [87]. CO-OFDM based polarization-division-multiplexing is discussed in the following.

Extending single-polarization transmission to dual-polarization transmission, the OFDM signal of Eq. (2.12) is extended from 1-dimension value to 2-dimension Jones vector $\vec{\boldsymbol{s}}(t)$

$$\vec{\boldsymbol{s}}(t) = \sum_{i=-\infty}^{+\infty} \sum_{k=-N_{sc}/2+1}^{N_{sc}/2} \vec{\boldsymbol{c}}_{ik} \prod (t - iT_s) e^{j2\pi f_k (t - iT_s)} \tag{2.32}$$

$$\vec{\boldsymbol{s}}(t) = \begin{bmatrix} s_x \\ s_y \end{bmatrix}, \quad \vec{\boldsymbol{c}}_{ik} = \begin{bmatrix} c_{ik}^x \\ c_{ik}^y \end{bmatrix} \tag{2.33}$$

$$f_k = \frac{k-1}{T_s} \tag{2.34}$$

$$\prod(t) = \begin{cases} 1, (0 < t < T_s) \\ 0, (t \le 0, t > T_s) \end{cases}, \tag{2.35}$$

where s_x and s_y are the x and y polarization components of the transmitted OFDM signal $\vec{\boldsymbol{s}}(t)$, $\vec{\boldsymbol{c}}_{ik}$ is the OFDM signal in frequency domain at the k-th subcarrier of the

i-th symbol with the *x* and *y* polarization components, c_{ki}^{x}, c_{ki}^{y}, $\Pi(t)$ is the pulse waveform of the symbol, f_k is the frequency of the *k*-th subcarrier, and T_s is the OFDM symbol period. The optical field of the up-converted optical OFDM signal is

$$E(t) = \vec{\boldsymbol{s}}(t) e^{j(\omega_{LD1} t + \phi_{LD1})}, \tag{2.36}$$

where ω_{LD1} and ϕ_{LD1} are the angular frequency and the phase of the laser source. Considering the impulse response $h(t)$ of the optical medium, the received optical signal $E_r(t)$ becomes

$$E_r(t) = \vec{\boldsymbol{r}_0}(t) e^{j(\omega_{LD1} t + \phi_{LD1})} \tag{2.37}$$

$$\vec{\boldsymbol{r}_0}(t) = \vec{\boldsymbol{s}}(t) \otimes h(t), \tag{2.38}$$

where $\otimes$ stands for convolution. After the down-conversion of the optical signal using the polarization diversity receiver and I/Q detection, the obtained baseband signal is given by

$$\vec{\boldsymbol{r}}(t) = \vec{\boldsymbol{r}_0}(t) e^{j(\omega_{off} t + \Delta\phi)}, \tag{2.39}$$

where ω_{off} and $\Delta\phi$ are the frequency and phase difference between the transmitter laser and LO.

The procedures of DFT window synchronization and frequency synchronization for PDM signal are the same as that for the single-polarization signal described in section 2.4. After DFT, the OFDM signal in frequency domain is given by

$$\vec{\boldsymbol{c}}_{ki}^{'} = e^{j\phi_i} \cdot \boldsymbol{H}_k \cdot \vec{\boldsymbol{c}}_{ki} + \vec{\boldsymbol{n}}_{ki} \tag{2.40}$$

$$\boldsymbol{H}_k = e^{j\Phi_D(f_k)} \cdot \boldsymbol{T}_k \tag{2.41}$$

$$\boldsymbol{T}_k = \prod_{l=1}^{N} \exp\left\{ \left(-\frac{1}{2} j \vec{\boldsymbol{\beta}}_l f_k - \frac{1}{2} \vec{\boldsymbol{\alpha}}_l \right) \vec{\boldsymbol{\sigma}} \right\} \tag{2.42}$$

$$\Phi_D(f_k) = \pi \cdot c \cdot D_t \cdot f_k^2 / f_{LD1}^2, \tag{2.43}$$

where $\vec{\boldsymbol{c}}_{ki}^{'}$ is the received signal in form of Jones vector at the *k*-th subcarrier in the *i*-th OFDM symbol with *x* and *y* polarization components, $c_{ki}^{'x}$ and $c_{ki}^{'y}$, $\vec{\boldsymbol{n}}_{ki}$ is the noise with the *x* and *y* polarization components n_{ki}^{x} and n_{ki}^{y}. ϕ_i is the OFDM symbol laser

phase noise, $\Phi_D(f_k)$ is the phase dispersion due to the fiber chromatic dispersion, $\boldsymbol{T}_k$ is the 2x2 Jones matrix for the fiber link, N is the number of cascading elements of PMD/PDL, $\vec{\boldsymbol{\beta}}_l$ is the birefringence vector, $\vec{\boldsymbol{\alpha}}_l$ is the PDL vector. and $\vec{\boldsymbol{\sigma}}$ is the Pauli matrix vector.

Different from the single-polarization transmission, the optical link in dual-polarization transmission is described by a 2x2 channel matrix $\boldsymbol{H}_k$. Special pilot symbols are designed for the estimation of the 2x2 channel matrix, where the 2m-th symbols are set to zeros for the pilots in x polarization, while the (2m+1)-th symbols are set to zeros for the pilots in y polarization (m=0 to half of the number of pilot symbols). As a result, the 2m-th pilot symbol has a form of $[0\ S_y]$, and the (2m+1)-th pilot symbol has a form of $[S_x\ 0]$. The received signals corresponding to the pilot symbols are

$$R_{2m} = S_y \begin{bmatrix} h_{12} \\ h_{22} \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} 0 \\ S_y \end{bmatrix} \tag{2.44}$$

$$R_{2m+1} = S_x \begin{bmatrix} h_{11} \\ h_{21} \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} S_x \\ 0 \end{bmatrix}, \tag{2.45}$$

where h_{11}, h_{12}, h_{21}, h_{22} are the four elements of the channel matrix $\boldsymbol{H}_k$. Therefore, the first column of the channel matrix can be estimated from the (2m+1)-th pilot symbols, and the second column of the channel matrix can be estimated from the 2m-th symbols of the pilot symbols. The laser phase noise is estimated using the pilot subcarriers similar to the case of single-polarization. After the channel matrix $\boldsymbol{H}_k$ and laser phase noise ϕ_i are estimated, the subcarriers can be recovered as

$$\vec{\boldsymbol{c}}_{ki} = \frac{\boldsymbol{H}_k^*}{\left|\boldsymbol{H}_k\right|^2} e^{-j\phi_i} \vec{\boldsymbol{c}}'_{ki}. \tag{2.46}$$

2.6 Conclusion

In this chapter, we review the high-performance coherent transmission system, where a linear channel is obtained with the help of coherent transmitter and receiver. Detailed discussion is provided for the CO-OFDM system, including the signal generation and the DSP procedure to remove the channel impact and recover the

signal. The PDM supported coherent transmission is also introduced, doubling the capacity of the fiber channel.

3 Space-Division-Multiplexing for Ultra-High-Capacity Optical Transmission

3.1 Introduction

The capacity of the optical networks has witnessed a huge growth during the last few decades, due to the exploration of the advanced optical communication technologies. To sustain the capacity increase, various multiplexing techniques have been intensely investigated to make full utilization of the available modulation freedoms in fiber channels. Nowadays, commercial coherent optical transmissions are able to multiplex the signal in dimensions of time, wavelength, phase and polarization, in order to pack up more information into a single fiber [22-30]. Using the multiplexing techniques, the optical system achieves a higher data rate with reduced cost per bit, making the multiplexing scheme a favorable choice when upgrading the capacity. With the development of various multiplexing technologies, the available freedom in single-mode fiber (SMF) has been maximum explored. As a result, the capacity of SMF is approaching its fundamental limit quickly [56]. To keep up with the increasing demand of data traffic, an option is to install new fibers to transmit more information through parallel channels. The cost will increase linearly with the capacity using parallel channels. The increased cost will undoubtedly put a pressure on the capacity growth in the future.

Another option is to find a new dimension for the signal multiplexing. Space-division-multiplexing (SDM) uses the space mode as an additional degree of modulation freedom, sustaining a constant growth of the channel capacity [57-64]. Different from the SMF, a multi-mode or multi-core fiber can support the propagation of multiple spatial modes. The different spatial modes can be treated as independent channels to transmit the signal simultaneously. Compared to the parallel channels, SDM can increase the channel capacity in a cost efficient way. The fiber capacity is greatly increased with different spatial modes closely packed in one fiber link. Meanwhile, the cost per bit is significantly reduced, since the cost of cable and its installation is averaged by the large capacity. A huge amount of effort has been devoted to develop the enabling technologies for the SDM transmission. With the innovations of the high-performance multi-mode fibers (MMFs), multiplexing and

demultiplexing components, SDM has become a preferred solution for the high-capacity transmission [89-91].

The crosstalk between the different spatial modes is an important issue in SDM system, which is also mentioned as modal dispersion (MD) [92]. Coherent transmission provides a solution for overcome this problem by compensating the MD using multiple-input multiple-output (MIMO) signal processing. Similar to the polarization mode dispersion (PMD) in SMF which can be compensated using 2x2 MIMO with coherent transmission, the MD in multi-mode fiber (MMF) can be compensated by 2*N*x2*N* MIMO, where *N* is the number of spatial modes supported in fiber. The signal processing becomes complicated when *N* is large. Therefore, few-mode fiber (FMF) based optical transmission is proposed [93-96]. With a smaller number of spatial modes, FMF can achieve a data rate beyond that of SMF while keeping the signal processing to a manageable level [97-100].

The FMF channel can be described by a channel matrix, which plays a critical role in receiver design and determines the ultimate performance of the FMF system. An in-depth understanding of the statistical properties of fiber channel subject to DGD is essential to assist the system design and achieve an optimal system performance. In this chapter, we will first review the current research on FMF transmission, including the fundamentals of FMFs, as well as the mode multiplexing and demultiplexing components. Then, we will investigate into the of the stochastic properties of FMF channel. The frequency dependence of channel matrix will be studied in the form of autocorrelation function (ACF), indicating the divergence of the channel changes in frequency domain. Analytical form of the ACF will be developed from the canonical stochastic differential equation (SDE) for the FMF channel matrix in the regime of strong coupling. The time evolution of the channel matrix will also be studied by measuring the fluctuation speed of MD. High-dimensional Stokes-space analysis is introduced into the measurement to eliminate the errors caused by laser phase noise.

3.2 Fundamentals of few-mode fiber

The guided modes in an optical fiber are referred to the modes that can be well-confined in the core area, and propagate for a long distance with low loss. The number of guided modes that can be supported in a fiber is dependent on the fiber design. The mode condition in a fiber is usually described by the normalized frequency parameter, which is also called as *V*-number [101-103]. The *V*-number is defined as

$$V = k_0 a\sqrt{n_1^2 - n_1^2}\,, \tag{3.1}$$

where n_1 and n_2 are the refractive indexes of the core and cladding layers, a is the radius of the core, $k_0 = 2\pi/\lambda$ is the wavenumber, and λ is the wavelength of the optical signal. For a step-index fiber with a *V*-member smaller than 2.405, only one guided mode is supported in the fiber. Therefore, it is called as single-mode fiber [104,105]. The core diameter of SMF is typically smaller than 10 μm [106]. When the core is enlarged, the *V*-number increases as well. As a result, more than one guided modes can be supported in the fiber. The MMF usually has a core diameter ranging from 50 μm to 65 μm, supporting the propagation of multiple spatial modes [107]. The channel capacity is enlarged by *N* times when *N* guided modes are propagated in the fiber. To recover the modes after the random mode coupling during the propagation, 2*N* by 2*N* MIMO DSP is needed at the receiver. To control the DSP complexity to a manageable level, FMF has been proposed, with a core diameter slightly larger than SMF but much smaller than MMF, leading to a small mode number *N*. The simplest case of FMF is two-mode fiber (TMF), which has a *V*-number above single-mode condition but below three-mode condition, 2.4<V<3.8. TMF supports the propagation of two spatial modes, LP_{01} and LP_{11}. Since LP_{11} mode includes two orthogonal modes, LP_{11a} and LP_{11b}, a TMF actually supports the propagation of 6 modes, considering the two polarization modes for each spatial mode [108-111]. By lighting all 6 modes, the channel capacity can be improved by 3 times compared to the SMF transmission.

The TMF used in our lab has the parameters summarized in Table. 3.1 [112-115]. The Ge-doped TMF has a step-index profile with a core diameter of 11.9 μm and refractive index step of 5.4×10^{-3}. The *V*-number of the TMF is calculated to be 3.62, which fits into the two-mode condition, 2.4<*V*<3.8. The chromatic dispersion of LP_{11} mode is measured to be 17 ps/nm·km, which is similar to that of the SMF, while the chromatic dispersion of LP_{01} mode is measured to be 22.1 ps/nm·km, slightly higher than the SMF. With the fiber parameters, the modal index is simulated for different wavelength. As shown in Fig. 3.1, three spatial modes are supported in the transmission window around 1550 nm, including the fundamental mode LP_{01}, and two degenerate modes, LP_{11a}, LP_{11b}. The degenerate mode LP_{11} is cut off at the wavelength of 2323 nm, so the fiber can be used as a SMF after the wavelength of

2323 nm. Great difference in the effective refractive index is observed between the modes, LP_{01} and LP_{11}. As a result, the two modes are subject to a small mode coupling and a large modal dispersion (MD) of 3 ps/m. The loss of the fiber is measured to be 0.26 dB/km.

Table 3.1 Parameters of TMF

	Value	**Unit**
Spool length	4500	m
Core diameter	11.9	µm
Cladding diameter	109	µm
Dispersion LP_{01}	22.1	ps/(nm*km)
Effective area LP_{01}	94.7	μm^2
Dispersion LP_{11}	17	ps/(nm*km)
Effective area LP_{11}	99.9	μm^2
MD LP_{01}-LP_{11}	3.0	ps/m
LP_{01} Cut-off Wavelength	2323	nm
Fibre Loss(LP_{01},LP_{11})	0.26	dB/km

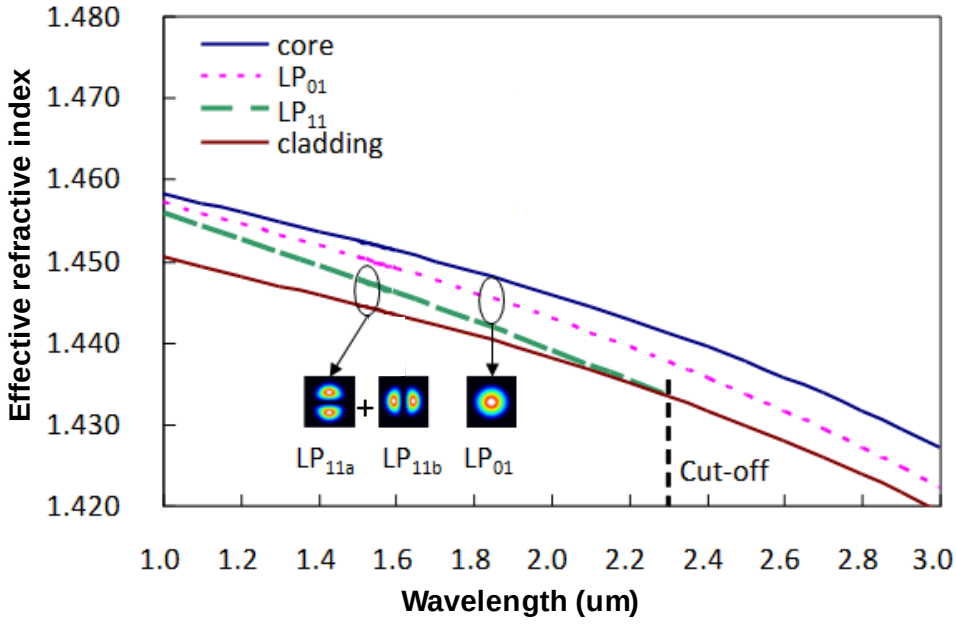

Fig. 3.1 The modal index of the TMF vs. wavelength.

3.3 Mode-multiplexing/demultiplexing components

For a FMF with *N* spatial modes, a total of 2*N* modes can be used for signal transmission, considering two polarization modes in each spatial mode. Fig. 3.2 shows the schematic diagram of the FMF transmission [112-115]. The transmitted signals are generated with 2*N* transmitters as signals for 2*N* independent channels. The two polarization modes are first combined using polarization beam combiner (PBC). Then, the signals are converted to the corresponding spatial modes and combined using a mode-multiplexing before launched into the FMF. After the transmission, the spatial modes are separated using a mode-demultiplexing component, and the high-order spatial modes are converted back to the fundamental mode. The polarization modes are separated by the polarization beam splitter (PBS), followed by coherent receivers to detect the signal in each mode. 2*N*x2*N* MIMO DSP is used to compensate the linear channel impairments, and recover the signals [116]. High-performance mode converter and mode multiplexing/demultiplexing device are the critical components in a FMF transmission system.

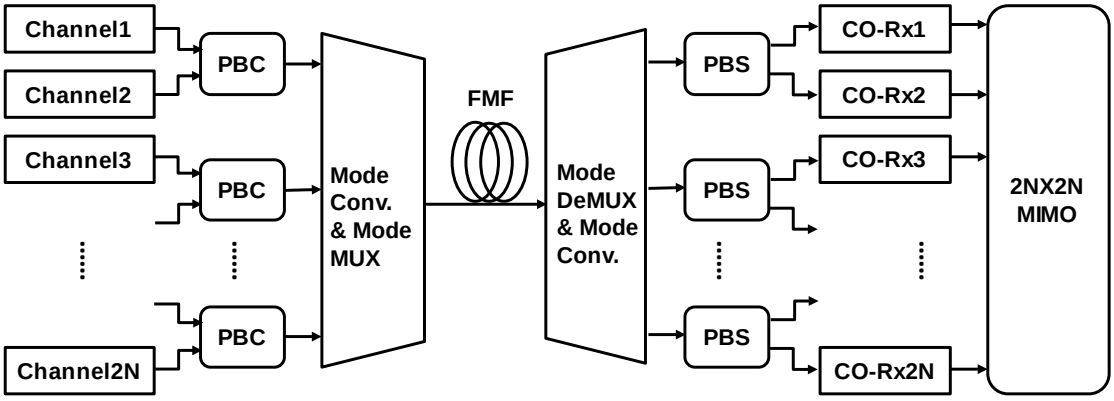

Fig. 3.2 Schematic diagram of FMF transmission.

Fig. 3.3 shows the free-space based setup to convert and multiplex the LP_{11} mode to a TMF. The transmitter generated the signal in fundamental mode. To convert the fundamental mode to a higher-order mode, SMF is spliced with TMF to transfer the power into TMF. A mode striper is then used to strip out the higher-order mode in TMF, leaving fundamental mode in the fiber. After that, mode converter is used to convert the fundamental mode into desired mode, LP_{11}. Modes LP_{11a} and LP_{11b} are combined together using a mode combiner. The mode combiner is carefully aligned to reduce the insert loss and launch the signal into the correct mode of the transmission fiber. The mode demultiplexing and converting at the receiver side can be achieved

using a setup constructed in an inverted order of Fig. 3.3. The mode striper, mode converter and mode combiner are discussed in detail below.

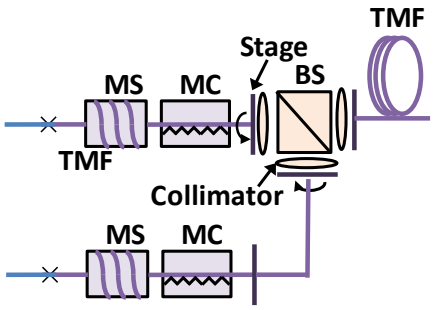

Fig. 3.3 Mode converting and multiplexing setup.

- **Mode striper**

Mode striper is fabricated based on the mechanism of macro-bending [117]. By tightly winding the TMF around a post with diameter of 8 mm for more than 20 turns, the higher-order mode, such as LP_{11} mode, experiences a huge attenuation, while the attenuation for the fundamental mode, LP_{01}, is smaller than 0.2 dB. A rejection ratio higher than 30 dB is obtained for LP_{11} mode. Therefore, the LP_{11} mode is stripped out, leaving only the LP_{01} mode.

- **Mode converter**

Mode converter can be obtained using different material and structures, such as long-period fiber grating (LPFG), spatial light modulator (SLM), free-space phase plate, or fused spatial mode coupler (SMC). In this section, we describe a LPFG based mode converter, where a grating structure is built in the fiber. Resonant coupling is induced between two spatial modes when the grating pitch Λ equals to the beating length of the two modes [118,119]. We use the TMF with parameters listed in Table 3.1 to build a mode converter for modes LP_{01} and LP_{11} modes. A metal grating with evenly-spaced grooves is pressed against the TMF, leading to the deformation of the fiber core. Therefore, the LPFG is built in the fiber. In order to decide the groove pitch for the metal grating, we measured the MD of the TMF. The beating length for modes LP_{01} and LP_{11} is expressed as [120-122]

$$L_B = 2\pi / (\beta_{01} - \beta_{11}) \qquad ,(3.2)$$

where β_{01} and β_{11} are the propagation constants of modes LP_{01} and LP_{11}. With the MD of 3 ns/km, the mode beating length is calculated to be 524 μm, which

corresponds to the groove pitch for the metal grating. The core deformation has an s-bend arc shape. With the core deformation radius of 0.2 μm, the coupling length of 8.1 mm is needed for the mode converter to achieve its maximum coupling efficiency. An extinction ration over 20 dB is measured around the wavelength of 1550 nm for a wavelength range more than 10 nm. The metal grating and TMF are mounted on a 3-axis stage to adjust the angle and pressure applied between the metal grating and fiber. After careful adjustment, we achieve the maximum extinction ratio at the wavelength of 1550 nm.

- **Mode combiner**

Fig. 3.4 shows a free-space based mode combiner to combine three spatial modes, LP_{11a}, LP_{11b} and LP_{01}. The combiner consists of three XY translation stages, two 50:50 beam splitters and four collimating lens. The three XY translation stages are connected to the input fiber, through which the modes LP_{11a}, LP_{11b} and LP_{01} are launched into the setup. The fiber ends are connected with FC type adapters. The orientations of the LP_{11a} and LP_{11b} modes are adjusted to be orthogonal to each other by rotating the fiber adapters axially. The input modes are then collimated to a spot size of 2 mm by the three collimating lens with numerical aperture NA of 0.25 and effective focal length f of 11.0 mm. The collimated beams pass through the beam splitters in either transmission or reflection direction, and arrive at the output collimator. The beam spot is focused by the output collimator and launched into the core of the output fiber. To achieve the maximum coupling ratio, the three input XY translation stages are carefully aligned to put the beam spots to the correct position. The free-space beam combiner is polarization insensitive. The minimum loss is 3 dB for the LP_{01} mode, and 6 dB for the LP_{11} modes. The setup can also be used as mode splitter by exchanging the input and output.

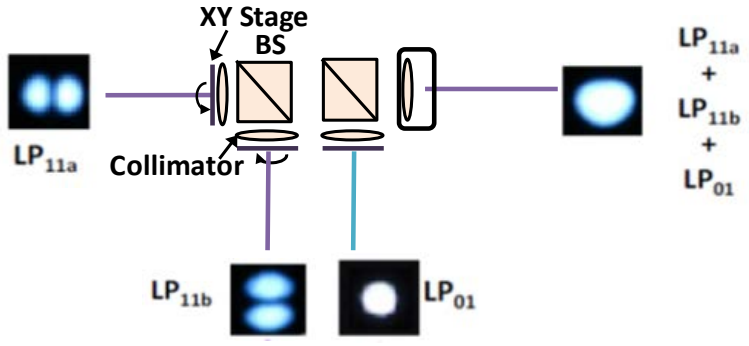

Fig. 3.4 Mode combiner for three spatial modes LP_{11a}, LP_{11b} and LP_{01}.

3.4 FMF channel matrix description and decomposition

For FMF transmission, the mode coupling caused by fiber index imperfections and environmental perturbations lead to signal dispersion, called differential group delay (DGD) [92]. While large DGD can improve the channel diversity and makes the system robust again mode dependent loss (MDL), excessive DGD can increase the computational complexity of the receiver DSP. An in-depth understanding of the statistical properties of channel subject to DGD is essential to assist system design for achieving an optimal system performance.

With respect to SMF, there exists a large body of study on the stochastic properties of polarization mode dispersion (PMD) in terms of both PMD vector in Stokes space [123,124] and channel matrix in Jones space [125,126]. Though Stokes space and Jones space are isomorphic [127], PMD vector is more suitable for the analysis of time-domain direct detection systems, whereas Jones matrix is more useful for the analysis of coherent detection systems. Based on white Gaussian noise model for the birefringence vector of strong coupling SMF, some important statistical properties for PMD vector and Jones matrix could be derived using standard tools of stochastic calculus [123-126].

The basic concepts and statistical study methods for SMF can be extended to FMF. Since each spatial mode supports two orthogonal polarization modes, FMF with N spatial modes supports totally $2N$ modes with different group velocities, which can be represented by a $2N$-dimensional vector. Then, the linear effect of FMF channel on the propagating modes can be described as a $2N$ x $2N$ channel matrix in a generalized Jones space, and the modal dispersion (MD) vector, a general form of PMD of SMF, is a real vector of dimension $4N^2-1$ in a generalized Stokes space [128,129]. As long as the modes are strongly coupled with each other during the propagation in FMF, the white Gaussian noise assumption would still be valid for modeling the mode coupling vector, a general form of birefringence vector of SMF [128,129]. Recently, some studies have been conducted on the statistical properties of MD vector and group delays of FMF [130,131]. On the other hand, the latest decade has seen the revival of coherent detection, where DSP is performed on channel matrix of fiber optic link [112-114]. In order to describe the properties of channel matrix, we decompose it over the generalized high-dimensional Gell-Mann matrices, an equivalent of the 2-

dimensional Pauli matrices for SMF. The decomposition of the FMF channel matrix provides a convenient tool for the further analysis of channel property.

The main notations used below are as follows: $|V\rangle$ denotes a column vector in generalized Jones space of dimension 2*N*, $\vec{V}$ denotes a column vector in generalized Stokes space of dimension $4N^2-1$, $\tilde{V}$ denotes a column vector of dimension $4N^2$, a column vector with elements v_i is denoted by $[\cdots v_{i-1}; v_i; v_{i+1}\cdots]$; Symbol '†' represents the transpose conjugate operation for a matrix, and 'T' represents the transpose operation; *E*[ ] denotes the average of an ensemble, and *Tr*{ } denotes the trace of a matrix.

The optical field in an *N*-mode FMF at length *z* and frequency ω can be expressed as a 2*N*-dimensional complex vector $|\psi(z,\omega)\rangle$. The linear channel of FMF can be treated as a *2N* input *2N* output linear system given by [128,129]

$$|\psi(z,\omega)\rangle = U(z,z_0,\omega)|\psi(z_0,\omega)\rangle, \tag{3.3}$$

where *U* is a channel matrix of dimension 2*N* x 2*N*, describing the random mode coupling and group delays in optical channel. Without consideration of MDL, the channel matrix *U* is a unitary matrix with unit determinant.

Similar to the Pauli matrices used as the basis for 2x2 Jones matrix of SMF [126], the basis for a 2*N* x 2*N* channel matrix includes an identity matrix $\Lambda_0 = I$, and $4N^2-1$ traceless Hermitian matrices Λ_i (*i*=1 to $4N^2-1$), satisfying the trace-orthogonal condition [132]

$$Tr\{\Lambda_i \Lambda_j\} = 2N\delta_{ij}, \tag{3.4}$$

where δ_{ij} is the Kronecker symbol. The channel matrix *U* could be expressed as a combination of the basis matrices

$$U = \sum_{i=0}^{4N^2-1} u_i \Lambda_i = u_0 I + \vec{u}\cdot\vec{\Lambda}, \tag{3.5}$$

where the vector $\vec{\Lambda} = [\Lambda_1; \Lambda_2; \cdots; \Lambda_{4N^2-1}]$ is an ensemble of the traceless Hermitian matrices, the complex coefficient u_0 is the weight of the identity matrix in *U*, and the complex vector $\vec{u} = [u_1; u_2; \cdots; u_{4N^2-1}]$ is an ensemble of the weights of each traceless Hermitian matrix. Since one set of weights u_i can exclusively identify a matrix *U*, the

properties of a 2*N* x 2*N* channel matrix *U* is fully represented by a vector $\tilde{u} = [u_0; \vec{u}]$ of dimension $4N^2$. Using the trace-orthogonal condition of Eq. (3.4), the weights u_i of basis matrices Λ_i could be easily extracted by

$$u_i = \frac{1}{2N} Tr\{\Lambda_i U\}, \tag{3.6}$$

In group theory, the traceless Hermitian matrices Λ_i (*i*=1 to $4N^2 - 1$) form the generators of special unitary group of degree 2*N* (SU(2*N*)). At the same time, they are a basis for the Lie algebra of SU(2*N*) satisfying [133]

$$\Lambda_m \Lambda_n = \delta_{mn} I + \sum_k (jf_{mnk} + d_{mnk}) \Lambda_k \tag{3.7}$$

$$\Lambda_n \Lambda_m = \delta_{mn} I + \sum_k (-jf_{mnk} + d_{mnk}) \Lambda_k, \tag{3.8}$$

where the indices *k*, *m*, *n* take values from 1 to $4N^2 - 1$, and the coefficients *f* and *d* are the structure constants. With the structure constants *f* and *d*, a cross product of vectors $\vec{A}$ and $\vec{B}$ in generalized Stokes space can be defined as

$$\vec{A} \otimes \vec{B} = \sum_{mnk} f_{mnk} A_m B_n \vec{e}_k = -\vec{B} \otimes \vec{A}, \tag{3.9}$$

where A_m is the *m*-th element in vector $\vec{A}$, B_n is the *n*-th element in vector $\vec{B}$, and $\vec{e}_k$ is a unit vector with 1 for the *k*-th element. A dot product can be defined as

$$\vec{A} \Theta \vec{B} = \sum_{mnk} d_{mnk} A_m B_n \vec{e}_k = \vec{B} \Theta \vec{A}. \tag{3.10}$$

Then, the matrices $\vec{A} \cdot \vec{\Lambda}$ and $\vec{B} \cdot \vec{\Lambda}$ satisfy the relations

$$(\vec{A} \cdot \vec{\Lambda})(\vec{B} \cdot \vec{\Lambda}) - (\vec{B} \cdot \vec{\Lambda})(\vec{A} \cdot \vec{\Lambda}) = 2j(\vec{A} \otimes \vec{B}) \cdot \vec{\Lambda} \tag{3.11}$$

$$(\vec{A} \cdot \vec{\Lambda})(\vec{B} \cdot \vec{\Lambda}) + (\vec{B} \cdot \vec{\Lambda})(\vec{A} \cdot \vec{\Lambda}) = 2(\vec{A} \cdot \vec{B}) I + 2(\vec{A} \Theta \vec{B}) \cdot \vec{\Lambda}. \tag{3.12}$$

Adding Eqs. (3.11) and (3.12) leads to

$$(\vec{A} \cdot \vec{\Lambda})(\vec{B} \cdot \vec{\Lambda}) = (\vec{A} \cdot \vec{B}) I + (\vec{A} \Theta \vec{B} + j\vec{A} \otimes \vec{B}) \cdot \vec{\Lambda}. \tag{3.13}$$

The choices for the trace-orthogonal matrices are multiple. The only guide for the choice is the simplicity of calculation. We choose the generalized Gell-Mann matrices as the basis matrices to decompose the channel matrix. The *n*-dimensional Gell-Mann matrices are constructed by the following algorithm:

The first *n*-1matrices are diagonal with the form

$$\Lambda_k = \sqrt{\frac{n}{k(k+1)}} \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & -k & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix} \quad (k=1 \text{ to } n\text{-}1). \tag{3.14}$$

where the first k diagonal entries are 1, and the k+1th diagonal entry is $-k$. The rest entries are all 0. The coefficient $\sqrt{n/(k(k+1))}$ is added to satisfy the trace-orthogonal condition $Tr\{\Lambda_i \Lambda_j\} = n\delta_{ij}$.

The rest $n^2 - n$ Gell-Mann matrices are off-diagonal. Half of them are symmetric

$$\Lambda_{pq} = \begin{bmatrix} & & 1 & \\ & & & \\ 1 & & & \\ & & & \end{bmatrix}, \tag{3.15}$$

where the entries at the p-th row, q-th column and the q-th row, p-th column are 1, and the rest entries are 0. The other half are antisymmetric

$$\Lambda_{pq} = \begin{bmatrix} & & -i & \\ & & & \\ i & & & \\ & & & \end{bmatrix}, \tag{3.16}$$

where the entry at the p-th row and q-th column is $-i$, the entry at the q-th row and p-th column is i, and the rest entries are 0. The coefficient for the off-diagonal matrices is $\sqrt{n/2}$.

Following this construction algorithm, we can easily prove out that $\sum_m \Lambda_m \Lambda_m = (n^2 - 1)I$. The Gell-Mann matrices coincide with the Pauli matrices when N=1.

3.5 Autocorrelation function of FMF channel matrix

In this section, we first derive a canonical stochastic differential equation (SDE) for channel matrix of FMF in the regime of strong coupling. From the SDE, we develop an analytical form of the autocorrelation function (ACF) of channel matrix, from which the channel correlation bandwidth is obtained. A comparison of simulation and analytical results demonstrates the accuracy of our analytical results.

The evolution of channel matrix *U* along fiber position z is governed by the mode coupling vector $\vec{\beta}$, a generalization of birefringence vector of SMF

$$\frac{dU(z,\omega)}{dz} = \frac{j}{2N}\left[\vec{\beta}(z,\omega)\cdot\vec{\Lambda}\right]U(z,\omega), \tag{3.17}$$

$\vec{\beta}\cdot\vec{\Lambda}$ is required to be a Hermitian matrix to ensure the unitarity of *U*. Such that, all the elements in $\vec{\beta}(z,\omega)$ are real [128]. For a 'long' fiber, where the overall fiber length is significantly larger than the correlation length of local modal dispersion and the coupling between various modes is strong, white Gaussian noise model can be adopted for the mode coupling vector [128]

$$\vec{\beta}(z,\omega) = \mu\omega\vec{n}(z), \tag{3.18}$$

where the elements of $\vec{n}(z)$ are statistically uncorrelated white Gaussian processes with zero mean and unit variance, and the parameter μ represents the mode coupling strength.

Substituting channel matrix *U* in Eq. (3.17) with its representation of Eq. (3.5), and using the product defined in Eq. (3.13), we get the evolution equation of $\tilde{u}$

$$\begin{aligned} du_0 &= \frac{j\mu\omega}{2N}\vec{u}\cdot d\vec{W} \\ d\vec{u} &= \frac{j\mu\omega}{2N}(u_0 d\vec{W} + \vec{u}\Theta d\vec{W} - j\vec{u}\otimes d\vec{W}) \end{aligned}, \tag{3.19}$$

where $d\vec{W} = \vec{n}dz$ is the differential of the Brownian motion $\vec{W}(z)$. Eq. (3.19) can be written in a canonical form of SDE

$$d\tilde{u}(z) = q(\tilde{u}(z))d\vec{W}(z), \tag{3.20}$$

where $\tilde{u} = [u_0;\vec{u}]$, and $q(\tilde{u})$ is the diffusion matrix of dimension $4N^2$ x $4N^2-1$ with entries to be polynomials of the components of $\tilde{u}$, defined as

$$q = \frac{j\mu\omega}{2N}\begin{bmatrix} \vec{u}^{\mathrm{T}} \\ u_0 I + \vec{u}\Theta - i\vec{u}\otimes \end{bmatrix}. \tag{3.21}$$

We should note that, normally a SDE is on a real-value space, whereas the dynamic vector $\tilde{u}$ in Eq. (3.20) is complex. However, we can consider Eq. (3.20) as a short hand for real-value SDEs. Namely, we can separate the real and imaginary part of the dynamic vector $\tilde{u}$, and treat the two parts individually as real elements. As long

as the Brownian motion term $\vec{W}(z)$ is real, the real and complex forms of SDE describe the same physics of stochastic process. For the sake of concision, we use the complex vector $\tilde{u}$ in our discussion.

From Eqs. (3.17) and (3.18), the *m*-th column $\tilde{q}_m(\tilde{u})$ in the diffusion matrix *q* satisfies

$$\tilde{q}_m \cdot \tilde{\Lambda} = \frac{j\mu\omega}{2N} \Lambda_m (\tilde{u} \cdot \tilde{\Lambda}) . \tag{3.22}$$

Denote $q_{x,y}$ as the entry at *x*-th row and *y*-th column in the diffusion matrix *q*. Using the trace-orthogonality of $\vec{\Lambda}$ in Eq. (3.22), $q_{x,y}$ can be extracted as

$$q_{x,y} = \frac{1}{2N} Tr\left\{ \Lambda_{x-1} (\tilde{q}_y \cdot \tilde{\Lambda}) \right\} . \tag{3.23}$$

Some important statistical properties of the dynamic vector $\tilde{u}$ can be derived from the SDE Eq. (3.20) using standard tools of stochastic calculus. However, those standard tools are usually applied for the Ito form of SDE, whereas the product involving the Brownian motion in Eq. (3.20) should be interpreted as Stratonovich product [125,126]. Before Eq. (3.20) can be directly used for stochastic calculus, it should be converted from Stratonovich form into its equivalent Ito form. In the following context, we will use two different approaches to carry out the conversion.

- Stratonovich-Ito Conversion Algorithm Approach

In the SDE theory, there is a standard algorithm to convert SDE of Eq. (3.20) from Stratonovich form to its equivalent Ito form [134]

$$d\tilde{u} = \tilde{c}(\tilde{u}) dz + q(\tilde{u}) d\vec{W} , \tag{3.24}$$

where the diffusion matrix $q(\tilde{u})$ is the same as that in Eq. (3.20), and $\tilde{c}(\tilde{u})$ is the drift correction vector of dimension $4N^2$, whose *l*-th component c_l is given by

$$c_l = \frac{1}{2} \sum_{n=1}^{4N^2} \sum_{m=1}^{4N^2-1} q_{n,m} \frac{\partial q_{l,m}}{\partial u_{n-1}} . \tag{3.25}$$

$q_{n,m}$ can be obtained using Eq. (3.23). Then, the product of vectors $\tilde{c}$ and $\tilde{\Lambda}$ can be written as

$$\tilde{c}\cdot\tilde{\Lambda}=\frac{1}{2}\sum_{n=1}^{4N^2}\sum_{m=1}^{4N^2-1}\frac{1}{2N}Tr\{\Lambda_{n-1}(\tilde{q}_m\cdot\tilde{\Lambda})\}\frac{\partial\tilde{q}_m\cdot\tilde{\Lambda}}{\partial u_{n-1}}, \tag{3.26}$$

where $\tilde{q}_m\cdot\tilde{\Lambda}$ is already known in Eq. (3.22). Noting $\partial\tilde{u}\cdot\tilde{\Lambda}/\partial u_n=\Lambda_n$, Eq. (3.26) can be rewritten as

$$\tilde{c}\cdot\tilde{\Lambda}=\frac{-1}{8N^2}\mu^2\omega^2[\sum_{m=1}^{4N^2-1}u_m\Lambda_m+\sum_{n=1}^{4N^2-1}\sum_{m=1}^{4N^2-1}\frac{1}{2N}Tr\{\Lambda_n\Lambda_m(\tilde{u}\cdot\tilde{\Lambda})\}\Lambda_m\Lambda_n]. \tag{3.27}$$

The product $\Lambda_n\Lambda_m$ and $\Lambda_m\Lambda_n$ in Eq. (3.27) can also be expanded as Eqs. (3.7) and (3.8). Using the trace-orthogonality of $\vec{\Lambda}$, the structure constants f_{mnk} and d_{mnk} can be extracted by

$$d_{mnk}=\frac{1}{2N}Tr\{\Lambda_k(\Lambda_m\Lambda_n+\Lambda_n\Lambda_m)\} \tag{3.28}$$

$$f_{mnk}=-\frac{i}{2N}Tr\{\Lambda_k(\Lambda_m\Lambda_n-\Lambda_n\Lambda_m)\}. \tag{3.29}$$

Switching indices *m* with *n* in the above functions, we have $d_{mnk}=d_{nmk}$, and $f_{mnk}=-f_{nmk}$. Substituting *n* with *m* in Eq. (3.29), we have $f_{mmk}=0$. Then, $\Lambda_m\Lambda_m=I+\sum_k d_{mmk}\Lambda_k$. Since we know $\sum_m\Lambda_m\Lambda_m=(4N^2-1)I$, we can easily prove out $\sum_m d_{mmk}=0$.

Substituting Eqs. (3.7) and (3.8) into the second term on the right side of Eq. (3.27), we have

$$\begin{aligned}\sum_{n=1}^{4N^2-1}\sum_{m=1}^{4N^2-1}\frac{1}{2N}Tr\{\Lambda_n\Lambda_m(\tilde{u}\cdot\tilde{\Lambda})\}\Lambda_m\Lambda_n=&\sum_{n=1}^{4N^2-1}\sum_{m=1}^{4N^2-1}[\delta_{mn}u_0+\sum_{k=1}^{4N^2-1}(-jf_{mnk}+d_{mnk})u_k]\\&\cdot[\delta_{mn}\Lambda_0+\sum_{l=1}^{4N^2-1}(-jf_{mnl}+d_{mnl})\Lambda_l]\end{aligned} \tag{3.30}$$

Using the properties of coefficients f_{mnk} and d_{mnk} we have discussed above, Eq. (3.30) can be simplified as

$$\sum_{n=1}^{4N^2-1}\sum_{m=1}^{4N^2-1}\frac{1}{2N}Tr\{\Lambda_n\Lambda_m(\tilde{u}\cdot\tilde{\Lambda})\}\Lambda_m\Lambda_n=(4N^2-1)u_0I+\sum_{kl}\sum_{mn}(f_{mnk}f_{mnl}+d_{mnk}d_{mnl})u_l\Lambda_k. \tag{3.31}$$

For the Gell-Mann matrices, we can also prove out that $\sum_{mn}(f_{mnk}f_{mnl}+d_{mnk}d_{mnl})=0$ for the case of $k\neq l$, and $\sum_{mn}(f_{mnk}{}^2+d_{mnk}{}^2)=(4N^2-2)$ for the case of $k=l$. Thus, Eq. (3.31) can be written as

$$\tilde{c}\cdot\tilde{\Lambda}=\frac{-(4N^2-1)}{8N^2}\mu^2\omega^2[u_0I+\sum_{m=1}^{4N^2-1}u_m\Lambda_m]. \tag{3.32}$$

As such, the drift correction vector $\tilde{c}$ in Eq. (3.24) becomes

$$\tilde{c} = -\frac{4N^2 - 1}{8N^2}\mu^2\omega^2\tilde{u} \,. \tag{3.33}$$

The *i*-th component in $\tilde{c}$ is proportional to the *i*-th component in $\tilde{u}$.

- Stochastic Integral Definition Approach

We can also develop the correction term $\tilde{c}$ from the first principle of definitions for the Stratonovich and Ito integral.

According to the definition of stochastic integral [135], the integral of SDE Eq. (3.20) can be interpreted as

$$\int_a^b q(\tilde{u}(z))d\vec{W}(z) = \lim_{\delta\to 0}\sum_{i=0}^{n-1} q(\lambda\tilde{u}(z_{i+1}) + (\lambda - 1)\tilde{u}(z_i))\Delta\vec{W}_i \,, \tag{3.34}$$

where z_i (*i*=0 to *n*-1) are *n* partitions in the intervals [*a*,*b*] with mesh $\delta = \max(z_{i+1} - z_i)$, $\Delta\vec{W}_i = \vec{W}(z_{i+1}) - \vec{W}(z_i)$, and λ is the parameter defining different methods of integral. The case of λ=0 corresponds to the Ito integral, whereas λ=1/2 corresponds to the Stratonovich integral.

Since SDE Eq. (3.20) is an equivalent of Eq. (3.17) with a different representation for the channel matrix *U*, Eq. (3.34) is equivalent to

$$\frac{j\mu\omega}{2N}\int_a^b [d\vec{W}(z)\cdot\vec{\Lambda}]U(z) = \frac{j\mu\omega}{2N}\lim_{\delta\to 0}\sum_{i=0}^{n-1}[\Delta\vec{W}_i\cdot\vec{\Lambda}](\lambda U(z_{i+1}) + (\lambda - 1)U(z_i)) \,. \tag{3.35}$$

From Eq. (3.35), the correction term, defined as the conversion between the Ito and Strantonovich integral, is a given by

$$\frac{i}{4N}\mu\omega[d\vec{W}\cdot\vec{\Lambda}]dU \,. \tag{3.36}$$

Substituting *dU* of Eq. (3.17) into Eq. (3.36), and using the rules of $dW_i dW_i = dz$, $dW_i dW_j = 0$ ($i \neq j$) [136], and $\vec{\Lambda}\cdot\vec{\Lambda} = (4N^2 - 1)I$, we can finally reach the correction term as

$$-\frac{4N^2 - 1}{8N^2}\mu^2\omega^2 U\cdot dz \,. \tag{3.37}$$

The correction term of Eq. (3.37) we get here is consistent with that we developed using the standard Stratonovich-Ito conversion algorithm in Eq. (3.33).

It follows that Eq. (3.17) in its Ito form can be expressed as

$$\frac{dU(z,\omega)}{dz} = -\frac{4N^2-1}{8N^2}\mu^2\omega^2 U + \frac{j\mu\omega}{2N}\left[\vec{n}(z,\omega)\cdot\vec{\Lambda}\right]U(z,\omega)\,. \tag{3.38}$$

Eq. (3.19) in its Ito form is

$$\begin{aligned} du_0 &= -\frac{(4N^2-1)\mu^2\omega^2}{8N^2}u_0 dz + \frac{j\mu\omega}{2N}\vec{u}\cdot d\vec{W} \\ d\vec{u} &= -\frac{(4N^2-1)\mu^2\omega^2}{8N^2}\vec{u}dz + \frac{j\mu\omega}{2N}(u_0 d\vec{W} + \vec{u}\Theta d\vec{W} - i\vec{u}\otimes d\vec{W}) \end{aligned}\,. \tag{3.39}$$

The canonical form for Eq. (3.39) is Eq. (3.24), with the diffusion matrix *q* defined by Eq. (3.21), and the drift vector $\tilde{c}$ defined as Eq. (3.33). The standard tools of stochastic calculus can be directly applied to Eq. (3.24).

The frequency correlation of channel matrix can be defined as

$$R(z,\omega_1,\omega_2) = U^{\dagger}(z,\omega_1)U(z,\omega_2)\,. \tag{3.40}$$

R is also a unitary matrix, describing the disparity between the channel matrices *U* at frequencies ω_1 and ω_2. The expectation *E*[*R*] is the ACF of channel matrix, giving a measure of the statistical dependence of *U*, and indicating how quickly *U* diverges in frequency domain. The bandwidth of *E*[*R*] is called channel correlation bandwidth, representing a frequency span over which *U* can be approximately considered as a constant matrix from the view of DSP.

Taking derivative of Eq. (3.40) with respect to length *z*, the evolution of *R* along *z* is

$$\frac{dR(z,\omega_1,\omega_2)}{dz} = \frac{dU^{\dagger}(z,\omega_1)}{dz}U(z,\omega_2) + U^{\dagger}(z,\omega_1)\frac{dU(z,\omega_2)}{dz}\,. \tag{3.41}$$

Substituting Eq. (3.17) into Eq. (3.41), we obtain

$$\frac{dR(z,\omega_1,\omega_2)}{dz} = \frac{j\mu(\omega_2-\omega_1)}{2N}[U^{\dagger}(z,\omega_1)(\vec{n}(z)\cdot\vec{\Lambda})U(z,\omega_1)]R(z,\omega_1,\omega_2)\,. \tag{3.42}$$

Term $U^{\dagger}(z,\omega_1)(\vec{n}(z)\cdot\vec{\Lambda})U(z,\omega_1)$ can be regarded as a rotated white Gaussian noise, and denoted as $\vec{n}'(z)\cdot\vec{\Lambda}$, where $\vec{n}'$ is also a vector of white Gaussian noise in a new coordinate system. From Eq. (3.42), we can observe that *R* is determined only by the frequency difference $\Delta\omega = \omega_2 - \omega_1$, such that Eq. (3.42) can be rewritten as

$$\frac{dR(z,\Delta\omega)}{dz} = \frac{j\mu\cdot\Delta\omega}{2N}[\vec{n}'(z)\cdot\vec{\Lambda}]R(z,\Delta\omega)\,. \tag{3.43}$$

Eq. (3.43) has the same form as Eq. (3.17), substituting ω with $\Delta\omega$. Following the steps in last section, we can also obtain the Ito equation for Eq. (3.43) in a similar form to Eq. (3.24) given by

$$d\tilde{r} = \tilde{c}(\tilde{r})dz + q(\tilde{r})d\vec{W}\,, \tag{3.44}$$

where $\tilde{r}$ is the vector representation of *R* on the basis matrices $R = \tilde{r}\cdot\tilde{\Lambda}$. The diffusion matrix $q(\tilde{r})$ is given by

$$q(\tilde{r}) = \frac{j\mu\cdot\Delta\omega}{2N}\begin{bmatrix} \vec{r}^{\,\mathrm{T}} \\ r_0 I + \vec{r}\Theta - i\vec{r}\otimes \end{bmatrix}. \tag{3.45}$$

The drift vector $\tilde{c}(\tilde{r})$ is given by

$$\tilde{c}(\tilde{r}) = -\frac{4N^2-1}{8N^2}\mu^2\cdot\Delta\omega^2\cdot\tilde{r}\,. \tag{3.46}$$

The Fokker-Planck equation (FPE) is a powerful tool of stochastic calculus, which can be used to derive the distribution of a dynamic process governed by SDE [136]. Applying FPE to the Eq. (3.44) gives the probability density function (pdf) *p* of vector $\tilde{r}$

$$\frac{\partial p(\tilde{r},z)}{\partial z} = \frac{1}{2}\sum_{n=1}^{4N^2}\sum_{m=1}^{4N^2}\frac{\partial^2 Q_{mn}\,p(\tilde{r},z)}{\partial r_{m-1}\partial r_{n-1}} - \sum_{k=1}^{4N^2}\frac{\partial c_k\,p(\tilde{r},z)}{\partial r_{k-1}}, \tag{3.47}$$

where $Q = qq^{\mathrm{T}}$. Though the distribution of the vector $\tilde{u}$ for channel matrix representation experiences an instant spread on unit sphere after fiber input due to the white Gaussian noise model adopted for the mode coupling vector [126], the distribution of the vector $\tilde{r}$ for ACF representation should experience a gradual spread on the unit sphere since the transpose conjugate operation in Eq. (3.40) mitigates the impact of random coupling.

On the other hand, the Dynkin formula could give the expectation of any smooth function *f* of the dynamic vector $\tilde{r}$ without the knowledge of the pdf [136]

$$\frac{dE[f(\tilde{r})]}{dz} = E[Gf(\tilde{r})]\,, \tag{3.48}$$

where *G* is the Ito generator

$$G = \frac{1}{2}\sum_{n=1}^{4N^2}\sum_{m=1}^{4N^2} Q_{mn}\frac{\partial^2}{\partial r_{m-1}\partial r_{n-1}} + \sum_{k=1}^{4N^2} c_k \frac{\partial}{\partial r_{k-1}}. \tag{3.49}$$

To obtain the ACF, we need to calculate $E[r_i]$. The smooth functions in this case are $f(\vec{r}) = r_i$. The second order differential operator in Eq. (3.49) to r_i equals to 0, and only the first order differential operator left. Such that, the evolution of $E[r_i]$ is only determined by the drift term c_k. Therefore, the Dynkin formula gives

$$\frac{dE[r_i(z)]}{dz} = -\frac{4N^2-1}{8N^2}\cdot\mu^2\cdot\Delta\omega^2 E[r_i(z)]. \tag{3.50}$$

The correlation matrix *R* at the input of fiber is an identity matrix *I* for any frequencies. With the initial condition $r_0(0) = 1$ and $\vec{r}(0) = \vec{0}$, we obtain

$$E[r_0] = \exp(-\frac{4N^2-1}{8N^2}\cdot\mu^2\cdot\Delta\omega^2\cdot z), \tag{3.51}$$

and $E[r_i] = 0$ ($i \neq 0$). Consequently, the ACF of channel matrix can be expressed as

$$E[R] = \exp(-\frac{4N^2-1}{8N^2}\cdot\mu^2\cdot\Delta\omega^2\cdot z)I. \tag{3.52}$$

From Eq. (3.52), the 3dB channel correlation bandwidth, where the correlation level falls to 0.5, is obtained [137]

$$B_{\Delta\omega} = \frac{4N}{\mu}\sqrt{-\frac{2\ln 0.5}{(4N^2-1)z}}. \tag{3.53}$$

Eq. (3.53) shows the frequency dependence of channel matrix decreases with the square root of fiber length *z*.

Taking derivative of Eq. (3.40) with respect to frequencies ω_1 and ω_2, we obtain

$$\frac{\partial^2 R}{\partial\omega_1\partial\omega_2} = \frac{\partial U^\dagger(\omega_1)}{\partial\omega_1}\cdot\frac{\partial U(\omega_2)}{\partial\omega_2}. \tag{3.54}$$

Setting $\omega_1 = \omega_2$ in Eq. (3.54), and noting that the Hermitian MD matrix is defined as $\vec{\Omega}\cdot\vec{\Lambda} = -j2NU_\omega U^\dagger$, where the subscript '$\omega$' stands for the derivative over ω [128], the frequency derivative of *R* can be related to the MD by

$$Tr\left\{\left.\frac{\partial^2 R}{\partial\omega_1\partial\omega_2}\right|_{\omega_1=\omega_2}\right\}=\frac{1}{4N^2}Tr\left\{(-j2NU_\omega U^+)(j2NUU_\omega^+)\right\}=\frac{\left|\vec{\Omega}\right|^2}{2N}\frac{n!}{r!(n-r)!}. \quad (3.55)$$

Provided that R is determined by the frequency difference $\Delta\omega$, $\partial^2 R/\partial\omega_1\partial\omega_2=-\partial^2 R/\partial^2\Delta\omega$. Substituting this into Eq. (3.55), we have

$$E\left[\left|\vec{\Omega}\right|^2\right]=-2N\left.\frac{\partial^2 Tr\{E[R]\}}{\partial(\Delta\omega)^2}\right|_{\omega=0}=(4N^2-1)\mu^2 z\,. \quad (3.56)$$

Eq. (3.56)is consistent with the study reported in [128]. The MD vector $\vec{\Omega}(z,\omega)$ of FMF undertakes a $4N^2-1$ dimensional isotropic random walk, and its root-mean-square value Ω grows as the square root of the fiber length *z*. [128]

From Eqs. (3.53) and (3.56), the MD and the ACF can be related by

$$Tr\{E[R]\}=2N\exp(-\frac{E[\vec{\Omega}^2]\cdot\Delta\omega^2}{8N^2})\,. \quad (3.57)$$

Our conclusions about the ACF of the FMF of Eqs. (3.53) and (3.57) coincide with those of the SMF studied in [126] when N=1. We extend the work of [126] to a more general case of arbitrary number of modes.

We conduct Monte-Carlo simulation to verify our theoretical result obtained in this section. A FMF is divided into 100 sections, assuming that the length of one section is longer than the correlation length. The local principal modes (PMs) in separate sections are considered independent and uniformly distributed. The propagation of each section is modeled as a 2*N* x 2*N* matrix with equivalent statistical properties given by [130,131]

$$V(\omega)=C_1T(\omega)C_2^{\dagger}\,, \quad (3.58)$$

where C_1 and C_2 are independent random unitary matrices corresponding to the random coupling of the PMs at the input and the output of the section, and *T* is a diagonal matrix representing the impact of group delays on the PMs

$$T(\omega)=diag[e^{-j\omega t_1},e^{-j\omega t_2},...,e^{-j\omega t_{2N}}]\,, \quad (3.59)$$

where t_i (*i*=1 to 2*N*) are the group delays of the local PMs. The average delay is a constant by setting $\sum t_i=0$. The channel matrix of the whole fiber is obtained by multiplexing the propagation matrices *V* for each section. The choice of the individual

group delays does not have significant impact on the statistical properties of the global fiber as long as the modes are strongly coupled [131]. We just need to ensure that the MD in one section $\tau = \sqrt{2N\sum t_i^2}$ satisfies the requirement of the global MD $\Omega = \tau\sqrt{100}$.

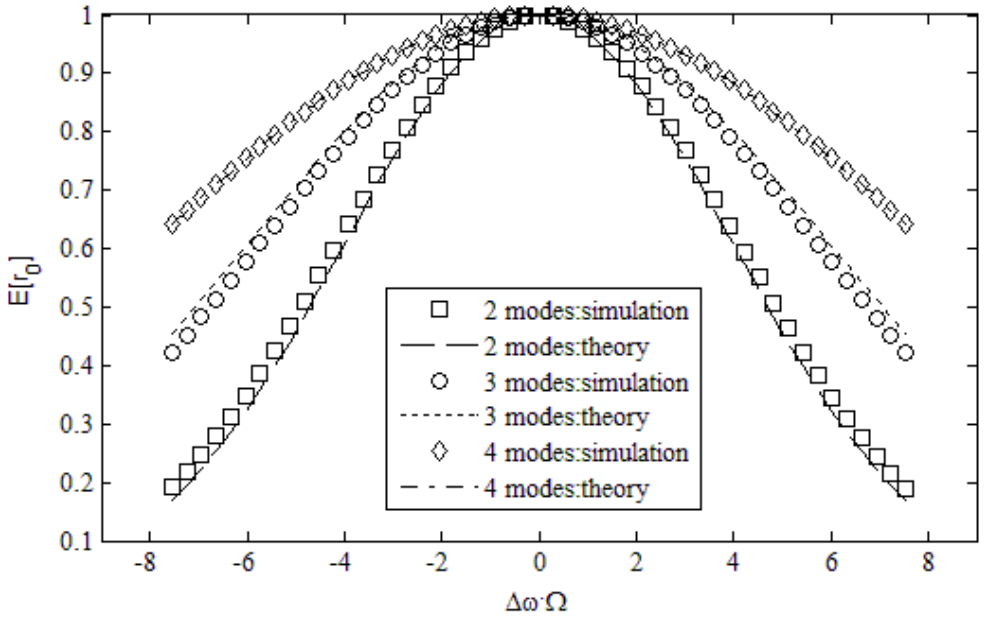

Fig. 3.5 The comparison of the analytical result and simulation result of the ACF of the Jones matrix for FMF with 2 to 4 modes.

After the channel matrices *U* at different frequencies are obtained using the Monte-Carlo simulation, their correlation *R* could be calculated using Eq. (3.40). The coordinate r_0 could be obtained by a trace operation $Tr\{R\}/2N$. We measured the ACF for FMFs with a 10-ps global MD in the bandwidth of 250 GHz. 10000 fiber configurations are used in the simulation to obtain the ensemble averaging. The simulation results and theoretical predictions of $E[r_0]$ for FMFs with 2 to 4 modes are plotted in Fig. 3.5. The accuracy of our theory result is justified by the excellent match between the simulation and the theoretical results.

3.6 Monitoring fast change of modal dispersion in FMF

Advanced DSP techniques in either single carrier or multicarrier modulation, can compensate the modal dispersion (MD) using blind channel estimation or pilot-symbol based channel estimation [89,90]. The signal equalization is performed based on the channel matrix to recover the signals affected by MD. MD in FMF is a stochastic property that varies randomly due to the temperature fluctuation and mechanical vibration [130]. A re-estimation of channel matrix at regular intervals is

needed to guarantee that MD can be properly compensated. The time interval of re-estimation is decided by the speed of the channel change. The analogy can be found in the fast polarization change of SMF, which has been investigated in [138] using a fast polarimeter. The dynamics of FMF channel has been reported in [139] by measuring the change of received signals using 4-dimensional Jones vector. Phase noise estimation is an important part in [139] for an accurate measurement. In this section, we employ high-dimensional Stokes-space analysis to monitor the speed of MD fluctuation. The change of MD in FMF channel can be described as the rotation of a Stokes vector on a high-dimensional sphere [128]. The phase noise is intrinsically eliminated by the Stokes-space analysis, which can reduce the complexity and inaccuracy of the measurement. By monitoring the speed of MD fluctuation, we can adjust the channel estimation intervals to obtain the optimal efficiency of information transmission.

For explanatory clarity, we use a dual-polarization, dual-spatial-mode multiplexed transmission system, where 4 orthogonal modes are transmitted, as an example of the high-dimensional Stokes-space analysis. A complex column vector $|\psi\rangle$ in the generalized Jones space of dimension 4 is used to denote the optical field in FMF. The transmitted signals and received signals can be related by Eq. (3.3). We send fixed $|\psi_t\rangle$ at the transmitter for the duration of the measurement. The changes of $|\psi_r\rangle$, which we record at the receiver, represent the dynamic properties of the channel matrix *U*. Considering the laser phase noise, and ignoring the equalization-enhanced phase noise (EEPN) [140], the received signal in experiment is denoted as $|\psi_r\rangle\exp(i\theta(t))$, where $\theta(t)$ is the time varying phase noise, which can potentially cause inaccuracy into the measurement when Jones vector approach is used [139].

As discussed in Section 3.5, a generalized Jones vector of dimension *n* is isomorphic to a generalized Stokes vector of dimension n^2-1. n^2-1 traceless Hermitian matrices, Λ_i (*i*=1 to n^2-1), form the basis in the Stokes space, satisfying the trace-orthogonal condition of Eq. (3.4). A normalized Jones vector is used in the following discussion, ignoring the effect of mode-dependent loss (MDL) in FMF. The Jones vector $|\psi\rangle$ is related to its corresponding Stokes vector $\vec{S}$ through a projection operator [140,141]

$$|\psi\rangle\langle\psi| = \frac{I}{n} + \sqrt{\frac{n-1}{2n}}\vec{S}\cdot\vec{\Lambda}, \tag{3.60}$$

where $\langle\psi|$ denotes the Hermitian conjugate of $|\psi\rangle$, I is the n-dimensional identity matrix, and the vector $\vec{\Lambda} = [\Lambda_1; \Lambda_2; \cdots; \Lambda_{n^2-1}]$ is an ensemble of the basis matrices. It follows from Eq. (3.60) that the phase noise $\exp(i\theta(t))$ is cancelled in the projection operator of $|\psi\rangle\langle\psi|$. Since $|\psi\rangle\langle\psi|$ produces a Hermitian matrix, the elements in $\vec{S}$ are all real [128]. They can be extracted using the trace-orthogonal condition of Eq. (3.4) as

$$S_i = \sqrt{\frac{n}{2(n-1)}} Tr\{\Lambda_i |\psi\rangle\langle\psi|\} = \sqrt{\frac{n}{2(n-1)}} \langle\psi|\Lambda_i|\psi\rangle. \tag{3.61}$$

Assume the Jones vector $|\psi\rangle$ has a form of $[\psi_1, \psi_2, \psi_3, \psi_4]^T$, where superscript '$T$' stands for transpose. Using Eq. (3.61), its corresponding Stokes vector can be written as

$$\vec{S} = \sqrt{n/(2(n-1))}[|\psi_1|^2 - |\psi_2|^2, \sqrt{1/3}(|\psi_1|^2 + |\psi_2|^2 - 2|\psi_3|^2), \sqrt{1/6}(|\psi_1|^2 + |\psi_2|^2 + |\psi_3|^2 - 3|\psi_4|^2),$$

$$2\,\mathrm{Re}\{\psi_1\psi_2^*\}, 2\,\mathrm{Re}\{\psi_1\psi_3^*\}, 2\,\mathrm{Re}\{\psi_1\psi_4^*\}, 2\,\mathrm{Re}\{\psi_2\psi_3^*\}, 2\,\mathrm{Re}\{\psi_2\psi_4^*\}, 2\,\mathrm{Re}\{\psi_3\psi_4^*\}, 2\,\mathrm{Im}\{\psi_1\psi_2^*\},$$

$$2\,\mathrm{Im}\{\psi_1\psi_3^*\}, 2\,\mathrm{Im}\{\psi_1\psi_4^*\}, 2\,\mathrm{Im}\{\psi_2\psi_3^*\}, 2\,\mathrm{Im}\{\psi_2\psi_4^*\}, 2\,\mathrm{Im}\{\psi_3\psi_4^*\}]^T. \tag{3.62}$$

It is easy to show that $\vec{S}$ has a magnitude of 1, and therefore it can be seen as a unit vector on a sphere of dimension n^2-1. The rotation of $\vec{S}$ on the Stokes sphere represents the change of the Jones vector $|\psi\rangle$. The instantaneous frequency of the Stokes vector rotation in the time interval of $[t, t+\Delta t]$ is given by

$$f = \frac{\arccos(\vec{S}(t)\cdot\vec{S}(t+\Delta t))}{2\pi\cdot\Delta t}. \tag{3.63}$$

This 15-dimensional Stokes vector in TMF is an extensional of conventional 3-dimensional Stokes vector in SMF. Our approach of high-dimensional Stoke-space analysis can be extended to any dimension.

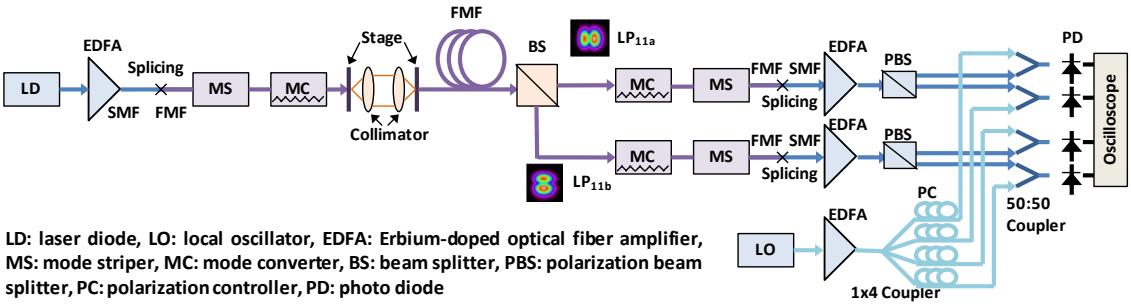

Fig. 3.6 Set-up for the measurement of TMF channel fluctuation.

Fig. 3.6 shows the set-up of our measurement. At the transmitter, a SMF and a FMF are first center-spliced to launch the power from laser into FMF. After that, a mode striper (MS) is used to strip out high order mode leaving only fundamental mode LP_{01}. The MS is realized by tightly winding the FMF over an 8-mm post. LP_{01} mode after the MS is then converted to LP_{11} mode using a mode converter (MC), which is realized by a long-period fiber grating (LPFG). The modal extinction ratio (ER) of the MC used in our measurement is beyond 12 dB. The generated LP_{11} mode is launched into the tested FMF using stages and collimating lenses. Maximum coupling ratio is obtained by carefully aligning the stages and the lenses. The input Stokes vector of LP_{11} mode could be arbitrary as long as it is stable for the duration of the measurement. After the transmission, the full information in the 4 orthogonal modes is collected by mode demultiplexing. A beam splitter (BS) is first used to split received signals into two paths. The fiber ends on the stages of the two paths can be axially rotated to make sure the orthogonal orientations of LP_{11a} and LP_{11b} mode. MC and MS are used in reverse order as they are used in transmitter to convert LP_{11} mode into LP_{01} mode. The two polarizations are further separated by polarization beam splitter (PBS). The received orthogonal modes are detected by heterodyne coherent detection, using a local oscillator (LO) with 40-MHz wavelength shifting from the transmitter laser. Four paths of LO is coupled with the received signal of each mode before the photo diode (PD). A polarization controller (PC) is placed in each path of LO to align the polarizations. Data is collected by an oscilloscope and processed offline.

We use the set-up described in the last section to measure the speed of MD fluctuation for the 26-km custom-designed circular-core TMF used in our lab. The fiber parameters for the TMF are as following: the loss is 0.26 dB/km for LP_{01} and LP_{11} mode, the chromatic dispersion is 22 ps/(nm·km) for LP_{01} mode, and 17

ps/(nm·km) for LP_{11} mode, the DMD between LP_{01} and LP_{11} mode is 3 ps/m, and the DMD within LP_{11} mode is 80 ps/km. Due to the large difference of effective refractive index between LP_{01} and LP_{11} mode, the coupling between these two modes can be ignored. A sampling rate of 125 MSa/s is used at the oscilloscope for a recording time of 20 ms. If a Stokes vector rotates to its orthogonal state in the sampling interval of 8 ns, the rotation angle on the Stokes sphere would be 0.6 π [140]. It follows that the fastest measurable frequency in our experiment is 3.8×10^6 revolution/s (RPS) according to Eq. (3.63).

To investigate the impact of environmental disturbance on the speed of MD fluctuation, we also compare the two situations when TMF is placed still and when it is hit with a screw driver. The evolutions of the 15 elements in the Stokes vector for the TMF, which can be obtained from Eq. (3.61), are shown in Fig. 3.7(a) and (b). Comparing Fig. 3.7(a) with (b), the large fluctuation in Fig. 3.7(b) indicates a faster rotation of the Stokes vector for the disturbed TMF. That means the outside disturbance would accelerate the MD change. It is inevitable that the field fiber would be affected by the environmental disturbance. Thus, monitoring the speed of MD fluctuation is a useful technique.

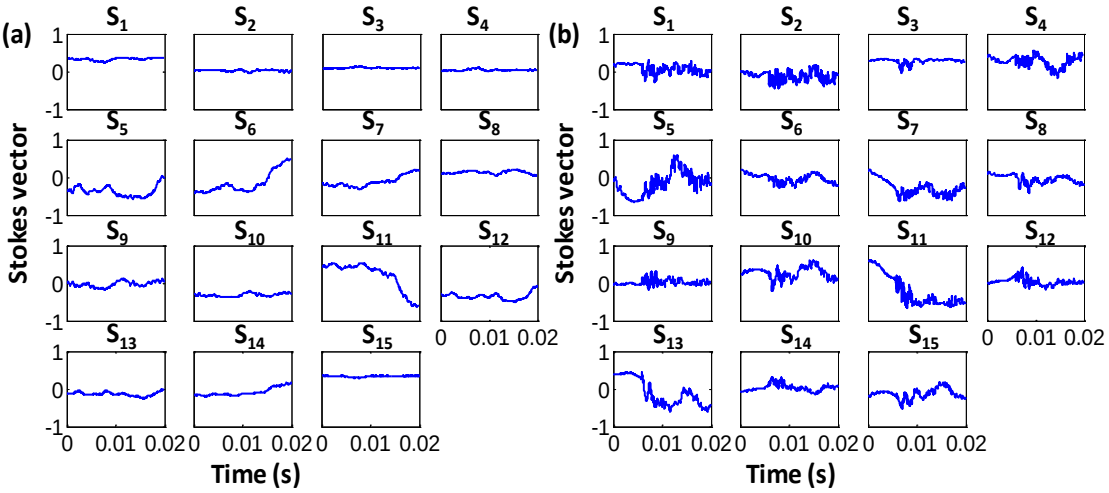

Fig. 3.7 (a) The evolution of 15-dimensional Stokes vector for still TMF. (b) The evolution of 15-dimensional Stokes vector for disturbed TMF.

We also calculate the MD evolution frequency using Eq. (3.63). Comparing Fig. 3.8(a) with (b), the highest evolution frequency can reach 7,500 RPS for disturbed TMF, while the highest evolution frequency for still TMF is less than 1,000 RPS. We plot the complementary cumulative distribution (CCDF) of the evolution frequency in

Fig. 3.8(c) and (d). The evolution frequency at the probability of 10^{-2} is 600 RPS for still TMF, and 5,000 RPS for disturbed TMF.

According to our measurement result, the Stokes vector evolution frequency in a mode multiplexing transmission channel may reach 10^4 RPS under strong disturbance. To correctly estimate the channel matrix and compensate MD, we recommend a re-estimation at timescale of microsecond. The technique described in this paper can also be used to monitor the speed of channel change. The re-estimation interval can be adjusted according to the channel condition.

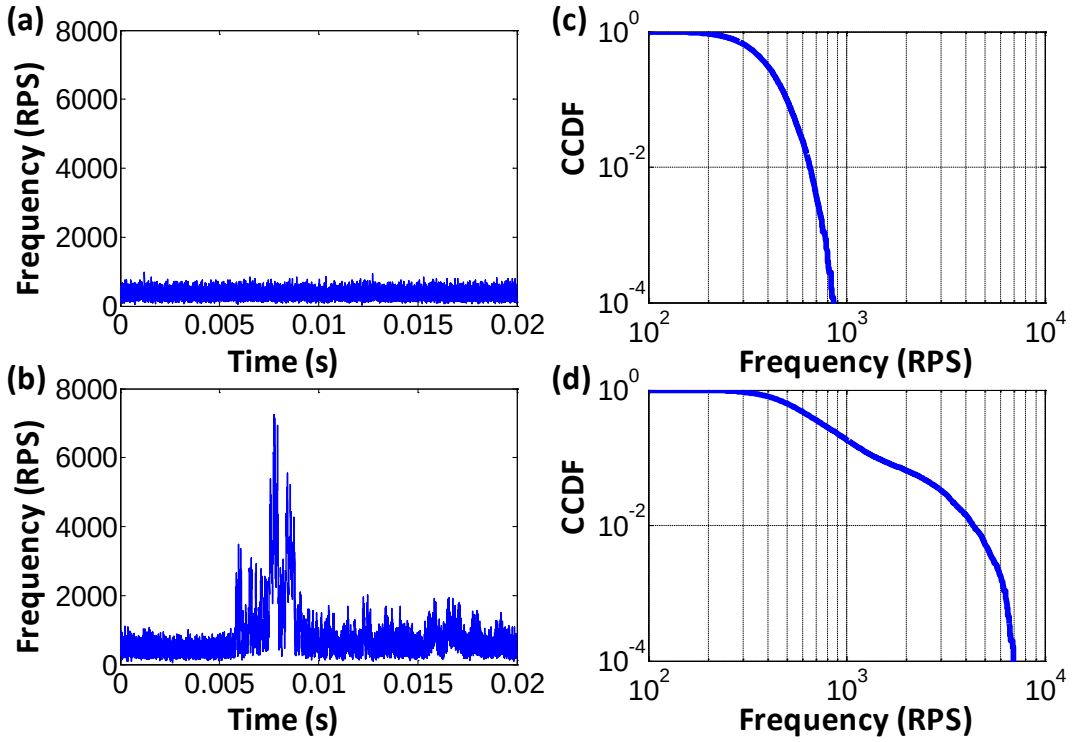

Fig. 3.8 (a) Rotation frequency of Stokes vector for still TMF. (b) Rotation frequency of Stokes vector for disturbed TMF. (c) Complementary cumulative distribution function (CCDF) of rotation frequency for still TMF. (d) CCDF of rotation frequency for disturbed TMF.

3.7 Conclusion

In this chapter, we review the fundamentals of FMF, and the mode multiplexing/demultiplexing devices for the FMF transmission. To investigate the frequency dependence of FMF channel, we have used two different approaches to derive the SDE for the channel matrix of a FMF in the regime of strong coupling. The channel matrix is decomposed over generalized high-dimensional Gell-Mann matrices,

an equivalent of two-dimensional Pauli matrices. We also develop the ACF of channel matrix for a strong coupling FMF. The channel correlation bandwidth developed from the ACF indicates that the frequency dependence of the channel matrix decreases with the square root of fiber length. The validity of our analytical result is confirmed by the Monte-Carlo simulation. To investigate the stochastic property of FMF channel, we propose a method of measuring the speed of MD fluctuation using high-dimensional Stokes-space analysis. The Stoke analysis could cancel the measurement inaccuracy caused by the phase noise, and provides an efficient monitoring for the channel.

4 Advanced Modulation Formats for High-Performance Short-Reach Networks with Direct Detection

4.1 Introduction

During the last decade, long-haul optical networks have witnessed a capacity advance to multi-Terabit with the revival of coherent communications [80, 142-143]. By employing polarization-diversity transceiver and sophisticated DSP, coherent communication achieves great robustness against the linear fiber impairments caused by chromatic dispersion (CD) and polarization mode dispersion (PMD). In the meantime, short-reach networks are also facing a challenge to upgrade their capacity per wavelength beyond 100 Gb/s to support the growing traffic demand of data center interconnections. Different from long-haul networks, short-reach networks accommodate wide geographic coverage by deploying a massive number of transceivers with reach smaller than hundreds of kilometers. As a result, the transceiver cost becomes the primary consideration for the short-reach network design. The costly polarization-diversity transceiver may prevent coherent detection from the widespread use in short-reach networks.

Direct modulation (DM) based direct detection (DM/DD) is the simplest DD scheme [144-146]. Since the CD introduces severe nonlinear distortion to the signal due to the square-law photo-detection, the transmission distance of DM/DD is limited to tens of kilometers. Furthermore, the phase and polarization dimensions are not utilized by DM/DD, leading to low electrical spectrum efficiency (SE). To support the transmission of hundreds of kilometers, a linear channel is needed where linear mapping can be obtained between the baseband and optical signal. For the high-performance short-reach optical networks, we aim to find promising solutions which can achieve a linear channel in a cost efficient way.

Various novel approaches have been proposed to bridge the gap between the low-cost low-performance IM/DD system and the high-performance high-cost coherent system [54-55,147-149]. The solutions can be classified into two categories. The first category uses directly modulated laser (DML) at the transmitter to reduce the system cost, and coherent detection at receiver to obtain the linear channel. The second category uses direct detection at the receiver to reduce the system cost, and coherent modulation at the transmitter to improve the spectral efficiency. To achieve a linear

channel with direct detection, advanced DD modulation formats have been proposed through a so-called self-coherent (SCOH) approach, where a carrier is transmitted together with the signal to down-convert the optical signal at the receiver. In that way, the impact of the CD can be easily removed from the received signal, greatly extending the transmission distance of the DD system.

One of the widely investigated SCOH schemes is single sideband (SSB) modulation format, where the signal is filled at one side of the carrier, enabling the recovery of the complex signal using a single-ended photodiode (PD) [54-55,147]. To avoid the interference of the signal-to-signal beat noise (SSBN) produced at the square-law detection, guard band or subcarrier interleaving is adopted, leading to the reduced electrical SE [54,55]. Gapless SSB modulation format is then proposed to improve the electrical SE by removing the guard band and employing iterative SSBN cancellation [54,55]. However, single-polarization SSB modulation format cannot achieve electrical SE higher than 1/2 since it utilizes only half spectrum of the baseband.

The electrical SE is a critical metric to evaluate the performance of short-reach networks. Compared to SSB modulation formats, double sideband (DSB) modulation formats utilize full spectrum of the baseband, presenting greater potential to achieve higher electrical SE. The technical challenge is to obtain the phase-diverse DSB signal using DD scheme. In the past few years, significant efforts have been made to improve the electrical SE of DD system using DSB modulation formats [150-152]. Block-wise phase shift (BPS) DD and signal carrier interleaved (SCI) DD have been proposed, achieving 1/2 and 2/3 electrical SE compared to single-polarization coherent detection [150,151] (throughout the text, the relative electrical SE is normalized to that of single-polarization coherent detection). Recently, DD with single-polarization modulation achieves 100% electrical SE for the first time by using Stokes vector direct detection (SV-DD) [152].

This chapter is organized as follows: Section 2 gives a brief background of conventional DD modulation format and the motivation of applying advanced DD modulation formats. Section 3 provides a review of various advanced DD modulation formats in the form of both SSB and DSB. Sections 4 and 5 present study on the mitigation of the PMD and PDL induced impairments in SV-DD systems,

respectively. Section 6 gives the comparison of the advanced DD modulation formats. The conclusion is drawn in Section 7.

4.2 Motivation of advanced modulation formats for direct detection

Conventional DD, referring to DM/DD, has the merits of simple implementation and low cost. As shown in Fig. 4.1, DM/DD modulates the non-negative signal directly onto a laser, and converts it to the photocurrent by a single-ended PD at the receiver [146]. There are two limiting factors for conventional DD. The first one is the nonlinear mapping between the transmitted baseband signal and the optical field, as well as the nonlinear mapping between the optical field and the received signal after the square-law detection. Due to the nonlinear mapping, the CD causes severe nonlinear distortion to the signal, significantly degrading the performance of the system. Therefore, the conventional DD is only suitable for very short application, where the dispersion is not significant. The second limiting factor is the reduced spectral efficiency. As DM/DD uses single-polarization intensity modulation without utilizing the phase and polarization dimensions, the spectral efficiency is significantly reduced.

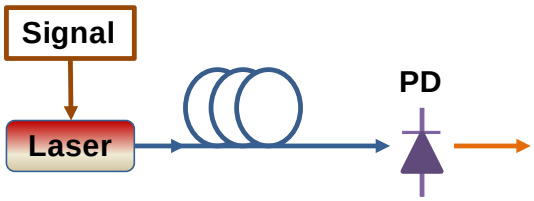

Fig. 4.1 Conceptual diagram of a conventional DD system.

Since the standardization of 100G Ethernet LAN interfaces in 2010, data center interconnects are also facing the increasing challenge to upgrade their capacity. High-performance DD system, with capacity per wavelength at 100 Gb/s and beyond and transmission distance over hundreds of kilometers, is urgently in demand. Various advanced modulation formats have been proposed to achieve this goal. We classify the advanced modulation formats in to two categories: SSB modulation formats and DSB modulation formats, where SSB modulation formats fill the signal into half side of the baseband, while the DSB modulation formats utilize the full band. For both formats, the baseband signal is linearly mapped to the optical field, leading to a linear communication channel under the impact of CD. A main carrier is transmitted together with the signal to down-convert the optical signal to the baseband at the

receiver. Since the CD introduces a frequency dependent phase shift to the optical field, the CD induced impairment can be easily compensated by removing the phase shift from the down-converted signal. The linear mapping between the baseband signal and the optical field contributes to high-performance DD system using advanced modulation formats.

4.3 High-performance advanced modulation formats for direct detection

In this section we will divide the high-performance advanced DD modulation formats into 2 categories: SSB formats and DSB formats. Their principles and transceiver architectures will be discussed in detail.

4.3.1 Single sideband modulation formats

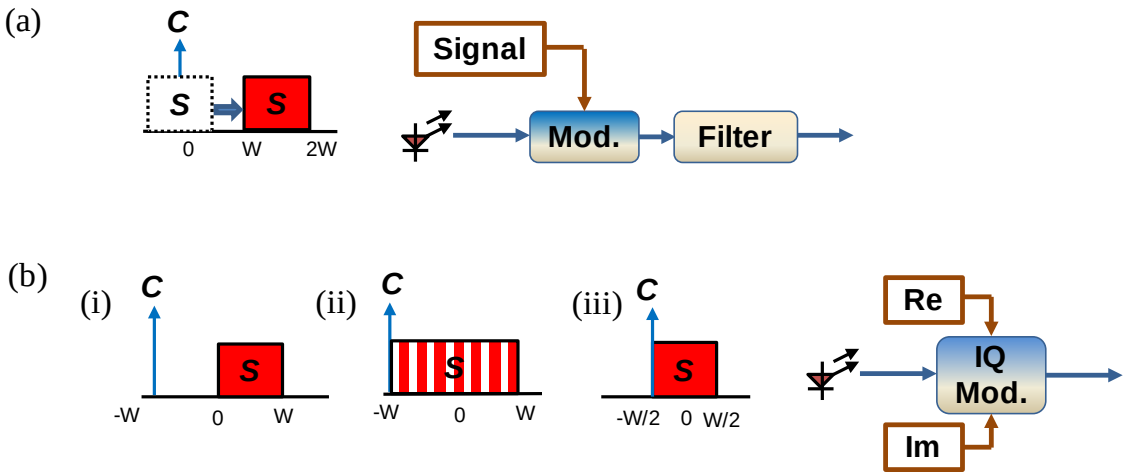

Fig. 4.2 SSB generation schemes: (a) offset SSB, and (b) RF tone assisted SSB which can be (i) with guard band, or (ii) subcarrier-interleaving, or (iii) without guard band.

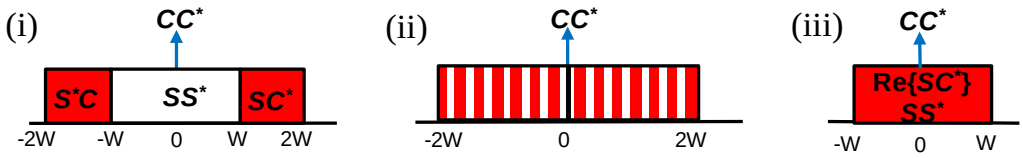

Fig. 4.3 Received baseband spectra for (i) SSB with guard band, (ii) subcarrier-interleaved SSB, and (iii) gapless SSB.

To detect the complex optical signal using a single-ended PD, SSB modulation format fills the signal at one side of the carrier. SSB signal can be generated by different approaches. As shown in Fig. 4.2(a), offset SSB signal is generated by combing a passband signal and the main carrier at DC [147]. The electrical signal is fed into an intensity modulator, generating output optical signal with Hermitian symmetric

spectrum around the main carrier. An optical filter is then used to filter out one side of the spectrum, yielding the SSB signal. Denoting the signal as *S* and carrier as *C*, the output of the square law photodiode is

$$|S+C|^2 = |C|^2 + |S|^2 + 2\,\mathrm{Re}\{SC^*\}. \tag{4.1}$$

It includes three beating terms from the SSB signal: (1) the linear signal term Re{SC^*}, (2) the nonlinear noise term SS^*, and (3) the DC term CC^*. The SSB signal can be fully recovered from half of the spectrum of the linear beating term Re{SC^*}. Besides the linear channel, the self-coherent approach has another merit: as the signal and carrier are provided by the same laser, the frequency and the phase noise in signal and carrier are the sample. Without the frequency offset compensation and phase noise compensation, DSP is simplified.

However, the signal term Re{SC^*} will be contaminated by the signal-to-signal beating noise (SSBN) SS^* if they are overlapped in frequency. To avoid this problem, a gap is reserved in frequency between the carrier and the modulated signal. For the SSB baseband signal with frequency range of [0, ω_N], the signal is only modulated in the frequency range [$\omega_{N/2}$, ω_N], where *N* denotes the number of subcarrier. The modulated optical signal can be written as $\sum_{i=N/2}^{N} E_i \exp[i(\omega_i+\omega_c)t]$, where ω_c is the frequency of carrier, ω_i is the frequency of *i*-th subcarrier. After the square law operation of single-ended PD, the recovered baseband signal can be written as

$$\begin{aligned}|S+C|^2 = |C|^2 &+ \sum_{i=N/2}^{N} E_i \exp(j\omega_i t)\cdot\sum_{i=N/2}^{N} E_i^* \exp(-j\omega_i t)\\ &+\sum_{i=N/2}^{N} E_i \exp(j\omega_i t)+\sum_{i=N//2}^{N} E_i \exp(-j\omega_i t)\end{aligned}. \tag{4.2}$$

The linear term falls into the frequency range [$-\omega_N$, $-\omega_{N/2}$] and [$\omega_{N/2}$, ω_N], while the SSBN term falls into the frequency range [$-\omega_{N/2}, \omega_{N/2}$]. Therefore, the signal is free from the contamination of SSBN. Another way to avoid the contamination from SSBN is to use frequency interleaving, where the signal is modulated only on the odd subcarriers. The optical signal after the modulation can be written as $\sum_{i=0}^{N} E_{2i+1} \exp[i(\omega_{2i+1}+\omega_c)t]$. After the detection of single-ended PD, the recovered baseband signal can be expressed as

$$\begin{aligned}\left|S+C\right|^2 = \left|C\right|^2 + \sum_{i=0}^{N} E_{2i+1}\exp(j\omega_{2i+1}t)\cdot\sum_{i=0}^{N} E_{2i+1}^{*}\exp(-j\omega_{2i+1}t) \\ +\sum_{i=0}^{N} E_{2i+1}\exp(j\omega_{2i+1}t)+\sum_{i=0}^{N} E_{2i+1}\exp(-j\omega_{2i+1}t)\end{aligned}. \quad (4.3)$$

The linear signal only falls into the odd subcarriers ω_{2i+1} , while the SSBN only falls into the even subcarriers ω_{2i}. For these two approaches, the offset SSB signal takes 1/4 spectrum of the baseband, leading to the electrical SE of 1/4 at both transmitter and receiver.

The electrical SE at the transmitter can be improved by using RF tone assisted SSB signal, where the RF tone introduces a virtual carrier at the frequency $\omega\neq0$ []. As shown in Fig. 4.2(b), RF tone assisted SSB signal reaches electrical SE of 1/2 at the transmitter when using (i) guard band, or (ii) subcarrier-interleaving [54,55] (see insets (i) and (ii) in Fig. 4.2(b)), and reaches electrical SE of 100% if no guard band is used (see inset (iii) in Fig. 4.2(b)). The obtained electrical signal is then fed into an IQ modulator. The electrical spectra at the receiver for the three types of SSB signal are shown in Fig. 4.3. For the methods of using (i) guard band, and (ii) subcarrier-interleaving, the nonlinear noise fills half of the Hermitian symmetric spectrum, leading to the electrical SE of 1/4 at the receiver. For the method (iii) without guard band, the electrical SE is improved to 1/2 while the signal is contaminated by the nonlinear noise. Iterative SSBN cancellation has to be used to remove the nonlinear noise, increasing the DSP complexity. For single-polarization SSB modulation format, the electrical SE at the receiver cannot be higher than 1/2 due to the Hermitian symmetry of the baseband spectrum.

Though higher electrical SE could be obtained by employing dual-polarization SSB signal [148,149], the implementation complexity of the transceiver approaches that of polarization multiplexed coherent systems. Electrical SE is an important metric to evaluate the performance of DD system. Considering the limited bandwidth of electronic components, higher electrical SE means higher data rate when the same QAM size is used, or means better receiver sensitivity when smaller QAM size is used to achieve the same data rate. Several DD schemes using DSB modulation formats have been proposed to improve the electrical SE. The technical challenge is to obtain the phase-diverse DSB signal using DD schemes. In the next subsection, we will introduce three novel DD schemes which can successfully solve the problem of phase diversity.

4.3.2 Double sideband modulation formats

- **Block-wise phase shift direct detection**

As indicated in Eq. (4.1), only the in-phase part of the signal can be recovered when single-ended PD is used. When a 90^o phase shift is induced between the signal and carrier, the quadrature part of the modulated signal can also be obtained. Based on this approach, the block-wise phase shift direct detection (BPS-DD) is proposed. BPS-DD uses a transmitter architecture shown in Fig. 4.4. The DSB signal is modulated on one path using an IQ modulator, and combined with the carrier from the other path. In order to achieve the phase diversity, the DSB signal is repeated in two consecutive time blocks, where a phase shift is introduced between them using three different approaches depicted in Fig. 4.5 [150].

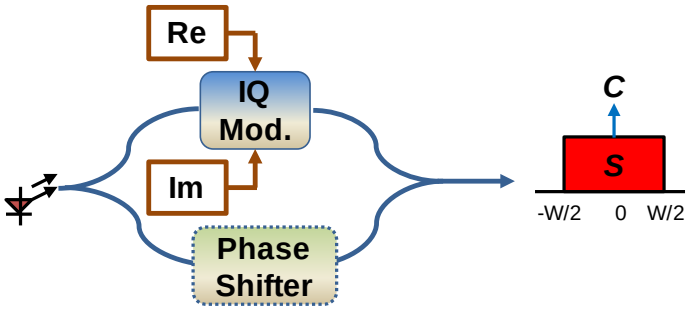

Fig. 4.4 Transmitter structure for block-wise phase switching (BPS) DD.

For the first approach shown in Fig. 4.5(a), the phase of the carrier C is switched by 90^o in two consecutive blocks, while the signal S remains the same. The outputs of the two blocks at the square-law detection are

$$I_1 = |S + C|^2 = |C|^2 + 2\,\mathrm{Re}\{SC^*\} + |S|^2 \tag{4.4}$$

$$I_2 = |S + jC|^2 = |C|^2 + 2\,\mathrm{Im}\{SC^*\} + |S|^2, \tag{4.5}$$

where I_1 and I_2 contain the real and imaginary parts of the DSB signal respectively. By combining them, we have

$$I = I_1 + jI_2 = (1+j)|C|^2 + 2SC^* + (1+j)|S|^2, \tag{4.6}$$

where the fading free signal is provided by the second term SC^* on the right side of the equation. The first term is a DC term, which can be simply ignored. The third term is the nonlinear noise term, by which the linear signal is contaminated. To obtain the linear signal SC^*, iterative SSBN cancellation is applied to remove the nonlinear noise

from the received signal *I*. High carrier-to-signal power ratio (CSPR) is needed for an effective SSBN cancellation. In the second approach shown in Fig. 4.5(b), the 90° phase shift is implemented in the signal *S*. The principle of the second approach is the same as the first one.

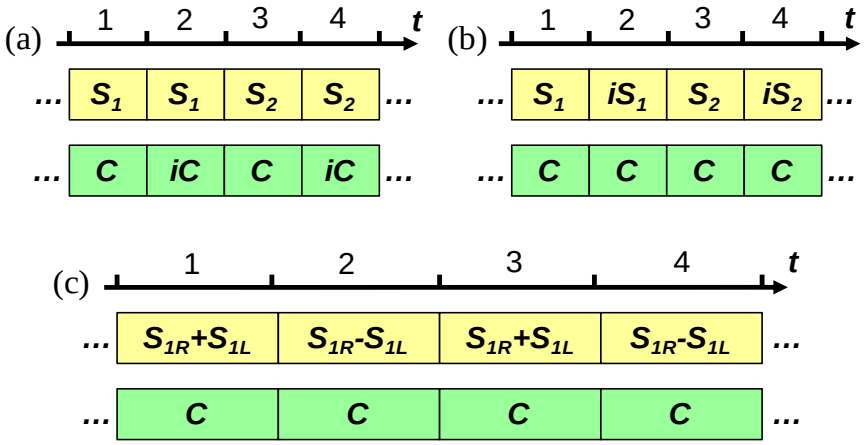

Fig. 4.5 Conceptual diagram of three phase switching approaches: (a) carrier phase switching (CPS) where carrier phase is switched by 90 degree, (b) signal phase switching (SPS) where signal phase is switched by 90 degree, and (c) (signal) set phase reversal (SPR) where the phase of lower sideband is switched by 180 degree.

In the third approach shown in Fig. 5(c), the phase shift is carried out by switching the lower sideband S_L of the signal by 180° and maintaining the upper sideband S_R the same. The outputs of the two consecutive blocks at the square-law detection are

$$I_1 = |S_R + S_L + C|^2 = |C|^2 + 2\,\mathrm{Re}\{(S_R + S_L)C^*\} + |S_R + S_L|^2 \tag{4.7}$$

$$I_2 = |S_R - S_L + C|^2 = |C|^2 + 2\,\mathrm{Re}\{(S_R - S_L)C^*\} + |S_R - S_L|^2 . \tag{4.8}$$

Combining I_1 and I_2, we have

$$I_3 = I_1 + I_2 = 2|C|^2 + 4\,\mathrm{Re}\{S_R C^*\} + 2|S_R|^2 + 2|S_L|^2 \tag{4.9}$$

$$I_4 = I_1 - I_2 = 4\,\mathrm{Re}\{S_L C^*\} + 4\,\mathrm{Re}\{S_R S_L^*\}. \tag{4.10}$$

The upper and lower sideband of the DSB signal is obtained from I_3 and I_4 respectively. Similar to the previous approaches, large CSPR and iterative SSBN cancellation are needed to mitigate the nonlinear noise.

For BPS scheme, full band is utilized to transmit the phase-diverse signal. Since the information is repeated in in two consecutive blocks, 1/2 electrical SE is achieved

by BPS scheme, which is the same as the gapless SSB-DD [54,55]. By using the first approach in Fig. 5, 49.4 Gb/s direct detection optical OFDM over 80-km standard single-mode fiber (SSMF) has been experimentally demonstrated [150].

The performance of self-coherent approach using single-ended PD is limited by the two factors: the limited spectrum efficiency and the contamination from SSBN. Since the single-ended PD produces real signal. Only the in-phase part or the quadrature part of the modulated signal can be recovered at a time. Either half of the spectrum is wasted to recover the signal from a real output, or half of the time is wasted to recover the in-phase and quadrature part separately.

- **Signal-carrier interleaved direct detection**

Coherent detection is naturally resistant to the SSBN. High spectral efficiency can be achieved by recovering the complex signal. In coherent detection, the signal and carrier (local oscillator) enter into a 90^o optical hybrid as two separate inputs. After the hybrid, the in-phase and quadrature parts of the signal are recovered by two balanced PDs, which remove the SSBN instinctively. If separate signal and carrier can be obtained at the receiver, and used as two inputs for the optical hybrid, the recovered signal will have the benefits of improved spectral efficiency and resistance to SSBN.

In fiber communication system, the optical signal can be multiplexed in three domains, the frequency domain, the time domain and the polarization domain. To separate the signal and carrier in frequency domain, a filter is needed to filter out the carrier from the received signal. A frequency gap has to be reserved between the signal and carrier to accommodate the bandwidth of the filter, reducing the spectral efficiency. Moreover, the system cost will be significantly increased with a filter. Therefore, the multiplexing of signal and carrier in frequency domain is an impractical solution.

In Signal-carrier interleaved direct detection (SCI-DD) scheme, the carrier and DSB signal are multiplexed in time domain. Separate time blocks are used for the transmission of signal and cattier. As shown in Fig. 4.6, the received signal is first divided into two paths. By introducing one block delay into one path, the signal and carrier blocks are aligned again. The two paths are then fed into the inputs of a standard balanced receiver, where the phase-diverse DSB signal gets recovered. The

electrical SE of SCI-DD is determined by how the signal and carrier blocks are arranged [151]. In the first approach shown in Fig. 4.7(a), a carrier block is inserted after each signal block. After the delay at the receiver, Block A contains signal S_1 in the upper path and carrier C in the delayed path, while Block B contains carrier C in the upper path and signal S_1 in the delayed path. Conjugate outputs, $S_1C^*$and $S_1{}^*C$ are obtained from Block A and B via balanced receiver, leading to the electrical SE of 1/2. The electrical SE can be improved by the second approach, where a carrier block is inserted after every two signal blocks, as shown in Fig. 4.7(b). After the delay at the receiver, the outputs, S_1C^*, $S_1S_2{}^*$, and $S_2{}^*C$, are obtained from Block A, B and C via the balanced receiver. The DSB signal S_1, S_2 can be recovered from Block A and C, leading to the electrical SE of 2/3.

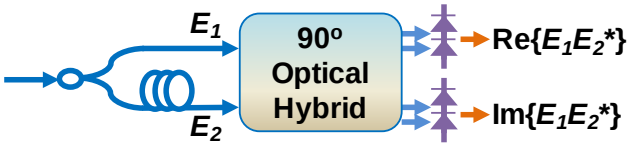

Fig. 4.6 Receiver structure for signal carrier interleaved (SCI) DD.

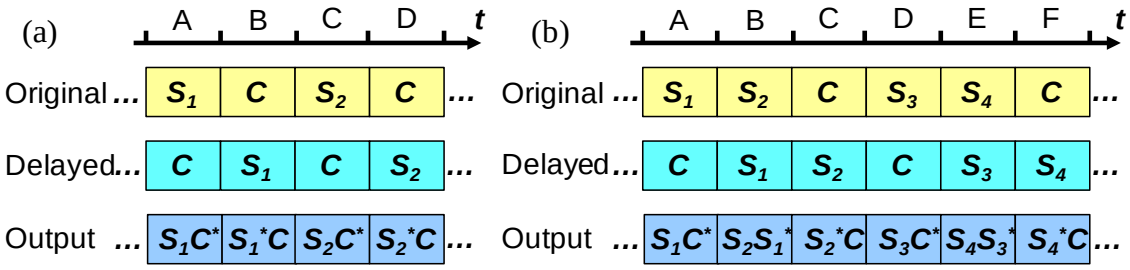

Fig. 4.7 Conceptual diagram of signal carrier interleaved (SCI) schemes: (a) SCI-DD with 1/2 SE, and (b) SCI-DD with 2/3 SE.

Beside the improved electrical SE, SCI-DD has another advantage due to the usage of the balanced receiver. The high common mode rejection ratio of the balanced PD makes SCI-DD naturally resistant to the impairment of SSBN. Without the need for SSBN cancellation, DSP becomes simpler, and low CSPR can be used to improve the receiver sensitivity. 102.4-Gb/s direct detection optical OFDM over 80 km SSMF has been successfully demonstrated using SCI scheme with 2/3 electrical SE in [153].

- **Stokes vector direct detection**

For Stokes vector direct detection (SV-DD) scheme, the DSB signal S and carrier C are multiplexed in two orthogonal polarizations as shown in Fig. 4.8(a) [152]. The

signal and carrier are provided by the same laser source. After modulating the signal on one path, the signal and carrier paths are combined using polarization beam combiner, loading the signal and carrier into two orthogonal polarizations. Along the transmission over fiber, the two polarization state rotates randomly, coupling with each other. For this reason, the signal and carrier cannot be separated directly using a polarization beam splitter at the receiver. The polarization state has to be recovered before the recovery of the signal. It is easy to recover the polarization state in the coherent system, where the polarization state is represented by a 2-dimension Jones vector. The elements of the Jones vector are the optical fields in the two polarizations, which can be easily obtained by coherent detection with the help of local oscillator. A 2x2 MIMO is then applied to the received Jones vector to recover the original polarization state. The optical fields in two polarizations cannot be directly obtained using direct detection. Therefore, it is a nature way to represent the polarization state in term of Stokes vector in the direct detection system. Given a polarization state $[E_x\ E_y]^T$ in Jones space, the corresponding representation in Stokes space is $[|E_x|^2+|E_y|^2\ |E_x|^2-|E_y|^2\ 2\mathrm{Re}\{E_xE_y^*\}\ 2\mathrm{Im}\{E_xE_y^*\}]^T$, where E_x and E_y are the electrical fields in the two orthogonal polarizations and the superscript 'T' represents transpose [127]. The first element is total power in two polarizations, the second element is the power difference between the two polarizations, the third and fourth elements are the in-phase and quadrature parts of the beating between the two polarizations. Since the first element can be calculated from the other three elements, the 4-dimensioanl vector usually is simplified as a 3-dimensional vector $[|E_x|^2-|E_y|^2\ 2\mathrm{Re}\{E_xE_y^*\}\ 2\mathrm{Im}\{E_xE_y^*\}]^T$. After the Stokes vector is obtained by a 3-dimensional detection at the receiver, the polarization recovery is then implemented in Stokes space.

The Stokes vector after fiber transmission is detected using the direct detection scheme shown in Fig. 4.8(b) [152]. After splitting the two orthogonal polarizations, X_{out} and Y_{out}, using a polarization beam splitter (PBS), each polarization is divided into two paths by a 3 dB coupler. One path from X_{out} and Y_{out} is fed into a 90° optical hybrid followed by two balanced detectors (BDs), producing the outputs of $\mathrm{Re}\{X_{out}Y_{out}^*\}$ and $\mathrm{Im}\{X_{out}Y_{out}^*\}$. The other path is fed into a BD directly, producing the output of $|X_{out}|^2-|Y_{out}|^2$. The three outputs provide us the three elements of the received Stokes vector $[|X_{out}|^2-|Y_{out}|^2\ 2\mathrm{Re}\{X_{out}Y_{out}^*\}\ 2\mathrm{Im}\{X_{out}Y_{out}^*\}]^T$.

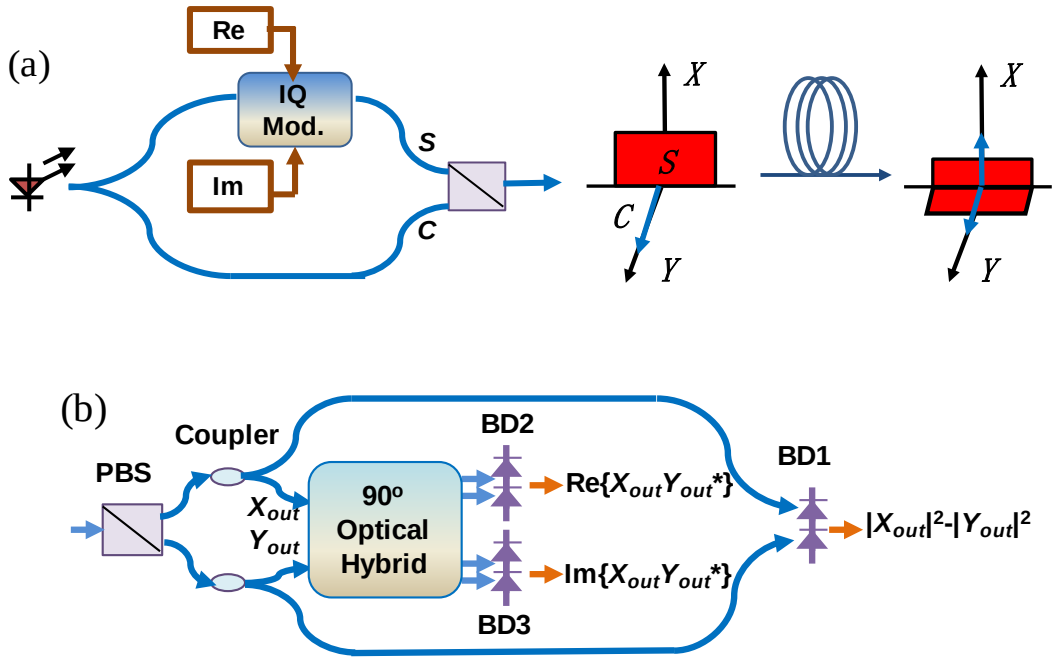

Fig. 4.8 Structures of (a) transmitter and (b) 3-dimensional receiver for the Stokes vector direct detection.

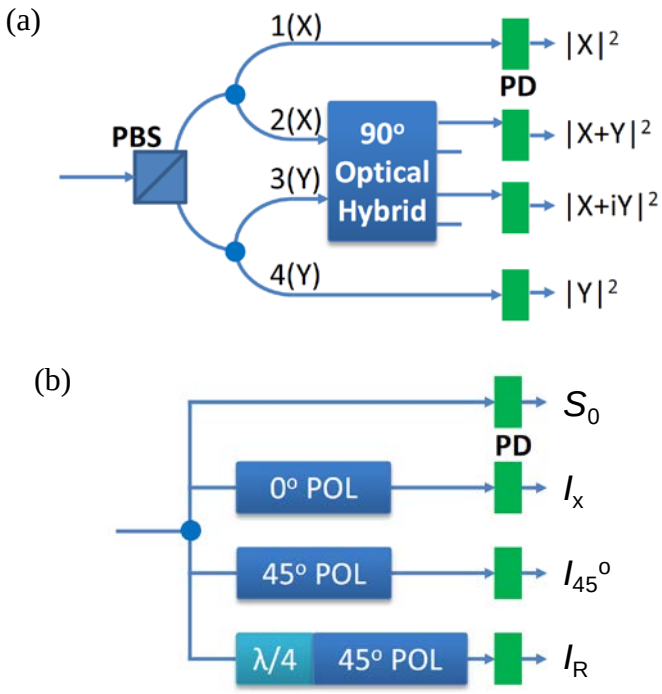

Fig. 4.9 Stokes vector detection schemes with single-ended PDs.

The Stokes vector can also be obtained by single-ended PD using the two schemes shown in Fig. 4.9. In Fig. 4.9(a), the outputs of the four PDs gives us $|X_{out}|^2$, $|Y_{out}|^2$, $|X_{out}+Y_{out}|^2$ and $|X_{out}+jY_{out}|^2$. The Stokes vector can be obtained from the four outputs as

$$\begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & -1 \\ -1 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} |X_{out}|^2 \\ |X_{out}+Y_{out}|^2 \\ |X_{out}+jY_{out}|^2 \\ |Y_{out}|^2 \end{bmatrix}. \tag{4.11}$$

In Fig. 4.9(b), the total power S_0 is detected by the first PD. The PD after the 0° polarizer detects the power in X polarization, the PD after the 45° polarization gives us $|X_{out}+Y_{out}|^2$, and the PD after $\lambda/4$ wave plate and 45° polarizer gives us $|X_{out}+jY_{out}|^2$. The Stokes vector can be obtained by

$$\begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ -1 & 0 & 2 & 0 \\ -1 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} S_0 \\ I_x \\ I_{45^0} \\ I_R \end{bmatrix}. \tag{4.12}$$

Assuming the transmitted signal of $[S\ C]^T$, the received Stokes vector $[|X_{out}|^2\text{-}|Y_{out}|^2\ 2\mathrm{Re}\{X_{out}Y_{out}^*\}\ 2\mathrm{Im}\{X_{out}Y_{out}^*\}]^T$ is related to the transmitted Stokes vector $[|S|^2\text{-}|C|^2\ 2\mathrm{Re}\{SC^*\}\ 2\mathrm{Im}\{SC^*\}]^T$ by a 3x3 channel matrix $\boldsymbol{V}$, given by

$$\begin{bmatrix} |X_{out}|^2-|Y_{out}|^2 \\ 2\,\mathrm{Re}\{X_{out}Y_{out}^*\} \\ 2\,\mathrm{Im}\{X_{out}Y_{out}^*\} \end{bmatrix} = \boldsymbol{V}\cdot \begin{bmatrix} |S|^2-|C|^2 \\ 2\,\mathrm{Re}\{SC^*\} \\ 2\,\mathrm{Im}\{SC^*\} \end{bmatrix}, \quad \boldsymbol{V} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}, \tag{4.13}$$

where h_{mn} is the element of the channel matrix $\boldsymbol{V}$ of m-th row and n-th column. The channel matrix $\boldsymbol{V}$ indicates the linear effect of fiber channel. Once $\boldsymbol{V}$ is known, the transmitted Stokes vector can be recovered by multiplying $\boldsymbol{V}^{-1}$ to the received Stokes vector.

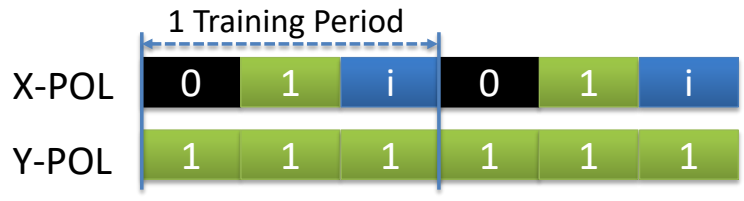

Fig. 4.10 Training symbols for Stokes channel estimation.

The channel matrix $\boldsymbol{V}$ can be acquired using polarization training symbols. As shown in Fig. 4.10, a training period is divided into three slots. For each slot, S is set to be 0, 1 or j, while C maintains to be 1. The polarization states of the three slots, [0

$1]^T$, $[1\ 1]^T$ and $[j\ 1]^T$, correspond to the Stokes vectors of $[1\ 0\ 0]^T$, $[0\ 1\ 0]^T$ and $[0\ 0\ 1]^T$ (ignoring the constant factors). Plugging them into Eq. (4.13), the three columns of the channel matrix $[h_{11}\ h_{21}\ h_{31}]^T$, $[h_{12}\ h_{22}\ h_{32}]^T$ and $[h_{13}\ h_{23}\ h_{33}]^T$ are obtained at the outputs. After recovering the transmitted Stokes vector, $[|S|^2\text{-}|C|^2\ \ 2\mathrm{Re}\{SC^*\}\ \ 2\mathrm{Im}\{SC^*\}]^T$, we arrive at SC^* by combining the last two elements in the vector, from which the phase-diverse DSB signal is fully developed. The SSBN is moved to the first element of the Stokes vector during the polarization recovery, leaving the signal free from the impact of SSBN. The high-performance of SV-DD is manifested by its two major advantages: (1) SV-DD can be used to detect signal with any polarization state, while achieving high electrical SE of 100%, and (2) the SSBN is automatically cancelled during the polarization recovery, without needing for high CSPR. A 160 Gb/s SV-DD systems over 160 km SSMF has been experimentally demonstrated in [152].

SV-DD provides a novel way to implement the DSP enabled polarization demultiplexing in DD system. The application of this scheme has been extended to the intensity modulation based direct detection (IM/DD) to improve its SE. By transmitting the intensity modulated signal in two polarizations and detecting the signal using SV-DD, polarization division multiplexed (PDM) IM/DD is achieved in [154] with the data rate of 224 Gb/s. By modulating the signal on 4 polarization states and tracking the Stokes vector at the receiver, [155] is able to obtain IM/DD link with the data rate of 128 Gb/s.

Other than the modulation scheme shown in Fig. 4.8, the signal can be modulated with different schemes in Stokes vector direct detection system. In Fig. 4.11(a), the real signals are modulated to the intensity of the two polarizations. In other words, the signal is modulated to the first two elements of the Stokes vector, $S_0 = |E_x|^2 + |E_y|^2$ and $S_1 = |E_x|^2 - |E_y|^2$. After polarization recovery, the signal is obtained from as $|E_x|^2 = (S_0 + S_1)/2$, $|E_y|^2 = (S_0 - S_1)/2$. In Fig. 4.11(b), one polarization is modulated with only intensity information, while the other polarization is modulated with both intensity and phase information. After polarization recovery, the intensity information of the two polarizations is acquired from the first two elements of the Stokes vector, $|E_x|^2 + |E_y|^2$ and $|E_x|^2 - |E_y|^2$, while the phase information is acquired from the beating

term $E_x E_y^*$ in the last two elements of the Stokes vector. In Fig. 4.11(c), the optical source is first modulated with an intensity modulator. Then, the optical source is divided into two paths with power ratio of $\cos(\theta/2)/\sin(\theta/2)$ using another intensity modulator. By modulating one polarization with phase $\exp(i\delta)$, the corresponding Stokes vector at the transmitter is $S_0[1 \quad \cos\theta \quad \sin\theta\cos\delta \quad \sin\theta\sin\delta]$.

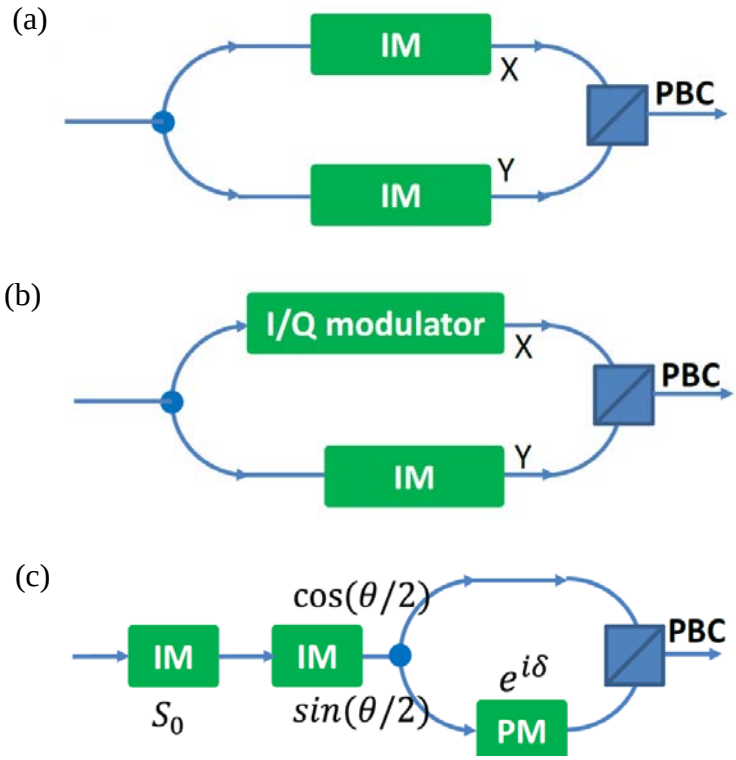

Fig. 4.11 Different modulation schemes for Stokes vector direct detection system.

4.4 Comparison of DD schemes using advanced modulation formats

The performance and complexity of DD schemes using advanced modulation formats are compared with the polarization multiplexed coherent detection in Table 1. Obviously, coherent detection has the best performance. The high performance of coherent detection is based on its expensive polarization-diversity transceiver and sophisticated DSP. On the contrary, cost-efficient transmission system can be obtained by using DD schemes. However, their performances are sacrificed by their simple transceiver and DSP. DD schemes using single-ended PD have the lowest electrical SE. Since a guard band has to be reserved to accommodate the nonlinear noise produced at sing-ended PD, offset SSB only achieves the electrical SE of 25%. The electrical SE is improved to 50% by virtual SSB and BPS-DD, where iterative SSBN cancellation is employed to eliminate the nonlinear noise. After turning to

balanced receiver, SCI-DD achieves the electrical SE of 66.7%. The highest electrical SE is achieved by SV-DD with its 3-dimensional detection.

Table 4.1 Comparison of advanced DD modulation formats. E-SE: ELECTRICAL SPECTRUM EFFICIENCY NORMALIZED TO SINGLE-POLARIZATION COHERENT DETECTION; MOD.: MODULATOR; IM: INTENSITY MODULATOR; IQ-M.: I/Q MODULATOR; PD: PHOTO-DETECTOR; B-PD: BALANCE PD.

Scheme	**E-SE**	**Mod.**	**PD**	**Required OSNR @ 10^{-2} BER**	**Additional DSP**
Offset SSB	25%	1 IM	1	18 dB for 10Gb/s 8QAM [55]	-
Virtual SSB	50%	1 IQ-M	1	13 dB for 10Gb/s 4QAM [55]	Iterative SSBN canc.
BPS-DD	50%	1 IQ-M	1	32 dB for 49.4 Gb/s 8QAM [150]	Iterative SSBN canc.
SCI-DD	66.7%	1 IQ-M	2 B-PD	22 dB for 43.2 Gb/s 16 QAM [151]	-
SV-DD	100%	1 IQ-M	3 B-PD	27 dB for 160 Gb/s 16 QAM [152]	3x3 real MIMO
POL-MUX-CO	200%	2 IQ-M	4 B-PD	12 dB for 107 Gb/s 4QAM [143]	2x2 complex MIMO Frequency offset Phase noise

4.5 Channel impairments compensation for Stokes vector direct detection

4.5.1 Mitigation of PMD impairment in SV-DD system

PMD effect induces a time delay $\Delta\tau$ between the two principal polarizations during fiber transmission. Under the first-order PMD approximation, the input polarization state $[X_{in}\ Y_{in}]^{\mathrm{T}}$ and the output polarization state $[X_{out}\ Y_{out}]^{\mathrm{T}}$ in Jones space are related by [156]

$$\begin{bmatrix} X_{out}(\omega) \\ Y_{out}(\omega) \end{bmatrix} = \begin{bmatrix} a_2 & b_2 \\ -b_2^* & a_2^* \end{bmatrix} \begin{bmatrix} e^{j\omega\Delta\tau/2} & 0 \\ 0 & e^{-j\omega\Delta\tau/2} \end{bmatrix} \begin{bmatrix} a_1 & b_1 \\ -b_1^* & a_1^* \end{bmatrix} \begin{bmatrix} X_{in}(\omega) \\ Y_{in}(\omega) \end{bmatrix}, \qquad (4.14)$$

where the first and third matrices at the right side of the equation denote the polarization rotations at the fiber input and output, and the second matrix denotes the

differential phase shift $\omega\Delta\tau$ between two principle polarizations caused by the time delay $\Delta\tau$. Eq. (4.14) indicates the frequency dependence of the polarization evolution.

Since the polarization training symbols in SV-DD scheme are signals at the carrier frequency ω_c, the channel matrix $V(\omega_c)$ obtained from the training symbols gives us the information of the polarization evolution at the frequency ω_c. Using the channel matrix $V(\omega_{c)}$ to recover the polarization state, only the polarization state at the frequency ω_c can get fully recovered. The misalignment of the polarization states between the transmitted signal and the recovered one at the other frequencies causes impairment to SV-DD systems. Denoting the transmitted signal as $S(t)=\sum_i S_i e^{j\omega_i t}$, where ω_i and S_i are the baseband frequency and modulation symbol for the *i*-th subcarrier, we calculate the recovered Stokes vector from Eq. (4.14). Combining the second and third components of the Stokes vector, we arrive at [157]

$$\begin{aligned} E_x E_y^* = &\sum_i [\cos(\omega_i \Delta\tau/2) + j(|b_1|^2 - |a_1|^2)\sin(\omega_i \Delta\tau/2)] S_i e^{j\omega_i t} + 2ja_1^* b_1 \\ &\cdot\{\sum_i [\cos(\omega_i \Delta\tau/2) + j(|b_1|^2 - |a_1|^2)\sin(\omega_i \Delta\tau/2)] S_i e^{j\omega_i t}\} \cdot \{\sum_i \sin(\omega_i \Delta\tau/2) S_i e^{j\omega_i t}\}^*, \end{aligned} \tag{4.15}$$

where the first term on the right side of the equation denotes the beating between the recovered signal in *X* polarization and the carrier in *Y* polarization, and the second term denotes the beating between the recovered signal in *X* polarization and the residual signal in *Y* polarization due to PMD. Under the condition of $|\omega i\Delta\tau| \ll 1$, which is valid in short- and medium-reach optical transmissions, the first and second terms can be simplified as $S(t)=\sum_i S_i e^{j\omega_i t}$ and $N_{NL} = F \cdot S(t) \cdot (\sum_i \omega_i S_i e^{j\omega_i t})^*$, where N_{NL} is the PMD-induced nonlinear noise and $F = ja_1^* b_1 \Delta\tau$ is a constant factor determined by the rotation matrix in Eq. (4.14). When *F* and *S*(*t*) are known, the PMD induced noise N_{NL} can be estimated and subtracted from the recovered signal. We propose two different algorithms to estimate the PMD induced noise N_{NL}. The flow charts of algorithms A and B are shown in Figs. 4.12 and 4.13.

For both algorithms, specially-designed training symbols are transmitted to calculate the factor F, where only the odd subcarriers are filled with data, the even subcarriers are left empty. For the received training symbols, the first term of Eq. (4.15) falls into the odd subcarriers, while the second term falls into the even subcarriers. As shown in the flow charts of Figs. 4.12(a) and 4.13(a), after performing FFT for the received training symbols, the first and second terms of Eq. (4.15) are

obtained by separating the odd and even subcarriers. Different approaches are adopted by algorithm A and B when calculating *F*.

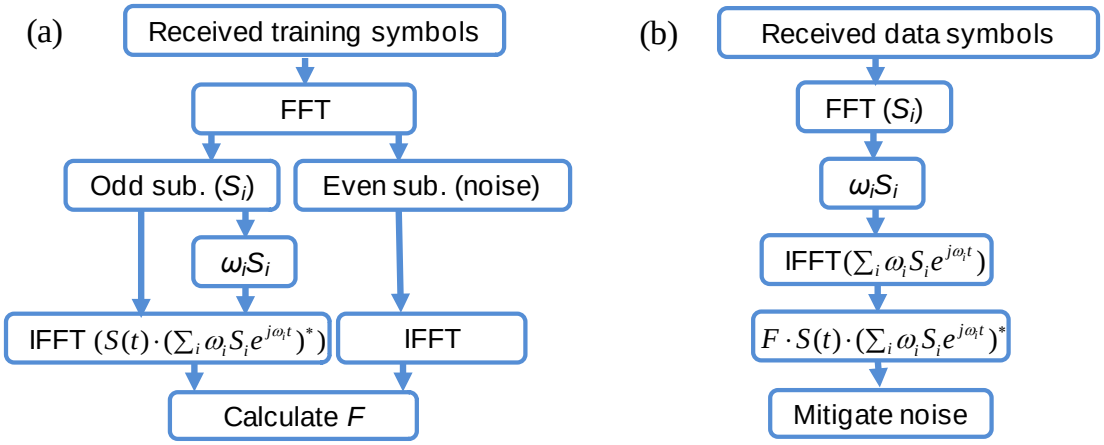

Fig. 4.12 Flow chart of algorithm A. (a) Calculate F using training symbols. (b) Mitigate the PMD induced noise in data symbols.

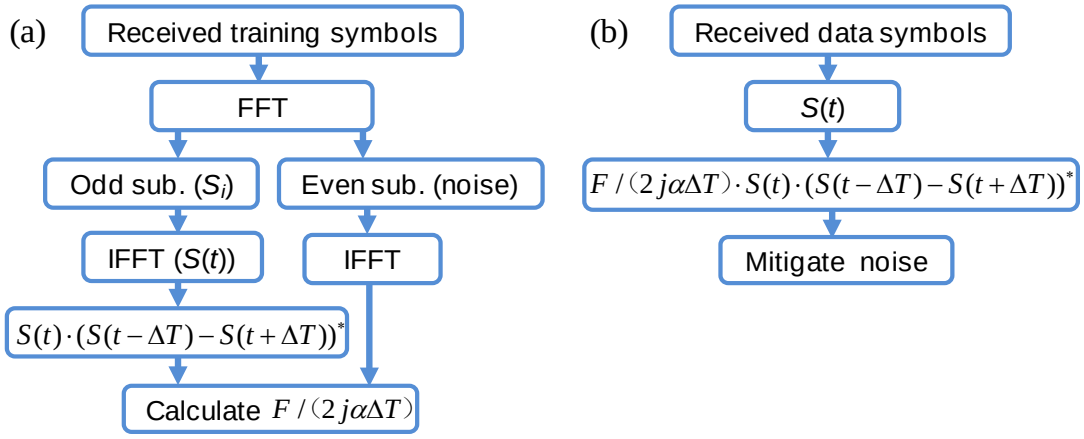

Fig. 4.13 Flow chart of algorithm B. (a) Calculate F using training symbols. (b) Mitigate the PMD induced noise in data symbols.

In algorithm A shown in Fig. 4.12(a), $\omega_i S_i$ is calculated in the frequency domain. S_i and $\omega_i S_i$ are then converted to $S(t)$ and $\sum_i \omega_i S_i e^{j\omega_i t}$ by IFFT. By comparing $S(t)\cdot(\sum_i \omega_i S_i e^{j\omega_i t})^*$ with the second term, *F* is obtained. In algorithm B shown in Fig. 4.13(a), signal in the frequency domain S_i is directly converted to $S(t)$ by IFFT. Then, $S(t)$ is advanced and delayed with one sampling interval ΔT to obtain $S(t+\Delta T)$ and $S(t-\Delta T)$. Subtracting $S(t+\Delta T)$ from $S(t-\Delta T)$, we obtain

$$S(t-\Delta T)-S(t+\Delta T)=\sum_i S_i e^{j\omega_i t}e^{-j\omega_i \Delta T}-\sum_i S_i e^{j\omega_i t}e^{j\omega_i \Delta T}=-2j\sum_i \sin(\omega_i \Delta T)S_i e^{j\omega_i t}. \quad (4.16)$$

Assuming that $\sin(\omega_i \Delta T) \approx \alpha \omega_i \Delta T$ under the condition of $|\omega_i \Delta T| \leq \pi/2$, where α is a constant, and $|\omega_i \Delta T| \leq \pi/2$ is satisfied when the oversampling rate is larger than or equals to 2, the second term in Eq. (4.15) can be rewritten as $F/(2j\alpha\Delta T) \cdot S(t) \cdot (S(t-\Delta T) - S(t+\Delta T))^*$. By comparing $S(t) \cdot (S(t-\Delta T) - S(t+\Delta T))^*$ with the second term, we obtain $F/(2j\alpha\Delta T)$. The PMD induced noise in the data symbols is mitigated following the flow charts of Figs. 4.12(b) and 4.13(b). In algorithm A, FFT is performed to calculate $\omega_i S_i$, and then IFFT is performed to obtain $F \cdot S(t) \cdot (\sum_i \omega_i S_i e^{j\omega_i t})^*$. In algorithm B, $F/(2j\alpha\Delta T) \cdot S(t) \cdot (S(t-\Delta T) - S(t+\Delta T))^*$ is obtained directly from the time domain signal $S(t)$. Though the assumption of $\sin(\omega_i \Delta T) \approx \alpha \omega_i \Delta T$ may induce slight inaccuracy during the noise estimation, algorithm B does not need an additional pair of FFT/IFFT for PMD mitigation in contrast with algorithms A. This leads to a much efficient DSP for algorithm B.

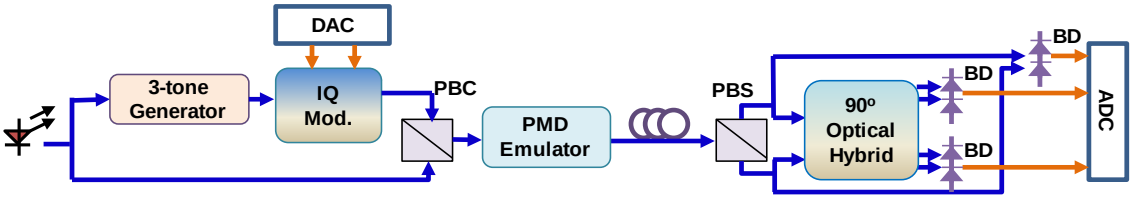

Fig. 4.14 Experimental setup to verify the PMD mitigation algorithms.

We use the two proposed algorithms to mitigate the PMD induced impairment in a 93-Gb/s SV-DD system. The experiment setup is illustrated in Fig. 4.14. A three-tone generator is applied in the signal path to obtain a broader optical bandwidth. 16 QAM OFDM signal is generated by an arbitrary waveform generator (AWG) and modulated onto the three tones. The FFT size is 4096 with the central 3420 subcarriers filled with data. A total of 100 data symbols are transmitted, where the first 4 symbols are used as the training symbols by setting the even subcarriers null. 128 point of cyclic prefix (CP) is added into each symbol. The polarization training symbols are added before the OFDM frame. The AWG operates at a sampling rate of 10 GSa/s, leading to an optical bandwidth of 8.35 GHz for one tone, and 25 GHz for three tones. The raw data rate is 4x25=100 Gb/s. Counting the training symbols and CP, the data rate is reduced to 93 Gb/s. The length of the carrier path is matched with the signal path to eliminate the phase noise. The CSPR is maintained at 0 dB. The signal is transmitted over an 80-km fiber. To adjust the PMD in the transmission link, a PMD emulator (FiberPro PE4200) is placed at the input of the fiber. A polarization controller (PC) is

inserted before the PMD emulator to control the input polarization angle. In our experiment, the angle is set to be 45°, corresponding to the worst impairment. The Stokes vector is detected using a 90° optical hybrid and three BDs at the receiver. The electrical signal is sampled by an oscilloscope at a sampling rate of 50 GSa/s. The collected data is processed offline. The PMD induced impairment is to be mitigated using the two proposed algorithms after the polarization recovery.

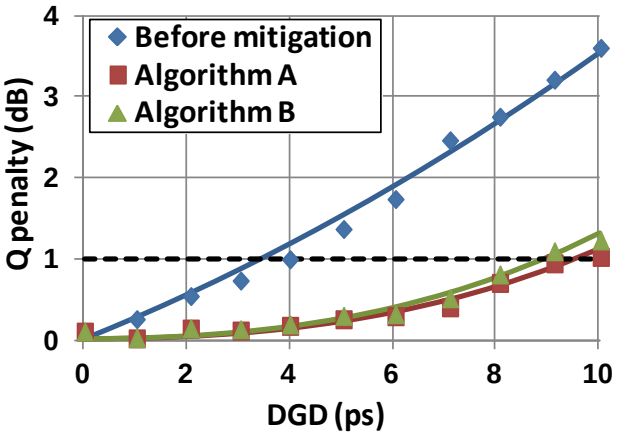

Fig. 4.15 Q penalty as a function of DGD before and after PMD mitigation.

The system performances before and after applying the algorithms are compared. Considering the small differential group delay (DGD) (smaller than 0.36 ps) in the 80-km fiber, the DGD value in the transmission link is dominated by the PMD emulator with adjustable DGD ranging from 0 to 10 ps. The Q penalty is measured as the PMD-induced Q degradation when the initial Q is set at 16.5 dB. The Q penalty as a function of DGD is plotted in Fig. 4.15. According to Eq. (4.15), the PMD induced noise increases rapidly with the DGD value. We can observe a significant increase of Q penalty with the increase of DGD before PMD mitigation. The Q penalty exceeds 1 dB (the dashed line in Fig. 4.15) when DGD is larger than 3.5 ps and reaches 3.5 dB at the DGD of 10 ps. After mitigating the PMD induced impairment using algorithms A and B, the Q penalty is greatly reduced. The Q penalty reaches 1 dB at DGD of 9.5 ps using algorithm A, and at DGD of 9 ps using algorithm B. Algorithms A and B extend the system's PMD tolerance from 3.5 to 9.5 and 9 ps, respectively. The system suffers from random noise when PMD is absent, which is regarded as additive white Gaussian noise (AWGN) with a constant power spectral density. Additional nonlinear noise is then induced into the system when PMD is present. Fig. 4.16(a) shows the spectra of signal and noise at DGD of 10 ps before applying PMD mitigation

algorithms. The noise level is significantly increased, and the noise power spectral density becomes uneven due to PMD induced noise. Figs. 4.16(b) and (c) show the spectra after applying algorithms A and B. The greatly reduced noise level in Figs. 4.16(b) and (c) demonstrates the effectiveness of the two proposed algorithms in mitigating the PMD induced impairment. Furthermore, algorithm B is almost as effective as algorithm A. Considering their DSP complexity, algorithm B would be the preferred choice for the practical implementation.

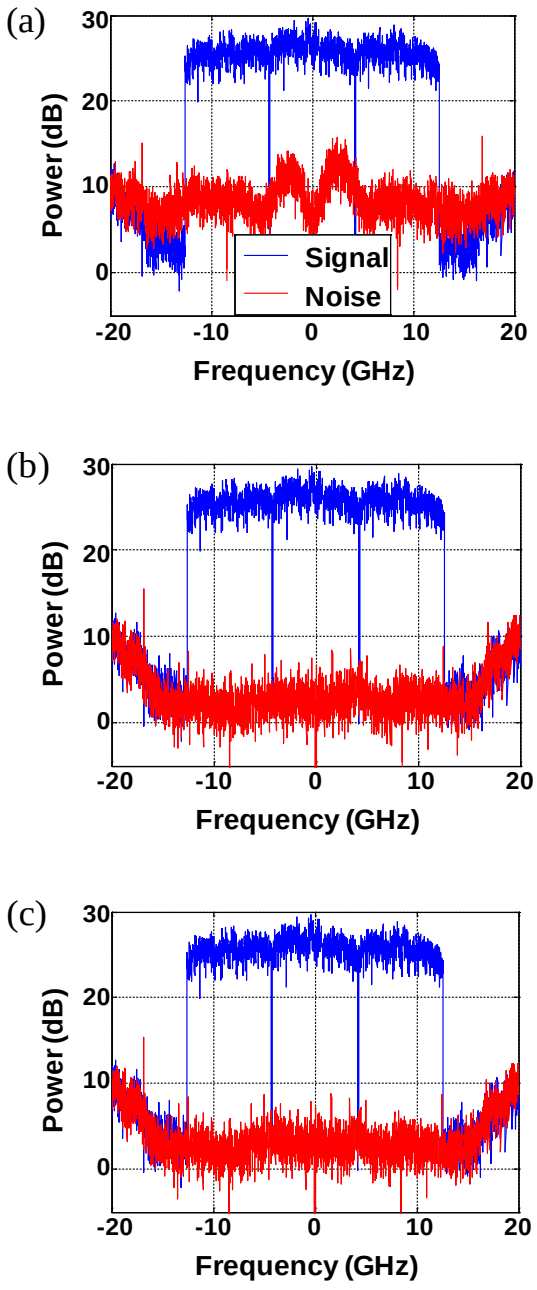

Fig. 4.16 Signal and noise spectra (a) before PMD mitigation, (b) after PMD mitigation using algorithm A, and (c) after PMD mitigation using algorithm B.

4.5.2 Mitigation of PDL impairment in SV-DD system

The asymmetric structure of optical components and the polarization dependent gain of erbium-doped fiber amplifiers (EDFAs) lead to unequal insert losses for two polarization modes, or so-called polarization-dependent loss (PDL). Assuming the attenuation difference between the two principal polarizations is 2α, the input polarization state $[S \;\; C]^{\mathrm{T}}$ and output polarization state $[X_{out} \;\; Y_{out}]^{\mathrm{T}}$ in Jones space are related by [158]

$$\begin{bmatrix} X_{out} \\ Y_{out} \end{bmatrix} = \begin{bmatrix} a_2 & b_2 \\ -b_2^* & a_2^* \end{bmatrix} \begin{bmatrix} \sqrt{1+\alpha} & 0 \\ 0 & \sqrt{1-\alpha} \end{bmatrix} \begin{bmatrix} a_1 & b_1 \\ -b_1^* & a_1^* \end{bmatrix} \begin{bmatrix} S \\ C \end{bmatrix}, \tag{4.17}$$

where the first and third matrices on the right side of the equation are the rotation matrices, similar to those in Eq. (4.14), and the second matrix indicates the differential loss between two principal polarizations due to PDL. The PDL value in decibel is $\Gamma = 10\log_{10}[(1+\alpha)/(1-\alpha)]$. Converting the output vector $[X_{out} \;\; Y_{out}]^{\mathrm{T}}$ into Stokes space, the three elements of the output Stokes vector can be calculated from Eq. (4.17) as

$$|X_{out}|^2 - |Y_{out}|^2 = A_1|S|^2 + B_1|C|^2 + C_1\,\mathrm{Re}\{SC^*\} + D_1\,\mathrm{Im}\{SC^*\} \tag{4.18}$$

$$\mathrm{Re}\{X_{out}Y_{out}^*\} = A_2|S|^2 + B_2|C|^2 + C_2\,\mathrm{Re}\{SC^*\} + D_2\,\mathrm{Im}\{SC^*\} \tag{4.19}$$

$$\mathrm{Im}\{X_{out}Y_{out}^*\} = A_3|S|^2 + B_3|C|^2 + C_2\,\mathrm{Re}\{SC^*\} + D_3\,\mathrm{Im}\{SC^*\}. \tag{4.20}$$

where symbols, A_n, B_n, C_n and D_n, are determined by the PDL coefficient α and the elements of rotation matrices in Eq. (4.17). The coefficients A_n, B_n, C_n and D_n (n=1, 2, 3) in Eqs. (4.18), (4.19) and (4.20) are given below:

$$A_1 = [(|a_2|^2 - |b_2|^2)(|a_1|^2(1+\alpha) - |b_1|^2(1-\alpha)) - 2(a_2b_2^*a_1b_1 + a_2^*b_2a_1^*b_1^*)\sqrt{(1+\alpha)(1-\alpha)}]$$

$$B_1 = [(|a_2|^2 - |b_2|^2)(|b_1|^2(1+\alpha) - |a_1|^2(1-\alpha)) + 2(a_2b_2^*a_1b_1 + a_2^*b_2a_1^*b_1^*)\sqrt{(1+\alpha)(1-\alpha)}]$$

$$C_1 = [2(|a_2|^2 - |b_2|^2)(a_1b_1^* + a_1^*b_1) + 2(a_2b_2^*(a_1^2 - b_1^2) + a_2^*b_2(a_1^{*2} - b_1^{*2}))\sqrt{(1+\alpha)(1-\alpha)}]$$

$$D_1 = [2i(|a_2|^2 - |b_2|^2)(a_1b_1^* - a_1^*b_1) + 2(a_2b_2^*(a_1^2 + b_1^2) - a_2^*b_2(a_1^{*2} + b_1^{*2}))\sqrt{(1+\alpha)(1-\alpha)}]$$

$$A_2 = \mathrm{Re}\{-a_2b_2(|a_1|^2(1+\alpha) - |b_1|^2(1-\alpha)) - (a_2^2a_1b_1 - b_2^2a_1^*b_1^*)\sqrt{(1+\alpha)(1-\alpha)}\}$$

$$B_2 = \mathrm{Re}\{a_2b_2(|a_1|^2(1-\alpha) - |b_1|^2(1+\alpha)) + (a_2^2a_1b_1 - b_2^2a_1^*b_1^*)\sqrt{(1+\alpha)(1-\alpha)}\}$$

$$C_2 = \mathrm{Re}\{-2a_2b_2(a_1b_1^* + a_1^*b_1) + (a_2^2(a_1^2 - b_1^2) + b_2^2(b_1^{*2} - a_1^{*2}))\sqrt{(1+\alpha)(1-\alpha)}\}$$

$$D_2 = \mathrm{Im}\{-2a_2b_2(a_1b_1^* - a_1^*b_1) + (a_2^2(a_1^2 + b_1^2) + b_2^2(b_1^{*2} + a_1^{*2}))\sqrt{(1+\alpha)(1-\alpha)}\}$$

$$A_3 = \mathrm{Im}\{-a_2b_2(|a_1|^2(1+\alpha) - |b_1|^2(1-\alpha)) - (a_2^2a_1b_1 - b_2^2a_1^*b_1^*)\sqrt{(1+\alpha)(1-\alpha)}\}$$

$$B_3 = \mathrm{Im}\{a_2b_2(|a_1|^2(1-\alpha) - |b_1|^2(1+\alpha)) + (a_2^2a_1b_1 - b_2^2a_1^*b_1^*)\sqrt{(1+\alpha)(1-\alpha)}\}$$

$$C_3 = \mathrm{Im}\{-2a_2b_2(a_1b_1^* + a_1^*b_1) + (a_2^2(a_1^2 - b_1^2) + b_2^2(b_1^{*2} - a_1^{*2}))\sqrt{(1+\alpha)(1-\alpha)}\}$$

$$D_3 = \mathrm{Re}\{-2a_2b_2(a_1b_1^* - a_1^*b_1) + (a_2^2(a_1^2 + b_1^2) + b_2^2(b_1^{*2} + a_1^{*2}))\sqrt{(1+\alpha)(1-\alpha)}\}$$

When no PDL exists in the system, α=0, we have A_n=-B_n (n=1, 2, 3). The first and second terms in Eqs. (4.18), (4.19) and (4.20) can be combined as $A_n(|S|^2-|C|^2)$. Therefore, the output Stokes vector $[|X_{out}|^2-|Y_{out}|^2\ 2\mathrm{Re}\{X_{out}Y_{out}^*\}\ 2\mathrm{Im}\{X_{out}Y_{out}^*\}]^T$ is linear transformation of the input Stokes vector $[|S|^2-|C|^2\ 2\mathrm{Re}\{SC^*\}\ 2\mathrm{Im}\{SC^*\}]^T$. The linear relationship is expressed by the 3x3 channel matrix in Eq. (4.13). On the other hand, when PDL exists in the system, we arrive at $|A_n|\neq|B_n|$. In that case, the $|S|^2$ and $|C|^2$ terms cannot be combined, leading to the invalidity of the linear relationship of Eq. (4.13). Therefore, the 3x3 channel estimation, described in Section 4.3.2, fails to acquire the channel information accurately. Impairment is induced to SV-DD systems during the polarization recovery.

We propose two different approaches to mitigate the impact of PDL. In the first approach, we adopt 4x4 channel estimation instead of 3x3 channel estimation. We separate the first element of the Stokes vector, $|E_x|^2-|E_y|^2$, into 2 individual elements, $|E_x|^2$ and $|E_y|^2$, and extend the 3-dimensional Stokes vector to a 4-dimensional vector $[|E_x|^2\ |E_y|^2\ 2\mathrm{Re}\{E_xE_y^*\}\ 2\mathrm{Im}\{E_xE_y^*\}]^T$. To detect the 4-dimensional vector, the 'BD1' in Fig. 4.8(b) is replaced with two single-ended PDs. The input 4-dimensional vector $[|S|^2\ |C|^2\ 2\mathrm{Re}\{SC^*\}\ 2\mathrm{Im}\{SC^*\}]^T$ and output 4-dimensional vector $[|\mathrm{X_{out}}|^2\ |\mathrm{Y_{out}}|^2\ 2\mathrm{Re}\{X_{out}Y_{out}^*\}\ 2\mathrm{Im}\{X_{out}Y_{out}^*\}]^T$ are related by a 4x4 channel matrix H

$$\begin{bmatrix} |X_{out}|^2 \\ |Y_{out}|^2 \\ \mathrm{Re}\{X_{out}Y_{out}\} \\ \mathrm{Im}\{X_{out}Y_{out}\} \end{bmatrix} = \boldsymbol{H} \cdot \begin{bmatrix} |S|^2 \\ |C|^2 \\ \mathrm{Re}\{SC^*\} \\ \mathrm{Im}\{SC^*\} \end{bmatrix}, \quad \boldsymbol{H} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \\ h_{41} & h_{42} & h_{43} & h_{44} \end{bmatrix}, \qquad (4.21)$$

where h_{nm} denotes the element of the channel matrix H at the n-th row and m-th column. By extending the Stokes vector to 4 dimensions, linear communication channel can be obtained even at the presence of PDL. To acquire the 4x4 channel matrix H, specially-designed polarization training symbols are used. A training symbol is divided into 4 slots, where the signal S is set to be 0, 1, -1 and i respectively, while the carrier C remains to be 1. The 4-dimensional vectors for the 4 slots are T_1=[0 1 0 0]^T, T_2=[1 1 1 0]^T, T_3=[1 1 -1 0]^T and T_4=[1 1 0 1]^T. Taking them to Eq. (4.21), we obtain the vectors R_1, R_2, R_3 and R_4 at the output. The columns of the 4x4 channel matrix can be calculated by

$$[h_{11} \quad h_{21} \quad h_{31} \quad h_{41}]^T = (\boldsymbol{R}_2 + \boldsymbol{R}_3)/2 - \boldsymbol{R}_1 \tag{4.22}$$

$$[h_{12} \quad h_{22} \quad h_{32} \quad h_{42}]^T = \boldsymbol{R}_1 \tag{4.23}$$

$$[h_{13} \quad h_{23} \quad h_{33} \quad h_{43}]^T = (\boldsymbol{R}_2 - \boldsymbol{R}_3)/2 \tag{4.24}$$

$$[h_{14} \quad h_{24} \quad h_{34} \quad h_{44}]^T = \boldsymbol{R}_4 - (\boldsymbol{R}_2 + \boldsymbol{R}_3)/2. \tag{4.25}$$

After the 4x4 channel matrix is obtained, the rest DSP procedures are the same as for 3-dimensional Stokes vector.

In the second approach, the 3x3 channel training described in Section 4.3.2 is still applied, while additional algorithm is used to mitigate the impairment caused by PDL. For a better understanding of the impact of PDL, we rewrite Eqs. (4.18), (4.19) and (4.20) as

$$\left|X_{out}\right|^2 - \left|Y_{out}\right|^2 = A_1(\left|S\right|^2 - \left|C\right|^2) + (A_1 + B_1)\left|C\right|^2 + C_1 \,\mathrm{Re}\{SC^*\} + D_1 \,\mathrm{Im}\{SC^*\} \tag{4.26}$$

$$\mathrm{Re}\{X_{out}Y_{out}^*\} = A_2(\left|S\right|^2 - \left|C\right|^2) + (A_2 + B_2)\left|C\right|^2 + C_2 \,\mathrm{Re}\{SC^*\} + D_2 \,\mathrm{Im}\{SC^*\} \tag{4.27}$$

$$\mathrm{Im}\{X_{out}Y_{out}^*\} = A_3(\left|S\right|^2 - \left|C\right|^2) + (A_3 + B_3)\left|C\right|^2 + C_2 \,\mathrm{Re}\{SC^*\} + D_3 \,\mathrm{Im}\{SC^*\}. \tag{4.28}$$

From Eqs. (4.26), (4.27) and (4.28), we can see that PDL introduces terms $(A_n+B_n)|C|^2$ (n=1, 2, 3) into the three elements of the received Stokes vector, disturbing the linear relationship between the input and output Stokes vector, where A_n, B_n are constant numbers determined by Eq. (4.17), and $|C|^2$=1 for SV-DD scheme. By estimating the DC terms $(A_n+B_n)|C|^2$ and removing them from the received Stokes vector, the linear relationship between the input and output Stokes vectors can be recovered. Training symbols with the input polarization states [S C]^T=[1 1]^T and [-1

1]^T are used to estimate the DC terms. Taking the training symbols into Eqs. (4.26), (4.27) and (4.28), the output Stokes vectors for the training symbols are [A_1+B_1+C_1 A_2+B_2+C_2 A_3+B_3+C_3]^T and [A_1+B_1-C_1 A_2+B_2-C_2 A_3+B_3-C_3]^T. Subtracting their sum [A_1+B_1 A_2+B_2 A_3+B_3]^T from the received Stokes vector, the linear relationship between the transmitted and received Stokes vector is recovered. After that, the channel training and the polarization recovery would be the same as described in Section 4.3.2.

The proposed PDL mitigation algorithms are evaluated by simulation. The simulation uses a similar setup as illustrated in Fig. 4.14, where the multi-band generator is removed, and PMD emulator is replaced by PDL emulator. 16 QAM OFDM signal is generated and modulated onto the CW laser source. The FFT size of OFDM signal is 4096, in which 2048 subcarriers are filled with signal. Different polarization training symbols for algorithms A and B are added before the OFDM frame. The sampling rate of DAC is set to be 50 GSa/s, resulting in a signal bandwidth of 25 GHz. Counting the CP and training symbols, the overall data rate for the described system is 95 Gb/s. The signal is transmitted over a 100 km SSMF with a chromatic dispersion (CD) of 17 ps/nm·km. The input polarization angle of the PDL emulator is set to be π/4, corresponding to the worst case. At the receiver, the signal is detected by 4-dimensional detection for algorithm A and by 3-dimensional detection for algorithm B. After the signal is received, algorithms A and B are carried out for the linear channel estimation and polarization recovery.

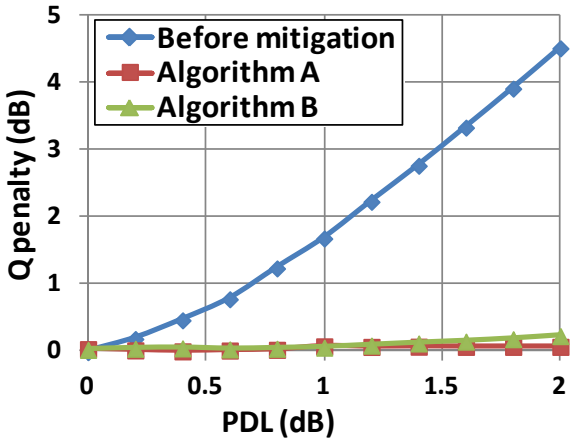

Fig. 4.17 Q penalty as a function of PDL before and after mitigation.

In simulation, optical signal to noise ratio (OSNR) is set to be 22.6 dB to achieve the Q value of 16.5 dB for a system without PDL. The system impairment caused by PDL

is measured by the Q penalty. The Q penalty as a function of the PDL, ranging from 0 to 2 dB, is plotted in Fig. 15. As explained in Eqs. (4.26), (4.27) and (4.28), PDL introduces DC terms into the three elements of the received Stokes vector, leading to inaccurate polarization recovery. Without applying PDL mitigation algorithm, the Q penalty increases dramatically with the PDL, and reaches 4.5 dB at PDL of 2 dB. On the other hand, the Q penalty maintains around 0 dB after either algorithm A or B is applied. The simulation result shows that both algorithms A and B can efficiently eliminate the PDL induced errors during the polarization recovery. Algorithm B is the preferred choice due to its simplicity without an additional PD.

4.6 Conclusion

Aiming to improve the electrical spectral efficiency and transmission reach of DD systems, various advanced modulation formats have been proposed for high-performance short-reach optical interconnects. Among these schemes, SV-DD achieves the highest electrical SE (100% SE compared to the single-polarization coherent detection), making it a promising candidate for future short-reach optical networks. Additionally, the performance of SV-DD systems can be impaired by the PMD and PDL effects. Novel algorithms have been proposed in this paper to mitigate the impairments caused by PMD and PDL. By applying these algorithms, high-performance SV-DD can be achieved even under severe impairments of PMD and PDL.

5 High-Performance Short-Reach Networks with Directly Modulated Laser

5.1 Introduction

With the rapid increase of data traffic, metro networks are facing the challenges of upgrading its capacity to 100 Gb/s and beyond in the near future. Shared by relatively fewer users, the cost becomes a primary consideration for the metro networks. Significant amount of efforts has been made to improve the performance of the metro networks while maintaining low cost [145-147, 159-161]. In Chapter 4, we have mentioned that the cost of short-reach networks can be reduced by using direct detain or direct modulation. The low-cost direct detection for short-reach application has been discussed in Chapter 4. In this chapter, we focus on the direct modulation scheme for high-performance short-reach optical networks.

In comparison with the in-phase/quadrature (IQ) modulator and external cavity laser (ECL) used for PDM-QAM, directly modulated laser (DML) has the merits of small form factor, low power consumption, and low cost. Therefore, it provides a promising solution for cost-efficient metro networks. In combination with DML, coherent detection is necessary to meet with the chromatic dispersion (CD) requirement in the metro networks [160,161]. 100 Gb/s transmission with pulse-amplitude-modulation (PAM) signal has been demonstrated in [161] using DML and coherent detection. For the conventional systems with DML, only the intensity modulation is considered during the signal recovery, despite the fact that both intensity and phase are acquired at the coherent receiver [160,161]. The conventional recovery scheme is referred as intensity-only recovery in the following text. Considering the chirp effect, the phase (or frequency) of the optical signal is modulated along with the intensity in DML [162-164]. By utilizing the additional information carried by phase (or frequency), the receiver sensitivity can be significantly improved. With increased OSNR margin, such systems can be upgraded to higher data rate or longer transmission distance. An approach to utilize the frequency modulation has been proposed in [161], which converts the frequency modulation into the intensity modulation with an optical filter. The distance between different modulation levels is enlarged after the conversion, improving the receiver sensitivity. However, the modulation information carried by phase (or frequency) is

not analyzed in detail. In this chapter, we demonstrate the recovery of polarization-division-multiplexed (PDM) directly modulated signal using combined information of intensity and phase. Detailed analysis will be provided for two intensity and phase combined signal recovery schemes.

Unlike the intensity, the phase is modulated with a more complicated modulation mechanism in DML. Special digital signal processing (DSP) has to be designed to use phase information for signal recovery [165-167]. In this chapter, we will introduce two intensity and phase combined signal recovery schemes. The organization of this chapter is as follows: section 5.2 introduces the conventional DML system. The phase modulation mechanism in DML will be analyzed in section 5.3, derived from the dynamic laser chirp equation. Based on that, the DSP used in the two schemes will be revealed in section 5.4, including the chirp characterization methods and signal recovery algorithms. The experimental demonstrations of the two schemes will be described in section 5.5, where the recovery of PDM directly modulated signal using intensity and phase combined recovery scheme is demonstrated for the first time with a data rate of 40 Gb/s over a transmission distance of 160 km. Improving the OSNR sensitivity by 5.5 dB and 10 dB respectively, the two intensity and phase combined recovery schemes show great advantages over the intensity-only recovery.

5.2 Conventional digital coherent system with DML

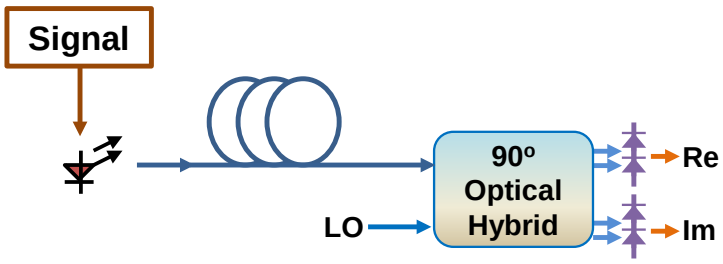

Fig. 5.1 Conceptual diagram of digital coherent system with DML.

Fig. 5.1 shows the conceptual diagram of the digital coherent system with DML. A low-cost DML is used as the transmitter, with the driven voltage encoded by the multi-level signal. The modulated output optical signal is transmitted through the fiber. A transmission distance over 100 km is enabled by the coherent detection at the receiver. At the coherent receiver, the optical signal is combined with local oscillator in a 90º optical hybrid, followed by two balanced PD. The local oscillator beats with the optical signal, down-converting the optical field to baseband. The in-phase and quadrature parts are given by the two balanced-PDs individually. After the coherent detection, the channel equalization is performed on the

obtained complex signal, including the chromatic dispersion (CD) compensation and polarization mode dispersion (PMD) compensation (if polarization-division-multiplexing is adopted at the transmitter).

A 100 Gb/s transmission over 400 km standard single mode fiber (SMF) has been demonstrated in [160], using polarization-division-multiplexed 4-level pulse amplitude modulation signal using directly modulated VCSEL. The symbol decision is based on the intensity of the recovered signal. Since the frequency and phase recovery are omitted during the DSP, the phase of the signal rotates randomly, forming 4 rings on the constellation plane. The space between the different levels can be enlarged using an optical filter, improving the system performance. Due to the chirp effect in directly modulated laser, the frequency of the symbols with high intensity is blue shifted compared to the symbols with low intensity. By using an optical filter with a centre frequency lower than that of the optical signal, the attenuation for the symbols with low intensity (red shifted) is larger than the symbols with high intensity (blue shifted). The optical filter converts the frequency modulation of the directly modulated laser into amplitude modulation, enlarging the intensity space between different modulation levels. By using the optical filter as the frequency to amplitude (FM/AM) modulation, an error free transmission of 83.33 Gb/s polarization division multiplexed PAM-4 signal over 400 km SSMF with 20% overhead FEC can be achieved in [161].

5.3 Phase modulation mechanism in DML

Controlled by the driving current, the variation of the electron density in DML leads to the modulation of the output signal [162,163]. As the gain increases proportionally with the electron density, the output intensity changes linearly with the modulation level. Due to the linear modulation response, the optical intensity provides a straightforward way to recover the signal in the transmission system with DML [160-161, 168-169]. On the other hand, the modulation response of output phase (or frequency) is more complicated. When the refractive index varies with the electron density, the frequency of the output signal is shifted as well. Derived from the rate equations governing the laser dynamics, the frequency chirp $\Delta\nu(t)$ is a function of the output power $P(t)$, and can be expressed as [162,163].

$$\Delta\nu(t) = \frac{\alpha}{4\pi}\left\{\frac{d}{dt}\ln P(t) + \kappa P(t)\right\}. \tag{5.1}$$

As indicated in Eq. (5.1), the frequency chirp consists of two components. The transient chirp, which is the first term in Eq. (5.1), evolves the time derivation of the logarithm of the output power, scaled by the linewidth enhancement factor α, and the

adiabatic chirp, which is the second term in Eq. (5.1), is proportional to the output power, scaled by the adiabatic chirp coefficient κ as well as the linewidth enhancement factor α.

The phase of the optical signal is an integral of its frequency. Taking the integral of Eq. (5.1), and considering the impact from the uncompensated frequency offset f_{off} and phase noise $\theta_{noise}(t)$, the phase $\varphi(t)$ of the output signal can be written as

$$\begin{aligned}\varphi(t) &= 2\pi\int \Delta\nu(t)dt + 2\pi f_{off}t + \theta_{noise}(t) \\ &= \frac{\alpha}{2}\left(\ln P(t) + \int \kappa P(t)dt\right) + 2\pi f_{off}t + \theta_{noise}(t).\end{aligned} \tag{5.2}$$

The first two terms in Eq. (5.2) are the phase modulation induced by the frequency chirp, where the first term is proportional to the logarithm of the output power, and the second term takes the time integral of the output power. Though the first term is related to the current modulation signal directly, the second term reflects the accumulated impact of the modulation signal, which makes it complicated to recover the signal using phase information.

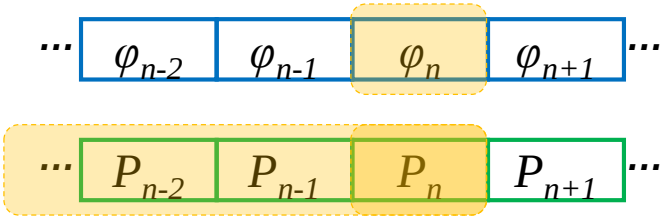

Fig. 5.2 The accumulated impact of the modulation signal on the phase information of the current symbol.

For two neighboring symbols in a sequence, i.e., the (n-1) and n-th (n=0,1,⋯) symbols, the phase of Eq. (5.2) can be rewritten as

$$\varphi(t_{n-1}) = \frac{\alpha}{2}\left(\ln P(t_{n-1}) + \int_{t=t_{n-2}}^{t_{n-1}} \kappa P(t)dt + \int_{t=t_0}^{t_{n-2}} \kappa P(t)dt\right) + 2\pi f_{off}t_{n-1} + \theta_{noise}(t_{n-1}), \tag{5.3}$$

$$\varphi(t_n) = \frac{\alpha}{2}\left(\ln P(t_n) + \int_{t=t_{n-2}}^{t_n} \kappa P(t)dt + \int_{t=t_0}^{t_{n-2}} \kappa P(t)dt\right) + 2\pi f_{off}t_n + \theta_{noise}(t_n). \tag{5.4}$$

where t_i is the time of the i-th sampling point. The first n-2 symbols lead to the same phase in Eqs. (5.3) and (5.4). To remove the accumulated impact of the modulation signal, phase rotation is used instead of the absolute value of phase. Considering the phase noise as a slow-varying parameter, the phase noise in two neighboring symbols

is assumed to be constant, $\theta_{noise}(t_{n-1}) = \theta_{noise}(t_n)$. Subtracting Eq. (5.4) with Eq. (5.3), the phase rotation between the two symbols is given by

$$\Delta\varphi(t_{n-1}, t_n) = \frac{\alpha}{2}\left(\ln P(t_n) - \ln P(t_{n-1}) + \int_{t_{n-1}}^{t_n} \kappa P(t) dt \right) + 2\pi f_{offset} \Delta t, \qquad (5.5)$$

where $\Delta t = t_n - t_{n-1}$ is the time interval of a symbol. The common phase in two neighboring symbols is removed by the subtraction operation. Except for a constant term $2\pi f_{offset}\Delta t$ due to the uncompensated frequency offset, the phase rotation of Eq. (5.5) is determined by the output power in the two neighbouring symbols. To a good approximation, we assume that the output power changes linearly between the symbols, the integral term in Eq. (5.5) can be further simplified, and we have

$$\Delta\varphi(t_{n-1}, t_n) = \frac{\alpha}{2}\left(\ln \frac{P(t_n)}{P(t_{n-1})} + \kappa \cdot \frac{P(t_{n-1}) + P(t_n)}{2} \cdot \Delta t \right) + 2\pi f_{offset} \Delta t. \qquad (5.6)$$

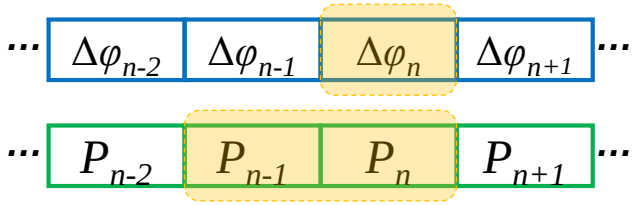

Fig. 5.3 The impact of the modulation signal on the phase rotation of the current symbol.

By using phase rotation $\Delta\varphi$, we obtain two benefits: (1) by removing the accumulated impact of the modulation signal, the phase rotation $\Delta\varphi$ is determined by the modulation signal in two neighboring symbols, which makes the signal recovery more manageable, and (2) the laser phase noise is removed, which results in simpler DSP, as opposed to being in need of sophisticated phase locked loop (PLL) algorithm.

5.4 Characterization of chirp coefficients in directly modulated laser

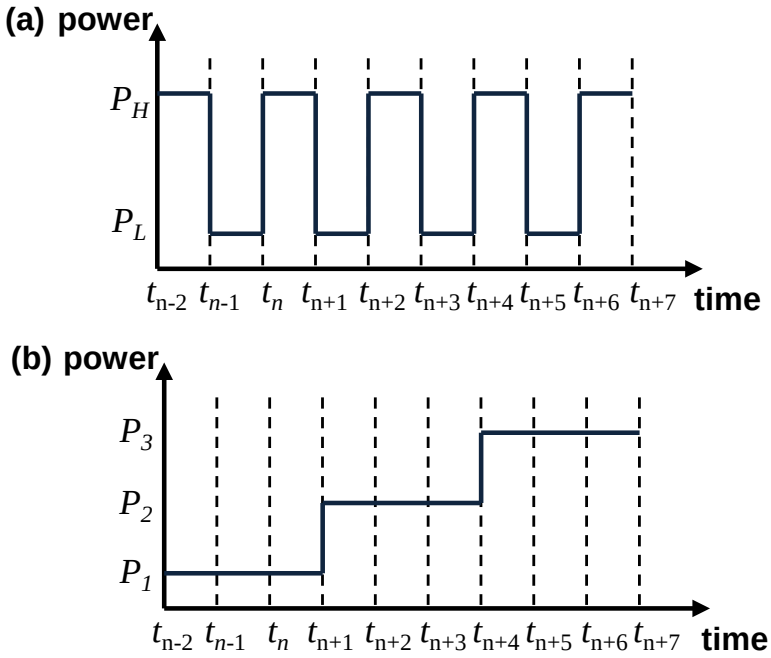

Fig. 5.4 Pilot symbols used for the characterization of the coefficients of (a) α and (b) κ.

As shown in Eq. (5.6), the phase rotation $\Delta\varphi$ is related to the modulation signal through the coefficients α and κ. Before the signal recovery, the coefficients, α and κ, need to be characterized using specially designed pilot symbols. To characterize the coefficient α, pilot symbols shown in Fig. 1(a) are used, where the modulation signal switches between high and low levels. For a transition from low to high level, the phase rotation is

$$\Delta\varphi(t_{n-1},t_n)=\frac{\alpha}{2}\left(\ln\frac{P_H}{P_L}+\kappa\cdot\frac{P_L+P_H}{2}\cdot\Delta t\right)+2\pi f_{offset}\Delta t, \tag{5.7}$$

where P_L and P_H are the output power corresponding to the low and high levels respectively. For a transition from low to high level, the phase rotation is

$$\Delta\varphi(t_n,t_{n+1})=\frac{\alpha}{2}\left(\ln\frac{P_L}{P_H}+\kappa\cdot\frac{P_L+P_H}{2}\cdot\Delta t\right)+2\pi f_{offset}\Delta t. \tag{5.8}$$

Only the first term is changed in Eqs. (5.7) and (5.8). By subtracting Eq. (5.8) with Eq. (5.7), the term related to coefficient κ is removed, and the coefficient α is given by

$$\alpha = \frac{\Delta\varphi(t_n, t_{n+1}) - \Delta\varphi(t_{n-1}, t_n)}{\ln(P_L / P_H)}. \tag{5.9}$$

To characterize the coefficient κ, pilot symbols shown in Fig. 5.4(b) are used. For three continuous symbols at time t_{n-2}, t_{n-1} and t_n, the signal is modulated at the same level with output power of P_1. Noting that the first term in Eq. (5.6) equals to zero when the output power maintains the same, the phase rotation within the three symbols is

$$\Delta\varphi(t_{n-2}, t_{n-1}) = \Delta\varphi(t_{n-1}, t_n) = \frac{\alpha\kappa\Delta t P_1}{2} + 2\pi f_{offset}\Delta t. \tag{5.10}$$

To remove the constant phase rotation due to the frequency offset in Eq. (5.10), we have three continuous pilot symbols, at time t_{n+1}, t_{n+2} and t_{n+3}, modulated at another level with output power of P_2. Then, we have

$$\Delta\varphi(t_{n+1}, t_{n+2}) = \Delta\varphi(t_{n+2}, t_{n+3}) = \frac{\alpha\kappa\Delta t P_2}{2} + 2\pi f_{offset}\Delta t. \tag{5.11}$$

Subtracting Eq. (5.10) with Eq. (5.11), the coefficient κ can be calculated as

$$\kappa = \frac{2(\Delta\varphi(t_{n-2}, t_{n-1}) - \Delta\varphi(t_{n+1}, t_{n+2}))}{\alpha\Delta t(P_1 - P_2)} = \frac{2(\Delta\varphi(t_{n-1}, t_n) - \Delta\varphi(t_{n+2}, t_{n+3}))}{\alpha\Delta t(P_1 - P_2)}. \tag{5.12}$$

5.5 Intensity and phase combined signal recovery

While the intensity of the received signal is simply proportional to the modulation signal, the phase of the received signal is related to the modulation signal through frequency chirp effect in DML. The symbol decision in conventional scheme is only based on the intensity information. In our scheme, the phase is combined with the intensity for the symbol decision. The receiver sensitivity is improved by the additional information carried IN the phase. Two approaches have been proposed for the symbol decision based on the combination of the intensity and phase.

5.5.1 Recovery scheme based on complex signal

As shown in Eq. (5.6), the phase rotation $\Delta\varphi(t_{n-1}, t_n)$ is determined by the modulation signal at time t_{n-1} and t_n. The part related to the modulation signal at time t_{n-1} can be estimated from the output power at time t_{n-1} as

$$\theta(t_{n-1}) = \frac{\alpha}{2}\left(-\ln P(t_{n-1}) + \kappa \cdot \frac{P(t_{n-1})}{2} \cdot \Delta t\right). \tag{5.13}$$

By removing $\theta(t_{n-1})$ from Eq. (5.6), the residual phase is related to the modulation signal at time t_n only, therefore the relationship between the phase and the modulation signal is further simplified. The residual phase is then combined with the intensity at time t_n to construct a complex signal $A(t_n)$ as

$$A(t_n) = \sqrt{P(t_n)} \cdot \exp\left\{ j \cdot \frac{\alpha}{2} \left(\ln P(t_n) + \kappa \cdot \frac{P(t_n)}{2} \cdot \Delta t \right) \right\}. \tag{5.14}$$

Noting that the amplitude and phase increase simultaneously with the modulation level, the constellation of the complex signal $A(t_n)$ follows the trajectory in Fig. 5.5, when the modulation level is increased gradually. By choosing N modulation levels properly, we can have N constellation points evenly spread on the plane, based on which the signal can be recovered. The described signal recovery scheme is referred as complex recovery in the following text, as it is based on the complex signal of Eq. (5.14).

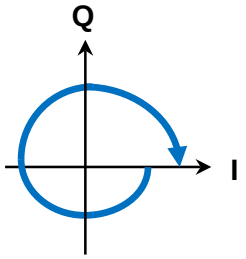

Fig. 5.5 Constellation trajectory for gradually increased modulation level.

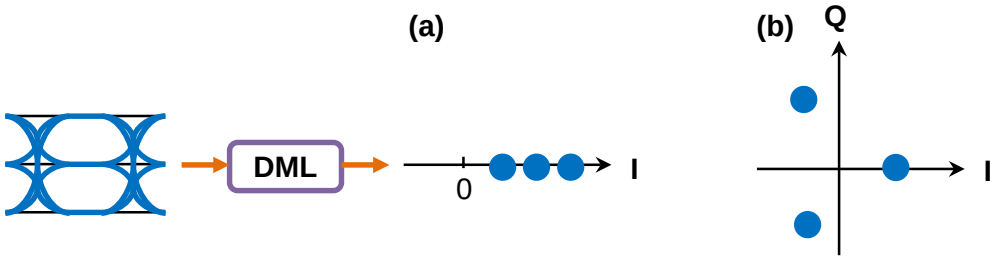

Fig. 5.6 Constellations of 3-PAM signal when (a) intensity-only recovery and (b) complex recovery are used.

When intensity-only recovery is used, the N modulation levels correspond to N constellation points on the real axis. By converting to complex recovery, the constellation points are rotated and spread on a two dimensional plane. Taking 3-level PAM (3-PAM) signal as an example, the constellations obtained by the two recovery schemes are shown in Figs. 3(a) and 3(b) respectively. By inducing a phase rotation

into the signal, the constellation points are separated further away. Therefore, the complex recovery shows a better tolerance to the noise compared to the intensity-only recovery, leading to better performance. However, the performance of complex recovery is limited by two factors: (1) the phase rotation of Eq. (5.6) is not fully utilized for signal recovery, part of the information is artificially removed for the simplicity of DSP; and (2) additional noise is induced into the phase during the estimation of Eq. (5.13).

5.5.2 Recovery scheme based on Viterbi algorithm

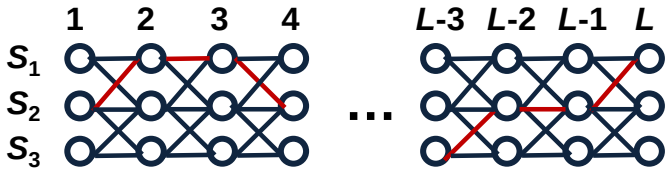

Fig. 5.7 Trellis for a signal sequence corresponding to 3-level modulation of a DML.

To avoid the drawbacks in complex recovery, Viterbi algorithm is applied to facilitate a more advanced signal recovery. For Viterbi algorithm, the possible modulations in a signal sequence are represented by a trellis shown in Fig. 5.7 [170, 171]. The L symbols in the sequence are denoted by the L nodes in the trellis, and N modulation levels are denoted by the N states ($S_1, S_2, \cdots, S_N$) at each node. Using a branch in the trellis to represent a modulation transition, an input sequence corresponds to a unique path (red line in Fig. 5.7) through the trellis. Viterbi algorithm searches the most likely path based on two metrics, the branch metric and the state metric [170].

The branch metric $m\left[S_i(K-1), S_j(K)\right]$ ($i, j = 1, 2, \cdots, N$) is a measurement of the likelihood for the transition from the i-th state at (K-1)-th node to the j-th state at K-th node. It is usually characterized by the distance between the ideal outputs and the actual observations at the receiver. For the transition $[S(K-1), S(K)]$, we have three observations, the intensity for the two symbols, $P(K-1)$ and $P(K)$, and the phase rotation between the two symbols $\Delta\varphi(K-1, K)$. The branch metric is then defined as

$$m\left[S_i(K-1), S_j(K)\right] = \left|P_i - P(K-1)\right| + \left|\sqrt{P_j}\cdot\exp(j\Delta\varphi_{ij}) - \sqrt{P(K)}\cdot\exp(j\Delta\varphi(K-1, K))\right|^2, \quad (5.15)$$

where P_i and P_j are the ideal intensity for the state S_i and S_j, and $\Delta\varphi_{ij}$ is the ideal phase rotation corresponding to the transition from S_i to S_j. A smaller branch metric

indicates a smaller distance between the ideal outputs and the observations, namely a larger possibility of the transition. The second term in Eq. (5.15) achieves its smallest value $\left|\sqrt{P_j}-\sqrt{P(K)}\right|^2$ when $\Delta\varphi_{ij}=\Delta\varphi(K-1,K)$ and its largest value $\left|\sqrt{P_j}+\sqrt{P(K)}\right|^2$ when $\left|\Delta\varphi_{ij}-\Delta\varphi(K-1,K)\right|=\pi$. We have $\left|\sqrt{P_j}-\sqrt{P(K)}\right|^2\le\left|P_j-P(K)\right|\le\left|\sqrt{P_j}+\sqrt{P(K)}\right|^2$, where $\left|P_j-P(K)\right|$ is the second term of the branch metric which uses intensity information only. Therefore, the variation range of the branch metric is enlarged by involving phase information, which consequently improves the accuracy of the decoding. A further improvement might be achieved by an optimization of the branch metric, which will be investigated in the future.

The state metric, $M\left[S_j(K)\right]$, records the minimum path weight over all the possible paths leading to the *j*-th state at *K*-th node. To minimize the path weight over *K* nodes, it is equivalent to minimize the path weight leading to each state at the first *K*-1 node $M\left[S_i(K-1)\right]$ ($i=1,2,\cdots,N$), and then minimize the sum of $M\left[S_i(K-1)\right]$ and the corresponding branch metric $m\left[S_i(K-1),S_j(K)\right]$ to the *j*-th state at the *K*-th node. Then, the state metric for the (*K*+1)-th node can be calculated from the *K*-th node. The iteration process is known as Viterbi algorithm, and can be expressed as [170]

$$M\left[S_j(K)\right]=\min_{\forall i}\left\{M\left[S_i(K-1)\right]+m\left[S_i(K-1),S_j(K)\right]\right\}. \qquad (5.16)$$

where $\forall i$ stands for operation for all the states *i*. The signal recovery scheme using Viterbi algorithm is referred as Viterbi recovery. The procedures of Viterbi recovery can be summarized as follows. The branch metrics for all the transitions in the trellis are calculated using Eq. (5.15). Starting from the given state $S_i(1)$ at the first node, the state metrics are initialized at the second node with the branch metrics $m\left[S_i(1),S_j(2)\right]$. After that, the state metrics at each node are calculated using Eq. (5.16). The paths corresponding to the state metrics are also recorded. The survivor path is the one with the minimum state metric at the last node, from which the signal is recovered.

Compared to the complex recovery, the Viterbi recovery can provide the following advantages: (1) the phase information in Eq. (5.6) is fully utilized by the Viterbi recovery; and (2) the phase noise induced by the phase estimation of Eq. (5.13) can be avoid using the Viterbi recovery. Therefore, the system performance can be further improved by the Viterbi recovery.

5.6 Experiments and results

Two intensity and phase combined recovery schemes are experimentally demonstrated in both single-polarization and dual-polarization systems. The advantages of the two schemes are verified by comparing their performances with the conventional intensity-only recovery.

The experiment setup for single-polarization system is shown in Fig. 5.8(a). A directly modulated distributed feedback (DFB) laser is used as transmitter with output wavelength of 1550 nm and 3-dB modulation bandwidth of 10 GHz. Biased by a current source, the DFB laser is driven by the amplified multi-level PAM signal, produced by an arbitrary waveform generator (AWG) working at the sampling rate of 10 GBaud. For the single-polarization transmission, a data rate of 15.88 Gb/s is achieved by using 3-PAM signal, which carries $\log_2(3)=1.588$ bits per symbol, and a data rate of 20 Gb/s is achieved by using 4-PAM signal, which carries $\log_2(4)=2$ bits per symbol. Different mapping schemes are used, when binary bits are mapped into 3-PAM and 4-PAM symbols. Two bits are converted into one symbol for 4-PAM signal, while m bits are converted into n symbols for 3-PAM signal. As stated in IEEE 802.3bp, with different combinations of m and n, different coding efficiencies can be achieved for 3-PAM signal.

For the two recovery schemes using the phase information, pilot symbols depicted in Figs. 5.4(a) and 5.4(b) are inserted before the transmitted signal for the characterization of the coefficients α and κ. The characterization only needs to be done for one time, and the signal can be transmitted without pilots after that. Therefore, the data rate will not be reduced by the pilot symbols. The drive signals are optimized separately for intensity-only and complexity recovery schemes. For the intensity-only recovery and Viterbi recovery schemes, bias current of 70 mA and electrical signal with Vpp of 4V are used to drive the DML. The drive voltage of each modulation level is adjusted to generate optical signals with equally spaced amplitude. For the complex recovery scheme, bias current is increased to 90 mA to enlarge the radius of the constellation circle. Phase rotation from the first to the last modulation level is controlled by the Vpp of the electrical signal. For 3-PAM and 4-PAM signal, Vpp of 2.9 V and 3.4 V are used respectively. The drive voltage of each modulation level is adjusted to have evenly spread constellations. The output signal from the DFB laser is transmitted over 160 km standard single mode fiber (SSMF). For the single-

polarization transmission, the launch power is set to be -3 dBm with an optical attenuator inserted before the SSMF.

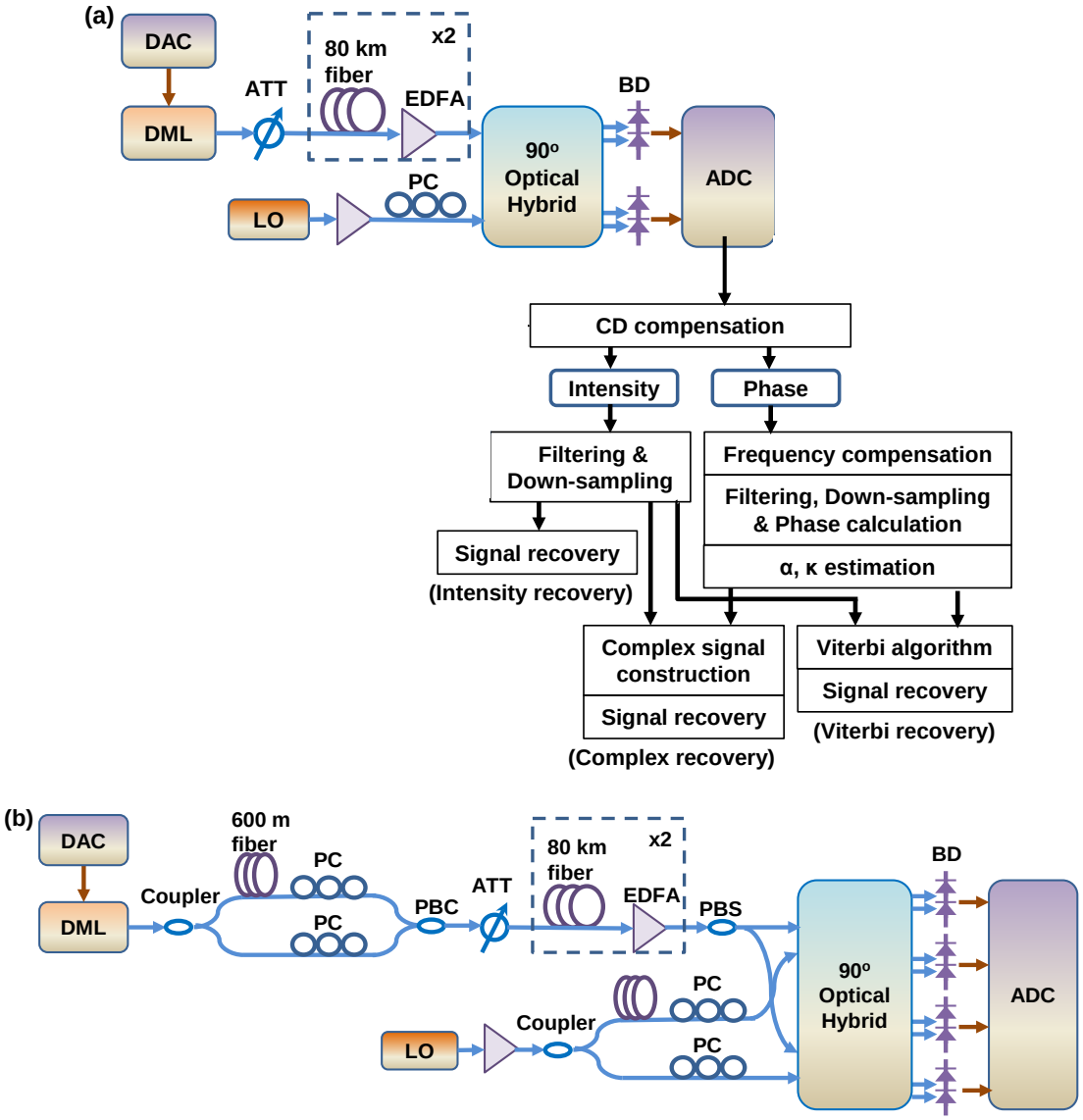

Fig. 5.8 Experiment setups for systems with (a) single polarization and (b) dual polarization. DML: directly modulated laser, ATT: attenuator, EDFA: erbium-doped-fiber amplifier, LO: local oscillator, PC: polarization controller, BD: balanced detector, PBC/PBS: polarization beam combiner/splitter, DAC: digital-to-analog convertor, ADC: analog-to-digital convertor.

The data rate is doubled when polarization-division-multiplexed (PDM) signal is transmitted. For the dual-polarization system, the PAM signal is generated using the same method as in the single-polarization system. The generated signal is split to two paths, with one path delayed by 600 m fiber, as shown in Fig. 5.8(b), so that the signal from the DFB laser with a linewidth of ~10 MHz is phase-de-correlated for the two

paths. The two de-correlated paths are multiplexed by a polarization beam combiner (PBC) into two orthogonal polarizations. The launch power is adjusted to 0 dBm for the PDM signal. With polarization-division-multiplexing, the data rate is increased to 31.76 Gb/s when 3-PAM signal is used, and 40 Gb/s when 4-PAM signal is used.

At the receiver, the signal is detected by a standard coherent receiver, where the local oscillator (LO) is provided by a tunable laser with a linewidth of 11 kHz. To reduce the electrical bandwidth requirement at the receiver, the wavelength of the LO is tuned to 1550 nm, close to the central wavelength of the optical signal. For single-polarization transmission as shown in Fig. 5.8(a), a polarization controller (PC) is used to align the polarization of the LO with the received signal before mixing them in a 90^o optical hybrid. The *I* and *Q* components of the received signal are provided by a pair balanced detectors (BDs) with 40 GHz bandwidth. The detected data is collected by a digital sampling oscilloscope working at the sampling rate of 50 GSa/s, and processed offline. After the chromatic dispersion compensation, the intensity and the phase of the received signal are separated for different processing. After a Bessel filter, the intensity information is down-sampled to 10 GSa/s. For intensity-only recovery, the symbol decision is made directly after that. For other two recovery schemes using phase information, the frequency offset needs to be compensated. After the Bessel filter and down-sampling, the phase rotation is calculated for the neighboring symbols. Coefficients α and κ are estimated from the pilot symbols using Eqs. (5.9) and (5.12). For the complex recovery, the phase of Eq. (5.13) is estimated and removed from the phase rotation. The residual phase is then combined with the intensity to construct the complex signal, based on which the symbol decision is made. For the Viterbi recovery, the branch metric is calculated using Eq. (5.15) with the combined information from phase and intensity. The state metric is then calculated to search the minimum path weight, based on which the signal is recovered. The channel memory for the Viterbi algorithm is 1, and the number of states is 3 for 3-PAM signal and 4 for 4-PAM signal. Good performance can be reached by storing 20 sequences of data.

For dual-polarization transmission, the received signal is separated into X and Y polarizations using a polarization beam splitter (PBS) as shown in Fig. 5.8(b). Two paths of LO are aligned with the X and Y polarizations respectively. The two polarizations are mixed with the LO in a 90^o optical hybrid followed by two pairs of

BDs, by which *I* and *Q* components of the two polarizations are provided. After the CD compensation, digital polarization demultiplexing is implemented using an adaptive butterfly equalizer with 41 taps [172]. From that, the DSP for each polarization follows the same procedures for single-polarization transmission.

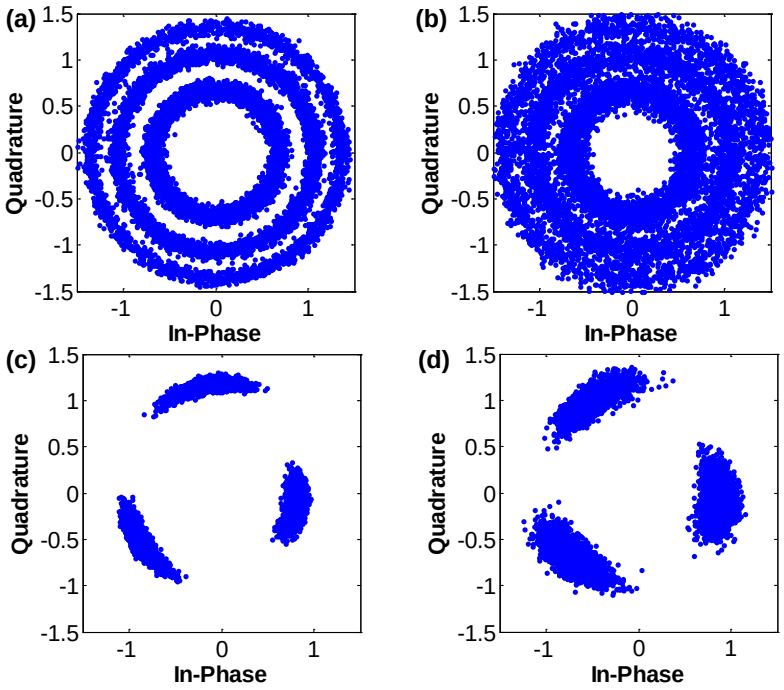

Fig. 5.9 Constellation of 3-PAM signal where (a) and (b) are for intensity-only recovery at OSNR of 25 dB and 19 dB respectively, and (c) and (d) are for complex recovery at OSNR of 25 dB and 19 dB respectively.

Since the phase is random in the intensity-only recovery, a ring is obtained for the constellation of each modulation level. Figs 5.9(a) and 5.9(b) are the constellations of 3-PAM signal when intensity-only recovery is used. In Fig. 5.9(a), three clearly separated rings can be found when OSNR equals to 25 dB. However, the boundary of each level becomes blurred in Fig. 5.9(b) when OSNR decreases to 19 dB. For complex recovery, the constellation of each modulation level is rotated by a fixed angle. Figures 5.9(c) and 5.9(d) are the constellations of 3-PAM signal when complex recovery is used. At OSNR of 25 dB, three ‘clouds’ can be found distantly separated from each other, as shown in Fig. 5.9(c). Though the ‘clouds’ become noisier when OSNR decreases to 19 dB, the three modulation levels are still clearly separated, as shown in Fig. 5.9(d). The phase rotation leads to an enlarged distance between different modulation levels, resulting in a better performance when complex recovery is used. Figures 5.10(a) and 5.10(b) are the constellations of 4-PAM signal at OSNR

of 28 dB and 22 dB, when intensity-only recovery is used. Figures 5.10(c) and 5.10(d) are the constellations of 4-PAM signal at OSNR of 28 dB and 22 dB, when complex recovery is used. Enlarged distance can also be observed for 4-PAM signal using complex recovery. For Viterbi recovery, the phase information is more accurate, by avoiding the phase estimation in complex recovery, leading to an even better performance. Total of $4x10^5$ bits have been collected for BER calculation.

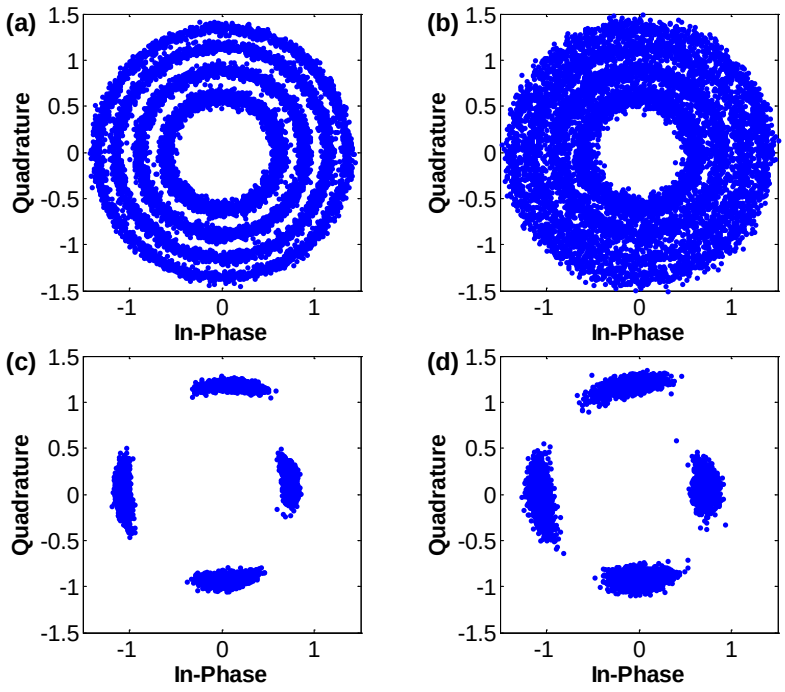

Fig. 5.10 Constellation of 4-PAM signal where (a) and (b) are for intensity-only recovery at OSNR of 28 dB and 22 dB respectively, and (c) and (d) are for complex recovery at OSNR of 28 dB and 22 dB respectively.

Fig. 5.11 shows the measured BER vs. OSNR for the single-polarization transmission using (a) 3-PAM and (b) 4-PAM signal. The system performances, using three different recovery schemes, are compared for both back-to-back and 160 km transmission. Compared with intensity-only recovery, the OSNR sensitivity is improved by around 5.5 dB by complex recovery and by around 10 dB by Viterbi recovery for both 3-PAM and 4-PAM signal. It is noted that the BER performance of Viterbi recovery is slightly improved, compared to our previous measurement of [166] for single-polarization transmission using 4-PAM signal. The improvement of around 0.5 dB is achieved by an optimization of DSP associated with the estimation of coefficients α and κ. The recovery of polarization-division-multiplexed signal using our proposed schemes is demonstrated for the first time.

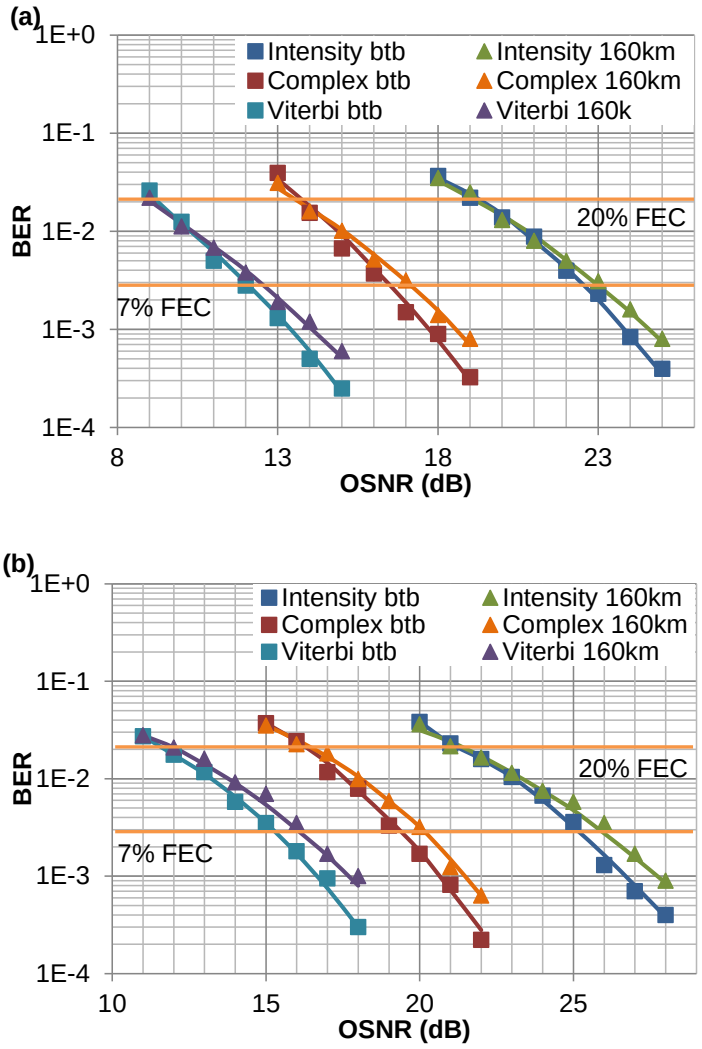

Fig. 5.11 BER vs. OSNR for single-polarization transmission with (a) 3-PAM and (b) 4-PAM signal. The baud rate is 10 Gbaud.

The measured BER vs. OSNR for dual-polarization transmission with 3-PAM and 4-PAM signal are shown in Figs. 5.12(a) and 5.12(b) respectively. The OSNR sensitivity of dual-polarization transmission is around 3.5 dB worse than single-polarization transmission. The OSNR improvements due to complex recovery and Viterbi recovery are similar to the case of single-polarization transmission. For the dual-polarization transmission, intensity-only recovery can only be used to recover 3-PAM signal, achieving the threshold of 20% FEC, and fails to recover 4-PAM signal. However, the 4-PAM signal can be successfully recovered by both complex recovery and Viterbi recovery schemes. By achieving the threshold of 7% FEC, signal with a net data rate of 37.38 Gb/s can be transmitted. The better performance of the Viterbi recovery scheme makes it more suitable for the transmission with higher data rate or

longer distance. By using a transmitter with higher electrical bandwidth, the intensity and phase combined recovery schemes can be readily extended to the application over 100 Gb/s.

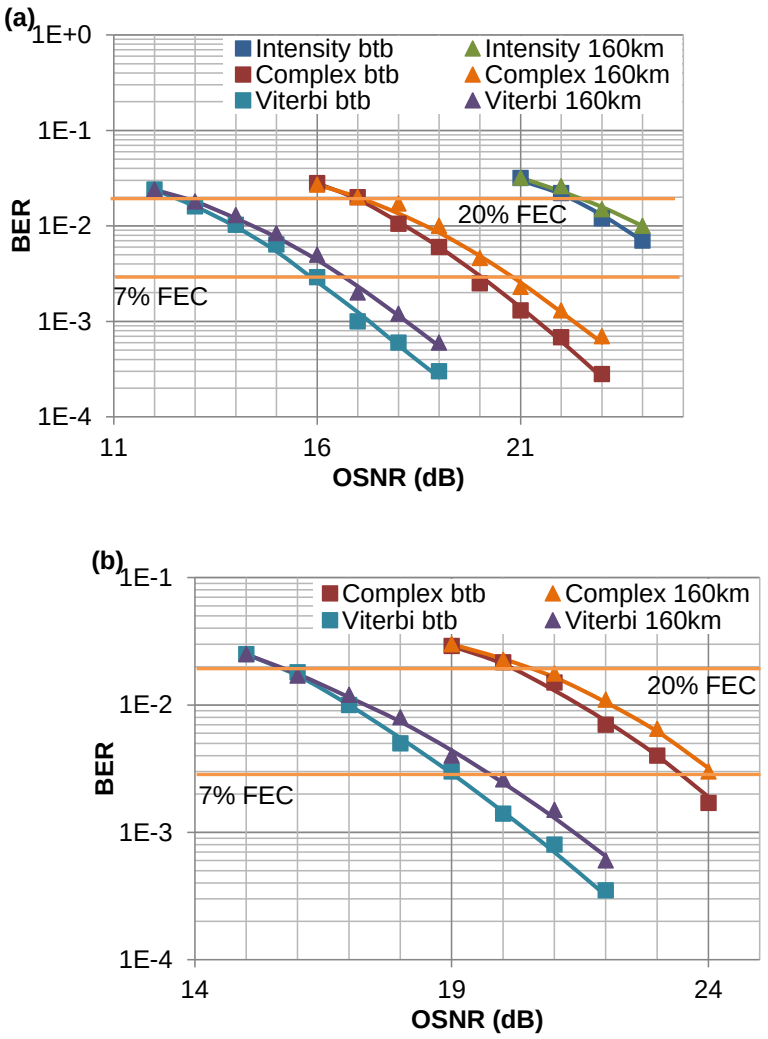

Fig. 5.12 BER vs. OSNR for dual-polarization transmission with (a) 3-PAM and (b) 4-PAM signal. The baud rate is 10 Gbaud.

5.7 Conclusion

We have demonstrated two recovery schemes which use combined information of intensity and phase in polarization-division-multiplexed system with DML and coherent detection. Besides the intensity, the output phase of DML is also modulated through the chirp effect. By exploiting the information carried by phase, the receiver sensitivity is greatly improved when the signal is recovered based on combined information of intensity and phase. For the scheme of complex recovery, the constellations of different modulation levels are distinctly separated due to the induced phase rotation, which makes the system more resistant to the noise. The system performance can be further enhanced using the scheme of Viterbi recovery,

where the phase information is utilized more efficiently, and the noise induced by the phase estimation is avoided. Compared to the conventional intensity-only recovery, the receiver sensitivity is improved by 5.5 dB using complex recovery and 10 dB using Viterbi recovery. With intensity and phase combined signal recovery, 40 Gb/s PDM-4-PAM signal can be transmitted over 160 km fiber with 7% FEC overhead. Considering its superior performance, higher data rate at longer distance can be achieved by means of Viterbi recovery if higher baud-rate DMLs are used.

6 Conclusions

6.1 Summary of this work

Both the long-haul and short-reach optical networks are faced with the great challenge of improving their capacities in order to keep up with increasing traffic demand. In this thesis, we have investigated the promising techniques to achieve high-capacity optical transmission in in a cost efficient way. Different solutions have been proposed for long-haul and short-reach systems, considering their different requirements in data rate, transmission distance and system cost. For long-haul systems, the few-mode fiber (FMF) based transmission is able to reach a channel capacity beyond that of the single-mode fiber (SMF), and reduce the cost per bit by transmitting parallel channels with multiple spatial modes. For short-reach systems, the direct modulation or direct detection is adopted to reduce the cost of the system. Various advanced modulation formats have been proposed to improve the performance of such systems.

6.1.1 Ultra high-capacity long-haul transmission based few-mode fiber

The capacity crunch in fiber optic communication has led to an active exploration of space-division multiplexed transmission system using FMF. FMF supports simultaneous propagation of multiple spatial modes and polarization modes, enhancing the channel diversity and information capacity. However, the random mode coupling and differential modal delay (DMD) in FMF lead to the signal dispersion, called mode dispersion (MD). Advanced digital signal processing (DSP) techniques in either single carrier or multicarrier modulation, can compensate MD using blind channel estimation or pilot-symbol based channel estimation. Then signal equalization can be performed based on the channel matrix to recover the signals affected by MD. The study on the property of channel matrix is critical for the FMF transmission. To fully describe the channel matrix of a FMF, we decompose it based on the generalized high-dimensional Gell-Mann matrices, an equivalent of the 2-dimensional Pauli matrices for a SMF. The frequency properties of channel matrix can be studied in terms of the autocorrelation function (ACF). In this thesis, we derive a canonical stochastic differential equation (SDE) for the channel matrix of a FMF in the regime of strong coupling. From the SDE, we develop an analytical form of the ACF of channel matrix, from which the channel correlation bandwidth is obtained, showing how fast channel changes in frequency domain. We have also adopted high-

dimensional Stokes-space analysis to measure the time evolution of the MD for FMF. One advantage of the method is that the laser phase noise is intrinsically cancelled.

6.1.2 High-performance short-reach networks with direct detection

The explosive growth of the traffic between data centers has led to an urgent demand for high-performance short-reach optical interconnects with data rate beyond 100G per wavelength and transmission distance over hundreds of kilometers. Since direct detection (DD) provides a cost-efficient solution for short-reach interconnects, various advanced modulation formats have been intensively studied to improve the performance of DD for high-performance short-reach optical interconnects. In this thesis, we report the recent progress on the advanced DD modulation formats that provide superior electrical spectral efficiency (SE) and transmission reach beyond that of simple direct modulation (DM) based direct detection (DM/DD). We first provide a review of the current advanced modulation formats for high-performance short-reach optical interconnects. Among these formats, Stokes vector direct detection (SV-DD) achieves the highest electrical spectrum efficiency, presenting itself as a promising candidate for future short-reach networks. We then expound some novel algorithms to achieve high-performance SV-DD systems under severe impairments of either polarization mode dispersion (PMD) or polarization dependent loss (PDL).

6.1.3 High-performance short-reach networks with directly modulated laser

Directly modulated laser (DML) provides an alternative choice to reduce the system cost of short-reach networks. Due to the chirp effect, the modulation information is carried by both intensity and phase of the optical signal generated by a DML. Compared to the signal recovery based on intensity alone, the receiver sensitivity is significantly improved when phase is also involved for signal recovery. In this thesis, we provide thorough analysis of two schemes to recover the directed modulated signal based on the combined information of intensity and phase. The intensity and phase combined recovery for polarization-division-multiplexed (PDM) directly modulated signal is demonstrated for the first time, achieving a data rate of 40 Gb/s over a transmission distance of 160 km. Compared with the conventional intensity-only recovery, the receiver sensitivity is improved by 5.5 dB and 10 dB respectively by using two intensity and phase combined recovery schemes.

6.2 Future work and perspectives

After a rapid development over the past decade, coherent technology has been widely deployed in long-haul optical networks to achieve high-performance high-capacity tramsmission. With the availability of high-oder modulation format and high-speed transceiver, the coherent system can reach a larger capacity than ever before. With the development of coherent communication, fiber nonlinearity has become an important issue, which limites the system performance. The research on fiber nonlinearity mitigation is crucial to improve the system performance and extend the transmission distance in the future. Besides that, the exploration of the techniques to narrow the gap between the achievable channel capacity and Shannon limit also becomes intensive, in order to approach the upper bound of the channel capacity. For these reasons, the research on advanced modulation formats and intelligent DSP algorithm will be important for the future development of the coherent communication.

By using space-division multiplexing (SDM) technique, the channel capacity can be increased dramatically. The success of SDM will depend on the highly integrated and cost-efficient components which can exploit the potentials to transmit and process parallel spatial modes simultaneously. Such components include the low modal dispersion, low loss few-mode fibers, the high-efficiency mode converter, the high extinction ratio multiplexing/de-multiplexing devices, the few-mode amplifier, the few-mode fiber compatible coupler, switch, filter, and ROADM. It is expected that with the maturity of SDM superchannel technique, ultra-high data rate beyond 1 Tb/s can be achieved with significantly improved spectral efficiency than that of the existing single-mode fiber systems.

In the meantime, the low-cost direct modulation/detection systems still are the favorable choices in the medium- to short-reach optical networks. To achieve a high capacity with such systems, it is critical to develop the novel modulation techniques which can improve the spectral efficiency by modulating the signal on available dimensions, such as intensity, phase and polarizations. The future research will focus on the powerful DSP algorithms which enable the direct modulation/detection systems to transmit and recover the signal with limited capability to modulate/detect the full dimension of the optical field. For short-reach optical transmission, the combination of direct modulation and direct detection is preferred in the aspect of cost efficiency. The chromatic dispersion in such system leads to the distortion of the

signal, which limits the signal bandwidth and the transmission distance significantly. A direct-modulation direct-detection system, which can overcome the signal distortion due to the chromatic distortion will be another interesting topic.

Bibliography

1. G. Bell, "On the production and reproduction of sound by light," Am. J. Sci., vol. 20, pp. 305–324 (1880).
2. J. L. Baird, British Patent 285,738, (1928).
3. C. W. Hansell, U. S. Patent 1,751,584, (1930).
4. H. Lamm, "Biegsame optische gerate," Z. Instrumenten., vol. 50, pp. 579 (1930).
5. A. C. S. van Heel, "A new method of transporting optical images without aberrations," Nature, vol. 173, pp. 39 (1954).
6. H. H. Hopkins and N. S. Kapany, "A flexible fiberscope using static scanning," Nature, vol. 173, pp. 39–41 (1954).
7. B. O'Brian, U. S. Patent 2,825,260 (1958).
8. B. I. Hirschowitz, U. S. Patent 3,010,357 (1961).
9. K. C. Kao and G. A. Hockham, "Dielectric-fiber surface waveguides for optical frequencies," IEE Proc., vol. 113, pp. 1151 (1966).
10. W. G. French, J. B. MacChesney, P. B. O'Connor, and G. W. Tasker, "Optical Waveguides with Very Low Losses," Bell Syst. Tech. J., vol. 53, pp. 951 (1974).
11. T. Miya, Y. Terunuma, T. Hosaka, and T. Miyashita, "Ultimate low-loss single-mode fiber at 1.55 μm," Electron. Lett., vol. 15, no. 4, pp. 106–108 (1979).
12. G. Charlet, S. Lanne, L. Pierre, C. Simonnean, P. Tran, H. Mardoyan, P. Brindel, M. Gorlier, J-C. Antona, M. Molina, P. Sillard, J. Godin, W. Idler, and S. Bigo, "Cost-optimized 6.3 Tbit/s-capacity terrestrial link over 17 × 100 km using phase-shaped binary transmission in a conventional all-EDFA SMF-based system," in Optical Fiber Communications Conference 2003, paper PD25-1.
13. S. Kawanishi, et al., "3 Tbit/s (160 Gbit/s 19 channel) optical TDM and WDM transmission experiment," Electronics Letters, vol. **35**, no. 10, pp. 826-827 (1999).

14. D. Qian, M. Huang, E. Ip, Y. Huang, Y. Shao, J. Hu, and T. Wang, "101.7-Tb/s (370×294-Gb/s) PDM-128QAM OFDM Transmission over 3×55-km SSMF using Pilot-based Phase Noise Mitigation," in Optical Fiber Communication Conference 2011, paper PDPB5.

15. A. Sano, T. Kobayashi, S. Yamanaka, A. Matsuura, H. Kawakami, Y. Miyamoto, K. Ishihara, and H. Masuda,"102.3-Tb/s (224 x 548-Gb/s) C- and Extended L-band All-Raman Transmission over 240 km Using PDM-64QAM Single Carrier FDM with Digital Pilot Tone," in Optical Fiber Communication Conference 2012, paper PDP5C.3.

16. M. Kavehrad and D. Zaccarin, “Optical code-division-multiplexed systems based on spectral encoding of noncoherent sources,” Journal of Lightwave Technology, vol. 13, no.3, pp. 534-545 (1995).

17. W. Xu, and K. Kitayama, “Analysis of beat noise in coherent and incoherent time-spreading OCDMA,” Journal of Lightwave Technology, vol. **22**, no. 10, pp. 2226-2235 (2004).

18. E. D. J. Smith, R. J. Blaikie, and D. P. Taylor, “Performance enhancement of spectral-amplitude-coding optical CDMA using pulse-position modulation,” IEEE Transactions on Communications, vol. 46**,** no. 9, pp. 1176-1185 (1998).

19. T.H. Shake, “Security performance of optical CDMA Against eavesdropping,” Journal of Lightwave Technology, vol. 23, no. 2, p. 655-670 (2005).

20. K.E. Stubkjaer, “Semiconductor optical amplifier-based all-optical gates for high-speed optical processing,” IEEE Journal of Selected Topics in Quantum Electronics, vol. 6, no. 6, pp. 1428-1435 (2000).

21. F. Parmigiani, et al., “Timing Jitter Tolerant All-Optical TDM Demultiplexing Using a Saw-Tooth Pulse Shaper,” Photonics Technology Letters, vol. 20, no. 23, pp. 1992-1994 (2008).

22. T. Morioka, et al., “1 Tbit/s (100 Gbit/s 10 channel) OTDM/WDM transmission using a single supercontinuum WDM source,” Electronics Letters, vol. 32, no. 10, pp. 906-907 (1996).

23. M. Saruwatari, “All-optical signal processing for terabit/second optical transmission,” IEEE Journal of Selected Topics in Quantum Electronics, vol. 6, no. 6, pp. 1363-1374 (2000).

24. I. Baldine, et al., “JumpStart: a just-in-time signaling architecture for WDM burst-switched networks,” Communications Magazine, vol. 40, no. 2, pp. 82-89(2002).

25. M. Koshiba, “Wavelength division multiplexing and demultiplexing with photonic crystal waveguide couplers,” Journal of Lightwave Technology, vol. 19, no. 12, pp. 1970-1975(2001).

26. M. Duser and P. Bayvel, *Analysis of a dynamically wavelength-routed optical burst switched network architecture.* Lightwave Technology, Journal of, 2002. **20**(4): p. 574-585.

27. J. Strand, A.L. Chiu, and R. Tkach, “Issues for routing in the optical layer,” Communications Magazine, vol. 39, no. 2, pp. 81-87 (2001).

28. Bosco, G., et al., “On the use of NRZ, RZ, and CSRZ modulation at 40 Gb/s with narrow DWDM channel spacing,” Journal of Lightwave Technology, vol. 20, no. 9, pp. 1694-1704 (2002).

29. J. Yu, X. Zhou, M. F. Huang, Y. Shao, D. Qian, T. Wang, M. Cvijetic, P. Magill, L. Nelson, M. Birk, S. Ten, H. B. Matthew, and S. K. Mishra, “17 Tb/s (161x 114 Gb/s) PolMux-RZ-8PSK transmission over 662 km of ultra-low loss fiber using C-band EDFA amplification and digital coherent detection,” in 34th European Conference on Optical Communication 2008, paper Th.3.E.2.

30. A. Sano, H. Masuda, T. Kobayashi, M. Fujiwara, K. Horikoshi, E. Yoshida, Y. Miyamoto, M. Matsui, M. Mizoguchi, H. Yamazaki, Y. Sakamaki, and H. Ishii, “69.1-Tb/s (432 x 171-Gb/s) C- and Extended L-Band Transmission over 240 Km Using PDM-16-QAM Modulation and Digital Coherent Detection,” in Optical Fiber Communication Conference 2010, paper PDPB7.

31. Y.-K. Huang, E. Ip, M.-F. Huang, B. Zhu, P. N. Ji, Y. Shao, D. W. Peckham, R. Lingle, Y. Aono, T. Tajima, and T. Wang, "10x456-Gb/s DP-16QAM transmission over 8 x 100 km of ULAF using coherent detection with a 30-GHz analog-to-digital converter,” in 15th OptoElectronics and Communications Conference 2010, Sapporo, Japan, paper PD3.

32. Salehi, J.A., “Code division multiple-access techniques in optical fiber networks. I. Fundamental principles,” IEEE Transactions on Communications, vol. 37 no. 8, pp. 824-833 (1989).

33. Brucnal, P.R., M. Santoro, and F. Ting, “Spread spectrum fiber-optic local area network using optical processing,” Journal of Lightwave Technology, vol. 4, no. 5, pp. 547-554 (1986).

34. V. Van, T. A. Ibrahim, and P. T. Ho, “All-optical time division demultiplexing and pulse routing using semiconductor microring resonators," in Optical Fiber Communication Conference and Exhibit 2002, paper TuU4.

35. U. Feiste, et al. “160 Gbit/s transmission over 116 km field-installed fiber using 160 Gbit/s OTDM and 40 Gbit/s ETDM,” in Optical Fiber Communication Conference and Exhibit 2001, paper ThF3.

36. K. L. Deng, et al., “A highly-scalable, rapidly-reconfigurable, multicasting-capable, 100-Gb/s photonic switched interconnect based upon OTDM technology,” Journal of Lightwave Technology, vol. 18, no. 12, pp. 1892-1904 (2000).

37. M. Nakazawa, et al., “TDM single channel 640 Gbit/s transmission experiment over 60 km using 400 fs pulse train and walk-off free, dispersion flattened nonlinear optical loop mirror,” Electronics Letters, vol. 34, no. 9, pp. 907-908 (1998).

38. H. Takara, T. Ohara, and K. Sato, “Over 1000 km DWDM transmission with supercontinuum multi-carrier source,” Electronics Letters, vol. 39, no. 14, pp. 1078-1079 (2003).

39. T. Tsuritani, et al., “Ultra-long-haul 40-Gbit/s-based DWDM transmission using optically prefiltered CS-RZ signals,” IEEE Journal of Selected Topics in Quantum Electronics, vol. 10, no. 2, pp. 403-411 (2004).

40. M. Suzuki, T. Tsuritani, and N. Edagawa. “Multi-terabit long haul DWDM transmission with high spectral efficiency in advanced semiconductor lasers and applications/ultraviolet and blue lasers and their applications/ultralong haul DWDM transmission and networking/WDM components,” Digest of the LEOS Summer Topica 2001.

41. Golovchenko, E.A., D.I. Kovsh, and S.M. Abbott, “Design tools and methods for high capacity DWDM long-haul transmission systems,” in The 4th Pacific Rim Conference on Lasers and Electro-Optics 2001.

42. C. Riziotis and M.N. Zervas, “Design considerations in optical add/drop multiplexers based on grating-assisted null couplers,” Journal of Lightwave Technology, vol. 19**,** no. 1, pp. 92-104 (2001).

43. A. Boskovic, et al. “Broadcast and select OADM nodes application and performance trade-offs,” in Optical Fiber Communication Conference and Exhibit 2002.

44. Eldada, L., et al., “Integrated multichannel OADMs using polymer Bragg grating MZIs,” Photonics Technology Letters, vol. 10, no. 10, pp. 1416-1418(1998).

45. I. Lelic, “Large multi-stage OXC,” in 27th European Conference on Optical Communication 2001.

46. L. Marazzi, et al, “Cross-talk-limited cascadability of different OXC architectures,” in Technical Digest. Summaries of papers presented at the Conference on Lasers and Electro-Optics 2001.

47. A. H. Gnauck, P. Winzer, C. Doerr, and L. Buhl, “10 x 112-Gb/s PDM 16-QAM Transmission over 630 km of Fiber with 6.2-b/s/Hz Spectral Efficiency,” in Optical Fiber Communication Conference 2009, paper PDPB8.

48. A. H. Gnauck, P. J. Winzer, S. Chandrasekhar, X. Liu, B. Zhu, and D. W. Peckham, “10 x 224-Gb/s WDM transmission of 28-Gbaud PDM 16-QAM on a 50-GHz grid over 1,200 km of fiber,” in Optical Fiber Communication Conference 2012, paper PDPB8.

49. X. Zhou, J. Yu, M. F. Huang, Y. Shao, T. Wang, L. Nelson, P. D. Magill, M. Birk, P. I. Borel, D. W. Peckham, and R. Lingle, “64-Tb/s (640x 107-Gb/s) PDM-36QAM transmission over 320 km using both pre- and post-transmission digital equalization,” in Optical Fiber Communication Conference 2010, paper PDPB9.

50. M. S. Alfiad, M. Kuschnerov, S. L. Jansen, T. Wuth, D. van den Borne, and H. de Waardt, “11 x 224-Gb/s POLMUX-RZ-16QAM transmission over 670 km of SSMF with 50-GHz channel spacing,” IEEE Photon. Technol. Lett., vol. 22, no. 15, pp. 1150–1152 (2010).

51. J.-X. Cai, Y. Cai, Y. Sun, C.R. Davidson, D.G. Foursa, A. Lucero, O. Sinkin, W. Patterson, A. Pilipetskii, G. Mohs, and N.S. Bergano, “112x112 Gb/s transmission over 9,360 km with channel spacing set to the baud rate (360% spectral efficiency),” in 36th European Conference and Exposition on Optical Communication 2010, paper PD2.1.

52. M. Nölle, J. Hilt, L. Molle, M. Seimetz, and R. Freund, “8x224 Gbit/s PDM 16QAM WDM transmission with real-time signal processing at the transmitter,” in 36th European Conference and Exposition on Optical Communication 2010, paper We.8.C.4.

53. B. Schmidt, A. J. Lowery, and J. Armstrong, “Experimental demonstrations of electronic dispersion compensation for long-haul transmission using direct-detection optical OFDM,” J. Lightwave Technol., vol. 26, no. 1, pp. 196–203 (2008).

54. W. R. Peng, X. Wu, V. Arbab, B. Shamee, L. Christen, J. Yang, K. Feng, A. Willner, and S. Chi, “Experimental demonstration of a coherently modulated and directly detected optical OFDM system using an RF-tone insertion,” in Optical Fiber Communication Conference 2008, San Diego, CA, paper OMU2.

55. W. R. Peng, B. Zhang, K. M. Feng, X. X. Wu, A. E. Willner, and S. Chi, “Spectrally efficient direct-detected OFDM transmission incorporating a tunable frequency gap and an iterative detection technique,” J. Lightwave Technol., vol. 27, no. 24, pp. 5723–5735 (2009).

56. P. P. Mitra and J. B. Stark, “Nonlinear limits to the information capacity of optical fiber communications,” Nature, vol. 411, pp. 1027–1030 (2001).

57. T. Hayashi, T. Taru, O. Shimakawa, T. Sasaki, and E. Sasaoka, “Ultra-Low-Crosstalk Multi-Core Fiber Feasible to Ultra-Long-Haul Transmission,” in Optical Fiber Communication Conference 2011, paper PDPC2.

58. J. Sakaguchi, Y. Awaji, N. Wada, A. Kanno, T. Kawanishi, T. Hayashi, T. Taru, T. Kobayashi, and M. Watanabe, “109-Tb/s (7×97×172-Gb/s SDM/WDM/PDM) QPSK transmission through 16.8-km homogeneous multi-core fiber,” in Optical Fiber Communication Conference 2011, paper PDPB6.

59. J. Sakaguchi, B.J. Puttnam, W. Klaus, Y. Awaji, N. Wada, A. Kanno, T. Kawanishi, K. Imamura, H. Inaba, K. Mukasa, R. Sugizaki, T. Kobayashi, M. Watanabe, “19-core fiber transmission of 19x100x172-Gb/s SDM-WDM-PDM-QPSK signals at 305Tb/s,” in Optical Fiber Communication Conference 2012, paper PDP5C.1.

60. R. Ryf, R. Essiambre, A. Gnauch, S. Randel, M.A. Mestre, C. Schmidl, P. Winzer, R. Delbue, P. Pupalaikis, A. Sureka, T. Hayashi, T. Taru, and T. Sasaki, “Space-division Multiplexed Transmission over 4200 km 3-Core mircrosctructured Fiber,” in Optical Fiber Communication Conference 2012, paper PDP5C.2.

61. B. Zhu, T. Taunay, M. Fishteyn, X. Liu, S. Chandrasekhar, M. Yan, J. Fini, E. Monberg, and F. Dimarcello, “Space-, Wavelength-, Polarization-Division Multiplexed Transmission of 56-Tb/s over a 76.8-km Seven-Core Fiber,” in Optical Fiber Communication Conference 2011, paper PDPB7.

62. B. Zhu, T.F. Taunay, M. Fishteyn, X. Liu, S. Chandrasekhar, M. F. Yan, J. M. Fini, E. M. Monberg, and F. V. Dimarcello, “112-Tb/s Space-division multiplexed DWDM transmission with 14-b/s/Hz aggregate spectral efficiency over a 76.8-km seven-core fiber,” Opt. Express, vol. 19, pp. 16665–16671 (2011).

63. S. Berdagué and P. Facq, “Mode division multiplexing in optical fibers,” Appl. Opt., vol. 21, pp. 1950–1955 (1982).

64. H. R. Stuart, "Dispersive Multiplexing in Multimode Optical Fiber," Science, vol. 289, pp. 281–283 (2000).

65. B. C. Thomsen, "MIMO enabled 40 Gb/s transmission using mode division multiplexing in multimode fiber," in Optical Fiber Communication Conference 2010, paper OThM6.

66. B. Franz, D. Suikat, R. Dischler, F. Buchali, and H. Buelow, “High speed OFDM data transmission over 5 km GI-multimode fiber using spatial multiplexing with 2x4 MIMO processing,” in European Conference and Exhibition On Optical Communication 2010, paper Tu3.C.4.

67. L. Gruner-Nielsen, Y. Sun, J. W. Nicholson, D. Jakobsen, R. Lingle, and B. Palsdottir, “Few Mode Transmission Fiber with low DGD, low Mode Coupling and low Loss,” in Optical Fiber Communication Conference 2012, paper PDP5A.1.

68. J. M. Tang, P. M. Lane, and K. A. Shore, “High-speed transmission of adaptive modulated optical OFDM signals over multimode fibers using directly modulated DFBs,” J. Lightwave Technol., vol. 24, no. 1, pp. 429–441 (2006).

69. C. Xie, P Dong, P. Winzer, C. Greus, M. Ortsiefer, C. Neumeyr, S. Spiga, M. Muller, and M. C. Amann, “960-km SSMF transmission of 105.7-Gb/s PDM 3-PAM using directly modulated VECSELs and coherent detection,” Opt. Express vol. 21, no. 9, pp. 11585-11589(2013).

70. C. Xie, S. Spiga, P Dong, P. Winzer, A Gnauck, C. Greus, M. Ortsiefer, C. Neumeyr, M. Muller, and M. C. Amann, “Generation and transmission of 100-Gb/s 4-PAM using directly modulated VCSELs and coherent detection,” in Optical Fiber Communication Conference 2014, San Francisco, CA, paper Th3K.7.

71. R. W. Chang, “Synthesis of band-limited orthogonal signals for multichannel data transmission,” Bell Sys. Tech. J., vol. 45, pp. 1775-1796 (1966).

72. B. R. Saltzberg, “Performance of an efficient parallel data transmission system,” IEEE Trans. Commun., vol. 15, pp. 805-813 (1967).

73. R. Nee and R. Prasad. OFDM for Wireless Multimedia Communications (1st ed.). Artech House, Inc., Norwood, MA, USA, 2000.

74. S. Hara and R. Prasad, Multicarrier Techniques for 4G Mobile Communications, Artech House, Inc., Norwood, MA, USA, 2003.

75. W. Shieh, X. Yi, Y. Ma, and Y. Tang, “Theoretical and experimental study on PMD-supported transmission using polarization diversity in coherent optical OFDM systems,” Optics Express, vol. **15**, pp. 9936–9947 (2007).

76. W. Shieh, X. Yi, and Y. Tang, “Transmission experiment of multi-gigabit coherent optical OFDM systems over 1000 km SSMF fiber,” Electron. Lett., vol. 43, pp. 183–185 (2007).

77. S. L. Jansen, I. Morita, N. Takeda, and H. Tanaka, “20-Gb/s OFDM transmission over 4,160-km SSMF enabled by RF-Pilot tone phase noise compensation,” in Optical Fiber Communication Conference 2007, paper PDP15.

78. W. Shieh and C. Athaudage, "Coherent optical orthogonal frequency division multiplexing," Electron. Lett., vol. 42, pp. 587–588 (2006).

79. Y. Tang, W. Shieh, X. Yi and R. Evans, “Optimum design for RF-to-optical up-converter in coherent optical OFDM systems,” IEEE Photon. Technol. Lett., vol. 19, pp. 483–485 (2007).

80. S. J. Savory, G. Gavioli, R. I. Killey, and P. Bayvel, “Electronic compensation of chromatic dispersion using a digital coherent receiver,” Optics Express, vol. 15, pp. 2120–2126 (2007).

81. D. S. Ly-Gagnon, S. Tsukarnoto, K. Katoh, and K. Kikuchi, “Coherent detection of optical quadrature phase-shift keying signals with carrier phase estimation,” J. Lightwave Technol., vol. 24, pp. 12–21 (2006).

82. S. B. Cohn and N. P. Weinhouse, “An automatic microwave phase measurement system,” Microwave Journal, vol. **7**, pp. 49–56 (1964).

83. C. A. Hoer and K.C. Roe, "Using an arbitrary six-port junction to measure complex voltage ratios", IEEE Trans. on MTT, vol. MTT-23, pp. 978–984 (1975).

84. W. Shieh and I. Djordjevic, Orthogonal Frequency Division Multiplexing for Optical Communications, Academic Press, 2009.

85. W. Shieh, H. Bao, and Y. Tang, “Coherent optical OFDM: theory and design,” Opt. Express, vol. 16, pp. 841–859 (2008).

86. W. Shieh, X. Yi, Y. Ma, and Q. Yang, “Coherent optical OFDM: has its time come? [Invited],” J. Opt. Netw., vol. **7**, pp. 234–255 (2008).

87. W. Shieh, “PMD-supported coherent optical OFDM systems,” IEEE Photon. Technol. Lett. vol. 19, pp. 134–136 (2007).

88. Q. Yang, Y. Tang, Y. Ma, and W. Shieh, “Experimental Demonstration and Numerical Simulation of 107-Gb/s High Spectral Efficiency Coherent Optical OFDM,” J. Lightwave Technol., vol. 27, pp. 168–176 (2009).

89. A. Li, A. A. Amin, X. Chen, and W. Shieh, “Reception of mode and polarization multiplexed 107-Gb/s CO-OFDM signal over a two-mode fiber,” in Optical Fiber Communication Conference 2011, paper PDPB8.

90. R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, A. Sierra, S. Mumtaz, M. Esmaeelpour, E. C. Burrows, R. Essiambre, P. J. Winzer, D. W. Peckham, A. H. McCurdy, and R. Lingle, “Mode-division multiplexing over 96 km of few-mode fiber using coherent 6x6 MIMO processing,” J. Lightw. Technol., vol. 30, pp. 521-531 (2012).

91. N. Bai, E. Ip, Y. Huang, E. Mateo, F. Yaman, M. Li, S. Bickham, S. Ten, J. Linares, C. Montero, V. Moreno, X. Prieto, V. Tse, K. M. Chung, A. P. T. Lau, H. Tam, C. Lu, Y. Luo, G. Peng, G. Li, and T. Wang, “Mode-division multiplexed transmission with inline few-mode fiber amplifier,” Opt. Express, vol. 20, pp. 2668-2680 (2012).

92. R. Olshansky, “Mode-coupling effects in graded-index optical fibers,” Appl. Opt., vol. 14, pp. 935-945 (1975).

93. M. Salsi, C. Koebele, D. Sperti, P. Tran, P. Brindel, H. Mardoyan, S. Bigo, A. Boutin, F. Verluise, P. Sillard, M. Bigot-Astruc, L. Provost, F. Cerou, and G. Charlet, “Transmission at 2x100Gb/s, over Two Modes of 40km-long Prototype Few-Mode Fiber, using LCOS based Mode Multiplexer and Demultiplexer,” in Optical Fiber Communication Conference 2011, paper PDPB9.

94. R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, R. Essiambre, P. Winzer, D. W. Peckham, A. McCurdy, and R. Lingle, “Space-division multiplexing over 10 km of three-mode fiber using coherent 6 × 6 MIMO processing,” in Optical Fiber Communication Conference 2011, paper PDPB10.

95. R. Ryf, S. Randel, A. Gnauck, C. Bolle, A. Sierra, S. Mumtaz, M. Esmaeelpour, E. Burrows, R. Essiambre, P. Winzer, D. Peckham, A. McCurdy, and R. Lingle, “Mode-Division Multiplexing Over 96 km of Few-Mode Fiber Using Coherent 6 × 6 MIMO Processing,” J. Lightwave Technol., vol. 30, pp. 521–531 (2012).

96. N. K. Fontaine, C. R. Doerr, M.A. Mestre, R. Ryf, P. Winzer, L. Buhl, Y. Sun, X. Jiang, and R. Lingle, “Space-division multiplexing and all-optical MIMO demultiplexing using a photonic integrated circuit,” in Optical Fiber Communication Conference 2012, paper PDP5B.1.

97. S. Randel, R. Ryf, A. Gnauck, M.A. Mestre, C. Schmidt, R. Essiambre, P. Winzer, R. Delbue, P. Pupalaikis, A. Sureka, Y. Sun, X. Jiang, and R. Lingle, “Mode-multiplexed 6x20-GBd QPSK Transmission over 1200-km DGD-Compensated Few-Mode Fiber,” in Optical Fiber Communication Conference 2012, paper PDP5C.5.

98. R. Ryf, M.A. Mestre, A. Gnauck, S. Randel, C. Schmidt, R. Essiambre, P. Winzer, R. Delbue, P. Pupalaikis, A. Sureka, Y. Sun, X. Liang, D. Peckham, A.H. McCurdy, and R. Lingle, “Low-Loss Mode Coupler for Mode-Multiplexed transmission in Few-Mode Fiber,” in Optical Fiber Communication Conference 2012, paper PDP5B.5.

99. P. Sillard, M. Astruc, D. Boivin, H. Maerten, and L. Provost, “Few-Mode Fiber for Uncoupled Mode-Division Multiplexing Transmissions,” in 37th European Conference and Exposition on Optical Communication 2011, paper Tu.5.LeCervin.7.

100. L. Gruner-Nielsen, Y. Sun, J. W. Nicholson, D. Jakobsen, R. Lingle, and B. Palsdottir, “Few Mode Transmission Fiber with low DGD, low Mode Coupling and low Loss,” in Optical Fiber Communication Conference 2012, paper PDP5A.1.

101. D. Marcuse, Light Transmission Optics, Van Nostrand Reinhold, New York, USA (1982).

102. A. W. Snyder and J. D. Love, Optical Waveguide Theory, Chapman and Hall, London, UK (1983).

103. M. J. Adams, An Introduction to Optical Waveguides, Wiley & Sons, Inc., New York, USA, Chap. 7 (1981).

104. G. P. Agrawal, Nonlinear Fiber Optics, Third Edition, Academic Press, San Diego, CA, USA (2001).

105. K. Okamoto, Fundamentals of Optical Waveguides, Second Edition, Academic Press, San Diego, CA, USA (2000).

106. ITU-T Recommendation G.652, "Transmission Media Characteristics: Characteristics of a Single-Mode Optical Fiber Cable," (2005).

107. ITU-T Recommendation G.651, "Transmission Media Characteristics: Characteristics of a 50/125 M Multimode Graded Index Optical Fiber Cable," (1993).

108. B. Y. Kim, J. N. Blake, S. Y. Huang, and H. J. Shaw, "Use of highly elliptical core fibers for two-mode fiber devices," Opt. Lett. vol. 12, pp. 729–731 (1987).

109. S. Y. Huang, J. N. Blake, and B. Y. Kim, "Perturbation effects on mode propagation in highly elliptical core two-mode fibers," J. Lightwave Technol., vol. 8, no. 1, pp. 23–33 (1990).

110. J. N. Blake, B. Y. Kim, and H. J. Shaw, "Fiber-optic modal coupler using periodic microbending," Opt. Lett., vol. 11, pp. 177–179 (1986).

111. J. N. Blake, S. Y. Huang, B. Y. Kim, and H. J. Shaw, "Strain effects on highly elliptical core two-mode fibers," Opt. Lett., vol. 12, pp. 732–734 (1987).

112. A. Li, A. A. Amin, X. Chen, and W. Shieh, "Transmission of 107-Gb/s mode and polarization multiplexed CO-OFDM signal over a two-mode fiber," Opt. Express, vol. 19, pp. 8808–8814 (2011).

113. A. Li, A. A. Amin, X. Chen, S. Chen, G. Gao, and W. Shieh, "Reception of Dual-Spatial-Mode CO-OFDM Signal Over a Two-Mode Fiber," J. Lightwave Technol., vol. 30, pp. 634–640 (2012).

114. A. A. Amin, A. Li, S. Chen, X. Chen, G. Gao, and W. Shieh, "Dual-LP11 mode 4x4 MIMO-OFDM transmission over a two-mode fiber," Opt. Express, vol. 19, pp. 16672–16679 (2011).

115. X. Chen, A. Li, J. Ye, A. Al Amin, and W. Shieh, “Reception of Dual-LP11-Mode CO-OFDM Signals through Few-mode Compatible Optical Add/Drop Multiplexer,” in Optical Fiber Communication Conference 2012, paper PDP5B.4.

116. T. Morioka, Y. Awaji, R. Ryf, P. Winzer, D. Richardson, and F. Poletti, “Enhancing optical communications with brand new fibers,” Communications Magazine, vol. 50, no. 2, pp. s31–s42 (2012).

117. J. Sakai and T. Kimura, “Bending loss of propagation modes in arbitrary-index profile optical fibers,” Appl. Opt., vol. 17, pp. 1499–1506 (1978).

118. R. C. Youngquist, J. L. Brooks, and H. J. Shaw, “Two-mode fiber modal coupler,” Opt. Lett., vol. 9, pp. 177–179 (1984).

119. S. Savin, M.J. F. Digonnet, G. S. Kino, and H. J. Shaw, “Tunable mechanically induced long-period fiber gratings,” Opt. Lett., vol. 25, pp. 710–712 (2000).

120. S. Y. Huang, J. N. Blake, and B. Y. Kim, “Perturbation effects on mode propagation in highly elliptical core two-mode fibers,” J. Lightwave Technol., vol. 8, no. 1, pp. 23–33 (1990).

121. J. N. Blake, B. Y. Kim, and H. J. Shaw, “Fiber-optic modal coupler using periodic microbending,” Opt. Lett., vol. 11, pp. 177–179 (1986).

122. J. N. Blake, S. Y. Huang, B. Y. Kim, and H. J. Shaw, “Strain effects on highly elliptical core two-mode fibers,” Opt. Lett., vol. 12, pp. 732–734 (1987).

123. P. K. A. Wai, and C. R. Menyuk, “Polarization mode dispersion, decorrelation, and diffusion in optical fiber with randomly varying birefringence,” J. Lightw. Technol. vol. 14, pp. 148-157 (1996).

124. G. J. Foschini, and C. D. Poole, “Statistical theory of polarization dispersion in single mode fibers,” J. Lightw. Technol. vol. 9, pp. 1439-1456 (1991).

125. A. Bononi, and A. Vannucci, “Statistics of the Jones matrix of fibers affected by polarization mode dispersion,” Opti. Letters, vol. 26, pp. 675-677 (2001).

126. A. Vannucci, and A. Bononi, “Statistical characterization of the Jones matrix of long fibers affected by polarization mode dispersion (PMD),” J. Lightw. Technol. vol. 20, pp. 783-793 (2002).

127. J. P. Gordon, and H. Kogelnik, “PMD fundamentals: polarization mode dispersion in optical fibers,” Proc. Natl. Acad. Sci. USA, vol. 97, pp. 4541-4550 (2000).

128. C. Antonelli, A. Mecozzi, Mark Shtaif, and P. J. Winzer, “Stokes-space analysis of modal dispersion in fiber with multiple mode transmission,” Opt. Express, vol. 20, pp. 11718-11733 (2012).

129. C. Antonelli, A. Mecozzi, M. Shtaif, “Modeling of linear and nonlinear coupling in multiple-mode fiber optic transmission with MIMO signal processing” Siganls, Systems and Computers (Asilomar 2012), pp. 645-649 (2012).

130. K.-P. Ho, and J. M. Kahn, “Statistics of group delays in multimode fiber with strong mode coupling,” J. Lightw. Technol., vol. 29, pp. 3119-3128 (2011).

131. K.-P. Ho, and J. M. Kahn, “Frequency diversity in mode-division multiplexing systems,” J. Lightw. Technol., vol. 29, pp. 3719-3726 (2011).

132. M. S. Byrd, and N. Khaneja, “Characterization of the positivity of the density matrix in terms of the coherent vector representation,” Phys. Rev. A, vol. 68, 062322 (2003).

133. S. Weigert, “Baker-Campbell-Hausdorff relation for special unitary groups SU(N),” J. Phys. A: Math Gen., vol. 30, pp. 8739-8749 (1997).

134. L. Arnold, “Stochastic differential equations theory and application,” New York: Wiley (1974).

135. T. C. Gard, “Introduction to stochastic differential equations,” New York: Marcel Dekker, INC. (1988).

136. B. Øksendal, “Stochastic differential equation,” New York: Springer-Verlag Berlin Heidelberg (1989).

137. D. C. Cox, R. Leck, “Correlation bandwidth and delay spread multipath propagation statistics for 91-MHz urban mobile radio channels,” IEEE Transact. on Communications, vol. 23, pp. 1271-1280 (1975).

138. P. M. Krummrich, K. Kotten, “Extremely fast (microsecond scale) polarization changes in high speed long haul WDM transmission systems”, in Optical Fiber Communication Conference 2004, paper FI3.

139. X. Chen, J. He, A. Li, J. Ye, and W. Shieh, “Characterization of Dynamic Evolution of Channel Matrix in Two-mode Fibers,” in Optical Fiber Communication Conference 2013, paper OM2C.3.

140. K. P. Ho, and W. Shieh, “Equalization-enhanced phase noise in mode-division multiplexed systems,” J. Lightwave Technol., vol. 31, pp. 2237–2243 (2013).

141. Q. Hu, and W. Shieh, “Autocorrelation function of channel matrix in few-mode fibers with strong mode coupling,” Opt. Express, vol. 21, no. 19, pp. 22153-22165 (2013).

142. X. Liu, F. Buchali, and R. W. Tkach, “Improving the nonlinear tolerance of polarization-division-multiplexed CO-OFDM in long-haul fiber transmission,” J. Lightwave Technol., vol. 27, no. 16, pp. 3632-3640 (2009).

143. W. Shieh, Q. Yang, and Y. Ma, “107 Gb/s coherent optical OFDM transmission over 1000-km SSMF fiber using orthogonal band multiplexing,” Opt. Express, vol. 16, no. 9, pp. 6378-6386 (2008).

144. A. J. Lowery and J. Armstrong, “Orthogonal-frequency-division multiplexing for dispersion compensation of long-haul optical systems,” Opt. Express, vol. 14, no. 6, pp. 2079-2084 (2006).

145. W. Peng, X. Wu, K. Feng, V. R. Arbab, B. Shamee, J. Yang, L. C. Christen, A. E. Willner, and S. Chi, “Spectrally efficient direct-detected OFDM transmission employing an iterative estimation and cancellation technique,” Opt. Express, vol. 17, no. 11, pp. 9099-9111 (2009).

146. J. M. Tang, P. M. Lane, and K. A. Shore, “High-speed transmission of adaptive modulated optical OFDM signals over multimode fibers using directly modulated DFBs,” J. Lightwave Technol., vol. 24, no. 1, pp. 429–441 (2006).

147. B. Schmidt, A. J. Lowery, and J. Armstrong, “Experimental demonstrations of electronic dispersion compensation for long-haul transmission using direct-detection optical OFDM,” J. Lightwave Technol., vol. 26, no. 1, pp. 196–203 (2008).

148. B. J. C. Schmidt, Z. Zan, L. B. Du, and A. J. Lowery, “120 Gbit/s over 500-km using single-band polarization-multiplexed self-coherent optical OFDM,” J. Lightwave Technol., vol. 28, no. 4, pp. 328-335 (2010).

149. M. Nazarathy and A. Agmon, “Doubling direct-detection data rate by polarization multiplexing of 16-QAM without a polarization controller,” ECOC 2013, London, UK, paper Mo4.C.4.

150. X. Chen, A. Li, D. Che, Q. Hu, Y. Wang, J. He, and W. Shieh, “Block-wise phase switching for double-sideband direct detected optical OFDM signals,” Opt. Express, vol. 12, no. 11, pp. 13436-13441 (2013).

151. X. Chen, D. Che, A. Li, J. He and W. Shieh, “Signal-carrier interleaved optical OFDM for direct detection optical communication,” Optics Express, vol. 21, no. 16, pp. 32501–32507 (2013).

152. D. Che, A. Li, X. Chen, Q. Hu, Y. Wang, and W. Shieh, “160-Gb/s Stokes vector direct detection for short reach optical communication,” in Optical Fiber Communication Conference 2014, San Francisco, CA, paper Th5C.7.

153. D. Che, X. Chen, J. He, A. Li and W. Shieh, “102.4-Gb/s single-polarization derect-detection reception using signal carrier interleaved optical OFDM,” in Optical Fiber Communication Conference 2014, San Francisco, CA, paper Tu3G.7.

154. M. Morsy-Osman, M. Chagnon, M. Poulin, S. Lessard, D.V. Plant, “1λx 224 Gb/s 10 km transmission of polarization division multiplexed PAM-4 signal using 1.3 μm SiP intensity modulator and a direct-detection MIMO-based receiver,” ECOC 2014, Cannes, France, paper PD.4.4.

155. J. Estaran, M. A. Usuga, E. Porto, M. Piels, M. I. Olmedo, and I. T. Monroy, “Quad-polarization transmission for high-capacity IM/DD links,” ECOC 2014, Cannes, France, paper PD.4.3.

156. G. J. Foschini and C. D. Poole, “Statistical theory of polarization dispersion in single mode fibers,” J. Lightwave Technol., vol. 9, no. 11, pp. 1439-1455 (1991).

157. Q. Hu, D. Che and W. Shieh, “Mitigation of PMD induced nonlinear noise in Stokes vector direct detection system,” ECOC 2014, Cannes, France, paper P.3.4.

158. C. Xie and L. F. Mollenauer, “Performance degradation induced by polarization-dependent loss in optical fiber transmission systems with and without polarization-mode dispersion,” J. Lightwave Technol., vol. 21, no. 9, pp. 1953-1957 (2003).

159. Q. Hu, D. Che, Y. Wang and W. Shieh, “Advanced modulation formats for high-performance short-reach optical interconnects,” Opt. Express, vol. 23, no. 3, pp. 3245-3259 (2015).

160. C. Xie, P Dong, P. Winzer, C. Greus, M. Ortsiefer, C. Neumeyr, S. Spiga, M. Muller, and M. C. Amann, “960-km SSMF transmission of 105.7-Gb/s PDM 3-PAM using directly modulated VECSELs and coherent detection,” Opt. Express, vol. 21, no. 9, pp. 11585-11589 (2013).

161. C. Xie, S. Spiga, P Dong, P. Winzer, A Gnauck, C. Greus, M. Ortsiefer, C. Neumeyr, M. Muller, and M. C. Amann, “Generation and transmission of 100-Gb/s 4-PAM using directly modulated VCSELs and coherent detection,” in Optical Fiber Communication Conference 2014, San Francisco, CA, paper Th3K.7.

162. R. S. Tucker, “High-speed modulation of semiconductor laser,” IEEE Trans. Electron Devices, vol. 32, no. 12, pp. 2572-2584(1985).

163. T. L. Koch and J. E. Bowers, “Factors affecting wavelength chirping in directly modulated semiconductor lasers,” in Proceedings Conference Laser and Electrooptics 1985, Baltimore, MD, 72-74.

164. G. P. Agrawal, N. A. Olsson, “Self-phase modulation and spectral broadening of optical pulses in semiconductor laser amplifiers,” J. Quantum Electron., vol. 25, no. 11, pp. 2297-2306 (1989).

165. Q. Hu, D. Che, Y. Wang, A. Li, J. Fang, and W. Shieh, “Beyond amplitude-only detection for digital coherent system using directly modulated laser,” Opt. Lett., vol. 40, no. 12, pp. 2762-2765 (2015).

166. D. Che, Q. Hu, F. Yuan, Q. Yang, and W. Shieh, “Enabling complex modulation and reception of directly modulated signals using laser frequency chirp,” Photon. Technol. Lett., vol. 27, no. 22, pp. 2407-2410 (2015).

167. D. Che, Qian Hu, Feng Yuan, and William Shieh, “Enabling complex modulation using the frequency chirp of directly modulated lasers,” in European Conference on Optical Communication 2015, Valencia, Spain, paper Mo.4.5.3.

168. A. Larsson, P. Westbergh, J. Gustavsson, A. Haglund and B. Kögel, “High-speed VCSELs for short reach communication,” Semicond. Sci. and Tech., vol. 26, no. 1, pp. 1-5 (2011).

169. W. Yan, T. Tanaka, B. Liu, M. Nishihara, L. Li, T. Takahara, Z. Tao, J. C. Rasmussen, and T. Drenski, “100 Gb/s optical IM-DD transmission with 10G-Class devices enabled by 65 GSamples/s CMOS DAC core,” in Optical Fiber Communication Conference 2013, Anaheim, CA, paper OM3H.1.

170. G.D. Forney, “The Viterbi algorithm,” in Proceedings of the IEEE, vol. 61, no. 3, pp. 268-278 (1973).

171. G. D. Forney, “Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference,” IEEE Trans. Inform. Theory, vol. 18, no. 3, pp. 363-378 (1972).

172. S. J. Savory, “Digital filter for coherent optical receiver,” Opt. Express, vol. 16, no. 2, pp. 804-817 (2008).

Appendix A

Acronyms

ACF	Autocorrelation function
ADC	Analogue-to-Digital Converter
ASE	Amplified Spontaneous Emission
AWG	Arbitrary Waveform Generator
BD	Balanced detector
BER	Bit Error Ratio
BPS-DD	Block-wise phase shift direct detection
BPSK	Binary Phase Shift Keying
BR	Balanced Receiver
BS	Beam Splitter
BW	Bandwidth
CD	Chromatic Dispersion
CDM	Code Division Multiplexing
CFO	Carrier Frequency Offset
CO-OFDM	Coherent Optical OFDM
CP	Cyclic Prefix
CPE	Common Phase Error
CSPR	Carrier-to-signal power ratio
CWDM	Coarse Wavelength Division Multiplexing
DAC	Digital to Analogue Converter
DC	Direct Current
DCD	Differential Chromatic Dispersion
DD	Direct detection
DDO-OFDM	Direct Detection Optical OFDM
DEMUX	De-multiplexer
DFB	Distributed Feedback
DFT	Discrete Fourier Transform

DGD	Differential Group Delay
DM	Direct modulation
DMD	Differential Modal Delay
DML	Directly modulated laser
DSB	Double sideband
DSP	Digital Signal Processing
DWDM	Dense Wavelength Division Multiplexing
EA	Effective Area
ECL	External-Cavity Laser
EDFA	Erbium-Doped Fibre Amplifier
ER	Extinction Ratio
FDE	Frequency Domain Equalization
FDM	Frequency Division Multiplexing
FEC	Forward Error Correction
FMF	Few-Mode Fibre
FPE	Fokker-Planck equation
FT	Fourier Transform
FT-OFDM	Fourier Transform based OFDM
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FSO	Free-Space Optical communication
GI	Guard Interval
ICI	Inter-Carrier Interference
IDFT	Inverse Discrete Fourier Transform
IF	Intermediate Frequency
IFFT	Inverse Fast Fourier Transform
IIR	Infinite Impulse Response
ISI	Inter-Symbol Interference
LAN	Local Area Network
LO	Local Oscillator

LPF	Low Pass Filter
LP	Linear Polarized
LPFG	Long Period Fibre Grating
MAN	Metropolitan Area Network
MB-DFTS-OFDM	Multi-Band DFTS-OFDM
MC	Mode Converter
MCF	Multi-Core Fibre
MCM	Multi-Carrier Modulation
MDG	Mode Dependant Gain
MDL	Mode Dependant Loss
MDM	Mode Division Multiplexing
MDM-CO-OFDM	Mode Division Multiplexed Coherent Optical OFDM
MIMO	Multiple Input Multiple Output
MMF	Multimode Fibre
MM-NLSE	Multimode Nonlinear Schrödinger Equation
MMSE	Minimum Mean Square Error
M-PSK	M-ary Phase Shift Keying
M-QAM	M-ary Quadrature Amplitude Modulation
MS	Mode Stripper
MSC	Mode Selective Coupler
MDEMUX	Mode Demultiplexer
MMUX	Mode Multiplexer
MUX	Multiplexer
MZM	Mach-Zehnder Modulator
NF	Noise Figure
NLSE	Nonlinear Schrödinger Equation
OADM	Optical Add/Drop Multiplexer
OCDMA	Optical code division multiple access
OFDM	Orthogonal Frequency Division Multiplexing
O-OFDM	Optical OFDM

OSA	Optical Spectrum Analyser
OSNR	Optical Signal to Noise Ratio
OSP	OFDM Symbol Phase
OTDM	Optical Time Division Multiplexing
OTR	Optical To RF
OXC	Optical cross connect
PAM	Pulse-amplitude-modulation
PAPR	Peak to Average Power Ratio
PBC	Polarization Beam Combiner
PBS	Polarization Beam Splitter
PC	Polarization controller
PD	Photodiode
PDL	Polarization Dependant Loss
PDM	Polarization Division Multiplexing
PMD	Polarization Mode Dispersion
PMF	Polarization Maintaining Fibre
PPLN	Periodically Poled Lithium Niobate
PSCF	Pure Silica Core Fibre
PSK	Phase Shift Keying
PSP	Principle State of Polarization
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RF	Radio Frequency
RIN	Relative Intensity Noise
ROADM	Reconfigurable Optical Add/Drop Multiplexer
RTO	RF To Optical
SC	Single Carrier
SC-FDE	Single Carrier Frequency Domain Equalization
SC-FDM	Single Carrier Frequency Division Multiplexing
SCI-DD	Signal carrier interleaved direct detection

SCM	Single-Carrier Modulation
SCOH	Self-coherent
SDE	Stochastic differential equation
SDM	Space Division Multiplexing
SE	Spectral Efficiency
SISO	Single Input Single Output
SLM	Spatial Light Modulator
SMC	Spatial Mode Combiner
SMF	Single Mode Fibre
SMS	Spatial Mode Splitter
SNR	Signal to Noise Ratio
SSB	Single Side Band
SSB-OFDM	Single Side Band OFDM
SSBN	Signal-to-signal beating noise
SSMF	Standard Single Mode Fibre
SV-DD	Stokes Vector Direct Detection
SW	Switch
TDM	Time Division Multiplexing
TE	Transverse Electric
TM	Transverse Magnetic
TITO	Two Input Two Output
TMF	Two Mode Fibre
TS	Training Sequence
ULAF	Ultra Large Area Fibre
WDM	Wavelength Division Multiplexing