thesis for the degree of Doctor of Philosophy

Advanced Digital Signal Processing for High Spectral Efficiency Multidimensional Transmission

Zonglong He

CHALMERS

Photonics Laboratory
Department of Microtechnology and Nanoscience (MC2)
Chalmers University of Technology
Göteborg, Sweden, 2024

Advanced Digital Signal Processing for High Spectral Efficiency Multidimensional Transmission
Zonglong He

ISBN 978-91-8103-054-9
Doktorsavhandlingar vid Chalmers tekniska högskola,
Ny serie nr 5512
ISSN 0346-718X

Photonics Laboratory
Department of Microtechnology and Nanoscience (MC2)
Chalmers University of Technology
SE–412 96 Göteborg
Sweden
Telephone: +46 (0)73–55 48 205

Front cover illustration: Multidimensional modulation and multichannel signal processing in the comb-based superchannel systems

Printed in Sweden by
Reproservice
Chalmers Tekniska Högskola
Göteborg, Sweden, May 2024

Advanced Digital Signal Processing for High Spectral Efficiency Multidimensional Transmission

Zonglong He
Photonics Laboratory
Department of Microtechnology and Nanoscience (MC2)
Chalmers University of Technology

Abstract

Modern coherent optical communication systems utilize all available physical dimensions for data modulation. In general, independent modulation and signal processing are performed in coherent systems, leaving a performance gap to the Shannon limit. This thesis focuses on advanced digital signal processing to achieve joint encoding and joint processing systems with high spectral efficiency (SE).

First, we investigate the low-complexity and practical digital predistortion approaches to eliminate the nonlinear transmitter distortion, which is one dominant impairment for the high-SE transponder. Joint processing of multiple wavelength channels in the receiver can eliminate the linear and nonlinear interchannel interference in long-haul superchannel systems. Using the frequency-locked received channels provided by the comb-based superchannel and receiver, two joint processing schemes are studied. We investigate the performance of multichannel equalization, which cancels linear interchannel crosstalk and reduces the guard bands, in the long-haul transmission constrained by the amplified spontaneous emission noise. In addition, we propose a perturbative-based compensation to eliminate both self-phase modulation and cross-phase modulation caused by the interference wavelength channels.

Multidimensional (MD) Vononoi constellations generated by a structured geometric shaping method provide shaping gain over conventional quadrature amplitude modulation (QAM) and have low-complexity encoding and decoding algorithms, which are suitable for high-SE applications. The use of MD formats closes the gap to the theoretical Shannon limit. By employing a 24-dimensional VC with a record constellation size of 7.9×10^{28} and lookup table based predistortion, we present a 12.2 bit/s/Hz C-band transmission over 40 km single-mode fiber and have the first experimental demonstration of a significant MD shaping gain over QAM formats in the soft-decision coded system.

Keywords: Coherent communication, frequency comb, superchannel, signal processing, multidimensional format, wideband transmission

Publications

This thesis is based on the work contained in the following papers:

[A] **Z. He**, J. Song, K. Vijayan, C. Häger, A. Graell i Amat, H. Wymeersch, P. A. Andrekson, M. Karlsson, and J. Schröder, "Periodicity enabled size reduction of symbol based predistortion for high-order QAM", *Journal of Lightwave Technology*, vol. 40, no. 18, pp. 6168-6178, 2022.

[B] J. Song*, **Z. He***, C. Häger, M. Karlsson, A. Graell i Amat, H. Wymeersch, and J. Schröder, "Over-the-fiber digital predistortion using reinforcement learning", *Proceedings of European Conference on Optical Communications (ECOC)*, 2021, paper Tu1C2.4.

[C] **Z. He**, K. Vijayan, A. Mirani, M. Karlsson, and J. Schröder, "Inter-channel interference Cancellation for long-haul superchannel system", *Journal of Lightwave Technology*, vol. 42, no. 1, pp. 48-56, 2024.

[D] **Z. He**, A. Mirani, M. Karlsson, and J. Schröder, "Perturbation based joint SPM and XPM compensation for superchannel system", submitted to *Photonics Technology Letters*, Apr 2024.

[E] **Z. He**, S. Li, E. Deriushkina, P. A. Andrekson, E. Agrell, M. Karlsson, and J. Schröder, "12.2 bits/s/Hz C-band transmission with high-gain low-complexity 24-dimensional geometric shaping", *Journal of Lightwave Technology*, accepted, Mar 2024.

Related publications and conference contributions by the author, not included in the thesis:

[F] **Z. He**, K. Vijayan, M. Mazur, M. Karlsson, and J. Schröder, "Look-up table based pre-distortion for transmitters employing high-spectral-efficiency modulation formats", *Proceedings of European Conference on Optical Communications (ECOC)*, 2020, paper Tu1D.6.

[G] K. Vijayan, **Z. He**, B. Foo, M. Karlsson, and P. A. Andrekson, "Modulation format dependence on transmission reach in phase-sensitively amplified fiber links", *Optics Express*, vol. 28, no. 23, pp. 34623-34638, 2020.

[H] K. Vijayan, **Z. He**, B. Foo, M. Karlsson, and P. A. Andrekson, "Nonlinearity mitigation dependence on modulation format in phase-sensitively amplified fiber links", *Frontiers in Optics / Laser Science*, 2020, paper FTh1E.1.

[I] **Z. He***, J. Song*, C. Häger, K. Vijayan, P. A. Andrekson, M. Karlsson, A. G. i. Amat, H. Wymeersch, and J. Schröder, "Symbol-based supervised learning predistortion for compensating transmitter nonlinearity", *Proceedings of European Conference on Optical Communications (ECOC)*, 2021, paper Tu2C2.3.

[J] K. Vijayan, **Z. He**, B. Foo, J. Schröder, M. Karlsson, and P. A. Andrekson, "Phase-sensitively amplified wavelength-division multiplexed optical transmission systems", *Optics Express*, vol. 29, no. 21, pp. 33086-33096, 2021.

[K] A. Mirani, K. Vijayan, **Z. He**, S. Li, E. Agrell, J. Schröder, P. A. Andrekson, and M. Karlsson, "Experimental demonstration of 8-dimensional Voronoi constellations with 65,536 and 16,777,216 symbols", *Proceedings of European Conference on Optical Communications (ECOC)*, 2021, paper Th2C1.3.

[L] **Z. He**, J. Song, C. Häger, A. G. i. Amat, H. Wymeersch, P. A. Andrekson, M. Karlsson, and J. Schröder, "Experimental demonstration of learned pulse shaping filter for superchannels", *Proceedings of Optical Fiber Communications Conference (OFC)*, 2022, paper W2A.33.

[M] A. Mirani, K. Vijayan, S. Li, **Z. He**, J. Schröder, P. Andrekson, E. Agrell, and M. Karlsson, "Comparison of physical realizations of multidimensional Voronoi constellations in single mode fibers", *Proceedings of European Conference on Optical Communications (ECOC)*, 2022, paper We5.39.

[N] P. Zhao*, **Z. He***, V. Shekhawat, M. Karlsson, and P. A. Andrekson, "100-Gbps per-channel all-optical wavelength conversion without pre-amplifiers based on an integrated nanophotonic platform", *Nanophotonics*, vol. 12, no. 17, pp. 3427-3434, 2023.

[O] A. Mirani, K. Vijayan, S. Li, **Z. He**, E. Agrell, J. Schröder, P. A. Andrekson, and M. Karlsson, "Physical realizations of multidimensional Voronoi constellations in optical communication systems", *Journal of Lightwave Technology*, vol. 41, no. 17, pp. 5557-5563, 2023.

[P] Z. Chen, **Z. He**, A. Mirani , J. Schröder, P. A. Andrekson , M. Karlsson, M. Xiang, Y. Yu, M. Tang, Y. Qin, and S. Fu, "Transmitter optimization for PS-QAM signal in high spectral efficiency metro-transmission", *Journal of Lightwave Technology*, vol. 41, no. 9, pp. 2736-2746, 2023.

[Q] **Z. He**, P. A. Andrekson, M. Karlsson, and J. Schröder, "Frequency-comb enabled interchannel crosstalk mitigation with reduced complexity", *Proceedings of European Conference on Optical Communications (ECOC)*, 2023, paper We.A.7.5.

Contents

Acknowledgement

First, I would like to thank my supervisor, Dr. Jochen Schröder, for giving me the opportunity to work in the Photonics Laboratory as a PhD student, for open discussion of interesting ideas, and for providing me with endless support in research as well as programming. My co-supervisor, Prof. Magnus Karlsson, deserves limitless thanks for the fruitful discussions of fiber-optic theory/experiments and for all the support during my PhD. I would like to thank my examiners, Prof. Peter Andrekson and Prof. Victor Torres Company, for discussing the individual study plan, laboratory organization, and possible future experiments.

I am grateful to my office mate, Ekaterina Deriushkina, for the enjoyable time, interesting talk, and your continuous support throughout my PhD. Many thanks to Dr. Mikael Mazur for sharing your knowledge on coherent communication, superchannel experiments, and digital signal processing. Thank you Dr. Jinxiang Song for the interesting collaboration and fruitful discussion on neural networks. Special thanks to Dr. Ping Zhao, Dr. Kovendhan Vijayan, and Dr. Ali Mirani for teaching me the basics of fiber communication, the collaboration, and for all the help in the lab. Thank you Dr. Shen Li for the collaboration on multidimensional transmission experiments, and for sharing your deep knowledge of Voronoi constellation and multilevel coding. Many thanks to Prof. Henk Wymeersch, Prof. Erik Agrell, Prof. Alexandre Graell i Amat, and Prof. Christian Häger for the fruitful discussion during the project meeting and the nice feedback on the collaboration. Thank you Yan Gao for all the discussion of the thin-film lithium niobate modulator. Special thanks to Rasmus Larsson for the funny time in the lab and for discussion on the fiber communication. I would like to thank Prof. Fuchuan Lei, Dr. Zhichao Ye, Dr. Óskar B Helgason, Dr. Marcello Girardi, Dr. Israel Rebolledo Salgado, and Dr. Krishna Sundar Twayan for sharing lots of knowledge on ultrafast photonics and frequency comb. I wish to thank Prof. Shumin Wang for the interesting talk and constructive discussion

during lunchtime. Many thanks to other Photonics group members for your friendship and all the help. Special thanks to Gunnel Berggren for giving great support on all the administrative work.

I wish to thank my friend, Tian Xu, for organizing the amazing trips to France and Italy, and for the countless phone calls during the Covid-19 pandemic period. Thank you Xinzhi Wu for being my roommate in the first year of my PhD and giving me endless support in life. Finally, I would like to express my gratitude to my parents and my girlfriend Feiwei Lu. This work would never have been possible without your love and support.

Zonglong He
Göteborg, Sweden
April 2024

Acronyms

ADC	analog-to-digital converter
AIR	achievable information rate
ASE	amplified spontaneous emission
ASIC	application-specific integrated circuit
AWGN	additive white Gaussian noise
BICM	bit interleaved coded modulation
BER	bit error rate
CD	chromatic dispersion
CDC	chromatic dispersion compensation
CMOS	complementary metal–oxide–semiconductor
CNN	convolutional neural network
CPE	carrier phase estimation
DA	data-aided
DAC	digital-to-analog converter
DBP	Digital backpropagation
DCF	dispersion compensation fiber
DLA	direct learning architecture
DRE	digital resolution enhancer
DSP	digital signal processing
DPD	digital predistortion
DPE	digital pre-emphasis
EDFA	erbium-doped fiber amplifier
ENOB	effective number of bits
EO	electro-optic
FEC	forward error correction
FFNN	feed-forward neural network
FIR	finite impulse response
FOE	frequency offset estimation
GBs	guard bands

GMI	generalized mutual information
GN	Gaussian noise
GS	geometric shaping
HD	hard-decision
ICI	interchannel interference
ILA	indirect learning architecture
IM/DD	intensity-modulation/direct-detection
IQ	in-phase and quadrature
ISI	intersymbol interference
LDPC	low-density parity-check
LMS	least-mean-square
LO	local oscillator
LUT	look-up table
MCE	Multichannel equalization
MD	multidimensional
MI	mutual information
MIMO	multiple-input multiple-output
MLC	multi-level coding
MMSE	minimum-mean-square-error
MSE	mean square error
MZM	Mach-Zehnder modulator
NLC	nonlinear compensation
NLSE	nonlinear Schödinger equation
NN	neural network
OFC	optical frequency comb
OH	overhead
OSNR	optical signal-to-noise ratio
PAPR	peak-to-average power ratio
PB	perturbation-based
PE	phase estimation
PMD	polarization mode dispersion
PS	probabilistic shaping
QAM	quadrature amplitude modulation
QPSK	quadrature phase shift keying
RF	radio frequency
RL	reinforcement learning
RRC	root-raised-cosine
SD	soft-decision
SDM	space division multiplexing

SE	spectral efficiency
SER	symbol error rate
SNDR	signal-to-noise and distortion ratio
SNR	signal-to-noise ratio
SPM	self-phase modulation
SSM	split-step method
TIA	trans-impedance amplifier
VS	Volterra series
WDM	wavelength division multiplexing
WH	Wiener-Hammerstein
XPM	cross-phase modulation
ZF	zero-forcing

Chapter 1

Introduction

We live in an information era where the internet provides indispensable services in daily life including messaging, email, remote meetings, online searching, low-latency gaming, and high-resolution videos. The reliable and endless access to the internet regardless of geographic location is supported by the worldwide communication infrastructure, in other words, the optical network, which connects the world with light. Today, the ever-increasing data traffic is supported by carrying information on the various dimensions of the optical field. Here, we review the history of optical communication and then discuss the outline of the thesis.

1.1 Historical review

The beginning of the optical communication era dates back to the fundamental theory of the stimulated emission effect first found by Albert Einstein in 1917 [1] and experimentally observed by Rudolf Ladenburg in 1928 [2]. This theoretical foundation inspired the invention of the maser by Charles H. Townes with his students James P. Gordon and Herbert J. Zeiger [3] based on the amplification of microwave radiation, leading to the study of the so-called optical maser, i.e., laser [4], which can provide coherent single-frequency light for carrying information. Later in 1960, relying on the stimulated emission of optical radiation, Theodore Maiman invented the first pulsed laser [5]. In parallel, the first continuous-wave laser, a gas laser, was produced by Ali Javan, William R. Bennett Jr., and Donald R. Herriott [6], while the first demonstration of a semiconductor laser was presented by Robert N. Hall in 1962 [7]. Together with the optical source and the phototransis-

tor invented by John N. Shive [8] for light detection, the advent of silica glass based fiber [9] with low attenuation, served as a desirable transmission medium, enabled practical optical fiber communication. In earlier systems, intensity-modulation/direct-detection was commonly used to transmit the data on the intensity of the light due to its low-complexity transmitter and receiver structure. To enable the transmission of light in a lossy fiber over long distances, a repeater is needed to amplify the optical signal. Initially, repeaters were based on electronic regeneration and relied on optical detection, electrical amplification, and optical signal generation. The invention of the erbium-doped fiber amplifier (EDFA) in the late 1980s [10] allowed the joint amplification of broadband optical signals, motivating the commercialization of the wavelength division multiplexing systems that utilize the wavelength dimension to scale up system capacity. Today, the C-band window from 1530 to 1565 nm is widely used for optical communication, due to the low attenuation of $<$0.2 dB/km in the single-mode fiber [11, 12] and the EDFA gain bandwidth.

Starting from the 2000s, the development of high-speed electronic devices including digital-to-analog converters (DAC), analog-to-digital converters (ADC), and application-specific integrated circuits facilitated the joint detection of amplitude, phase, and the two polarizations of the optical signal in low-complexity polarization-diversity coherent receivers and enabled powerful digital signal processing to track the fast polarization rotation and compensate for other channel impairments. The high-resolution DACs and ADCs enabled the generation and detection of advanced multilevel modulation formats [13], respectively, to boost spectral efficiency (SE). The use of increasing symbol rate and modulation order of the quadrature amplitude modulation formats achieved the line rate of 400 Gb/s [14], 800 Gb/s [15, 16], and 1.6 Tb/s [17] per wavelength. The further increase in the line rate is inhibited by the effective number of bits of the DACs/ADCs, which decreases for higher frequency due to the electrical noise and distortion [18], and bandwidth limitation of state-of-the-art transponders. Parallel to the development of high-symbol-rate transceivers to achieve high per-channel throughput, multiband transmission has attracted significant attention for increasing the maximum total throughput of $<$100 Tb/s in the conventional C band to $>$300 Tb/s [19]. This was achieved by extending the data transmission to O-, E-, S-, L-, and U-bands with the assistance of various doped-fiber amplifiers as well as Raman amplifier. In contrast to C-band EDFA with

a low noise figure (NF) of around 5.5 dB, the rare–earth doped amplifiers used in other bands have relatively high NF and limited gain. Therefore, additional Raman amplifier with high-pump power are needed to enable high amplification gain in multiband transmission, which significantly increases the power consumption and system complexity. The unavailable multiband coherent transceivers further limit the use of other wavelength bands. Recently, much effort has been devoted to space division multiplexing with multi-mode [20–22] or multi-core [23–25] fibers, facilitating Pb/s applications [21,22,24,25] through transmitting the data over multiple spatial channels simultaneously.

In general, independent data is modulated in each physical dimension, leaving a performance gap to the Shannon limit due to the unoptimized constellation points in the multidimensional space. After signal detection, each received wavelength channel is processed individually, which limits the cancellation of linear and nonlinear interference originating from other wavelength channels. Given the limited spectral and fiber resources, the full capacity of the optical communication system is not yet exploited, and in fact, not even theoretically known due to the nonlinear characteristic of the channel.

1.2 Thesis outline

This thesis focuses on high-spectral-efficiency multidimensional modulation (MD) formats as well as joint digital signal processing to mitigate system impairments and increase spectral efficiency. In contrast to conventional two-dimensional QAM, constellation shaping can close the gap to the Shannon limit by changing the probability and geometry distribution of the constellation, commonly referred to as probabilistic shaping and geometric shaping (GS). MD constellations as a GS format have more degrees of freedom to achieve a smaller minimum Euclidean distance, providing shaping gain over QAM by modulating the data on multiple dimensions of the optical field. A class of structured GS formats, i.e., Voronoi constellations (VC), are investigated in the thesis for the realization of multidimensional transmission. This thesis also discusses various linear and nonlinear pre-distortion methods to mitigate transceiver imperfections that is dominant for high-cardinality constellations, as well as joint processing in the receiver to exploit the full potential of the digital compensation.

The outline of this thesis is as follows. Chapter 2 describes the ba-

sic building blocks of coherent optical communication systems and the corresponding system impairments, including distortion caused by the transceiver imperfection and the fiber channel. The popular receiver-side digital signal processing is briefly discussed, focusing on pilot-based algorithms. In Chapter 3, linear and nonlinear pre-distortion are introduced to compensate for the bandwidth limitation effect and pattern-dependent nonlinear distortion, respectively, in particular for high-symbol-rate high-SE applications. Chapter 4 discusses various joint receiver-side processing schemes such as joint carrier recovery, multi-channel equalization, and multi-channel digital backpropagation, which take advantage of the phase and frequency coherence property of the comb-based superchannel to investigate the full capacity of a joint-detection receiver. Finally, in Chapter 5, the MD modulation formats are introduced, primarily focusing on the structured VC with low-complexity encoding and decoding algorithms. Since the MD signal can be modulated over amplitude, phase, time, polarization, and wavelength components of the optical field, the related physical impairments as well as possible compensation schemes are briefly discussed. Finally, Chapter 6 presents future outlooks of the work.

The summary of the included papers is listed as follows. In Paper A, a periodicity-based size reduction approach for the look-up table (LUT) pre-disotion is experimentally demonstrated. Being an extension of Paper F and Paper I, this method is compared with the neural network pre-distortion trained with the LUT output, in terms of size reduction efficiency and performance. Paper B experimentally presents a reinforcement learning-based pre-distortion scheme to avoid the cumbersome channel modeling. In contrast with the conventional deep learning approach that backpropagates the gradient through the surrogate channel model for training, the proposed method estimates the gradient in the receiver by adding known Gaussian-distributed noise with a small variance to the transmitted signal. In Paper C, the multi-channel equalization is investigated in the comb-based superchannel to cancel the linear interchannel crosstalk under optical-back-to-back and long-haul transmission, primarily studying the performance penalty caused by the amplified spontaneous emission noise. Utilizing the same experimental setup, in Paper D, we verify the perturbation-based nonlinear compensation algorithm with signal detection in three separate receivers. This method is compared with the digital backpropagation regarding the achievable information rate. Paper E investigates the performance of a

high-SE (i.e., 12.2 bit/s/Hz) 24-dimensional VC in the C-band transmission. The design of a soft-decision multilevel coding scheme enables a significant shaping gain over the QAM format at SE $>$ 12 bit/s/Hz, which is demonstrated here for the first time.

Chapter 2

Coherent Transceivers and Transmission System

In communications theory we typically characterize a system by its channel response. The channel of a modern optical communication system has a complex response composed of different subsystems. In general, it is difficult to separate the individual responses as many contribute both linearly and nonlinearly. The main components are the transmitter, fiber channel, and receiver. In the following, I will discuss each component and how they impair the system's performance. In addition, the limiting factors of the systems are analyzed in various conditions applying different signal formats.

2.1 Transmitter

Coherent fiber-optic communication systems transmit data over the amplitude, phase, and polarization of the optical carrier using a coherent transmitter as illustrated in Fig. 2.1. The main steps in the process are digital preprocessing, digital-to-analog conversion, and modulation onto the optical carrier. At the input of the transmitter, a digital stream of bits from the data traffic first passes through a forward error correction (FEC) encoder, which adds redundant bits for protecting the transmitted information in the noisy channel. A detailed description of the coded system including FEC encoding and decoding is given in Section 2.4. To yield efficient data transmission, the digital bits are then mapped to constellation points before pulse shaping. The digital samples are converted to analog waveforms through the digital-to-analog converter (DAC) and

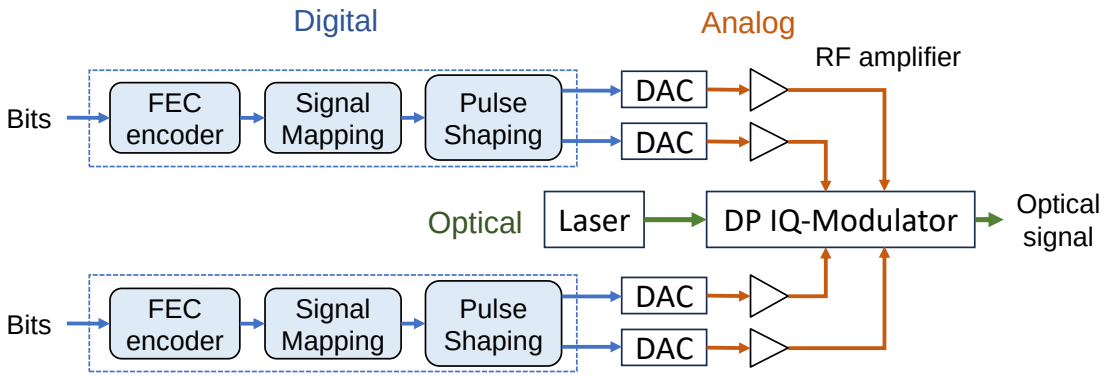

Figure 2.1: The diagram of the dual-polarization (DP) transmitter. The information bits are digitally pre-processed, converted to analog signals by digital-to-analog converters (DACs), amplified by RF drivers, and finally fed to a DP IQ modulator for optical modulation.

enlarged by a radio-frequency (RF) amplifier to drive the IQ modulator for the electrical-to-optical conversion.

2.1.1 Modulation format

Digital signal mapping converts the information bits to symbols selected from a certain constellation for transmitting the data over the channel in an efficient manner. The constellation (or modulation format) is important as it determines the energy efficiency of the data transmission. A wide range of modulation formats can be used depending on the system applications. One important candidate for coherent systems is quadrature amplitude modulation (QAM) which comprises two independent and orthogonal dimensions, namely in-phase and quadrature components of the carrier wave. The simplest QAM signal is quadrature phase shift keying (QPSK) with only one amplitude level, showing a high tolerance to Gaussian noise and making it suitable for long-haul transmission [26]. Today, advanced modulation formats (e.g., 16-, 64-, and 256-QAM) with multiple levels in each dimension are commonly used in modern high spectral-efficiency (SE) coherent systems to increase capacity for example in metro Networks [27, 28]. As shown in Fig. 2.2, the square M-QAM signal has $\sqrt{M}$ constellation points in each dimension and leads to a SE of $\log_2(M)$ bits/symbol/dimension-pair. Given a fixed signal power, the higher-order QAM formats have a lower Euclidean distance between the neighboring constellation points and therefore are more sensitive to additive Gaussian noise and other impairments. Modulating both polarizations further doubles the capacity regardless of

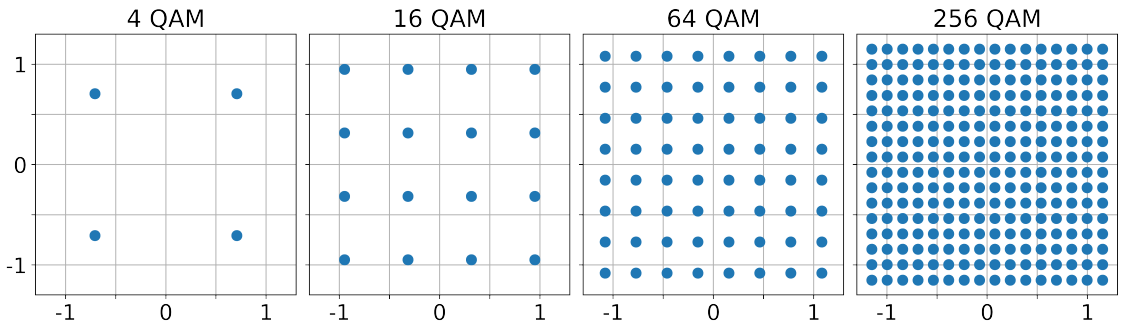

Figure 2.2: The illustration of QAM constellations.

polarization-dependent fiber channel impairments.

2.1.2 Pulse shaping

Pulse-shaping is the process of minimizing signal bandwidth without introducing intersymbol interference (ISI). For an input signal x, the pulse shaped signal y_{PS} can be expressed by

$$y_{\text{PS}}[n] = \sum_{k=-(m-1)/2}^{(m-1)/2} x[n] \cdot h(t - kT) \tag{2.1}$$

where $h(t)$ is the impulse response of the pulse shaping filter, k is the time index, and m is the memory length of the filter. The Nyquist ISI criterion defines the theoretical condition that no ISI is induced in the communication channel. An ideal pulse shaping satisfying the Nyquist ISI criterion can be realized by a sinc filter with a rectangular shape in the frequency domain. However, it is not achievable in practical systems due to the required infinite memory length, the limited resolution of the hardware, and the overall frequency response of the transmitter. Pulse shaping typically employs a raised-cosine filter [29] to eliminate ISI. Neglecting the channel response, the same root-raised-cosine (RRC) filter can be employed in the transmitter-side pulse shaping and receiver-side matched filtering to generate an overall raised-cosine response [30]. When practical transceivers and fiber channels lying between the two RRC filters break the perfect pulse shaping, the matched filtering is usually performed by adaptive equalization [31] described in Section 2.3.2.

The impulse response of the RRC filter is given by [32]

$$h_{\mathrm{RRC}}(t) = \begin{cases} \frac{1}{T_s}\left(1+\beta(\frac{4}{\pi}-1)\right) & t=0 \\ \frac{\beta}{T_s\sqrt{2}}\left[(1+\frac{2}{\pi})\sin\frac{\pi}{4\beta}+(1-\frac{2}{\pi})\cos\frac{\pi}{4\beta}\right] & t=\pm\frac{T_s}{4\beta} \\ \frac{1}{T_s}\frac{\sin\left[\pi\frac{t}{T_s}(1-\beta)\right]+4\beta\frac{t}{T_s}\cos\left[\pi\frac{t}{T_s}(1+\beta)\right]}{\pi\frac{t}{T_s}\left[1-(4\beta\frac{t}{T_s})^2\right]} & \text{otherwise} \end{cases} \tag{2.2}$$

where β is the roll-off factor and T_s is the symbol period.

2.1.3 Digital-to-analog converter

After signal mapping and pulse shaping, the discrete-time digital signals are converted to continuous-time analog waveforms using DACs. By scaling the digital signal to the maximum input range, the DACs with n-bits nominal resolution quantize the input signal into 2^n amplitude levels ranging from - $V_{\max}$ to $+V_{\max}$, with $V_{\max}$ being the maximum output voltage of the DACs. Ideally, the DACs can be modeled as a n-bits uniform quantizer. In practice the DAC exhibits jitter, noise, and other distortion originating from the electronic circuits. The effective number of bits (ENOB) describes the realistic resolution of the DACs, which are obtained from [33, 34]

$$\mathrm{ENOB(dB)} = \frac{\mathrm{SNDR(dB)} - 1.76}{6.02} \tag{2.3}$$

where the signal-to-noise and distortion ratio (SNDR) can be measured in the experimental setup. A typical SNDR of 35 dB at 0 Hz [35] corresponds to an ENOB of around 5.5 bits. It should be noted that the SNDR reduces with frequency [18], leading to a frequency-dependent ENOB and putting additional penalties on high-symbol-rate applications. When the ENOB is ≥ 4 bits, the effect of SNDR can be accurately modeled as additive Gaussian noise originating from the effective resolution. The noise variance is given by [36]

$$\sigma_q^2 \approx \frac{\Delta^2}{12} \tag{2.4}$$

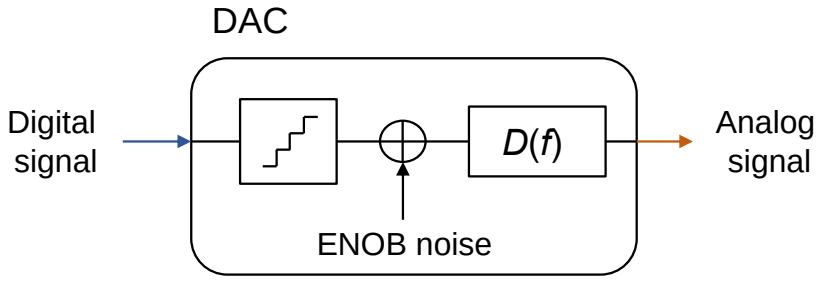

Figure 2.3: The block diagram of the digital-to-analog converter (DAC) consisting of an ideal quantizer, quantization noise corresponding to the effective number of bits (ENOB), and a filter to model the linear frequency response of the component.

where Δ is the quantization step size and expressed as [36]

$$\Delta = \frac{x_{\max}}{2^{\mathrm{ENOB}} - 1} = \frac{\sqrt{\sigma_x^2 \cdot 10^{\frac{\mathrm{PAPR}}{10}}}}{2^{\mathrm{ENOB}} - 1} \tag{2.5}$$

where $x_{\max}$ is the maximum input amplitude of the DAC. Assuming a constant input signal variance σ_x^2, we note that an increased peak-to-average power ratio (PAPR) results in larger quantization noise. This is because the majority of the signals are constrained to a smaller range, which leads to reduced effective resolution for the quantization. Therefore, any digital pre-processing that induces a higher PAPR should be optimized to eliminate the quantization noise. In general, proper clipping can be applied to the signal to reduce the PAPR [37] and therefore alleviate the penalty arising from the quantization effects. When we take into account the practical impairments, the realistic block diagram of the DAC consists of an ideal quantizer, additive Gaussian noise, and a linear filter corresponding to the frequency response of the circuits.

Recently, a digital pre-compensation approach was proposed to increase the effective resolution of the DAC, forming a so-called "digital resolution enhancer (DRE)" [38–40]. With a DRE, the required number of bits of the DAC to yield 400Gb/s 64-QAM was reduced to 4 [38], leading to less hardware complexity and power consumption. In addition, the joint optimization of DRE and clipping was presented to alleviate the quantization effects by further reducing the PAPR [40]. The DRE eliminates the in-band quantization noise by pushing it out-of-band, which broadens the signal spectrum and may impair the neighboring wavelength channels in wavelength division multiplexed systems. A narrow-

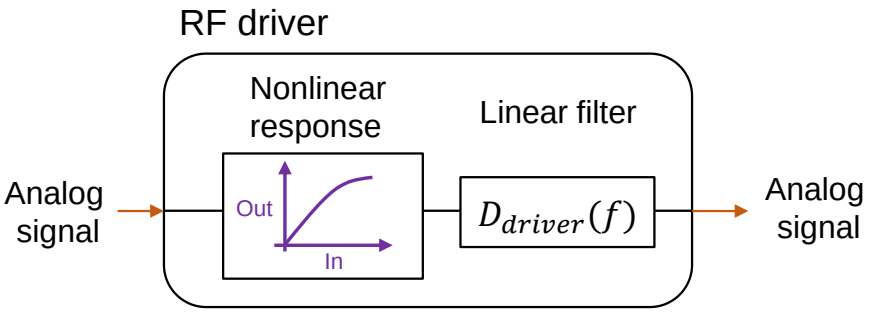

Figure 2.4: The block diagram of the RF driver consisting of a nonlinear function and a linear filter.

band filter can be applied to each channel to remove the out-of-band quantization noise but may induce an extra penalty in superchannel systems where several channels are densely packed.

2.1.4 Radio-frequency driver

The maximum output voltage of DAC is usually limited at about 0.5 V, which is significantly less than the V_π (explained in Section 2.1.5) of the modulator, resulting in strong optical noise originating from the booster EDFA and less transmission reach. Therefore, RF amplifiers need to be used at the modulator input, which boosts the signal power at the modulator output given a constant optical input power. The schematic of the RF driver is expressed as a combination of a nonlinear response caused by power saturation and a linear filter. For extremely small input signals, the RF amplifier has an ideal linear response with almost constant gain regardless of the input voltage. However, the gain of the driver decreases significantly for higher input power, which leads to amplitude-dependent nonlinear distortion for the electrical signals. Furthermore, the nonlinearity has memory [41], due to the combination of the nonlinear distortion with the limited bandwidth of the amplifier, which needs to be taken into account to fully describe the driver nonlinearity. As we primarily operate the driver in the high-gain configuration (i.e., nonlinear regime) to provide significant amplification for the electrical signal, the characterization of the linear and nonlinear properties of the driver become increasingly important for modern optical communication systems employing high-order modulation formats. Nonlinear pre-distortion can linearize the driver by using the inverse of the nonlinear response. More details can be found in Section 3.3.

2.1.5 In-phase/quadrature modulator

After digital pre-processing, digital-to-analog conversion, and electrical signal amplification, the baseband analog signal is ready for being modulated on the optical carrier. In coherent fiber-optic systems, the well-known Mach-Zehnder modulator (MZM) is the basic component being used for optical modulation. The schematic of an MZM is illustrated in Fig. 2.5 (a). The input of the MZM consists of the continuous-time optical carrier $E_{\text{in}}(t) = \sqrt{P}cos(2\pi vt/\lambda_0 + \phi_0(t)) = \sqrt{P}cos(w_0 t + \phi_0(t))$ determining the central wavelength λ_0 of the modulation, the electrical signal being modulated $r_1(t)/r_2(t)$, and a direct-current signal r_{bias} that is used to change the operation point of the MZM. The incoming optical signal is equally divided into two branches, where the phase-modulated signal at each arm is combined to form interference. For simplicity, the $\pi/2$ phase shift at the two outputs of the 3dB coupler is neglected as the bias signal r_{bias} will artificially modify the phase difference between these two arms. Assuming an ideal linear phase modulation at each arm with the drive voltage, the relationship between the output field $E_{\text{out}}(t)$ and the input field $E_{\text{in}}(t)$, in other words, the field transfer function, of the MZM can be expressed as

$$\frac{E_{\text{out}}}{E_{\text{in}}} = \frac{1}{2}\left[\exp\left(j\frac{\pi}{V_\pi}\big(r_1(t) + r_{\text{bias}}\big)\right) + \exp\left(j\frac{\pi}{V_\pi}r_2(t)\right)\right] \tag{2.6}$$

$$= \exp\left(j\frac{\pi}{V_\pi}\frac{r_1(t) + r_2(t) + r_{\text{bias}}}{2}\right)\cos\left(\frac{\pi}{V_\pi}\frac{r_1(t) - r_2(t) + r_{\text{bias}}}{2}\right), \tag{2.7}$$

where the first term in Eq. (2.7) describes the phase modulation property and the second term denotes the intensity modulation. By setting $r_{\text{RF}}(t) = r_1(t) = -r_2(t)$, we operate in push-pull mode of the MZM which yields an efficient pure intensity modulation without inducing chirp. Here, we neglect the relative phase rotation $\exp(j\frac{\pi r_{\text{bias}}}{2V_\pi})$. Considering the differential RF driving signal as $\Delta r_{\text{RF}}(t) = r_1(t) - r_2(t) = 2r_{\text{RF}}(t)$, the transfer function can be re-written as

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

where $r_{\text{in}}(t) = \Delta r_{\text{RF}}(t) + r_{\text{bias}}$ is the total electrical input signal and V_π is defined as the voltage of bias signal that cause π phase shift for the upper arm, thereby leading to a deconstructive interference.

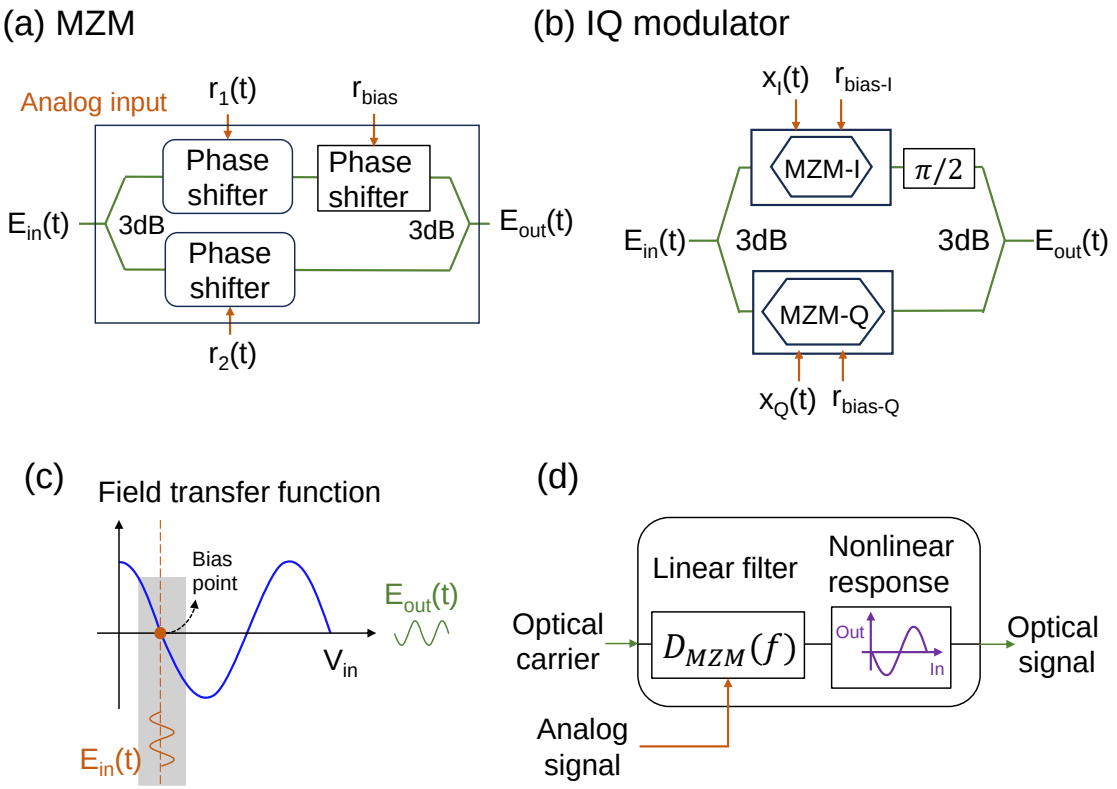

Figure 2.5: (a) Schematic of the Mach-Zehnder modulator (MZM), (b) Schematic of the IQ modulator, (c) transfer function of the MZM, (d) Block diagram of the modulator consisting of a linear filter and a nonlinear transfer function.

Given the drive voltage V_{in}, the field transfer function of MZM is depicted in Fig. 2.5 (c). The bias voltage is selected at the null point ($r_{\text{bias}} = (2n+1)V_\pi, n = 0, 1, 2, ...$) corresponding to zero power transmission such that we have a large swing for the differential RF driving signal $\Delta r_{\text{RF}}(t)$. To avoid nonlinear distortion, we would like to operate in the linear region of the MZM marked with grey color in Fig. 2.5 (c), where the transfer function can be approximated with $E_{\text{out}}/E_{\text{in}} \approx \pm\frac{\pi}{2V_\pi}\Delta r_{\text{RF}}(t)$ regardless of the nonlinearity. The sign of the output is plus for odd n and minus for even n, leading to an inverted output signal in the latter case. To modulate the data on the two quadratures, we use the IQ modulator as shown in Fig. 2.5 (b). It consists of two MZMs to perform individual amplitude modulations of each dimension and a $\pi/2$ phase shifter to make them orthogonal. Without considering the nonlinearity and the sign of the transfer function, the output field of the IQ modulator is

given by

$$E_{\text{out}(t)} = \frac{\pi}{V_\pi}\Big[r_{\text{I}}(t) + jr_{\text{Q}}(t)\Big]\sqrt{P}\cos(w_0 t + \phi_0(t)) \tag{2.9}$$

where P is the input optical power and ϕ_0 denotes the laser phase noise. The electrical signal for the in-phase/quadrature branch is expressed as $r_{\text{I/Q}}(t) = V_p x_{\text{I/Q}}(t) cos(w_{\text{RF}} t + \phi_{\text{RF}}(t))$ with V_p being the peak value of the driving voltage, $x_{\text{I/Q}}(t)$ the normalized encoded amplitude envelope and w_{RF} the angular frequency of the electrical signal. We observe from Eq. (2.9) that a small driving voltage (e.g., $V_{\text{DAC}} \leq 0.5V$) includes significant loss for the output optical signal given a typical $V_\pi = 3.5V$ for the popular $LiNbO_3$-based MZMs [42], explaining the need for electrical amplifiers before the optical modulation. Figure. 2.5 (d) illustrates the block diagram of the MZM consisting of a linear filter and a sine function to model the linear electrical response and the nonlinear optical transfer function, respectively.

2.1.6 Linear transmitter impairments

Linear transmitter impairments are primarily caused by the frequency response of each electrical and electro-optical component. It induces ISI in the signal, especially when the symbol rate is close to or larger than the 3 dB bandwidth of the transceiver. As shown in [43], the experimentally measured S_{21} response of the DAC and RF driver can be modeled as 2-order and 4-order Bessel function, while the S_{21} of the IQ modulator is approximated by 0.5-order Gaussian filter. The overall 3dB bandwidth of the transmitter jointly arises from all the devices, leading to strong ISI for high-symbol-rate applications. To alleviate the linear impairments, static or dynamic digital pre-emphasis (DPE) can be applied to the transmitter side for ISI cancellation. A detailed description of the DPE can be found in Section 3.2.

2.1.7 Nonlinear transmitter impairments

The nonlinear impairments of the transmitter comprise the nonlinearity caused by the IQ modulators and RF amplifiers. The sinusoidal transfer function of the MZMs limits the driving amplitude, inducing a trade-off between nonlinearity and signal-to-noise ratio (SNR) for high-order QAM formats. More importantly, the combination of linear ISI and nonlinear power saturation of the RF driver results in pattern-dependent

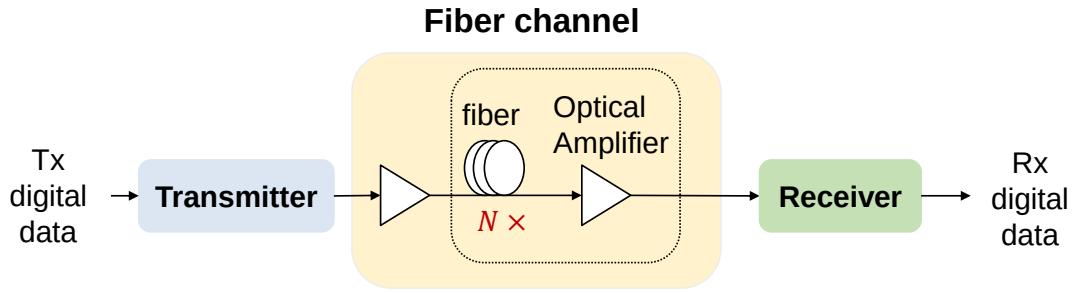

Figure 2.6: The diagram of the fiber-optic channel.

distortion, significantly degrading the transmitter performance. QPSK with only one amplitude level has a high tolerance to the transmitter nonlinearity. On the other hand, multilevel modulation formats are extremely sensitive to nonlinear impairments. A simple arcsine function can be applied to eliminate the memoryless nonlinear impairments caused by MZMs. Whereas more complex digital pre-distortion needs to be implemented to compensate for the transmitter's nonlinearity with memory. Various commonly used nonlinear pre-distortion approaches are described in Section 3.3.

2.2 Fiber Channel

The extremely low loss of optical fibers enables broad applications of long-haul fiber-optic transmission systems. A standard single-mode fiber (SMF) has an attenuation of <0.2 dB/km. A diagram of a fiber-optic communication system is shown in Fig. 2.6. The modulated optical signal is amplified by a booster before being transmitted in the fiber link, which includes multiple spans of fibers and inline amplifiers. The gain of the amplifier is equivalent to the signal attenuation to cancel the fiber loss and maintain the same launch power at each span. The amplifiers are typically erbium-doped fiber amplifiers (EDFAs), and commercial amplifiers have around 5.5 dB noise figure (NF) in the C-band [44] and a slightly higher 6 or 6.5 dB NF in the L-band [45]. The noisy optical signal with similar power to the transmitted one is detected in the receiver for demodulation.

2.2.1 The AWGN channel

In the simplest case, the fiber channel is fully described by the additive amplified spontaneous emission (ASE) noise originating from the inline EDFA disregarding fiber nonlinearity and dispersion. The fiber link can then be modeled with the additive white Gaussian noise (AWGN) channel $y = x + z_{\mathrm{ASE}}$, where x is the transmitted signal, y is the received signal, and z_{ASE} is the ASE noise. The received signal has an optical signal-to-noise ratio (OSNR) of

$$\mathrm{OSNR(dB)} = \frac{P_{\mathrm{sig}}}{P_{\mathrm{ASE}}} = 58 + P_{\mathrm{sig}} - L_{\mathrm{fiber}} - \mathrm{NF} - 10 \cdot \log 10(N) \quad (2.10)$$

where P_{sig} is the signal launch power, P_{ASE} is the power of the ASE noise, L_{fiber} is the fiber loss, and N is the number of transmission spans. Assuming an ideal noiseless transceiver, the OSNR is the fundamental evaluation metric for such linear fiber channels without dispersion.

2.2.2 The nonlinear fiber channel

For a single-polarization system, the signal propagation in an optical fiber can be described by the nonlinear Schödinger equation (NLSE). A dual-polarization signal propagating in the nonlinear fiber channel is given by solving the Manakov equation [46, 47]

$$\frac{\alpha}{2}U_x + \frac{\partial}{\partial z}U_x + j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}U_x = j\frac{8}{9}\gamma\left[|U_x|^2 + |U_y|^2\right]U_x \quad (2.11)$$

$$\frac{\alpha}{2}U_y + \frac{\partial}{\partial z}U_y + j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}U_y = j\frac{8}{9}\gamma\left[|U_x|^2 + |U_y|^2\right]U_y \quad (2.12)$$

where $U_{x/y}$ is the complex amplitude of the x/y polarization, α is the fiber attenuation, β_2 is the second-order dispersion parameter, and γ is the nonlinear coefficient. Here, the third-order dispersion is not included due to its negligible influence. The nonlinear term on the right side of the equation describes the Kerr nonlinear effects including self-phase modulation (SPM), cross-phase modulation (XPM), and four-wave mixing (FWM). As one of the most dominant impairments in long-haul transmission systems, the inherent intensity-sensitive nonlinear interaction limits the further increase of launch power and significantly degrades the SNR, resulting in less transmission reach and capacity.

2.2.3 Gaussian noise model

To simulate the fiber nonlinearity for performance estimation in long-haul systems, the split-step method is popularly used to numerically solve the NLSE. While solving the NLSE is reasonably straightforward on modern desktop PCs, the computational complexity is high when simulating multiple channels, and prohibitive for real-time applications. Therefore, many recent studies have focused on characterizing the Kerr nonlinear effect with simple models. The Gaussian noise model (GN-model) [48–51] is one of the most popular candidates in dispersion uncompensated systems where no dispersion compensation fiber is deployed after each span for optical dispersion compensation. In such systems with high accumulated dispersion, the nonlinear distortion can be considered as an added perturbation that is accurately approximated by Gaussian noise. Together with the ASE noise originating from the inline EDFA, the resulting total signal-to-noise ratio ($\mathrm{SNR}_{\mathrm{tot}}$) is written as [51]

$$\mathrm{SNR}_{\mathrm{tot}} = \frac{P_{\mathrm{sig}}}{P_{\mathrm{ASE}} + P_{\mathrm{NL}}} = \frac{P_{\mathrm{sig}}}{P_{\mathrm{ASE}} + \alpha_{\mathrm{NLI}} P_{\mathrm{sig}}^3} \tag{2.13}$$

where P_{NL} is the power of the nonlinear noise and α_{NLI} is a nonlinear parameter depending on the transmitted signal and fiber transmission. It should be noted that the GN-model is valid assuming an input Gaussian signal, thereby being insensitive to multilevel modulation formats. A more complicated extended GN model [52] can be used to include modulation-dependent effects, which precisely model the fiber nonlinearity in particular for high-order constellations.

2.2.4 The linear dispersive fiber channel

When the fiber nonlinearity is neglected, a linear dispersive model can describe the signal propagation by setting the nonlinear coefficient $\gamma = 0$ in Eq. (2.11)(2.12) [53]

$$\frac{\partial}{\partial z} U_x + \frac{\alpha}{2} U_x + j \frac{\beta_2}{2} \frac{\partial^2}{\partial t^2} U_x = 0 \tag{2.14}$$

$$\frac{\partial}{\partial z} U_y + \frac{\alpha}{2} U_y + j \frac{\beta_2}{2} \frac{\partial^2}{\partial t^2} U_y = 0 \tag{2.15}$$

Using the exponential attenuation of the amplitude and solving the Eq. (2.14)(2.15) in the frequency domain, the analytical solution of the

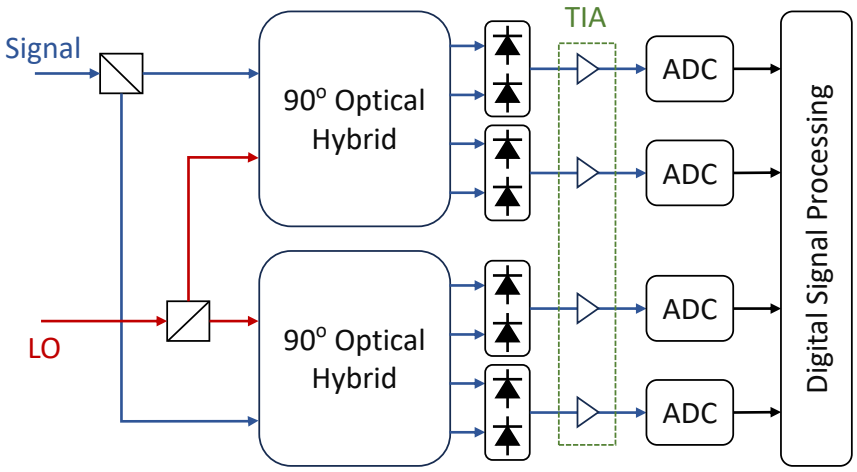

Figure 2.7: The diagram of the polarization diverse coherent receiver.

complex amplitude in a linear dispersive channel is given by

$$U_{x/y}(z,w) = U(0,w)\exp\left(j\frac{\beta_2 w^2 z}{2}\right) \tag{2.16}$$

We can observe from Eq. (2.16) that the dispersion induces a frequency-domain parabolic phase shift. Performing an inverse Fourier transform to Eq. (2.16), the phase shift in the frequency domain corresponds to a wavelength-dependent time delay. Therefore, a signal pulse will be broadened after transmission due to dispersion.

2.3 Receiver

In modern dual-polarization systems, a coherent receiver with polarization diversity is commonly used to detect the signal. The schematic diagram of a coherent receiver is shown in Fig. 2.7.

2.3.1 Optical front-end

The signal and local oscillator (LO) are sent to a polarization beam splitter before being combined in a 90° optical hybrid separately for each polarization. The balanced photo-detectors perform optical-to-electrical conversion and trans-impedance amplifiers (TIA) boosts the signal for the in-phase and quadrature of each polarization. Finally, the electrical signals are quantized in the analog-to-digital converters. Similar to the DACs, the frequency-dependent ENOB of the ADCs is an important

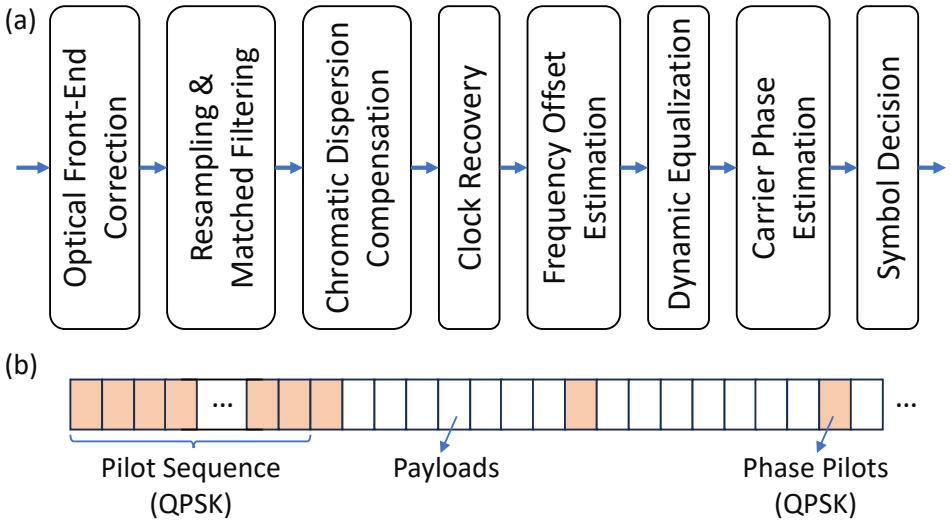

Figure 2.8: (a) The diagram of the receiver-side DSP, (b) The frame structure of the pilot-based DSP.

parameter for the receivers, especially for high-order modulation formats. It should be noted that polarization demultiplexing is not performed in the optical front end due to the unknown polarization of the input signal. A polarization controller can be added to a single-polarization input signal to align the polarization. However, for a dual-polarization signal, the digital signal processing (DSP) on the receiver side can efficiently demultiplex the polarization components.

2.3.2 Receiver-side digital signal processing

In modern coherent communication systems, the rapid development of the CMOS technique enables high-speed ADC and application-specific integrated circuits for DSP. Most importantly, this avoids a complicated phase-locked loop and allows the use of a free-running laser as the local oscillator, which significantly simplifies the receiver hardware structure. In addition, various dominant impairments such as transceiver imperfections, dispersion, carrier offset, and fiber nonlinearity are efficiently compensated in the DSP. Today, advanced DSP [54, 55] (main block diagram shown in Fig. 2.8 (a)) facilitates the implementation of high-SE multilevel modulation formats to achieve high-capacity transmission.

The received signal may experience IQ phase/gain mismatch, which is caused by the sub-optimum bias of the MZM, non-ideal 90° phase shift between the in-phase and quadrature components, the difference in the RF drivers as well as imperfections in the 90° optical hybrid,

photo-detectors, and transimpedance amplifiers. Higher-order modulation formats are more sensitive to IQ imbalance due to the increased number of constellation points. Therefore, orthogonalization techniques such as the Gram-Schmidt algorithms [56] or Löwdin [57, 58] algorithms are used. The former approach fixes one vector and forces the other to be orthogonal, while the latter one rotates both vectors in different directions to obtain orthogonality. This symmetric operation makes Löwdin algorithms more suitable for high-cardinality constellations.

Static equalization is needed to compensate for the time-invariant impairments. As previously discussed in Section 2.1.2, the RRC filter is employed in the receiver to remove out-of-band noise and perform matched filtering with the assumption of neglecting the channel response. The practical matched filtering is implemented in the dynamic equalizer with more details being provided later in this section. Additionally we compensate for the dispersion using static chromatic dispersion compensation (CDC) which avoids the need for adding dispersion compensation fiber after each transmission span to cancel the dispersion. The CDC acts by applying a filter with an impulse response which is equivalent to the inverse transfer function of the chromatic dispersion. While the popular finite-impulse response (FIR) filter can efficiently perform the time-domain dispersion equalization for short-reach transmission [59] in which a small number of taps is enough to cancel the small accumulated dispersion, frequency-domain dispersion equalization is preferred for long-haul transmission system [60] due to its relatively low implementation complexity for large tap number compared to the time-domain filtering. The frequency-domain equalizer can be implemented in an overlap-add approach [61] or an overlap-save scheme [62, 63] with lower complexity than the former one.

Another important DSP is the carrier offset recovery that compensates for the laser frequency fluctuation and phase noise caused by the transmitter laser source and the LO. While blind algorithms based on fourth-power operation [64–66] (e.g., the well-known Viterbi-Viterbi algorithm for phase estimation [67]) are efficient for performing carrier recovery for QPSK with constant amplitude, more advanced approaches such as blind phase search [68] or pilot-aided algorithms [69–71] should be employed for high-order constellations with various modulation phase components. The last impairments that need to be considered are the residual static distortions (originating from imperfect matched filtering, CDC, and other time-invariant system responses) as well as other time-

variant phenomenons including fast polarization rotation and polarization mode dispersion (PMD). These residual static and dynamic impairments can be mitigated in the blind adaptive multiple-input multiple-output (MIMO) equalizer whose coefficients are iteratively updated by stochastic gradient descent-based adaptation algorithms such as the constant modulus algorithm [72, 73], radially directed algorithm [74, 75], and decision-directed least-mean-square algorithm [76]. We should emphasize that, in contrast to blind methods, data-aided equalization is more stable, particularly for high-symbol-rate and high-SE applications. Therefore, we will concentrate on the discussion of pilot-based DSP used throughout the research presented in this thesis.

Pilot-based digital signal processing

The schematic of the frame structure used in one popular pilot-based DSP [77] is shown in Fig. 2.8(b). Importantly, the data-aided dynamic equalization and frequency offset estimation (FOE) are performed by a long QPSK pilot sequence located at the beginning of the frame. It should be emphasized that the periodically inserted known QPSK symbols serve as phase pilots and facilitate the pilot-enabled carrier phase estimation (CPE). With the frame length of l_{frame}, the pilot sequence length of l_{ps}, and the insertion ratio (defined by the ratio between the number of payload/information symbols and the number of phase pilots) of R_{ins}, the total pilot ratio is expressed as

$$\alpha_{\text{pilot}} = \frac{(l_{\text{frame}} - l_{\text{ps}})/R_{\text{ins}} + l_{\text{ps}}}{l_{\text{frame}}} \tag{2.17}$$

An optimized pilot ratio for a high-symbol-rate high-SE system (e.g., 24Gbaud 256-QAM [78]) can be as low as 1.4% by setting $l_{\text{frame}} = 104704$, $l_{\text{ps}} = 1024$, and $R_{\text{ins}} = 256$. However, when the symbol rate decreases to 5 Gbaud or 10 Gbaud, the insertion ratio of the phase pilots needs to be reduced to track the larger sensitivity to phase fluctuations originating from the increased symbol duration. For the 12.2 bits/s/Hz C-band transmission of 5 Gbaud 24-dimensional Voronoi constellations (with record $2^{96} = 7.9 \times 10^{28}$ total constellation points in the 24-dimensional space and around 540 unique points in each two-dimensional plane) presented in Paper V, we increase the phase pilot ratio be setting $R_{\text{ins}} = 128$ and reduce the pilot sequence length to 512. The resulting relatively high overall pilot ratio of 3.8% is primarily due to the low frame length of 16896 symbols, which is constrained by the

combined effects of limited DAC storage memory and the high upsampling ratio of 12 (required from the 60GS/s DACs). In such conditions, we find that the small DAC memory significantly constrains the pilot overhead regardless of the short pilot sequence length and large phase insertion ratio being used.

For coherent optical systems, the pilot-based adaptive equalization is usually performed by the 2×2 MIMO FIR filter with a butterfly structure expressed as [55]

$$\begin{bmatrix} \mathbf{r}_X[n] \\ \mathbf{r}_Y[n] \end{bmatrix} = \begin{bmatrix} \mathbf{w}_{XX}^H[n] & \mathbf{w}_{XY}^H[n] \\ \mathbf{w}_{YX}^H[n] & \mathbf{w}_{YY}^H[n] \end{bmatrix} \begin{bmatrix} \mathbf{r}_X^{'}[n] \\ \mathbf{r}_Y^{'}[n] \end{bmatrix} \tag{2.18}$$

where $\mathbf{r}^{'}/\mathbf{r}$ represents the input/output complex signal vector with N samples, $\mathbf{w} = [w(n), w(n-1), ..., w(n-N)]^T$ are the complex coefficients of the N-tap FIR filter, n is the time index, and the superscript H indicates the conjugate transpose. To compensate for dynamic impairments, the filter coefficients need to be continuously optimized over time such that time-varying effects can be tracked. The basic principle of the updates can be understood as minimizing the mean square error denoted by the cost function $J[n] = \mathrm{E}[|e[n]|^2]$ with $e[n]$ being the error signal. For data-aided equalization [79], the error functions simply calculate the difference between the known transmitted symbol c and the received noisy symbol r separately for each polarization according to

$$\begin{aligned} e_X[n] &= c_X[n] - r_X[n] \\ e_Y[n] &= c_Y[n] - r_Y[n] \end{aligned} \tag{2.19}$$

The coefficients can be efficiently updated by the gradient descent algorithms, which approach the optimum solution through calculating the gradient of the cost function for each update. The most well-known candidate is the least-mean-square (LMS) algorithm [80] which uses a filter coefficient update according to

$$\begin{aligned} \mathbf{w}_{XX}[n+1] &= \mathbf{w}_{XX}[n] + \mu e_X^* \mathbf{r}_X^{'}[n] \\ \mathbf{w}_{XY}[n+1] &= \mathbf{w}_{XY}[n] + \mu e_X^* \mathbf{r}_Y^{'}[n] \\ \mathbf{w}_{YX}[n+1] &= \mathbf{w}_{YX}[n] + \mu e_Y^* \mathbf{r}_X^{'}[n] \\ \mathbf{w}_{YY}[n+1] &= \mathbf{w}_{YY}[n] + \mu e_Y^* \mathbf{r}_Y^{'}[n]. \end{aligned} \tag{2.20}$$

Following these update rules, this filter forms a so-called "data-aided LMS equalizer". It is worth emphasizing that the use of a longer pilot

sequence provides better training for the equalizer and leads to improved performance at the expense of transmitting more redundant information. The balance between the equalization performance and the information loss due to the inserted pilot symbols needs to be taken care of to fully optimize the system performance.

For blind carrier offset recovery algorithms, the fourth power operation is usually used to remove the modulation phase to facilitate the phase noise and frequency offset compensation. The data-aided algorithms take advantage of the known training symbol for the estimation and therefore support the mitigation of stronger carrier offset [81]. Assuming a perfect synchronization, the transmitted and received symbols are aligned to each other without timing delay. The synchronized training sequence is then ready for the frequency offset estimation with the first step to remove the disturbance caused by the modulation phase. The signal z with modulation phase mitigation is then expressed as

$$z[n] = y[n]c^*[n] \tag{2.21}$$

where $y[n] = c[n]\exp\left(j(2\pi\Delta f n T_s + \phi[n])\right)$ is the complex received signal with uncompensated carrier offset, c is the known pilot symbol, Δf is the frequency offset, and T_s is the symbol duration. By inserting the full expression of $y(n)$ in Eq. (2.21), it can be re-written as

$$z[n] = |c[n]|^2 \exp\Big(j\big(2\pi\Delta f n T_s + \phi[n]\big)\Big) \tag{2.22}$$

From Eq. (2.22) we see that the modulation phase from c(n) is fully removed. Therefore, the estimation of frequency offset is obtained as

$$\Delta f_{\text{est}} = \frac{1}{2\pi T_s}\arg\left\{\sum_{n=1}^{L}\Big[z[n]z^*[n-1]\Big]\right\} \tag{2.23}$$

After the compensation of FOE, the received signal is re-written as $y_{\text{FOE}}[n] = c[n]\exp\left(j\phi[n]\right)$. Following a similar process according to Eq. (2.21), the signal z_{cpe} with modulation phase compensation can be given by

$$z_{\text{cpe}}[n] = \begin{cases} y_{\text{FOE}}c[n], & n = l_{\text{PS}} + kR_{\text{ins}} \\ 0, & n \neq l_{\text{PS}} + kR_{\text{ins}} \end{cases} \tag{2.24}$$

where $k = 0, 1, 2, ...$, the residual phase term corresponds to the phase noise $\phi[n]$ if $n = l_{\text{PS}} + kR_{\text{ins}}$, l_{PS} and R_{ins} is the pilot sequence length and phase pilot insertion ratio defined previously, respectively. Since the

transmitted symbols remain unknown except for the periodically inserted pilots, we need to apply an interpolation filter [69, 70] to the signal z_{cpe} before estimating the phase noise for the whole sequence including the information and pilots symbols. More details regarding the optimization of the pilot overhead were investigated for varying simulated impairments including frequency offset, phase noise, and PMD [77].

Digital backpropagation

In modern coherent communication systems, fiber nonlinearity is one of the most detrimental effects, which inhibits the increase of capacity and transmission reach, in particular for superchannel systems employing densely spaced wavelength channels. State-of-the-art receiver DSP can perform compensation of the nonlinear distortion caused by the channel of interest itself as well as other interference channels. Digital backpropagation (DBP) is an efficient digital compensation approach that sends the received signal through the inverse nonlinear Schödinger equation, which is normally emulated by the split-step method (SSM), to reverse the nonlinear distortion. For single-carrier systems, the single-channel DBP can be applied to the traditional coherent receiver structure and DSP for compensating the SPM. However, the SPM becomes insignificant in WDM systems where the XPM is the dominant effect, requiring multichannel DBP to fully mitigate the SPM and XPM. More details of the multichannel DBP are described in Section 4.4.

2.4 Forward error correction

FEC coding is one of the most important building blocks in modern fiber communication systems. It protects the information from noise by adding redundant bits or symbols in the transmitted sequence. One common and powerful coding structure is the concatenated code [82, 83] consisting of one outer hard-decision (HD) code and another inner soft-decision (SD) code. In the coded modulation system depicted in Fig. 2.9, the bit error rate (BER) before decoding, also known as pre-FEC BER, is used to estimate the system performance without FEC coding. Error-free transmission requires a BER of 10^{-15}-10^{-9}, which is difficult to achieve in uncoded systems. With the assistance of FEC coding, communication with negligible error rates are possible is possible. Typically, while the inner SD code can reduce the BER to 10^{-5}-10^{-2} if the SNR reaches a code-dependent threshold, the output HD code brings the BER down to

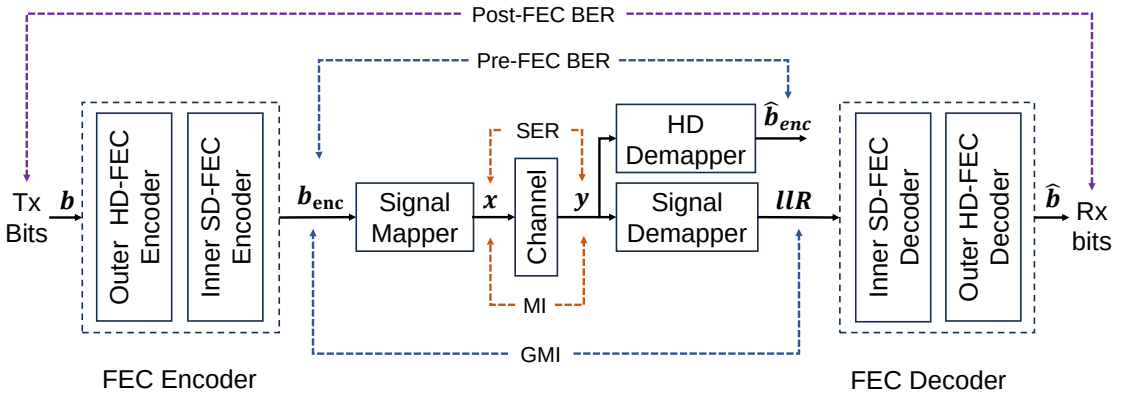

Figure 2.9: The schematic of the coded modulation system.

10^{-15}-10^{-9} after HD-FEC decoding. For long-haul transmission, inner SD-FECs with 20% overhead (OH) and a BER threshold of around 2.7×10^{-2} [84] are commonly used together with an outer $< 1\%$ OH HD-FEC. A popular candidate for SD-FEC is the low-density parity-check (LDPC) code, which has been included in the DVB-S2 standard [85]. HD codes with 7% OH and a BER threshold of 3.8×10^{-3} [86] are widely employed for short-reach transmission in which the system impairments are usually significantly smaller.

2.5 Modulation formats and limiting factor

Similar to the noise degradation stemming from nonlinear fiber propagation (see Section 2.2.3) we can also model the residual linear and nonlinear degradation (after pre-distortion) from the transceivers as an additive Gaussian noise term. The total SNR of the system is given by

$$\mathrm{SNR}_{\mathrm{tot}} = \frac{P_{\mathrm{sig}}}{P_{\mathrm{TRx}} + P_{\mathrm{ASE}}} = \frac{1}{\frac{1}{\mathrm{SNR}_{\mathrm{TRx}}} + \frac{1}{\mathrm{OSNR}}} \tag{2.25}$$

where P_{TRx} is the power of the transceiver noise. Figure 2.10 illustrates the total SNR ($\mathrm{SNR}_{\mathrm{tot}}$) and BER as a function of OSNR assuming a fixed transceiver SNR ($\mathrm{SNR}_{\mathrm{TRx}}$) of 23 dB measured in our optical back-to-back experimental setup for a 20 GBaud signal. For OSNR>20 dB, the transceiver noise starts to have a significant contribution to overall SNR for high-order QAM formats. Given a 20% SD OH and a BER threshold of 2.7×10^{-2}, from Fig. 2.10 (b) we observe a larger required

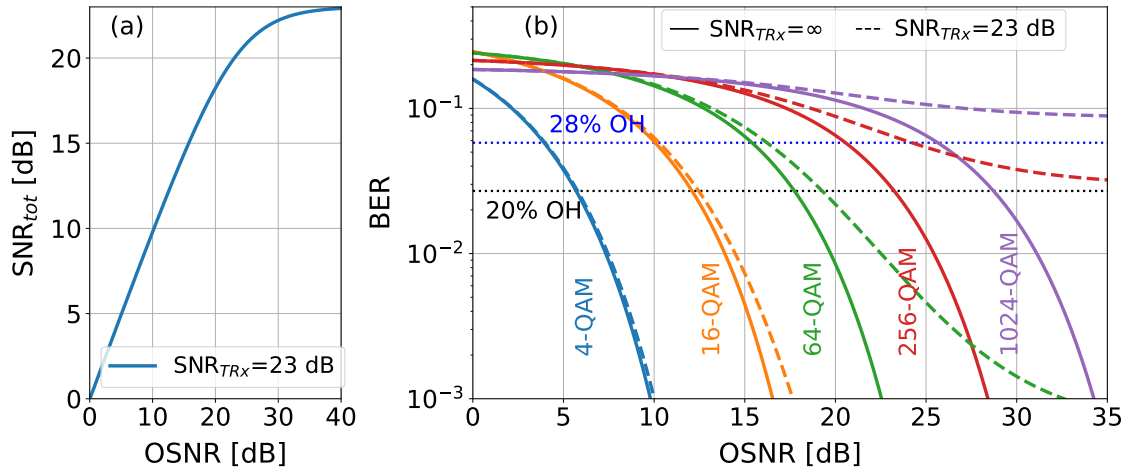

Figure 2.10: Given a fixed $\mathrm{SNR_{Tx}} = 23$ dB, (a) the simulated SNR and (b) BER performance, versus OSNR for various QAM formats.

OSNR at $\mathrm{SNR_{TRx}} = \infty$ (i.e., no transceiver noise) for higher-order formats due to their reduced euclidean distance between the constellation points. Furthermore, at $\mathrm{SNR_{TRx}} = 23$ dB, the transceiver noise significantly limits 64-QAM, 256-QAM, and 1024-QAM. A higher SD OH of 28% [87] is required for 256-QAM to achieve error-free transmission at OSNR=35 dB. The BER of 1024-QAM is around 1×10^{-2} even at large OSNRs, making it challenging to perform forward error correction with the commercially available SD code.

Chapter 3

Digital Pre-distortion

Today, the majority of the impairments in coherent communication systems are compensated after signal detection. The development of the CMOS technique in the 2000s enabled the fabrication of high-speed analog-to-digital converters (ADCs) and application-specific integrated circuits (ASICs) [88, 89], which were first used in intensity-modulation direct-detection (IM/DD) systems and then applied to the commercial coherent transceiver in 2008 [90]. The ADCs and ASICs allow the applications of powerful receiver-side digital-to-signal processing (DSP) [91, 92] to mitigate impairments such as chromatic dispersion (CD) [93, 94], polarization mode dispersion (PMD) [95], polarization rotation, carrier offset, transceiver imperfections, and even fiber nonlinearity [96].

Alternatively, pre-compensation can be implemented in the transmitter to alleviate the impairments with CD being the popular one. In earlier WDM systems, each transmitted wavelength channel was propagated in individual dispersion compensation fiber (DCF) before wavelength multiplexing [97]. The inverse dispersion provided in the DCF reduces the CD effect during the signal propagation. Similarly, a frequency modulation technique was used to generate a pre-distorted signal and reduce the total dispersion of the whole system [98]. The aforementioned pre-compensation is achieved in an optical manner and changes with the experimental configuration, which significantly increases the implementation cost and complexity.

Digital-to-analog converters (DACs) and transmitter DSP facilitate high-spectral efficiency transmission by enabling advanced modulation and pulse-shaping. At the same time, we can take advantage of DSP capabilities to implement digital pre-compensation (DPC) to compen-

sate for various transceiver distortions and channel impairments [99–104]. Chromatic dispersion and fiber nonlinearity can be pre-compensated digitally in the transmitter [102–104]. In contrast to the optical approach, the DPC can be updated in the transmitter ASIC without changing the system configuration. For modern coherent systems employing multi-level modulation formats, the generated signal suffers from significant transceiver imperfections such as the frequency response of each component, the nonlinear response of the RF amplifier, and the sinusoidal transfer function of the IQ modulator. The compensation for these transceiver impairments in the receiver would incur an extra penalty due to the mixing of transceiver and fiber channel distortions. Therefore, it is important to apply digital pre-distortion (DPD) to mitigate the transceiver distortion for state-of-the-art coherent communication systems.

This chapter describes the transmitter-side DSP to compensate for linear and nonlinear transmitter impairments of high-spectral efficiency systems. Various popular linear and nonlinear pre-distortion approaches are discussed and compared in terms of performance and complexity.

3.1 Learning structure of pre-distortion

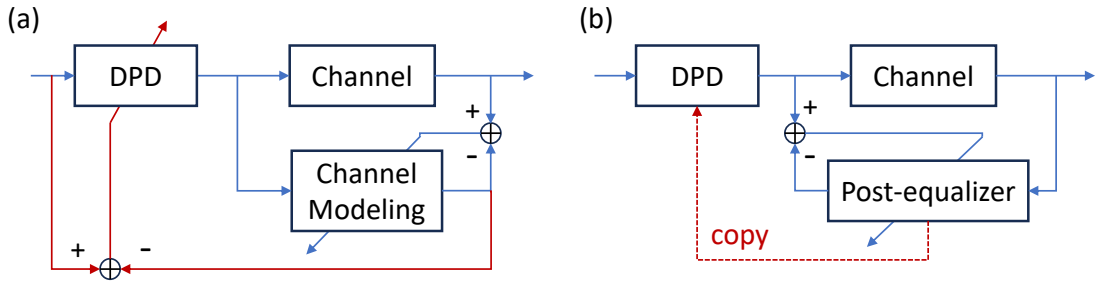

Figure 3.1: The illustration of (a) direct learning architecture and (b) indirect learning architecture.

A static filter can perform DPD to cancel the linear or nonlinear transmitter's distortion by generating an inverse response in the transmitter DSP. However, it needs to derive the filter response from the characterization of each hardware component, channel modeling, and analytical calculation, which results in significant implementation complexity. In contrast, a dynamic filter can achieve pre-equalization without the time-consuming hardware characterization and the complex mathematical processing. In addition, it adaptively updates the compensation if the

channel response changes. The popular structures for training the adaptive pre-distortion can be classified as the direct learning architecture (DLA) [105, 106] and the indirect learning architecture (ILA) [107–110], which is illustrated in Fig. 3.1 (a) and (b), respectively. Generally, the former method requires identifying the realistic channel with a differential surrogate model to estimate the gradient for the coefficient update in the DPD, whereas the ILA first learns a post-equalizer (see Section 2.3.2) and then copies the optimized coefficients to the pre-distorting filter. Ideally, the DLA outperforms the ILA, which is trained with a noisy feedback signal, at the expense of increased complexity used for the channel modeling [111, 112]. It should be emphasized that not all nonlinear systems have perfect theoretically inverted expressions [113, pp. 123], limiting the widespread application of the DLA. In practice, the inaccurate identification of the channel response would lead to performance degradation for the DPD. Therefore, the low-complexity ILA has become more popular in practical systems.

3.2 Linear pre-distortion

As previously stated in Section 2.1.6, the frequency response of each component in the transmitter induces inter-symbol interference and degrades system performance. To alleviate this effect, a static or dynamic finite impulse response (FIR) filter, commonly referred to as a digital pre-emphasis (DPE) filter, can be applied to the transmitter DSP to compensate for the bandwidth limitation, where the majority of the degradation in commercial transponders is from the DACs and ADCs. The system diagram illustrated in Fig. 3.2 describes the implementation of a linear FIR filter $P(f)$ for linear pre-emphasis. For simplification, we only include the dominant transceiver components (i.e., DACs and ADCs) and a linear fiber channel (i.e., $H_{\text{ch}}(f) = 1$) in the model.

3.2.1 Zero-forcing pre-emphasis filter

For the simplest case, a zero-forcing (ZF) filter that yields the inverse linear response of the DAC can be applied after pulse shaping to achieve pre-emphasis. Given the frequency response $D_{\text{DAC}}(f)$ and the desired transfer function $D_{\text{DAC, des}}(f)$ of the DAC, the ZF pre-emphasis filter is given by $P_{\text{ZF}}(f) = D_{\text{DAC, des}}(f)/D_{\text{DAC}}(f)$. However, the ZF generates a large gain for small-value frequency components. This phenomenon amplifies out-of-band noise and leads to a significant performance penalty.

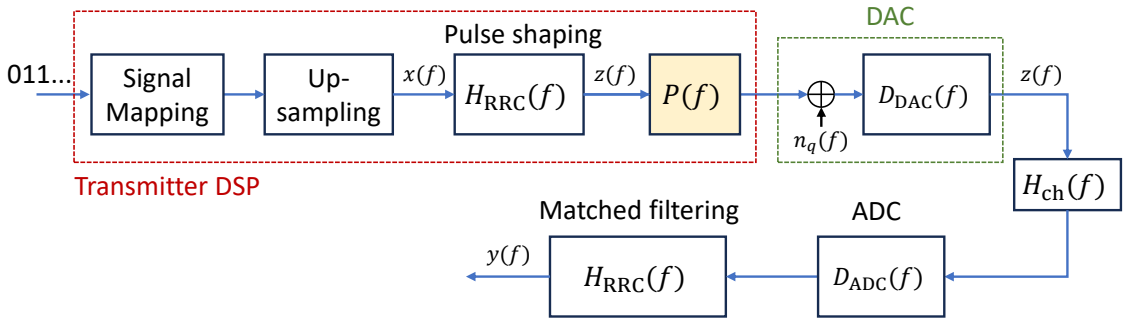

Figure 3.2: The illustration of transmitter-side digital signal processing with linear pre-distortion, $P(f)$: the linear pre-emphasis filter.

To solve the overshot problem, a modified ZF filter was proposed in [34] with the optimized filter response given by

$$P_{\mathrm{ZF}}(f) = \frac{D_{\mathrm{DAC,\,des}}(f)}{D_{\mathrm{DAC}}(f) + n_o} \tag{3.1}$$

where n_o is a constant noise-related term and needs to be optimized individually for a given modulation format, bandwidth, and the effective number of bits to minimize the BER. The experimentally measured frequency response of the DAC closely matches a third-order Bessel function, while the desired transfer function can be modeled by the Gaussian function. The detailed optimization procedure can be found in [34].

3.2.2 Minimum-mean-square-error pre-emphasis filter

The ZF filter provides sub-optimum performance because it does not take quantization noise into account. From Fig. 3.2 we can see that, assuming $D_{\mathrm{ADC}}(f) = 1$, the received signal $y(f)$ can be expressed in the frequency domain as

$$y(f) = x(f)H_{\mathrm{RRC}}^2(f)P(f)D_{\mathrm{DAC}}(f) + n_q(f)H_{\mathrm{RRC}}D_{\mathrm{DAC}}(f) \tag{3.2}$$

where $x(f)$ is the frequency-domain expression of the upsampled signal before pulse shaping. To obtain the desired received signal $y_{\mathrm{des}}(f) = x(f)H_{\mathrm{RRC}}^2(f)$, a minimum-mean-square-error (MMSE) approach can minimize the square error given by [114]

$$e_{\mathrm{MMSE}} = \int_{-B}^{B} E\left[\left|y(f) - x(f)H_{\mathrm{RRC}}^2(f)\right|^2\right] df \tag{3.3}$$

where $B = (1 + \beta_{RRC})R_s$, β_{RRC} is the roll-off factor of RRC filter, R_s is the symbol rate, and $E[\cdot]$ corresponds to the expectation operator. Using a partial derivative approach, the analytical solution of the DPE is calculated as [114]

$$P_{\mathrm{MMSE}}(f) = \frac{H_{\mathrm{RRC}}^2(f) D_{\mathrm{DAC}}^*(f)}{H_{\mathrm{RRC}}^2(f) \left|D_{\mathrm{DAC}}(f)\right|^2 + \varepsilon} \tag{3.4}$$

with ε including the quantization noise for the optimization of DPE filter. For an ENOB> 4 bits, the quantization noise can be replaced with AWGN according to [36]. Given a fixed ENOB and peak-to-average power ratio (PAPR), ε is analytically expressed as

$$\varepsilon = \frac{10^{\frac{\mathrm{PAPR}}{10}}}{6R_s \cdot 2^{2\mathrm{ENOB}}} \int_{-B}^{B} \left|D_{\mathrm{DAC}}(f)\right|^2 H_{\mathrm{RRC}}^2(f)\, df. \tag{3.5}$$

In contrast to the modified ZF filter, the MMSE-based DPE has an analytical solution for the noise-related term and avoids the heuristic optimization.

3.2.3 Dynamic time-domain pre-emphasis filter

The implementation of static DPE filters requires a priori knowledge of the DAC. Therefore, the time-consuming measurement of the frequency response and ENOB inhibits its application in elastic optical networks [115, 116], where the transceivers from different vendors would be used in different system configurations. An approximation of the DAC characteristics can be used but results in a performance penalty. Adaptive DPE filters outperform their static counterparts in terms of compatibility, allowing it to be dynamically updated for a wide range of transceivers without adding implementation complexity.

For the compensation of weak bandwidth limitations, a short (few-tap) DPE filter can be efficiently implemented in the time domain. In the simplest case, the time-domain dynamic DPE filter was obtained in an electrical back-to-back link by directly connecting the converted analog signal to an oscilloscope [99, 100]. Similarly, it was trained in an optical back-to-back setting assuming perfect compensation of the fiber channel impairments with DSP [99]. The ILA method depicted in Fig. 3.1 (b) can be applied to learn the filter coefficients. A post-equalizer is first trained with pilot symbols to achieve an inverse linear response of the transceiver. Then the same coefficients are copied to the DPE filter for

pre-compensation. Generally, the dynamic DPE filter can be trained by the least mean square (LMS) algorithms by minimizing the mean square error (MSE) between the transmitted and received signals. The LMS-type approaches can be categorized into blind or pilot-based methods by choosing different error functions. For the data-aided LMS (DA-LMS) algorithms with error function $\varepsilon_{\text{TD-LMS}}(n) = \widehat{y}(n) - x(n)$ with n being the time index, x being the transmitted signal, and $\widehat{y}$ being the output of the post-equalizer, the optimization of DA-LMS is expressed as [80]

$$\min_{p(n)} \text{MSE} = E\left[|\varepsilon_{\text{TD-LMS}}|^2\right] \tag{3.6}$$

where $p(n)$ is the coefficient of the time-domain DPE filter. As a static gradient descent algorithm, the update of the filter coefficients using the LMS method is given by [80]

$$p_{k+1}(n) = p_k(n) - \mu \frac{\partial \text{MSE}_k}{\partial p_k^*(n)} = p_k(n) + \mu \widehat{y}_k(n) \varepsilon_{\text{TD-LMS}}^*(n) \tag{3.7}$$

where μ is the step size used for controlling the update speed. This approach can perform an overall pre-emphasis on DACs, RF drivers, and the ADCs. Besides the electrical and optical noise, the strong nonlinear distortion caused by the power saturation of the driver may significantly affect the identification of the DPE coefficients and degrade the pre-compensation performance.

3.2.4 Dynamic frequency-domain pre-emphasis filter

The adaptive DPE filter can be implemented in the frequency domain to compensate for linear transmitter responses with large memory. Similar to the time-domain approach, the DA-LMS algorithm can be applied to optimize the filter coefficients. The error function is given by $\varepsilon_{\text{FD-LMS}}(f) = \widehat{y}(f) - x(f)$ in the frequency domain. The optimization is provided by minimizing a weighted mean square error (WMSE) [117]

$$\min_{P(f)} \text{WMSE} = \int_{-B}^{B} E\left[|x(f)|^2\right] E\left[|\varepsilon_{\text{FD-LMS}}(f)|^2\right] df \tag{3.8}$$

where the term $E\left[|x(f)|^2\right]$ constrains the optimization on the high-power frequency components. Adaptive updating of the coefficients can be performed by stochastic gradient descent until the normalized error

$|\varepsilon_{\text{FD-LMS}}(f)|^2/|x(f)|^2$ is below a given threshold. The coefficients update is expressed as

$$P_{k+1}(f) = P_k(f) - \mu \frac{\partial \text{WMSE}_k}{\partial P_k^*(f)} \tag{3.9}$$

where the detailed expression of Eq. (3.9) can be found in [117, Eq. 4]. To reduce the implementation complexity, the overlap-and-add method [118] was used for the LMS implementation.

3.3 Nonlinear pre-distortion

The pattern-dependent distortion caused by the RF amplifier and IQ modulator significantly limits the performance of advanced modulation formats with multiple amplitude levels. To alleviate this effect and facilitate high-SE constellations, nonlinear pre-distortion needs to be implemented. The diagram of the overall linear and nonlinear DPD is illustrated in Fig. 3.3. Various widely employed nonlinear DPD approaches, including look-up table (LUT)/neural networks (NNs) implemented on the symbol basis and arcsine function/Volterra series (VS)/NN applied on the sample basis, are introduced in this section.

3.3.1 Arcsine function

The sinusoidal transfer function of the IQ modulation generates non-equidistantly spaced constellation points and reduces the noise tolerance for the high-energy constellation points. An input signal voltage close to or larger than the inherent V_π of the modulator causes strong nonlinear distortion (See Fig. 2.5 (c)). Given a fixed transmitter electrical noise, an optimum output voltage of DACs exists as a trade-off between the SNR and the achievable linearity of the transmitter. An arcsine function can be applied after the pulse shaping to pre-compensate for the nonlinear response of the MZM. It should be noted that the arcsine function significantly increases the PAPR for large input signal swings. Therefore, an optimization of the signal amplitude is required for the arcsine-based pre-distortion. From Eq. (2.5) we can see that an increased PAPR induces larger quantization noise, requiring a clipping function added as the last step of the transmitter DSP to eliminate this degradation.

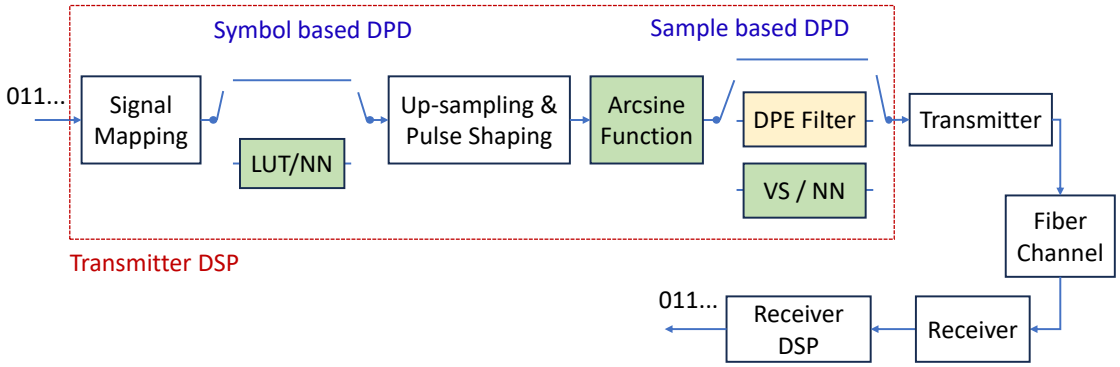

Figure 3.3: The illustration of transmitter-side digital signal processing with linear and nonlinear pre-distortion. LUT: look-up table; DPE: digital pre-emphasis; VS: Volterra series; NN: neural network.

3.3.2 Look-up table

Generally, DACs have limited output power, which leads to high modulation loss when the converted analog signal is directly fed to the IQ modulator (see Eq. (2.9)). Therefore, RF amplifiers are required to boost the signal before optical modulation such that its amplitude is close to the V_π of the modulator. However, given a large input power, the output of the RF driver is saturated and the signal is significantly distorted due to the inherent nonlinear behavior, resulting in a reduced output power and increased modulation loss.

In wireless communication, so-called "amplifier linearization" applies a digital predictor to the baseband signal to mitigate the nonlinear effects of the RF power amplifier [119–122]. A look-up table (LUT) is one popular candidate to compensate for this nonlinear behavior. As a low-complexity solution, it simply adds a correction term to the transmitted signal depending on the signal itself (memoryless LUT) or the signal pattern of a given length (memory LUT) such that the received signal is closer to the ideal constellation. Historically, LUT-based DPD was first employed to alleviate the memoryless effects caused by the instantaneous nonlinearity [119], which means that the symbol at a given time is fully distorted by the symbol itself but not one at a different time. However, the effect of bandwidth limitation becomes non-negligible for an increased symbol rate and leads to intersymbol interference. This linear memory together with instantaneous nonlinearity induces nonlinear distortion with memory, requiring an implementation of a memory DPD to fully remove all the nonlinearity.

The size of the LUT will exponentially grow with modulation order

and memory length (explained later), requiring a large storage memory. Several size-reduction approaches have been proposed to alleviate this effect. Initially, a two-dimensional LUT was applied to pre-compensate for the complex constellations [120]. Since the in-phase and quadrature dimensions are generally independent in the coherent transmitter, two independent LUTs can be employed for the in-phase and quadrature parts, respectively, providing size reduction [123]. Due to the power sensitivity of the nonlinear effects, the size of the LUT can be significantly reduced by only considering the pre-distortion of the high-energy constellation points [124] or the signal pattern with strong nonlinearity [123]. Assuming a symmetric signal distribution with the same number of amplitude levels in each dimension, the LUT-based DPD is suitable for other signal formats including probabilistically-shaped (PS) and/or geometrically-shaped (GS) constellations.

Here, we assume that two independent (and similar) LUTs are implemented for the in-phase and quadrature components, respectively. The pattern errors (PEs), which characterize the nonlinear distortion for each signal pattern, are first estimated in the training stage. Then we apply the obtained PEs in the compensation stage to perform nonlinear DPD. Given a memory$-n$ LUT (i.e., LUT-n) with n denoting the memory of the nonlinear effects, the size of the LUT is given by $2 \cdot N^n$, where $N = \sqrt{M}$ for even $\log_2(M)$ and $N = (\sqrt{2 \cdot M} + \sqrt{M/2})/2$ for odd $\log_2(M)$. For PS-QAM and GS-QAM formats, N is the number of unique constellation points in each dimension. Importantly, the size will exponentially increase with N and the memory length of the LUT. All length-n pattern sequences $\boldsymbol{s}$ of the employed signal need to be stored in the LUT as an entry address for the identification of the input signal, while $\boldsymbol{s}(p)$ indicates the p^{th} pattern sequence. During the training process, a length-n sliding window is first applied to extract the transmitted signal sequence $\boldsymbol{x}_k^{\text{I/Q}} = \{x_{k-\frac{n-1}{2}}^{\text{I/Q}}, ..., x_k^{\text{I/Q}}, ..., x_{k+\frac{n-1}{2}}^{\text{I/Q}}\}$ separately for the in-phase/quadrature to identify the input pattern $\boldsymbol{s}(p_{i/q})$, where k corresponds to the central index in the extracted sequence. For the transmitted symbol $x_k^{\text{I/Q}}$ and the corresponding post-processed symbol $y_k^{\text{I/Q}}$, the distorted error is calculated as $e_k^{\text{I/Q}} = x_k^{\text{I/Q}} - y_k^{\text{I/Q}}$, which includes Gaussian noise except for the nonlinear impairments. To eliminate the influence of random noise fluctuation on the error estimation, all the errors corresponding to the same input pattern $\boldsymbol{s}(p_{i/q})$ are averaged

following

$$\delta^{\mathrm{I/Q}}(p_{i/q}) = \frac{1}{N_{p_{i/q}}} \sum_{\boldsymbol{x}_k^{\mathrm{I/Q}} = \boldsymbol{s}(p_{i/q})} e_k^{\mathrm{I/Q}} \tag{3.10}$$

where $N_{p_{i/q}}$ is the number of signal pattern $\boldsymbol{x}_k^{\mathrm{I/Q}}$ in the training signal sequence that is equivalent to $\boldsymbol{s}(p_{i/q})$. Then, the estimated PEs are stored in the LUT using the corresponding entry address. In the compensation process, the same sliding window as the training stage is used for pattern identification before performing nonlinear mitigation with the pre-distorted PEs in the LUT. For each extracted sequence $\boldsymbol{x}_k^{\mathrm{I/Q}}$, the central symbol is corrected by

$$x_k^{\mathrm{tx}} = x_k^{\mathrm{tx-I}} + j \cdot x_k^{\mathrm{tx-Q}} = x_k^{\mathrm{I}} + \delta^{\mathrm{I}}(p_i) + j \left(x_k^{\mathrm{Q}} + \delta^{\mathrm{Q}}(p_q) \right) \tag{3.11}$$

For a 64-QAM signal, the unique constellation points in the in-phase or quadrature dimensions are expressed as pulse-amplitude modulation (PAM) with 8 levels (i.e., $\mathrm{PAM}_8 = \{-7, -5, -3, -1, +1, +3, +5, +7\}$). It is interesting to investigate the amplitude distribution of the estimated PEs in terms of the pattern index $p_{i/q}$, which shows the relationship between the nonlinear effects and a given signal pattern. This helps us to have a better understanding of the nonlinear behavior in a straightforward sense. An arbitrary or unordered index, depending on the order of occurrence of the input sequence pattern, was used in [123] for patterns with the same central symbol. It can not illustrate any nonlinear distribution except the overall variance of the nonlinearity (see [123, Fig. 4]). We apply the indexing shown in Table 3.1 for a memory-3 LUT with $\{-7, -7, -7\}$ being the first pattern and $\{+7, +7, +7\}$ being the last pattern. From the beginning, every continuous 8 patterns have the same first and second symbols, while every continuous 64 patterns share the same first symbol. Following this structured indexing, the nonlinear behavior among the signal patterns can be clearly illustrated. We observe that every group of 64 patterns sharing a common first symbol has a similar pattern error distribution except for a mean difference, resulting in a periodic shape. It should be noted that the distinct PEs of in-phase and quadrature LUT are caused by the difference between the employed RF amplifiers, the IQ imbalance, and the bias operation point of the modulator. This confirms that independent LUTs need to be applied to the in-phase and quadrature components of the signal to avoid performance penalty. The mechanism of the periodicity can be explained by modeling the nonlinear effects by a third-order Volterra filter.

Table 3.1: Order of the pattern indexing in a memory-3 LUT for 64-QAM.

Index p	Pattern Sequence $\boldsymbol{s}$			Index p	Pattern Sequence $\boldsymbol{s}$		
1	-7	-7	-7	65	-5	-7	-7
2	-7	-7	-5	66	-5	-7	-5
3	-7	-7	-3	67	-5	-7	-3
4	-7	-7	-1	68	-5	-7	-1
5	-7	-7	1	69	-5	-7	1
6	-7	-7	3	70	-5	-7	3
7	-7	-7	5	71	-5	-7	5
8	-7	-7	7	72	-5	-7	7
9	-7	-5	-7	.	.	.	.
10	-7	-5	-5	.	.	.	.
11	-7	-5	-3	.	.	.	.
12	-7	-5	-1	.	.	.	.
.	.	.	.	505	7	7	-7
.	.	.	.	506	7	7	-5
.	.	.	.	507	7	7	-3
.	.	.	.	508	7	7	-1
61	-7	7	1	509	7	7	1
62	-7	7	3	510	7	7	3
63	-7	7	5	511	7	7	5
64	-7	7	7	512	7	7	7

As the majority of the nonlinearity is contributed by the memoryless term, the signal pattern with the same central symbol will experience similar nonlinear distortion.

In paper A, we utilize this periodic nonlinear property to generate reduced-size LUTs with significantly smaller storage memory requirements. The size reduction and DPD performance are compared among various modulation formats such as 64-QAM, 256-QAM, and 1024-QAM. A higher size-reduction efficiency is achieved for higher-order constellations due to the increased periodic patterns in each dimension, which is attractive for modern high-SE optical communication.

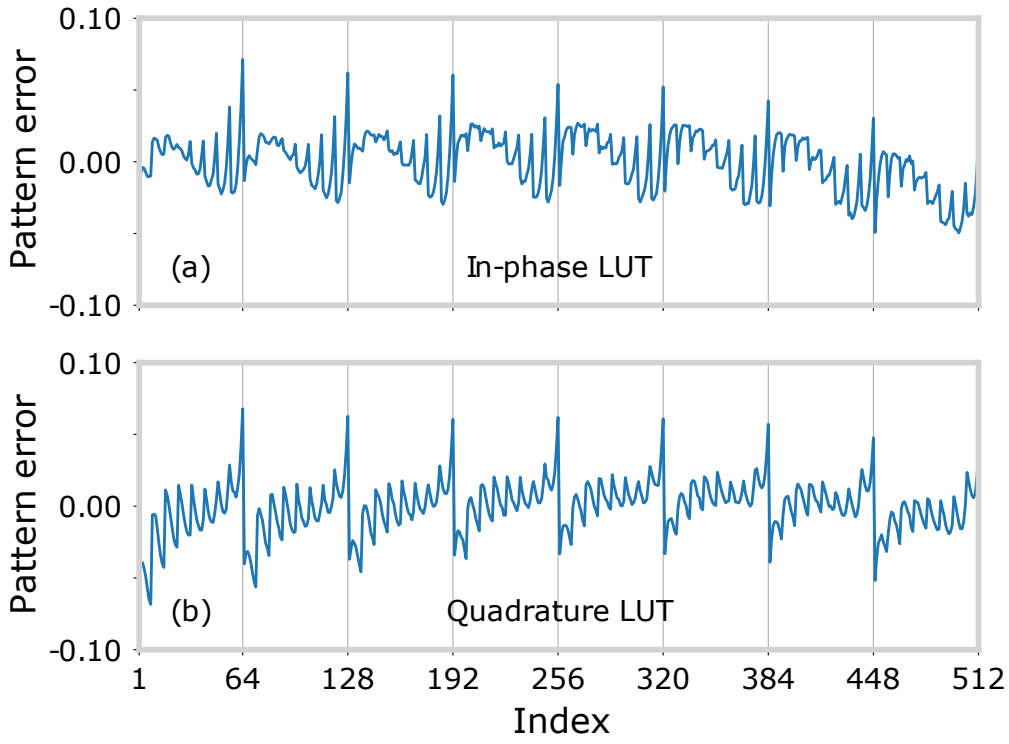

Figure 3.4: (a) In-phase and (b) quadrature pattern error of the memory-3 LUT estimated for a single-polarization 20 Gbaud 64-QAM.

3.3.3 Volterra series

The Volterra series is widely used to mathematically represent linear and nonlinear system in a single step, while a filter comprising an inverse response can perform the pre-distortion for the whole system. Compared to the linear and LUT-based DPD, it performs a joint compensation of the linear and nonlinear impairments at the expense of increased complexity. Similar to the dynamic time-domain DPE filter, the ILA architecture can efficiently compute the inverse response of the characterized nonlinear system model. Ignoring the quantization effect that can be adequately expressed by additive white Gaussian noise, the combined response of DAC and RF driver can be compensated by a truncated p-th order Volterra series given by [101]

$$y[n] = h_0 + \sum_{k_1=0}^{m_{p-1}} \cdots \sum_{k_p=k_{p-1}}^{m_{p-1}} h_p[k_1, \ldots, k_p] \cdot \prod_{i=1}^{p} x[n - k_i - \tau_p] \qquad (3.12)$$

where h_0 and h_1 indicate the DC and linear composition of the system, respectively. h_p is the kernel related to the p-th order nonlinearity, m_p is the corresponding memory length, and x/y is the input/output signal

of the Volterra filter. The possible delay between the 1-st and p-th component is characterized by τ_p. Given a large memory and high-order nonlinearity, the full expression in Eq. (3.12) is complex and thus unattractive for real-time practical systems. To reduce the complexity, the Eq. (3.12) can be simplified to a memory polynomial given by [125]

$$y_{\mathrm{MP}}[n] = h_0 + \sum_{i=0}^{p-1} \sum_{k=0}^{m-1} h[k] \cdot x[n-k] \, |x[n-k]|^i \tag{3.13}$$

Eq. (3.12) indicates the Volterra series output as a sum of the linearly-weighted nonlinear samples. The coefficients of such post-equalizers in the ILA can be approximated by least-square algorithms. Cholesky decomposition and Newton algorithm yield block-based estimations of the coefficients [125], requiring an extremely high computational complexity. The sampled-based estimation using stochastic gradient descent is preferred due to the low storage memory requirements. The coefficients of the pre-distorted Volterra filter can be iteratively updated by the LMS algorithm [80]. However, the nonlinearity becomes much smaller for high-order components, leading to a poor condition of the covariance matrix. The low convergence speed in such conditions inhibits its application in practical systems. The alternative recursive least-square algorithm provides a fast-convergence solution with higher computational complexity than the LMS [80]. After training, the learned coefficients are copied to the pre-distorter for transmitter-side nonlinear compensation.

3.3.4 Neural networks

Today, NNs have attracted significant attention in optical communication. End-to-end learning, which includes a transmitter and receiver NN, performs a joint optimization of the whole system and provides an alternative approach for mitigating various impairments in coherent systems [126]. NN-based DPDs have been studied to compensate for nonlinear impairments caused by different components of the transmitter [108, 127]. Feed-forward NNs (FFNNs) without memory were used to mitigate the nonlinearity of the MZM [128] or the low-resolution DAC [129]. Similarly, the memoryless nonlinear effect of the RF amplifier was canceled by a FFNN [130], which outperforms memory polynomials. To take into account the memory effects, time-delay NNs [131, 132], convolutional NNs (CNNs) [133], and recurrent NNs [134] were used to perform nonlinear DPD. Considering the low-pass filtering of the components and the nonlinear distortion of RF drivers, the coherent transmitter

was efficiently modeled by a Wiener-Hammerstein (WH) structure [135], which comprises two linear FIR filters and a memoryless nonlinear function inserted between them. Note that the same WH structure was used in the DPD block to achieve the inverse response. A cascaded WH structure including multiple instantaneous nonlinear functions can include the nonlinear transfer function of the MZM and even the nonlinearity of the DAC [135]. However, it requires significantly higher computational complexity and complicates the training process of the DPD, which makes this model less attractive. Inspired by the WH model, a simple NN-based DPD was proposed by replacing the two linear filters with CNNs and the nonlinear function with a fully connected FFNN [136]. Even though it significantly outperforms the Volterra-based DPD learned in ILA, this approach uses the DLA for training the NNs and therefore requires a numerical model including linear and nonlinear response of each component.

The required channel modeling in the NN-based DPDs becomes cumbersome in the practical system because some of the components have non-differentiable responses. A surrogate model can approximate the channel response to assist the gradient estimation with possible mismatch to the real system [137], leading to degradation of the DPD performance. In such conditions, reinforcement learning (RL) [138] can alleviate this problem by adding a random Gaussian perturbation to the transmitted signal for gradient estimation. In Paper B, the RL-based DPD was directly trained in an experimental system without channel modeling, showing its great potential in state-of-the-art coherent communication applications.

Chapter 4

Joint Digital Signal Processing

In this chapter, we describe receiver-based digital signal processing. In contrast to traditional digital signal processing (DSP) operates on channels separately, we will discuss here how multi-channel processing can increase spectral efficiency (SE) and system throughput. Linear processing such as joint phase estimation and multi-channel equalization performs better laser phase noise estimation and cancels the interchannel interference, respectively. While multi-channel nonlinear compensation eliminates both the intra-channel and inter-channel nonlinear crosstalk.

4.1 Comb-based superchannels

Current commercially available single-carrier WDM transponders achieved data rates of 400 Gb/s in 2016 [14] and 800 Gb/s in 2020 [15,16]. Considering annual capacity increases of 40% for router blades, the required bit rates per interface will reach 10 Tb/s by 2024 [139]. State-of-the-art coherent transponders utilize the four available dimensions (i.e., in-phase, quadrature, and two polarizations) and high-order modulation formats to achieve high capacity in each wavelength channel, whereas the limited effective number of bits (ENOB) and nonlinear transceiver distortion inhibit the applications of higher-order QAMs such as 256-QAM, 512-QAM, and 1024-QAM. Given a dual-polarization 64-QAM, the required symbol rate and 3-dB transceiver bandwidth for a 10 Tb/s transponder is 833 Gbaud and >400 GHz, respectively. Such symbol rates are not feasible with the current generation of transponders. Even in the future,

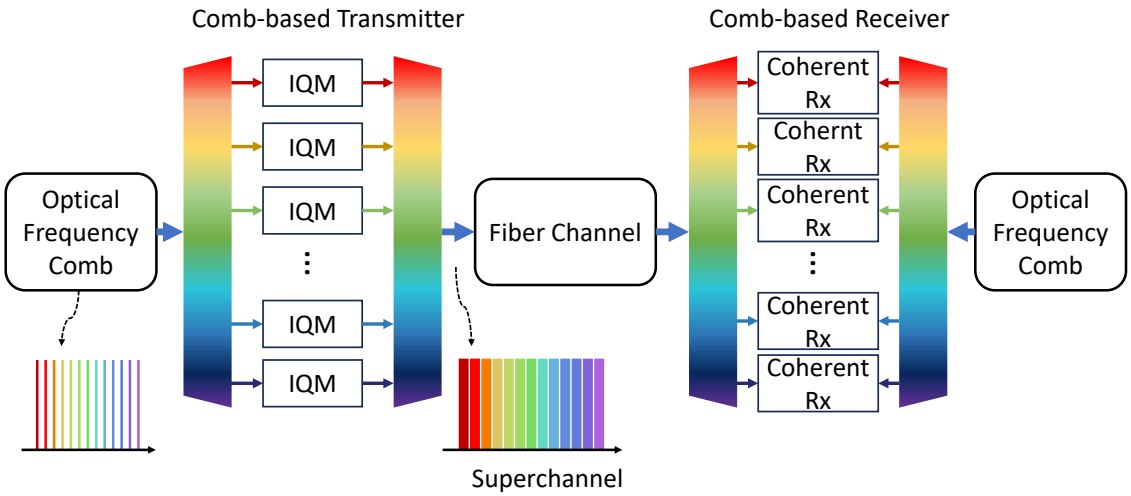

Figure 4.1: The schematic of the superchannel system assisted by frequency comb.

there is likely no avenue to get to such symbol rates. In addition, ENOB, the dominant performance characteristic of the digital-to-analog converters (DACs)/analog-to-digital converters (ADCs), decreases dramatically with frequency as stated in Section 2.1.3, resulting in more degradation for high-symbol-rate applications. These stringent requirements make single-carrier solutions unrealistic.

In order to attain the 10 Tb/s line rate with commercially available components, multiple wavelength carriers, which are independently modulated, could be combined to form so-called spectral superchannels. In this scenario, the transmitter's capacity is easily scaled up without upgrading the DACs, amplifier drivers, and IQ modulators. Using a state-of-the-art 1 Tb/s single-carrier transponder enabled by 90 Gbaud 64-QAM signal [140] and not considering the digital overhead, ten wavelength channels can offer the required 10 Tb/s bit rate. Given the limited spectrum such as the commonly used C-band in commercial fiber-optic communication systems, SE denotes the achievable bit rate per unit frequency, deciding the upper limit of the system capacity. The SE of the superchannels is strongly related to guard bands (GBs) between the modulated wavelength carriers [139]. The selection of large GBs avoids linear interchannel interference (ICI) arising from the spectral overlap between the neighboring channels. In contrast, densely packed channels could improve the SE and maximize the channel capacity. Due to the random frequency drift, the conventional system using an independent laser for each channel requires GBs of a few GHz to alleviate ICI [141].

An optical frequency comb (OFC), a multi-carrier laser source with frequency and phase coherence, yields evenly spaced carrier lines, resulting in a sub-GHz-GBs superchannel and significantly improved SE [54, 142]. Compared to the system employing laser arrays constructed by independent lasers, only a required extra wavelength demultiplexer is added after the OFC to separate the carriers.

For a given targeted symbol rate R_s, the bandwidth of the signal needs to be minimized to improve the SE by using a reduced GB. Pulse shaping therefore becomes important as it limits the signal bandwidth. Aggressive optical filtering, as a popular implementation of pulse shaping before the digital transponder era, was used to confine the signal bandwidth and reduce the guard bands. Due to the implementation of ASIC-based transmitter DSP, digital pulse shaping has become increasingly popular. However, as stated in Section 2.1.3, the limited resolution of the DAC and the transmitter frequency response inhibits the generation of an ideal rectangular signal spectrum. This hardware imperfection leads to a spectral gap in the optimized superchannels, which inhibits Nyquist WDM transmission with zero GBs. The so-called super-Nyquist-WDM with negative GBs can be achieved by special pulse shaping to eliminate ISI. Alternatively, in the comb-based superchannels, multichannel equalization can overcome ISI and facilitate super-Nyquist transmission with more details being described in Section 4.3.

Electro-optic frequency comb

An electro-optic frequency comb (EO-comb) is one popular approach for the OFC generation. A common block diagram of an EO-comb is illustrated in Fig. 4.2. The optical line provided by a continuous-wave laser is fed to a cascade of two phase modulators and one intensity modulator, driven by a common radio-frequency clock. Each modulator generates multiple sidebands of the input signal. If the phases are adjusted appropriately, the cascaded modulation will yield a wideband frequency comb. The center wavelength and frequency spacing of the output comb is determined by the wavelength of the input carrier and the frequency of the RF clock, respectively. The phase shift induced by the phase modulators determines the number of generated comb lines and thus the bandwidth of the comb. A larger phase shift can be achieved by either driving the phase modulator with a higher voltage or increasing the number of phase modulators. From Fig. 4.2 we see that the EO-comb can generate around 50 lines. Given an input carrier frequency of f_c and a RF frequency of

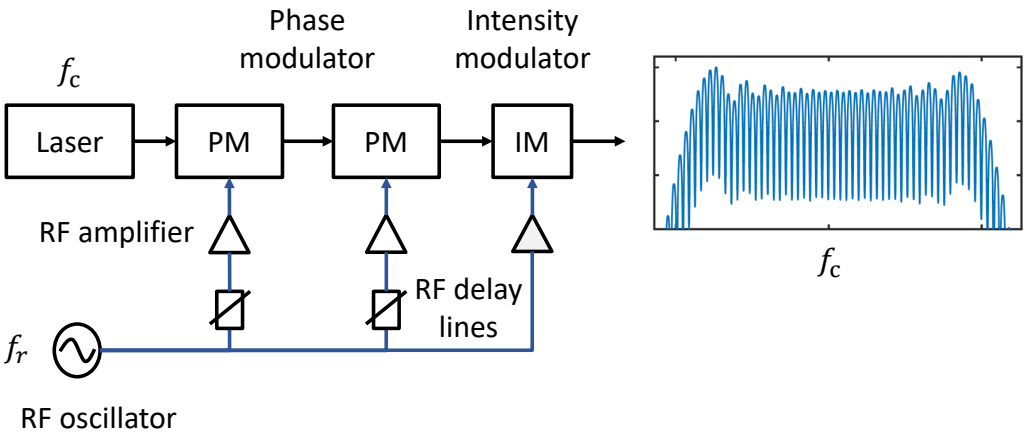

Figure 4.2: The schematic of an EO-Comb. The output signal spectrum indicates the generated comb with around 50 high-power lines spaced at 25GHz.

f_c, the frequency of the EO-comb is expressed as $f_n = f_c + nf_r$. Taking the inherent phase coherence into account, the phase noise of each comb line is given by

$$\phi_n(t) = \phi_c(t) + n\Delta\phi_r(t) \tag{4.1}$$

where $\phi_c(t)$ denotes the random phase fluctuation of the carrier and $\Delta\phi_r(t)$ corresponds to the phase noise of the RF oscillator. Importantly, the delay lines added to each phase modulator should be optimized to maintain a flat output comb. Ideally, the SE of the comb could be optimized with a water-filling [143, 144], i.e. modulating lines with different modulation formats according to their comb power and thus OSNR after modulation. In practical low-complexity transponders employing the same modulation formats in each wavelength channel, we perform proper optical shaping to flatten the comb lines, enabling similar performance among the channels. Generally, the comb has to be amplified after modulation to obtain sufficient launch power. Thus, after amplification, the OSNR of the individual sub-channels is dependent on the optical power of the corresponding line before modulation.

At the receiver, we can either use a bank of individual lasers to serve as local oscillators (LOs) for coherent detection or replace the receiver lasers with a second comb. When we are using individual lasers, the phase and frequency coherence of the superchannel is broken, as both the transmitter laser and receiver LO determine the phase noise of each channel. In contrast, a comb-based LO can maintain the phase-locked

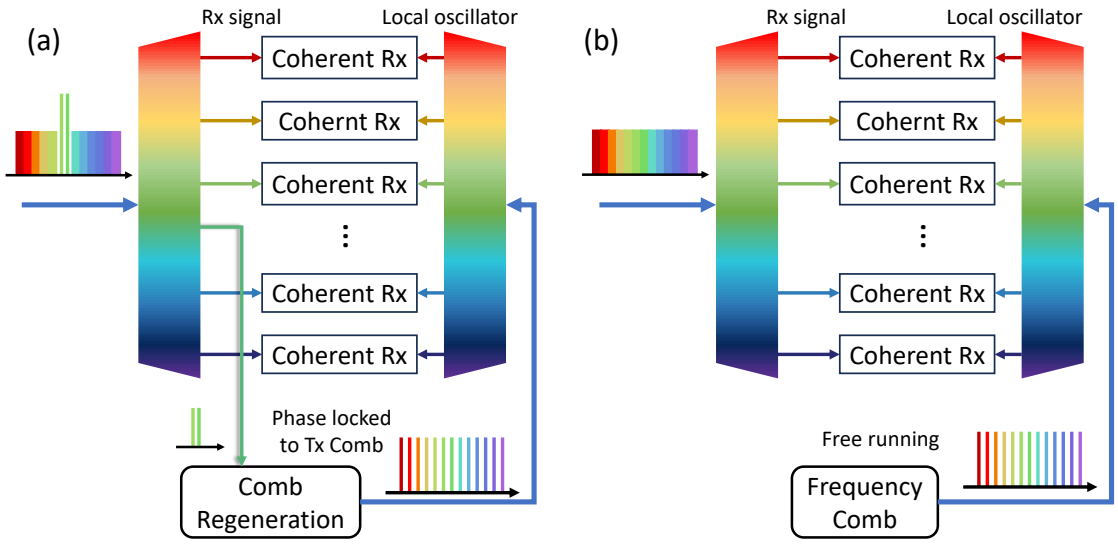

Figure 4.3: The schematic of the comb-based receiver for joint phase recovery (a) Comb regeneration, (b) Free running LO comb.

channels after detection. The coherence properties preserved by the comb-based transmitter and receiver enable a wide range of receiver-side joint DSP such as joint carrier recovery [145,146], multi-channel equalization [147, 148], and multi-channel nonlinearity compensation [149, 150].

4.2 Joint carrier recovery

The comb-based superchannel providing frequency stability and phase coherence among the wavelength channels facilitates the joint processing of carrier offset in the receiver. We therefore primarily discuss the joint carrier recovery approaches employing comb-based LOs.

The joint carrier recovery can be achieved by two different receiver structures as shown in Fig. 4.3. The first approach relies on comb regeneration assisted by optical pilot tones and duplicates an identical transmitter comb for the receiver LOs, enabling multi-wavelength self-homodyne detection. The comb generation using two optical pilots [151, 152] can be performed by a mode-locked laser, four-wave mixing, or an EO-comb with an electrical phase-locked loop to recover the phase-locked RF clock. Alternatively, when the RF phase noise is much smaller than the optical one, a single optical pilot can regenerate the comb and reduce the optical pilot overhead, enabling a high-SE superchannel experiment illustrated in [78,153]. Since the frequency offset and phase noise are fully removed in an optical manner, the receiver DSP is

significantly simplified. Generally, the self-homodyne method leads to better performance at a cost of hardware complexity and SE, especially for high-order modulation formats which are more sensitive to the carrier offset, compared to intradyne detection utilizing a digital approach for the compensation. The second method implements a free-running comb as LO and performs the digital joint processing for carrier offset compensation [145, 146]. In modern fiber-optic communication systems, low hardware complexity is preferable in terms of cost. We therefore focus on the discussion of the digital joint carrier recovery.

When employing an unsynchronized receiver comb, the carrier offset remains uncompensated after signal detection, requiring a two-stage process in the conventional receiver DSP to mitigate the frequency and phase fluctuations. The frequency offset estimation algorithm is first applied to cancel the slowly drifting frequency offset, then the phase estimation algorithm compensates for the phase variation [146]. We assume that the common frequency offset estimation is used without joint processing, the schematic building block of the joint phase estimation is illustrated in Fig. 4.4 (a). First, the phase noise of each wavelength channel is individually estimated regardless of the processing in other channels. Then, the estimated phase noise of the targeted N channels is averaged to generate a more accurate phase estimation by alleviating the noise during the phase estimation. Finally, the phase correction obtained by jointly processing N channels is applied to each subchannel for phase recovery. This approach improves the system's phase noise tolerance at the expense of computational complexity. On the other hand, the master-slave method, depicted in Fig. 4.4 (b), performs phase estimation of the master channel and utilizes the same phase correction for other slave channels to reduce the DSP complexity. As stated in Section 4.1, the additional phase noise caused by the RF oscillator linearly increases with the order of the comb lines from the center, meaning that the performance of this approach reduces with the increased number of processed channels. To extend the master-slave phase noise compensation approach, the same procedure can apply frequency offset estimation of the master channel to the slave channel by taking the spacing difference between the transmitter and LO comb into account, which further simplifies the implementation of carrier offset recovery.

Importantly, the aforementioned master-slave approach offers a low-complexity alternative for joint carrier recovery. It should be noted that these phase compensation methods have stringent path-matching

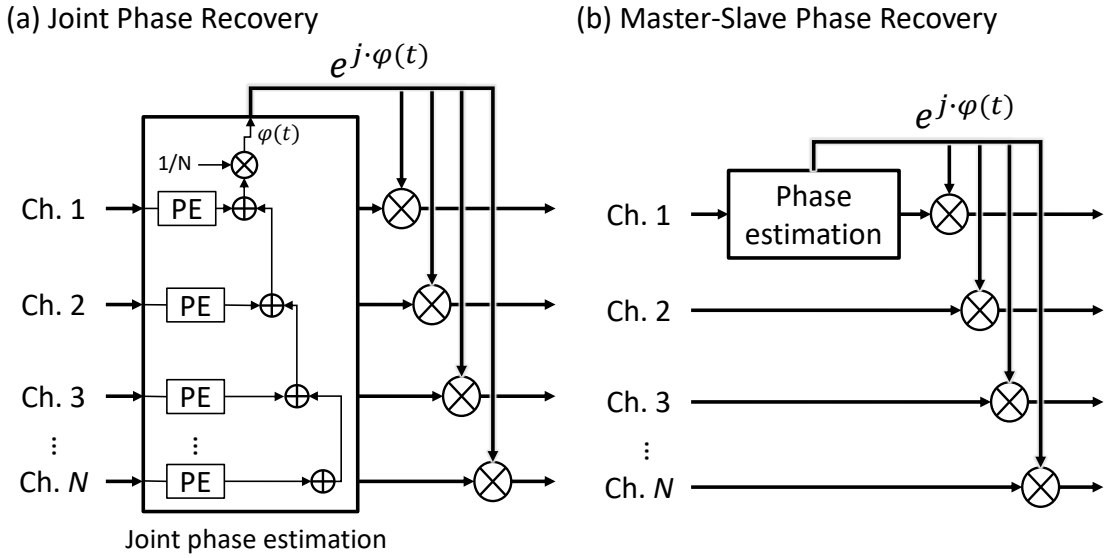

Figure 4.4: The schematic of the (a) Joint phase recovery, (b) Master-Slave phase recovery, PE: phase estimation.

requirements to eliminate the phase difference in the transceiver. Furthermore, the relatively small channel-dependent phase is challenging to compensate for before the cancellation of the phase noise of the laser. A feedback loop is needed in this condition to align the small phase fluctuation independently for each channel [146]. Either Blind or pilot-based algorithms described in Section 2.3.2 can perform carrier recovery. The experimental demonstration of record 12-bit/s/Hz C-band transmission using a comb-based superchannel [78] utilizes 1% digital pilot overhead to compensate for all impairments, where only 0.4% overhead is used for phase recovery. This highlights the remarkable performance of pilot-aided DSP and makes the master-slave phase recovery less attractive due to the relatively low phase pilot overhead being used.

4.3 Linear multi-channel equalization

The motivation for multi-channel equalization comes from the reduction of guard bands and maximization of the spectral efficiency as discussed in Section 4.1. Multi-channel equalization was first proposed in a simulated dual-polarization superchannel system considering the joint processing of three neighboring channels in the receiver [147]. The same approach was extended to an experimental superchannel consisting of two subchannels [148], which neglected the ICI from the other side. In

this experiment, free-running transmitter and LO lasers are used without frequency locking, which may induce frequency misalignment between the wavelength channels during joint processing and lead to suboptimum performance. The updated joint equalization was presented in [154] which utilizes a single equalizer to mitigate all impairments including polarization rotation, polarization mode dispersion, ISI, and ICI. The proposed scheme was experimentally investigated in a comb-based superchannel system and three-receiver structure to estimate the ICI cancellation performance when ENOBs and other transceiver distortion are dominant impairments. In paper C, besides the transmitter imperfection and electrical noise, we studied the influence of ASE noise originating from the inline EDFA in the noise-loading measurements as well as long-haul transmission experiments. In paper Q (not included in the thesis), we further alleviate the receiver complexity of the multichannel equalization by reducing the required number of receivers from 3 to 2.

Given a comb-based superchannel system with N channels, the signal of each channel is commonly shaped by a root-raised-cosine (RRC) filter in the transmitter DSP before the wavelength multiplexer. The theoretical filtered signal bandwidth, defined as the spectral width of the non-zero frequency components, is given by

$$B_{\text{ch}} = \frac{R_s}{2}(1+\beta) \tag{4.2}$$

where β is the roll-off factor of the RRC filter. The pulse-shaped signals are combined to form the superchannel and transmitted over the optical fiber. In the receiver, another RRC filter serves as the matched filter and removes the out-of-band noise and neighboring channels. When the channel spacing f_r of the frequency comb is selected such that ICI occurs, i.e. $f_r < B_{\text{ch}}$, the post-processed signal on l-th channel can be expressed as

$$r_l(t) = \sum_{i=1}^{N} \sum_{k=-\infty}^{\infty} s_{i,k} g_{i,l}(t-kT) e^{j(\Delta\Omega_i t + \phi_i)} + n(t) \tag{4.3}$$

where $s_{i,k}$ indicates the complex QAM signal of i-th channel on time index k, and $n(t)$ is the additive noise of the whole system. Specifically, $\Delta\Omega_i$ and ϕ_i is the frequency offset and phase noise of i-th carrier. It should be noted that $g_{i,l}(t)$ corresponds to the overall linear impulse response induced from channel i to channel l, where the detailed expression can be found in [147, Eq. 3]. Linear ICI is caused by the spectral overlap from the adjacent wavelength channels. Therefore, the Eq. (4.3) can be

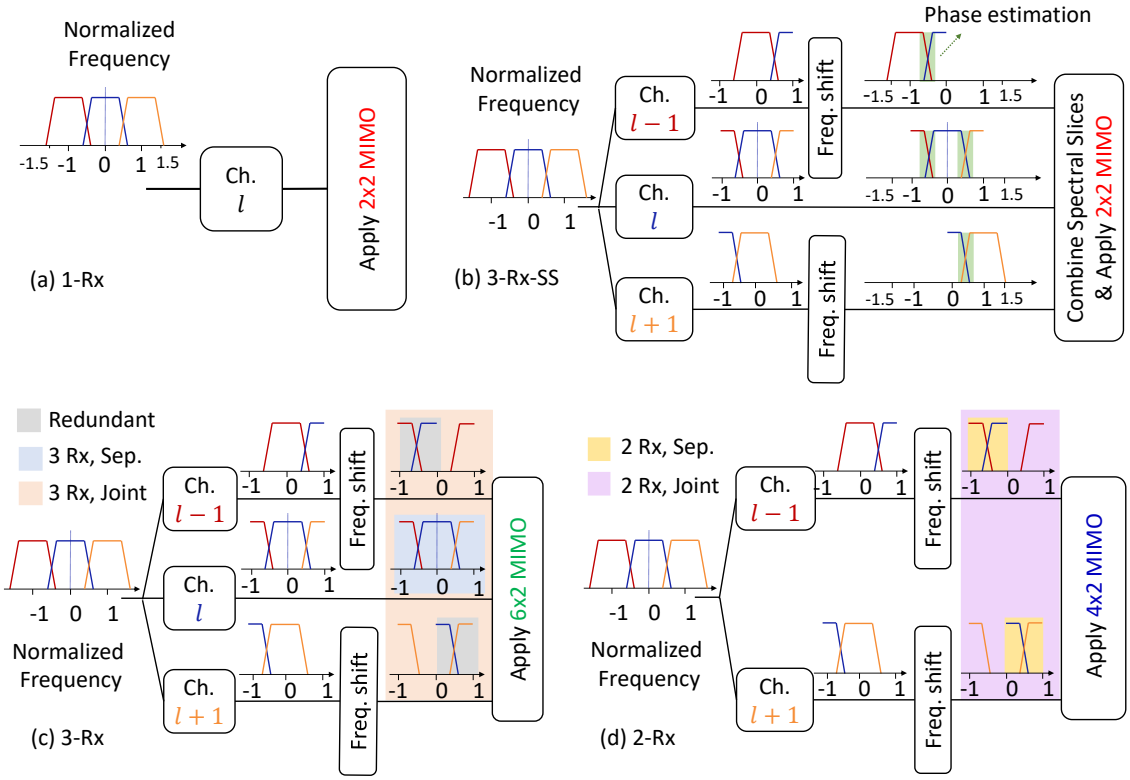

Figure 4.5: The schematic of the comb-based receiver for joint phase recovery (a) Comb regeneration, (b) Free running LO comb.

simplified by considering three channels $l-1, l, l+1$ and given by

$$\begin{aligned} r_l(t) = &\sum_{k=-m}^{m} s_{l,k} g_{l,l}(t-kT) e^{j(\Delta\Omega_l t + \phi_l)} + n(t) \\ &+ \sum_{k=-m}^{m} s_{l-1,k} g_{l-1,l}(t-kT) e^{j(\Delta\Omega_{l-1} t + \phi_{l-1})} \\ &+ \sum_{k=-m}^{m} s_{l+1,k} g_{l+1,l}(t-kT) e^{j(\Delta\Omega_{l+1} t + \phi_{l+1})} \end{aligned} \tag{4.4}$$

where m is the memory length of the channel response. From Eq. (4.4) we observe that the linear crosstalk is characterized by a linear convolution of the channel's impulse response and the transmitted signal. Theoretically, the ICI can be fully compensated by a MIMO equalizer. However, in practice, the frequency misalignment among the channels, the additive Gaussian noise, channel impairments, and the transceiver's imperfection may degrade the ICI cancellation.

Depending on the receiver structure, the receiver's bandwidth re-

quirements, and the computational complexity of the DSP several multi-channel equalization approaches can sufficiently mitigate the ICI. In the simplest scenario, illustrated in Fig. 4.5 (a), the three neighboring channels can be detected by a broadband coherent receiver, whose bandwidth $B_{\text{1-Rx}}$ needs to be larger than the three-channel signal bandwidth, i.e. $B_{\text{1-Rx}} > f_r + B_{\text{ch}}$. To simplify the analysis, the symbol rate is selected to be equivalent to the channel spacing. In the signal spectra, the frequency f_{norm} is given by $f_{\text{norm}} = f/R_s$ with f being the absolute frequency. After signal detection, a 2×2 MIMO equalizer operating at 4 samples per symbol (SPS) can cover the whole spectrum, i.e. all three channels with non-DC center caused by frequency offset, to efficiently perform the ICI cancellation.

Alternatively, the spectral slicing technique [155], which provides broadband detection by combining several individually measured spectral splices, can be used to detect the channels with three low-bandwidth receivers. The principle of the spectral-slicing receiver is depicted in Fig. 4.5 (b). The three channels are first separated into three independent receivers measuring at 2 SPS before being resampled to an oversampling ratio of 4. Importantly, for the side channels $l-1$ and $l+1$, the signals need to be shifted in the frequency domain such that they are aligned to the corresponding high-frequency location before the channel separation. The phase difference that exists between the individually measured channels can be estimated using the overlapping spectra (e.g., 3 GHz spectra were used in [155]). After correcting the phase mismatch, the measured signals are added to create the full-bandwidth signal. Finally, a 2×2 MIMO equalizer is applied to cancel linear crosstalk. Disregarding the noise disturbance and hardware imperfection, the three-receiver scheme enabled by manual spectral slicing should perform similarly to the single-receiver approach as the accessible spectral information is the same.

In contrast to the spectrally sliced receiver which manually stitches the channels together by subtracting the phase difference, a more advanced approach using a 6×2 equalizer superimposes the spectra and eliminates the ICI simultaneously in a dynamic fashion. This method operates at 2 SPS and requires a minimum bandwidth $B_{\text{3-Rx}}$ of around B_{ch} for each receiver, which is the lowest among all the included methods. It should be noted that, in contrast to the manual spectrally slicing, a similar amount of frequency shifting needs to be added to the signal of the side channel $l-1$ and $l+1$. Since the MIMO equalizer can only mit-

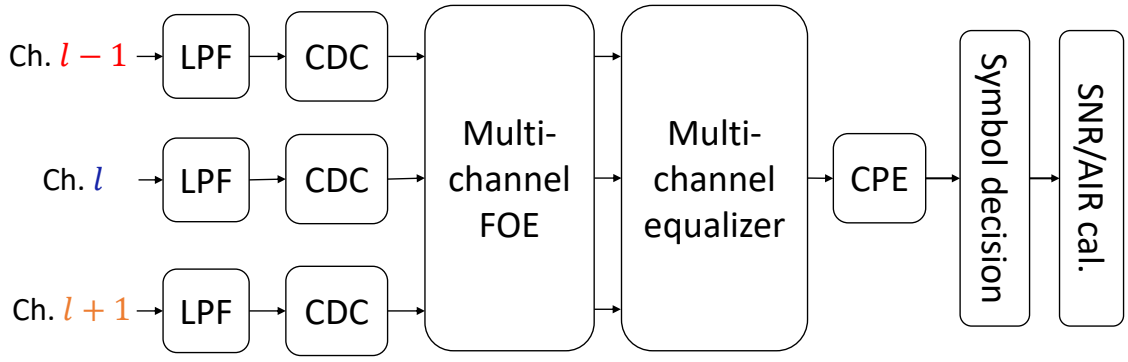

Figure 4.6: Joint digital signal processing diagram for superchannel receiver. LPF: lowpass filtering; CDC: chromatic dispersion compensation; FOE: frequency offset estimation; CPE: carrier phase estimation; SNR: signal-to-noise ratio; AIR: achievable information rate.

igate linear impairments, the frequency mismatch, which is caused by a nonlinear process, should be corrected before multi-channel equalization to avoid the performance penalty.

In the three-receiver scheme shown in Fig. 4.5 (c), the signal in the central channel is repeatedly detected outside the channel l. While half of the central channel spectra are received by the channel $l-1$, the other half are measured in channel $l+1$. This indicates that we can remove the central channel receiver without the loss of information. Therefore, we proposed a two-receiver structure by maintaining the side channel $l-1$ and $l+1$ for ICI cancellation. Different from the three-receiver scheme, this simplified scheme requires a receiver bandwidth $B_{\text{2-Rx}}$ of around f_r to fully recover the signal in the central channel since nearly half of the spectra is detected in each receiver. More importantly, data-aided algorithms need to be performed in the equalization because the target signal (i.e., the signal of the central channel) is strongly degraded. In contrast, blind multichannel equalizer suffers convergence issues, which perform worse than the independent processing that detects and equalizes the central channel itself.

The aforementioned receiver structures are compared in terms of the receiver bandwidth, MIMO DSP structure, number of SPS required for the ICI cancellation, and the equalizer taps number. The three-receiver scheme using 6×2 MIMO holds the minimum bandwidth requirement for the receiver, while only a small number of taps are needed. In contrast, the single-receiver structure requires a maximum receiver bandwidth and a relatively high 4-time oversampling to capture the full signal with fre-

Table 4.1: Comparison of various receiver structures for inter-channel interference cancellation, n-Rx:n-receiver structure, SS: spectral splicing, Eq.: equalizer.

	B_{Rx}	MIMO DSP	# of SPS	# of Eq. taps
1-Rx	$> f_r + B_{\mathrm{ch}}$	2×2	4	large
2-Rx	$> f_r$	4×2	2	small
3-Rx	$> B_{ch}$	6×2	2	small
3-Rx-SS	$> B_{ch}$	2×2	4	large

quency offset. Due to the increased SPS, a large tap number is needed to remove the ICI. Theoretically, the performance of the single-receiver, two-receiver, and three-receiver structures should have the same ICI mitigation performance regardless of the small noise difference included in the equalizer and the slightly different filtering effect. In a realistic coherent receiver, an optimum input power exists due to the trade-off between the noise and the nonlinearity of the TIA. For a larger signal's bandwidth, the power density of the input signal is reduced to avoid strong nonlinear distortion and preserve the highest SNR, resulting in a penalty to each subchannel. This effect can be explained by the system employing a single receiver to detect multiple subcarriers. When the number of jointly measured channels increases, the obtained SNR of each channel reduces [156]. Depending on the available hardware and computational resources as well as the required performance, one of the ICI cancellation approaches can be selected in the practical systems.

4.4 Multi-channel digital back-propagation

In an ideal condition, Multi-channel DBP needs to operate in the full field of the propagated signal to achieve the full compensation of intra-channel and inter-channel nonlinearity. It had been demonstrated numerically that every additionally included wavelength channel in the full-bandwidth DBP results in 0.1 dB extra gain when more than five channels are jointly processed [157]. In conventional WDM systems, each wavelength channel is individually detected by a separate coherent receiver and post-processed by independent DSP. It had been the-

oretically studied in [149] and experimentally verified in [150] that for multi-channel DBP one needs essentially frequency-locked channels to really unlock the gains. A comb-based transmitter and receiver structure can maintain fixed channel spacing among the wavelength channels and therefore are desired for implementing the multichannel DBP. For simplicity, similar to the linear crosstalk cancellation in Section 4.3, one single receiver can simultaneously detect several narrowband channels and perform joint nonlinear processing at the requirement of high-bandwidth optical hybrids and ADCs, which becomes challenging for high-capacity WDM systems with large symbol rate per channel. To alleviate the hardware limitation, spectral splicing can perform digital stitching of individually measured channels to yield wideband detection using several low-bandwidth receivers. Using the spectrally spliced receiver, the multi-channel DBP was experimentally demonstrated in [158] for a 1.2 Tb/s five-channel superchannel system over 176 GHz bandwidth.

Many practical issues including the number of steps per span in the SSM, the sampling rate of the receiver, and polarization mode dispersion (PMD) will impact the performance of multichannel DBP [159,160]. Given a small number of steps per span, the single-channel DBP may outperform the multi-channel one due to the performance penalty caused by the imprecise inversion of the nonlinear propagation. For DBP that operates within a large bandwidth, a significantly larger number of steps per span are required to maintain the accuracy of the nonlinear compensation [159]. Following the Nyquist–Shannon sampling theorem [161], the sampled signal can be reconstructed without aliasing when more than two-time oversampling is performed. In addition, the Kerr nonlinearity induced spectral broadening requires a higher sampling rate to avoid aliasing and eliminate performance degradation, especially when more channels are jointly processed. For a limited sampling rate, the backpropagated bandwidth should be optimized to achieve the best nonlinear mitigation. Finally, the channel fluctuations, especially in combination with PMD cause issues, significantly degrade the multichannel DBP performance due to an inaccurate inversion of the nonlinearity when more channels are included. Specifically, a PMD parameter as high as 1 $\mathrm{ps}/\sqrt{\mathrm{km}}$ results in a negligible gain for multi-channel DBP compared to the single-channel counterpart [159]. It should be noted that all the strong random effects in the fiber limits the performance of the DBP. The high computational complexity of the single- or multi-channel DBP hinders its application in practical systems, whereas low-complexity implementa-

tion of DBP should be investigated for high-capacity systems [162, 163].

4.5 Perturbation-based nonlinear compensation

Perturbation theory is a common approach for the analysis of nonlinearity in the nonlinear Schödinger equation (for single-polarized signals) as well as the coupled Manakov equation (for dual-polarized signals). For simplicity, it treats the nonlinear interference as a linear additive perturbation ΔU. Given the solution U_0 in the linear channel, the received signal is given by $U = U_0 + \Delta U$, where the nonlinear perturbation in a dual-polarization discrete-time system at time T_k can be expressed as [164]

$$\Delta U_k^x = P_0^{3/2} \sum_{m,n} C_{m,n} s_{k+m}^x (s_{k+n}^x \overline{s}_{k+l}^x + s_{k+n}^y \overline{u}_{k+l}^y) \tag{4.5}$$

$$\Delta U_k^y = P_0^{3/2} \sum_{m,n} C_{m,n} s_{k+m}^y (s_{k+n}^y \overline{s}_{k+l}^y + s_{k+n}^x \overline{u}_{k+l}^x) \tag{4.6}$$

where $l = m + n$, k is the time index, and the overbar indicates the complex conjugate operation. The nonlinear distortion is determined by the signal's peak power P_0 at the fiber input, the transmitted signal s, and most importantly the nonlinear coefficients $C_{m,n}$, where the analytical solution of $C_{m,n}$ was given in [164, Eq. 11a, 11b and 11c]. In superchannel systems where complex interchannel interference dominates, it is challenging to obtain the analytical expression of the nonlinear perturbation. Following the perturbation theory, the detected signal r in the channel of interest c can be considered as a sum of the transmitted signal and nonlinear terms from interference channels [165]

$$r_k^{x,c} = s_k^{x,c} + \Delta U_k^{x,c} \tag{4.7}$$

$$\begin{aligned} \Delta U_k^{x,c} = & \sum_{m,n} C_{m,n}^{x,c} s_{k+m}^{x,c} (s_{k+n}^{x,c} \overline{s}_{k+m+n}^{x,c} + s_{k+n}^{y,c} \overline{s}_{k+m+n}^{y,c}) \\ & + \sum_{l \neq c} \sum_{m,n} C_{m,n}^{x,l} (2 s_{k+m}^{x,c} s_{k+n}^{x,l} \overline{s}_{k+m+n}^{x,l} \\ & + s_{k+m}^{x,c} s_{k+n}^{y,l} \overline{s}_{k+m+n}^{y,l} + s_{k+m}^{y,c} s_{k+n}^{x,l} \overline{s}_{k+m+n}^{y,l}). \end{aligned} \tag{4.8}$$

A similar nonlinear perturbation for y-polarization $\Delta U_k^{y,c}$ is given by switching the polarization x and y in Eq. 4.7 and Eq.4.8. Specifically, $C_{m,n}^{x/y,c}$ corresponds to the intra-channel nonlinear crosstalk caused by

x/y polarization, whereas $C_{m,n}^{x/y,l}$ denotes the inter-channel nonlinearity such as XPM or four-wave-mixing effects.

To determine the total nonlinearity and facilitate the nonlinear post-equalization, the coefficients $C_{m,n}^{x/y,c/l}$ need to be obtained first. Considering the multiplication of the three signals as the input and assisting with the training sequence, the coefficients $C_{m,n}^{x/y,c/l}$ can be learned by linear regression. Provided the joint detection of the superchannel as well as an estimation of the transmitted signal s based on the noisy and distorted signal r, the complicated nonlinear term ΔU is obtained according to Eq. (4.8). Therefore, the equalized signal is given by $z = r_k^{x,c} - \Delta U_k^{x,c}$. In contrast to DBP, the perturbation nonlinear compensation does not require the Fourier transform operation, leading to much lower computational complexity. This approach requires an approximation of the transmitted signal. For large nonlinear perturbation that induces a high error rate, the estimation error causes a significant performance penalty for the nonlinear compensation.

Similar to the DBP, the inverse perturbation theory can be applied to the received signal for nonlinearity mitigation. For the single-carrier channel with dual polarization, the inverse Manakov equation can be used to equalize the signal. Then the recovered signal is given by [166]

$$z_k^x = C^x r_k^x + \sum_{m,n} \widetilde{C}_{m,n}^x r_{k+m}^x (r_{k+n}^x \overline{r}_{k+l}^x + r_{k+n}^y \overline{r}_{k+l}^y) \tag{4.9}$$

$$z_k^y = C^y r_k^y + \sum_{m,n} \widetilde{C}_{m,n}^y r_{k+m}^y (r_{k+n}^y \overline{r}_{k+l}^y + r_{k+n}^x \overline{r}_{k+l}^x). \tag{4.10}$$

For superchannel systems which suffer intra-channel and inter-channel nonlinearity, the nonlinear perturbation from interference channels needs to be taken into account. According to the perturbative analysis in Eq. 4.7 and Eq. 4.8, the equalized x-polarization signal can be extended to [167]

$$z_k^{x,c} = C r_k^{x,c} + \Delta r_k^{x,c} \tag{4.11}$$

$$\begin{aligned} \Delta r_k^{x,c} =& \sum_{m,n} \widetilde{C}_{m,n}^{x,c} r_{k+m}^{x,c} (r_{k+n}^{x,c} \overline{r}_{k+m+n}^{x,c} + r_{k+n}^{y,c} \overline{r}_{k+m+n}^{y,c}) \\ &+ \sum_{l \neq c} \sum_{m,n} \widetilde{C}_{m,n}^{x,l} (2 r_{k+m}^{x,c} r_{k+n}^{x,l} \overline{r}_{k+m+n}^{x,l} \\ &+ r_{k+m}^{x,c} r_{k+n}^{y,l} \overline{r}_{k+m+n}^{y,l} + r_{k+m}^{y,c} r_{k+n}^{x,l} \overline{r}_{k+m+n}^{y,l}). \end{aligned} \tag{4.12}$$

Similarly, the equalized y-polarization signal is calculated by switching the x and y polarization in these expressions.

Importantly, the nonlinear coefficient $\widetilde{C}_{m,n}^{x,l}$ caused by the interfering wavelength channels is sensitive to the relative polarization state of each channel. As the signals among the subchannels are not co-polarized in practical systems, the random polarization rotation induces time-varying nonlinear interference for the superchannel. In paper D, we employed the perturbation-enabled nonlinear compensation in the comb-based superchannel and joint-detection system. Both SPM and XPM were compensated in the three-channel systems.

Chapter 5

Full-dimensional Optical Transmission System

Today, state-of-the-art coherent systems utilize all the available physical dimensions including time, amplitude, phase, polarization, wavelength, and even space to maximize the capacity [168]. However, independent modulation and receiver digital signal processing (DSP) in each dimension prevent the system from exploiting the maximum system performance. In contrast, joint encoding and DSP are required to achieve the full-dimensional capacity.

This chapter describes the implementation of multi-dimensional formats in an experimental optical communication system and the main challenges. Various practical issues such as limited physical impairments, signal modulation, and receiver-side signal processing are discussed.

5.1 Multidimensional modulation formats

Conventional coherent systems utilize 2-dimensional (2D) modulation formats such as quadrature amplitude modulation (QAM) which are independently modulated to each polarization component. In contrast to the squared-shape constellations, constellation shaping provides means to close the 1.53 dB gap to the Shannon limit by generating constellations with Gaussian distribution [169]. This can be achieved by changing either the probability distribution or geometric location of the constellation points, commonly referred to as probabilistic shaping (PS) [170,171] and geometric shaping (GS) [172], respectively. The former approach modifies the constellation's probability and provides efficient rate-adaptive

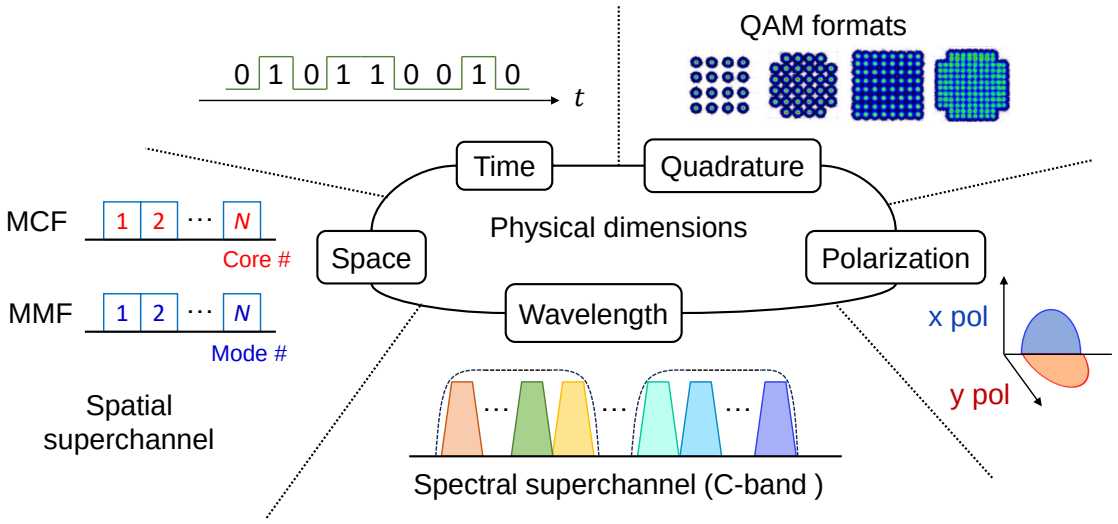

Figure 5.1: The schematic of the full-dimensional fiber-optic communication system. MCF: multi-core fiber; MMF: multi-mode fiber.

means for long-haul transmission without switching the modulation order and forward error correction code rate [173], thus simplifying the structure for rate-varying applications. In addition, it is fully compatible with typical mapping/de-mapping algorithms and DSP, since the geometric positions of the constellation points are not changed. However, a distribution matcher is needed to convert the uniformly distributed data bits to non-uniformly distributed symbols [174], which significantly increases the computational complexity and adds delay to the system, making it unsuitable for low-latency applications. In general, the GS pushes more constellation points to the low-energy regime for achieving the Gaussian signaling. 2D-GS constellations with a circular geometric shape can provide high shaping gain for large-cardinality constellations [175]. This is because the large constellation set enables better emulation of Gaussian signaling, provided that the geometric distribution is fully optimized. The GS signal, normally with an irregular constellation shape, requires the implementation of properly designed labeling/de-mapping such as end-to-end learning [176], which needs to be re-optimized for constellations with different cardinality. More importantly, high-cardinality GS signals suffer from transceiver impairments including quantization noise and the transmitter's nonlinearity, leading to a reduced shaping gain.

The multidimensional (MD) space provides a greater degree of freedom to optimize the constellation and therefore allows the generation of more power-efficient GS constellations compared to 2D constellations. Given an n-dimensional constellation $\mathbf{C} = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, ..., \mathbf{c}_M\}$ with M

points, the average symbol energy is given by

$$E_s = \sum_{k=1}^{M} P_{\mathbf{C}}(\mathbf{c}_k)||\mathbf{c}_k||^2 \tag{5.1}$$

where $P_{\mathbf{C}}(\cdot)$ is the probability mass function of the constellation. The minimum Euclidean distance among the constellation points is calculated as

$$d_{\min} = \min_{l \neq k} d_{k,l} \tag{5.2}$$

with $d_{k,l}$ being the Euclidean distance between the constellation points $\mathbf{c}_k$ and $\mathbf{c}_l$. Under the well-known "union-bound" condition, the decision error for every unique symbol pair (e.g., $\mathbf{c}_k$ and $\mathbf{c}_l$) is independently estimated and summed to obtain the total symbol error rate (SER), simplifying the performance estimation and forming an upper bound of the SER. Assuming a uniform probability distribution for the constellation, the upper limit of the SER of the MD constellation $\mathbf{C}$ in the AWGN channel is given by [29]

$$SER \leq \frac{1}{M} \sum_{k=1}^{M} \sum_{l=1, l \neq k}^{M} \frac{1}{2} \mathrm{erfc}\left(\frac{d_{k,l}}{2\sqrt{N_0}}\right) \tag{5.3}$$

where $\mathrm{erfc}(\cdot)$ is the complementary error function and $N_0/2$ is the noise variance for each dimension. Considering the static nature of Gaussian noise, the constellation set with minimum Euclidean distance will have the highest contribution to the SER. Therefore, provided the same minimum Euclidean distance, we need to optimize the geometric shape of the constellation such that the total power is minimized. This optimization is similar to solving the N-dimensional sphere packing problem. In this case, the most noise-tolerant MD format is generated. Various approaches can be used for generating MD formats, including set-partitioning [177], block coding [169,178], lattices' spherical cutting [179], switching in different components of each dimension such as polarization [180], frequency [181], and time [182], or relying on the biorthogonal plane [183]. Four-dimensional (4D) modulation formats have attracted significant attention in coherent systems where the data can be jointly modulated to two quadratures and two polarizations. Among the widely studied 4D constellations, the polarization-switched quadrature phase shift keying (PS-QPSK) with 1.5 bits/2-dimension SE was proved to be

the most energy-efficient format [180]. Later on, high-dimensional (more than 4D) formats were studied in [179] providing the SE of < 4 bits/2-dimension, showing a better energy efficiency than PS-QPSK. Nevertheless, the generation of these high-SE MD formats usually requires look-up tables to store the constellation points. In addition, even for moderate SEs, complex iterative search algorithms need to be applied to optimize the bit-labeling of the constellations. Even if the densely packed MD formats provide better SER performance than conventional 2D formats due to increased minimum Euclidean distance, the selection of unoptimized bit-labeling will significantly degrade the uncoded bit error rate (BER), especially in the low-SNR regimes. For the high-dimensional formats, proper optimization of labeling enables shaping gain at the uncoded BER of 2.7×10^{-2} (i.e, threshold of the 20% overhead SD-FEC required for metro and long-haul applications) for SE of 1-1.5 bits/2D-symbol, whereas the improvement is limited to the uncoded BER range between 1×10^{-3} and 1×10^{-2} for SE of 1.75-2 bits/2D-symbol [179]. This may be explained by the lack of joint design of MD constellations and FEC coding. More importantly, the complexity of the decoding algorithm grows significantly with increased SE, in which large-cardinality constellations are generated, making the MD formats infeasible for high-SE applications.

Voronoi constellations (VCs) [184,185] generated by a structured GS approach are of interest since no LUT is required to store the whole constellation set. In addition, the corresponding encoding [184] and decoding algorithms [186, 187] require relatively low complexity even at high SE and thus facilitate practical implementation of high-cardinality VCs using reasonable computational resources. Recently, 32-dimensional (32D) VCs with 2^{32} (i.e., 2-bit/symbol/dimension-pair) and 2^{64} (i.e., 4-bit/symbol/dimension-pair) constellation points were simulated in the AWGN channel and verified in experiments [188, 189], outperforming QAM in terms of uncoded BER between 10^{-3} and 10^{-2}. Later, a record cardinality of 2^{96} (i.e., 7.9×10^{28}) was numerically demonstrated with 24-dimensional (24D) VCs with a high SE of 8 bits/s/dimensioan-pair [190].

Similar to other MD formats, the aforementioned VCs have worse performance than the conventional formats (e.g., QAM) at BER threshold for 20% OH SD-FEC (i.e., 2.7×10^{-2}). In contrast, VCs with better mutual information (MI) performance than QAM were presented in the AWGN channel [191, 192], indicating an ideal shaping gain of VCs provided a perfect coded modulation. The achievable information rate

(AIR) indicates the amount of information per second that can be reliably transmitted over the channel for a given coded system. In the commonly used bit interleaved coded modulation (BICM), the upper limit of AIR is referred to as the generalized mutual information (GMI). Whereas the corresponding upper bound for the multi-level coding (MLC) scheme is MI. When the coded system is sub-optimum, a capacity gap exists between the GMI and MI, explaining the worse VC performance compared to QAM in terms of GMI presented in [191,192]. To remove this penalty, a SD MLC scheme employing DVB-S2 standard LDCP codes was proposed and applied to high-cardinality VCs, showing a high shaping gain of 0.4 dB for 8-dimensional (8D) VCs and 0.8 dB for 24D VCs [190], respectively, in simulations of the AWGN channel. The joint design of VCs and low-complexity FEC coding makes this scheme attractive for high-SE state-of-the-art fiber networks.

Energy-efficient 8D and 24D Voronoi constellations with a SE of 8 bit/symbol/dimension-pair are illustrated in Fig. 5.2, where the signals are projected onto the 2D plane to ease comparison to conventional QAM and presented in the 1D axis to highlight its probability distribution in each dimension. Aside from the quadrature components, the residual dimensions are spread over 2 time slots, indicated by markers with different colors in Fig. 5.2, for 8D VC and 6 time slots for 24D VC, respectively. As a GS format, the VCs are designed to approximate the Gaussian distributed constellations to reach the Shannon limit. From the 1D projection, we can observe that the probability distribution of the 24D VC is closer to the Gaussian shape, explaining the additional 0.4 dB shaping gain. For VCs with higher dimensions, the inherent greater degree of freedom makes the approximation of the Gaussian distribution easier and therefore yields larger improvements over QAM.

5.2 Physical channel impairments

When we consider modulating data on all available dimensions, including the time, quadrature, polarization, wavelength, and space components, high-SE MD formats suffer from system impairments such as carrier offset, transceiver impairments, and fiber nonlinearity. This was investigated in [188, 189], where the 32D VC with a low SE of 2 and 4 bit/2D-symbol was modulated over the wavelength, polarization, time, and quadrature components of the electromagnetic field [188, 189]. The impact of the channel impairments was investigated using various phys-

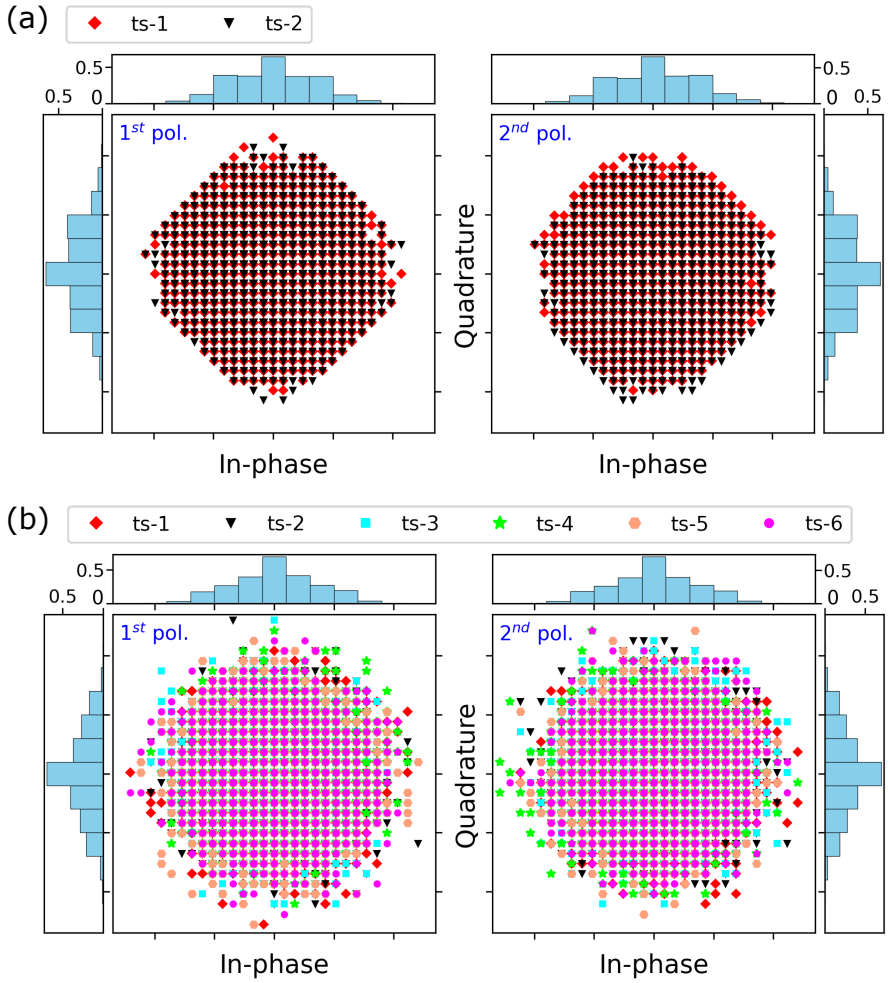

Figure 5.2: The 1D and 2D projection of the (a) 8D and (b) 24D Voronoi constellation.

ical realizations. For example, the nonlinear distortion originating from the fiber Kerr effect can be studied by placing more signal dimensions in the wavelength components, where the resulting interchannel interference will generate an extra performance penalty. From [189] we observed that the highest reach was achieved by using the time slots as much as possible. This is because the dominant impairments in the time dimension, i.e. intersymbol interference (ISI), can be efficiently canceled by a static and dynamic equalizer. In contrast, the polarization-dependent distortion caused by the imperfection of the transceiver and nonlinear fiber impairments will significantly degrade the shaping gain of VCs, unless applying perfect pre-distortion/post-equalization and nonlinear compensation algorithms (e.g., multi-channel digital backpropagation, and perturbation-based compensation) to fully remove the distortion. For

large-symbol-rate systems and high-SE modulation formats, the transmitter's nonlinearity and the strong filtering of the electronic devices generate linear and nonlinear ISI, respectively, causing degradation to signal in different time slots. More importantly, the ENOB of the DACs and ADCs results in low-resolution quantization of the large constellations, in particular for large symbol rates as the ENOB drops quickly with the increasing frequency. Furthermore, the compensation of frequency offset and phase noise becomes increasingly important for high-SE MD constellations with large cardinality. Therefore, even if the jointly encoded MD constellations provide significant shaping gain over conventional QAM formats, it should be emphasized that advanced DSP including digital pre-distortion and receiver-side digital processing is required to alleviate the system impairments and preserve the shaping gain of MD formats.

5.3 Joint encoding and digital signal processing

For the above-mentioned discussion, we realize that advanced DSP should be implemented together with the MD formats to yield the optimum shaping gain in the presence of system impairments. In addition, as previously discussed in Section 4, joint digital processing in the receiver assisted with comb-based superchannel provides a better compensation for the carrier offset, linear crosstalk, and fiber Kerr nonlinearity, unblocking the full potential of the advanced DSP to achieve higher capacity. Importantly, the joint modulation with MD formats and multi-channel receiver-side DSP form multidimensional superchannel systems, as depicted in Fig. 5.3, maximizing the capacity of the fiber-optic communication system. An optical frequency comb (OFC) is used to provide phase-locked evenly spaced lines as the optical carrier, which enables the generation of a tightly packed superchannel minimizing the guard band thus increasing the SE. For the joint digital signal modulation, the information bits are first fed to a MD coding block for joint FEC coding and MD signal mapping. Before the digital-to-analog conversion in the DAC, the created constellation is upsampled, and pulse-shaped in the transmitter-side DSP. Linear and nonlinear pre-distortion can be added to alleviate the transmitter imperfection. The electrical signals drive the corresponding dual-polarization IQ modulators for joint optical modulation across the polarization and wavelength channels, which are finally combined to generate the spectral superchannel. During signal detection, several coherent receivers are applied to detect all the subchannels

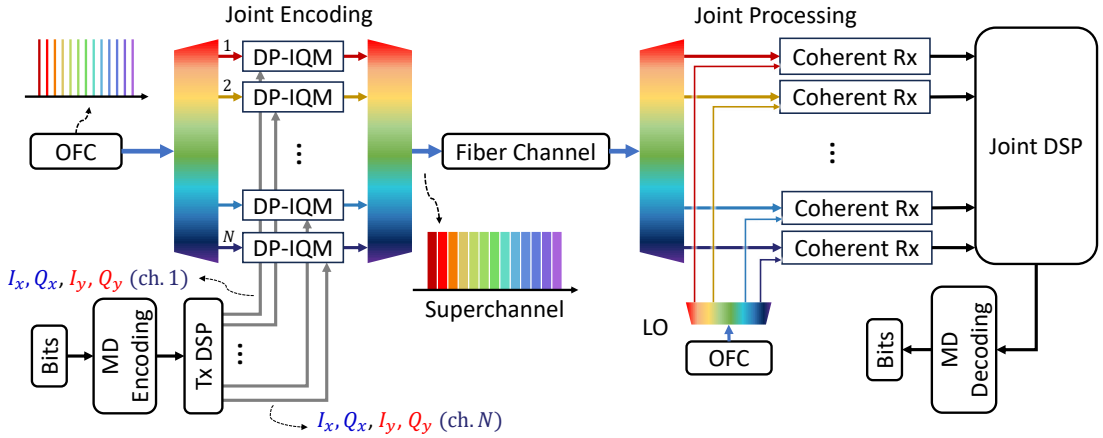

Figure 5.3: The schematic of joint encoding and digital signal processing for multidimensional superchannel system.

to facilitate the joint DSP. Importantly, OFC-based local oscillators are needed to maintain the frequency among the detected wavelength channels, enabling the efficient implementation of joint processing. The post-processed signals are finally decoded and demapped into bits. It should be noted that such schemes can be easily extended to space division multiplexing, e.g. multi-core or multi-mode fiber, system by adding an extra spectral-superchannel transmitter for each core or mode to form hybrid superchannels including the wavelength and space dimensions to further scale up the capacity.

In paper E, we experimentally present a high-gain high-SE 24D VC in the coded modulation system spanning the C-band. Receiver-side joint processing is not included due to the limited hardware resources. A soft-decision multilevel coded scheme is used to preserve the shaping gain of VCs. Here, we have the first experimental demonstration of high-SE MD modulation formats with a high shaping gain in the soft-decision coded systems, showing its potential in high-capacity long-haul transmission applications.

Chapter 6

Future outlook

Many future research topics have attracted significant attention in coherent communication systems. First of all, higher-symbol-rate coherent transponders are of interest due to the ever-increasing demand for data traffic. The current state-of-the-art >100 GHz transceivers enable the generation of broadband signals (>170 Gbaud) and thus yield more than 2 Tb/s line rate per wavelength [193]. It would be interesting to develop higher-bandwidth coherent transceivers and optical equalization/filtering to facilitate multi-terabit applications. Another important research area is the design of a low-V_π modulator, which relaxes the high-gain requirement of electrical amplification and the strong nonlinear distortion from the RF drivers. Furthermore, a relatively low V_π of <1 V can enable efficient optical modulation without electrical amplifiers, which improves the signal-to-noise ratio and reduces power consumption. Candidates for high-bandwidth sub-volt-V_π modulators are silicon-organic hybrid [194] modulators and thin-film lithium niobate [195, 196] modulators. Besides the high capacity per wavelength channel, it is of interest to investigate multiband transmission that utilizes the spectra outside the conventional C-band for data transmission. One key component in multiband coherent communication is the optical amplifier. While erbium-doped fiber amplifiers are primarily limited to the C- and L-bands, rare–earth doped fiber amplifiers [197] or Raman amplifiers [198] open up the transmission in other wavelength bands. The development of a low-noise-figure amplification technique for multiband systems is necessary as it can significantly improve transmission performance. Other important components required for multiband transmission are the high-performance multiband transceivers and high-efficiency wavelength converters.

In the following section, we review several future research areas connected to transmitter- and receiver-side digital signal processing in multidimensional transmission systems.

Time-domain Volterra-based nonlinear post-equalizer

The Volterra equalizer is commonly used to perform efficient nonlinear digital pre-distortion, which can significantly improve the transmitter performance, particularly for large-cardinality constellations required for high-spectral-efficiency coherent systems. Similar to the pre-compensation, the Volterra-based post-equalizer has been experimentally presented in the receiver to mitigate the transmitter's nonlinearity. In Paper V, we use half of the transmitted sequence for training the coefficients required in perturbative nonlinear compensation and apply the other half to estimate the performance due to the included polarization-dependent cross-phase modulation (XPM) mitigation. A long training sequence is required in the linear regression approach used for learning perturbative coefficients to preserve accurate estimation, which leads to a 50% overhead in our experiments. To eliminate this high information loss without sacrificing performance, a Volterra-based multichannel post-equalizer can replace the perturbative methods to mitigate the self-phase modulation and XPM. Similar to the single-channel Volterra equalizer, the training of the nonlinear post-equalizer can be performed by least mean squares or recursive least squares algorithms, requiring a significantly smaller overhead and thus increasing the capacity.

Multidimensional Voronoi constellations in space division multiplexing system

In paper V, we employ a 24-dimensional (24D) Voronoi constellation (VC) in the spectral superchannel system and achieve high spectral efficiency spanning the C-band. The 24D VC is modulated over amplitude, phase, two polarizations, and six time slots. To exploit the full-dimensional capacity achieved by considering all available physical dimensions, it will be interesting to include the space dimension in the multidimensional (MD) modulation and investigate the performance of MD VC formats in the multi-more or multi-core fiber.

Joint modulation and joint signal proccesisng

The MD constellations provide shaping gain over conventional quadrature amplitude modulation formats due to the increased minimum Euclidean distance, which eliminates the capacity gap to the Shannon limit. Joint processing such as joint carrier recovery, and multichannel linear/nonlinear equalization can further boost the system capacity by canceling various impairments. There is no demonstration of jointly modulating the transmitted signals and jointly detecting them in the receiver to obtain the highest achievable capacity given the available resources. Therefore, it will be interesting to apply MD formats to the superchannel systems with a joint-processing receiver structure.

Chapter 7

Summary of papers

Paper A

Periodicity-Enabled Size Reduction of Symbol Based Predistortion for High-Order QAM, *Journal of Lightwave Technology*, vol. 40, no.18, pp. 6168-6178, 2022.

A look-up table (LUT) is a low-complexity and popular digital predistortion (DPD) approach to mitigate the transmitter's nonlinearity with memory. In this paper we experimentally demonstrate novel reduced-size LUTs based on the periodicity property and supervised learning neural networks (NNs) in the coherent systems, while the output of the LUT-based DPD is used for training the NN. We investigate the size-reduction efficiency and bit error rate performance for various single-polarization high-order quadrature amplitude modulation (QAM) constellations including 64-QAM, 256-QAM, and 1024-QAM in the optical back-to-back transmission. The results show that the proposed LUTs and NNs achieve similar performance as the full-size LUTs and enable significant size reduction. Among these methods, NNs perform the best in terms of size-reduction efficiency and performance improvement.

My contributions: I developed the concept of size reduction and wrote the pre-distortion DSP. I built the experimental setup and performed all the measurements. I wrote the paper with assistance from co-authors.

Paper B

Over-the-fiber Digital Predistortion using Reinforcement Learning, *Proceedings of European Conference on Optical Communications (ECOC)*, 2021.

In this paper we propose novel reinforcement learning neural networks (RL-NN) for mitigating the coherent transmitter's nonlinearity. Compared to the NNs learned in the direct learning architecture, the proposed RL-NN is directly trained in the experimental systems, which avoids complex and possibly inaccurate channel modeling. We compare the performance of the RL-NN with the memoryless arcsine-based predistortion in the optical back-to-back configuration with self-homodyne detection. The results show that the proposed RL-NN outperforms the arcsine-based compensation with a bit error rate reduction as high as 60%.

My contributions: I designed the experimental setup and the DSP except for neural networks. I performed all the measurements and helped in writing the paper.

Paper C

Inter-Channel Interference Cancellation for Long-Haul Superchannel System, *Journal of Lightwave Technology*, vol. 42, no. 1, pp.48-56, 2024.

A frequency comb is a multi-carrier optical source with frequency and phase coherence among the comb lines, which is suitable for superchannel systems requiring densely packed wavelength channels. Multichannel equalization (MCE) is an efficient digital approach to cancel interchannel interference and improve spectral efficiency. In this work we experimentally demonstrate the MCE enabled by a three-receiver detection architecture in the comb-based superchannel systems. Importantly, we investigate the influence of the amplified spontaneous emission (ASE) noise in the noise loading measurements and long-haul transmission experiments. The results show that the ASE originating from in-line amplification significantly degrades the MCE performance, particularly for high-order constellations, in long-haul

systems. We highlight that an optimization of the equalizer tap number is needed for the MCE to eliminate the noise coupling from the neighboring wavelength channel and reduce the performance penalty.

My contributions: I built the experimental setup, designed the multi-channel equalization DSP, and performed all the measurements. I wrote the paper.

Paper D

Perturbation based Joint SPM and XPM Compensation for Superchannel Systems, submitted to *Photonics Technology Letters*, Apr. 2024

In this paper we demonstrate a perturbation-based nonlinear compensation (PB-NLC) approach to mitigate the fiber Kerr nonlinearity in comb-based superchannel systems. After long-haul transmission, the superchannel consisting of three wavelength channels is detected by three independent coherent receivers with comb-based local oscillators. Both the self-phase modulation from the channel of interest and the cross-phase modulation caused by the interfering channel are eliminated in the proposed PB-NLC.

My contributions: I developed the nonlinear compensation DSP and carried out the simulation. I built the experimental setup and performed all the transmission measurements assisted by co-authors. I wrote the paper.

Paper E

12.2 bits/s/Hz C-band Transmission with High-Gain Low-Complexity 24-Dimensional Geometric Shaping, *Journal of Lightwave Technology*, accepted, Mar. 2024.

Voronoi constellations (VCs) as a structured multi-dimensional (MD) geometric shaping approach have low-complexity encoding and decoding algorithms. We experimentally present a high-spectral-efficiency (high-SE) 24-dimensional VC with a record constellation size of 7.9×10^{28}

in the C-band transmission systems. With the properly designed soft-decision (SD) multilevel coding scheme, we have the first experimental demonstration of high-SE MD constellations with a high shaping gain in the SD coded modulation systems, highlighting the potential of MD VCs for high-capacity applications.

My contributions: I built the experimental setup and conducted all the measurements. I processed the data and wrote the paper with assistance from co-authors.

References

[1] A. Einstein, "Zur quantentheorie der strahlung," *Physikalische Zeitschrift*, vol. 18, pp. 121–128, 1917.

[2] J. Hecht, "Short history of laser development," *Optical engineering*, vol. 49, no. 9, p. 091002, 2010.

[3] J. P. Gordon, H. J. Zeiger, and C. H. Townes, "The maser—New type of microwave amplifier, frequency standard, and spectrometer," *Physical Review*, vol. 99, pp. 1264–1274, 1955.

[4] A. L. Schawlow and C. H. Townes, "Infrared and optical masers," *Physical Review*, vol. 112, pp. 1940–1949, 1958.

[5] T. H. Maiman, "Stimulated optical radiation in ruby," *Nature*, vol. 187, pp. 493–494, 1960.

[6] A. Javan, W. R. Bennett, and D. R. Herriott, "Population inversion and continuous optical maser oscillation in a gas discharge containing a He-Ne mixture," *Physical Review Letters*, vol. 6, pp. 106–110, 1961.

[7] R. N. Hall, G. E. Fenner, J. D. Kingsley, T. J. Soltys, and R. O. Carlson, "Coherent light emission from GaAs junctions," *Physical Review Letters*, vol. 9, pp. 366–368, 1962.

[8] J. N. Shive, "Photoresistive translating device," Patent, US2560606A, 1949.

[9] F. P. Kapron, D. B. Keck, and R. D. Maurer, "Radiation losses in glass optical waveguides," *Applied Physics Letters*, vol. 17, no. 10, pp. 423–425, 1970.

[10] R. Mears, L. Reekie, I. Jauncey, and D. Payne, "Low-noise erbium-doped fibre amplifier operating at 1.54μm," *Electronics Letters*, vol. 23, pp. 1026–1028, 1987.

[11] R. Csencsits, P. J. Lemaire, W. A. Reed, D. S. Shenk, and K. L. Walker, "Fabrication of low-loss single-mode fibers," in *Proceedings of Optical Fiber Communication Conference*, 1984.

[12] H. Yokota, H. Kanamori, Y. Ishiguro, G. Tanaka, S. Tanaka, H. Takada, M. Watanabe, S. Suzuki, K. Yano, M. Hoshikawa, and H. Shimba, "Ultra-low-loss pure-silica-core single-mode fiber and transmission experiment," in *Proceedings of Optical Fiber Communication Conference*, 1986.

[13] X. Zhou and J. Yu, "Multi-level, multi-dimensional coding for high-speed and high-spectral-efficiency optical transmission," *Journal of Lightwave Technology*, vol. 27, no. 16, pp. 3641–3653, 2009.

[14] "Nokia PSE-2 super coherent technology," 2016. [Online]. Available: https://networks.nokia.com/products/pse-2-super-coherent-technology.

[15] "Infinera 800G live trial," 2020. [Online]. Available: https://www.infinera.com/ press-release/infinera-breaks-industry-record-with-800g-transmission-over-950-kilometers-in-a-live-netwo rk-trial.

[16] "Huawei 800G optical module," 2020. [Online]. Available: https://www.huawei.com/ en/news/2020/2/800g-tunable-ultra-hi gh-speed-optical-module.

[17] "Ciena wavelogic 6 extreme," 2023. [Online]. Available: https://www.ciena.com/insights/blog/2023/introducing-wavelogi c-6.

[18] A. Napoli, P. W. Berenguer, T. Rahman, G. Khanna, M. M. Mezghanni, L. Gardian, E. Riccardi, A. C. Piat, S. Calabrò, S. Dris, A. Richter, J. K. Fischer, B. Sommerkorn-Krombholz, and B. Spinnler, "Digital pre-compensation techniques enabling high-capacity bandwidth variable transponders," *Optics Communications*, vol. 409, pp. 52–65, 2018.

[19] B. Puttnam, R. Luis, Y. Huang, I. Phillips, D. Chung, N. Fotaine, G. Rademacher, M. Mazue, L. Dallachiesa, H. Chen, W. Forysiak, R. Man, R. Ryf, D. Nielson, and H. Furukawa, "301 Tb/s E, S, C+L-band transmission over 212 nm bandwidth with E-band bismuth-doped fiber amplifier and gain equalizer," in *Proceedings of European Conference on Optical Communication*, 2023.

[20] G. Rademacher, R. S. Luís, B. J. Puttnam, T. A. Eriksson, R. Ryf, E. Agrell, R. Maruyama, K. Aikawa, Y. Awaji, H. Furukawa, and N. Wada, "High capacity transmission with few-mode fibers," *Journal of Lightwave Technology*, vol. 37, no. 2, pp. 425–432, 2019.

[21] G. Rademacher, B. J. Puttnam, R. S. Luís, T. A. Eriksson, N. K. Fontaine, M. Mazur, H. Chen, R. Ryf, D. T. Neilson, P. Sillard, F. Achten, Y. Awaji, and H. Furukawa, "1.01 Peta-bit/s C+L-band transmission over a 15-mode fiber," in *Proceedings of European Conference on Optical Communication*, 2020.

[22] G. Rademacher, R. S. Luís, B. J. Puttnam, N. K. Fontaine, M. Mazur, H. Chen, R. Ryf, D. T. Neilson, D. Dahl, J. Carpenter, P. Sillard, F. Achten, M. Bigot, J. Sakaguchi, and H. Furukawa, "1.53 Peta-bit/s C-band transmission in a 55-mode fiber," in *Proceedings of European Conference on Optical Communication*, 2022.

[23] G. Rademacher, R. S. Luís, B. J. Puttnam, R. Ryf, S. van der Heide, T. A. Eriksson, N. K. Fontaine, H. Chen, R.-J. Essiambre, Y. Awaji, H. Furukawa, and N. Wada, "172 Tb/s C+L band transmission over 2040 km strongly coupled 3-core fiber," in *Proceedings of Optical Fiber Communication Conference*, 2020.

[24] B. J. Puttnam, R. S. Luis, G. Rademacher, Y. Awaji, and H. Furukawa, "1 Pb/s transmission in a 125μm diameter 4-core MCF," in *Conference on Lasers and Electro-Optics*, 2022.

[25] G. Rademacher, M. van den Hout, R. S. Luís, B. J. Puttnam, G. Di Sciullo, T. Hayashi, A. Inoue, T. Nagashima, S. Gross, A. Ross-Adams, M. J. Withford, J. Sakaguchi, C. Antonelli, C. Okonkwo, and H. Furukawa, "Randomly coupled 19-core multicore fiber with standard cladding diameter," in *Proceedings of Optical Fiber Communication Conference*, 2023.

[26] J. Renaudier, G. Charlet, O. Bertran Pardo, H. Mardoyan, P. Tran, M. Salsi, and S. Bigo, "Experimental analysis of 100Gb/s coherent

PDM-QPSK long-haul transmission under constraints of typical terrestrial networks," in *Proceedings of European Conference on Optical Communication*, 2008.

[27] P. J. Winzer, "High-spectral-efficiency optical modulation formats," *Journal of Lightwave Technology*, vol. 30, no. 24, pp. 3824–3835, 2012.

[28] P. J. Winzer and R.-J. Essiambre, "Advanced modulation formats for high-capacity optical transport networks," *Journal of Lightwave Technology*, vol. 24, no. 12, pp. 4711–4728, 2006.

[29] S. Benedetto and E. Biglieri, *Principles of Digital Transmission: With Wireless Applications*. Norwell, MA: Kluwer, 1999.

[30] J. Proakis, *Digital communications*. McGraw-Hill, 1995.

[31] C. Zhu, B. Corcoran, A. V. Tran, and A. J. Lowery, "Nyquist-WDM with low-complexity joint matched filtering and adaptive equalization," *IEEE Photonics Technology Letters*, vol. 26, no. 23, pp. 2323–2326, 2014.

[32] S. Daumont, B. Rihawi, and Y. Lout, "Root-raised cosine filter influences on PAPR distribution of single carrier signals," in *Proceedings of International Symposium on Communications, Control and Signal Processing*, 2008, pp. 841–845.

[33] W. Kester, "Understand SINAD, ENOB, SNR, THD, THD + N, and SFDR," in *Analog Devices*, 2009, Tutorial MT-003.

[34] D. Rafique, A. Napoli, S. Calabro, and B. Spinnler, "Digital preemphasis in optical communication systems: On the DAC requirements for Terabit transmission applications," *Journal of Lightwave Technology*, vol. 32, no. 19, pp. 3247–3256, 2014.

[35] F. Buchali, W. Idler, N. Rastegardoost, T. Drenski, R. Ward, and L. Zhao, "Preemphased prime frequency multicarrier bases ENOB assessment and its application for optimizing a dual-carrier 1-Tb/s QAM transmitter," in *Proceedings of Optical Fiber Communication Conference*, 2016.

[36] N. Benvenuto and G. Cherubini, *Algorithms for Communications Systems and Their Applications*, 2002, ch. 5, pp. 331–436.

[37] C. Laperle and M. O'Sullivan, "Advances in high-speed DACs, ADCs, and DSP for optical coherent transceivers," *Journal of Lightwave Technology*, vol. 32, no. 4, pp. 629–643, 2014.

[38] Y. Yoffe, E. Wohlgemuth, and D. Sadot, "Digitally enhanced DAC: Low-resolution digital pre-compensation for high speed optical links," in *Proceedings of Optical Fiber Communication Conference*, 2018, pp. 1–3.

[39] Y. Yoffe, G. Khanna, E. Wohlgemuth, E. de Man, B. Spinnler, N. Hanik, A. Napoli, and D. Sadot, "Low-resolution digital pre-compensation enabled by digital resolution enhancer," *Journal of Lightwave Technology*, vol. 37, no. 6, pp. 1543–1551, 2019.

[40] M. van den Hout, S. van der Heide, and C. Okonkwo, "Digital resolution enhancer employing clipping for high-speed optical transmission," *Journal of Lightwave Technology*, vol. 38, no. 11, pp. 2897–2904, 2020.

[41] J. Vuolevi, T. Rahkonen, and J. Manninen, "Measurement technique for characterizing memory effects in RF power amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 8, pp. 1383–1389, 2001.

[42] "40Gbps $LiNbO_3$ external modulator," . [Online]. Available: https://www.fujitsu.com/downloads/OPTCMP/lineup/40gln/ln40gnrz-catalog-e.pdf.

[43] D. Rafique, H. Griesser, and J.-P. Elbers, "Enabling 64Gbaud coherent optical transceivers," in *Proceedings of Optical Fiber Communication Conference*, 2017.

[44] "AEDFA-23 Amonics C-band inline EDFA," . [Online]. Available: https://www.amonics.com/product/10.

[45] "AEDFA-L-DWDM Amonics L-band inline EDFA," . [Online]. Available: https://www.amonics.com/product/5.

[46] D. Marcuse, C. Manyuk, and P. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *Journal of Lightwave Technology*, vol. 15, no. 9, pp. 1735–1746, 1997.

[47] C. Menyuk and B. Marks, "Interaction of polarization mode dispersion and nonlinearity in optical fiber transmission systems," *Journal of Lightwave Technology*, vol. 24, no. 7, pp. 2806–2826, 2006.

[48] P. Serena and A. Bononi, "An alternative approach to the Gaussian noise model and its system implications," *Journal of Lightwave Technology*, vol. 31, no. 22, pp. 3489–3499, 2013.

[49] P. Poggiolini, A. Carena, V. Curri, G. Bosco, and F. Forghieri, "Analytical modeling of nonlinear propagation in uncompensated optical transmission links," *IEEE Photonics Technology Letters*, vol. 23, no. 11, pp. 742–744, 2011.

[50] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications," *Journal of Lightwave Technology*, vol. 32, no. 4, pp. 694–721, 2014.

[51] P. Poggiolini, G. Bosco, A. Carena, V. Curri, and F. Forghieri, "A simple and accurate model for non-linear propagation effects in uncompensated coherent transmission links," in *Proceedings of International Conference on Transparent Optical Networks*, 2011.

[52] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Optics Express*, vol. 22, no. 13, pp. 16 335–16 362, 2014.

[53] G. Agrawal, in *Nonlinear Fiber Optics (Fifth Edition)*. Academic Press, 2013.

[54] D. S. Millar, R. Maher, D. Lavery, T. Koike-Akino, M. Pajovic, A. Alvarado, M. Paskov, K. Kojima, K. Parsons, B. C. Thomsen, S. J. Savory, and P. Bayvel, "Design of a 1 Tb/s superchannel coherent receiver," *Journal of Lightwave Technology*, vol. 34, no. 6, pp. 1453–1463, 2016.

[55] M. S. Faruk and S. J. Savory, "Digital signal processing for coherent transceivers employing multilevel formats," *Journal of Lightwave Technology*, vol. 35, no. 5, pp. 1125–1141, 2017.

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

[57] P. Löwdin, "On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals," *The Journal of Chemical Physics*, vol. 18, no. 3, pp. 365–375, 1950.

[58] I. Mayer, "On Löwdin's method of symmetric orthogonalization," *International Journal of Quantum Chemistry*, vol. 90, no. 1, pp. 63–65, 2002.

[59] A. Eghbali, H. Johansson, O. Gustafsson, and S. J. Savory, "Optimal least-squares FIR digital filters for compensation of chromatic dispersion in digital coherent optical receivers," *Journal of Lightwave Technology*, vol. 32, no. 8, pp. 1449–1456, 2014.

[60] J. Shynk, "Frequency-domain and multirate adaptive filtering," *IEEE Signal Processing Magazine*, vol. 9, no. 1, pp. 14–37, 1992.

[61] T. Xu, G. Jacobsen, S. Popov, M. Forzati, J. Mårtensson, M. Mussolin, J. Li, K. Wang, Y. Zhang, and A. T. Friberg, "Frequency-domain chromatic dispersion equalization using overlap-add methods in coherent optical system," *Journal of Optical Communications*, vol. 32, no. 2, pp. 131–135, 2011.

[62] J. C. Geyer, C. R. S. Fludger, T. Duthel, C. Schulien, and B. Schmauss, "Efficient frequency domain chromatic dispersion compensation in a coherent Polmux QPSK-receiver," in *Proceedings of Optical Fiber Communication Conference*, 2010.

[63] R. Kudo, T. Kobayashi, K. Ishihara, Y. Takatori, A. Sano, and Y. Miyamoto, "Coherent optical single carrier transmission using overlap frequency domain equalization for long-haul optical systems," *Journal of Lightwave Technology*, vol. 27, no. 16, pp. 3721–3728, 2009.

[64] A. Leven, N. Kaneda, U.-V. Koc, and Y.-K. Chen, "Frequency estimation in intradyne reception," *IEEE Photonics Technology Letters*, vol. 19, no. 6, pp. 366–368, 2007.

[65] S. Hoffmann, S. Bhandare, T. Pfau, O. Adamczyk, C. Wordehoff, R. Peveling, M. Porrmann, and R. Noe, "Frequency and phase estimation for coherent QPSK transmission with unlocked DFB lasers," *IEEE Photonics Technology Letters*, vol. 20, no. 18, pp. 1569–1571, 2008.

[66] M. G. Taylor, "Phase estimation methods for optical coherent detection using digital signal processing," *Journal of Lightwave Technology*, vol. 27, no. 7, pp. 901–914, 2009.

[67] A. J. Viterbi and A. M. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission," *IEEE Transactions on Information Theory*, vol. 29, no. 4, pp. 543–551, 1983.

[68] T. Pfau, S. Hoffmann, and R. Noe, "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for M-QAM constellations," *Journal of Lightwave Technology*, vol. 27, no. 8, pp. 989–999, 2009.

[69] M. Magarini, L. Barletta, A. Spalvieri, F. Vacondio, T. Pfau, M. Pepe, M. Bertolini, and G. Gavioli, "Pilot-symbols-aided carrier-phase recovery for 100-G PM-QPSK digital coherent receivers," *IEEE Photonics Technology Letters*, vol. 24, no. 9, pp. 739–741, 2012.

[70] A. Spalvieri and L. Barletta, "Pilot-aided carrier recovery in the presence of phase noise," *IEEE Transactions on Communications*, vol. 59, no. 7, pp. 1966–1974, 2011.

[71] M. Pajovic, D. S. Millar, T. Koike-Akino, R. Maher, D. Lavery, A. Alvarado, M. Paskov, K. Kojima, K. Parsons, B. C. Thomsen, S. J. Savory, and P. Bayvel, "Experimental demonstration of multi-pilot aided carrier phase estimation for DP-64QAM and DP-256QAM," in *Proceedings of European Conference on Optical Communication*, 2015.

[72] D. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Transactions on Communications*, vol. 28, no. 11, pp. 1867–1875, 1980.

[73] V. Zarzoso and P. Comon, "Optimal step-size constant modulus algorithm," *IEEE Transactions on Communications*, vol. 56, no. 1, pp. 10–13, 2008.

[74] M. Ready and R. Gooch, "Blind equalization based on radius directed adaptation," in *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, 1990, pp. 1699–1702.

[75] D. Lavery, M. Paskov, R. Maher, S. J. Savory, and P. Bayvel, "Modified radius directed equaliser for high order QAM," in *Proceedings of European Conference on Optical Communication*, 2015.

[76] S. Qureshi, "Adaptive equalization," *IEEE Communications Magazine*, vol. 20, pp. 9–16, 1982.

[77] M. Mazur, J. Schröder, A. Lorences-Riesgo, T. Yoshida, M. Karlsson, and P. A. Andrekson, "Overhead-optimization of pilot-based digital signal processing for flexible high spectral efficiency transmission," *Optics Express*, vol. 27, no. 17, pp. 24 654–24 669, 2019.

[78] M. Mazur, J. Schröder, A. Lorences-Riesgo, T. Yoshida, M. Karlsson, and P. A. Andrekson, "12 b/s/Hz spectral efficiency over the C-band based on comb-based superchannels," *Journal of Lightwave Technology*, vol. 37, no. 2, pp. 411–417, 2019.

[79] M. Kuschnerov, M. Chouayakh, K. Piyawanno, B. Spinnler, E. de Man, P. Kainzmaier, M. S. Alfiad, A. Napoli, and B. Lankl, "Data-aided versus blind single-carrier coherent receivers," *IEEE Photonics Journal*, vol. 2, no. 3, pp. 387–403, 2010.

[80] S. O. Haykin, *Adaptive Filter Theory (5th Edition).* Pearson, 2013.

[81] X. Zhou, X. Chen, and K. Long, "Wide-range frequency offset estimation algorithm for optical coherent systems using training sequence," *IEEE Photonics Technology Letters*, vol. 24, no. 1, pp. 82–84, 2012.

[82] J. G. D. Forney, *Concatenated Codes.* MIT Press, 1965.

[83] M. Barakatain, D. Lentner, G. Böecherer, and F. R. Kschischang, "Performance-complexity tradeoffs of concatenated FEC for higher-order modulation," *Journal of Lightwave Technology*, vol. 38, no. 11, pp. 2944–2953, 2020.

[84] B. J. Puttnam, R. S. Luís, W. Klaus, J. Sakaguchi, J.-M. Delgado Mendinueta, Y. Awaji, N. Wada, Y. Tamura, T. Hayashi, M. Hirano, and J. Marciante, "2.15 Pb/s transmission using a 22 core homogeneous single-mode multi-core fiber and wideband optical comb," in *Proceedings of European Conference on Optical Communication*, 2015.

[85] "Digital video broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications (DVB-S2)," vol. ETSI EN 302 307, V1.2.1, pp. 3247–3256, 2009.

[86] Z. Xing, M. Xiang, E. El-Fiky, X. Li, M. G. Saber, L. Xu, P.-C. Koh, and D. V. Plant, "Experimental demonstration of 600 Gb/s net rate PAM4 transmissions over 2 km and 10 km with a 4-λ CWDM TOSA," *Journal of Lightwave Technology*, vol. 38, no. 11, pp. 2968–2975, 2020.

[87] F. Buchali, A. Klekamp, L. Schmalen, and T. Drenski, "Implementation of 64QAM at 42.66 GBaud using 1.5 samples per symbol DAC and demonstration of up to 300 km fiber transmission," in *Proceedings of Optical Fiber Communication Conference*, 2014.

[88] K. Poulton, R. Neff, B. Setterberg, B. Wuppermann, T. Kopley, R. Jewett, J. Pernillo, C. Tan, and A. Montijo, "A 20 GS/s 8b ADC with a 1 MB memory in 0.18μm CMOS," in *IEEE International Solid-State Circuits Conference*, vol. 1, 2003, pp. 318–496.

[89] K. Roberts, M. O'Sullivan, K.-T. Wu, H. Sun, A. Awadalla, D. J. Krause, and C. Laperle, "Performance of dual-polarization QPSK for optical transport systems," *Journal of Lightwave Technology*, vol. 27, no. 16, pp. 3546–3559, 2009.

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

[91] S. Tsukamoto, K. Katoh, and K. Kikuchi, "Coherent demodulation of optical multilevel phase-shift-keying signals using homodyne detection and digital signal processing," *IEEE Photonics Technology Letters*, vol. 18, no. 10, pp. 1131–1133, 2006.

[92] K. Roberts, D. Beckett, D. Boertjes, J. Berthold, and C. Laperle, "100G and beyond with digital coherent signal processing," *IEEE Communications Magazine*, vol. 48, no. 7, pp. 62–69, 2010.

[93] A. C. Singer, N. R. Shanbhag, and H.-m. Bae, "Electronic dispersion compensation," *IEEE Signal Processing Magazine*, vol. 25, no. 6, pp. 110–130, 2008.

[94] H. Bulow, F. Buchali, and A. Klekamp, "Electronic dispersion compensation," *Journal of Lightwave Technology*, vol. 26, no. 1, pp. 158–167, 2008.

[95] X. Li, X. Chen, G. Goldfarb, E. Mateo, I. Kim, F. Yaman, and G. Li, "Electronic post-compensation of WDM transmission impairments using coherent detection and digital signal processing," *Optics Express*, vol. 16, no. 2, pp. 880–888, 2008.

[96] E. M. Ip and J. M. Kahn, "Fiber impairment compensation using coherent detection and digital signal processing," *Journal of Lightwave Technology*, vol. 28, no. 4, pp. 502–519, 2010.

[97] T. Naito, T. Terahara, T. Chikama, and M. Suyama, "Four 5-Gbit/s wdm transmission over 4760-km straight-line using pre- and post-dispersion compensation and FWM cross talk reduction," in *Proceedings of Optical Fiber Communication Conference*, 1996, pp. 182–183.

[98] T. Koch and R. Alferness, "Dispersion compensation by active predistorted signal synthesis," *Journal of Lightwave Technology*, vol. 3, no. 4, pp. 800–805, 1985.

[99] G. Khanna, B. Spinnler, S. Calabrò, E. De Man, and N. Hanik, "A robust adaptive pre-distortion method for optical communication transmitters," *IEEE Photonics Technology Letters*, vol. 28, no. 7, pp. 752–755, 2016.

[100] G. Khanna, S. Calabrò, B. Spinnler, E. De Man, and N. Hanik, "Joint adaptive pre-compensation of transmitter I/Q skew and frequency response for high order modulation formats and high baud rates," in *Proceedings of Optical Fiber Communication Conference*, 2015.

[101] P. W. Berenguer, M. Nölle, L. Molle, T. Raman, A. Napoli, C. Schubert, and J. K. Fischer, "Nonlinear digital pre-distortion of transmitter components," *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1739–1745, 2016.

[102] R. Killey, "Dispersion and nonlinearity compensation using electronic predistortion techniques," in *IEEE Seminar on Optical Fibre Communications and Electronic Signal Processing*, 2005.

[103] R. Killey, P. Watts, M. Glick, and P. Bayvel, "Electronic precompensation techniques to combat dispersion and nonlinearities in optical transmission," in *Proceedings of European Conference on Optical Communication*, vol. 2, 2005, pp. 251–254.

[104] R. Killey, P. Watts, M. Glick, and P. Bayvel, "Electronic dispersion compensation by signal predistortion," in *Proceedings of Optical Fiber Communication Conference*, 2006.

[105] J. Kim and K. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory," *Electronics Letters*, vol. 37, no. 23, pp. 1417–1418, 2001.

[106] D. Zhou and V. E. DeBrunner, "Novel adaptive nonlinear predistorters based on the direct learning algorithm," *IEEE Transactions on Signal Processing*, vol. 55, no. 1, pp. 120–133, 2007.

[107] C. Eun and E. Powers, "A new volterra predistorter based on the indirect learning architecture," *IEEE Transactions on Signal Processing*, vol. 45, no. 1, pp. 223–227, 1997.

[108] D. Psaltis, A. Sideris, and A. Yamamura, "A multilayered neural network controller," *IEEE Control Systems Magazine*, vol. 8, no. 2, pp. 17–21, 1988.

[109] A. Zhu and T. Brazil, "An adaptive volterra predistorter for the linearization of RF high power amplifiers," in *IEEE MTT-S International Microwave Symposium Digest*, vol. 1, 2002, pp. 461–464.

[110] L. Ding, G. Zhou, D. Morgan, Z. Ma, J. Kenney, J. Kim, and C. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Transactions on Communications*, vol. 52, no. 1, pp. 159–165, 2004.

[111] H. Paaso and A. Mammela, "Comparison of direct learning and indirect learning predistortion architectures," in *IEEE International Symposium on Wireless Communication Systems*, 2008, pp. 309–313.

[112] Z. Yu and E. Zhu, "A comparative study of learning architecture for digital predistortion," in *Asia-Pacific Microwave Conference*, 2015.

[113] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems.* New York: John Wiley & Sons, 1980.

[114] A. Napoli, M. M. Mezghanni, T. Rahman, D. Rafique, R. Palmer, B. Spinnler, S. Calabro, C. Castro, M. Kuschnerov, and M. Bohn, "Digital compensation of bandwidth limitations for high-speed DACs and ADCs," *Journal of Lightwave Technology*, vol. 34, no. 13, p. 3053 – 3064, 2016.

[115] M. Dallaglio, A. Giorgetti, N. Sambo, L. Velasco, and P. Castoldi, "Routing, spectrum, and transponder assignment in elastic optical networks," *Journal of Lightwave Technology*, vol. 33, no. 22, pp. 4648–4658, 2015.

[116] O. Gerstel, M. Jinno, A. Lord, and S. B. Yoo, "Elastic optical networking: A new dawn for the optical layer?" *IEEE Communications Magazine*, vol. 50, no. 2, pp. s12–s20, 2012.

[117] A. Napoli, S. Calabrò, D. Rafique, R. Palmer, B. Spinnler, and M. Bohn, "Adaptive digital pre-emphasis for high speed digital analogue converters," in *Proceedings of Optical Fiber Communication Conference*, 2016.

[118] M. H. Hayes, *Digital Signal Processing.* New York: McGraw Hill, 1999.

[119] J. Grabowski and R. C. Davis, "An experimental M-QAM modem using amplifier linearization and baseband equalization techniques," in *National Telesystems Conference*, 1982, p. E3.2.

[120] A. A. M. Saleh and J. Salz, "Adaptive linearization of power amplifiers in digital radio systems," *The Bell System Technical Journal*, vol. 62, no. 4, pp. 1019–1033, 1983.

[121] Y. Nagata, "Linear amplification technique for digital mobile communications," in *IEEE Vehicular Technology Conference*, vol. 1, 1989, pp. 159–164.

[122] J. Cavers, "Amplifier linearization using a digital predistorter with fast adaptation and low memory requirements," *IEEE Transactions on Vehicular Technology*, vol. 39, no. 4, pp. 374–382, 1990.

[123] S. Zhalehpour, J. Lin, W. Shi, and L. A. Rusch, "Reduced-size lookup tables enabling higher-order QAM with all-silicon IQ modulators," *Optics Express*, vol. 27, no. 17, pp. 24 243–24 259, 2019.

[124] J. Zhang, P. Gou, M. Kong, K. Fang, J. Xiao, Q. Zhang, X. Xin, and J. Yu, "PAM-8 IM/DD transmission based on modified lookup table nonlinear predistortion," *IEEE Photonics Journal*, vol. 10, no. 3, pp. 1–9, 2018.

[125] D. Morgan, Z. Ma, J. Kim, M. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of RF power amplifiers," *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3852–3860, 2006.

[126] B. Karanov, M. Chagnon, F. Thouin, T. A. Eriksson, H. Bülow, D. Lavery, P. Bayvel, and L. Schmalen, "End-to-end deep learning of optical fiber communications," *Journal of Lightwave Technology*, vol. 36, no. 20, pp. 4843–4855, 2018.

[127] A. Bernardini, M. Carrarini, and S. De Fina, "The use of a neural net for copeing with nonlinear distortions," in *European Microwave Conference*, vol. 2, 1990, pp. 1718–1723.

[128] M. Schaedler, M. Kuschnerov, S. Calabrò, F. Pittalà, C. Bluemm, and S. Pachnicke, "AI-based digital predistortion for IQ Mach-Zehnder modulators," in *Asia Communications and Photonics Conference*, 2019.

[129] M. Abu-Romoh, S. Sygletos, I. D. Phillips, and W. Forysiak, "Neural-network-based pre-distortion method to compensate for low resolution DAC nonlinearity," in *Proceedings of European Conference on Optical Communication*, 2019.

[130] C. Tarver, A. Balatsoukas-Stimming, and J. R. Cavallaro, "Design and implementation of a neural network based predistorter for enhanced mobile broadband," in *IEEE International Workshop on Signal Processing Systems*, 2019, pp. 296–301.

[131] T. Gotthans, G. Baudoin, and A. Mbaye, "Digital predistortion with advance/delay neural network and comparison with Volterra derived models," in *IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communication*, 2014, pp. 811–815.

[132] R. Hongyo, Y. Egashira, T. M. Hone, and K. Yamaguchi, "Deep neural network-based digital predistorter for doherty power amplifiers," *IEEE Microwave and Wireless Components Letters*, vol. 29, no. 2, pp. 146–148, 2019.

[133] X. Hu, Z. Liu, X. Yu, Y. Zhao, W. Chen, B. Hu, X. Du, X. Li, M. Helaoui, W. Wang, and F. M. Ghannouchi, "Convolutional neural network for behavioral modeling and predistortion of wideband power amplifiers," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 8, pp. 3923–3937, 2022.

[134] J. Sun, W. Shi, Z. Yang, J. Yang, and G. Gui, "Behavioral modeling and linearization of wideband RF power amplifiers using BiLSTM networks for 5G wireless systems," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 11, pp. 10 348–10 356, 2019.

[135] T. Sasai, M. Nakamura, E. Yamazaki, A. Matsushita, S. Okamoto, K. Horikoshi, and Y. Kisaka, "Wiener-Hammerstein model and its learning for nonlinear digital pre-distortion of optical transmitters," *Optics Express*, vol. 28, no. 21, pp. 30 952–30 963, 2020.

[136] V. Bajaj, F. Buchali, M. Chagnon, S. Wahls, and V. Aref, "Deep neural network-based digital pre-distortion for high baudrate optical coherent transmission," *Journal of Lightwave Technology*, vol. 40, no. 3, pp. 597–606, 2022.

[137] S. Dörner, S. Cammerer, J. Hoydis, and S. t. Brink, "Deep learning based communication over the air," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 132–143, 2018.

[138] F. A. Aoudia and J. Hoydis, "Model-free training of end-to-end communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 11, p. 2503–2516, 2019.

[139] P. J. Winzer and D. T. Neilson, "From scaling disparities to integrated parallelism: A decathlon for a decade," *Journal of Lightwave Technology*, vol. 35, no. 5, pp. 1099–1115, 2017.

[140] "Baud rate, modulation, and maximizing coherent optical performance," 2021. [Online]. Available: https://www.infinera.com/wp-content/uploads/Baud-Rate-Modulation-and-Maximizing-Coherent-Optical-Performance-0294-WP-RevA-0921.pdf.

[141] J. Rahn, L. Dardis, D. Krause, M. Rice, C. Berry, A. Kumpera, A. Nilsson, X. Xu, K. Croussore, P. Samra, K. Weidner, Z. Morbi, S. DeMars, A. Vasilyev, C. Chen, and P. Freeman, "DSP-enabled frequency locking for near-Nyquist spectral efficiency superchannels utilizing integrated photonics," in *Proceedings of Optical Fiber Communication Conference*, 2018.

[142] M. Mazur, M.-G. Suh, A. Fülöp, J. Schröder, V. Torres-Company, M. Karlsson, K. Vahala, and P. Andrekson, "High spectral efficiency coherent superchannel transmission with soliton microcombs," *Journal of Lightwave Technology*, vol. 39, no. 13, pp. 4367–4373, 2021.

[143] D. Che and W. Shieh, "Approaching the capacity of colored-SNR optical channels by multicarrier entropy loading," *Journal of Lightwave Technology*, vol. 36, no. 1, pp. 68–78, 2018.

[144] D. Che, A. Li, and W. Shieh, "Entropy loading for multi-carrier optical systems," in *Next-Generation Optical Communication: Components, Sub-Systems, and Systems VII*, vol. 10561. SPIE, 2018, pp. 146–152.

[145] L. Lundberg, M. Mazur, A. Lorences-Riesgo, M. Karlsson, and P. A. Andrekson, "Joint carrier recovery for DSP complexity reduction in frequency comb-based superchannel transceivers," in *Proceedings of European Conference on Optical Communication*, 2017.

[146] L. Lundberg, M. Karlsson, A. Lorences-Riesgo, M. Mazur, V. Torres-Company, J. Schröder, and P. A. Andrekson, "Frequency comb-based WDM transmission systems enabling joint signal processing," *Applied Sciences*, vol. 8, no. 5, 2018.

[147] J. Pan, C. Liu, T. Detwiler, A. J. Stark, Y.-T. Hsueh, and S. E. Ralph, "Inter-channel crosstalk cancellation for Nyquist-WDM superchannel applications," *Journal of Lightwave Technology*, vol. 30, no. 24, pp. 3993–3999, 2012.

[148] C. Liu, J. Pan, T. Detwiler, A. Stark, Y.-T. Hsueh, G.-K. Chang, and S. E. Ralph, "Joint digital signal processing for superchannel coherent optical communication systems," *Optics Express*, vol. 21, no. 7, pp. 8342–8356, 2013.

[149] N. Alic, E. Myslivets, E. Temprana, B. P.-P. Kuo, and S. Radic, "Nonlinearity cancellation in fiber optic links based on frequency referenced carriers," *Journal of Lightwave Technology*, vol. 32, no. 15, pp. 2690–2698, 2014.

[150] E. Temprana, E. Myslivets, B.-P. Kuo, L. Liu, V. Ataie, N. Alic, and S. Radic, "Overcoming Kerr-induced capacity limit in optical fiber transmission," *Science*, vol. 348, no. 6242, pp. 1445–1448, 2015.

[151] A. Lorences-Riesgo, T. A. Eriksson, A. Fülöp, P. A. Andrekson, and M. Karlsson, "Frequency-comb regeneration for self-homodyne superchannels," *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1800–1806, 2016.

[152] A. Lorences-Riesgo, M. Mazur, T. A. Eriksson, P. A. Andrekson, and M. Karlsson, "Self-homodyne 24×32-QAM superchannel receiver enabled by all-optical comb regeneration using Brillouin amplification," *Opt. Express*, vol. 24, no. 26, pp. 29 714–29 723, 2016.

[153] M. Mazur, A. Lorences-Riesgo, J. Schröder, P. A. Andrekson, and M. Karlsson, "High spectral efficiency PM-128QAM comb-based superchannel transmission enabled by a single shared optical pilot tone," *Journal of Lightwave Technology*, vol. 36, no. 6, pp. 1318–1325, 2018.

[154] M. Mazur, J. Schröder, M. Karlsson, and P. A. Andrekson, "Joint superchannel digital signal processing for effective inter-channel interference cancellation," *Journal of Lightwave Technology*, vol. 38, no. 20, pp. 5676–5684, 2020.

[155] N. K. Fontaine, G. Raybon, B. Guan, A. Adamiecki, P. J. Winzer, R. Ryf, A. Konczykowska, F. Jorge, J.-Y. Dupuy, L. L. Buhl, S. Chandrashekhar, R. Delbue, P. Pupalaikis, and A. Sureka, "228-GHz coherent receiver using digital optical bandwidth interleaving and reception of 214-GBd (856-Gb/s) PDM-QPSK," in *Proceedings of European Conference on Optical Communication*, 2012.

[156] B. Geiger, E. Sillekens, F. Ferreira, R. Killey, L. Galdino, and P. Bayvel, "On the performance limits of high-speed transmission using a single wideband coherent receiver," *Journal of Lightwave Technology*, vol. 41, no. 12, pp. 3816–3824, 2023.

[157] R. Dar and P. J. Winzer, "On the limits of digital back-propagation in fully loaded WDM systems," *IEEE Photonics Technology Letters*, vol. 28, no. 11, pp. 1253–1256, 2016.

[158] N. K. Fontaine, X. Liu, S. Chandrasekhar, R. Ryf, S. Randel, P. Winzer, R. Delbue, P. Pupalaikis, and A. Sureka, "Fiber nonlinearity compensation by digital backpropagation of an entire 1.2-Tb/s superchannel using a full-field spectrally-sliced receiver," in *Proceedings of European Conference on Optical Communication*, 2013.

[159] G. Liga, T. Xu, A. Alvarado, R. I. Killey, and P. Bayvel, "On the performance of multichannel digital backpropagation in high-capacity long-haul optical transmission," *Optics Express*, vol. 22, no. 24, pp. 30 053–30 062, 2014.

[160] F. P. Guiomar, S. B. Amado, R. M. Ferreira, J. D. Reis, S. M. Rossi, A. Chiuchiarelli, J. R. F. de Oliveira, A. L. Teixeira, and A. N. Pinto, "Multicarrier digital backpropagation for 400G optical superchannels," *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1896–1907, 2016.

[161] C. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, 1949.

[162] A. Napoli, Z. Maalej, V. A. J. M. Sleiffer, M. Kuschnerov, D. Rafique, E. Timmers, B. Spinnler, T. Rahman, L. D. Coelho, and N. Hanik, "Reduced complexity digital back-propagation methods for optical communication systems," *Journal of Lightwave Technology*, vol. 32, no. 7, pp. 1351–1362, 2014.

[163] E. F. Mateo, F. Yaman, and G. Li, "Efficient compensation of inter-channel nonlinear effects via digital backward propagation in WDM optical transmission," *Optics Express*, vol. 18, no. 14, pp. 15 144–15 154, 2010.

[164] Z. Tao, L. Dou, W. Yan, L. Li, T. Hoshida, and J. C. Rasmussen, "Multiplier-free intrachannel nonlinearity compensating algorithm operating at symbol rate," *Journal of Lightwave Technology*, vol. 29, no. 17, pp. 2570–2576, 2011.

[165] E. Averyanov, A. Redyuk, O. Sidelnikov, M. Soroklna, M. Fedoruk, and S. Turitsyn, "Perturbative machine learning technique for nonlinear impairments compensation in WDM systems," in *Proceedings of European Conference on Optical Communication*, 2018.

[166] A. Redyuk, E. Averyanov, O. Sidelnikov, M. Fedoruk, and S. Turitsyn, "Compensation of nonlinear impairments using inverse perturbation theory with reduced complexity," *J. Lightw. Technol.*, vol. 38, no. 6, pp. 1250–1257, 2020.

[167] I. Kozulin and A. Redyuk, "Interchannel nonlinearity compensation using a perturbative machine learning technique," *Optics Communications*, vol. 493, pp. 1250–1257, 2021.

[168] P. Winzer, "Making spatial multiplexing a reality," *Nature Photonics*, vol. 8, no. 5, pp. 345–348, 2014.

[169] G. Forney, R. Gallager, G. Lang, F. Longstaff, and S. Qureshi, "Efficient modulation for band-limited channels," *IEEE Journal on Selected Areas in Communications*, vol. 2, no. 5, pp. 632–647, 1984.

[170] G. Böcherer, P. Schulte, and F. Steiner, "Probabilistic shaping and forward error correction for fiber-optic communication systems," *Journal of Lightwave Technology*, vol. 37, no. 2, pp. 230–244, 2019.

[171] J. Cho and P. J. Winzer, "Probabilistic constellation shaping for optical fiber communications," *Journal of Lightwave Technology*, vol. 37, no. 6, pp. 1590–1607, 2019.

[172] O. Jovanovic, F. Da Ros, D. Zibar, and M. P. Yankov, "Geometric constellation shaping for fiber-optic channels via end-to-end learning," *Journal of Lightwave Technology*, vol. 41, no. 12, pp. 3726–3736, 2023.

[173] F. Buchali, F. Steiner, G. Böcherer, L. Schmalen, P. Schulte, and W. Idler, "Rate adaptation and reach increase by probabilistically shaped 64-QAM: An experimental demonstration," *Journal of Lightwave Technology*, vol. 34, no. 7, pp. 1599–1609, 2016.

[174] P. Schulte and G. Böcherer, "Constant composition distribution matching," *IEEE Transactions on Information Theory*, vol. 62, no. 1, pp. 430–434, 2016.

[175] E. Sillekens, G. Liga, D. Lavery, P. Bayvel, and R. I. Killey, "High-cardinality geometrical constellation shaping for the nonlinear fibre channel," *Journal of Lightwave Technology*, vol. 40, no. 19, pp. 6374–6387, 2022.

[176] R. T. Jones, M. P. Yankov, and D. Zibar, "End-to-end learning for GMI optimized geometric constellation shape," in *Proceedings of European Conference on Optical Communication*, 2019.

[177] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Transactions on Information Theory*, vol. 28, no. 1, pp. 55–67, 1982.

[178] G. Forney and L.-F. Wei, "Multidimensional constellations. I. Introduction, figures of merit, and generalized cross constellations," *IEEE Journal on Selected Areas in Communications*, vol. 7, no. 6, pp. 877–892, 1989.

[179] D. S. Millar, T. Koike-Akino, S. Ö. Arık, K. Kojima, K. Parsons, T. Yoshida, and T. Sugihara, "High-dimensional modulation for coherent optical communications systems," *Opt. Express*, vol. 22, no. 7, pp. 8798–8812, 2014.

[180] E. Agrell and M. Karlsson, "Power-efficient modulation formats in coherent transmission systems," *J. Lightw. Technol.*, vol. 27, no. 22, pp. 5115–5126, 2009.

[181] T. A. Eriksson, P. Johannisson, M. Sjödin, E. Agrell, P. A. Andrekson, and M. Karlsson, "Frequency and polarization switched QPSK," in *Proceedings of European Conference on Optical Communication*, 2013.

[182] X. Liu, T. H. Wood, R. W. Tkach, and S. Chandrasekhar, "Demonstration of record sensitivity in an optically pre-amplified receiver by combining PDM-QPSK and 16-PPM with pilot-assisted digital coherent detection," in *Proceedings of Optical Fiber Communication Conference*, 2011.

[183] T. A. Eriksson, P. Johannisson, E. Agrell, P. A. Andrekson, and M. Karlsson, "Biorthogonal modulation in 8 dimensions experimentally implemented as 2PPM-PS-QPSK," in *Proceedings of Optical Fiber Communication Conference*, 2014.

[184] J. Conway and N. Sloane, "A fast encoding method for lattice codes and quantizers," *IEEE Transactions on Information Theory*, vol. 29, no. 6, pp. 820–824, 1983.

[185] G. Forney, "Multidimensional constellations. II. Voronoi constellations," *IEEE Journal on Selected Areas in Communications*, vol. 7, no. 6, pp. 941–958, 1989.

[186] J. H. Conway and N. J. A. Sloane, "Fast quantizing and decoding algorithms for lattice quantizers and codes," *IEEE Transactions on Information Theory*, vol. IT-28, no. 2, pp. 227–232, 1982.

[187] B. M. Kurkoski, "Encoding and indexing of lattice codes," *IEEE Transactions on Information Theory*, vol. 64, no. 9, pp. 6320–6332, 2018.

[188] A. Mirani, K. Vijayan, S. Li, Z. He, J. Schröder, P. Andrekson, E. Agrell, and M. Karlsson, "Comparison of physical realizations of multidimensional Voronoi constellations in single mode fibers," in *Proceedings of European Conference on Optical Communication*, 2022.

[189] A. Mirani, K. Vijayan, S. Li, Z. He, E. Agrell, J. Schröder, P. Andrekson, and M. Karlsson, "Physical realizations of multidimensional Voronoi constellations in optical communication systems," *Journal of Lightwave Technology*, vol. 41, no. 17, pp. 5557–5563, 2023.

[190] S. Li, A. Mirani, M. Karlsson, and E. Agrell, "Coded modulation schemes for Voronoi constellations," *arxiv:2308.00407*, pp. 1–13, 2023.

[191] S. Li, A. Mirani, M. Karlsson, and E. Agrell, "Low-complexity Voronoi shaping for the Gaussian channel," *IEEE Transactions on Communications*, vol. 70, no. 2, pp. 865–873, 2022.

[192] S. Li, A. Mirani, M. Karlsson, and E. Agrell, "Power-efficient Voronoi constellations for fiber-optic communication systems," *Journal of Lightwave Technology*, vol. 41, no. 5, pp. 1298–1308, 2023.

[193] M. Nakamura, M. Nagatani, T. Jyo, F. Hamaoka, M. Mutoh, Y. Shiratori, H. Wakita, T. Kobayashi, H. Takahashi, and

Y. Miyamoto, "Over 2-Tb/s net bitrate single-carrier transmission based on >130-GHz-bandwidth InP-DHBT baseband amplifier module," in *Proceedings of European Conference on Optical Communication*, 2022.

[194] C. Eschenbaum, A. Mertens, C. Füllner, A. Kuzmin, A. Schwarzenberger, A. Kotz, G. Ramann, M. Chen, J. Drisko, B. Johnson, J. Zyskind, J. Marcelli, M. Lebby, W. Freude, S. Randel, and C. Koos, "Thermally stable silicon-organic hybrid (SOH) Mach-Zehnder modulator for 140 GBd PAM4 transmission with sub-1 v drive signals," in *Proceedings of European Conference on Optical Communication*, 2022.

[195] H. Mardoyan, S. Almonacil, F. Jorge, F. Pittalà, M. Xu, B. Krueger, F. Blache, B. Duval, L. Chen, Y. Yan, X. Ye, A. Ghazisaeidi, S. Rimpf, Y. Zhu, J. Wang, M. Goix, Z. Hu, M. Duthoit, M. Gruen, X. Cai, and J. Renaudier, "First 260-GBd single-carrier coherent transmission over 100 km distance based on novel arbitrary waveform generator and thin-film lithium niobate I/Q modulator," in *Proceedings of European Conference on Optical Communication*, 2022.

[196] M. Xu, Y. Zhu, F. Pittalà, J. Tang, M. He, W. C. Ng, J. Wang, Z. Ruan, X. Tang, M. Kuschnerov, L. Liu, S. Yu, B. Zheng, and X. Cai, "Dual-polarization thin-film lithium niobate in-phase quadrature modulators for terabit-per-second transmission," *Optica*, vol. 9, no. 1, pp. 61–62, 2022.

[197] D. Soma, T. Kato, S. Beppu, D. J. Elson, H. Muranaka, H. Irie, S. Okada, Y. Tanaka, Y. Wakayama, N. Yoshikane, T. Hoshida, and T. Tsuritani, "25-Thz O+S+C+L+U-band digital coherent DWDM transmission using a deployed fibre-optic cable," in *Proceedings of European Conference on Optical Communication*, 2023.

[198] P. Hazarika, M. Tan, A. Donodin, S. Noor, I. Phillips, P. Harper, J. S. Stone, M. J. Li, and W. Forysiak, "E-, S-, C- and L-band coherent transmission with a multistage discrete Raman amplifier," *Optics Express*, vol. 30, no. 24, pp. 43 118–43 126, 2022.