Advanced Digital Signal Processing for Next-Generation Coherent Optical Communication Transceivers

Medeiros Diniz, Júlio César

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Technical University of Denmark

Ph.D. Thesis

Advanced Digital Signal Processing for Next-Generation Coherent Optical Communication Transceivers

Júlio César Medeiros Diniz

Supervisors:
Darko Zibar, Assoc. Prof.
Molly Piels, PhD
Francesco Da Ros, PhD

February 2019

Ørsteds Plads, Byg. 343
DK-2800 Kongens Lyngby
Denmark

DTU Fotonik
Department of Photonics Engineering

Preface

The work presented in this Thesis was carried out as a part of my Ph.D. project in the period of 1st of October, 2015 to 26th February, 2019. The work took place at the Department of Photonics Engineering of the Technical University of Denmark (DTU Fotonik), with a visiting stay of 14 weeks at the Photonics Research Center of The Hong Kong Polytechnic University (PolyU).

The Ph.D. project was financed by the Villum Foundation Young Investigator Program and supervised by

- Darko Zibar (main supervisor), Associate Professor, Department of Photonics Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark.
- Molly Piels (co-supervisor), Hardware Engineer, Juniper Networks, Goleta, CA, USA. (Until December 2016: Researcher, Department of Photonics Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark)
- Francesco Da Ros (co-supervisor), Researcher, Department of Photonics Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark.

Abstract

For many years, research on optical communication technologies have been driven by the ever-increasing demand for higher capacity, lower costs and more energy efficiency. To avoid a capacity crunch, the design of future systems needs to be constantly upgraded. Thus, new advanced digital signal processing (DSP) systems need to be developed in order to meet the requirements of coherent optical communication systems of next generations.

In this context, the contributions presented in this thesis relate to the main topic of DSP for coherent optical communication systems and more specifically in the subsequent subjects: (a) clock recovery; (b) transceiver calibration; and (c) carrier phase recovery (CPR). Regarding (a), original contributions to the study of fully digital clock recovery are presented. Numerical performance investigations are shown for both polarization division multiplexing (PDM) and spatial division multiplexing (SDM) systems. For (b), it is demonstrated novel methods for calibration of both transmitters and receivers. At the transmitter-side, an application of a cooperative coevolutionary genetic algorithm (CC-GA) is discussed. The original contribution comprises the calibration of time skews between electrical components, bias voltages and amplitude imbalances, and it presents novel parameters that can be used for calibration of transmitters. Also, it is demonstrated a joint chromatic dispersion (CD) and time skew estimator for coherent optical receivers. Numerical and experimental performance evaluations are carried out for both methods. Finally, concerning (c), a new algorithm based on principal component analysis (PCA) is proposed for hardware-efficient CPR. The method is compared to state-of-the-art methods by means of simulations and experiments, and outperforms both in computational complexity and overall performance.

Resumé på Dansk

I årevis har forskning i optisk kommunikation været drevet af kravet om øget kapacitet, reducerede udgifter samt stigende energi effektivitet. For at undgå et kapacitetsnedbrud, skal designet af fremtidens systemer konstant opgraderes. Dette dikterer at nye avancerede digitale signalbehandlings (DSB) systemer skal udvikles for at sikre at kravene dikteret af næste generations kohærente optiske systemer kan opfyldes.

I denne sammenhæng vil bidragene præsenteret i denne afhandling overordnet være relateret til DSB møntet på kohærente optiske kommunikations systemer og mere specifikt adresserer (a) klok gendannelse, (b) sender/modtager kalibrering samt (c) fase gendannelse af bærebølgen (CPR). Hvad angår (a) så bliver der ved studiet af et komplet digital klok gendannelse præsenteret nye originale bidrag. Numeriske undersøgelser af ydeevnen af både polarisations sammenflettede (PMD) og rummelig sammenflettede (SDM) systemer bliver præsenteret. I (b) bliver unikke nye metoder til kalibrering af både sender og modtager demonstreret. I sender siden bliver en applikation af den såkaldte kooperative ko-evolutionære generisk algoritme (CC-GA) diskuteret. Det nye bidrag består både af en kalibrering af tidsforvridningen mellem elektriske komponenter, bias spændinger og amplitude ubalancer, samt introduktion af en ny parameter som kan benyttes til kalibrering af senderen. Desuden bliver en estimeringsparameter til optisk kohærente modtagere som kombinerer kromatisk dispersion (CD) og tidsforvridning demonstreret. Evaluering af begge teknikker er foretaget både numerisk og eksperimentelt. Afslutningsvis vil der for emnet (c) blive foreslået en ny algoritme baseret på en principal komponent analyse (PCA) der medfører en hardware effektiv CPR. Denne metode er blevet sammenlignet med state-of-the-art teknikker både ved hjælp af simuleringer samt eksperimenter, og den nye foreslåede algoritme yder langt bedre både hvad angår beregningsmæssig kompleksitet samt ydeevne generelt.

Acknowledgements

First of all, I would like to thank my parents, my family and specially my Wife, Camila, for all the support they gave me over the time I was here in Denmark.

I would like to thank Darko Zibar for his guidance during the time of this Ph.D. project.

I would like to thank my friend Edson Porto da Silva, who was the first person encouraging me to applying for a Ph.D. position at DTU Fotonik.

Thanks to Francesco Da Ros and Molly Piels for all I have learned, the help I got from them, and for a productive cooperation, which also included many dinners and liters of coffee.

Talking about coffee, I would like to thanks my "340" friends Peter, Frederik, Shajeel, Ayman, Beatrice, Daniele, Lisbeth, Davide, Mads, Pierre, Ahmed, Leif, Michael, Mikkel, Kjeld and the others that may be missing now but will be included in the final version of this thesis.

Thanks to my colleages and friends, the "DSP cowboys": Simone, Rasmus, Stefano, Stenio, Martin and Nicola.

Thanks to Anders Clausen for letting me be part of the Experimental Course in Optical Communications team and for helping me many times with the Danish language.

I am grateful to Alan, Faisal, Remi and all the other people from The Hong Kong Polytechnic University, who kindly received me as visiting Ph.D. student during my external stay in the Spring of 2018. I really enjoyed the discussions on Machine Learning for Optical Communication Systems and the opportunity to learn new ways to approach research problems. Furthermore, the Hong Kong food is something that I am going to miss.

Moreover, I would like to thank the High Speed Optical Communications (HSOC) and the Machine Learning in Photonics Systems (MLPS) groups. It

was my pleasure to be part of this really nice, diverse, and motivating work environment.

Ph.D. Publications

The following publications have resulted from this Ph.D. project.

Contributions to peer-reviewed conferences

[C1] **J. C. M. Diniz**, F. Da Ros, R. T. Jones, and D. Zibar, "Time skew estimator for dual-polarization QAM transmitters", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Gothenburg, Sweden, Sep. 2017, paper P1.SC3.41 (cited on pages 4, 50, 51).

[C2] **J. C. M. Diniz**, M. Piels, and D. Zibar, "Performance evaluation of clock recovery for coherent mode division multiplexed systems", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, CA, USA, Mar. 2017, paper W2A.50. DOI: 10.1364/ofc.2017.w2a.50 (cited on pages 4, 23, 27, 44, 46).

[C3] **J. C. M. Diniz**, E. P. da Silva, M. Piels, and D. Zibar, "Joint IQ skew and chromatic dispersion estimation for coherent optical communication receivers", in *Signal Processing in Photonic Communications (part of Advanced Photonics Congress)*, Vancouver, Canada, Jul. 2016, paper SpTu2F.2. DOI: 10.1364/SPPCOM.2016.SpTu2F.2 (cited on pages 4, 21, 23, 78, 80–83).

[C4] S. Savian, **J. C. M. Diniz**, A. P. T. Lau, F. N. Khan, S. Gaiarin, R. Jones, and D. Zibar, "Joint estimation of IQ phase and gain imbalances using convolutional neural networks on eye diagrams", in *Conference on Lasers and Electro-Optics*, San Jose, CA, USA, May 2018. DOI: 10.1364/cleo_si.2018.sth1c.3 (cited on page 21).

[C5] E. P. da Silva, F. Klejs, M. Lillieholm, S. Iqbal, M. P. Yankov, **J. C. M. Diniz**, T. Morioka, L. K. Oxenløwe, and M. Galili, "Experimental characterization of 10 x 8 GBd DP-1024QAM transmission with 8-bit

DACs and intradyne detection", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2018.

[C6] J. Wass, J. Thrane, M. Piels, **J. C. M. Diniz**, R. Jones, and D. Zibar, "Machine learning for optical performance monitoring from directly detected PDM-QAM signals", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Düsseldorf, Germany, Sep. 2016, pp. 1082–1084.

[C7] R. T. Jones, **J. C. M. Diniz**, M. P. Yankov, M. Piels, A. Doberstein, and D. Zibar, "Prediction of second-order moments of inter-channel interference with principal component analysis and neural networks", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Gothenburg, Sweden, Sep. 2017, paper W.1.D.4. DOI: 10.1109/ecoc.2017.8346176.

Articles in peer-reviewed journals

[J1] **J. C. M. Diniz**, F. Da Ros, E. P. da Silva, R. T. Jones, and D. Zibar, "Optimization of DP-M-QAM transmitter using cooperative coevolutionary genetic algorithm", *J. Lightw. Technol.*, vol. 36, no. 12, pp. 2450–2462, Jun. 2018. DOI: 10.1109/jlt.2018.2815347 (cited on pages 4, 10, 13, 26, 37, 51, 53–59, 61, 63–75).

[J2] **J. C. M. Diniz**, F. Da Ros, and D. Zibar, "Clock recovery challenges in DSP-based coherent single-mode and multi-mode optical systems", *Future Internet*, vol. 10, no. 7, Art. no. 59, Jun. 2018. DOI: 10.3390/fi10070059 (cited on pages 4, 23, 27, 29, 30, 35, 42–44, 46).

[J3] J. Thrane, J. Wass, M. Piels, **J. C. M. Diniz**, R. Jones, and D. Zibar, "Machine learning techniques for optical performance monitoring from directly detected PDM-QAM signals", *J. Lightw. Technol.*, vol. 35, no. 4, pp. 868–875, Feb. 2017. DOI: 10.1109/jlt.2016.2590989.

Unpublished works

[U1] J. C. M. Diniz, Q. Fan, S. M. Ranzini, F. N. Khan, F. Da Ros, D. Zibar, and A. P. T. Lau, *Principal component based technique for carrier phase recovery of square M-QAM systems*, to be submitted (cited on pages 85, 90, 91, 93, 96, 97, 102, 106).

Contents

1

Introduction

RECENT PROGRESS in information and digital communication technologies has been continuously driving deep transformations in how our society works [1, 2]. Reliable communications enable the development of high-bandwidth applications such as on-demand media streaming, cloud storage and distributed computing, as well as low-latency applications such as voice-over-IP, automatic stock trading and online gaming. One of the key factors of these changes is the development of communications through optical fibers, due to its high bandwidth and the low loss of modern fibers. Optical fibers are becoming ubiquitous as they move from the early applications in long and ultra-long haul intercontinental communications into even shorter distances such as data center networks [3].

The evolution of optical communication systems differs from other well known and broadly used types of communication systems. Due to strong bandwidth and power limitations, a wide range of sophisticated modulation, coding and signal detection techniques were developed for wired (copper) and wireless communication systems [4–6]. These methods relied on modulating the information into degrees of freedom such as the amplitude and the phase of a high frequency electrical signal, as well as a complex scheme of reception based on a local oscillator that would interfere with the received signal creating an amplified and coherent copy of the signal in a much lower frequency, easier to process. While some of these techniques have been adapted for optical communications systems during late 1980's and early 1990's with the purpose of reach extension [7–9], the coherent detection lost space, both in academia and industry, to the breakthroughs represented by the development of signal amplification by erbium-doped fiber amplifiers (EDFAs) [10], and techniques for wavelength division multiplexing (WDM)

[11, 12], solving temporarily power and bandwidth limitations, respectively. Together with the development of optical devices, these technologies enabled systems with up to 80 channels per fiber operating at data rates up to 40 Gb/s per channel and with a simple on-off keying (OOK) modulation technique, direct detection and chromatic dispersion (CD) compensated optically by dispersion compensating fibers (DCFs) or fiber Bragg gratings (FBGs) [13].

In the late 2000's, advances in microelectronics contributed to the development of high-speed and accurate analog-to-digital converters (ADCs) and digital-to-analog converters (DACs), allowing the comeback of the coherent detection in optical communications. It had no longer the main purpose of only signal amplification, but to allow for compensation of all the linear (and even some nonlinear) impairments in the digital domain through the employment of advanced digital signal processing (DSP) techniques [14–16]. The digital compensation of CD enabled substantial cost reductions by the elimination of passive optical signal processing components from the transmission link [17], while the digital compensation of polarization mode dispersion (PMD) enabled the utilization of signals employing polarization division multiplexing (PDM), which doubled the capacity of standard single-mode fibers (SMFs) [18–20]. Coherent detection also enabled the use of higher-order modulation formats such as quadrature phase shift keying (QPSK) and m-ary quadrature amplitude modulation (QAM) [21–24]. At present, field trials of a single-carrier 600 Gb/s signal, employing modulation formats up to 64QAM and symbol rates up to 69 GBd, were shown for metropolitan distances [25].

Since the advent of internet, the demand for digital information has been growing in an exponential fashion, and it is expected to continue growing, reaching 399 exabytes per month of overall IP traffic by 2022 [26]. This will be mainly driven by the improvement of already existing services, as well as emerging applications such as internet of things (IoT), smart cities and big data. In order to keep track of the ever-increasing demand, next generations of optical communication technologies need to improve data throughput reducing the cost per bit and increasing the energy efficiency. The paths towards these important objectives require the solution of a plethora of practical problems. On the one hand, low power and highly integrated optical devices are extremely necessary. On the other hand, advanced signal

processing tools and novel concepts are essential to improve the use of resources available in the optical networks.

The approaches to upgrade the capacity of the current optical communication systems relate to the enhancement of spectral efficiency and rely on increasing the number of levels in one or more of the dimensions available for modulation and multiplexing of electromagnetic waves [22, 27, 28]. Increasing the modulation order is typically associated to fiber nonlinearities that appear due to the stronger launch power needed to counterbalance the amplification noise, providing a better signal-to-noise ratio (SNR) [29, 30]. It also imposes other practical limitations as the requirement for more precise DACs and ADCs, lasers with lower linewidths [31] and better control of the modulator voltages [32], while the transceiver calibration becomes more stringent [33–35].

Alternatively, space can be added as another multiplexing dimension in spatial division multiplexing (SDM) systems, multiplying the capacity of the optical channel by the number of spatial channels being used [36–39]. These systems rely in highly complex multiple-input multiple-output (MIMO) signal processing techniques at the receiver-end in order to compensate for the random coupling between cores and/or modes [40–42]. Therefore, the success of these systems greatly depends on the design of reliable multi-core fibers (MCFs) and multi-mode fibers (MMFs), SDM-compatible amplification and routing schemes, and the feasibility of the advanced DSP techniques that are needed.

In this context, the aim of this thesis is to advance the field of coherent optical communications, to explore the feasibility of current generation algorithms in next generation systems and to investigate, create and demonstrate new, state-of-the-art algorithms for next generation DSP-based coherent optical transceivers. Furthermore, methods from artificial intelligence (AI) and machine learning areas are also investigated.

1.1 Structure and scope

A significant portion of the advancements in the coherent optical communication field can be associated to the progress of DSP subsystems. Critical functionalities such as synchronization, channel equalization and decoding are performed effectively and efficiently with DSP. Moreover, in recent years, AI and machine learning have gained more and more attention in the optical

communications community [43–48] because of their endless application possibilities and apparently lower complexity in solving some problems.

The subject matter of this thesis entails advanced DSP techniques to be used in next generation coherent optical communication transceivers.

This thesis interpolates material from three conference papers [C1–C3], two journal papers [J1, J2] and one still unpublished journal submission, researched by the author within the duration of the Ph.D. project. This thesis is structured as follows.

Chapter 2 provides a brief theoretical foundation for the basics of coherent optical communications that will be necessary in the remainder of the thesis. Transmitter and receiver concepts are discussed and a linear fiber model comprising random spatial coupling is presented. Furthermore, receiver-side DSP blocks are discussed.

Chapter 3 investigates the performance and feasibility of clock recovery in both PDM and mode division multiplexing (MDM) systems. MDM is a particular case of SDM in which few-mode fibers (FMFs) are used. Furthermore, the foundations of clock tone amplitude, which will be used in the next chapter is presented.

Chapter 4 proposes a method for joint optimization of transmitter inter-polarization (XY) and in-phase/quadrature (IQ) time skews, amplitude mismatches, and bias voltages. The method is based on a cooperative coevolutionary genetic algorithm (CC-GA) with fitness functions extracted from a reference directly-detected QAM signal generated at the transmitter. It works as a calibration method, and it can find the parameters that generate the best possible constellation.

Chapter 5 proposes a low-complexity scanning method for joint estimation of receiver IQ time skews and CD. The proposed method is experimentally demonstrated and show good time skew and CD accuracy.

Chapter 6 proposes a low complexity, modulation order independent, non-data-aided (NDA), carrier phase recovery method for synchronous decoding of arbitrary QAM constellations, suitable for parallel implementation. The proposed method is based on principal component analysis (PCA) and outperforms the well known and widely used blind phase search (BPS) algorithm at low SNR values, showing much lower cycle slip rate.

Finally, Chapter 7 concludes this thesis and proposes ideas for future works.

2

Fundamentals of Coherent Optical Communication Systems

A SHORT SUMMARY of coherent optical communication systems is presented. The concepts shown here emphasize the main topics present in the subsequent chapters of this thesis. Detailed discussion on the subject can be found in the references [15, 19, 22, 49–53].

2.1 Coherent optical communication systems

A typical communication system, as shown in Figure 2.1, can be described by three basic elements: the transmitter, who sends a message; the receiver, who receives a message; and the channel, the physical media connecting them [54]. The main purpose of the transmitter is to convert the message signal from an information source into a format that is possible to transmit through the channel. When this signal is propagated, it is distorted due to channel imperfections, interference from other sources, and additive noise. Thus, the received signal is a corrupted version of the transmitted signal. Therefore, the receiver acts on the received signal in order to reconstruct a recognizable form of the original message that was generated by the information source.

Optical communication systems are particular systems where the information is carried from the transmitter to the receiver using light. Consequently, the transmitter needs to encode the message into an optical signal, and the receiver needs to be able to extract the information carried by the optical carrier, compensating for possible distortions caused by channel propagation. Due to the immunity for electromagnetic interference, low attenuation, and

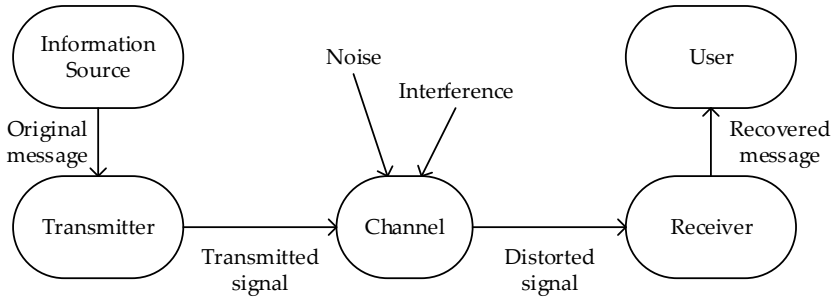

Figure 2.1.: Basic elements of a communication system.

high bandwidth, the optical fiber is the preferred channel for long distance optical communications. Although other media such as the air could be also used as a media for optical communications, in this thesis only the case where the optical fiber is the media is considered. Therefore, "optical communications" and "fiber-optic communications" are two terms that will be understood as synonyms.

The first implementations of optical communication systems uses intensity modulation and direct detection (IM-DD), encoding information on the amplitude of the optical signal, typically by encoding 1's and 0's into the presence or absence of light, respectively. These systems use photodetectors as receivers that convert the received optical signal amplitude into electrical current. They get deteriorated from propagation impairments such as chromatic dispersion (CD) and polarization mode dispersion (PMD), which need to be compensated before optical-to-electrical signal conversion.

In contrast, the use of a coherent front-end together with a local oscillator laser enables the recovery of the full electrical field of the received signal, *i.e.*, all the available orthogonal dimensions can be used to encode the data. This is particularly important because it allows the compensation of propagation impairments through digital signal processing. Such systems are shown in Figure 2.2. At the transmitter, the information is encoded with forward error correction (FEC) codes, and the bits are mapped into a modulation format. Pulse shaping controls the spectrum to increase spectral efficiency and reduce nonlinear impairments. Afterwards, pre-equalization corrects for the overall response of the digital-to-analog converter (DAC) and electrical drivers, and possibly for the effects the signal could face during propagation through the optical fiber. The signal is then converted to the electrical domain by a DAC and fed to an external optical modulator, in which the

signal will be modulated in an optical carrier provided by a continuous-wave (CW) laser. The signal is propagated through the optical channel, which consists basically of optical fiber spans, and possibly optical amplifiers and filters, until it reaches the receiver. There, the signal is converted to electrical baseband by beating with a local oscillator (LO) laser and photodetected. Then, the signal is sampled to the logical domain by a analog-to-digital converter (ADC). Digital signal processing (DSP) functions of equalization, synchronization, demodulation and error correction are performed in order to compensate for the channel propagation impairments and retrieve the information originally sent.

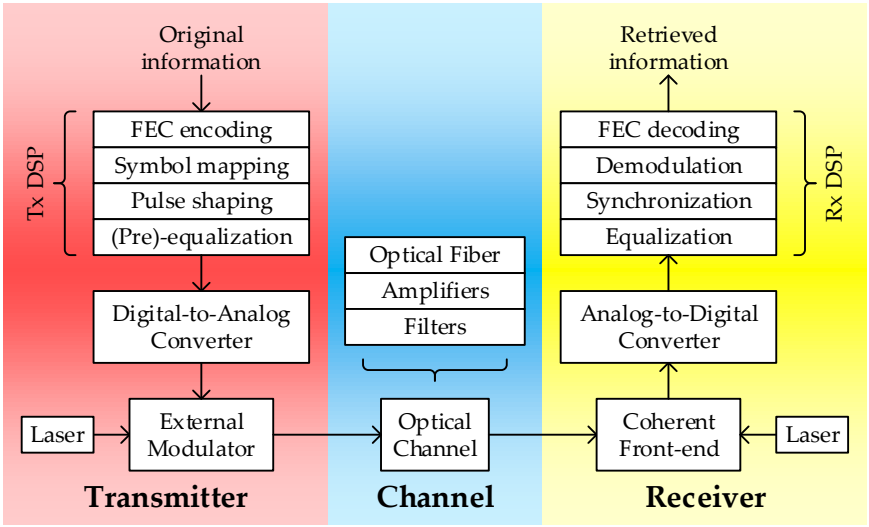

Figure 2.2.: Block diagram of a DSP-based coherent optical communication system.

2.2 Transmitters

2.2.1 Continuous-wave lasers

A laser is a device that emits light through a process of optical amplification based on the stimulated emission of electromagnetic radiation [51]. Therefore, a CW laser is a laser that produces a continuous output beam in a defined wavelength.

Considering phase and amplitude noises, the output electrical field, $\vec{s}_{CW}(t)$, of a CW laser can be written as

$$\vec{s}_{CW}(t) = \sqrt{P_s + \delta P(t)} \exp\left\{ j\left[\omega_s t + \phi_s + \phi_{ns}(t) \right] \right\} \vec{p}_s, \tag{2.1}$$

where P_s means the mean output power of the laser, $\delta P(t)$ is the power fluctuation induced by the spontaneous emission photons, ω_s is the laser angular frequency, ϕ_s is the starting phase of the laser, $\phi_{ns}(t)$ is the phase noise, and $\vec{p}_s$ is a bi-dimensional vector indicating the direction of polarization in space. For simplicity, only linear polarization is considered.

The phase noise can be modelled by a Wiener process [55],

$$\phi_{ns}(t) = \phi_{ns}(t - T) + \Delta\phi_{ns}(t), \tag{2.2}$$

where T is the time period between consecutive samples and $\Delta\phi_{ns}$ is the random phase increment that can be modelled by a zero-mean Gaussian process with variance given by

$$\mathbb{E}\left\{ \Delta\phi_{ns}^2 \right\} = 2\pi\Delta\nu|T|, \tag{2.3}$$

where $\Delta\nu$ is the laser linewidth (LW).

2.2.2 Digital modulation formats

The transmitter maps a sequence of binary digits (the original information) into a set of corresponding signal waveforms. With the number of possible waveforms being 2^b, where b is the number of bits. Each of these waveforms may differ in parameters such as phase, frequency, amplitude, polarization state, time, and spatial state [28], and can be described by a vector of orthogonal basis [56]. For example, an intensity modulated signal such as on-off keying (OOK), can be represented by a 1-dimensional vector representing amplitude, while a signal modulated with quadrature amplitude modulation (QAM) can be represented by a 2-dimensional vector (or a complex number) representing the in-phase and quadrature components of the signal, in which its angle mean the phase of the signal and its norm is the amplitude of the signal. Although polarization multiplexed (PM) signals can be represented by 4-dimensional vectors representing the in-phase and quadrature of each of the orthogonal polarization, they are typically represented by two separated

2-dimensional vectors due to the difficulty to graphically represent a 4-dimensional space. Figure 2.3 shows some examples of modulation formats used in optical communications.

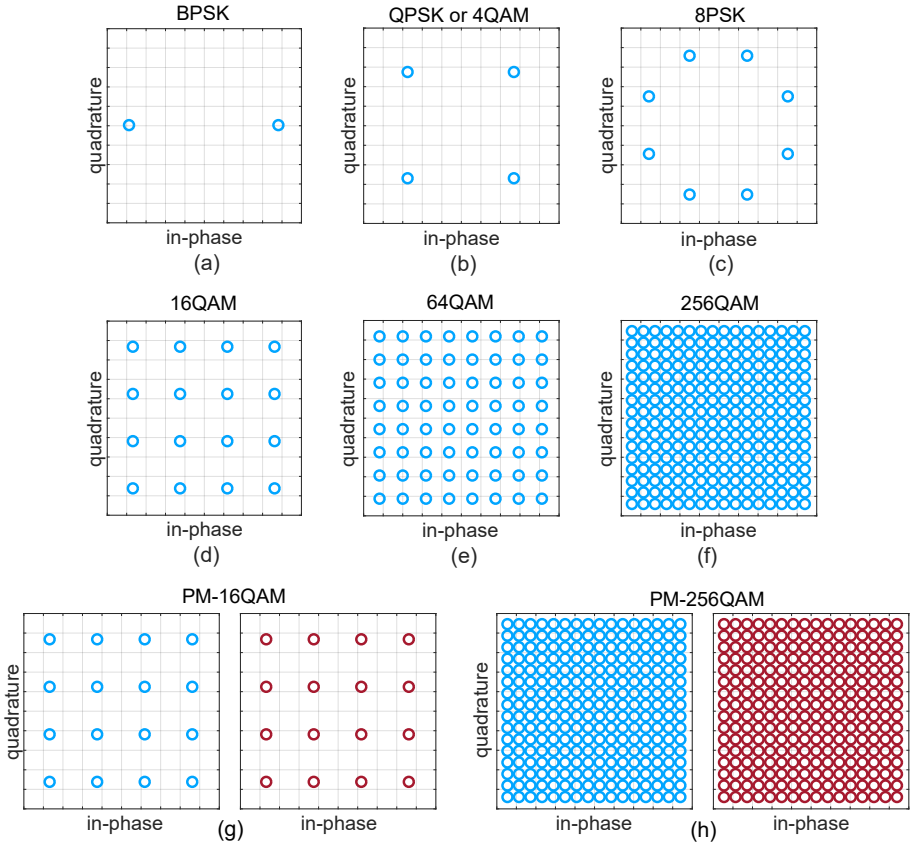

Figure 2.3.: Example of modulation formats commonly used in coherent optical communications. (a) binary phase-shift keying (BPSK), 1 bit/symbol; (b) quadrature phase shift keying (QPSK), or alternatively, 4-*ary* QAM (4QAM), 2 bit/symbol; (c) 8-*ary* phase-shift keying (PSK) (8PSK), 3 bit/symbol; (d) 16QAM, 4 bit/symbol; (e) 64QAM, 6 bit/symbol; (f) 256QAM 8 bit/symbol; (g) PM-16QAM, 8 bit/symbol; (h) PM-256QAM, 16 bit/symbol.

2.2.3 External modulators

An in-phase/quadrature (IQ) modulator is commonly employed to generate high-order modulation formats in high symbol rate optical communication systems. This kind of modulator is typically based on two parallel

Mach-Zehnder modulators (MZMs) embedded inside another MZM [52] (Figure 2.4).

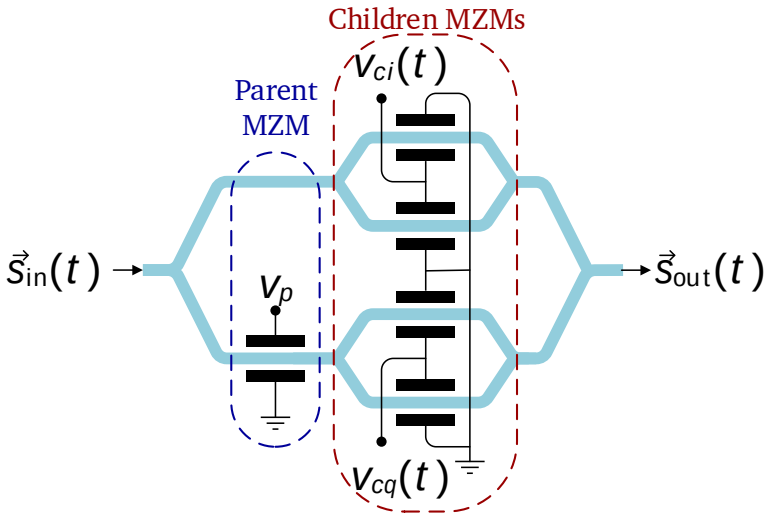

Figure 2.4.: In-phase/quadrature modulator. Based on [J1].

The two inner MZMs are commonly labeled as children modulators while the outer MZM is known as the parent modulator. They are used, respectively, to generate the in-phase (I) and quadrature (Q) components of the signal and to control the phase between these components, orthogonalizing them. If attenuation is omitted, the modulated signal coming from the IQ modulator, $\vec{s}_{\text{out}}(t)$, relates to its input, $\vec{s}_{\text{in}}(t)$, by

$$\vec{s}_{\text{out}}(t) = \frac{\vec{s}_{\text{in}}(t)}{2}\left[\cos\left(\frac{\pi v_{\text{ci}}(t)}{2V_\pi}\right) + \exp\left(j\frac{\pi}{V_\pi}v_p\right)\cos\left(\frac{\pi v_{\text{cq}}(t)}{2V_\pi}\right)\right], \quad (2.4)$$

where V_π is the halfwave voltage, v_p is the parent bias voltage, $v_{\text{ci}}(t)$ and $v_{\text{cq}}(t)$ are electrical modulating driving signals, and the indexes i and q indicate which are related to the in-phase and quadrature components, respectively. The halfwave voltage is the voltage necessary to phase-delay an optical signal by π rad. The parent bias is the voltage which controls the phase between I and Q components, being responsible to guarantee a $\pi/2$ phase shift orthogonality, ideally set at $V_\pi/2$. $v_{\text{ci}}(t)$ and $v_{\text{cq}}(t)$ are given by

$$\begin{aligned} v_{\text{ci}}(t) &= \dot{v}_{ci} + \ddot{v}_{ci}(t), \\ v_{\text{cq}}(t) &= \dot{v}_{cq} + \ddot{v}_{cq}(t), \end{aligned} \quad (2.5)$$

where $\ddot{v}_{ci}(t)$ and $\ddot{v}_{cq}(t)$ are the electrical waveforms carrying information, and $\dot{v}_{ci}$ and $\dot{v}_{cq}$ are the children bias voltages, also known as operation points.

A polarization diversity modulator can be used to produce a PM signal. This kind of modulator, shown in Figure 2.5 can be realized in an integrated manner by two IQ modulators and a 90° polarization rotator [57]. Each of the inner IQ modulators has its bias voltage inputs to control them. Then,

$$\begin{aligned}\vec{s}_{\text{out}}(t) = \frac{\vec{s}_{\text{in}}(t)}{4}\left\{\left[\cos\left(\frac{\pi v_{\text{cix}}(t)}{2V_\pi}\right)+\exp\left(j\frac{\pi}{V_\pi}v_{\text{px}}\right)\cos\left(\frac{\pi v_{\text{cqx}}(t)}{2V_\pi}\right)\right]\vec{x}\right.\\ \left.+\left[\cos\left(\frac{\pi v_{\text{ciy}}(t)}{2V_\pi}\right)+\exp\left(j\frac{\pi}{V_\pi}v_{\text{py}}\right)\cos\left(\frac{\pi v_{\text{cqy}}(t)}{2V_\pi}\right)\right]\vec{y}\right\},\end{aligned} \tag{2.6}$$

where $\vec{x} = \vec{p}_s$ and $\vec{y} = \begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\vec{p}_s$ are orthogonal polarization direction vectors. The bias voltages and the signals are distinguished in relation to which IQ modulator they relate by the inclusion of the additional indexes x and y, indicating the horizontal polarization (X) and the vertical polarization (Y), respectively.

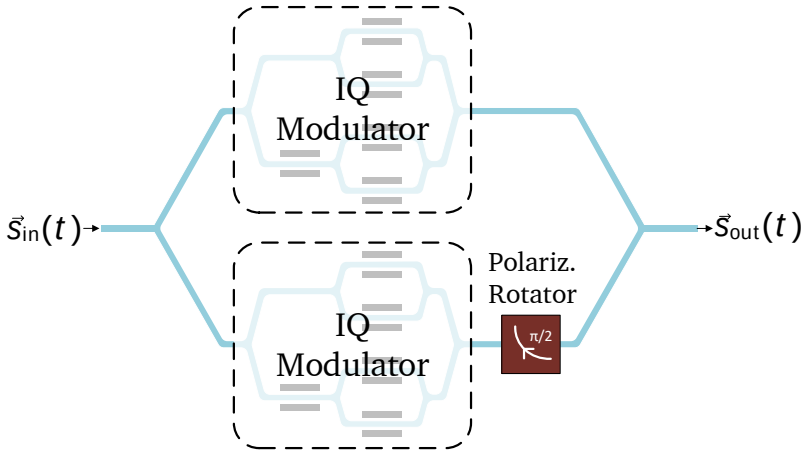

Figure 2.5.: Polarization-diversity optical modulator.

2.2.4 Signal generation with high-order modulation formats

In an ideal scenario, the parent bias should be set as $v_p = V_\pi/2$, with the driving signals swinging with peak-to-peak voltages of $2V_\pi$ around the children bias voltages of $\dot{v}_{ci} = \dot{v}_{cq} = -V_\pi$, in order to generate a standard single polarization QPSK signal. In this manner, the output signal would have the highest power efficiency and the maximum extinction ratio. The electrical field transfer function of the children MZM for this configuration is shown in Figure 2.6(A-a), and the resulting constellation is shown in Figure 2.6(B-a).

For higher-order QAM modulation formats, the sinusoidal transfer characteristics of the IQ modulator generate an undesired nonlinear distortion on the output signal if driving peak-to-peak voltages of $2V_\pi$ are applied. Figure 2.6(b) shows the electrical field transfer function and the resultant constellation for a 16QAM signal under these conditions.

Three approaches are typically used in order to avoid this non-linearity. The first approach would be to generate the components with low peak-to-peak voltages, so the modulator would be operating in the quasi-linear region. Figure 2.6(c) shows an example of this approach. Using the approximation $\cos(\theta + \pi/2) \approx \theta$, when $\theta \ll 1$, and rewriting $s_i(t) = \pi\left[\dot{v}_{ci}(t) + V_\pi\right]/2V_\pi$, $s_q(t) = \pi\left[\dot{v}_{cq}(t) + V_\pi\right]/2V_\pi$, and $v_p = V_\pi/2$, then, for small $s_i(t)$ and $s_q(t)$ peak-to-peak voltages,

$$\vec{s}_{\text{out}}(t) \propto \vec{s}_{\text{in}}(t)\left[s_i(t) + js_q(t)\right]. \tag{2.7}$$

The second approach is to digitally pre-distort the driving signals in the sense that the input signals revert the sinusoidal transfer function of the optical modulator. Figure 2.6(d) show an example of this approach. Considering a parent bias voltage of $v_p = V_\pi/2$ and children bias voltage of $\dot{v}_{ci} = \dot{v}_{cq} = -V_\pi$, and the pre-distorted signals as

$$v_{\text{ci}}(t) = \frac{2V_\pi}{\pi}\arcsin\left\{s_i(t)\right\}, \tag{2.8}$$

$$v_{\text{cq}}(t) = \frac{2V_\pi}{\pi}\arcsin\left\{s_q(t)\right\}, \tag{2.9}$$

where $s_i(t)$ and $s_q(t)$ are the desired signal components at the IQ modulator output, the same output signal as Eq. (2.7) would be achieved. The

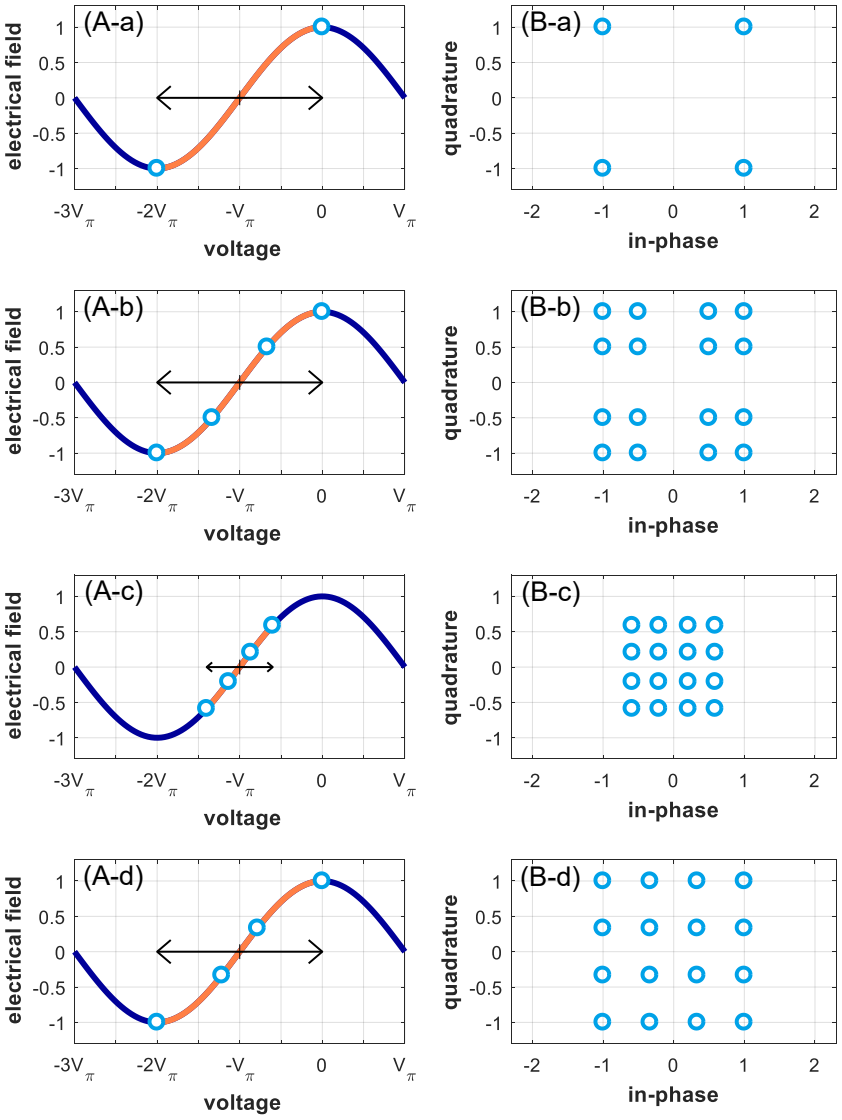

Figure 2.6.: Example of single-polarization QAM generation. (A) Electrical field transfer functions. (B) Constellation diagrams. (a) QPSK with maximum swing voltage. (b) 16QAM with maximum swing voltage. (c) 16QAM with smaller swing voltage operating in the quasi-linear region. (d) Pre-distorted 16QAM with maximum swing voltage. Based on [J1].

advantage of the first approach is that no digital pre-distortion is needed, but it would increase the insertion loss in the modulator. On the other hand, the second approach would be more efficient but the electrical drivers would suffer from nonlinearities due to the high output voltage. A possible

third approach would be to generate the signal at intermediate peak-to-peak voltages, avoiding the nonlinear region of the electrical drivers but using pre-distortion to compensate the nonlinearity from the modulator, benefiting from both approaches.

Using a polarization diversity modulator, a PM-QAM signal linearized by either pre-distorting the electrical input signal or operating in the quasi-linear region, is

$$\vec{s}_{\text{out}}(t) \propto \vec{s}_{\text{in}}(t)\left\{\left[s_{ix}(t)+js_{qx}(t)\right]\vec{x}+\left[s_{iy}(t)+js_{qy}(t)\right]\vec{y}\right\}. \tag{2.10}$$

2.3 The fiber-optic channel

An optical fiber is a dielectric cylindrical waveguide used to transmit light between its two ends. It is typically made of silica (SiO_2), with its transversal dimensions in the order of magnitude of a human hair. These fibers are commonly used for communications due to the lower loss and higher bandwidth in comparison with other forms of communication. The fiber is divided into two regions, the core and the cladding. Each region is doped with different chemical elements during the fabrication process in order to force the refractive-index of the core, n_1, to be slightly higher than the refractive-index of the cladding, n_2, therefore, making it possible for the light to be confined inside the core of the optical fiber, accordingly to the total reflection principle [51].

2.3.1 Fiber modes

An optical mode is a particular solution to the wave propagation equation that satisfies the appropriate boundary conditions and has the property that its spatial distribution does not change during propagation [51].

As the transverse refractive index profiles of the majority of commercial optical fibers are radially symmetric, and their small refractive index contrast between core and cladding, it could be assumed the propagating modes to be weakly guided. Thus, the calculation of the fiber modes can be greatly simplified and the modes can be described as linearly polarized (LP) modes [58].

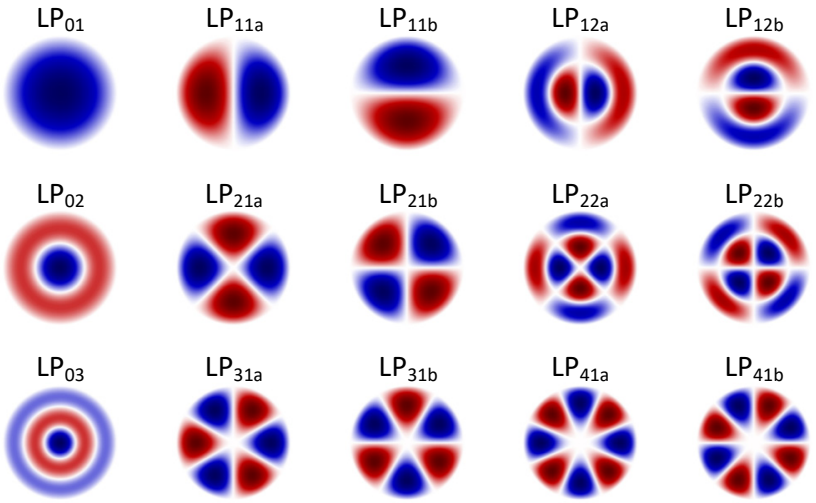

Figure 2.7.: Electrical field amplitude profile representation of all the 15 LP modes of a step-index fiber with $V = 7.5$. The blue and red colors mean positive and negative values of the electrical field, respectively.

The normalized frequency, V, is an important dimensionless parameter commonly used for determining the number of modes that an optical fiber can support. For an step-index fiber, V is given by

$$V = \frac{4\pi a}{\lambda}\sqrt{n_1^2 - n_2^2}, \tag{2.11}$$

where λ is the vacuum wavelength, and a is the radius of the fiber core. If a fiber is designed so that $V < 2.405$, it will only support the LP_{01} mode, and therefore this fiber is called single-mode fiber (SMF). Alternatively, if $V > 2.405$, the fiber will support more modes, being called multi-mode fiber (MMF). In this context, a few-mode fiber (FMF) is a MMF that supports only a small number of modes. Tab. 2.1 shows the cutoff frequencies for the first 15 LP modes [58–61], and Figure 2.7 shows the electrical field amplitude profiles of all the modes of a fiber with normalized frequency $V = 7.5$. It is important to note that each of the LP modes support two orthogonal polarization states (degenerate modes). Therefore, the number of degrees of freedom in a fiber is equal to two times the number of modes. This is particularly important for the analyses presented in Chapter 3.

LP modes	Cutoff V
LP_{01}	0
$LP_{11a} + LP_{11b}$	2.405
$LP_{21a} + LP_{21b} + LP_{02}$	3.832
$LP_{31a} + LP_{31b}$	5.136
$LP_{12a} + LP_{12b}$	5.520
$LP_{41a} + LP_{41b}$	6.380
$LP_{22a} + LP_{22b} + LP_{03}$	7.016

Table 2.1.: Cutoff frequencies for linearly polarized modes. The LP modes of the left column are supported, together with the above modes for the V-values greater than the right column.

2.3.2 Dispersion

Generically, dispersion is any phenomenon caused by distinct components of a signal travelling through a media with different velocities. This can broaden the signal being transmitted causing inter-symbol interference (ISI) and can be caused by different reasons. In this subsection, the effects and the main causes of the different kinds of dispersion in optical fibers will be introduced and discussed.

2.3.2.1 Chromatic dispersion

Chromatic dispersion is a particular type of dispersion in which the components that travel with different velocities are the spectral components of the optical signal. A given spectral component at angular frequency ω arrives at the output of the fiber with group velocity, v_g, defined as

$$v_g = \left(\frac{\partial \beta}{\partial \omega}\right)^{-1} = \frac{1}{\beta'}. \tag{2.12}$$

The time delay difference between two spectral components of the signal, ΔT, is given by

$$\Delta T = \frac{\partial T}{\partial \omega}\Delta\omega = \frac{\partial}{\partial \omega}\left(\frac{L}{v_g}\right)\Delta\omega = L\frac{\partial^2 \beta}{\partial \omega^2}\Delta\omega = L\beta''\Delta\omega, \tag{2.13}$$

This actually has two reasons. The first is the frequency dependence of the refractive index due to the material (known as material dispersion). And

the second is that the wave's phase velocity depends on its frequency due to the waveguide geometry. In terms of transfer function, $H_{\mathrm{CD}}(\omega)$, neglecting noise and nonlinearities, the output signal is related to the input signal due to CD by

$$\vec{S}_{\mathrm{out}}(\omega) = H_{\mathrm{CD}}(\omega)\vec{S}_{\mathrm{in}}(\omega), \tag{2.14}$$

$$H_{\mathrm{CD}}(\omega) = \exp\left\{-jD_{\mathrm{C}}\frac{\lambda^2 L}{4\pi c}\omega^2 + j\frac{D_{\mathrm{S}}\lambda^4 L}{24\pi^2 c^2}\omega^3\right\}, \tag{2.15}$$

where D_{C} is the CD parameter, D_{S} is the CD slope, related to third-order CD, L is the length of the fiber, and λ is the central wavelength of the signal.

The parameters D_{C} and D_{S} are related to the propagation constant, β, by

$$D_{\mathrm{C}} = -\frac{2\pi c}{\lambda^2}\frac{\partial^2\beta}{\partial\omega^2}, \tag{2.16}$$

$$D_{\mathrm{S}} = \frac{4\pi^2 c^2}{\lambda^4}\frac{\partial^3\beta}{\partial\omega^3} + \frac{4\pi c}{\lambda^3}\frac{\partial^2\beta}{\partial\omega^2}. \tag{2.17}$$

It is important to note that for a MMF, as the propagation constant is different for each mode, they will have different CD values.

2.3.2.2 Polarization-mode dispersion and modal dispersion

SMFs support two polarization modes. If the fiber core is perfectly symmetric, the refractive index is exactly the same in the two orthogonal polarization directions. However, asymmetries in the fiber core lead to birefringence, thus, the signal velocities in each of the orthogonally polarized components of the fundamental mode will be different. This is commonly referred to as PMD.

If the birefringence along the fiber is ideally constant, then the differential group delay, ΔT, is given by

$$\Delta T = L\left|\frac{1}{v_{\mathrm{gx}}} - \frac{1}{v_{\mathrm{gy}}}\right| = L\left|\beta'_{\mathrm{x}} - \beta'_{\mathrm{y}}\right| = L\Delta\beta', \tag{2.18}$$

where v_{gx} and v_{gy} are the group velocity in the vertical and horizontal polarization directions and $\Delta\beta'$ is related to the fiber birefringence. However,

due to random coupling between modes, induced by random perturbations of birefringence, and due to slow variations in the polarization states over time, the average time delay between polarization states is given by

$$\mathbb{E}\{\Delta T\} = D_P\sqrt{L}, \tag{2.19}$$

where D_P is the PMD parameter, given by

$$D_P = \Delta\beta'\sqrt{h}, \tag{2.20}$$

where h is the correlation length, *i.e.*, the length over which the complex polarization fields remain correlated.

Similarly, in MMFs the difference in the propagation constant between modes will result in the spread of the optical pulse in time. This spread will increase linearly with the distance for short length fibers operating in weak coupling and with the square-root of the distance for long fibers operating in the strong coupling regime. This is commonly referred to as modal dispersion (MD).

2.3.3 Optical loss and gain

The power of an optical waveform decreases while travelling through an optical media due to absorption, scattering, dirt in the connectors, fiber splicing, etc [51]. The attenuation, α, is defined in units of decibels (dB). The transfer function considering only attenuation is given by

$$\vec{S}_{\text{out}}(\omega) = 10^{-\alpha/20}\vec{S}_{\text{in}}(\omega). \tag{2.21}$$

Some passive optical components can exhibit loss that depends on the polarization or on the mode, these are referred as polarization-dependent loss (PDL) and mode-dependent loss (MDL), respectively.

The losses in optical fibers are compensated with optical amplifiers. These amplifiers are devices that transfers the energy from pump lasers into the signal. They are generally made with fibers with some kind of doping material, being the most common the erbium-doped fiber amplifiers (EDFAs). Alternatively, stimulated Raman scattering can also be used for optical signal

amplification. The gain, G, is also measured in units of dB, and the transfer function considering only gain is given by

$$\vec{S}_{\text{out}}(\omega) = 10^{G/20}\vec{S}_{\text{in}}(\omega). \tag{2.22}$$

2.3.4 Spatial multiplexers and demultiplexers

In mode division multiplexing (MDM) systems, the signal coming from the transmitter needs to be multiplexed into the MMF for the transmission and demultiplexed after it. This can be done by a photonic lantern [62].

The photonic lanterns are low-loss devices used to either multiplex signals into a MMF or demultiplex signals from a MMF. They are made from adiabatically merging several single-mode cores into one multimode core. Thus, providing an interface between multi-mode and single-mode systems, where the precise optical mapping is unimportant [62], or even desirable [63].

2.4 Coherent optical receivers

In principle, a coherent receiver is a set of devices that tries to convert all the orthogonal information present in an optical signal wave into the electrical domain by beating it with a waveform generated in a LO laser. Thus, allowing it for further DSP treatment for proper extraction of the information bits.

This section will shortly introduce the components of a coherent front-end for optical communications, as well as the DSP used to properly recover the signal.

2.4.1 Photodetector diode

A photodetector diode (PD) is an optoelectronic device that converts the power intensity of an input optical signal, $S_{\text{in}}(t)$, into an output electrical current, $i_{\text{out}}(t)$, given by

$$i_{\text{out}}(t) = R\left|\vec{s}_{\text{in}}(t)\right|^2 + i_s(t) + i_T(t), \tag{2.23}$$

where R is the PD responsivity and $i_s(t)$ and $i_T(t)$ are the shot and thermal noises, respectively [51].

PDs are used in direct detection systems to extract the information carried in the amplitude of optical signals.

2.4.2 Polarization diversity coherent receiver

A basic schematic of a polarization diversity coherent receiver is shown in Figure 2.8. In this type of receiver, a polarization multiplexed signal is converted into four lines of digital data, representing the in-phase and quadrature of each orthogonal polarization. For this, the incoming signals, both from the fiber and from the LO laser are polarization separated with polarization beam splitters (PBSs) and mixed in a pair of 90° optical hybrids. The hybrid is a device that allows the detection of the in-phase and quadrature components of an optical signal [52].

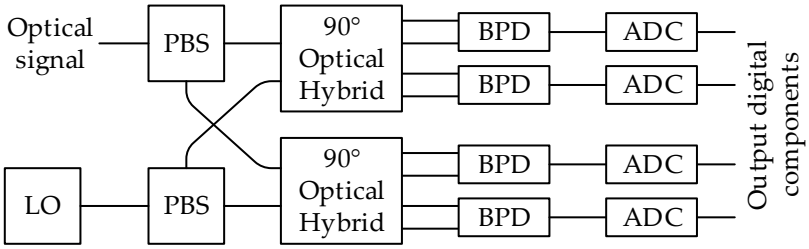

Figure 2.8.: Schematic of a polarization diversity coherent receiver. LO: local oscillator; PBS: polarization beam splitter; BPD: balanced photodetector; ADC: analog-to-digital converter.

If the two inputs of a polarization diversity coherent receiver are the received optical signal, $\vec{s}_{\text{in}}(t)$, and a signal from a CW laser acting as a local oscillator, $\vec{s}_{LO}(t)$, then, its outputs are

$$i_{out,ix}(t) = \Re\left\{\vec{s}_{LO}^{H}(t)\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\vec{s}_{\text{in}}(t)\right\}, \tag{2.24}$$

$$i_{out,qx}(t) = \Im\left\{\vec{s}_{LO}^{H}(t)\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\vec{s}_{\text{in}}(t)\right\}, \tag{2.25}$$

$$i_{out,iy}(t) = \Re\left\{\vec{s}_{LO}^{H}(t)\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\vec{s}_{\text{in}}(t)\right\}, \tag{2.26}$$

$$i_{out,qy}(t) = \Im\left\{\vec{s}_{LO}^{H}(t)\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\vec{s}_{\text{in}}(t)\right\}. \tag{2.27}$$

The output currents of the balanced photodetector diodes (BPDs) are sampled in an ADC bank for further DSP treatment of the signal.

2.4.3 Receiver DSP

The DSP at the receiver should perform three basic functions [3]:

Equalization: Its main purpose is to try to revert and mitigate the ISI accumulated by the signal along the transmission through the optical system. This is caused by linear fiber transmission effects such as CD and PMD (or MD), filtering from optical add-drop multiplexers (OADMs), and imbalances from the optical and electrical components composing the transmitter and the receiver front-ends. As some of these channel impairments are quasi-static and others are dynamic, it is beneficial to separate the equalization into static and dynamic filters.

Synchronization: Both the electrical and the optical oscillators from the transmitter and receiver are free-running. Synchronization is then needed to match the phase and frequency of these oscillators. This includes digital clock and timing recovery, as well as carrier frequency and phase recovery.

Decision and decoding: The recovered signal needs now to be converted into bits, retrieving the original information by decoding the signal with error correcting codes.

2.4.3.1 DSP chain

A common structure of signal processing in coherent optical receivers is shown in Figure 2.9.

Front-end corrections: It is necessary to compensate mismatches of amplitude, phase and timing between all the received signal components. Methods for estimation, compensation and/or mitigation of these impairments are included in this function [C3, 64, 65, C4].

CD estimation and static equalization: CD causes ISI that is typically long in time and has a time-invariant source, meaning that long static finite impulse response (FIR) filters can be used for compensation [15, 19, 66]. These are all-pass filters with length proportional to the amount of accumulated CD (therefore, scaling linearly with the transmission distance), and proportional to the square of the symbol rate. Although

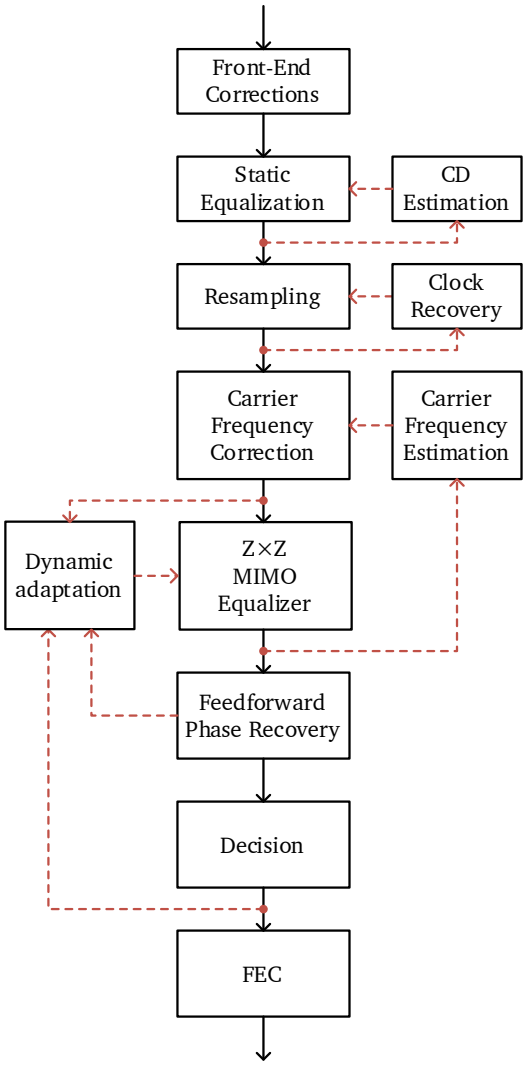

Figure 2.9.: Example of a DSP sequence at a coherent optical receiver. The black arrows represent the flow of information data, while the red arrows represent feedback and feedforward data for algorithms update. CD: chromatic dispersion; FEC: forward error correction; MIMO: multiple-input multiple-output; Z: number of spatial degrees of freedom.

FIR filters can be implemented in time-domain, long filters are preferably implemented in frequency-domain, using overlap-and-save or overlap-and-add methods. The complexity of these filters are a main concern in application specific integrated circuit (ASIC) design and can be responsible for up to 70% of area usage in a chip [67]. To reduce complexity,

operation in rates lower than 2 samples per symbol (undersampling) [68], and sub-band filtering are possible [69]. The exact accumulated CD is unknown prior to transmission and can be estimated during initialization [70]. In Chapter 5, a method for jointly estimate CD and time skew is proposed [C3].

Timing synchronization and resampling: The signal needs to be resampled to exactly 2 samples per symbol for further processing. A feedback control loop is responsible to estimate and compensate for the clock deviance and timing jitter from the transmitter and receiver clock sources [71], as well as the timing jitter imposed by equalization-enhanced phase noise (EEPN) [72–74]. The main algorithms for this task is the Gardner timing error detector (TED) [75], and others that are either equivalent [76] or based on it [77–80]. Parallel implementation is needed for operation in reduced ASIC clock frequency [81]. The theoretical formulation and analysis of these algorithms in both polarization division multiplexing (PDM) and spatial division multiplexing (SDM) systems is presented in Chapter 3 [C2, J2].

Dynamic MIMO equalization: For PDM systems, PMD causes a time-varying mixture of the two polarization components, and also a small dispersion compared to CD. Dynamic 2×2 multiple-input multiple-output (MIMO) equalizers composed of 4 FIR filters in a butterfly configuration are used to separate the information from the two polarizations while tracking state of polarization (SOP) changes, compensating for residual CD, and performing matched filtering. These filters are typically much shorter than the FIR filters for CD compensation. Equalizing the signal with 2 samples per symbol allows for compensation of residual timing mismatches [82]. The equalizer is updated by algorithms such as constant modulus algorithm (CMA), that is a fully blind algorithm that is better suitable for QPSK signals [83, 84], and frequently used for start-up process; radially directed equalizer (RDE), that is similar to CMA, but use multiple rings [85]; and least mean square (LMS), that is more precise than RDE but needs a signal without phase noise, thus requiring feedback from the phase estimator [22, 86].

For SDM systems, higher-order MIMO equalization is expected to be necessary [42, 87–92]. In these systems, the main concerns are the number of FIR filters, that can scale up with the square of the number of

spatial degrees of freedom, and the length of the filters that are expected to be much larger than in the PDM case.

Carrier frequency synchronization: Free-running lasers are used in coherent optical communications to reduce cost. Thus, an offset between the central frequency of the transmitter and receiver lasers will exist and be time-variant. A coarse estimation of the frequency offset, to reduce the complexity of further fine estimation and extend its range, can be made by analyzing the spectrum of the received signal [93, 94]. Fine frequency offset estimation can be made either in time domain [95] or in frequency domain [96, 97], extracting the information from the equalized signal.

Carrier phase recovery: The combination of laser phase noise from both the transmitter and receiver lasers needs to be compensated for proper decoding of the received signal. Although some feedback methods were proposed in literature [48, 85], the feedback latency and need for parallel processing make them prohibitive. Feedforward methods such as Viterbi-Viterbi (V&V) [98, 99] and blind phase search (BPS) [31] are commonly used in practical implementations. These methods usually suffer from cycle slips, in which $\pi/2$ phase rotations can lead to burst errors. They can be avoided by using FEC [100], pilot-tones [101], or filtering [102, 103]. In Chapter 6, a novel carrier recovery method based on principal component analysis (PCA), with reduced cycle slip rate and performance improved at low signal-to-noise ratio (SNR) values is proposed.

3

Clock Recovery for Coherent Single-Mode and Multi-Mode Optical Systems

THE DATA need to be processed synchronously in coherent optical communication systems. This can be made by transmitting through the optical fiber a pilot clock tone together with the signal. This would be an inefficient approach to clock recovery due to the increased overhead. Alternatively, extracting the clock information from the data itself requires that channel and transceiver impairments such as chromatic dispersion (CD), polarization mode dispersion (PMD) and time skew are at least partially compensated prior to clock recovery [104–108]. Therefore, the clock recovery process in these systems needs to be performed in the digital domain rather than in the analog domain.

Digital signal processing (DSP)-based coherent optical receivers started to be deployed targeting systems operating at 100 Gb/s per channel, employing polarization division multiplexing (PDM) and advanced modulation formats, and using single-mode fibers (SMFs) as transmission media [15, 109]. In order to continue meeting the ever growing demand for data rates, it is expected that future systems will exploit the space dimension for data transmission in spatial division multiplexing (SDM) systems [3].

SDM is a promising technology that may provide interface data rates for future 10 Tb/s per channel and beyond systems [110, 111]. It can be realized either by multiplexing the signal in several cores using multi-core fibers (MCFs) as the transmission channel, several modes using few-mode fiber (FMF), or even by a combination of the two using MCFs with multimode

cores. SDM systems using only FMFs as transmission channel are better known in literature as mode division multiplexing (MDM) systems. For systems employing MCF, a relatively low degree of crosstalk between cores can be achieved. In contrast, this is a challenging task for systems employing FMF, where a high degree of mode crosstalk may occur [63].

The scientific community has mainly prioritized research on the equalization schemes for mode mixing and delay spread mitigation [42, 87–92]. Strong mode coupling regime has been shown to be preferable in terms of equalizer impulse response length [63] and nonlinearity tolerance [112, 113]. Nevertheless, most experimental demonstrations have been in the weak or intermediate coupling regimes [90]. Very limited attention has been paid to the impact of mode mixing on clock recovery performance and its practicability in such systems.

In PDM systems, the clock recovery is generally performed before dynamic multiple-input multiple-output (MIMO) equalization by extracting a clock tone from the signal and them resampling it into the correct clock frequency. It is done before the MIMO equalization rather than after it in order to avoid the long feedback delay introduced by the equalization that would reduce the capacity to track fast clock jitter, reducing the overall performance. In MDM systems, it is expected that the computational complexity of MIMO equalization is even higher than in PDM systems, making it essential for the clock recovery be performed before MIMO equalization. If the clock tone cannot be extracted due transmission and transceiver front-end impairments, the remaining part of the DSP chain may fail [104, 106, 107]. Therefore, it is necessary to investigate the tolerance of clock recovery for MDM systems.

In PDM systems through SMFs, the clock tone vanishing can be mitigated by a simple polarization rotation method [108] and compensation of the transmitter time-skew between polarizations [J1]. However, in MDM systems, the combination of mode coupling and inter-modal dispersion in a FMF can significantly affect the performance of the timing synchronization methods, degrading them. In the strong coupling regime, even with low group delay (GD) spread, the clock tone completely vanishes, making clock recovery challenging.

A detailed numerical analysis of fully digital clock recovery in high-speed coherent optical communications systems is presented. The results consider both PDM signal transmission in a SMF and MDM signal transmission in

a FMF. The results and analyses presented here are based on the results originally presented in the author's work [C2, J2].

This Chapter is structured as follows. Section 3.1 presents a review of clock recovery in coherent optical receivers and it shows common algorithm implementations. Section 3.2 defines the propagation model for single-mode and multi-mode optical fibers, as well as a model considering inter-polarization (XY), inter-modal and in-phase/quadrature (IQ) time skews. From Sec. 3.4.1 to 3.4.5, a numerical analysis of the performance of clock recovery in single-mode fibers is presented. This analysis considers the impact of pulse shaping, residual chromatic dispersion, time skews (both XY and IQ) and PMD. Section 3.4.6 analyzes through simulations the clock recovery performance in MDM systems under different coupling regimes considering modal dispersion (MD). The Chapter is summarized in Section 3.5.

3.1 Clock recovery in coherent optical receivers

Synchronization between the clocks generated at the transmitter and the receiver is mandatory for the appropriate extraction of the transmitted information in communication systems [15]. The primary purpose of symbol synchronization is to detect the best decision instant in which the signal has the maximum signal-to-noise ratio (SNR) and energy, performing it with the best achievable precision.

A number of solutions available in the research literature today are based in data-aided (DA) algorithms [114], involving the transmission of a pre-determined training sequence or a reference clock signal together with the transmitted signal. Nonetheless, solutions of this kind demand either a higher power signal, in order to transmit an off-band pilot clock tone, or transmission at higher symbol rates, in order to include a symbol overhead, hence penalizing the overall system performance. Differently, non-data-aided (NDA) solutions extract the inherent timing information directly from the received signal. For that reason, the performance of this kind of solutions depends strongly on the quality of the clock information that is derivable from the received signal. High noise, inter-symbol interference (ISI), crosstalk and bandwidth constraints are the main impairments for these methods.

3.1.1 Feedback timing synchronization methods

One of the most well-known time synchronization algorithms was proposed 3 decades ago by Gardner [75]. His proposed algorithm is based in a feedback loop and needs only two samples per symbol, which implies low computational complexity. Figure 3.1 presents a common implementation of the method.

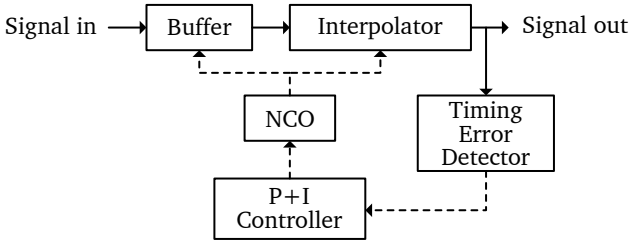

Figure 3.1.: Timing recovery implementation. NCO: numerically controlled oscillator; P+I: proportional-plus-integral. Continuous lines: data signal flow; Dashed lines: control signal flow.

The assumption behind Gardner's method is that as m-ary phase-shift keying (PSK) signals ideally have constant modulus, then, the signal energy is at its maximum when the signal is decided at the optimum instant, while it is attenuated during the transitions between its constellation points. In order to depict it, Figure 3.2(a) illustrates a quadrature phase shift keying (QPSK) constellation in red circles and the non-return-to-zero (NRZ) transitions between the symbols in blue lines, Figure 3.2(b) shows the eye diagram of the in-phase component of the QPSK signal and Figure 3.2(c) shows the eye diagram of the QPSK signal energy. It is possible to observe that the average signal energy is lower at the transitions than at the optimum decision instant.

By regulating the sampling instant of the signal, the algorithm maximizes the energy of the received signal. Therefore, the timing error detector block shown in Figure 3.1 computes, for every symbol, the Gardner timing error

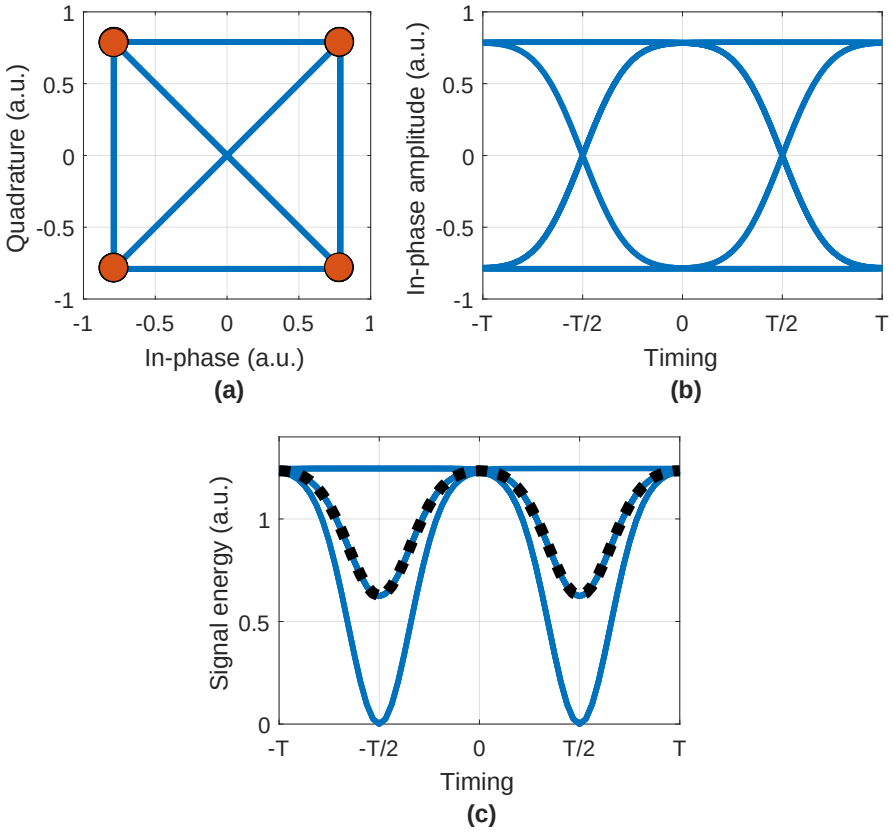

Figure 3.2.: Ideal QPSK signal with NRZ pulse shape. (a) Constellation (red circles) with its transitions between symbols (blue lines); (b) Eye-diagram of the in-phase component; (c) Eye-diagram of the signal energy (blue lines) and average energy in function of time (black dashed line). T is the symbol period. Based on [J2].

function, $e_{\text{Gard}}[n]$, which is the derivative of the energy of the interpolator output, $b(nT)$. Then,

$$\begin{aligned} e_{\text{Gard}}[n] &= \left[\bar{b}(nT)b(nT)\right]' = \bar{b}'(nT)b(nT) + \bar{b}(nT)b'(nT) \\ &= \left[b_i'(nT) - jb_q'(nT)\right]\left[b_{\text{i}}(nT) + jb_{\text{q}}(nT)\right] \\ &\quad + \left[b_{\text{i}}(nT) - jb_{\text{q}}(nT)\right]\left[b_i'(nT) + jb_q'(nT)\right] \\ &= 2b_{\text{i}}(nT)b_i'(nT) + 2b_{\text{q}}(nT)b_q'(nT), \end{aligned} \tag{3.1}$$

where n is a discrete index indicating the n-th symbol, T is the symbol period, $P_b(nT)$ is the power of the interpolator output, $b_{\text{i}}(nT)$ and $b_{\text{q}}(nT)$ are, respectively, the real and imaginary parts of $b(nT)$. The prime symbol

$(')$ means the first derivative, and the overbar $(\bar{})$ express the complex conjugate. Approximating $b'(nT)$ numerically,

$$b'(nT) = \frac{1}{2}\left[b\left(nT + \frac{T}{2}\right) - b\left(nT - \frac{T}{2}\right)\right], \tag{3.2}$$

and making $b[n] = b(nT/2 + T/2)$, then

$$\begin{aligned} e_{\text{Gard}}[n] &= b_{\text{i}}[2n-1]\big(b_{\text{i}}[2n] - b_{\text{i}}[2n-2]\big) \\ &\quad + b_{\text{q}}[2n-1]\Big(b_{\text{q}}[2n] - b_{\text{q}}[2n-2]\Big) \\ &= \Re\left\{\bar{b}[2n-1]\left(b[2n] - b[2n-2]\right)\right\}, \end{aligned} \tag{3.3}$$

where $b[n]$ is the n-th sample of $b(t)$ with $2/T$ sampling rate, and $\Re\{\}$ indicates the real part. Averaging through N symbols, the timing error function,

$$\tilde{e}_{\text{Gard}}[n] = \frac{1}{N}\sum_{i=0}^{N-1} e_{\text{Gard}}[n-i]. \tag{3.4}$$

is known in the literature as the S-curve [75, 115] and it is used to control if the signal is early or late in comparison to the optimum decision instant. Figure 3.3 displays, for an ideal NRZ-QPSK signal, all the possible $e_{\text{Gard}}[n]$ outputs in blue, and the S-curve, $\hat{e}_{\text{Gard}}[n]$, in black.

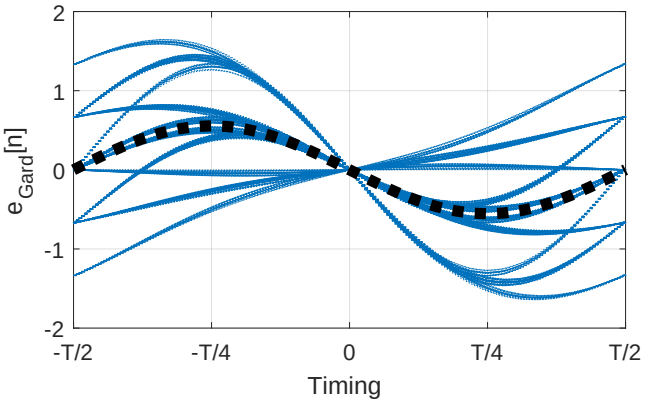

Figure 3.3.: Gardner's S-curve, in black, and the values that $e_{\text{Gard}}[n]$ assume for an ideal QPSK signal, in blue. Based on [J2].

A proportional-plus-integral (P+I) controller is responsible to filter the timing error output, controlling the frequency and phase of a numeric controlled

oscillator (NCO) that drives a timing signal to the interpolators, locking when the timing error function reaches zero (in the middle of the S-curve).

The Gardner's timing error detector function can also be computed in frequency domain by [115]

$$\tilde{e}_{\text{Gard}}[n] = \frac{2}{K}\sum_{k=1}^{K/2}\left[\sin\left(\frac{2\pi k}{K}\right)\Im\left\{B[k,n]\bar{B}[k+K/2,n]\right\}\right], \tag{3.5}$$

where $\Im\{\}$ indicates the imaginary part, $B[k,n]$ is the discrete Fourier transform (DFT) of a block of samples that starts in $b[2n]$ and ends in $b[2n+K]$, and K is the DFT block length. The equation above is analogous to the one presented by Godard [76],

$$\tilde{e}_{\text{God}}[n] = \frac{2}{K}\sum_{k=1}^{K/2}\left[\Im\left\{B[k,n]\bar{B}[k+K/2,n]\right\}\right], \tag{3.6}$$

and has been proven to produce comparable results in relation to the Gardner clock recovery method [115].

The performance of the timing synchronization process is strongly related to the amplitude of the S-curve, with higher values yielding better performance and values close to zero indicating loss of synchronization. The S-curve amplitude is known as clock tone amplitude (CTA) and it is possible to compute it by the absolute value instead of the imaginary part either in Eq. 3.5 or Eq. 3.6 [108]. For simplicity, in this Chapter, the CTA is computed as

$$\text{CTA} = \left|\sum_{k=1}^{K/2} B[k]\bar{B}\left[k+K/2\right]\right|. \tag{3.7}$$

As the CTA value computed from Equation 3.7 depends on K and the signal power, the normalized CTA, given by the CTA value divided by the maximum CTA value possible, will be used as performance metric.

3.1.2 Modified Gardner phase detector

The method proposed by Gardner and explained in the previous section relies that the signal has an NRZ or return-to-zero (RZ) pulse shape. The CTA is directly proportional to the roll-off of signals with raised cosine (RC) or root raised cosine (RRC) pulse shape, as for these pulse shapes, the signal

is filtered more at the frequency where the clock tone is, therefore, the Gardner method is not suitable for signals with spectrally-efficient schemes that rely on *quasi*-Nyquist pulse shape. A common modification on the phase detector used to circumvent this limitation is to consider its inputs as the energy of the input signal instead of considering its inputs as the raw signal [78, 79]. The modified phase detector is given by

$$\begin{aligned} e_{mod}[n] &= |b[2n-1]|^2 \left(|b[2n]|^2 - |b[2n-2]|^2\right), \\ &= b_{\mathrm{i}}^2[2n-1]\left(b_{\mathrm{i}}^2[2n] - b_{\mathrm{i}}^2[2n-2]\right) + \\ &\qquad b_{\mathrm{q}}^2[2n-1]\left(b_{\mathrm{q}}^2[2n] - b_{\mathrm{q}}^2[2n-2]\right). \end{aligned} \tag{3.8}$$

The assumption behind this modified method is similar to Gardner's method. For a low roll-off RC or RRC signal, the average signal energy tends to be constant, therefore, the energy maximization is unreliable in this case. If the input signal is squared, the optimum decision instant is the point of minimum energy. Figure 3.4 and Figure 3.5 depicts curves for a QPSK signal using a RC pulse shape with 0.1 roll-off. Figure 3.4(a) illustrates the ideal constellation in red circles and, in black lines, the transitions between the symbols. Figure 3.4(b) shows the eye diagram of the in-phase component. Figure 3.4(c) shows the eye diagram of the signal energy. Figure 3.4(d) shows the eye diagram of the absolute squared signal energy. Figure 3.5(a) and 3.5(b) show, respectively, the S-curves for the Gardner's method and the modified method. It is possible to observe that the average of the absolute squared signal energy (Figure 3.4(d)) has its lower point at the optimum decision instant, hence the inverted S-curve in 3.5(b) when in comparison to the one obtained in Figure 3.3.

3.2 Numerical fiber propagation model

In this Section, the numerical fiber propagation model used in the analyses is presented. This model considers the random couplings between polarizations and modes during fiber transmission.

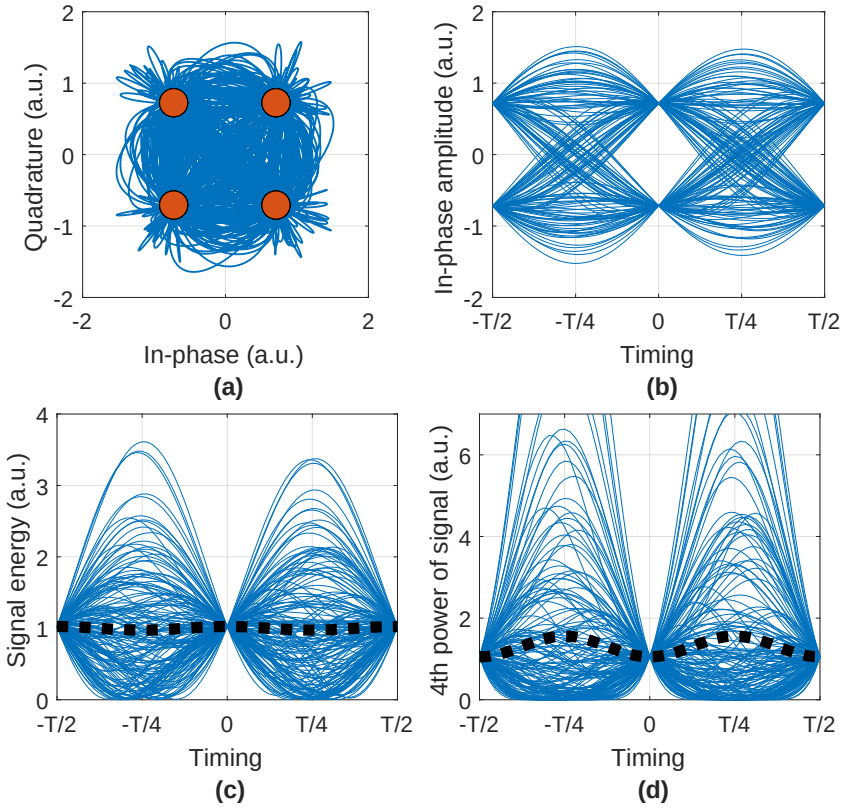

Figure 3.4.: Ideal QPSK signal with RC pulse shape and roll-off of 0.1. (a) Constellation (red circles) with some transitions between symbols (blue lines); (b) Eye-diagram of the in-phase component; (c) Eye-diagram of the signal energy (blue lines) and average energy in function of time (black dashed line). (d) Eye-diagram of the squared signal energy (blue lines) and average squared energy in function of time (black dashed line). T is the symbol period.

3.2.1 Single-mode fibers without coupling between polarizations

If nonlinearities and noise are ignored, the propagation of an optical signal through a SMF can be described, in discrete frequency domain, by

$$\vec{S}_{\text{out}}[\Omega] = \boldsymbol{M}[\Omega]\vec{S}_{\text{in}}[\Omega], \tag{3.9}$$

where $\boldsymbol{M}[\Omega]$ is a discrete frequency domain 2×2 matrix that expresses the transfer characteristics of the optical fiber supporting 2 polarizations, and

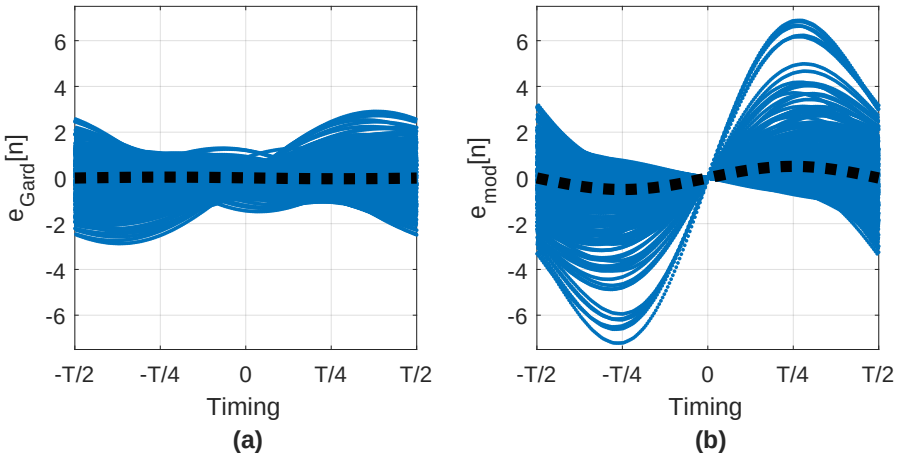

Figure 3.5.: S-curve, in black, and the values that it could assume, in blue; QPSK signal, RC pulse shape, roll-off of 0.1; (a) Gardner's method; (b) Modified method.

$\vec{S}_{\text{in}}[\Omega]$ and $\vec{S}_{\text{out}}[\Omega]$ are, respectively, the input and output electrical field vectors in the discrete frequency domain.

Since the fabrication process of SMFs are not perfect, orthogonal polarization states propagate through the optical fiber at different speeds, a phenomenon known as birefringence. Without coupling between polarization modes, the fiber transfer characteristic can be described by [116, 117]

$$\boldsymbol{M}[\Omega] = \boldsymbol{V}\boldsymbol{\Lambda}[\Omega]\boldsymbol{U}^{\mathsf{H}}, \tag{3.10}$$

where $\boldsymbol{V}$ and $\boldsymbol{U}$ are random unitary matrices that express coordinate base transformations that are necessary due to the unknown spatial direction of the birefringence. The H symbol indicates the matrix conjugate transpose. $\boldsymbol{V}$ and $\boldsymbol{U}$ are defined by

$$\boldsymbol{V}, \boldsymbol{U} \stackrel{\text{def}}{=} \begin{bmatrix} \cos(\theta) & \sin(\theta)e^{j\phi} \\ -\sin(\theta) & \cos(\theta)e^{j\phi} \end{bmatrix}, \tag{3.11}$$

where θ and ϕ are, respectively, random azimuth and ellipticity angles of the unitary rotation matrices. $\boldsymbol{\Lambda}[\Omega]$ is a diagonal matrix that describes the linear propagation in each of the polarization directions, and includes

polarization-dependent loss (PDL), differential group delay (DGD) and CD. $\boldsymbol{\Lambda}[\Omega]$ is expressed by

$$\boldsymbol{\Lambda}[\Omega] = \mathrm{diag}\left[\Lambda_1[\Omega], \Lambda_2[\Omega]\right], \tag{3.12}$$

where

$$\Lambda_1[\Omega] = \exp\left\{\frac{g_1}{2} - j\frac{\tau L}{2}\Omega - jD_C\frac{\lambda^2 L}{4\pi c}\Omega^2 + jD_S\frac{\lambda^4 L}{24\pi^2 c^2}\Omega^3\right\}, \tag{3.13}$$

$$\Lambda_2[\Omega] = \exp\left\{\frac{g_2}{2} + j\frac{\tau L}{2}\Omega - jD_C\frac{\lambda^2 L}{4\pi c}\Omega^2 + jD_S\frac{\lambda^4 L}{24\pi^2 c^2}\Omega^3\right\}, \tag{3.14}$$

where $\vec{g} = \left[g_1, g_2\right]^\top$ is the PDL vector, τ is the DGD, related to PMD, D_C is the CD parameter, D_S is the fiber dispersion slope, related to third-order CD, and L is the total optical fiber length.

3.2.2 Single-mode fibers with strong coupling between polarizations

Birefringence vary randomly throughout the length of an optical fiber because of bends and stretches. This makes the design of long optical fibers without coupling between polarizations unattainable. Therefore, DGD can be a limiting factor for optical communications. To overcome this issue, SMFs are fabricated with intentional birefringence in order to average out the GD throughout the signal propagation [51]. A common numerical approximation for the fiber propagation model is to express the optical fiber as the concatenation of shorter fiber segments (Figure 3.6), with slightly longer length than the length over which the complex polarization fields would remain correlated [117, 118]. Each of these segments have no coupling between polarizations, but both random rotations at the input and at the output occur and the DGD for each fiber segment is kept constant and not null.

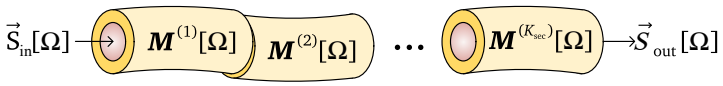

Figure 3.6.: Depiction of the fiber model used throughout the Chapter. The fiber is considered as a concatenation of shorter fiber segments. Based on [J2].

The signal at w-th fiber segment output, $\vec{S}_{\text{out}}^{(w)}(\Omega)$, is related to its input signal, $\vec{S}_{\text{in}}^{(w)}(\Omega)$, by

$$\vec{S}_{\text{out}}^{(w)}[\Omega] = \boldsymbol{M}^{(w)}[\Omega]\vec{S}_{\text{in}}^{(w)}[\Omega], \tag{3.15}$$

where w is the fiber segment index and $\boldsymbol{M}^{(w)}(\Omega)$ is the transfer characteristics matrix of the w-th segment. The full fiber transfer characteristic is, then, given by

$$\boldsymbol{M}[\Omega] = \prod_{k=1}^{K_{\text{sec}}} \boldsymbol{V}^{(w)}\boldsymbol{\Lambda}^{(w)}[\Omega]\boldsymbol{U}^{(k)\mathsf{H}}. \tag{3.16}$$

3.2.3 Multi-mode and multi-core fibers

In order to expand the model to include multi-mode fibers (MMFs) and MCFs, the transfer function characteristic, $\boldsymbol{M}[\Omega]$, can be extended, taking the form of a $Z \times Z$ matrix, where Z is the number of all possible spatial degrees of freedom. In this case, degrees of freedom can include both the two polarizations in a SMF, in which $Z = 2$, as well as the spatial degrees of freedom in either a MMF or a MCF, in which $Z = 2P > 2$, where P is the number of fiber modes or fiber cores. $\vec{S}_{\text{in}}[\Omega]$ and $\vec{S}_{\text{out}}[\Omega]$ become Z-dimensional input and output electrical field vectors in frequency domain. The matrix $\boldsymbol{\Lambda}^{(w)}[\Omega]$ can be expressed as

$$\boldsymbol{\Lambda}^{(w)}[\Omega] = \text{diag}\left[\Lambda_1[\Omega], \Lambda_2[\Omega], \cdots, \Lambda_Z[\Omega]\right], \tag{3.17}$$

$$\Lambda_z[\Omega] = \exp\left\{\frac{g_z^{(w)}}{2} - j\tau_z^{(w)}\Omega - j\frac{D_{C,z}\lambda^2 L_{\text{sec}}}{4\pi c}\Omega^2 + j\frac{D_{S,z}\lambda^4 L_{\text{sec}}}{24\pi^2 c^2}\Omega^3\right\}, \quad z \in [1, 2, \cdots, Z], \tag{3.18}$$

where L_{sec} is the length of each fiber segment, $\vec{g}^{(w)} = \left[g_1^{(w)}, g_2^{(w)}, \ldots, g_Z^{(w)}\right]^\top$ is the uncoupled mode-dependent loss (MDL) vector, $\vec{\tau}^{(w)} = \left[\tau_1{}^{(w)}, \tau_2{}^{(w)}, \ldots, \tau_Z{}^{(w)}\right]^\top$ is the GD vector related to the MD, $\vec{D}_C = \left[D_{C,1}, D_{C,2}, \ldots, D_{C,Z}\right]^\top$ is the CD, and $\vec{D}_S = \left[D_{S,1}, D_{S,2}, \ldots, D_{S,Z}\right]^\top$ is the fiber CD slope vector. The aforementioned vectors represent the respective impairment effect at each degree of freedom during the signal propagation through the w-th fiber segment. $\boldsymbol{V}^{(w)}$ and $\boldsymbol{U}^{(w)}$ are random $Z \times Z$ unitary matrices.

In order to produce $Z \times Z$ random unitary matrices, Givens' rotation matrices [119], $\boldsymbol{G}(i,k,\theta_{i,k},\phi_{i,k})$, are used. These matrices basically describe rotations relative to an axis, consequently, random rotations relative to all possible axes are able to generate random unitary matrices. Therefore, $\boldsymbol{U}^{(w)}$ and $\boldsymbol{V}^{(w)}$ are defined by

$$\boldsymbol{V}^{(w)},\boldsymbol{U}^{(w)} \stackrel{\text{def}}{=} \prod_{i=1}^{Z-1}\prod_{l=i+1}^{Z} \boldsymbol{G}\left(i,l,\theta_{i,l},\phi_{i,l}\right), \tag{3.19}$$

where $\theta_{i,l}$ and $\phi_{i,l}$ are, respectively, random azimuth and ellipticity rotation angles related to the rotation along the axis created by the i-th and the l-th degree of freedom, and each element $g_{m,n}$ of $\boldsymbol{G}(i,l,\theta_{i,l},\phi_{i,l})$ is given by

$$g_{m,n} = \begin{cases} \cos\left(\theta_{i,l}\right), & \text{if } m=n=i \\ \cos\left(\theta_{i,l}\right)e^{j\phi_{i,l}}, & \text{if } m=n=l \\ \sin\left(\theta_{i,l}\right)e^{j\phi_{i,l}}, & \text{if } m=i \text{ and } n=l \\ -\sin\left(\theta_{i,l}\right), & \text{if } m=k \text{ and } n=i \\ 1, & \text{if } m=n\neq i \text{ or } m=n\neq l \\ 0, & \text{otherwise.} \end{cases} \tag{3.20}$$

Observe that for the single-mode case, $Z=2$, the rotation matrix from Eq. (3.19) assumes the same form of Eq. (3.11).

3.3 Time skew model

Another important parameter that may impair the clock recovery performance is the time skew between signal components [J1]. So, apart from the DGD, the signal components can be time mismatched due to propagation differences in the electrical cables or other sources both at the transmitter and the receiver. Time skews can be of two types, IQ time skews and XY, or inter-modal, time skews.

3.3.1 Inter-polarization or inter-modal time skew

It is possible to model the XY or inter-modal time skews, $\boldsymbol{W}[\Omega]$, as

$$\boldsymbol{W}[\Omega] = \text{diag}\left[\exp\{-j\Omega\tau_1\},\exp\{-j\Omega\tau_2\},\cdots,\exp\{-j\Omega\tau_Z\}\right], \tag{3.21}$$

where $\vec{\tau}_{XY} = [\tau_1, \tau_2, \cdots, \tau_Z]^\top$ is the time skew vector representing the time skew in each degree of freedom. The received signal could be related to the input signal by,

$$\vec{S}_{\text{out}}[\Omega] = \boldsymbol{W}_{\text{Rx}}[\Omega]\boldsymbol{M}[\Omega]\boldsymbol{W}_{\text{Tx}}[\Omega]\vec{S}_{\text{in}}[\Omega]. \tag{3.22}$$

3.3.2 In-phase/quadrature time skew

Contrastingly, the IQ time skews are not possible to represent by matrix multiplications, and could be considered as a nonlinear impairment in the complex domain. If

$$\vec{s}(t) = \begin{bmatrix} s_{i1}(t) + js_{q1}(t) \\ s_{i2}(t) + js_{q2}(t) \\ \vdots \\ s_{iZ}(t) + js_{qZ}(t) \end{bmatrix} \tag{3.23}$$

is the time domain vector representing the signal, then, the IQ time-skewed signal, $\vec{\mathfrak{s}}(t, \vec{\tau}_{\text{IQ}})$, can be described by

$$\vec{\mathfrak{s}}(t, \vec{\tau}_{\text{IQ}}) = \begin{bmatrix} s_{i1}(t + \tau_1) + js_{q1}(t - \tau_1) \\ s_{i2}(t + \tau_2) + js_{q2}(t - \tau_2) \\ \vdots \\ s_{iZ}(t + \tau_Z) + js_{qZ}(t - \tau_Z) \end{bmatrix}. \tag{3.24}$$

where $\vec{\tau}_{\text{IQ}} = [\tau_1, \tau_2, \cdots, \tau_Z]^\top$ represents the time skew between in-phase and quadrature components in each of the degrees of freedom.

3.4 Clock recovery performance analysis

Throughout this section, the clock recovery performance is evaluated by the CTA for the QPSK modulation format. Similar performance is expected from higher-order quadrature amplitude modulation (QAM) modulation formats as the spectral shape of these formats are invariant to the modulation order [120].

3.4.1 Impact of pulse shaping

A back-to-back (B2B) system with an ideal QPSK modulation format and either RC or RRC pulse shapes are considered. The roll-off is swept from 0 to 1 and the CTA values using the gardner method and the modified method are plotted in Figure 3.7. For this simulation, no other linear or nonlinear impairment was considered.

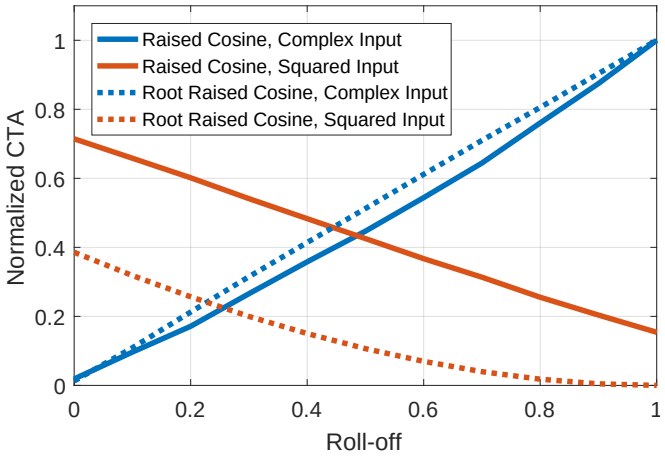

Figure 3.7.: Normalized clock tone amplitude for a QPSK signal employing either raised cosine (continuous lines) or root raised cosine (dashed lines) pulse shapes. CTA computed by the Gardner method (blue curves) and the modified method (red curves).

As expected, for very low roll-off factors, the CTA of the Gardner's method tends to zero. The modified algorithm shows much higher CTA for RC pulse shape in comparison with the RRC pulse shape for all roll-off factor values. It is possible to observe that the modified method still has non-negligible CTA for roll-off factor of 1. This result will be important on the development of the transmitter optimization method presented in Chapter 4.

3.4.2 Tolerance to chromatic dispersion

A B2B system with an ideal QPSK modulation format, NRZ and RC with 0.1 roll-off factor, and 16, 32, 48, and 64 GBd symbol rates were considered. A residual CD was swept and the normalized CTA are plotted in Figure 3.8.

The tolerance to CD is proportional to the square of the symbol period, and it is higher for lower roll-off signals due to the increased spectral efficiency.

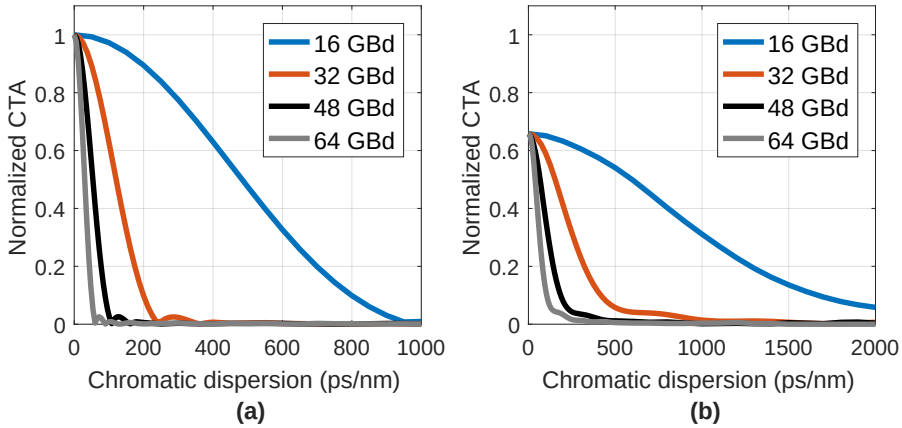

Figure 3.8.: Normalized clock tone amplitude for a QPSK signal with different symbol rates, employing: (a) NRZ pulse shape, computed by the Gardner's method; (b) RC with 0.1 roll-off factor, computed by the modified method.

3.4.3 Time skew between polarizations in single-mode fibers

A B2B system with time skews at both the transmitter and receiver was considered for the purpose of evaluation of the impact of time skews between polarizations. Moreover, and for simplicity, only an azimuth rotation was considered between the transmitter and the receiver. The transfer function for this simulation is given by

$$\frac{\vec{S}_{\text{out}}[\Omega]}{\vec{S}_{\text{in}}[\Omega]} = \begin{bmatrix} e^{-j\Omega\frac{\tau_{\text{Rx}}}{2}} & 0 \\ 0 & e^{j\Omega\frac{\tau_{\text{Rx}}}{2}} \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} e^{-j\Omega\frac{\tau_{\text{Tx}}}{2}} & 0 \\ 0 & e^{j\Omega\frac{\tau_{\text{Tx}}}{2}} \end{bmatrix}, \tag{3.25}$$

where τ_{Tx} is the time skew between polarizations at the transmitter, τ_{Rx} is the time skew between polarizations at the receiver, and θ the rotation angle considered. The received signal in both polarizations are

$$\begin{aligned} S_{\text{out,X}}[\Omega] &= e^{-j\Omega\frac{\tau_{\text{Rx}}}{2}} \left[e^{-j\Omega\frac{\tau_{\text{Tx}}}{2}} \cos(\theta) S_{\text{in,X}}[\Omega] + e^{j\Omega\frac{\tau_{\text{Tx}}}{2}} \sin(\theta) S_{\text{in,Y}}[\Omega] \right] \\ S_{\text{out,Y}}[\Omega] &= e^{j\Omega\frac{\tau_{\text{Rx}}}{2}} \left[-e^{-j\Omega\frac{\tau_{\text{Tx}}}{2}} \sin(\theta) S_{\text{in,X}}[\Omega] + e^{j\Omega\frac{\tau_{\text{Tx}}}{2}} \cos(\theta) S_{\text{in,Y}}[\Omega] \right]. \end{aligned} \tag{3.26}$$

Observe that in each polarization the received signal is a combination of the transmitted signal in both polarizations with different timings due to the

transmitter-side XY time skew. On the other hand, the receiver-side time skew delays the transmitted signal components equally. Thus, it is expected that the clock recovery performance should be invariant to the receiver time skew but should be impaired by the time skew at the transmitter depending on the polarization rotation during transmission. To show this behavior, an NRZ-QPSK with symbol rate of 32 GBd was considered as the signal input. The rotation angle was swept from 0 to 90 degrees, the transmitter time skew was swept from 0 to 31.125 ps (equivalent to 1 symbol period) and the receiver-side time skew was set to zero. The normalized CTA was extracted from the signal from polarization X and the results are shown in Figure 3.9.

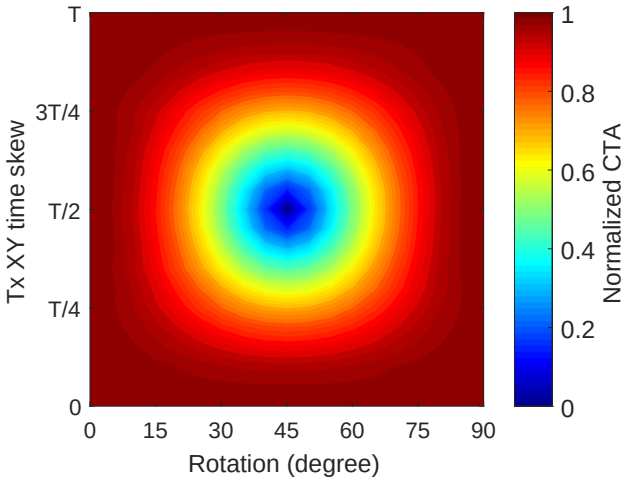

Figure 3.9.: Normalized clock tone amplitude as a function of rotation angle and transmitter-side inter-polarization time skew.

The CTA value fades completely when the rotation angle is at 45 degrees and the transmitter-side inter-polarization time skew is half of the symbol period. This behavior is similar to the impact from first-order DGD previously shown by Zibar *et al.* [104, 106]. It is possible to conclude then that the mitigation of XY time skews at the transmitter side is essential.

3.4.4 In-phase/quadrature time skew

A single polarization NRZ-QPSK signal in B2B with time skew between the in-phase (I) and quadrature (Q) components was considered. The normalized CTA is plotted in Figure 3.10.

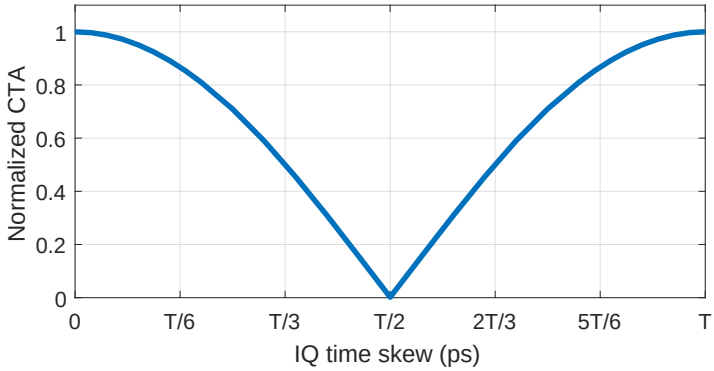

Figure 3.10.: Normalized clock tone amplitude as a function of the in-phase/quadrature time skew for a single polarization NRZ-QPSK signal.

The clock tone decreases to zero when the IQ time skew is equal to half of the symbol period.

3.4.5 Polarization mode dispersion

In order to analyze the impact in clock recovery performance due to PMD in SMFs, a simulation of a transmission link comprising a transmitter, a SMF and a coherent receiver was considered. Signals at 32 GBd and NRZ-QPSK modulation format, multiplexed in two orthogonal polarizations were generated at the transmitter. The transmission link was numerically simulated and consisted of a 100-km SMF with strong coupling between polarizations which was divided in 10000 sections of 10 m each [51]. No noise or linearities were considered. The simulation setup is shown in Figure 3.11

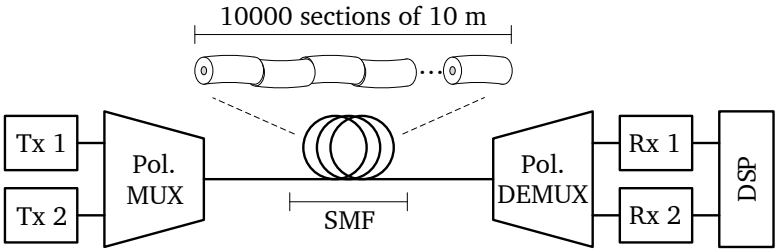

Figure 3.11.: General set-up for numerical simulation model of polarization multiplexed systems. Based on [J2].

The combination of different random Givens' rotation matrices and GD values for each fiber section was considered in order to evaluate the behavior of timing recovery in relation to DGD. Each segment had a random rotation with an angle uniformly distributed in the interval $\theta \sim [0, 2\pi)$. The uncoupled DGD was swept between 0.005 and 500 ps/km. The coupled DGD, *i.e.*, the resulting DGD after the signal propagation through the whole fiber, was measured by the maximum difference of the two eigenvalues of the matrix $\boldsymbol{H}[\Omega]$, defined by [63]

$$\boldsymbol{H}[\Omega] = j\frac{\partial \boldsymbol{M}[\Omega]}{\partial \Omega}\boldsymbol{M}^{\mathrm{H}}[\Omega], \tag{3.27}$$

where the differentiation was computed numerically.

Similarly to the case with transmitter time skew, it is expected that the CTA would vanish depending on the DGD value and the rotation angle. A possible solution to maximize the CTA value is to execute a polarization rotation in the digital domain prior to the clock recovery process [104, 108]. Therefore, it was considered both the CTA computed from a signal received at the horizontal polarization, $S_{\mathrm{out,X}}[\Omega]$ and the CTA computed from a combination of the horizontal and vertical polarizations, $S_{\mathrm{out,XY}}[\Omega]$, given by

$$S_{\mathrm{out,XY}}[\Omega] = S_{\mathrm{out,X}}[\Omega]cos(\phi) + S_{\mathrm{out,Y}}[\Omega]\sin(\phi), \tag{3.28}$$

where ϕ is the angle that maximized the CTA, emulating the methods proposed in [104, 108]. The results are shown in Figure 3.12.

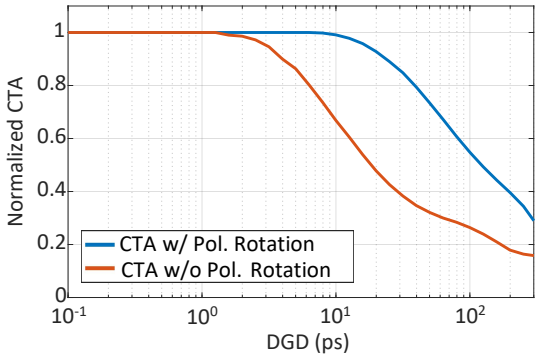

Figure 3.12.: Normalized clock tone amplitude of a 32 GBd NRZ-QPSK signal as function of the resulting (coupled) DGD for a 100-km single-mode fiber. Based on [J2].

While using the digital polarization rotation method, the CTA remain at the maximum level for coupled DGD values below 10 ps. Since a typical modern standard SMF has PMD parameter around 0.1 ps/$\sqrt{\text{km}}$, the CTA would only be degraded after 10000 km, requiring at least ten times this transmission distance in order to the CTA drop to half of its maximum value. Therefore, for practical applications of PDM systems in SMFs, the vanishing of clock tone should be a problem with a known solution. It is important to note that these distances were computed for a signal with 32 GBd, and the requirements would scale inversely with the symbol rate. For instance, if the symbol rate was 16 GBd, the distance needed for starting degrading the CTA would be 20000 km, while for a 32 GBd signal, this distance would be 5000 km.

3.4.6 Modal dispersion

In order to analyze the effects to the clock recovery performance due to MD in MMFs, a simulation of a transmission link comprising 6 transmitters, a mode multiplexer, a 3-mode FMF, a mode demultiplexer and 6 coherent receivers was considered. Signals at 32 GBd and NRZ-QPSK modulation format, multiplexed in 3 modes with 2 degrees of freedom per mode, were generated at the transmitter. The transmission link was divided in sections of 10 km each. At the receiver DSP, the clock extraction and resampling was considered before MIMO equalization and the clock extraction was performed per mode. The simulation setup is shown in Figure 3.13.

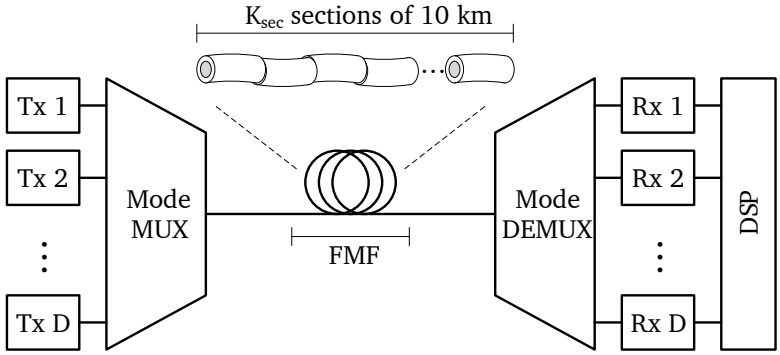

Figure 3.13.: General set-up for numerical simulation model of multi-mode multiplexed systems. Based on [C2, J2].

In order to evaluate the behavior of timing synchronization at different coupling regimes, we swept the rotation distribution of the unitary matrices $\boldsymbol{V}^{(w)}$ and $\boldsymbol{U}^{(w)}$ in Eq. (3.19), measuring the crosstalk, XT, into the fundamental mode group due to high order mode groups by

$$\mathrm{XT} = \frac{\sum_{i=3}^{Z} P_i}{P_1 + P_2}, \tag{3.29}$$

where $P_i[\Omega]$ is the power transferred from the first degree of freedom to the i-th degree of freedom, and is computed by

$$P_i = \sum_{k=1}^{N} \left\{ M_{1,i}[k]\bar{M}_{1,i}[k] + M_{2,i}[k]\bar{M}_{2,i}[k] \right\}, \tag{3.30}$$

where $M_{r,i}[k]$ is the element at the r-th column and i-th row of $\boldsymbol{M}[k]$, which is the transfer matrix in discrete frequency domain. For simplicity inter- and intra-mode group coupling were not distinguished, and no losses, non-linearities, time skew or chromatic dispersion were considered.

The results for a 1000-km transmission in a 3-mode FMF (100 sections of 10 km) is shown in Figure 3.14. This fiber has two mode groups, two degrees of freedom (one mode) on fundamental group and four degrees of freedom (two modes) in the second group. The strong coupling regime was considered as when the crosstalk was approximately 3 dB, the weak coupling regime when the crosstalk was lower than −18 dB and refer to intermediate coupling regime for the region in-between. The results shown are the average of 10 random simulations for each crosstalk value in a system with no losses or chromatic dispersion.

The curves were parameterized by the uncoupled GD, $\Delta\tau$, which is the time difference between the fastest and the slowest group velocity of the modes. As expected for zero GD, the clock tone remained unchanged at the maximum value irrespective of the crosstalk. Nevertheless, while increasing the uncoupled GD, a dramatic drop in the clock tone quality was observed. In the strong coupling case, an uncoupled GD of less than 0.03 ps/km was required in order to still have a detectable clock tone higher than 50% of the maximum value.

The behavior of CTA for strong coupling regime as a function of uncoupled GD at different transmission distances was also evaluated. The CTA drops

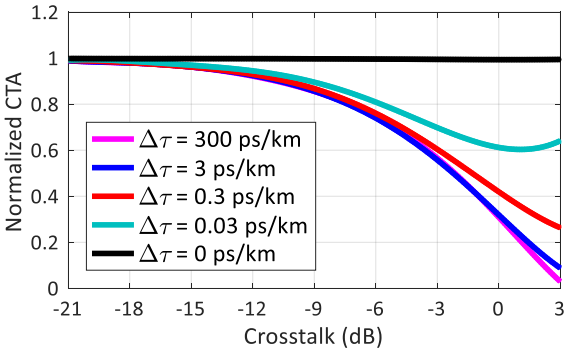

Figure 3.14.: Normalized clock tone amplitude as a function of crosstalk between mode groups in a 1000 km transmission of a 3-mode FMF for different values of uncoupled group delay. Based on [C2, J2].

for lower values of uncoupled GD as the transmission distance increases, therefore, the delay spread requirements are even more stringent for long-haul MDM transmission. Results are shown in Figure 3.15.

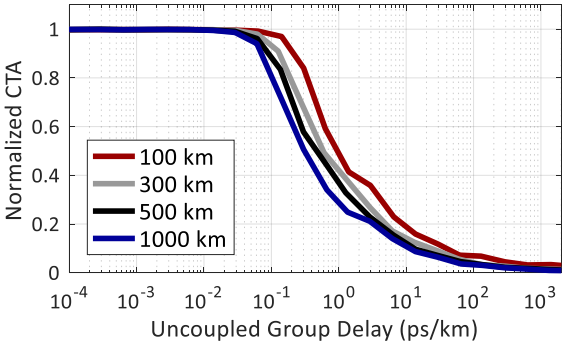

Figure 3.15.: Normalized clock tone amplitude as a function of uncoupled group delay between mode groups in a 3-mode FMF for distinct transmission distances. Based on [C2, J2].

3.5 Summary

This Chapter analyzed through simulations the performance and feasibility of NDA fully digital clock recovery algorithms for coherent optical communication systems. Tolerances to transceiver time skew, residual chromatic dispersion, different pulse shapes, polarization rotation, and PMD were demonstrated. In the next Chapter, an application of genetic algorithms

for optimization of transmitters for use in high-order modulation formats, which the cost functions are based on the CTA is presented.

4

Transmitter Optimization using Genetic Algorithm

THE GROWING DEMAND for higher bit rates and the necessity in reducing the cost per bit in optical communication systems have driven research in high-order modulation formats at high symbol rates combined with coherent detection. Furthermore, high capacity optical transmission has been widely investigated for long-haul links [3]. The continuous progress in electronics technology allowed state-of-the-art optical transceivers to evolve from legacy intensity-modulation and direct-detection based schemes to an association of multilevel modulation formats and coherent detection employing digital signal processing (DSP) [15, 22, 121].

The quality of coherent optical communication transceivers relies on the complexity involved in its manufacturing process. Front-end imperfections generally impair the signal quality both during generation at the transmitter and during reception at the receiver. Particularly, on the transmitter side, these imperfections may be caused by possible length and attenuation mismatches in the electrical components, a gain mismatch in transimpedance amplifiers (TIAs), and improper biasing of the optical modulator operation points, causing timing misalignments of several picoseconds in between the in-phase (I) and quadrature (Q) paths for both orthogonal polarizations, gain and phase imbalances between components, arbitrary DC levels, as well as nonlinearities caused by the sinusoidal response of the optical modulator. Furthermore, the penalty on coherent systems due to these imperfections scale up when using higher-order modulation formats and higher symbol rates [22, 122, 123] and can heavily impact multi-subcarrier transmission

systems [124, 125]. Similar problems could also arise on the receiver-side for analogous reasons.

It is accepted that receiver-side complex-valued 2×2 multiple-input multiple-output (MIMO) digital dynamic equalizers are able to compensate for inter-polarization (XY) time skew, but could not compensate for the in-phase/quadrature (IQ) imbalances either caused by malformed modulation formats or IQ time skews [126]. Recently, some solutions were presented based on adaptive MIMO equalizers. They are based in real-valued 4×4 [34] equalizers, complex-valued 4×2 [33, 126] equalizers, and receiver-side processing prior to MIMO equalization [35, 127]. However, these approaches would step up the receiver complexity that would lead to increased power consumption, being also impractical for already installed systems. None of these methods take into consideration the impacts into the timing synchronization process due to the IQ time skews and IQ imbalances existent at the received signal [128]. Furthermore, effects such as phase noise, chromatic dispersion (CD) and polarization mode dispersion (PMD) make it very difficult to distinguish between the impairments originated at the transmitter from the those caused by the receiver front-end, weakening the suitability of the use of such methods as just calibration processes.

Regarding time skews and IQ imbalances as static or very slow drifting impairments, adaptive equalization for the purpose of mitigation of such impairments is avoidable. In this manner, an alternative choice should be to perform calibration of the transmitter's front-end imbalances isolating it from the receiver rather than trying to mitigate its impairments into the receiver-side DSP. Then, rerunning the calibration process only periodically, in order to handle the aging of the equipment.

With this approach in mind, a method based on reconfigurable interference was proposed [129] for time skew estimation. This method requires a special apparatus based on photonic integrated circuits. A method based on the generation of arbitrary sine waves for self-calibration [130] and a method based on clock tone amplitude (CTA) extraction of a direct-detected signal, with the searching procedure performed by a simple genetic algorithm [C1] were also proposed. These methods may have worse performance if the transmitter bias voltages are not well set.

In this Chapter, a method to estimate and compensate for the front-end imperfections present in a polarization-diversity optical transmitter for high-order

quadrature amplitude modulation (QAM) modulation formats is proposed. This method is intended for joint calibration of transmitter IQ and XY time skew, amplitude mismatch, and bias voltages. Moreover, the presented method makes use of a cooperative coevolutionary genetic algorithm (CC-GA) which converges to the best solutions through fitness functions extracted from a directly detected reference signal while controlling the transmitter during the calibration process. The results and analyses, both at simulation and experimental levels that are presented here were based on the results in the author's original work in [C1, J1].

While this Chapter addresses the calibration of time skews and other impairments in a polarization-diversity transmitter, the next Chapter will address the time skew estimation and compensation on coherent receivers.

The structure of this Chapter is the following. Section 4.1 presents and reviews a theoretical basis for the impairments that usually affect a polarization diversity IQ transmitter and presents which information could be extracted by direct detection of the generated signal and then used for calibration. Section 4.2 introduces the proposed transmitter front-end calibration method based on CC-GA. Section 4.3 analyzes the method by means of simulations and demonstrates its estimating capabilities through experiments. Finally, Section 4.4 summarizes the Chapter.

4.1 Theoretical framework

This Section presents common impairments that may occur in the process of optical signal generation process and discusses how information on these impairments can be extracted by direct detection and subsequent compensation methods.

Section 4.1.1 provides a discussion of the effects of the transmitter-side time skew. Section 4.1.2 discusses the effects of incorrect bias voltages in the generation of QAM signals. Section 4.1.3 presents a definition of the amplitude mismatch and an estimation method.

4.1.1 Impact of time skew in QAM signals

The timing error between two components of a signal may be defined as a time skew. The time skew between the I and Q components of each polarization in high-order QAM transmitters is often called the IQ time skew,

whilst the time skew between orthogonal polarizations is generally referred to as the XY time skews [129].

A typical origin of time skews is a discrepancy in the length of the electrical lines at the transmitters. In Section 3.3, it was shown that both XY and IQ time skews can potentially degrade the transmission performance into a catastrophic level.

Considering that the I and Q components of a single-polarization optical signal are time-skewed, then

$$S_{\text{out}}(t) \propto e^{j\phi(t)}\Big[s_{\text{i}}(t) + js_{\text{q}}(t-\tau)\Big], \tag{4.1}$$

where $e_{\text{out}}(t)$ is a single polarization optical signal at the modulators output, $s_{\text{i}}(t)$ and $s_{\text{q}}(t)$ are the I and Q components of the signal, τ is the time skew between the components I and Q, $\phi(t)$ is the phase of the transmitter laser. For instance, Figure 4.1 displays power eye diagrams and constellation/transition diagrams in a single-polarization quadrature phase shift keying (QPSK) signal with non-return-to-zero (NRZ) pulse shaping for a few selected IQ time skew values. The power eye diagrams are eye diagrams made with the optical signal passed through a photodetector diode (PD), therefore showing its power.

In Figure 4.1, the transitions between consecutive symbols can be seen to cross farther away from the origin if IQ time skew values are increased. Since the clock tone used to synchronize the receiver indirectly depends on these transitions [128], the clock tone information can therefore be used as an effective error function, so that pre-compensation time skew values can be found that would maximize the clock tone.

The clock tone characteristics can be extracted in a directly-detected high-order QAM signal [79], as shown in section 3.1.2. Therefore, it would avoid interference from the laser phase noise if the time-skewed signal were received by a single PD. If the responsivity of the PD is ignored, the output current, $i(t)$, from a single-polarization signal can be written as

$$i(t) \propto \Big[|s_{\text{i}}(t)|^2 + |s_{\text{q}}(t-\tau)|^2\Big]. \tag{4.2}$$

The signal directly detected (Eq. 4.2) is proportional to the sum of the power of each optical signal component, *i.e.*, the I and Q components.

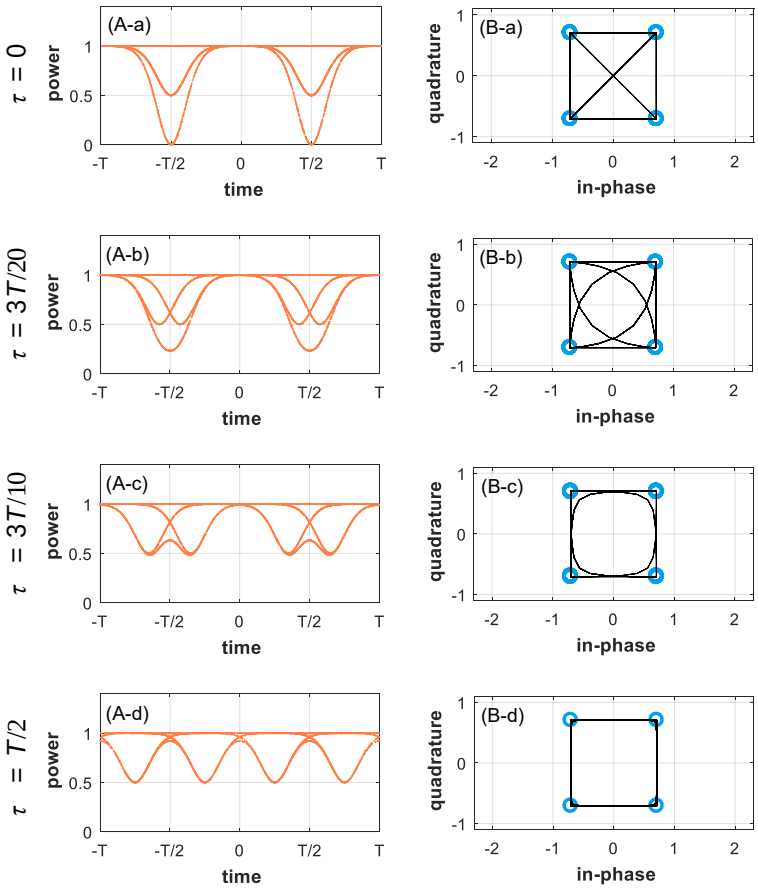

Figure 4.1.: IQ time skew effects in a single-polarization NRZ QPSK signal. (A) Power eye diagrams. (B) Constellation diagrams (blue circles) and symbol transition paths (black lines). (a) No time skew. (b) Time skew equals to 15% of symbol period. (c) Time skew equals to 30% of symbol period. (d) Time skew equals to half symbol period. Based on [J1].

The CTA is the maximum value of the timing error detector characteristics and it can be computed for the PD output current by [108]

$$\mathrm{CTA} = \left| \sum_{k=1}^{L/S} I[k]\bar{I}[k + L - L/S] \right|, \tag{4.3}$$

where $I(k)$ is the L-sized discrete Fourier transform of the received photocurrent, $i(t)$ from Eq. 4.2, and $S \geq 2$ is the number of samples per symbol. The CTA of the photodetected current is proportional to the sum of the clock tones from each of the components of the optical signal. Ideally, the relation

between CTA and the time skew, τ, normalized by the maximum possible CTA, is

$$\frac{\mathrm{CTA}(\tau)}{\max(\mathrm{CTA}(\tau))} = \left|\frac{1}{2} + \frac{1}{2}\exp\left(j\frac{\tau}{T}\right)\right|, \tag{4.4}$$

where T is the symbol period. This relation is illustrated in Figure 4.2. The CTA will have its maximum values for time skew $\tau = nT$, $\forall n \in \mathbb{Z}$, and its minimum values for time skew $\tau = (n + 1/2)T$, $\forall n \in \mathbb{Z}$. Consequently, a time skew estimator based on the CTA maximization have its estimation range limited by the interval $-T/2 < \tau < T/2$.

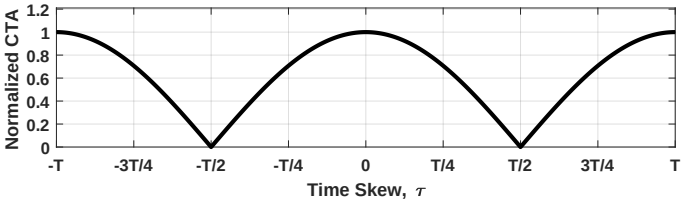

Figure 4.2.: Theoretical curve of CTA relative to transmitter time skew in a single-polarization signal. Based on [J1].

Equivalently, for a dual-polarization signal with time-skewed components, the field can be written as

$$\vec{s}_{\mathrm{out}}(t) \propto e^{j\phi(t)}\left\{\left[s_{ix}(t) + js_{\mathrm{qx}}\left(t-\tau_{\mathrm{x}}\right)\right]\vec{x} + \left[s_{iy}\left(t-\tau_{\mathrm{xy}}\right) + js_{\mathrm{qy}}\left(t-\tau_{\mathrm{xy}}-\tau_{\mathrm{y}}\right)\right]\vec{y}\right\}, \tag{4.5}$$

where $\vec{x} = \begin{bmatrix}1 & 0\end{bmatrix}^{\top}$ and $\vec{y} = \begin{bmatrix}0 & 1\end{bmatrix}^{\top}$ are, respectively, the direction vectors of the horizontal polarization (X) and vertical polarization (Y), τ_{x} and τ_{y} are, respectively, the IQ time skews between the components of the horizontal and vertical polarizations, τ_{xy} is the XY time skew, *i.e.*, the time skew between these two polarizations, and $s_{ix}(t)$, $s_{\mathrm{qx}}(t)$, $s_{iy}(t)$ and $s_{\mathrm{qy}}(t)$ are, respectively, the I and Q components of the horizontal polarization and the I and Q components of the vertical polarization.

Rewriting the time skews in relation to the I component of the horizontal polarization, $\tau_1 = \tau_{\mathrm{x}}$, $\tau_2 = \tau_{\mathrm{xy}}$, and $\tau_3 = \tau_{\mathrm{xy}} + \tau_{\mathrm{y}}$, then the photocurrent generated after passing the dual-polarization signal through a PD is

$$i(t) \propto |s_{ix}(t)|^2 + |s_{\mathrm{qx}}(t-\tau_1)|^2 + |s_{iy}(t-\tau_2)|^2 + |s_{\mathrm{qy}}(t-\tau_3)|^2. \tag{4.6}$$

For the dual-polarization signal, the CTA behavior is similar to the single-polarization case, as illustrated in Figure 4.3. The maximum CTA will only be achieved when all the values of the time skews are null, as shown in Figure 4.3(a).

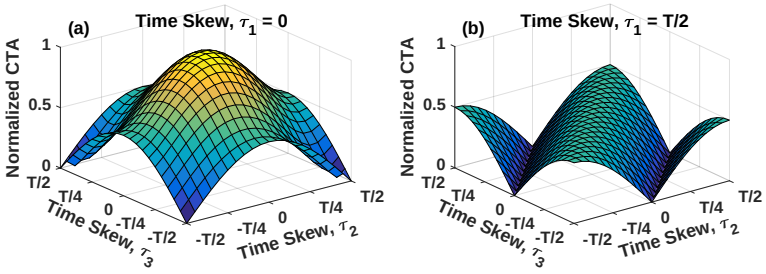

Figure 4.3.: Theoretical curves of CTA relative to transmitter time skew in a dual-polarization signal varying two values of time skews, while maintaining the remaining constant. (a) $\tau_1 = 0$; (b) $\tau_1 = T/2$. Based on [J1].

4.1.2 Impact of improper biasing on QAM signals

In this subsection, it is discussed how poorly-calibrated modulator bias voltages can impact the generated signal and what can be extracted from the signal for calibration purposes.

4.1.2.1 Operation points variation

The children bias voltages are responsible to move the center location of the signal constellation. They move it in a non-linear fashion due to the sinusoidal transfer function of the IQ modulator. For small peak-to-peak swing voltages, non-optimal children bias voltages change not only the constellation points but also the symbol transition paths, as can be seen in Figure 4.4(c). The transition paths cross farther from the zero and then, analogously to the time skew case, affect the CTA absolute value. This bias-dependent behavior of the CTA can be explored for the calibration of optimum values for the children bias voltages.

Moreover, variations in the children bias voltages change the power of the output optical signal, affecting the CTA in a different fashion. To avoid

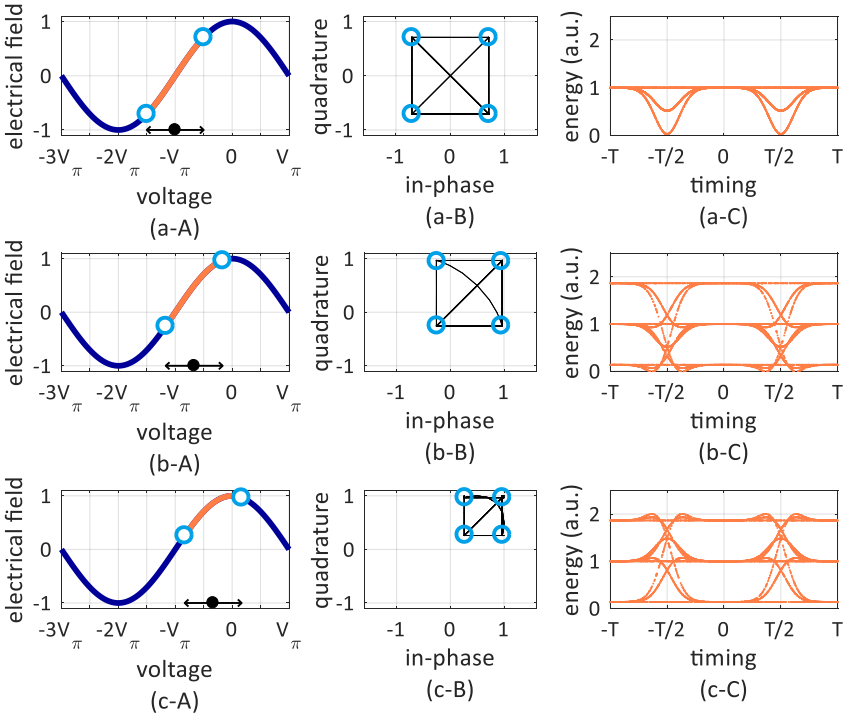

Figure 4.4.: Different children bias operation points in a single-polarization NRZ QPSK signal. (A) Electrical field transfer functions; (c) Constellation diagrams (blue circles) and symbol transition paths (black lines); (C) Eye diagrams of the signal energy; Children bias voltages: (a) $-V_\pi$; (b) $-2V_\pi/3$; (c) $-V_\pi/3$. Based on [J1].

interference from the signal power, a new metric is introduced, the modified clock tone amplitude (MCTA),

$$\mathrm{MCTA} = \frac{\left| \sum_{k=1}^{L/S} \left\{ I[k]\bar{I}[k+L-L/S] \right\} \right|}{\sum_{k=1}^{L} \left\{ I[k]\bar{I}[k] \right\}}. \tag{4.7}$$

The MCTA behavior for different values of the parent bias voltage, while varying the children bias voltages for a single polarization signal is demonstrated in Figure 4.5. The maximum MCTA value appears when the children bias voltages are optimum, and maintain a concave shape for different values of the parent bias voltages, being robust to its variation. For the polariza-

tion multiplexed (PM) case, the MCTA behaves similarly, having its global maximum values for the optimum values of all four children bias voltages.

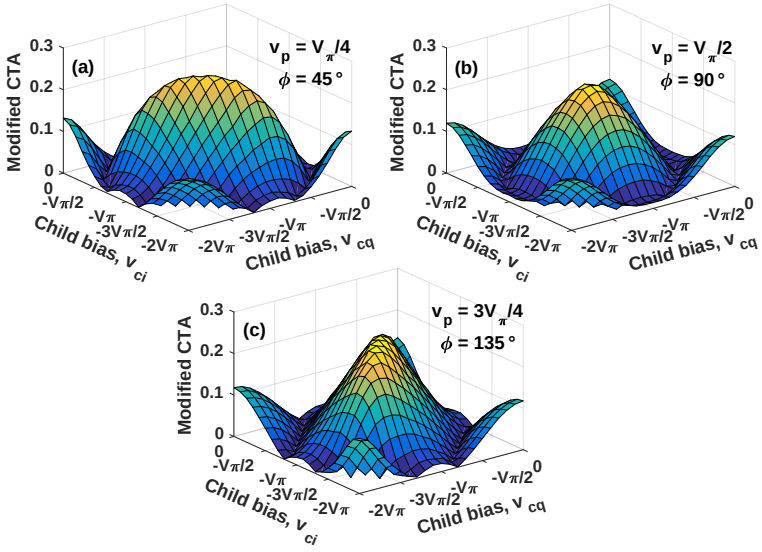

Figure 4.5.: Theoretical curves of MCTA relative to children bias voltages in a single-polarization NRZ QPSK signal with $4V_\pi/5$ peak-to-peak swing voltages. Parent bias voltage: (a) $V_\pi/4$. (b) $V_\pi/2$. (c) $3V_\pi/4$. Based on [J1].

4.1.2.2 Adjustment of IQ phase

The parent bias voltages adjusts the constellation phases in each polarization. It is possible to observe that the zero crossings are independent of the parent bias voltage (Figure 4.6). Therefore, in order to find the optimum values for these voltages, the CTA is not a suitable metric. Nonetheless, it is clear by Figure 4.6 that when v_p is correctly adjusted, *i.e.*, there is perfect orthogonality between I and Q components, the optical power of all symbols symmetric symbols are equal. Meanwhile, if the parent bias voltage diverge from its optimum value, the power corresponding to each of the symbols disperse, increasing the power variance. In such manner, it is possible to use the variance of the optical power signal, *i.e.*, the variance of the photodetected signal, as an indicator of the optimum value for the parent bias voltage. So, if the children bias are correctly set, the variance will be minimized when the parent bias is in its optimum value. Alternatively, the

inverse of the variance could be maximized in order to find the optimum parent bias values.

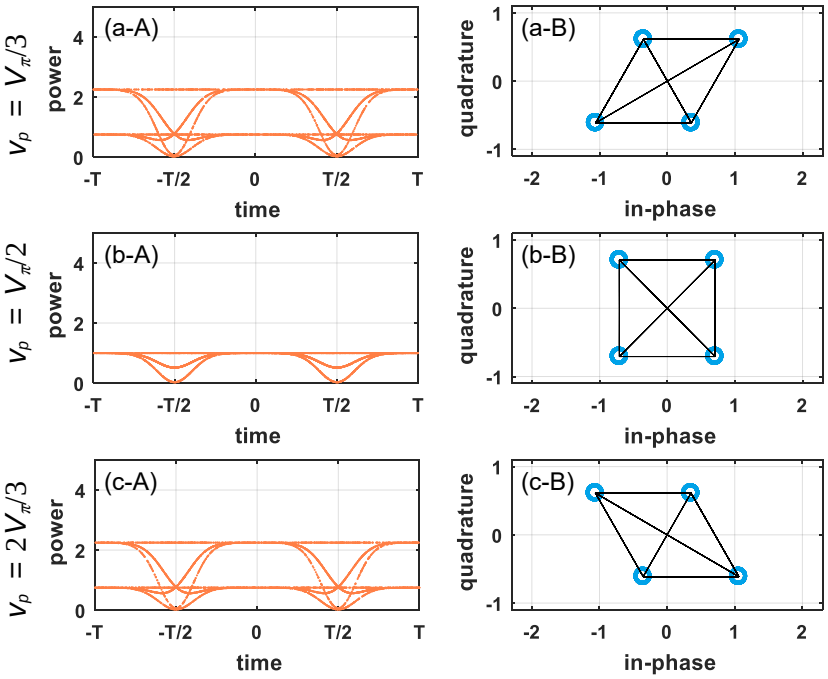

Figure 4.6.: Different parent bias in a single-polarization NRZ QPSK signal. (A) Power eye diagrams. (B) Constellation diagrams (blue circles) and symbol transition paths (black lines). (a) $V_\pi/3$ parent bias. (b) $V_\pi/2$ parent bias. (c) $2V_\pi/3$ parent bias. Based on [J1].

Figure 4.7(a) and Figure 4.7(c) show, respectively for the QPSK and 16QAM modulation formats, the variance versus parent bias voltage curve for a PM signal with optimum children bias voltages. Figure 4.7(b) and Figure 4.7(d) show the inverse of the variance for the same cases. Although not shown here, the variance curve have similar shape for QAM modulation formats of higher orders.

4.1.3 Amplitude mismatch in QAM signals

An amplitude mismatch between components of the optical signal may arise at the transmitter due to different attenuation in electrical paths and mismatched gains in electrical driver amplifiers. For a transmitter without time skew, operating at optimum bias voltages for QAM transmission and

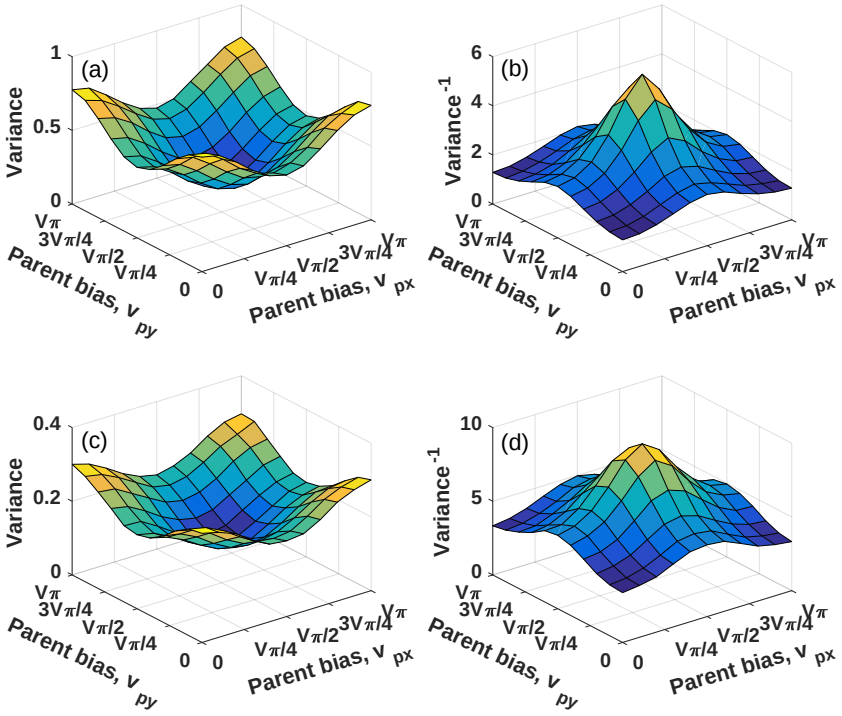

Figure 4.7.: Theoretical curve from the photodetected signal relative to parent bias voltages in a dual-polarization signal with $4V_{\pi}/5$ peak-to-peak swing voltages and optimum children bias voltages. (a) Variance for NRZ QPSK. (b) Inverse of variance for NRZ QPSK. (c) Variance for NRZ 16QAM. (d) Inverse of variance for NRZ 16QAM. Based on [J1].

swing voltages inside the linear region, the amplitude mismatched dual-polarization optical signal can be written as

$$\vec{s}_{\text{out}}(t) = e^{j\phi(t)} \left\{ \left[A_{\text{ix}} s_{\text{ix}}(t) + jA_{\text{qx}} s_{\text{qx}}(t) \right] \vec{x} + \left[A_{\text{iy}} s_{\text{iy}}(t) + jA_{\text{qy}} s_{\text{qy}}(t) \right] \vec{y} \right\}, \quad (4.8)$$

where A_{ix}, A_{qx}, A_{iy}, and A_{qy} are the amplitudes for each of the signal components, and $s_{\text{ix}}(t)$, $s_{\text{qx}}(t)$, $s_{\text{iy}}(t)$, and $s_{\text{qy}}(t)$ have unitary maximum amplitude.

An iterative method to estimate and compensate for these amplitude imbalances can be applied. It starts by defining amplitude imbalance correction factors, C_{qx}, C_{iy}, and C_{qy}, for each of the components, except the I component from the vertical polarization, that is used as a reference value. These

correction factors are multiplied by the signal components in the digital domain, before digital-to-analog conversion, such that the output signal is approximated by

$$\vec{s}_{\text{out}}(t) = e^{j\phi(t)} \left\{ \left[A_{\text{ix}} s_{\text{ix}}(t) + jC_{\text{qx}} A_{\text{qx}} s_{\text{qx}}(t) \right] \vec{x} + \left[C_{\text{iy}} A_{\text{iy}} s_{\text{iy}}(t) + jC_{\text{qy}} A_{\text{qy}} s_{\text{qy}}(t) \right] \vec{y} \right\}, \quad (4.9)$$

A signal only with the I component of one of the polarizations is then generated by "turning off" the other components, *i.e.*, reducing their swing voltages to zero. This turn-off procedure is possible only if the modulator is correctly biased for a QAM transmission, so a zero voltage on the input would result in minimum power on the output. Then, this single component signal is photodetected and sampled, and its power is computed by

$$P_{\text{ix}} = \sum_{k=1}^{N} i(k), \quad (4.10)$$

where $i(k)$ is the sampled photodetected current, and N is the number of samples acquired. Then, the power of each of the other components is computed. The correction factor for the amplitude imbalance is updated by

$$C_{h,\text{new}} = C_{h,\text{old}} \sqrt{\frac{P_{\text{ix}}}{P_h}}, \; h \in \{\text{qx, iy, qy}\}. \quad (4.11)$$

Due to the sinusoidal response of the modulators, these new correction factors are not immediately the best values, needing some iterations to converge to optimum values. After iterating the algorithm, the correction factors will converge to assure that $A_{\text{ix}} = C_{\text{qx}} A_{\text{qx}} = C_{\text{iy}} A_{\text{iy}} = C_{\text{qy}} A_{\text{qy}}$.

4.2 GA-based method for transmitter parameters calibration

To estimate and compensate for the PM transmitter front-end imperfections presented in the last section, it is possible to use the information extracted from the directly detected signal to optimize the aforementioned transmitter using a genetic algorithm and a turn-on/turn-off procedure. In this sec-

tion the genetic algorithm is discussed and then the proposed method for transmitter optimization is introduced.

4.2.1 Genetic algorithm for parameters optimization

Genetic algorithm (GA) is a particular class of evolutionary algorithms that has been successfully used to optimize a great variety of problems [131–134]. A typical genetic algorithm uses techniques inspired by evolutionary biology, as heredity, mutation, natural selection and crossover, being notably efficient to find good solutions in problems with many variables, and in the presence of noise. Thus, the GA is a suitable solution for transmitter parameters optimization, due to the quantity of variables to optimize and the noise present in CTA, MCTA and variance extraction processes. Additionally, it enables avoiding local extrema (minima or maxima), if the population used is big enough.

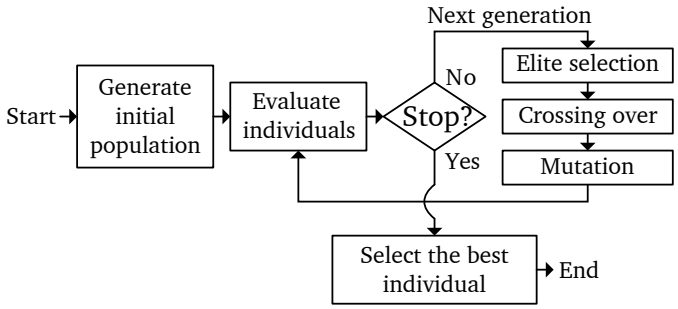

Figure 4.8.: Genetic algorithm block diagram. Based on [J1].

Figure 4.8 shows the basic implementation of a GA. The idea behind a GA is to optimize solutions, as living beings would evolve over generations in wild nature. It begins by creating randomly a set of initial solutions. These solutions are treated as individuals in a population and each variable of a solution is their chromosomes. Each of the individuals is assessed by a fitness function and genetic operations are then performed. The strongest individuals are labeled the elite group while the weakest ones die and are eliminated. A new offspring is then produced consisting of a pure copy of the elite group, crossover of pairs of elite chromosomes and elite group mutations. A new generation is then started and the process continues being repeated until a stopping criterion is met. The stopping criterion may be if the improvement between successive generations becomes insignificant or, more often, if a certain generation is achieved.

Tab. 4.1 shows a list of parameters of the PM optical modulator along with the information that can be extracted from the the directly detected signal and be used as fitness functions for optimization.

Table 4.1.: List of parameters along fitness functions

Parameter	Variables	Fitness function
Time skew	τ_1, τ_2, τ_3	CTA
Child bias voltage	$\dot{v}_{cix}$, $\dot{v}_{cqx}$, $\dot{v}_{ciy}$, $\dot{v}_{cqy}$	MCTA
Parent bias voltage	v_{px}, v_{py}	Variance
Amplitude mismatch	C_{qx}, C_{iy}, C_{qy}	Power per quadrature and pol. component

Then, as the transmitter impairment optimization is a problem with multiple fitness functions a multi-objective genetic algorithm (MO-GA) seems appropriate [134]. There could nevertheless be multiple solutions to a single problem leading to ambiguities and unsuitable solutions. On the other hand, a cooperative coevolutionary approach of genetic algorithms (CC-GA) can use partial fitness independence and provide faster and correct convergence to the best solution [135].

A larger population is divided by the CC-GA into subpopulations that are sequentially and iteratively resolved. Each subpopulation is referred to as species and an individual from a species is referred to as a specimen. A chromosome of a specimen consists of a subset of variables in the larger population. In the cooperative evaluation of each specimen, the only interaction between species is when the specimens evaluated are combined with randomly selected specimens from the elite group of other species. The combination of one specimen of each species forms a full chromosome with all parameters. The CC-GA thus simulates the cooperative evolution of different species in wildlife.

A larger problem can be reduced into three different species in the case of optimization of transmitter parameters: the time skews (τ_1, τ_2, τ_3), the children bias voltages ($\dot{v}_{cix}$, $\dot{v}_{cqx}$, $\dot{v}_{ciy}$, $\dot{v}_{cqy}$), and the parent bias voltages (v_{px}, v_{py}). The fitness functions from each of the species are the CTA, MCTA, and the variance of the photodetected current, respectively. Figure 4.9 shows a block diagram of the CC-GA implementation. It starts by creating a random population for each of the species. The data processing is done

for one species at a time. For the first species, each specimen is randomly associated with one specimen from each of the other species elite group. For instance, if time skews are the first species, each time skews' specimen will be associated to one elite children bias voltages' specimen and one elite parent bias voltages' specimen. These cooperative combinations are evaluated though the first fitness function, and genetic operations of elite selection, crossover and mutation are performed. It is important to note that a one-to-one mapping is not necessary, so one specimen from a different species' elite group can be associated with more than one specimens of the species being evaluated. The elite group for the first species is then updated and the process is repeated regarding the next species. After the last species evaluation, a new offspring is generated and then the whole process is repeated until a stopping criteria is met. At that point the best specimen of each species is selected, forming the final solution.

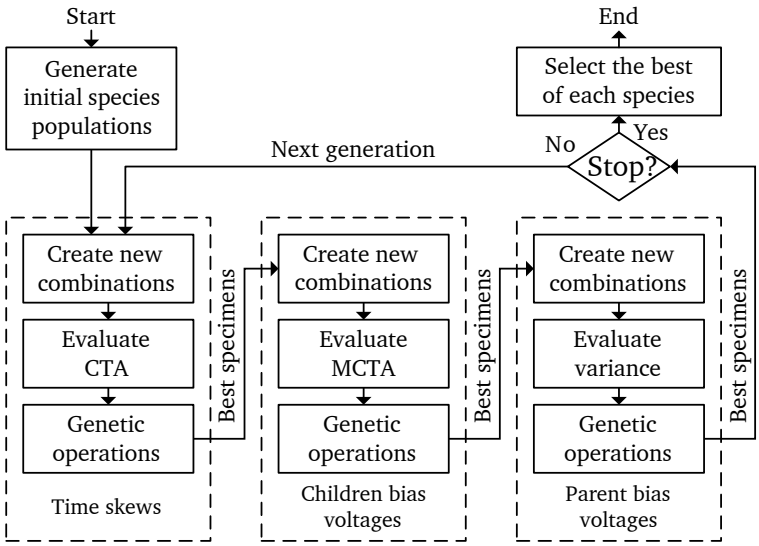

Figure 4.9.: Cooperative coevolutionary genetic algorithm block diagram. Based on [J1].

4.2.2 Proposed method: CC-GA + turn-on/turn-off procedure

An M-QAM reference signal is generated in the transmitter and used to estimate the transmitter impairments. This reference signal should be a PM signal operating in the quasi-linear region of the modulator, and should have a certain pulse-shape, modulation format and symbol rate. To avoid the

influence of transmitter laser phase noise and coherent receiver impairments, the signal is detected by employing direct detection. The output photocurrent is sampled in an analog-to-digital converter (ADC) with a sampling frequency greater than twice the symbol rate being used.

The fitness functions are then calculated from the sampled signal. These fitness functions are supplied to the CC-GA, which iterates the bias voltages and time skews by evaluating the associated fitness functions. Between each CC-GA generation, the best time skews and bias voltages are set and an iteration of a turn-on/turn-off procedure as presented in Section 4.1.3 is done for amplitude mismatch estimation and compensation. The best solution is chosen as the finishing parameters when a stop criterion is met. The schematic for the transmitter front-end calibration is presented in Figure 4.10. Since all transmitter parameters are optimized after the calibration ends, it is possible to change the reference signal to the signal to be transmitted with a different pulse shape, modulation format, and/or symbol rate.

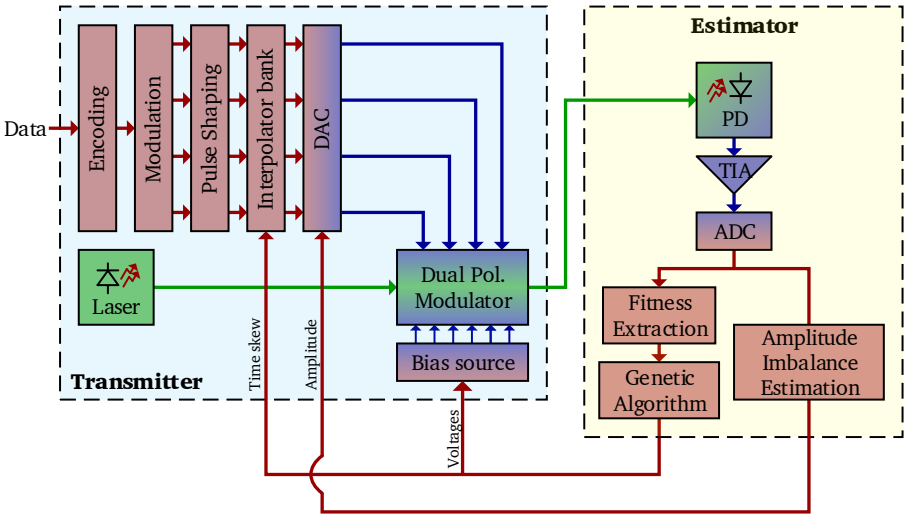

Figure 4.10.: Proposed algorithm scheme. Based on [J1].

4.3 Results and analysis

In this Section, the performance and operation of the proposed CC-GA method is analyzed by means of simulations and a demonstration is done by an experiment.

4.3.1 Simulation analysis

In order to evaluate the performance of the proposed method, the simulation setup of Figure 4.11 was used. First, sequences of bits are generated at a pseudo random bit sequence (PRBS) generator with length 31. These sequences are mapped to a modulation format at 1 sample per symbol and then filtered with a raised cosine (RC) pulse shape at 2 samples per symbol and roll-off 1, emulating an NRZ pulse shaping. The signal is then quantized with 8-bit resolution to emulate the limitations of a digital to analog converter. The signal is then resampled to a new sampling rate with each of the signal components being time delayed in order to account for the time skews. The signal components are then low-pass filtered and have their peak-to-peak values adjusted, emulating an electrical driver. The components are finally fed to a PM modulator, with variable bias voltages. Noise is added to the signal, which is then received in a PD. The cost functions are computed from the photodetected signal and fed to a parameter controller that will control the time skew, amplitude and bias voltage values accordingly to the proposed method.

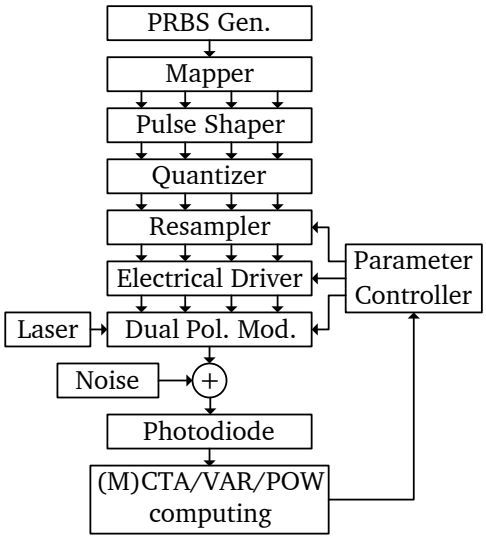

Figure 4.11.: Simulation setup for transmitter optimization evaluation. Based on [J1].

Unless stated otherwise, all the simulation results were extracted by the aforementioned setup with a PM NRZ 16QAM reference signal operating at 16 GBd. The quantity of symbols used for each fitness function computation

was 16384. The optical modulator was an ideal PM modulator with $V_\pi = 4\ V$, and no additional time skew. The reference peak-to-peak input signal voltages were selected to be 1.6 V, in order to operate inside the quasi-linear region of the optical modulator when correctly biased. Noise was added to the signal to guarantee an 18 dB OSNR (0.1 nm resolution) at the output of the modulator. The number of generations considered for the genetic algorithm was 50, and the time skew, child bias voltage, and parent bias voltage populations were 60, 80, and 50 specimens, respectively, in which, after each generation, 40% of the specimens were selected as elite, 50% mutated, and 10% passed through a crossing-over process. The number of simulation runs to compute the accuracy of the method was 100.

First, to evaluate the convergence speed of the proposed method, the above - mentioned simulation parameters were used. The estimated parameters after each generation are depicted in Figures 4.12 to 4.15. These estimated values are the average of the values of all elite specimens. The expected values for this simulation were −4 V for the children bias voltages (equivalent to $-V_\pi$), 90° for the IQ phase, 0 for the time skews and 1 for the relative amplitudes.

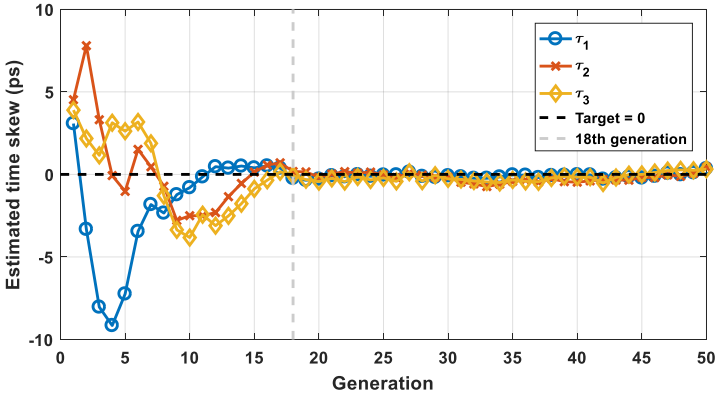

Figure 4.12.: Evolution through generations of the estimated time skew values. Based on [J1].

It is possible to observe in Figures 4.12 to 4.15 that after the 18^{th} generation all values seem to have converged to their expected value. This means that 3420 fitness-functions calculations and 54 power amplitude measurements were needed to converge to the final estimated values. Thus, the total calibration time would be limited by the amount of time needed for the IQ modulator stabilization and the optimum population sizes for the proposed

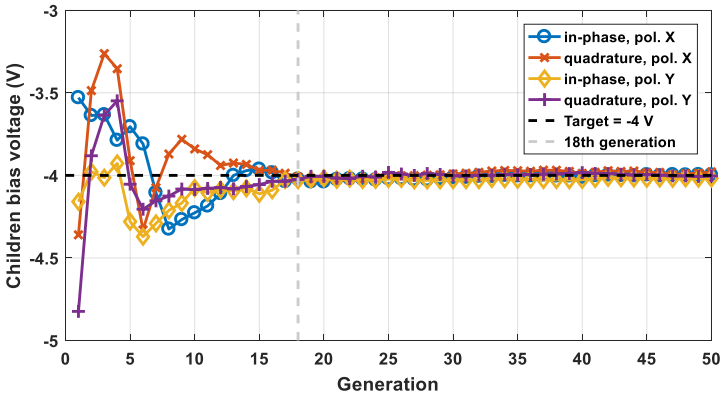

Figure 4.13.: Evolution through generations of the estimated child bias voltage values. Based on [J1].

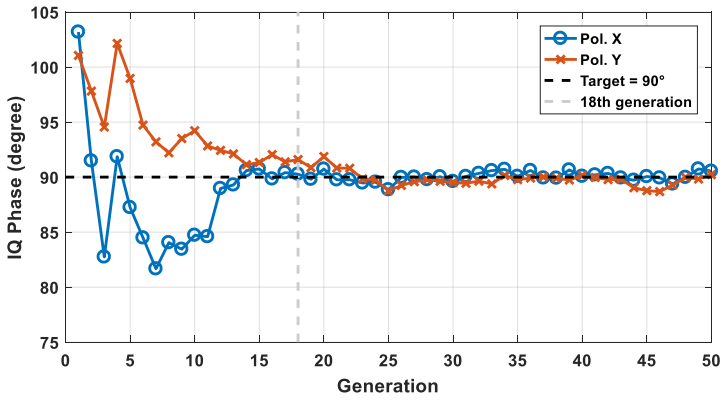

Figure 4.14.: Evolution through generations of the estimated IQ phase imbalance values. Based on [J1].

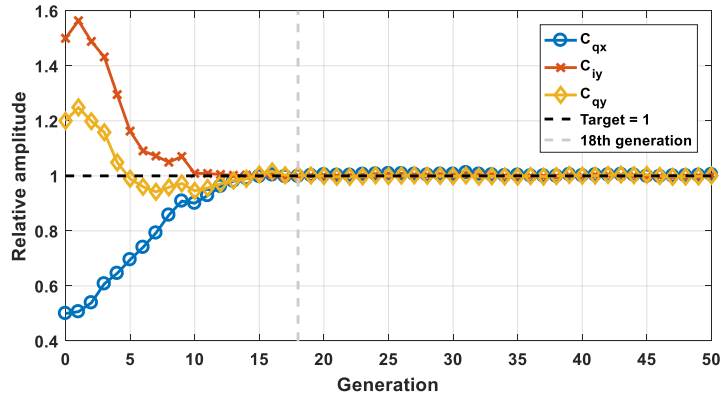

Figure 4.15.: Evolution through generations of the estimated amplitude imbalance values. Based on [J1].

algorithm. Nevertheless, if compared with full BER-based optimization, the proposed method is believed to significantly reduce calibration time since it significantly reduces the required data processing.

Another approach for convergence verification is to evaluate the evolution of the fitness values used at the CC-GA algorithm. The aforementioned evolution is depicted in Figures 4.16 to 4.18.

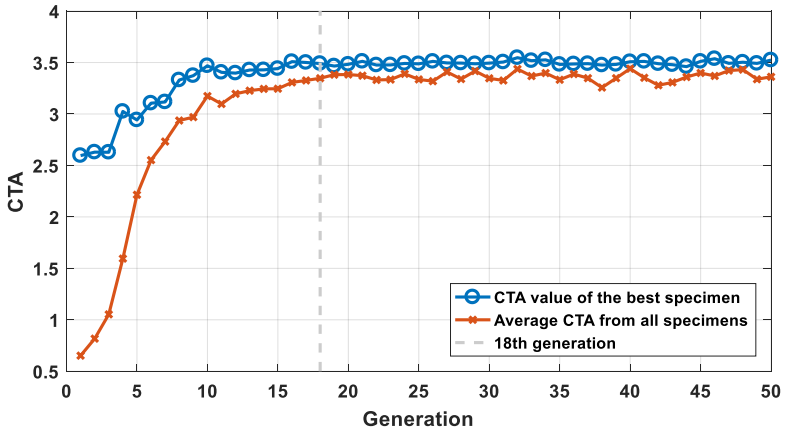

Figure 4.16.: Evolution through generations of the CTA fitness values. Based on [J1].

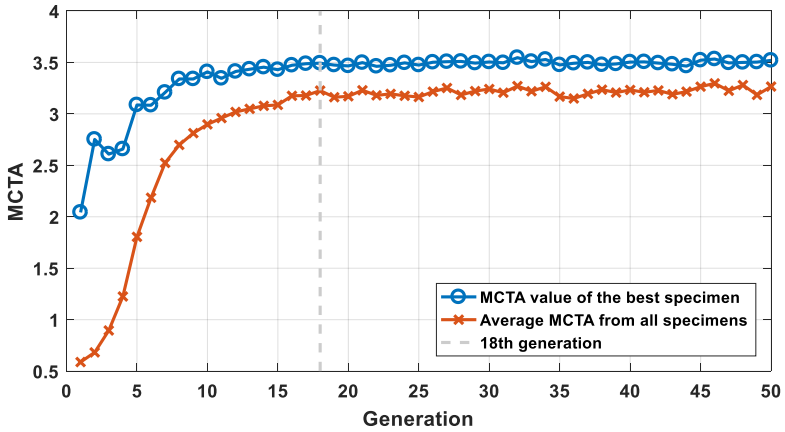

Figure 4.17.: Evolution through generations of the MCTA fitness values. Based on [J1].

If the average fitness values of all specimens approximate the fitness value of the best specimen, all specimens are therefore close to the optimal solution. A stopping criteria may be set when the average fitness function of all

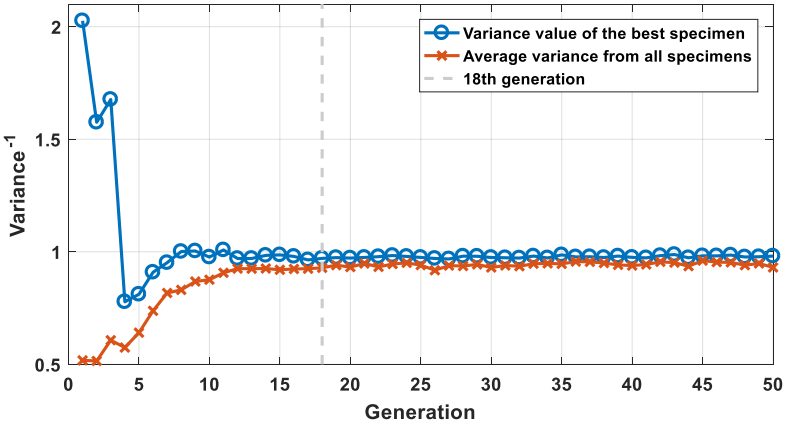

Figure 4.18.: Evolution through generations of the inverse of the variance fitness values. Based on [J1].

specimens exceeds a percentage of the fitness value of the best specimen. In this case, the stopping criteria would be 95.1%, 92.3%, and 95.5% for the CTA, MCTA and inverse variance fitness functions.

To evaluate the proposed method performance the accuracy and precision of the method was measured through Monte Carlo simulations. The accuracy and the precision were assessed by the absolute mean estimation error and the standard deviation relative to the target values, respectively. First, the performance using reference signals with different modulation formats was assessed by running the proposed method 100 times and then analyzing the last 10 generations from each of the iterations. In this way, the data size used to compute the mean and standard deviation was 1000. The results for the absolute mean estimation error and for the standard deviation are shown in Figure 4.19 and in Figure 4.20, respectively.

The standard deviations are one order of magnitude larger than the mean estimation error. This means that the method can be considered very accurate, as the errors in the estimation process are more random than systematic. The best reference signal in this case was the NRZ 16QAM that had standard deviations of 0.56° for the IQ phases, 0.019 V for the children bias voltages, 0.24 ps for the time skews and 0.003 for the amplitude imbalance. The distribution of the estimated values was Gaussian shaped, meaning that 99.7% of the estimations are expected to fall inside an interval of 3 times the standard deviation.

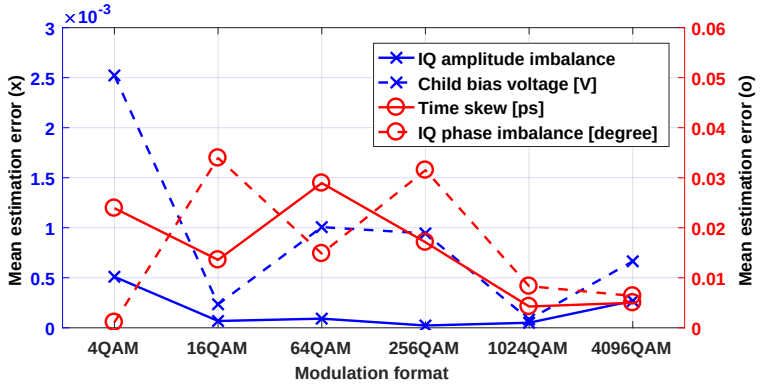

Figure 4.19.: Absolute mean estimation error compared to target values for different modulation formats. Based on [J1].

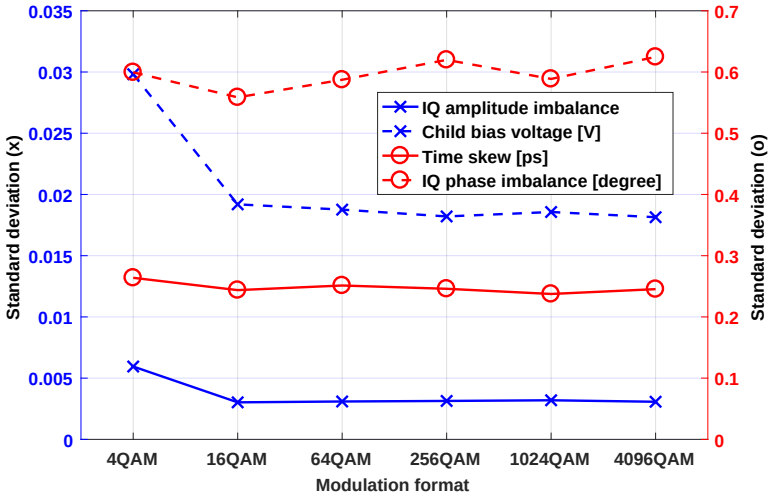

Figure 4.20.: Standard deviations for different modulation formats. Based on [J1].

The proposed method was also characterized by testing the number of symbols used to calculate each fitness function. The results are shown in Figure 4.21 and Figure 4.22.

Again, the standard deviations are one order of magnitude larger than the mean estimation error. Increasing the number of symbols used to compute every fitness value will increase the precision of the method. The number of symbols used in each fitness function computation will influence how fast the method convergence speed can be. The results show that doubling the number of symbols from 65536 to 131072 has a higher impact than doubling further on.

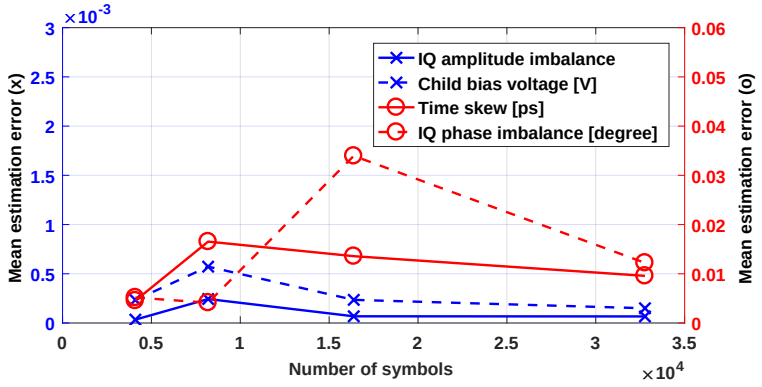

Figure 4.21.: Absolute mean estimation error compared to target values for different number of symbols for each cost-function calculation with NRZ 16QAM reference signal. Based on [J1].

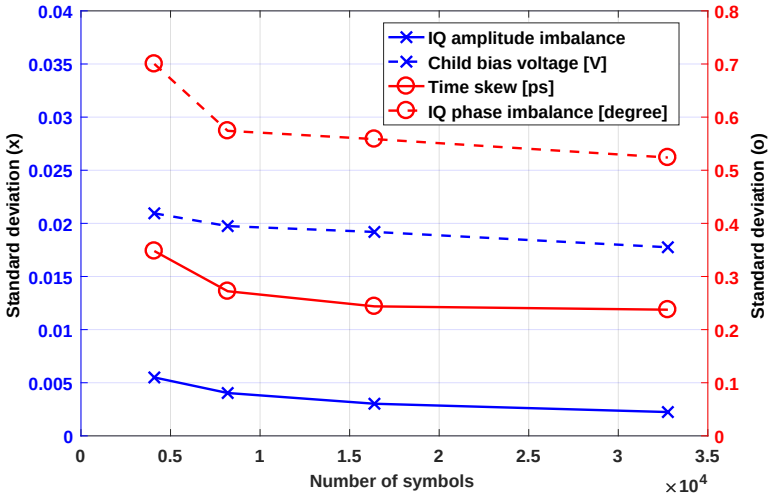

Figure 4.22.: Standard deviations for different number of symbols for each cost-function calculation with NRZ 16QAM reference signal. Based on [J1].

The penalty in the optical signal-to-noise ratio (OSNR) required to achieve a given bit error ratio (BER) due to each of the transmitter impairments was measured as a performance metric. Signals operating at 16 GBd were generated with QPSK, 16QAM and 64QAM as modulation formats and then received in a common DSP-based PM coherent receiver with decision-directed least-mean-square MIMO equalization, blind phase search carrier recovery, and standard decision regions for bit demapping[15]. The OSNR penalty was numerically measured when the signals were transmitted through an

additive white Gaussian noise channel by varying the noise variance. The results are shown in Figure 4.23.

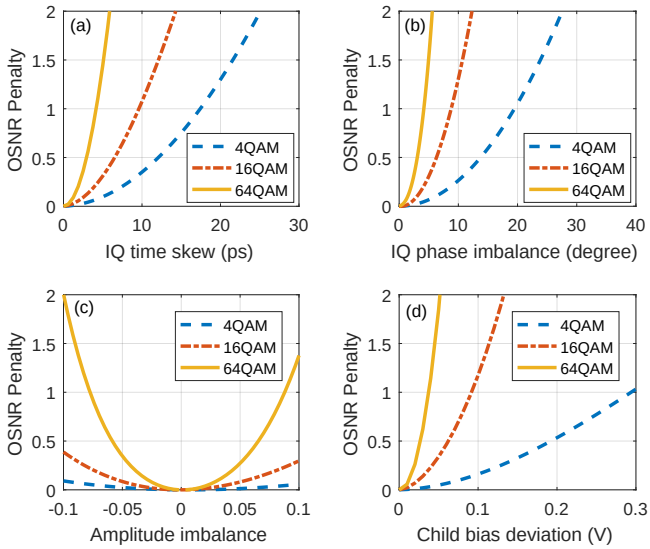

Figure 4.23.: OSNR penalty for a 16-GBd signal, $V_\pi = 4$ V, and peak to peak voltages of 1.6 V at BER $= 3.8\times10^{-3}$ for QPSK and 16QAM and at BER $= 10^{-2}$ for 64QAM, due to: (a) IQ time skews; (b) IQ phase imbalance (c) Amplitude imbalance; (d) Child bias voltage error. Based on [J1].

Then, considering the worst case scenario as each of the transmitter impairments calibrated with an error of 3 times the standard deviation, the OSNR penalties at BER of 3.8×10^{-3} was numerically measured for QPSK and 16QAM signals as 0.05 dB and 0.5 dB, respectively. For the 64QAM signal a 1.9 dB OSNR penalty at BER equals to 10^{-2} was measured. These penalties were mainly due to the child bias voltages and could be drastically reduced to 0.05 dB and 0.2 dB OSNR penalty for the 16QAM and 64QAM modulation formats, respectively, if a simple change in the decision regions at the receiver DSP is considered.

In comparison to other algorithms, IQ time skew can be calibrated using the method presented by Fludger *et al.* [130] with typical accuracy of 0.5 ps. Additionally to the IQ time skews, the method presented by Yue *et al.* [129] could find also XY time skews in the range of 0.5 ps. The method presented here have similar performance compared to these alternative methods while also calibrating the IQ phase imbalance, the IQ amplitude imbalance, and the bias operation voltages.

4.3.2 Experimental demonstration of time skew and operation point calibration with CC-GA

Finally, to demonstrate the method behavior an experiment with a PM modulator is reported. Four output channels of a 64-GSa/s arbitrary waveform generator (AWG) were applied to a PM Mach-Zehnder (MZ)-based IQ modulator, used to generate a reference signal at 16 GBd and NRZ-PM-16QAM modulation format. The generated signal was amplified by an erbium-doped fiber amplifier (EDFA), directly-detected in a 45-GHz bandwidth PD and then sampled by a digital storage oscilloscope (DSO) operating at 160 GSa/s. The fitness functions for the CC-GA were calculated on a personal computer that was also used to automatically control the time skew pre-compensation values in the AWG and the modulator bias voltages. The time skews and the correct operation points for this setup were previously unknown, with initial voltages being random and not resulting into recoverable constellations. The number of generations considered for the CC-GA was 30, and the time skew, child bias voltage, and parent bias voltage populations were 60, 80, and 50 specimens, respectively, in which, after each generation, 40% of the specimens were selected as elite, 50% mutated, and 10% passed through a crossing-over process. Due to time constraints, the turn-on/turn-off procedure was not assessed. The experimental setup is shown in Figure 4.24.

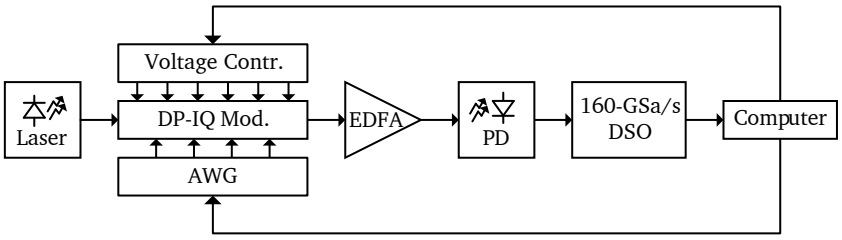

Figure 4.24.: Experimental setup for time skew and operation point calibration. Based on [J1].

The experiment was ran with the parameters stated above in order to assess the convergence. The estimated parameters after each generation are depicted in Figure 4.25, while the evolution of the fitness functions is depicted in Figure 4.26.

The estimated values are the average of the values of all specimens selected as elite and the fitness values are the CTA, MCTA and inverse variance

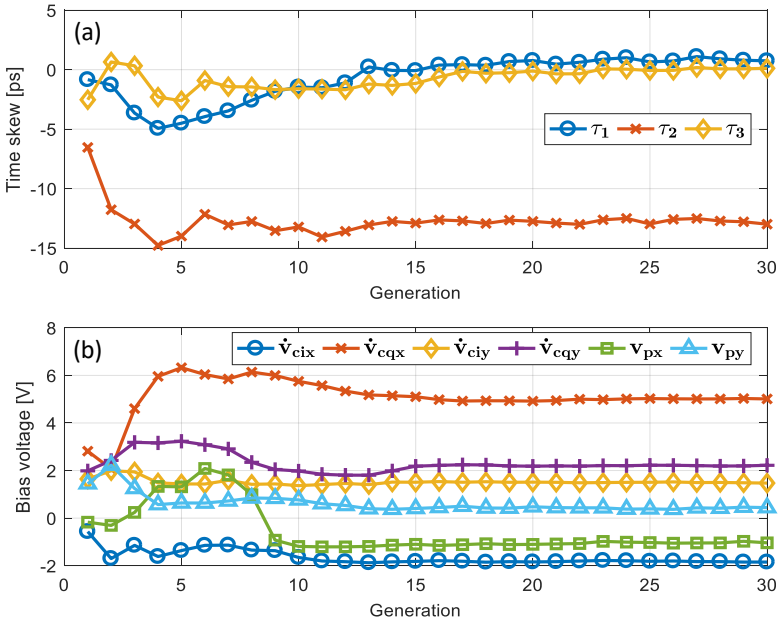

Figure 4.25.: Experimental evolution through generations of the estimated values. (a) Time skews. (b) Children and parent bias voltages. Based on [J1].

computed after each generation from the best specimen and an average of all specimens. As in the simulations, it can be seen in Figure 4.25 and Figure 4.26 that after the 18^{th} generation all variables have converged to their final values. With these final values, a 16 GBd NRZ-PM-16QAM was once again generated and received in a coherent receiver. After DSP offline processing, consisting of resampling to 2 samples per symbol, adaptive equalization based on common 2×2 MIMO decision-directed least mean squares algorithm, and carrier phase recovery based on blind phase search (BPS), it has been obtained the constellations shown in Figure 4.27.

From the recovered constellation, it can be qualitatively observed that the bias voltages were correctly estimated, while the IQ phase had a small error of about 1° in Y, which was within the accuracy of the simulations. This IQ phase error gives an insignificant OSNR penalty for a 16QAM signal at BER $= 10^{-2}$ threshold and would still give a penalty of less than 1-dB in OSNR if the modulation format used is changed to 256QAM[22].

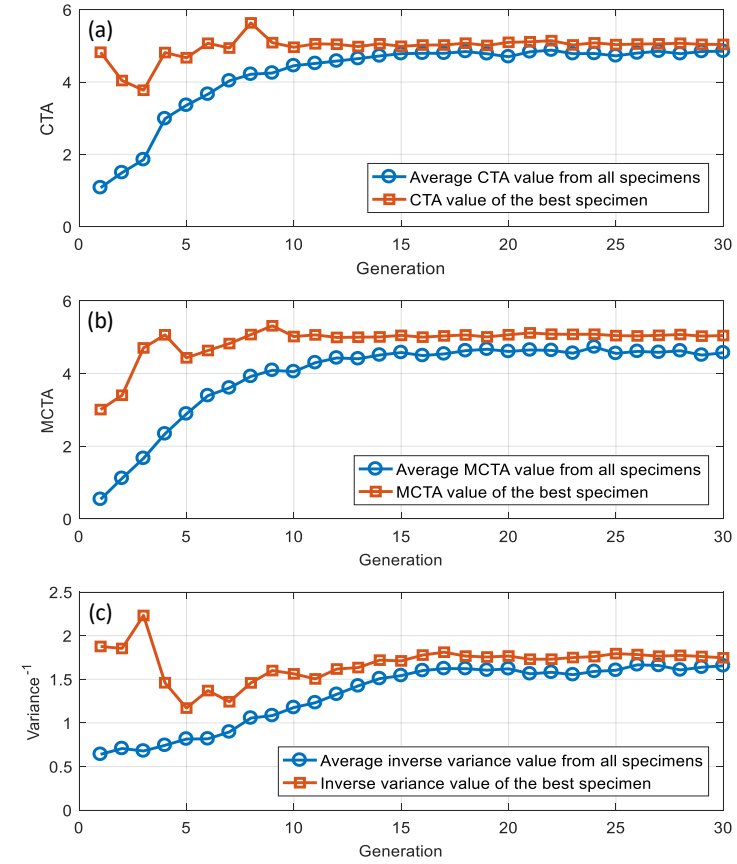

Figure 4.26.: Experimental evolution through generations of the fitness values. (a) Clock tone amplitude. (b) Modified clock tone amplitude. (c) Inverse of the variance. Based on [J1].

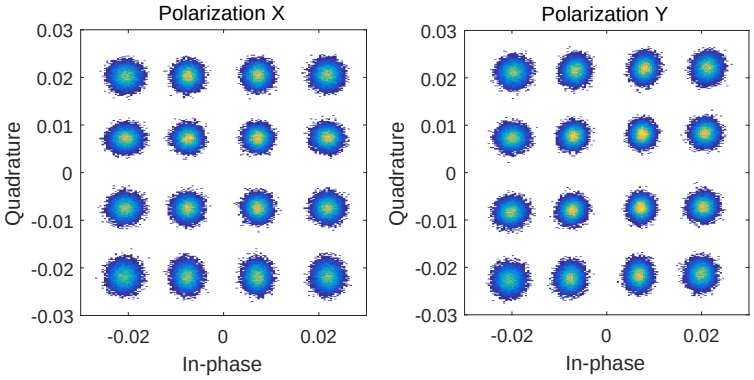

Figure 4.27.: Obtained 16 GBd PM-16-QAM constellations after optimization using the proposed method. Based on [J1].

4.4 Summary

A new and simple method has been suggested to optimize the common transmitter's front-end imperfections such as suboptimal biasing, time skews, and amplitude, and phase imbalances between I and Q components. This was achieved by using a cooperative coevolutionary genetic algorithm. This method is carried out on the transmitter side, so that the transmitter can be optimized automatically for the best operation regardless of the coherent receiver and thus prevent the complexity of the stressed DSP at the receiver from increasing. Simulations numerically evaluated the performance of the presented method and experiments were carried out to show the behavior of the method. The results also demonstrate the potential of the cooperative coevolutionary genetic algorithm as a rapid optimization method for fine-tuning and mitigating the transmitter impairments. In the next Chapter, a method for jointly estimating CD and IQ time skews of a coherent receiver is presented.

5

Joint IQ Time Skew and Chromatic Dispersion Estimator

DUE TO THE COMPLEXITY of a typical coherent front-end, it is common to have a timing misalignment (time skew) of several picoseconds in between the in-phase (I) and quadrature quadrature (Q) paths. These time skews are commonly calibrated during the manufacturing process and compensated through a bank of interpolators at the input of an application specific integrated circuit (ASIC) [15, 121, 136]. However, some residual time skew can remain. At higher symbol-rate, this post-calibration residual skew has a more severe impact than on lower symbol-rate transmissions due to the larger ratio of skew to symbol period. Thus, for high symbol-rate transmissions, an algorithm robust to in-phase/quadrature (IQ) time skew or a method to estimate the IQ time skew is required.

As discussed in the previous Chapter, common polarization demultiplexing algorithms that are based in 2×2 complex-valued multiple-input multiple-output (MIMO) architectures [15] are able to compensate only time skew between polarizations. The skew-robust MIMO equalizers proposed recently [33, 34, 126] introduce more complexity and demand more area in practical ASIC implementations in comparison to common complex-valued 2×2 MIMO equalizers.

Analogous to the transmitter case, adaptive equalization with the purpose of time skews would be avoidable. Then, a non-data-aided (NDA), joint chromatic dispersion (CD) and receiver front-end-IQ time skew estimator is proposed. This algorithm is a two-dimensional scanning algorithm based on well-known scanning CD estimation methods [70, 137–140]. It relies on the fact that the output of the CD compensation in the presence of IQ

time skew is a weighted combination of the CD compensated signal and its complex conjugate carrying twice as much dispersion. Hence the signal remains impaired by CD, by an amount proportional to the skew [126]. The proposed algorithm does not require any change in the signal path, thus it can be implemented without any major increase in complexity, saving power and space in practical implementations. A 32 GBd polarization multiplexed (PM)-16QAM experiment is carried out to evaluate the algorithm's performance. The results and analyses that are presented here were based on the results in the author's original work in [C3].

The structure of this Chapter is the following. Section 5.1 presents and reviews the mathematical foundations of the method. Section 5.2 presents the method. Section 5.3 shows an experimental analysis of the method.

5.1 CD compensation with IQ time skew at the receiver

If a single-polarization signal has time skew between its IQ components, $s_{\mathrm{i}}(t)$ and $s_{\mathrm{q}}(t)$, and assuming τ as the relative time skew between the I and Q components of a non-dispersed signal, $s_{\tau}(t)$ then

$$s_{\tau}(t) = s_{\mathrm{i}}(t) * \delta\left(t - \frac{\tau}{2}\right) + j s_{\mathrm{q}}(t) * \delta\left(t + \frac{\tau}{2}\right), \tag{5.1}$$

$$\begin{aligned} s_{\tau}(t) = \frac{s(t)}{2} * \left[\delta\left(t - \frac{\tau}{2}\right) + \delta\left(t + \frac{\tau}{2}\right)\right] + \\ \frac{\bar{s}(t)}{2} * \left[\delta\left(t - \frac{\tau}{2}\right) - \delta\left(t + \frac{\tau}{2}\right)\right], \end{aligned} \tag{5.2}$$

where $s(t) = s_{\mathrm{i}}(t) + j s_{\mathrm{q}}(t)$ is the complex transmitted signal, $*$ represents the convolution operation, and the over-bar ($\bar{}$) represents the complex conjugate.

Considering now a transmission with distance L and CD parameter D_C, the impulse response of the optical fiber channel is $h_{D_C L}(t)$, and finally the IQ time-skewed dispersed signal $s_{D_C L,\tau}(t)$ can be obtained as

$$\begin{aligned} s_{D_C L,\tau}(t) = \frac{s(t)}{2} * \left[h_{D_C L}\left(t - \frac{\tau}{2}\right) + h_{D_C L}\left(t + \frac{\tau}{2}\right)\right] + \\ \frac{\bar{s}(t)}{2} * \left[\bar{h}_{D_C L}\left(t - \frac{\tau}{2}\right) - \bar{h}_{D_C L}\left(t + \frac{\tau}{2}\right)\right], \end{aligned} \tag{5.3}$$

In frequency domain, the IQ time-skewed signal is expressed as

$$S_{D_C L,\tau}(\omega) = S(\omega) H_{D_C L}(\omega) \cos\left(\omega \frac{\tau}{2}\right) - j\bar{S}(-\omega) \bar{H}_{D_C L}(\omega) \sin\left(\omega \frac{\tau}{2}\right), \tag{5.4}$$

where $S(\omega)$ and $H_{D_C L}(\omega)$ are the transmitted signal in frequency domain and the CD transfer function, respectively. Assuming zero-forcing CD compensation, one obtains the following

$$\begin{aligned} \hat{S}_{D_C L,\tau}(\omega) &= H^{-1}_{D_C L}(\omega) S_{D_C L,\tau}(\omega) \\ &= S(\omega) \cos\left(\omega \frac{\tau}{2}\right) - j\bar{S}(-\omega) \left[\bar{H}_{D_C L}(\omega)\right]^2 \sin\left(\omega \frac{\tau}{2}\right). \end{aligned} \tag{5.5}$$

It is important to point out, considering the equations above, that if the time skew is not null ($\tau \neq 0$), the compensated signal contains a component which consists of the complex conjugate of the signal, nonetheless with doubled CD [126]. This component acts as an interference, which degrades the performance of clock synchronization algorithms [141], and thus may cause MIMO equalizer approaches to skew compensation [33, 34, 126] to fail.

5.2 Joint IQ time skew and CD estimation method

The proposed algorithm is based on bi-dimensionally scanning CD and IQ time skew values, similarly to the algorithms presented in [70, 137–140] for CD estimation. This is done by controlling a set of interpolator banks and the static equalizer, as shown in Figure 5.1, where the output of the static equalizer is followed by the timing recovery, MIMO equalization, and finally carrier recovery blocks.

In the algorithm sequence used, the data from the static equalizer is also received by the skew and CD estimator, which in turn controls the de-skew process and CD equalization. The goal of the estimator is to locate the optimum values, which minimizes a certain cost-function. One possible cost-function is based on the delay-tap sampling estimator (DTSE) [70, 138].

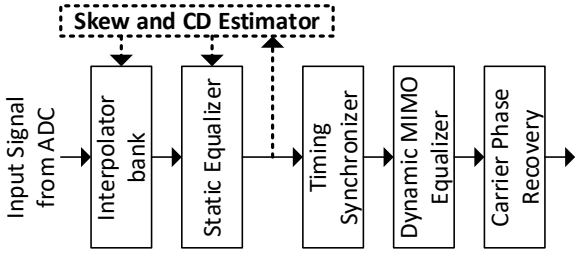

Figure 5.1.: Variation of common coherent optical communication receiver DSP algorithms chain with skew and CD estimator placement in dotted lines. Based on [C3].

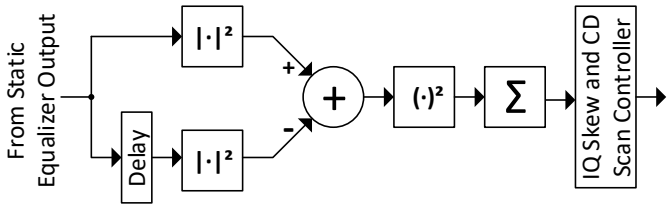

Figure 5.2.: Block diagram of the joint IQ time skew and CD estimation using DTSE cost-function. Based on [C3].

The DTSE cost-function block diagram is shown in Figure 5.2 and can be written as:

$$c\left(D_C L, \tau\right) = \sum_{i=d}^{M} \left[\left| \hat{s}_{D_C L, \tau}\left(i\right) \right|^2 - \left| \hat{s}_{D_C L, \tau}\left(i-d\right) \right|^2 \right]^2, \tag{5.6}$$

where M is the number of samples used for each point, d is a sample delay, and $\hat{s}_{D_C L, \tau}\left(i\right)$ is the received signal in time-domain after de-skewing process by a set of interpolator banks and CD compensation.

Figure 5.3 shows a typical surface of this cost-function for a received signal with $D_C L = 10^4$ ps/nm. The main target of the estimator is to find the minimum points shown of the surface. In Figure 5.4(a) it is possible to see a slice of this cost-function for CD compensation set to $D_C L = 10^4$ ps/nm. While Figure 5.4(b) presents a slice of this cost-function for zero time skew.

5.3 Experimental setup and results

The experimental setup shown in Figure 5.5 was used to benchmark the algorithms. The optical signal was modulated by an IQ modulator driven by two outputs of a 64 GSa/s arbitrary waveform generator (AWG) gen-

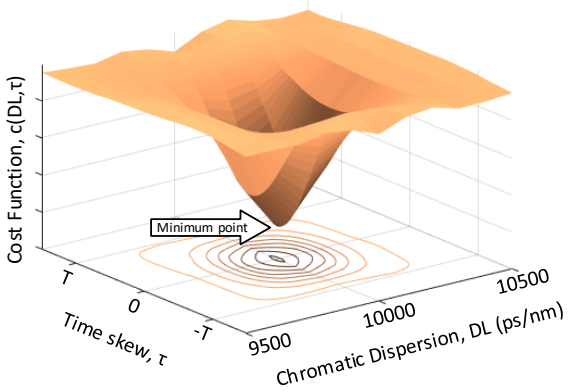

Figure 5.3.: Typical surface for DTSE cost-function while sweeping through IQ time skews and CD values. Based on [C3].

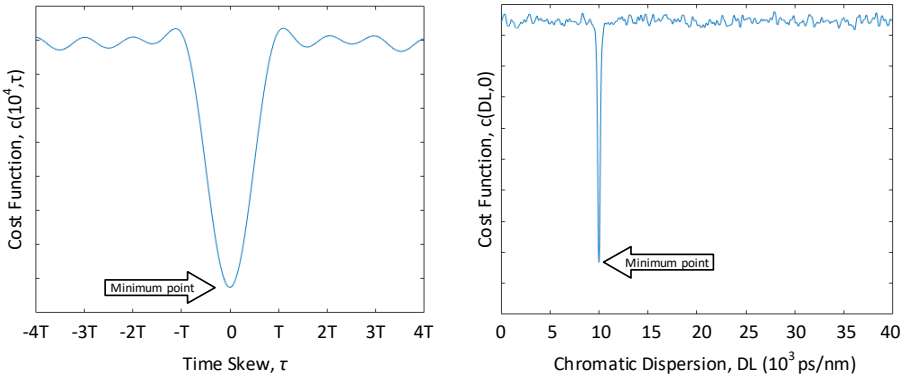

Figure 5.4.: Typical cost-function curves: (a) Only IQ time skew values are swept, with optimal CD compensation; (b) Only CD values are swept, with optimal skew compensation. Based on [C3].

erating 32 GBd 16-QAM signals. The modulated signal was then polarization multiplexed to generate a PM-16QAM signal. This signal transmitted through a recirculating loop with two 50-km Corning SMF-8 fiber spans with $D = 16.9$ ps/(nm.km), and Raman amplifiers to compensate fiber and switches losses. At the polarization-diversity receiver side, the electrical outputs were sampled by a real-time oscilloscope at 80 GSa/s. The output data was finally processed offline using the previously presented algorithm.

The analysis of the time skew estimation performance of the proposed method was focused here, since the DTSE was already extensively evaluated for CD estimation [70, 138–140]. It was considered 5 traces that were acquired in a experiment consisting of a 32 GBd PM-16QAM signal transmitted

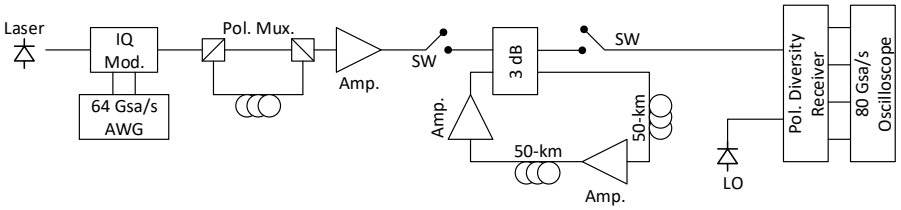

Figure 5.5.: Experimental setup for a 32 GBd PM-16QAM transmission with 100-km optical fiber per loop. AWG: arbitrary waveform generator; Based on [C3].

through 1200 km of a standard single-mode fiber (SMF) with an optical signal-to-noise ratio (OSNR) of approximately 20 dB. The IQ time skew of the received signal was swept in the digital domain for both polarizations at the same time from -62.5 to 62.5 ps with steps of 1.25 ps. The proposed method was used for estimating both CD and IQ time skew. For accumulated CD, all estimated values were between the interval of 20200 and 20360 ps/nm, which is consistent with the expected value of 20280 ps/nm for accumulated CD of a 1200-km standard SMF transmission. Figure 5.6 shows the estimated skew as a function of the actual time skew. The accuracy of the method is more clearly visible in Figure 5.7, which shows the absolute error between estimated value and actual values in function of actual time skew values. The average value of estimations for each time skew sweep was plotted in Figures 5.6 and 5.7. It is possible to observe that the maximum time skew estimation error for a single estimation, and the average of 10 estimations for each true IQ time skew value, was 3.1 ps, and 0.95 ps, respectively, representing 10% and 3% symbol period of a 32 GBd signal. So, the accuracy can be improved by computing and averaging consecutive estimations of the joint IQ time skew and CD estimator.

5.4 Summary

A low-complexity joint IQ time skew and CD estimator was proposed and experimentally demonstrated for a 32 GBd PM-16QAM transmission. The results show good accuracy for both IQ time skew and CD estimation. Furthermore, the proposed method has a small additional complexity since it is performed before MIMO equalization and timing recovery, avoiding complexity increase of skew tolerant MIMO equalizers [33, 34]. Also, the proposed method is able to deliver nearly free CD and time skew signals to a subsequent timing recovery block, benefiting its performance.

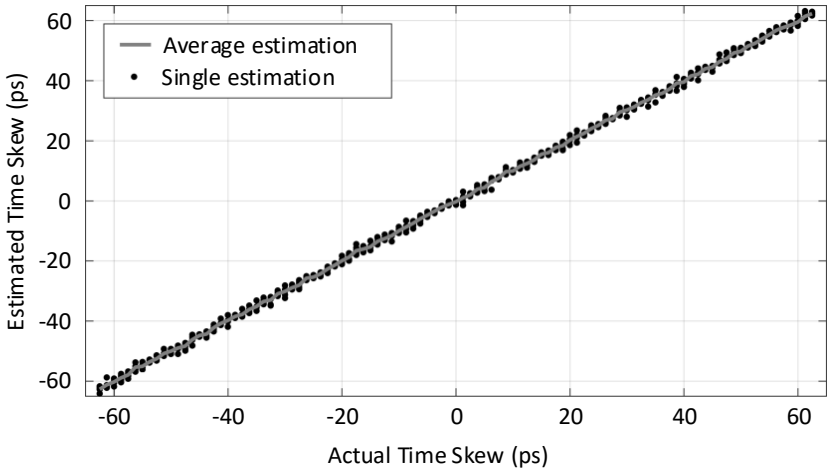

Figure 5.6.: Estimated IQ time skew versus actual IQ time skew. Based on [C3].

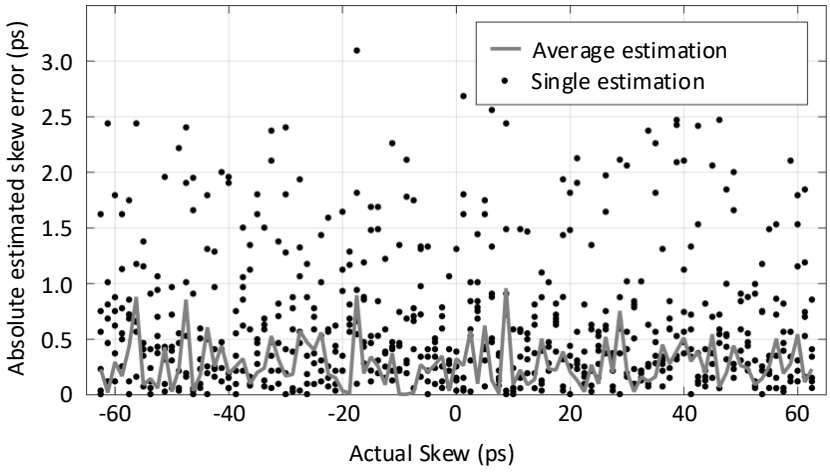

Figure 5.7.: Absolute error of estimation versus actual IQ time skew. Based on [C3].

6

Carrier Phase Estimation based on Principal Components

LASER PHASE NOISE is a major limiting factors in coherent optical communications. The conventional method for demodulating optical signals with a digital phase-locked loop (PLL) is highly sensitive to feedback delays, thereby limiting the laser linewidth (LW) tolerance of PLL-based carrier phase recovery (CPR) methods [142]. In addition, since the receiver-side digital signal processing (DSP) is performed at a rate much lower than the symbol rate, the processing must be carried out in parallel, making PLL-based procedures even less viable.

Unsupervised feed-forward CPR methods generally extract the phase noise either by applying a nonlinear relationship such as 4^{th} power [98, 103, 143], by using the constellation symmetry to fold it [144], or by massively testing different phase values in order to determine the symbols [31].

In this Chapter, a hardware-efficient and modulation order independent NDA feed-forward carrier recovery method for synchronous decoding of arbitrary QAM constellations is proposed. The proposed method is based on extraction of the principal component (PC) of the squared constellation and it is suitable for low signal-to-noise ratio (SNR) signal transmission.

This Chapter is based in the author's original work that will be submitted for publication [U1].

This Chapter is structured as follows. Section 6.1 reviews common feed-forward CPR methods, first presenting a parallelized structure for real world implementation and then presenting an ideal block-wise estimator and the blind phase search (BPS) algorithm. Section 6.2 presents both the

proposed PC-based method and a hybrid scheme with the proposed method as a coarse estimator and BPS as a fine estimator. Section 6.3 presents the computational complexity analysis and comparison between the CPR algorithms. Section 6.4 show the simulation analyses and the experimental demonstration. Finally, Section 6.5 summarizes the Chapter. Moreover, a mathematical proof of the presented method and the description of the complexity analysis are presented in the Appendices A and C.

6.1 Feed-forward phase estimation techniques

In a DSP-based coherent receiver, considering perfect equalization and timing synchronization, and an additive white Gaussian noise (AWGN) channel, the n-th received constellation symbol, $x[n]$, is rotated by an angle relative to the phase noise due to the transmitter and receiver lasers. So,

$$x[n] = x_{Tx}[n]e^{j\phi[n]} + w[n], \tag{6.1}$$

where $x_{Tx}[n]$ is a complex value representing the n-th transmitted symbol, $w[n]$ is an additive noise modeled by a zero-mean Gaussian random variable with variance σ_w^2, and $\phi[n]$ is the phase noise on the n-th symbol. Disregarding equalization-enhanced phase noise (EEPN), the phase noise can be described by a random walk process,

$$\phi[n] = \phi[n-1] + \Delta\phi[n], \tag{6.2}$$

where $\Delta\phi[n]$ is the random phase increment modeled by a zero-mean Gaussian random variable with variance given by

$$\sigma_\phi^2 = 2\pi\Delta\nu T_S, \tag{6.3}$$

where $\Delta\nu$ is the combined laser LW, *i.e.*, the sum of the transmitter and receiver lasers LWs, and T_S is the symbol period.

In practice, the DSP unit of the receiver is implemented in an ASIC that by the current technology cannot operate at the same clock frequency as the symbol rate, therefore requiring to process the data in m parallel stages at reduced clock frequencies [146]. Therefore, phase estimation algorithms

should be able to process the data in parallel with blocks of at least m sequential symbols being processed at each ASIC clock pulse.

Phase recovery algorithms rely on the fact that phase noise varies slowly over time, then it would be possible to estimate and compensate a unique phase value for each incoming signal block. Taking this into consideration, the parallelized input data block of $m+h$ symbols, is

$$\vec{x}[k] = \begin{bmatrix} x[m(k-1)-\frac{h}{2}+1] \\ x[m(k-1)-\frac{h}{2}+2] \\ \vdots \\ x[mk+\frac{h}{2}] \end{bmatrix}, \tag{6.4}$$

where h is the symbol overlap length between consecutive blocks.

6.1.1 Ideal block-wise phase estimation

The ideal estimation for the phase recovered signal is

$$x_{rec}[n] = x[n]e^{j\bar{\phi}[\lceil n/m \rceil]}, \tag{6.5}$$

where n is the time-index of each received symbol, $k = \lceil n/m \rceil$ is the index of each parallelized input data block, and $\bar{\phi}[k]$ is the average phase noise for the symbols of a block,

$$\bar{\phi}[k] = \frac{1}{m+h} \sum_{n=m(k-1)-\frac{h}{2}+1}^{mk+\frac{h}{2}} \phi[n]. \tag{6.6}$$

For a sufficiently low $m\Delta\nu T_S$ product, the residual phase noise $\phi_{res}[n] = \phi[n] - \bar{\phi}[\lceil n/m \rceil]$, although not Gaussian distributed (*i.e.* not stationary), has a probability distribution function that resembles a Gaussian shape, with variance $\sigma_{\phi_{res}}$ depending on this product.

6.1.2 Minimum-distance blind phase search

The minimum-distance BPS algorithm [31] is based on testing B phase values by rotating the received symbols and comparing them to their respective estimates after decision. The angle that minimizes the mean squared distance

between the rotated and the decided symbols is then the estimated phase to be compensated. Figure 6.1 shows a block diagram representation of the BPS algorithm.

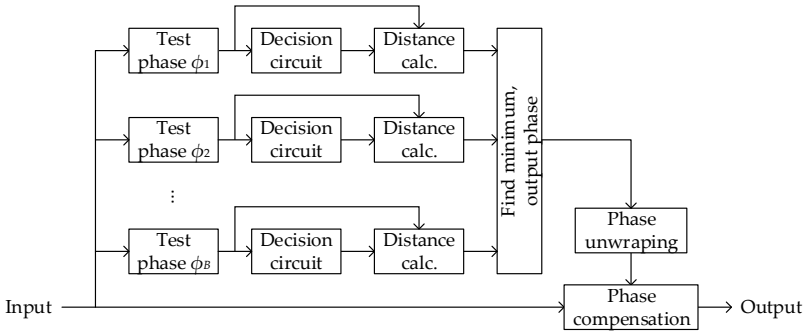

Figure 6.1.: Block diagram of the minimum distance blind phase search carrier phase recovery method.

For a parallelized input given by Eq. 6.4, the b-th test constellation, rotated by the b-th test phase, ϕ_b, is

$$\vec{x}_b[k] = \vec{x}[k]e^{j\phi_b}, \tag{6.7}$$

where ϕ_b is given by

$$\phi_b = \frac{(2b-1)\pi}{4B} - \frac{\pi}{4}, \qquad b \in \{1, 2, \ldots, B\}. \tag{6.8}$$

Then, the test constellations are fed into a decision circuit. The sum of the squared distances between these test constellations and their respective decisions,

$$d_b[k] = \sum_{i=1}^{m+h} \left| x_b[k,i] - \hat{x}_b[k,i] \right|^2, \tag{6.9}$$

is computed. In this equation, $x_b[k,i]$ is the i-th position of the rotated input vector $X_b[k]$, and $\hat{x}_b[k,i]$ is the i-th position of the rotated input vector after decision. The estimated phase is the phase associated to the minimum computed distance,

$$\hat{\phi}[k] = \arg\min_{\phi_b} d_b[k] \tag{6.10}$$

Then, to avoid the $\pi/4$ ambiguity, a phase unwrapping process is done to update the estimated phase,

$$\hat{\phi}[k] := \hat{\phi}[k] + \left\lfloor \frac{1}{2} + \frac{\hat{\phi}[k-1] - \hat{\phi}[k]}{\pi/2} \right\rfloor \frac{\pi}{2}, \tag{6.11}$$

that is therefore applied to the input block to give the output phase compensated block,

$$\hat{\vec{x}}[k] = \vec{x}[k] e^{j\hat{\phi}[k]}. \tag{6.12}$$

6.1.2.1 Two-stage minimum-distance BPS

A common implementation for complexity reduction of the BPS is to break it into two stages with B_1 and B_2 test phases in the first and the second stages, respectively. In this way, instead of testing B phases, the algorithm test up to $B_1(B_2+1)$ different phase values while the complexity is proportional only to $B_T = B_1 + B_2$. The algorithm starts similarly to the one-stage BPS, with the first stage performing a coarse search for the phase by rotating the input block by B_1 test phases,

$$\phi_{b,1} = \pi \left\lfloor \frac{2b-1}{4B_1} - \frac{1}{4} \right\rceil, \qquad b \in \{1, 2, \ldots, B_1\}. \tag{6.13}$$

Then, the second stage performs a fine phase search with test phases given by

$$\phi_{b,2}[k] = \hat{\phi_1}[k] + \frac{\pi}{B_2} \left\lfloor \frac{(2b-1)}{4B_1} - \frac{1}{4} \right\rceil, \qquad b \in \{1, 2, \ldots, B_2\}, \tag{6.14}$$

where $\hat{\phi}_1[k]$ is the phase estimated at the first stage. A phase unwrapping process and the rotation of input block follows. Figure 6.2 shows a block diagram representation of the two-stage blind phase search (BPS-BPS) algorithm.

6.2 Phase tracking based on principal components

Principal component analysis (PCA) is a well-known algorithm that converts a set of observations of a likely correlated random variable into a set of variables linearly decorrelated, the PCs [147]. It is usually applied as a

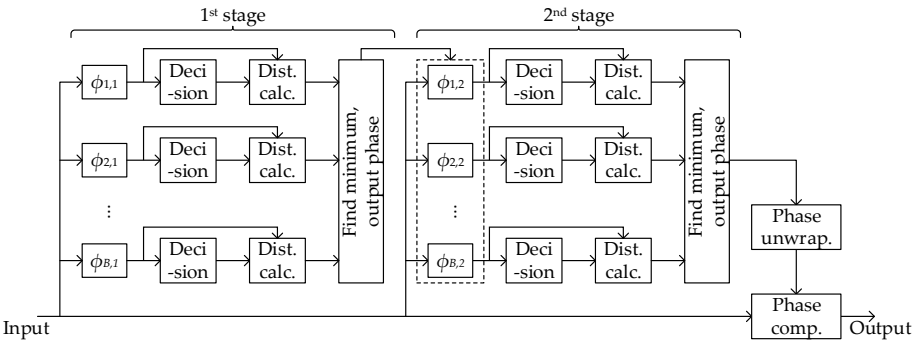

Figure 6.2.: Block diagram of the minimum distance blind phase search carrier phase recovery method. Based on [U1].

pre-processing tool in pattern recognition applications for extraction of the most critical data features, projecting the original data into a new feature space in which the first components keep the most of information about the original data.

Basically, PCA considers the features that display the highest variation in the data as the most informative, with their degree of variation being used to distinguish the most relevant data from the less relevant. Geometrically, PCA can be understood as a rotation of the original coordinate system axes into a new set of orthogonal axes in which they are ordered based on the amount of variation of the original data they account for.

More specifically, for an input data set $\{\vec{a}_1, \vec{a}_2, \vec{a}_3, \cdots, \vec{a}_L\}$ with L data vectors of N dimensions, the $N \times N$ covariance matrix, $\boldsymbol{C}$, given by

$$\boldsymbol{C} = \frac{1}{L} \sum_{i=1}^{L} \{(\vec{a}_i - \mathbb{E}\{\vec{a}\})(\vec{a}_i - \mathbb{E}\{\vec{a}\})^{\top}\}, \tag{6.15}$$

where $\mathbb{E}\{\vec{a}\}$ is the average vector of the input data set, has up to N eigenvalues, each with an eigenvector associated. The eigenvalue magnitude is proportional to the amount of information that is carried on the direction of its associated eigenvector. Thus, the most important PC is the eigenvector associated with the eigenvalue of highest magnitude. This PC is also known as the first PC.

6.2.1 Phase estimation method based on principal component

The PCs extracted directly from a QAM constellation are invariant to constellation rotation, as shown in Appendix A.1. But, if the received constellation is squared, the PCs extracted from this new signal have the PC angle proportional to the original constellation rotation, as shown in Appendix A.2. Figures 6.3(a) and 6.3(b) show 64QAM constellations without and with a $\pi/6$ rad phase rotation, respectively. Figures 6.3(c) and 6.3(d) show the same 64QAM constellations with and without $\pi/6$ rad phase rotation, but after the squaring process.

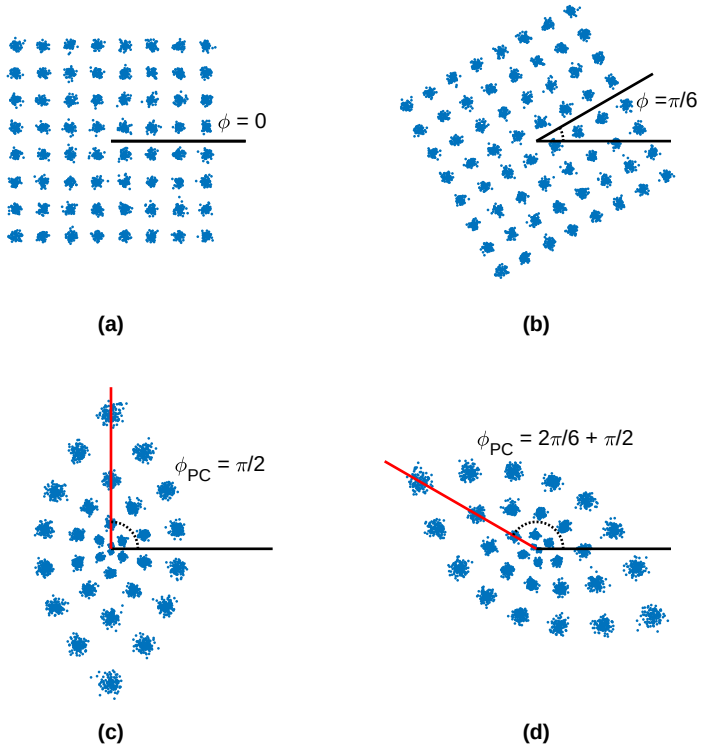

Figure 6.3.: 64QAM constellations: (a) before squaring and without phase rotation; (b) before squaring and with $\pi/6$ rad phase rotation; (c) after squaring and without phase rotation; (d) after squaring and with $\pi/6$ rad phase rotation. The principal component is marked in red. Based on [U1].

The angle of the first PC of the squared constellation is related to the rotation of the original constellation by the relation

$$\phi_{\text{PC}} = 2\phi + \frac{\pi}{2}. \tag{6.16}$$

Therefore, it is only necessary to extract the first PC of the squared constellation to have the phase estimation. A very useful numerical method, the power iteration method (PIM), can be used. The idea behind the PIM is that as the covariance matrix C maps the unit sphere to an ellipse, its longest axis corresponds to the first PC. Then, starting by a random vector $\vec{v}[0]$, and updating it by multiplication by $\boldsymbol{C}$ over and over, at some point $\vec{v}[k]$, $k \gg 0$, would have been stretched so much in the first PC direction that $\vec{v}[k]$ would have assumed its direction [148].

Applying it for the carrier phase estimation problem, it starts guessing the first PC as $\vec{v}_0 = \begin{bmatrix} 1 & 0 \end{bmatrix}^\top$. Supposing the k-th input data block with $N = m+h$ symbols of equation 6.4, the k-th matrix, $\boldsymbol{A}[k]$, of the squared input, is an $2\times N$ matrix given by

$$\boldsymbol{A}[k] = \begin{bmatrix} \Re\{x^2[k,1]\} & \Re\{x^2[k,2]\} & \cdots & \Re\{x^2[k,N]\} \\ \Im\{x^2[k,1]\} & \Im\{x^2[k,2]\} & \cdots & \Im\{x^2[k,N]\} \end{bmatrix}, \tag{6.17}$$

where $x[k,i]$ is the i-th position of the vector $\vec{x}[k]$. Then, as the input signals are supposed to be zero-mean, the covariance matrix is a 2×2 matrix updated for each incoming block as

$$\boldsymbol{C}[k] = \boldsymbol{A}[k]\boldsymbol{A}[k]^\top. \tag{6.18}$$

The PC $\vec{v}[k] = \begin{bmatrix} v[k,1] & v[k,2] \end{bmatrix}^\top$ is updated and normalized by

$$\vec{v}[k] = \boldsymbol{C}^d[k]\vec{v}[k-1], \tag{6.19}$$

$$\vec{v}[k] := \frac{\vec{v}[k]}{\sqrt{v^2[k,1] + v^2[k,2]}}. \tag{6.20}$$

where d is an integer representing the number of times in which the covariance matrix is multiplied. Finally, the estimated phase noise to be compensated at each block is given by

$$\hat{\phi}[k] = \frac{1}{2}\arctan\left\{\frac{\vec{v}[k,2]}{\vec{v}[k,1]}\right\} - \frac{\pi}{4}. \tag{6.21}$$

Figure 6.4 shows a block diagram representation of the proposed algorithm, which for the rest of this Chapter will be referred as the principal component-based phase estimation (PCPE).

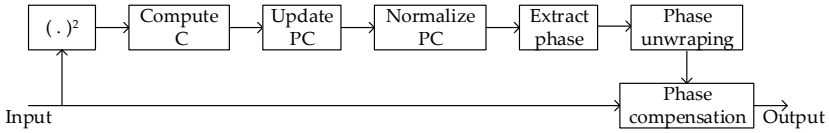

Figure 6.4.: Block diagram of carrier phase recovery method based on PCA. Based on [U1].

6.2.2 PCPE and BPS hybrid scheme

In the analysis and discussion section it will be shown that the PCPE has a considerably lower cycle slip rate (CSR) for low SNR signals, but worse performance in phase tracking when compared to BPS in low noise systems. Therefore, another possible approach is to try to join the cycle slip robustness of the PCPE with the accuracy of the BPS while still reducing its complexity. Figure 6.5 shows a block diagram for the PCPE and BPS hybrid scheme (PCPE-BPS).

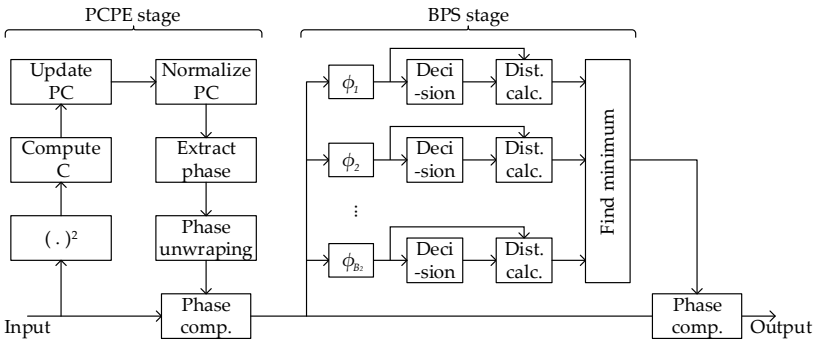

Figure 6.5.: Block diagram of carrier dual-stage hybrid phase recovery method based on coarse estimation by PCPE and fine tuning by BPS. Based on [U1].

In this scheme, the first stage performs the PCPE, while the second stage performs a BPS algorithm with test phases given by

$$\phi_b = \eta\pi\left[\frac{(2b-1)}{4B_2} - \frac{1}{4}\right], \qquad b \in \{1,2,\ldots,B_2\}, \tag{6.22}$$

where η is the phase aperture of the BPS stage, *i.e.*, the range in which fine test phase search will be performed. The phase unwrapping process is done still in the PCPE stage, therefore ensuring the same robustness to cycle slips as the PCPE algorithm alone.

6.3 Computational complexity analysis and comparison

In order to analyze and compare the computational complexity of the proposed PC-based methods and BPS, the number of operations necessary to implement the algorithms was calculated. The operations were divided in real additions (or subtractions), real multiplications (or divisions), computation of square root, accesses to a look up tables (LUTs), decisions and comparisons. The values obtained are based on specific implementations and could be different if different implementations were considered, serving for comparison purposes. The values are shown in Tab. 6.1, where $K = m+h$ is the number of symbols processed at each ASIC clock pulse. The method to compute each of these values is shown in the Appendix B.

	Algorithm			
Operation	BPS	BPS-BPS	PCPE	PCPE-BPS
Addition	5KB+2K+3	$5KB_T$+2K+3	6K+3d	$5KB_2$+8K+3d
Multiplication	6KB+4K	$6KB_T$+4K	10K+4d+4	$6KB_2$+14K+4d+4
Square root	0	0	1	1
LUT	2	B_2+2	3	3
Decision	KB	KB_T	0	KB_2
Comparison	B–1	B_T–2	0	B_2–1

Table 6.1.: Computational complexity for carrier phase estimation algorithms.

As an example, considering an ASIC operating at a clock frequency 64 times lower than the symbol rate and no overlap between consecutive blocks ($P = 64$), a single-stage BPS algorithm with a resolution of 100 test phases would require 32131 and 38656 real additions and real multiplications,

respectively. To have the same phase estimation resolution, a BPS-BPS algorithm would need each stage having 10 test phases ($B_T = 20$, $B_1 B_2 = 100$), and the number of additions and multiplications would be 6531 and 7936, respectively. The PCPE algorithm (with $d = 1$) would require 387 real additions and 648 real multiplications. This would represent a computational complexity reduction of at least 70 and 14 times in comparison to the single and dual-stage BPS implementations, respectively. Alternatively, if the PCPE-BPS method is considered, it would require 3715 real additions and 4744 real multiplications, representing roughly 40% of computational complexity reduction in comparison to BPS-BPS.

6.4 Algorithm evaluation

6.4.1 Simulation setup

In order to evaluate the performance of the proposed carrier recovery schemes, a single polarization transmission system was simulated. The modulation formats considered were square-M-QAM with $M = 4$, 16, 64 and 256. The laser LW was emulated by a random walk process [55]. The data were generated with 1 sample per symbol and fed directly into the carrier recovery scheme in order to avoid influence from other algorithms in the receiver DSP chain. The achievable information rates (AIRs) were estimated by mutual information (MI) computations between the transmitted symbols and the symbols after the carrier recovery, while the bit error ratio (BER) was computed between transmitted and decided bits. The CSR and the mean square error (MSE) of the phase estimation were computed comparing the estimated and the actual phase noise. The simulation setup is depicted in Figure 6.6.

In the next subsections the performance analysis and comparison is presented based on the CSR, MSE between the actual phase noise and the estimated, BER and AIR.

6.4.1.1 Cycle-slip rates

Due to the $\pi/2$ phase symmetry of square-M-QAM modulation formats, cycle slips may occur during the phase unwrapping process [66]. This is a highly nonlinear phenomenon that leads to a catastrophic failure if no special coding is being employed. Ideally, the maximum information rate

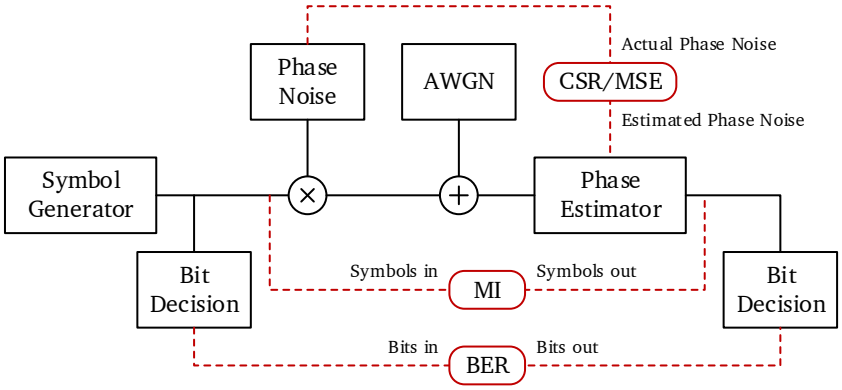

Figure 6.6.: Simulation setup. CSR: cycle slip rate; MSE: mean square error; BER: bit error rate; MI: mutual information. Based on [U1].

can only be achieved if no cycle slip occur, therefore it is very important that the CSR is as low as possible.

In our analysis, CSR is defined as the number of cycle slip occurrences per block. The first cycle slip occur at the first time the phase difference between the actual and the estimated phase noises becomes greater than $\pi/2$. Then, the following cycle slip only occurs if the difference between the actual and the estimate phase noises increase to more than π or return to be lower than $\pi/2$. Therefore, the CSR is computed as

$$\mathrm{CSR} = \sum_{k=2}^{K} \frac{\left|\tilde{\phi}[k] - \tilde{\phi}[k-1]\right|}{K-1}, \tag{6.23}$$

where K is the number of blocks simulated, and $\tilde{\phi}[k]$ is the rounded difference between the estimated phase noise and the average of the actual phase noise per block,

$$\tilde{\phi}[k] = \left\lfloor \frac{\hat{\phi}[k] - \bar{\phi}[k]}{\pi/2} \right\rceil. \tag{6.24}$$

The algorithms were analyzed under different noise conditions, sweeping both SNR and laser LW. For each simulated point, it was considered 1000 realizations of a 32 GBd ($T_S = 31.25$ ps) signal with 256 blocks of 64 symbols each (a total of 16384 symbols) and no overlap between consecutive blocks. Also, combined LWs of $\Delta\nu = 200$, 500, 1000 and 2000 MHz were considered. It was only considered the single-stage PCPE method and the BPS-BPS with 6 test phases in each stage for the quadrature phase shift keying (QPSK)

modulation format and 11 test phases in each stage for the higher-order square-M-QAM modulation formats. These values for the number of test phases were previously optimized for best performance and the results are omitted. The PCPE-BPS method was also ran but it is not being shown because it has the exact same CSR performance of the PCPE method alone. The results are presented in Figures 6.7 and 6.8.

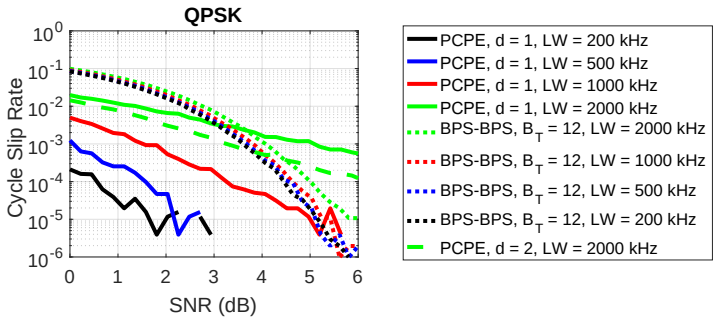

Figure 6.7.: Cycle slip rate for a 32 Gbd QPSK signal. BPS-BPS was run with 6 test phases in each stage. Based on [U1].

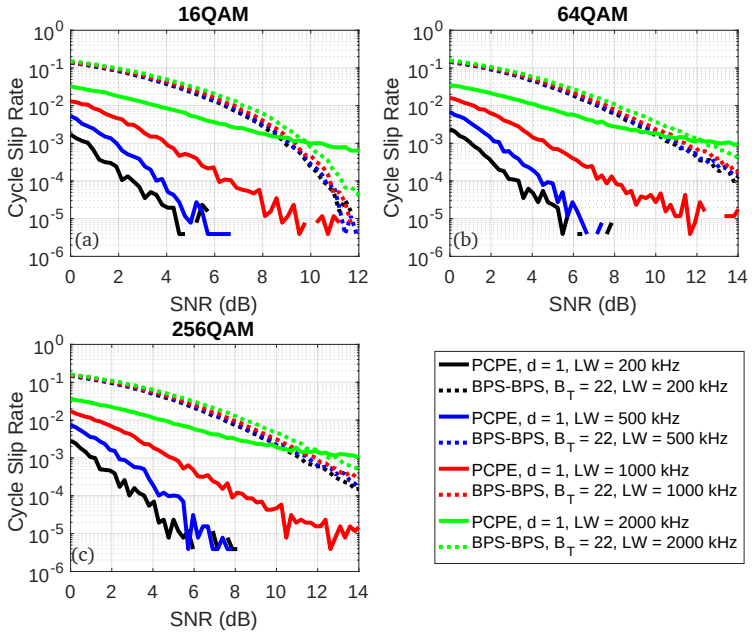

Figure 6.8.: Cycle slip rate for a 32 Gbd signal with square-M-QAM modulation format. (a) 16QAM; (b) 64QAM; (c) 256QAM. BPS-BPS was run with 11 test phases in each stage. Based on [U1].

For the QPSK modulation format at low LW values, the CSR was much smaller for the PCPE method than for the BPS-BPS. However, for high LWs (greater than 1000 MHz), the PCPE method performed better than BPS-BPS only for extremely low SNR values. It is important to note that high LWs affect more the PCPE performance than the BPS-BPS performance.

Similarly, the PCPE method shows an overall greater robustness to cycle slips than the BPS-BPS method for square-M-QAM modulation formats ($M = 16$, 64, 256). The CSR is higher for the PCPE method than for the BPS-BPS for high LW values. However, it is important to note that these values are greater than the maximum LW recommended for systems employing such modulation formats

6.4.1.2 Block size evaluation

The performance of a carrier phase estimation algorithm can be assessed by how precise the phase estimation is. It was considered as performance indicators, the BER and the MSE of phase estimation, defined by

$$\text{MSE} = \frac{1}{N}\sum_{n=1}^{N}\left\{\left(\hat{\phi}[\lceil n/m \rceil] - \phi[n]\right)^2\right\}, \tag{6.25}$$

where N is the number of symbols per realization of simulation.

The SNR was swept from 0 to 50 dB and the block size from 8 to 128 symbols. 100 realizations of 16384 symbols were simulated for each point of a 32 GBd 256QAM signal without residual frequency offset and combined laser LW of 200 kHz. The BPS-BPS method was considered with 11 test phases in each stage, while the PCPE method and the PCPE-BPS was considered with 11 test phases in the BPS stage and aperture $\eta = 1/11$ to match the same test phase range as the BPS-BPS. No overlap between consecutive blocks was considered and the cycle slip occurrences were not artificially compensated. The MSE and BER colorplots are shown in Figures 6.9 and 6.10. Due to the number of symbols and realizations, the counted BER floor was considered as 10^{-5}.

As expected, for low SNR values, the PCPE method outperforms the BPS-BPS method, however, for high SNR values, the PCPE method hits a BER and MSE floor higher than the other methods. In this case, the PCPE-BPS scheme combines the best of both algorithms, showing a bigger region with low

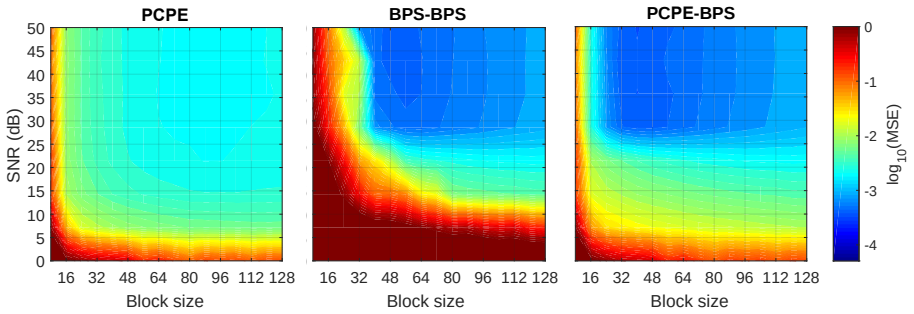

Figure 6.9.: Mean square error of phase estimation for a 32 GBd 256QAM signal without residual frequency offset and combined laser linewidth of 200 kHz. MSE was measured in rad^2.

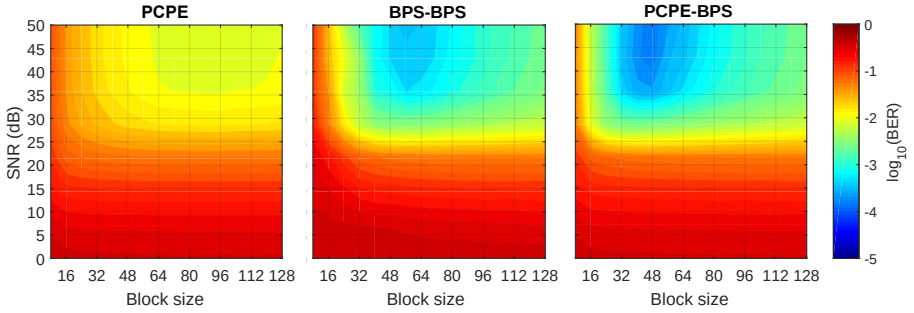

Figure 6.10.: Bit error ratio for a 32 GBd 256QAM signal without residual frequency offset and combined laser linewidth of 200 kHz.

MSE, with required SNR of 25 dB for BER of 10^{-2} with block sizes around 60 symbols per block. The block size is a parameter that depends on the internal clock in which an ASIC executes operations, therefore it is very important for an algorithm showing good performance for long block lengths. In this case, for a 32 GBd signal, the optimum performance would be achieved with an ASIC operating at around 533 MHz.

Then, the effect of a residual frequency offset that would remain being uncompensated was from a frequency offset compensation process prior to the carrier phase estimation [15, 93] was simulated. Figures 6.11 and 6.12 show MSE and BER results, respectively, for a 32 GBd signal with 200 kHz and 5 MHz of residual frequency offset.

Here, the block size plays a quite important role, as the overall performance of all the methods are impaired for long block lengths. With a 5 MHz residual frequency offset, the optimum block size is reduced to 40 symbols per block. The PCPE method is highly impaired in comparison to the case with no

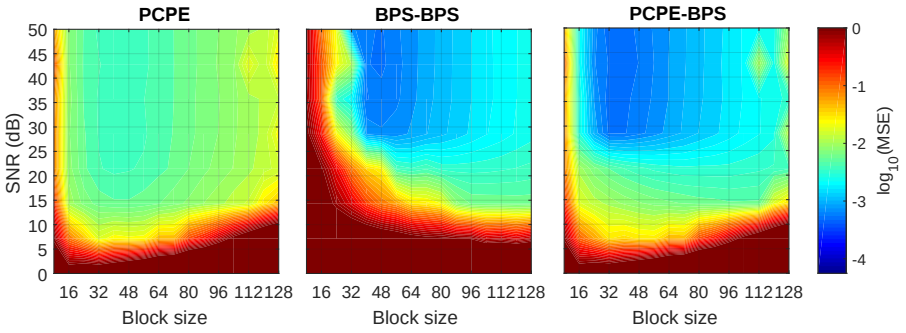

Figure 6.11.: Mean square error for a 32 GBd 256QAM signal with 5 MHz residual frequency offset and combined laser linewidth of 200 kHz. MSE was measured in rad^2

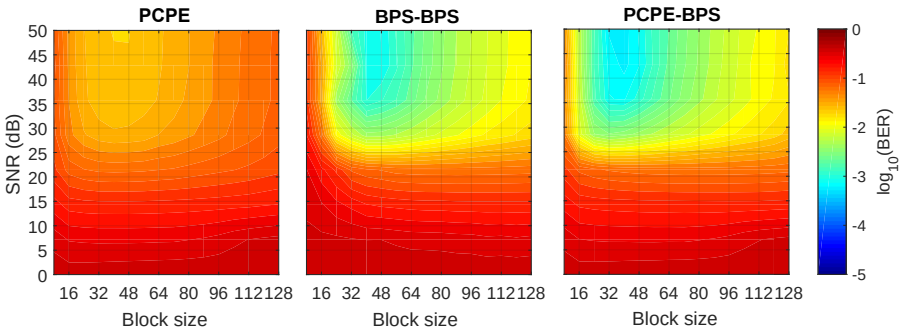

Figure 6.12.: Bit error ratio for a 32 GBd 256QAM signal with 5 MHz residual frequency offset and combined laser linewidth of 200 kHz.

residual frequency offset. Again, the PCPE-BPS method performs better than the other two methods, in regards to both MSE and BER around the optimum block length.

6.4.1.3 Mutual information penalties

For high-order modulation formats, forward error correction (FEC) are generally applied to achieve the maximum capacity of the channel. Therefore, the pre-FEC BER is not the most indicated metric for system evaluation [149, 150]. It is preferable to use a higher-order modulation format and FEC codes instead of operating with a lower-order modulation format in the saturation, *i.e.* in an SNR in which the MI is the maximum for that modulation format. In this section, the analyses of AIRs assume an optimal encoder and decoder are employed by calculating the MI between the transmitted and received

symbols. The MI computation is a Monte Carlo approach considering two different channels, one in which there is only AWGN and the other where the main source of noise is the residual phase noise. The MI computation is better explained in Appendix C.

The algorithms in a system with 200 kHz of LW and no residual frequency offset between the transmitter and receiver lasers were analyzed. Cycle slips were ideally compensated to simplify the MI computation. The SNR values were swept from 0 to 10 dB for QPSK, from 0 to 18 dB for 16QAM, from 0 to 30 dB for 64QAM, and from 0 to 45 dB for 256QAM. Each MI value is an average of the MI values computed from 10000 realizations with 1024 blocks of 64 symbols each. It was considered the PCPE, the BPS-BPS with 11 test phases in each stage, and the PCPE-BPS methods with 11 test phases and aperture of $\eta = 1/11$. It was considered the ideal case when there was no phase noise and also, as a comparison upper-bound, the best hypothetical performance that would be achieved in a parallelized receiver was considered, *i.e.* the ideal block-wise phase estimator (presented in Section 6.1.1). The results are shown in Figure 6.13. The MI versus the SNR (left side of Figure 6.13) and the penalty in MI in comparison with the ideal case without phase noise (right side of Figure 6.13) were plotted.

For all the modulation formats, the BPS-BPS method have a higher MI penalty than the other algorithms. This is because the BPS algorithm has a harder time distinguishing the test phases in the presence of high AWGN. Certain combinations of symbols can lead to ambiguities, and therefore estimation errors. Instead, as the PCPE method is not based on testing phases, it does not suffer from the same problem, performing better at low SNR values. The penalty introduced by the testing of phases can also be seen to a less extent in the performance of the PCPE-BPS, where the BPS stage is only the fine estimation step. For this method, the fine tuning is actually impairing the system, compared to the PCPE alone and therefore the performance is worse with this step. None of the methods, however achieve the same performance of the ideal block-wise phase estimator (BWPE) at low SNRs, showing that there are still some AWGN related impairments on the phase noise estimation process.

Analyzing each modulation format presented, the PCPE performs better than the other algorithms for all SNR values when operating with QPSK, while for the other modulation formats it performs better for SNR values lower than 11.5, 16 and 22 dB for 16, 64 and 256QAM, respectively. For values of

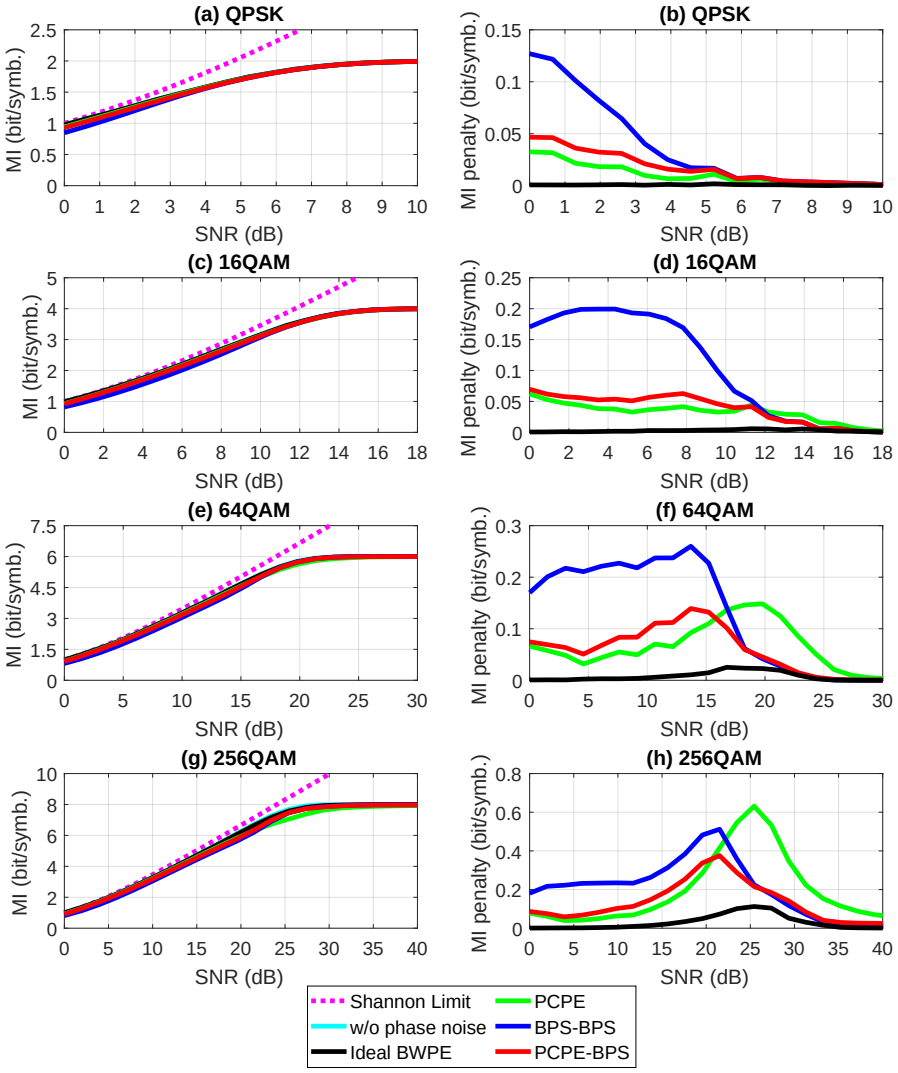

Figure 6.13.: Left side: Mutual information versus SNR for different modulation formats. Right side: Mutual information penalty from the no phase noise case for different modulation formats. (a,b) QPSK; (c,d) 16QAM (e,f) 64QAM (g,h) 256QAM. Based on [U1].

SNR higher than these thresholds, the PCPE-BPS have the best or the same performance as the BPS-BPS. As the computational complexity of PCPE-BPS is lower than the BPS-BPS, it makes the first a better choice.

6.4.1.4 Laser phase noise tolerance

To evaluate the LW tolerance the minimum SNR required to achieve an error-free transmission considering 20% of FEC overhead was computed. The target AIR, I_{target}, for this case was computed using the following equation:

$$I_{target} = \frac{\log_2(M)}{1 + \text{OH}}, \tag{6.26}$$

where M is the modulation order and OH is the FEC overhead. This leads to a target AIR of 1.6667, 3.333, 5 and 6.667 bits/symbol for QPSK, 16QAM, 64QAM and 256QAM, respectively.

The LW values were swept from 0 to 4000 kHz for QPSK, from 0 to 2000 kHz for 16QAM, from 0 to 1000 kHz for 64QAM and from 0 to 600 kHz for 256QAM. Cycle slips were ideally compensated. Each SNR required value were generated through interpolation of three points of SNR with AIR values computed from 1000 realizations with 1024 blocks of 64 symbols each. The symbol rate was 32 GBd. The results are shown in Figure 6.14.

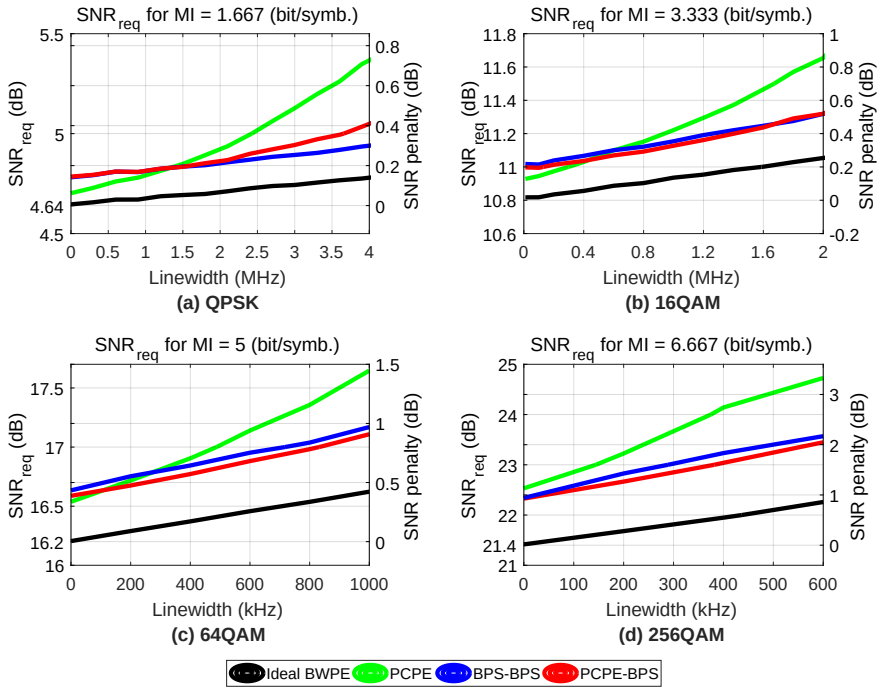

Figure 6.14.: Minimum SNR required for error-free transmission considering 20% FEC overhead. (a) QPSK; (b) 16QAM; (c) 64QAM; (d) 256QAM.

For the QPSK modulation format (Figure 6.14a), the PCPE has the lowest penalty for LW values under 1.3 MHz. For very high LW values, the performance of PCPE degrades faster than the BPS-BPS method, and the PCPE-BPS method suffers from the degraded performance of the PCPE stage. For intermediate LW values (between 1.3 and 2 MHz), both PCPE-BPS and BPS-BPS have similar performance.

For the 16QAM and 64QAM modulation formats (Figure 6.14b-c), the PCPE has the lowest penalty for LW values under 550 and 100 kHz, respectively. For higher LW values, both PCPE-BPS and BPS-BPS have similar performance, with PCPE-BPS marginally outperforming BPS-BPS.

For the 256QAM modulation format (Figure 6.14d), the PCPE has higher SNR required than the other methods for all the LW values considered. Also, the SNR penalty for all the methods are at least 1 dB worse than the hypothetical ideal BWPE.

In particular cases of systems using soft-decision FEC, moving to a higher modulation format order and considering more overhead helps to achieve more information rate for the same SNR [151, 152]. Then, the LW tolerance for 256QAM considering 50% of FEC overhead ($I_{\text{target}} = 5.33$ bit/symbol) was also considered. The results are shown in Figure 6.15.

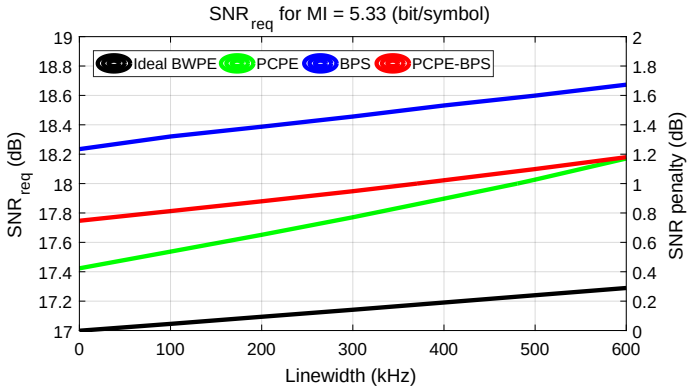

Figure 6.15.: Minimum SNR required for error-free transmission considering 50% FEC overhead for 256QAM modulation format.

In this case, the PCPE method outperforms the other methods for LW values under 600 kHz. Also, a 1 dB SNR penalty from the no phase noise case is achieved for a LW of 500 kHz.

6.4.2 Experimental validation

An 84-GSa/s AWG generated two electrical lines of a 28 GBd signal with a 0.4 roll-off root raised cosine pulse shape. The electrical signals were amplified in electrical drivers and fed to a $LiNbO_3$ external optical modulator together with a 100-kHz LW laser. The signal was PM emulated and received in a coherent receiver with a 100-kHz LW local oscillator laser. The received electrical signals were sampled by a 80 GS/s oscilloscope. The DSP was performed offline and consisted of a radially directed equalizer (RDE) algorithm with 12 taps for polarization demultiplexing and inter-symbol interference (ISI) compensation, a frequency domain 4-th power frequency offset estimator and the carrier phase estimator being evaluated. The modulation formats considered were 16QAM and 64QAM. The results are presented in Figures 6.16 and 6.17.

An 84-GSa/s AWG generated two electrical lines of a 28 GBd signal with a 0.4 roll-off root raised cosine pulse shape. The electrical signals were amplified in electrical drivers and fed to a $LiNbO_3$ external optical modulator together with a 100-kHz LW laser. The signal was PM emulated and combined by the noise generated in a erbium-doped fiber amplifier (EDFA). Then, it was received by a coherent receiver with another 100-kHz LW laser as local oscillator. The received electrical signals were sampled by a 80 GSa/s oscilloscope. The DSP was performed offline and consisted of a RDE algorithm with 12 taps for polarization demultiplexing and ISI compensation, a frequency domain 4^{th}-power frequency offset estimator and the CPR being evaluated. The modulation formats considered were 16 and 64QAM. Cycle slips were not previously compensated. The OSNR required to achieve an error-free transmission considering 20% FEC overhead was computed.

The PCPE algorithm performed better than the other two algorithms, with 0.1 and 0.15 dB gain in comparison to PCPE-BPS and BPS, respectively, for 16QAM. For 64QAM, the gains were 0.2 and 0.43 dB, respectively. These gains were higher than expected and are possibly due to IQ imbalances in the generation of the modulation formats. The PCPE is agnostic to this, therefore, performing better. For very low OSNR values, the PCPE and PCPE-BPS algorithms performed better due to lower cycle slip rates, as expected.

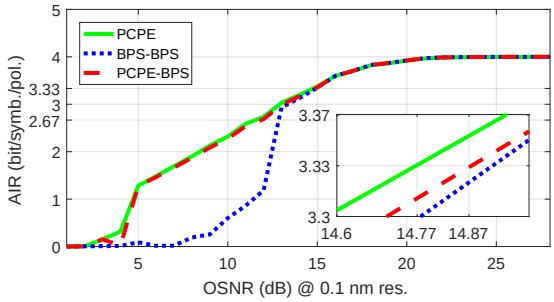

Figure 6.16.: Experimental results for 28 GBd PM-16QAM with combined LW of 200 kHz. Inset: Zoom around the 20% FEC overhead ($I_{\text{target}} = 3.333$ bit/symbol/polarization). Based on [U1].

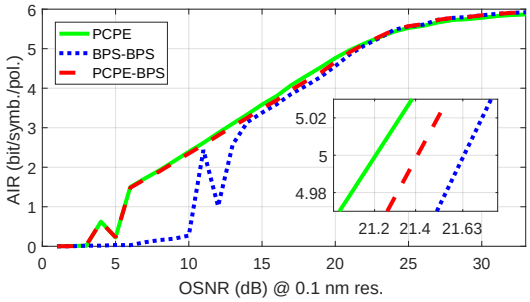

Figure 6.17.: Experimental results for 28 GBd PM-64QAM with combined LW of 200 kHz. Inset: Zoom around the 20% FEC overhead ($I_{\text{target}} = 5$ bit/symbol/polarization). Based on [U1].

6.5 Summary

A hardware-efficient non-data-aided carrier phase estimation method that is independent of modulation format and it is based on extracting the principal component of the received signal was proposed. This method is intended for implementation in ASICs operating in clocks several times slower than the symbol rate. PCPE have good performance for low SNR signals with much lower cycle slip rate and computational complexity in comparison to BPS. For intermediate to high SNR values, PCPE presents high residual phase noise. Consequently, a two-stage hybrid scheme was also presented and demonstrated. This scheme, comprised of a PCPE stage for coarse estimation and a BPS stage for fine tuning, combines the cycle slip robustness of the first with the accuracy of the former. This hybrid scheme shall be the best choice for flexibility, using only the PCPE stage for low SNR or the full scheme otherwise.

7

Conclusions and Outlook

THIS THESIS presented new contributions to the state-of-the-art digital signal processing (DSP) techniques for mitigation and compensation of transmission impairments in coherent optical systems. A summary of the results and an outlook on possible future research directions are provided in the following sections for each part of the thesis.

7.1 Clock recovery for PDM and MDM systems

In Chapter 3, the performance and feasibility of non-data-aided (NDA) fully digital clock recovery for coherent optical receivers was analyzed through simulations. Both current generation polarization division multiplexing (PDM) and possible next generation mode division multiplexing (MDM) systems were considered. Tolerances to transceiver time skew, residual chromatic dispersion, different pulse shapes, polarization rotation, polarization mode dispersion (PMD) and modal dispersion (MD) with weak to strong coupling were analyzed.

Specifically for PDM systems, more than 10000-km of a modern standard single-mode fiber (SMF) with differential group delay (DGD) value around 0.1 ps/$\sqrt{\text{km}}$ would be necessary in order to impair the clock recovery performance considerably, assuming polarization tracking methods are performed in the receiver DSP [128]. For systems using spectrally efficient pulse shapes such as raised cosine (RC) and root raised cosine (RRC), the clock tone amplitude (CTA) would vanish for very low roll-off factors, needing a modification on the algorithm [108] in order to have a detectable CTA. Using this modification, it was shown that RC has a better performance compared to RRC. On the other hand, it is shown that the mode coupling and delay spread have an impact much worse in the performance of timing synchronization

for MDM systems. Although strong coupling could be induced to reduce both the coupled group delay (GD) spread [63] and nonlinearities [113], the clock tone would completely vanishes even for very low GD values, making timing synchronization impractical. In order to have a detectable CTA, higher than 50% of its maximum value, the few-mode fiber (FMF) would need to have an unrealistic uncoupled GD of less than 0.03 ps/km. Increasing the transmission distance reduce the GD tolerance of MDM systems even more under strong coupling regime.

Possible solutions for the clock recovery in MDM systems could be to use pilot tones for timing synchronization. Another possible solution would be the design of FMFs with lower coupling between modes rather than trying to increase the coupling on purpose as proposed in [63]. The operation in weak coupling regime would benefit from the implementation of sparsity managed multiple-input multiple-output (MIMO) equalizers [88, 89]. Other transmission impairments such as the equalization-enhanced phase noise (EEPN) [153], a random timing jitter caused by the random phase noise in a coherent receiver, and fiber nonlinearities [154] could impact the clock recovery performance, and this impact could scale up with the dimensionality of the receiver. Future works should contemplate these areas.

7.2 Transceiver optimization

7.2.1 Transmitter IQ imbalances and time skews calibration

In Chapter 4, a new and simple method has been suggested to optimize the common transmitter's front-end imperfections such as suboptimal biasing, time skews, and amplitude, and phase imbalances between in-phase (I) and quadrature (Q) components. This was achieved by using a cooperative coevolutionary genetic algorithm (CC-GA). This method is carried out on the transmitter side, so that the transmitter can be optimized automatically for the best operation regardless of the coherent receiver and thus prevent the complexity of the stressed DSP at the receiver from increasing. The isolation from the receiver impairments was achieved by the use of a photodetector, in which the cost-functions for the CC-GA were extracted.

The performance of the method was first analyzed through numerical simulations, followed by an experimental demonstration of the method behavior. In the numerical analysis, all the parameters converged to their final values

in 18 generations of the CC-GA, being able to achieve the most accurate results when using non-return-to-zero (NRZ)-PM-16QAM as the reference signal.

Artificial intelligence methods, based on natural evolution are a prominent research area and could be used for optimization of many optical communication devices. The results presented in this Chapter demonstrated the potential of the CC-GA as a rapid optimization method for fine-tuning and mitigating transmitter impairments. Other optimization methods should be studied in the future in order to find which can converge to the most precise and fastest result. Dynamic and on-line optimization should be also researched, as well as the applications for calibration of other transmitter impairments such as the frequency response of the transmitter electrical drivers, digital-to-analog converters (DACs) and cables.

7.2.2 Receiver time skews and CD estimation

In Chapter 5, a low-complexity estimator was proposed for estimation of in-phase/quadrature (IQ) time skews in DSP-based coherent optical receivers, jointly with chromatic dispersion (CD) estimation. The proposed method was based in a common sweeping CD estimation algorithm and relied in the fact that if the receiver signal has time skew between the quadrature electrical signals, the signal after CD compensation will still have a component with double the CD.

The method was experimentally demonstrated for a 32 GBd polarization multiplexed (PM)-16QAM transmission. Both IQ time skew and CD estimation have demonstrated good accuracy in the results. The advantage of the proposed method is that it has a small additional complexity in comparison to no other methods since it is performed before MIMO equalization and timing recovery, avoiding the complexity increase necessary for skew tolerant MIMO equalizers [33, 34]. Also, the proposed method was able to deliver CD and IQ time skew free signals to a subsequent timing recovery block, benefiting its performance.

It is important to observe that the method proposed in this Chapter is less complex than the method presented in Chapter 4, but also less capable. Other optimization methods such as the ones based on evolutionary computation that would take care of other impairments such as IQ phase imbalances, and the frequency response of the drivers should be analyzed in future works.

7.3 Carrier recovery

In Chapter 6, a hardware-efficient, modulation-order independent, NDA and feedforward carrier phase recovery (CPR) algorithm, based on extracting the principal component (PC) of the received constellation after squaring process. The proposed method, known as principal component-based phase estimation (PCPE), was designed for block-wise parallel processing in order to contemplate a possible hardware implementation, taking into consideration application specific integrated circuit (ASIC) operation in much lower clock frequencies than the signal symbol rate. Simulation analyzes and an experiment were carried out in order to compare the proposed method with the traditional blind phase search (BPS) algorithm [31]. PCPE showed good performance for low signal-to-noise ratio (SNR) signals with much lower cycle slip rate and computational complexity in comparison to BPS. For intermediate to high SNR values, PCPE presents high residual phase noise. Consequently, a two-stage hybrid scheme was also presented and demonstrated. This scheme is comprised of a PCPE stage for coarse estimation and a BPS stage for fine tuning, combining the cycle slip robustness of the first with the accuracy of the former. This hybrid scheme shall be the best choice for flexibility, using only the PCPE stage when the received signal has low SNR or the full scheme otherwise.

The proposed method was analysed using mean square error (MSE), bit error ratio (BER), cycle slip rate (CSR) and mutual information (MI) as performance indicators under various scenarios and square QAM modulation formats up to 256QAM. The CSR of PCPE was much lower than BPS at low SNR signals when the combined laser linewidth (LW) values were reasonably low (< 500 kHz for 32 GBd signals). The PCPE method showed a higher MSE at high SNR signals in comparison with the other methods, while a lower MSE for low SNR signals. This behavior was more evident in the MI analysis in which PCPE alone was the best algorithm for low SNR signals, and the PCPE and BPS hybrid scheme (PCPE-BPS) was better for the high SNR signals.

The quest for capacity achieving transmission has been recently leading to the analysis and implementation of both non-uniformly distributed signals using probabilistic shaping [155] and non-trivial modulation formats using geometric shaping [156]. These approaches need novel CPR algorithms that are suitable for these applications [157]. Considering that principal

component analysis (PCA)-based techniques have better results than traditional methods in low SNR and also has low complexity, it is expected other applications, such as the ones aforementioned, could benefit from it and should be analysed in future works.

A

Analytical calculation of principal components

A.1 Calculation of PCs from QAM constellation

Let's consider $\vec{Q} = \begin{bmatrix} Q_i & Q_q \end{bmatrix}^\top$ as a random vector describing a square-M-quadrature amplitude modulation (QAM) constellation, where Q_i and Q_q represent the in-phase and quadrature components of the constellation, respectively, and are zero-mean independent, identically and uniformly distributed random variables with variance $\sigma_i^2 = \sigma_q^2 = \sigma^2$ and fourth moment $\mathbb{E}\{Q_i^4\} = \mathbb{E}\{Q_q^4\} = \mu_4$. The phase-rotated constellation, $\vec{Q}_\theta$, is given by

$$\vec{Q}_\theta = \begin{bmatrix} Q_{\theta_i} \\ Q_{\theta_q} \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} Q_i \\ Q_q \end{bmatrix} = \begin{bmatrix} \cos(\theta)Q_i - \sin(\theta)Q_q \\ \sin(\theta)Q_i + \cos(\theta)Q_q \end{bmatrix}. \quad \text{(A.1)}$$

The covariance matrix of $\mathbf{Q}_\theta$ is then

$$\boldsymbol{C}_{\vec{Q}_\theta} = \mathbb{E}\{\vec{Q}_\theta \vec{Q}_\theta^\top\} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}, \quad \text{(A.2)}$$

where

$$\begin{aligned} c_{11} &= \mathbb{E}\{\cos^2(\theta)Q_i^2 + \sin^2(\theta)Q_q^2 - 2\sin(\theta)\cos(\theta)Q_iQ_q\} \\ &= \cos^2(\theta)\sigma^2 + \sin^2(\theta)\sigma^2 = \sigma^2, \end{aligned} \quad \text{(A.3)}$$

$$
\begin{aligned}
c_{21} = c_{12} &= \mathbb{E}\left\{\sin(\theta)\cos(\theta)\left(Q_i^2 - Q_q^2\right) + Q_i Q_q\left(\cos^2(\theta) - \sin^2(\theta)\right)\right\} \\
&= \sin(\theta)\cos(\theta)(\sigma^2 - \sigma^2) = 0,
\end{aligned} \tag{A.4}
$$

$$
\begin{aligned}
c_{22} &= \mathbb{E}\left\{\sin^2(\theta)Q_i^2 + \cos^2(\theta)Q_q^2 + 2\sin(\theta)\cos(\theta)Q_i Q_q\right\} \\
&= \sin^2(\theta)\sigma^2 + \cos^2(\theta)\sigma^2 = \sigma^2.
\end{aligned} \tag{A.5}
$$

Then, $\boldsymbol{C}_{\vec{Q}_\theta}$ is invariant to the phase rotation, and consequently, the eigenvectors of $\boldsymbol{C}_{\vec{Q}_\theta}$ are also invariant to the phase rotation.

A.2 Calculation of PCs from QAM constellation after squaring

Let's now consider the phase rotated constellation squared,

$$
\begin{aligned}
\vec{S}_\theta &= \begin{bmatrix} S_i \\ S_q \end{bmatrix} = \begin{bmatrix} Q_{\theta i}^2 - Q_{\theta q}^2 \\ 2Q_{\theta i}Q_{\theta q} \end{bmatrix} \\
&= \begin{bmatrix} \left(\cos^2(\theta) - \sin^2(\theta)\right)\left(Q_i^2 - Q_q^2\right) - 4\cos(\theta)\sin(\theta)Q_i Q_q \\ 2\left[\cos(\theta)\sin(\theta)\left(Q_i^2 - Q_q^2\right) + \left(\cos^2(\theta) - \sin^2(\theta)\right)Q_i Q_q\right] \end{bmatrix},
\end{aligned} \tag{A.6}
$$

replacing the trigonometric relations,

$$
\begin{aligned}
\sin(2\theta) &= 2\sin(\theta)\cos(\theta), \\
\cos(2\theta) &= \cos^2(\theta) - \sin^2(\theta),
\end{aligned}
$$

in Eq. A.6, then,

$$
\vec{S}_\theta = \begin{bmatrix} \cos(2\theta)\left(Q_i^2 - Q_q^2\right) - 2\sin(2\theta)Q_i Q_q \\ \sin(2\theta)\left(Q_i^2 - Q_q^2\right) + 2\cos(2\theta)Q_i Q_q \end{bmatrix}. \tag{A.7}
$$

The covariance matrix of $\vec{S}_\theta$ is, then,

$$
\boldsymbol{C}_{\vec{S}_\theta} = \mathbb{E}\left\{\vec{S}_\theta \vec{S}_\theta^\top\right\} = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}, \tag{A.8}
$$

where

$$\begin{aligned} d_{11} &= \mathbb{E}\left\{\left[\cos(2\theta)\left(Q_i^2 - Q_q^2\right) - 2\sin(2\theta)Q_iQ_q\right]^2\right\} \\ &= \mathbb{E}\Big\{ \cos^2(2\theta)\left(Q_i^2 - Q_q^2\right)^2 + 4\sin^2(2\theta)Q_i^2Q_q^2 \\ &\qquad - 4\cos(2\theta)\sin(2\theta)Q_iQ_q\left(Q_i^2 - Q_q^2\right)\Big\} \\ &= \cos^2(2\theta)\left(2\mu_4 - 2\sigma^4\right) + 4\sin^2(2\theta)\sigma^4, \end{aligned} \tag{A.9}$$

$$\begin{aligned} d_{21} = d_{12} &= \mathbb{E}\Big\{ \left[\cos(2\theta)\left(Q_i^2 - Q_q^2\right) - 2\sin(2\theta)Q_iQ_q\right] \\ &\qquad \left[\sin(2\theta)\left(Q_i^2 - Q_q^2\right) + 2\cos(2\theta)Q_iQ_q\right]\Big\} \\ &= \mathbb{E}\Big\{ \cos(2\theta)\sin(2\theta)\left[\left(Q_i^2 - Q_q^2\right)^2 - 4Q_i^2Q_q^2\right] \\ &\qquad + 2\left(\cos^2(2\theta) - \sin^2(2\theta)\right)Q_iQ_q\left(Q_i^2 - Q_q^2\right)\Big\} \\ &= \cos(2\theta)\sin(2\theta)\left(2\mu_4 - 6\sigma^4\right), \end{aligned} \tag{A.10}$$

$$\begin{aligned} d_{22} &= \mathbb{E}\left\{\left[\sin(2\theta)\left(Q_i^2 - Q_q^2\right) + 2\cos(2\theta)Q_iQ_q\right]^2\right\} \\ &= \mathbb{E}\Big\{ \sin^2(2\theta)\left(Q_i^2 - Q_q^2\right)^2 + 4\cos^2(2\theta)Q_i^2Q_q^2 \\ &\qquad + 4\cos(2\theta)\sin(2\theta)Q_iQ_q\left(Q_i^2 - Q_q^2\right)\Big\} \\ &= \sin^2(2\theta)\left(2\mu_4 - 2\sigma^4\right) + 4\cos^2(2\theta)\sigma^4. \end{aligned} \tag{A.11}$$

The eigenvalues, λ, of $\boldsymbol{C}_{\vec{S}_\theta}$ can be computed for

$$\det\left(\boldsymbol{C}_{\vec{S}_\theta} - \lambda \boldsymbol{I}_{2\times 2}\right) = 0. \tag{A.12}$$

Using the following change of parameters

$$\begin{aligned} s_1 &= 2\mu_4 - 2\sigma^4, \\ s_2 &= 4\sigma^4, \\ s_3 &= 2\mu_4 - 6\sigma^4, \end{aligned} \tag{A.13}$$

then,

$$
\begin{aligned}
0 &= \det\left(\boldsymbol{C}_{\vec{S}_\theta} - \lambda \boldsymbol{I}_{2\times 2}\right) \\
&= \left(s_1 \cos^2(2\theta) + s_2 \sin^2(2\theta) - \lambda\right) \\
&\quad \left(s_1 \sin^2(2\theta) + s_2 \cos^2(2\theta) - \lambda\right) \\
&\quad - s_3^2 \cos^2(2\theta) \sin^2(2\theta) \\
&= \lambda^2 - \lambda(s_1 + s_2) \\
&\quad + \left(s_1^2 + s_2^2 - s_3^2\right) \sin^2(2\theta) \cos^2(2\theta) \\
&\quad + s_1 s_2 \left(\sin^4(2\theta) + \cos^4(2\theta)\right) \\
&= \lambda^2 - \lambda(s_1 + s_2) + s_1 s_2 + \\
&\quad \left(s_1^2 + s_2^2 - s_3^2 + 2 s_1 s_2\right) \sin^2(2\theta) \cos^2(2\theta).
\end{aligned}
\tag{A.14}
$$

Replacing Eq. A.13 in Eq. A.14,

$$
\lambda^2 - \lambda\left(2\mu_4 + 2\sigma^2\right) + 8\mu_4\sigma^4 - 8\sigma^8 = 0, \tag{A.15}
$$

which has roots given by

$$
\lambda_1 = 4\sigma^4, \tag{A.16}
$$

$$
\lambda_2 = 2\mu_4 - 2\sigma^4, \tag{A.17}
$$

where $\lambda_1 > \lambda_2$. The eigenvector, $\vec{v}_{\lambda_1} = \begin{bmatrix} v_1 & v_2 \end{bmatrix}^\top$, associated with the eigenvalue λ_1, for the matrix $\boldsymbol{C}_{\vec{S}_\theta}$ can be calculated by

$$
\begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} - \lambda_1 \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{A.18}
$$

$$
(d_{11} - \lambda_1) v_1 + d_{12} v_2 = 0 \tag{A.19}
$$

$$
v_1 = \frac{d_{12}}{d_{11} - \lambda_1} v_2 \tag{A.20}
$$

Replacing Eqs. A.9 and A.16 in Eq. A.20, then

$$\begin{aligned} v_1 &= \frac{-\cos(2\theta)\sin(2\theta)\left(2\mu_4 - 6\sigma^4\right)}{\cos^2(2\theta)\left(2\mu_4 - 2\sigma^4\right) + \left(\sin^2(2\theta) - 1\right)4\sigma^4} v_2 \\ &= \frac{-\cos(2\theta)\sin(2\theta)\left(2\mu_4 - 6\sigma^4\right)}{\cos^2(2\theta)\left(2\mu_4 - 6\sigma^4\right)} v_2 \\ &= -\frac{\sin(2\theta)}{\cos(2\theta)} v_2, \end{aligned} \tag{A.21}$$

then the eigenvector $\vec{v}_{\lambda_1}$ is

$$\vec{v}_{\lambda_1} = \begin{bmatrix} -\frac{\sin(2\theta)}{\cos(2\theta)} \\ 1 \end{bmatrix} v_2, \qquad \forall v_2 \in \mathbb{R}. \tag{A.22}$$

The angle of the eigenvector $\vec{v}_{\lambda_1}$ is

$$\angle \vec{v}_{\lambda_1} = \arctan\left(-\frac{\cos(2\theta)}{\sin(2\theta)}\right) = 2\theta + \frac{\pi}{2}. \tag{A.23}$$

B

Computational complexity calculation

In this Appendix, the computational complexity calculation used in Section 6.3 is described. For these calculations real subtractions and real divisions are considered to have the same complexity as real additions and real multiplications, respectively.

B.1 Principal component based phase estimator (PCPE)

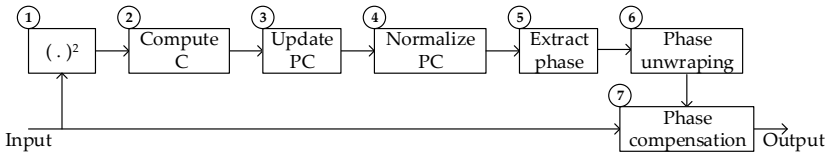

Figure B.1.: Block diagram of PCPE with steps indicated with numbers.

1. For the squaring computation, it is needed K square operations. If each input has the form $x = a + jb$, then $z = x^2 = a^2 - b^2 + j2ab$. Since real and imaginary parts are stored separately, and the multiplication by 2 is just a bit-wise shift, therefore, it is needed 1 real addition and 3 real multiplications to compute each square operation. Then, it is needed K real additions and $3K$ real multiplications for this step.
2. The correlation matrix $\boldsymbol{C}$ is the result of the multiplication of a $2 \times K$ matrix $\boldsymbol{A}$ by its conjugate. If the matrix $\boldsymbol{A}$ has the form of Eq. 6.17, then $\boldsymbol{C}$ has the form

$$\boldsymbol{C}(k) = \begin{bmatrix} \sum(\mathbb{R}(z))^2 & \sum \mathbb{R}(z)\mathbb{I}(z) \\ \sum \mathbb{R}(z)\mathbb{I}(z) & \sum(\mathbb{I}(z))^2, \end{bmatrix} \tag{B.1}$$

then, it is necessary $3K$ real multiplications and $3K-3$ real additions. If $d-1$ is the number of times in which the matrix $\boldsymbol{C}$ is multiplied by itself, then, it is needed more $3d-3$ real additions and $4d-4$ real multiplications.

3. The update of the principal component is a multiplication of a 2×2 matrix of real values by a 2×1 matrix of real values, yielding 4 real multiplications and 2 real additions.
4. The normalization of the principal component is to divide the principal component by its norm. If the principal component has the form $v = a + jb$, then

$$\begin{aligned} v_{norm} &= v/|v|, \\ v_{norm} &= \frac{a+jb}{\sqrt{a^2-b^2}}. \end{aligned} \tag{B.2}$$

   Therefore, 4 real multiplications, 1 real addition and 1 square root computation are needed for this step.
5. For the phase extraction, it is considered that 1 access to a look up table (LUT) is enough.
6. The phase unwrapping process can be realized with 3 real additions and 2 LUT accesses.
7. The phase compensation process is done by phase rotations, where each is a complex multiplication comprised of 4 real multiplications and 2 real additions. For the K symbols, then it is needed $4K$ real multiplications and $2K$ real additions.

The complexity of the whole PCPE algorithm can be expressed as $6K+3d$ real additions, $10K+4d+4$ real multiplications, 1 square root computation and 3 accesses to a LUT.

B.2 Single-stage blind phase search (BPS)

1. Each value of the input block is rotated by B test phases. Each rotation is a complex multiplication, yielding 2 additions and 4 multiplications. Therefore, it is needed in this step, $2KB$ real additions and $4KB$ real multiplications.
2. Each point is decided and the decisions are assigned to them. Therefore, KB decisions are needed in this step.
3. The distance is computed for each point. If the input rotated by the test phase has the form $a+jb$ and the decided signal has the form

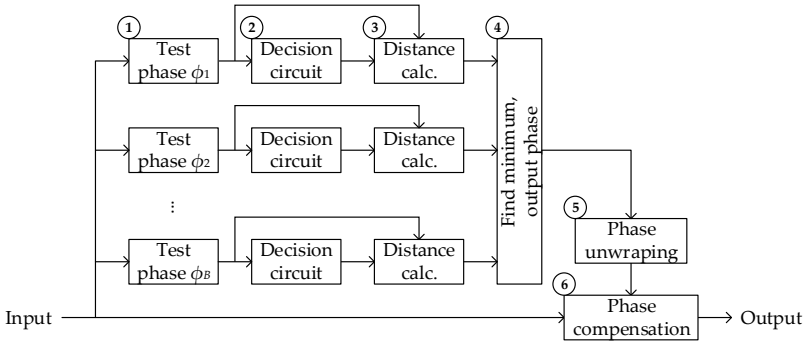

Figure B.2.: Block diagram of BPS with steps indicated with numbers.

$\hat{a}+j\hat{b}$, then the squared distance will have the form $(a-\hat{a})^2+(b-\hat{b})^2$, therefore $3KB$ real additions and $2KB$ real multiplications are needed in this step.

4. $B-1$ comparisons are needed to find the test phase associated with the minimum distance.
5. The same as the 6^{th} step of PCPE. Therefore, 3 real additions and 2 accesses to a LUT are needed.
6. The same as the 7^{th} step of PCPE. Therefore, $2K$ real additions and $4K$ real multiplications are needed.

The complexity of the whole single-stage BPS algorithm can be expressed as $5KB+2K+3$ real additions, $6KB+4K$ real multiplications, KB decisions, $B-1$ comparisons and 2 accesses to a LUT.

B.3 Two-stage blind phase search (BPS-BPS)

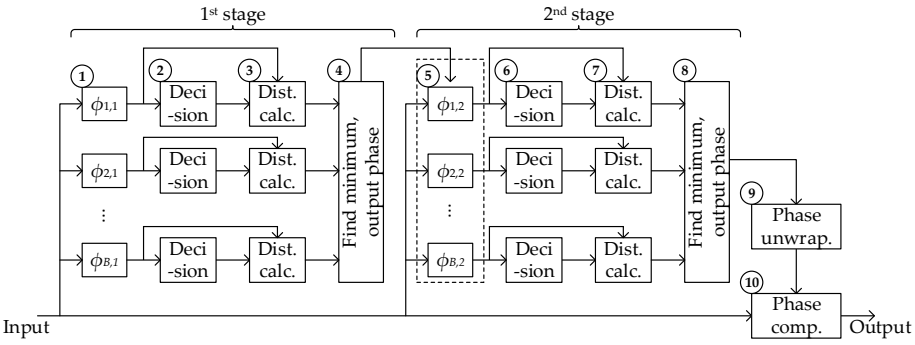

Figure B.3.: Block diagram of two-stage BPS with steps indicated with numbers.

1. Each value of the input block is rotated by B_1 test phases. Therefore, $2KB_1$ real additions and $4KB_1$ real multiplications are needed in this step.
2. Each point is decided and the decisions are assigned to them. Therefore, KB_1 decisions are needed in this step.
3. The distance is computed for each point. Therefore, $3KB_1$ real additions and $2KB_1$ real multiplications are needed in this step.
4. $B_1 - 1$ comparisons are needed to find the test phase associated with the minimum distance.
5. B_2 accesses to a LUT are made in this step to find the fine tuning test phases. Moreover, $2KB_2$ real additions and $4KB_2$ real multiplications are needed in order to perform the phase rotations.
6. Each point is decided again and the decision are assigned to them. Therefore, KB_2 decisions are needed in this step.
7. The distance is computed for each point. Therefore, $3KB_2$ real additions and $2KB_2$ real multiplications are needed in this step.
8. $B_2 - 1$ comparisons are needed to find the test phase associated with the minimum distance.
9. The same as the 6^{th} step of PCPE. Therefore, 3 real additions and 2 accesses to a LUT are needed.
10. The same as the 7^{th} step of PCPE. Therefore, $2K$ real additions and $4K$ real multiplications are needed.

The complexity of the whole two-stage blind phase search (BPS-BPS) algorithm can be expressed as $5KB_T + 2K + 3$ real additions, $6KB_T + 4K$ real multiplications, KB_T decisions, $B_T - 2$ comparisons and $B_2 + 2$ accesses to a LUT, where $B_T = B_1 + B_2$.

B.3.1 Hybrid PCPE and BPS scheme

1–7. These steps have the same complexity as the PCPE alone. Therefore, $6K + 3d$ real additions, $10K + 4d + 4$ real multiplications, 1 square root computation and 3 accesses to a LUT are needed in these steps.
8. Each value of the input block is rotated by B_2 test phases. Therefore, $2KB_2$ real additions and $4KB_2$ real multiplications are needed in this step.
9–11. These steps have the same complexity as the 6^{th}, 7^{th} and 8^{th} steps of BPS-BPS algorithm. Therefore, KB_2 decisions, $B_2 - 1$ comparisons,

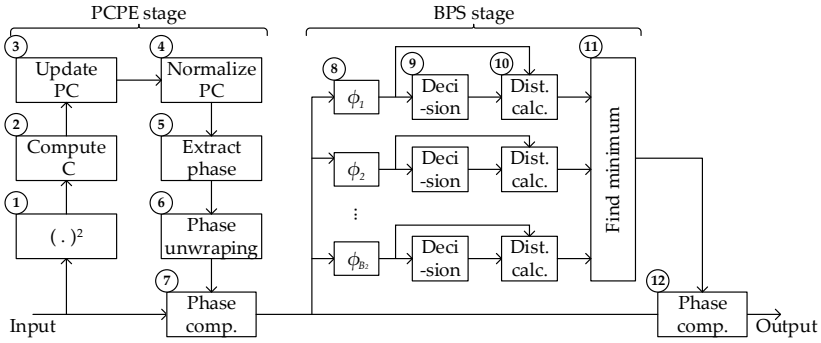

Figure B.4.: Block diagram of PCPE and BPS hybrid scheme with steps indicated with numbers.

$3KB_2$ real additions and $2KB_2$ real multiplications are needed in these steps.

12. The same as the 7^{th} step of PCPE. Therefore, $2K$ real additions and $4K$ real multiplications are needed.

The complexity of the whole PCPE-BPS algorithm can be expressed as $5KB_2 + 8K + 3d$ real additions, $6KB_2 + 14K + 4d + 4$ real multiplications, KB_2 decisions, $B_2 - 1$ comparisons and 3 accesses to a LUT.

C

Achievable information rates calculation

The MI between two random variables express the amount of information that can be extracted about the first random variable from the second random variable, and *vice versa*. The MI, $I(A;B)$, between the random variables A and B, is given by

$$I(A;B) = \sum_{a,b} p(a,b) \log_2 \frac{p(b|a)}{p(b)}, \tag{C.1}$$

where $p(a,b)$ is the joint probability distribution of A and B, $p(b)$ is the probability distribution of B, and $p(b|a)$ is the conditional probability distribution of B given the event $A = a$.

If A and B are, respectively, the input and output of an optical communication system, then the conditional probability $p(B|A)$ describes the communication channel. The actual channel statistics are generally unknown and the receiver needs to make decisions based on a possibly mismatched channel law [150], represented by an auxiliary channel model $q(B|A)$. The MI computed for this receiver, $\hat{I}(A;B)$, is a lower bound for the MI that would be achieved if the receiver knew the exact channel statistics [158], therefore,

$$\begin{aligned} \hat{I}(A;B) &= \sum_{a,b} p(a,b) \log_2 \frac{q(b|a)}{q_p(b)} \\ &= \mathbb{E}\left[\log_2 q(B|A) - \log_2 q_p(B)\right] \\ &\leq I(A;B), \end{aligned} \tag{C.2}$$

where $\mathbb{E}$ is the expectation operator, and $q_p(b)$ is the output distribution of the auxiliary channel, defined by

$$q_p(b) \triangleq \sum_a p(a)q(b|a). \tag{C.3}$$

The MI, $\hat{I}(A;B)$, is said to be an achievable information rate (AIR) [150, 158]. This means that by using an ideal forward error correction (FEC) code, it is possible to encode $\hat{I}(A;B)$ information bits in each symbol transmitted through the channel and decode them from the received symbols using a receiver optimized for the auxiliary channel model.

For the analyses made in Chapter 6, two channel models are considered. The first model considers a pure additive white Gaussian noise (AWGN) channel in which the received signal is only impaired by AWGN. The *a posteriori* probability for this model is given by

$$q_{AWGN}(b|a) = \frac{1}{\pi N_0} \exp\left\{-\frac{|(b-a)|^2}{N_0}\right\}, \tag{C.4}$$

where b and a are complex numbers, and N_0 is the noise estimated from the signal.

By the other hand, the second model considers that the signal is impaired by Gaussian noise on the amplitude of the signal, but differently from the previous model, it is impaired by a residual Gaussian phase noise on the phase of the signal. Then, for a change of variables given by

$$s_1 = \angle\frac{b}{a}, \tag{C.5}$$

$$s_2 = |b| - |a| \tag{C.6}$$

where s_1 and s_2 relate, respectively, to the phase and the amplitude differences between the received and transmitted signals, the *a posteriori* probability for the residual phase noise (RPN) model is a zero-mean bivariate Gaussian distribution given by

$$q_{RPN}(b|a) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{s_1^2}{\sigma_1^2} + \frac{s_2^2}{\sigma_2^2} - \frac{2\rho s_1 s_2}{\sigma_1\sigma_2}\right]\right\}, \tag{C.7}$$

where σ_1 and σ_2 are the standard deviations of s_1 and s_2, respectively, and ρ is the correlation between s_1 and s_2.

In order to exemplify these two models, AIR curves for a 64QAM signal at 32 GBd and combined LW of 4 MHz are shown in Figure C.1. The phase noise was compensated block by block ideally (as shown in Sec. 6.1.1), with each block having 64 symbols and no overlap between blocks. An AIR curve for a signal without phase noise was also plotted for comparison. For low SNR values, the AWGN contribution is more important than the residual phase noise, while for high SNR it is the opposite, therefore, the AWGN and RPN auxiliary channels approximate the actual AIR for the low and high SNR values, respectively.

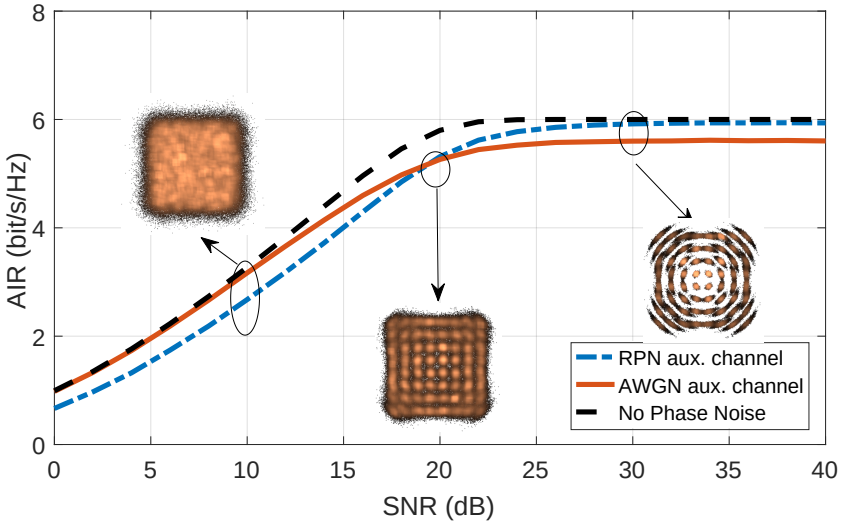

Figure C.1.: AIR for a 32 GBd 64QAM signal with 4 MHz combined linewidth and ideal block-wise phase compensation. Insets: constellations at 10, 20 and 30 dB SNR.

Acronyms

ADC	analog-to-digital converter
AI	artificial intelligence
AIR	achievable information rate
ASIC	application specific integrated circuit
AWG	arbitrary waveform generator
AWGN	additive white Gaussian noise
B2B	back-to-back
BER	bit error ratio
BPD	balanced photodetector diode
BPS	blind phase search
BPS-BPS	two-stage blind phase search
BPSK	binary phase-shift keying
BWPE	block-wise phase estimator
CC-GA	cooperative coevolutionary genetic algorithm
CD	chromatic dispersion
CMA	constant modulus algorithm
CPR	carrier phase recovery
CSR	cycle slip rate
CTA	clock tone amplitude

CW	continuous-wave
DA	data-aided
DAC	digital-to-analog converter
DCF	dispersion compensating fiber
DFT	discrete Fourier transform
DGD	differential group delay
DSO	digital storage oscilloscope
DSP	digital signal processing
DTSE	delay-tap sampling estimator
EDFA	erbium-doped fiber amplifier
EEPN	equalization-enhanced phase noise
FBG	fiber Bragg grating
FEC	forward error correction
FIR	finite impulse response
FMF	few-mode fiber
GA	genetic algorithm
GD	group delay
I	in-phase
IM-DD	intensity modulation and direct detection
IoT	internet of things
IQ	in-phase/quadrature
ISI	inter-symbol interference
LMS	least mean square
LO	local oscillator
LP	linearly polarized
LUT	look up table
LW	linewidth

MCF multi-core fiber

MCTA modified clock tone amplitude

MD modal dispersion

MDL mode-dependent loss

MDM mode division multiplexing

MI mutual information

MIMO multiple-input multiple-output

MMF multi-mode fiber

MO-GA multi-objective genetic algorithm

MSE mean square error

MZ Mach-Zehnder

MZM Mach-Zehnder modulator

NCO numeric controlled oscillator

NDA non-data-aided

NRZ non-return-to-zero

OADM optical add-drop multiplexer

OOK on-off keying

OSNR optical signal-to-noise ratio

P+I proportional-plus-integral

PBS polarization beam splitter

PC principal component

PCA principal component analysis

PCPE principal component-based phase estimation

PCPE-BPS PCPE and BPS hybrid scheme

PD photodetector diode

PDL polarization-dependent loss

PDM polarization division multiplexing

PIM	power iteration method
PLL	phase-locked loop
PM	polarization multiplexed
PMD	polarization mode dispersion
PRBS	pseudo random bit sequence
PSK	phase-shift keying
Q	quadrature
QAM	quadrature amplitude modulation
QPSK	quadrature phase shift keying
RC	raised cosine
RDE	radially directed equalizer
RPN	residual phase noise
RRC	root raised cosine
RZ	return-to-zero
SDM	spatial division multiplexing
SMF	single-mode fiber
SNR	signal-to-noise ratio
SOP	state of polarization
TED	timing error detector
TIA	transimpedance amplifier
V&V	Viterbi-Viterbi
WDM	wavelength division multiplexing
X	horizontal polarization
XY	inter-polarization
Y	vertical polarization

Bibliography

[1] T. A. Oliveira, A. C. Barbosa, H. Ramalhinho, and M. Oliver, "Citizens and information and communication technologies", in *2018 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2018, pp. 1–7. DOI: 10.1109/IJCNN.2018.8489031 (cited on page 1).

[2] M. Castells, *Information Technology, Globalization and Social Development*, ser. Discussion paper - United Nations Research Institute for Social Development no. 114. Geneva, Switzerland: UNRISD, Sep. 1999. [Online]. Available: http://www.unrisd.org/80256B3C005BCCF9/(httpAuxPages)/F270E0C066F3DE7780256B67005B728C (cited on page 1).

[3] E. Agrell, M. Karlsson, A. R. Chraplyvy, D. J. Richardson, P. M. Krummrich, P. Winzer, K. Roberts, J. K. Fischer, S. J. Savory, B. J. Eggleton, M. Secondini, F. R. Kschischang, A. Lord, J. Prat, I. Tomkos, J. E. Bowers, S. Srinivasan, M. Brandt-Pearce, and N. Gisin, "Roadmap of optical communications", *J. Opt.*, vol. 18, no. 6, Art. no. 063002, May 2016. DOI: 10.1088/2040-8978/18/6/063002 (cited on pages 1, 21, 25, 49).

[4] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*. New York: John Wiley & Sons, Inc., Oct. 2001. DOI: 10.1002/0471200573 (cited on page 1).

[5] T. Rappaport, *Wireless communications : principles and practice*. Upper Saddle River, N.J: Prentice Hall PTR, 2002, ISBN: 9780130422323 (cited on page 1).

[6] M. Alencar and V. C. da Rocha Jr., *Communication systems*. Boston, MA: Springer, 2005, ISBN: 978-0-387-27097-5. DOI: 10.1007/b138483 (cited on page 1).

[7] L. Kazovsky, "Multichannel coherent optical communications systems", *J. Lightw. Technol.*, vol. 5, no. 8, pp. 1095–1102, Aug. 1987. DOI: 10.1109/JLT.1987.1075618 (cited on page 1).

[8] F. Derr, "Coherent optical QPSK intradyne system: Concept and digital receiver realization", *J. Lightw. Technol.*, vol. 10, no. 9, pp. 1290–1296, Sep. 1992. DOI: 10.1109/50.156881 (cited on page 1).

[9] S. Betti, G. D. Marchis, and E. Iannone, "Applications of coherent techniques to optical transmission: State of the art and perspectives", *Eur. Trans. Telecomm.*, vol. 4, no. 6, pp. 641–650, Nov. 1993. DOI: 10.1002/ett.4460040613 (cited on page 1).

[10] E. Desurvire, J. R. Simpson, and P. C. Becker, "High-gain erbium-doped traveling-wave fiber amplifier", *Opt. Lett.*, vol. 12, no. 11, pp. 888–890, Nov. 1987. DOI: 10.1364/OL.12.000888 (cited on page 1).

[11] T. Miki and H. Ishio, "Viabilities of the wavelength-division-multiplexing transmission system over an optical fiber cable", *IEEE Trans. Commun.*, vol. 26, no. 7, pp. 1082–1087, Jul. 1978. DOI: 10.1109/tcom.1978.1094172 (cited on page 2).

[12] C. Brackett, "Dense wavelength division multiplexing networks: Principles and applications", *IEEE J. Sel. Areas Commun.*, vol. 8, no. 6, pp. 948–964, Aug. 1990. DOI: 10.1109/49.57798 (cited on page 2).

[13] D. Chen, S. Wheeler, D. Nguyen, A. Färbert, A. Schöpflin, A. Richter, C.-J. Weiske, K. Kotten, P. M. Krummrich, A. Schex, and C. Glingener, "3.2 Tb/s field trial (80 × 40 Gb/s) over 3 × 82 km SSMF using FEC, raman and tunable dispersion compensation", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Anaheim, CA, USA, Mar. 2001. DOI: 10.1364/ofc.2001.pd36 (cited on page 2).

[14] J. Kahn and K.-P. Ho, "Spectral efficiency limits and modulation/detection techniques for DWDM systems", *IEEE J. Sel. Top. Quantum Electron.*, vol. 10, no. 2, pp. 259–272, Mar. 2004. DOI: 10.1109/jstqe.2004.826575 (cited on page 2).

[15] S. J. Savory, "Digital coherent optical receivers: Algorithms and subsystems", *IEEE J. Sel. Top. Quantum Electron.*, vol. 16, no. 5, pp. 1164–1179, Sep. 2010, ISSN: 1077-260X. DOI: 10.1109/JSTQE.2010.2044751 (cited on pages 2, 5, 21, 25, 27, 49, 71, 77, 99).

[16] E. Ip, A. P. T. Lau, D. J. F. Barros, and J. M. Kahn, "Coherent detection in optical fiber systems", *Opt. Express*, vol. 16, no. 2, pp. 753–791, 2008. DOI: 10.1364/oe.16.000753 (cited on page 2).

[17] S. J. Savory, G. Gavioli, R. I. Killey, and P. Bayvel, "Electronic compensation of chromatic dispersion using a digital coherent receiver", *Opt. Express*, vol. 15, no. 5, pp. 2120–2126, Mar. 2007. DOI: 10.1364/oe.15.002120 (cited on page 2).

[18] M. Taylor, "Coherent detection method using DSP for demodulation of signal and subsequent equalization of propagation impairments", *IEEE Photonics Technol. Lett.*, vol. 16, no. 2, pp. 674–676, Feb. 2004. DOI: 10.1109/lpt.2003.823106 (cited on page 2).

[19] S. J. Savory, "Digital filters for coherent optical receivers", *Opt. Express*, vol. 16, no. 2, pp. 804–817, Feb. 2008. DOI: 10.1364/oe.16.000804 (cited on pages 2, 5, 21).

[20] S. J. Savory, A. D. Stewart, S. Wood, G. Gavioli, M. G. Taylor, R. I. Killey, and P. Bayvel, "Digital equalisation of 40Gbit/s per wavelength transmission over 2480km of standard fibre without optical dispersion compensation", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Cannes, France, Sep. 2006. DOI: 10.1109/ecoc.2006.4800978 (cited on page 2).

[21] P. J. Winzer and R.-J. Essiambre, "Advanced modulation formats for high-capacity optical transport networks", *J. Lightw. Technol.*, vol. 24, no. 12, pp. 4711–4728, Dec. 2006. DOI: 10.1109/jlt.2006.885260 (cited on page 2).

[22] M. S. Faruk and S. J. Savory, "Digital signal processing for coherent transceivers employing multilevel formats", *J. Lightw. Technol.*, vol. 35, no. 5, pp. 1125–1141, Mar. 2017, ISSN: 0733-8724. DOI: 10.1109/JLT.2017.2662319 (cited on pages 2, 3, 5, 23, 49, 74).

[23] P. Winzer and R.-J. Essiambre, "Advanced optical modulation formats", *Proc. IEEE*, vol. 94, no. 5, pp. 952–985, May 2006. DOI: 10.1109/jproc.2006.873438 (cited on page 2).

[24] M. Seimetz, "Performance of coherent optical square-16-QAM-systems based on IQ-transmitters and homodyne receivers with digital phase estimation", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Anaheim, CA, USA, Mar. 2006. DOI: 10.1109/ofc.2006.215620 (cited on page 2).

[25] *Telia carrier and Infinera demonstrate first 600G transmission in a production network*, Press Release, Sunnyvale, CA, USA: Infinera Corp. [Online]. Available: http://investors.infinera.com/new-releases/press-release-details/2018/Telia-Carrier-and-Infinera-Demonstrate-First-600G-Transmission-in-a-Production-Network/ (cited on page 2).

[26] Cisco, "Cisco visual networking index: Forecast and trends, 2017—2022", Cisco, White Paper, Nov. 2018. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html (cited on page 2).

[27] P. J. Winzer, D. T. Neilson, and A. R. Chraplyvy, "Fiber-optic transmission and networking: The previous 20 and the next 20 years [invited]", *Opt. Express*, vol. 26, no. 18, pp. 24 190–24 239, Sep. 2018. DOI: 10.1364/OE.26.024190 (cited on page 3).

[28] T. J. Xia and G. A. Wellbrock, "Commercial 100-Gbit/s coherent transmission systems", in *Optical Fiber Telecommunications*, I. Kaminow, T. Li, and A. Willner, Eds. Oxford, UK: Academic Press, 2013, vol. VIB, ch. 2, pp. 491–568. DOI: 10.1016/B978-0-12-396960-6.00002-X (cited on pages 3, 8).

[29] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks", *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010. DOI: 10.1109/jlt.2009.2039464 (cited on page 3).

[30] G. P. Agrawal, *Nonlinear Fiber Optics*. Academic Press, Oct. 2012, ISBN: 0123970237 (cited on page 3).

[31] T. Pfau, S. Hoffmann, and R. Noe, "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for *m*-QAM constellations", *J. Lightw. Technol.*, vol. 27, no. 8, pp. 989–999, Apr. 2009. DOI: 10.1109/jlt.2008.2010511 (cited on pages 3, 24, 85, 87, 110).

[32] S. M. Bilal and G. Bosco, "Automatic bias control of mach-zehnder modulators for QPSK and QAM systems", *J. Opt. Tech.*, vol. 81, no. 7, pp. 403–407, Jul. 2014. DOI: 10.1364/jot.81.000403 (cited on page 3).

[33] R. Rios-Müller, J. Renaudier, and G. Charlet, "Blind receiver skew compensation and estimation for long-haul non-dispersion managed systems using adaptive equalizer", *J. Lightw. Technol.*, vol. 33, no. 7, pp. 1315–1318, Apr. 2015, ISSN: 0733-8724. DOI: 10.1109/JLT.2014.2377582 (cited on pages 3, 50, 77, 79, 82, 109).

[34] M. Paskov, D. Lavery, and S. J. Savory, "Blind equalization of receiver in-phase/quadrature skew in the presence of nyquist filtering", *IEEE Photonics Technol. Lett.*, vol. 25, no. 24, pp. 2446–2449, Dec. 2013, ISSN: 1041-1135. DOI: 10.1109/LPT.2013.2288114 (cited on pages 3, 50, 77, 79, 82, 109).

[35] T. H. Nguyen, P. Scalart, M. Gay, L. Bramerie, O. Sentieys, J. C. Simon, C. Peucheret, and M. Joindot, "Blind transmitter IQ imbalance compensation in M-QAM optical coherent systems", *IEEE J. Opt. Commun. Netw.*, vol. 9, no. 9, pp. D42–D50, Sep. 2017, ISSN: 1943-0620. DOI: 10.1364/JOCN.9.000D42 (cited on pages 3, 50).

[36] P. J. Winzer, "Energy-efficient optical transport capacity scaling through spatial multiplexing", *IEEE Photonics Technol. Lett.*, vol. 23, no. 13, pp. 851–853, Jul. 2011. DOI: 10.1109/lpt.2011.2140103 (cited on page 3).

[37] G. M. Saridis, D. Alexandropoulos, G. Zervas, and D. Simeonidou, "Survey and evaluation of space division multiplexing: From technologies to optical networks", *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2136–2156, Aug. 2015. DOI: 10.1109/comst.2015.2466458 (cited on page 3).

[38] R.-J. Essiambre and R. W. Tkach, "Capacity trends and limits of optical communication networks", *Proc. IEEE*, vol. 100, no. 5, pp. 1035–1055, May 2012. DOI: 10.1109/jproc.2012.2182970 (cited on page 3).

[39] D. J. Richardson, J. M. Fini, and L. E. Nelson, "Space-division multiplexing in optical fibres", *Nature Photonics*, vol. 7, no. 5, pp. 354–362, May 2013. DOI: 10.1038/nphoton.2013.94 (cited on page 3).

[40] P. J. Winzer and G. J. Foschini, "MIMO capacities and outage probabilities in spatially multiplexed optical transport systems", *Opt. Express*, vol. 19, no. 17, pp. 16 680–16 696, Aug. 2011. DOI: 10.1364/oe.19.016680 (cited on page 3).

[41] S. O. Arik, J. M. Kahn, and K.-P. Ho, "MIMO signal processing for mode-division multiplexing: An overview of channel models and signal processing architectures", *IEEE Signal Process. Mag.*, vol. 31, no. 2, pp. 25–34, Mar. 2014. DOI: 10.1109/msp.2013.2290804 (cited on page 3).

[42] R. Asif, F. Ye, and T. Morioka, "Equalizer complexity for 6-LP mode 112 Gbit/s m-ary DP-QAM space division multiplexed transmission in strongly coupled few-mode-fibers", in *Proc. Eur. Conf. Netw. and Commun. (EuCNC)*, Paris, France, Jun. 2015. DOI: 10.1109/eucnc.2015.7194090 (cited on pages 3, 23, 26).

[43] D. Zibar, M. Piels, R. Jones, and C. G. Schaeffer, "Machine learning techniques in optical communication", *J. Lightw. Technol.*, vol. 34, no. 6, pp. 1442–1452, Mar. 2016. DOI: 10.1109/jlt.2015.2508502 (cited on page 4).

[44] F. N. Khan, Q. Fan, C. Lu, and A. P. T. Lau, "An optical communication's perspective on machine learning and its applications", *J. Lightw. Technol.*, Early Access, Feb. 2019. DOI: 10.1109/jlt.2019.2897313 (cited on page 4).

[45] F. N. Khan, Y. Zhou, A. P. T. Lau, and C. Lu, "Modulation format identification in heterogeneous fiber-optic networks using artificial neural networks", *Opt. Express*, vol. 20, no. 11, May 2012. DOI: 10.1364/oe.20.012422 (cited on page 4).

[46] F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini, and M. Tornatore, "An overview on application of machine learning techniques in optical networks", *IEEE Commun. Surveys Tuts.*, Early Access, 2018. DOI: 10.1109/comst.2018.2880039 (cited on page 4).

[47] J. Mata, I. de Miguel, R. J. Durán, N. Merayo, S. K. Singh, A. Jukan, and M. Chamania, "Artificial intelligence (AI) methods in optical networks: A comprehensive survey", *Opt. Switching Netw.*, vol. 28, Apr. 2018. DOI: 10.1016/j.osn.2017.12.006 (cited on page 4).

[48] D. Zibar, L. Carvalho, M. Piels, A. Doberstein, J. Diniz, B. Nebendahl, C. Franciscangelis, J. Estaran, H. Haisch, N. G. Gonzalez, J. R. F. de Oliveira, and I. T. Monroy, "Bayesian filtering for phase noise characterization and carrier synchronization of up to 192 Gb/s PDM 64-QAM", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Cannes, France, Sep. 2014, paper Tu.1.3.1. DOI: 10.1109/ecoc.2014.6963844 (cited on pages 4, 24).

[49] M. Kuschnerov, F. Hauske, K. Piyawanno, B. Spinnler, M. Alfiad, A. Napoli, and B. Lankl, "DSP for coherent single-carrier receivers", *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3614–3622, Aug. 2009. DOI: 10.1109/jlt.2009.2024963 (cited on page 5).

[50] K.-P. Ho and J. M. Kahn, "Statistics of group delays in multimode fiber with strong mode coupling", *J. Lightw. Technol.*, vol. 29, no. 21, pp. 3119–3128, Nov. 2011. DOI: 10.1109/jlt.2011.2165316 (cited on page 5).

[51] G. P. Agrawal, *Fiber-Optic Communication Systems*. John Wiley & Sons, Inc., Oct. 2010, ISBN: 0470505117. DOI: 10.1002/9780470918524 (cited on pages 5, 7, 14, 18, 19, 35, 42).

[52] M. Seimetz, *High-Order Modulation for Optical Fiber Transmission*. Heidelberg: Springer, 2009, ch. 2. DOI: 10.1007/978-3-540-93771-5 (cited on pages 5, 10, 20).

[53] D. van den Borne, "Robust optical transmission systems: Modulation and equalization", eng, PhD thesis, Technische Universiteit Eindhoven, 2008. DOI: 10.6100/ir633535 (cited on page 5).

[54] S. Haykin, *Communication Systems*. Wiley, Mar. 2009, ISBN: 0471697907 (cited on page 5).

[55] M. G. Taylor, "Phase estimation methods for optical coherent detection using digital signal processing", *J. Lightw. Technol.*, vol. 27, no. 7, pp. 901–914, Apr. 2009. DOI: 10.1109/jlt.2008.927778 (cited on pages 8, 95).

[56] J. G. Proakis, *Digital Communications*. Boston: McGraw Hill Higher Education, Dec. 2000, ISBN: 0071181830 (cited on page 8).

[57] *Optical DP-QAM transmitter*, SHF 46215 B, V003, SHF Communication Technologies AG, Dec. 2012 (cited on page 11).

[58] D. Gloge, "Weakly guiding fibers", *Appl. Opt.*, vol. 10, no. 10, p. 2252, Oct. 1971. DOI: 10.1364/ao.10.002252 (cited on pages 14, 15).

[59] Y. H. Wang and X. Zhang, "Elliptical Fourier series expansion method together with cutoff frequencies in elliptical optical waveguides", *J. Lightw. Technol.*, vol. 16, no. 10, pp. 1933–1941, Oct. 1998. DOI: 10.1109/50.721083 (cited on page 15).

[60] I. Gómez-Castellanos and R. M. Rodríguez-Dagnino, "Intensity distributions and cutoff frequencies of linearly polarized modes for a step-index elliptical optical fiber", *Opt. Eng.*, vol. 46, no. 4, p. 045 003, Apr. 2007. DOI: 10.1117/1.2719698 (cited on page 15).

[61] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*. Hoboken: Wiley-Interscience, Mar. 2007, ISBN: 0471358320 (cited on page 15).

[62] T. A. Birks, I. Gris-Sánchez, S. Yerolatsitis, S. G. Leon-Saval, and R. R. Thomson, "The photonic lantern", *Adv. Opt. Photonics*, vol. 7, no. 2, p. 107, Apr. 2015. DOI: 10.1364/aop.7.000107 (cited on page 19).

[63] S. O. Arik, K.-P. Ho, and J. M. Kahn, "Group delay management and multiinput multioutput signal processing in mode-division multiplexing systems", *J. Lightw. Technol.*, vol. 34, no. 11, pp. 2867–2880, Jun. 2016. DOI: 10.1109/jlt.2016.2530978 (cited on pages 19, 26, 43, 108).

[64] I. Fatadin, S. J. Savory, and D. Ives, "Compensation of quadrature imbalance in an optical QPSK coherent receiver", *IEEE Photonics Technol. Lett.*, vol. 20, no. 20, pp. 1733–1735, Oct. 2008. DOI: 10.1109/lpt.2008.2004630 (cited on page 21).

[65] M. S. Faruk and K. Kikuchi, "Compensation for in-phase/quadrature imbalance in coherent-receiver front end for optical quadrature amplitude modulation", *IEEE Photon. J.*, vol. 5, no. 2, pp. 7 800 110–7 800 110, Apr. 2013. DOI: 10.1109/jphot.2013.2251872 (cited on page 21).

[66] E. Ip and J. M. Kahn, "Feedforward carrier recovery for coherent optical communications", *J. Lightw. Technol.*, vol. 25, no. 9, pp. 2675–2692, Sep. 2007. DOI: 10.1109/jlt.2007.902118 (cited on pages 21, 95).

[67] V. Parahyba, F. Simões, E. Rosa, J. Oliveira, J. Oliveira, E. Filho, E. Schneider, S. Ranzini, L. Carvalho, J. Reis, and J. Diniz, "Performance against implementation of digital backpropagation for high-speed coherent optical systems", *Electron. Lett.*, vol. 51, no. 14, pp. 1094–1096, Jul. 2015. DOI: 10.1049/el.2015.0576 (cited on page 22).

[68] C. Malouin, B. Zhang, A. Wagner, S. Khatana, E. Ibragimov, H. Jiang, and T. Schmidt, "Sub-rate sampling in 100 Gb/s coherent optical receivers", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, San Diego, CA, USA, Mar. 2010. DOI: 10.1364/ofc.2010.otht3 (cited on page 23).

[69] I. Slim, A. Mezghani, L. G. Baltar, J. Qi, F. N. Hauske, and J. A. Nossek, "Delayed single-tap frequency-domain chromatic-dispersion compensation", *IEEE Photonics Technol. Lett.*, vol. 25, no. 2, pp. 167–170, Jan. 2013. DOI: 10.1109/lpt.2012.2231064 (cited on page 23).

[70] S. M. Ranzini, E. C. Magalhães, V. B. Ribeiro, V. V. Nascimento, and J. C. R. F. Oliveira, "Accurate blind chromatic dispersion estimation in long-haul 112Gbit/s PM-QPSK WDM coherent systems", in *Signal Processing in Photonic Communications (part of Advanced Photonics Congress)*, Colorado Springs, CO, USA, Jun. 2012, paper SPTh2B. DOI: 10.1364/sppcom.2012.spth2b.3 (cited on pages 23, 77, 79, 81).

[71] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*. New York: Springer, 1997. DOI: 10.1007/978-1-4899-1807-9 (cited on page 23).

[72] A. Kakkar, J. R. Navarro, R. Schatz, H. Louchet, X. Pang, O. Ozolins, G. Jacobsen, and S. Popov, "Comprehensive study of equalization-enhanced phase noise in coherent optical systems", *J. Lightw. Technol.*, vol. 33, no. 23, pp. 4834–4841, Dec. 2015. DOI: 10.1109/jlt.2015.2491363 (cited on page 23).

[73] K.-P. Ho, A. P. T. Lau, and W. Shieh, "Equalization-enhanced phase noise induced timing jitter", *Opt. Lett.*, vol. 36, no. 4, p. 585, Feb. 2011. DOI: 10.1364/ol.36.000585 (cited on page 23).

[74] A. P. T. Lau, W. Shieh, and K.-P. Ho, "Equalization-enhanced phase noise for 100Gb/s transmission and beyond with coherent detection", in *Proc. IEEE Int. Conf. Comm. Sys.*, Singapore, Nov. 2010. DOI: 10.1109/iccs.2010.5686612 (cited on page 23).

[75] F. Gardner, "A BPSK/QPSK timing-error detector for sampled receivers", *IEEE Trans. Commun.*, vol. 34, no. 5, pp. 423–429, May 1986. DOI: 10.1109/tcom.1986.1096561 (cited on pages 23, 28, 30).

[76] D. Godard, "Passband timing recovery in an all-digital modem receiver", *IEEE Trans. Commun.*, vol. 26, no. 5, pp. 517–523, May 1978. DOI: 10.1109/tcom.1978.1094107 (cited on pages 23, 31).

[77] W.-P. Zhu, Y. Yan, M. Ahmad, and M. Swamy, "Feedforward symbol timing recovery technique using two samples per symbol", *IEEE Trans. Circuits Syst. I*, vol. 52, no. 11, pp. 2490–2500, Nov. 2005. DOI: 10.1109/tcsi.2005.853902 (cited on page 23).

[78] N. Stojanovic, C. Xie, yu zhao, B. Mao, N. Gonzalez, J. Qi, and N. Binh, "Modified gardner phase detector for nyquist coherent optical transmission systems", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Anaheim, CA, USA, Mar. 2013, paper JTh2A.50. DOI: 10.1364/nfoec.2013.jth2a.50 (cited on pages 23, 32).

[79] M. Yan, Z. Tao, L. Dou, L. Li, Y. Zhao, T. Hoshida, and J. C. Rasmussen, "Digital clock recovery algorithm for nyquist signal", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Anaheim, CA, USA, Mar. 2013, paper OTu2I.7. DOI: 10.1364/OFC.2013.OTu2I.7 (cited on pages 23, 32, 52).

[80] K.-T. Wu and H. Sun, "Frequency-domain clock phase detector for nyquist WDM systems", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, San Francisco, CA, USA, Mar. 2014, paper Th3E.2. DOI: 10.1364/ofc.2014.th3e.2 (cited on page 23).

[81] X. Zhou and X. Chen, "Parallel implementation of all-digital timing recovery for high-speed and real-time optical coherent receivers", *Opt. Express*, vol. 19, no. 10, pp. 9282–9295, Apr. 2011. DOI: 10.1364/oe.19.009282 (cited on page 23).

[82] K. Kikuchi, "Clock recovering characteristics of adaptive finite-impulse-response filters in digital coherent optical receivers", *Opt. Express*, vol. 19, no. 6, pp. 5611–5619, Mar. 2011. DOI: 10.1364/oe.19.005611 (cited on page 23).

[83] D. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems", *IEEE Trans. Commun.*, vol. 28, no. 11, pp. 1867–1875, Nov. 1980. DOI: 10.1109/tcom.1980.1094608 (cited on page 23).

[84] V. B. Ribeiro, J. C. Oliveira, J. C. Diniz, E. Rosa, R. Silva, E. P. Silva, L. H. Carvalho, and A. C. Bordonalli, "Enhanced digital polarization demultiplexation via CMA step size adaptation for PM-QPSK coherent receivers", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, CA, USA, Mar. 2012, OW3H.5. DOI: 10.1364/ofc.2012.ow3h.4 (cited on page 23).

[85] I. Fatadin, D. Ives, and S. Savory, "Blind equalization and carrier phase recovery in a 16-QAM optical coherent system", *J. Lightw. Technol.*, vol. 27, no. 15, pp. 3042–3049, Aug. 2009. DOI: 10.1109/jlt.2009.2021961 (cited on pages 23, 24).

[86] J. Shynk, "Frequency-domain and multirate adaptive filtering", *IEEE Signal Process. Mag.*, vol. 9, no. 1, pp. 14–37, Jan. 1992. DOI: 10.1109/79.109205 (cited on page 23).

[87] S. O. Arik, D. Askarov, and J. M. Kahn, "Adaptive frequency-domain equalization in mode-division multiplexing systems", *J. Lightw. Technol.*, vol. 32, no. 10, pp. 1841–1852, May 2014, ISSN: 0733-8724. DOI: 10.1109/jlt.2014.2303079 (cited on pages 23, 26).

[88] K. Shi and B. C. Thomsen, "Sparse adaptive frequency domain equalizers for mode-group division multiplexing", *J. Lightw. Technol.*, vol. 33, no. 2, pp. 311–317, Jan. 2015. DOI: 10.1109/jlt.2014.2374837 (cited on pages 23, 26, 108).

[89] D. Lee, K. Shibahara, T. Kobayashi, T. Mizuno, H. Takara, A. Sano, H. Kawakami, T. Nakagawa, and Y. Miyamoto, "A sparsity managed adaptive MIMO equalization for few-mode fiber transmission with various differential mode delays", *J. Lightw. Technol.*, vol. 34, no. 8, pp. 1754–1761, Apr. 2016. DOI: 10.1109/jlt.2015.2511178 (cited on pages 23, 26, 108).

[90] S. Randel, R. Ryf, A. Sierra, P. J. Winzer, A. H. Gnauck, C. A. Bolle, R.-J. Essiambre, D. W. Peckham, A. McCurdy, and R. Lingle, "6×56-Gb/s mode-division multiplexed transmission over 33-km few-mode fiber enabled by 6×6 MIMO equalization", *Opt. Express*, vol. 19, no. 17, pp. 16 697–16 707, Aug. 2011. DOI: 10.1364/oe.19.016697 (cited on pages 23, 26).

[91] S. Randel, A. Sierra, S. Mumtaz, A. Tulino, R. Ryf, P. Winzer, C. Schmidt, and R. Essiambre, "Adaptive MIMO signal processing for mode-division multiplexing", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, CA, USA, Mar. 2012, paper OW3D.5. DOI: 10.1364/ofc.2012.ow3d.5 (cited on pages 23, 26).

[92] S. Randel and P. Winzer, "DSP for mode division multiplexing", in *Proc. Optoelecton. Commun. Conf. / Phot. Switching (OECC/PS)*, Kyoto, Japan, Jun. 2013, paper ThR3-1 (cited on pages 23, 26).

[93] J. C. M. Diniz, J. C. R. F. de Oliveira, E. S. Rosa, V. B. Ribeiro, V. E. S. Parahyba, R. da Silva, E. P. da Silva, L. H. H. de Carvalho, A. F. Herbster, and A. C. Bordonalli, "Simple feed-forward wide-range frequency offset estimator for optical coherent receivers", *Opt. Express*, vol. 19, no. 26, B323, Nov. 2011. DOI: 10.1364/oe.19.00b323 (cited on pages 24, 99).

[94] T. Nakagawa, K. Ishihara, T. Kobayashi, R. Kudo, M. Matsui, Y. Takatori, and M. Mizoguchi, "Wide-range and fast-tracking frequency offset estimator for optical coherent receivers", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Torino, Italy, Sep. 2010, We.7.A.2. DOI: 10.1109/ecoc.2010.5621111 (cited on page 24).

[95] A. Leven, N. Kaneda, U.-V. Koc, and Y.-K. Chen, "Frequency estimation in intradyne reception", *IEEE Photonics Technol. Lett.*, vol. 19, no. 6, pp. 366–368, Mar. 2007. DOI: 10.1109/lpt.2007.891893 (cited on page 24).

[96] 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 Technol. Lett.*, vol. 20, no. 18, pp. 1569–1571, Sep. 2008. DOI: 10.1109/lpt.2008.928846 (cited on page 24).

[97] M. Morelli and U. Mengali, "Feedforward frequency estimation for PSK: A tutorial review", *Eur. Trans. Telecomm.*, vol. 9, no. 2, pp. 103–116, Mar. 1998. DOI: 10.1002/ett.4460090203 (cited on page 24).

[98] A. J. Viterbi and A. M. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission", *IEEE Trans. Inf. Theory*, vol. 29, no. 4, pp. 543–551, Jul. 1983. DOI: 10.1109/tit.1983.1056713 (cited on pages 24, 85).

[99] M. Taylor, "Accurate digital phase estimation process for coherent detection using a parallel digital processor", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Vienna, Austria, Sep. 2005. DOI: 10.1049/cp:20050471 (cited on page 24).

[100] A. Bisplinghoff, S. Langenbach, E. Vercelli, R. Pastorelli, and T. Kupfer, "Cycle slip tolerant, differential encoding aware soft-decision FEC", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, CA, USA, Mar. 2015, paper Tu3B.2. DOI: 10.1364/ofc.2015.tu3b.2 (cited on page 24).

[101] H. Cheng, Y. Li, F. Zhang, J. Wu, J. Lu, G. Zhang, J. Xu, and J. Lin, "Pilot-symbols-aided cycle slip mitigation for DP-16qam optical communication systems", *Opt. Express*, vol. 21, no. 19, pp. 22 166–22 172, Sep. 2013. DOI: 10.1364/oe.21.022166. [Online]. Available: https://doi.org/10.1364/oe.21.022166 (cited on page 24).

[102] Y. Gao, E. Ha, A. P. T. Lau, C. Lu, X. Xu, and L. Li, "Non-data-aided and universal cycle slip detection and correction for coherent communication systems", *Opt. Express*, vol. 22, no. 25, pp. 31 167–31 179, Dec. 2014. DOI: 10.1364/oe.22.031167 (cited on page 24).

[103] V. N. Rozental, D. Kong, B. Foo, B. Corcoran, and A. J. Lowery, "Cycle-slip-less low-complexity phase recovery algorithm for coherent optical receivers", *Opt. Lett.*, vol. 42, no. 18, pp. 3554–3557, Sep. 2017. DOI: 10.1364/ol.42.003554 (cited on pages 24, 85).

[104] D. Zibar, J. C. R. F. de Olivera, V. B. Ribeiro, A. Paradisi, J. C. Diniz, K. J. Larsen, and I. T. Monroy, "Experimental investigation and digital compensation of DGD for 112 Gb/s PDM-QPSK clock recovery", *Opt. Express*, vol. 19, no. 26, pp. B429–B439, Nov. 2011. DOI: 10.1364/oe.19.00b429 (cited on pages 25, 26, 41, 43).

[105] H. Sun and K.-T. Wu, "A novel dispersion and pmd tolerant clock phase detector for coherent transmission systems", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, CA, USA, Mar. 2011, paper OMJ.4 (cited on page 25).

[106] D. Zibar, A. Bianciotto, Z. Wang, A. Napoli, and B. Spinnler, "Analysis and dimensioning of fully digital clock recovery for 112 Gb/s coherent polmux QPSK systems", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Vienna, Austria, Sep. 2009, paper 7.3.4 (cited on pages 25, 26, 41).

[107] D. Zibar, J. C. R. de Oliviera, V. B. Ribeiro, A. Paradisi, J. C. Diniz, K. J. Larsen, and I. T. Monroy, "Experimental investigation of digital compensation of DGD for 112 Gb/s PDM-QPSK clock recovery", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Geneva, Switzerland, Sep. 2011, paper Tu.6.A.4. DOI: 10.1364/ecoc.2011.tu.6.a.4 (cited on pages 25, 26).

[108] N. Stojanović, C. Xie, Y. Zhao, B. Mao, and N. Guerrero Gonzalez, "A circuit enabling clock extraction in coherent receivers", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Amsterdam, Netherlands, Sep. 2012, paper P3.08. DOI: 10.1364/eceoc.2012.p3.08 (cited on pages 25, 26, 31, 43, 53, 107).

[109] M. Birk, P. Gerard, R. Curto, L. E. Nelson, X. Zhou, P. Magill, T. J. Schmidt, C. Malouin, B. Zhang, E. Ibragimov, S. Khatana, M. Glavanovic, R. Lofland, R. Marcoccia, R. Saunders, G. Nicholl, M. Nowell, and F. Forghieri, "Coherent 100 Gb/s PM-QPSK field trial", *IEEE Commun. Mag.*, vol. 48, no. 7, pp. 52–60, Jul. 2010. DOI: 10.1109/MCOM.2010.5496878 (cited on page 25).

[110] P. J. Winzer and D. T. Neilson, "From scaling disparities to integrated parallelism: A decathlon for a decade", *J. Lightw. Technol.*, vol. 35, no. 5, pp. 1099–1115, Mar. 2017, ISSN: 0733-8724. DOI: 10.1109/jlt.2017.2662082 (cited on page 25).

[111] P. Winzer, "Spatial multiplexing: The next frontier in network capacity scaling", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, London, UK, Sep. 2013, paper We.1.D.1. DOI: 10.1049/cp.2013.1397 (cited on page 25).

[112] C. Antonelli, A. Mecozzi, and M. Shtaif, "Scaling of inter-channel nonlinear interference noise and capacity with the number of strongly coupled modes in SDM systems", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Anaheim, CA, USA, Mar. 2016, paper W4I.2. DOI: https://doi.org/10.1364/OFC.2016.W4I.2 (cited on page 26).

[113] F. Ferreira, N. M. Suibhne, C. Sánchez, S. Sygletos, and A. D. Ellis, "Advantages of strong mode coupling for suppression of nonlinear distortion in few-mode fibers", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Anaheim, CA, USA, Mar. 2016, paper Tu2E.3. DOI: 10.1364/ofc.2016.tu2e.3 (cited on pages 26, 108).

[114] K. Mueller and M. Muller, "Timing recovery in digital synchronous data receivers", *IEEE Trans. Commun.*, vol. 24, no. 5, pp. 516–531, May 1976. DOI: 10.1109/tcom.1976.1093326 (cited on page 27).

[115] L. Huang, D. Wang, A. P. T. Lau, C. Lu, and S. He, "Performance analysis of blind timing phase estimators for digital coherent receivers", *Opt. Express*, vol. 22, no. 6, pp. 6749–6763, Mar. 2014. DOI: 10.1364/oe.22.006749 (cited on pages 30, 31).

[116] A. D. Forno, A. Paradisi, R. Passy, and J. von der Weid, "Experimental and theoretical modeling of polarization-mode dispersion in single-mode fibers", *IEEE Photonics Technol. Lett.*, vol. 12, no. 3, pp. 296–298, Mar. 2000. DOI: 10.1109/68.826919 (cited on page 34).

[117] K.-P. Ho and J. M. Kahn, "Mode coupling and its impact on spatially multiplexed systems", in *Optical Fiber Telecommunications*, I. Kaminow, T. Li, and A. Willner, Eds. Oxford, UK: Academic Press, 2013, vol. VIB, ch. 11, pp. 491–568. DOI: 10.1016/b978-0-12-396960-6.00011-0 (cited on pages 34, 35).

[118] P. Wai and C. Menyak, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence", *J. Lightw. Technol.*, vol. 14, no. 2, pp. 148–157, Feb. 1996. DOI: 10.1109/50.482256 (cited on page 35).

[119] W. Givens, "Computation of plain unitary rotations transforming a general matrix to triangular form", *J. Soc. Ind. Appl. Math.*, vol. 6, no. 1, pp. 26–50, Mar. 1958. DOI: 10.1137/0106004 (cited on page 37).

[120] E. Ip and J. M. Kahn, "Power spectra of return-to-zero optical signals", *J. Lightw. Technol.*, vol. 24, no. 3, pp. 1610–1618, Mar. 2006. DOI: 10.1109/JLT.2005.863328 (cited on page 38).

[121] C. Rasmussen, Y. Pan, M. Aydinlik, M. Crowley, J. Geyer, P. Humblet, F. Liu, B. Mikkelsen, P. Monsen, N. Nadarajah, G. Pendock, B. Shah, and B. Shah, "Real-time dsp for 100+Gb/s", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Anaheim, CA, USA, Mar. 2013, paper OW1E.1. DOI: 10.1364/OFC.2013.OW1E.1 (cited on pages 49, 77).

[122] Y. Yue, Q. Wang, and J. Anderson, "Transmitter skew tolerance and spectral efficiency tradeoff in high baud-rate QAM optical communication systems", *Opt. Express*, vol. 26, no. 12, pp. 15 045–15 058, May 2018. DOI: 10.1364/oe.26.015045 (cited on page 49).

[123] ——, "Coherent transmitter skew limitation on spectrally efficient optical communication systems", in *SPIE Int. Conf. Opt. Comm. Netw. (ICOCN)*, Zhuhai, China, Feb. 2019. DOI: 10.1117/12.2520481 (cited on page 49).

[124] G. Bosco, S. M. Bilal, A. Nespola, P. Poggiolini, and F. Forghieri, "Impact of the transmitter IQ-skew in multi-subcarrier coherent optical systems", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Anaheim, CA, USA, Mar. 2016, paper W4A.5. DOI: 10.1364/ofc.2016.w4a.5 (cited on page 50).

[125] E. P. Silva and D. Zibar, "Widely linear blind adaptive equalization for transmitter iq-imbalance/skew compensation in multicarrier systems", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Düsseldorf, Germany, Sep. 2016, pp. 25–27 (cited on page 50).

[126] E. P. da Silva and D. Zibar, "Widely linear equalization for iq imbalance and skew compensation in optical coherent receivers", *J. Lightw. Technol.*, vol. 34, no. 15, pp. 3577–3586, Aug. 2016, ISSN: 0733-8724. DOI: 10.1109/JLT.2016.2577716 (cited on pages 50, 77–79).

[127] C. R. S. Fludger and T. Kupfer, "Transmitter impairment mitigation and monitoring for high baud-rate, high order modulation systems", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Düsseldorf, Germany, Sep. 2016, pp. 256–258 (cited on page 50).

[128] N. Stojanović and X. Chuan, "Clock recovery in coherent optical receivers", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, CA, USA, Mar. 2015, paper Th3G.4. DOI: 10.1364/OFC.2015.Th3G.4 (cited on pages 50, 52, 107).

[129] Y. Yue, B. Zhang, Q. Wang, R. Lofland, J. O'Neil, and J. Anderson, "Detection and alignment of dual-polarization optical quadrature amplitude transmitter IQ and XY skews using reconfigurable interference", *Opt. Express*, vol. 24, no. 6, pp. 6719–6734, Mar. 2016. DOI: 10.1364/oe.24.006719 (cited on pages 50, 52, 72).

[130] C. R. S. Fludger, T. Duthel, P. Hermann, and T. Kupfer, "Low cost transmitter self-calibration of time delay and frequency response for high baud-rate QAM transceivers", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, CA, USA, Mar. 2016, paper Th1D.3. DOI: 10.1364/ofc.2017.th1d.3 (cited on pages 50, 72).

[131] L. H. H. Carvalho, C. Franciscangelis, U. R. Duarte, V. N. Rozental, J. D. Reis, F. B. Fideles, G. J. Suzigan, F. D. Simoes, V. E. Parahyba, N. G. Gonzalez, A. C. Bordonalli, and J. C. R. F. Oliveira, "Multidimensional optimization of optical spectral shaping for fiber nonlinearities mitigation in high baud-rate systems", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Cannes, France, Sep. 2014, paper P.5.5. DOI: 10.1109/ecoc.2014.6963915 (cited on page 61).

[132] C. de A. S. Diniz, M. Garrich, G. J. Suzigan, J. S. Assine, J. D. Reis, J. R. F. de Oliveira, and D. A. A. Mello, "Embedded system for optical spectral optimization based on a genetic algorithm", in *Proc. SBMO/IEEE MTT-S Int. Microw. Optoelectron. Conf. (IMOC)*, Porto de Galinhas, Brazil, Nov. 2015, pp. 1–4. DOI: 10.1109/imoc.2015.7369050 (cited on page 61).

[133] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading: Addison-Wesley, 1989, ISBN: 0201157675 (cited on page 61).

[134] O. Kramer, *Genetic Algorithm Essentials*. Heidelberg: Springer, 2017, ch. 2. DOI: 10.1007/978-3-319-52156-5 (cited on pages 61, 62).

[135] M. A. Potter and K. A. D. Jong, "Cooperative coevolution: An architecture for evolving coadapted subcomponents", *Evol. Comput.*, vol. 8, no. 1, pp. 1–29, Mar. 2000. DOI: 10.1162/106365600568086 (cited on page 62).

[136] T. Tanimura, S. Oda, T. Tanaka, T. Hoshida, Z. Tao, and J. C. Rasmussen, "A simple digital skew compensator for coherent receiver", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Vienna, Austria, Sep. 2009, paper 7.3.2 (cited on page 77).

[137] R. A. Soriano, F. N. Hauske, N. G. Gonzalez, Z. Zhang, Y. Ye, and I. T. Monroy, "Chromatic dispersion estimation in digital coherent receivers", *J. Lightw. Technol.*, vol. 29, no. 11, pp. 1627–1637, Jun. 2011. DOI: 10.1109/jlt.2011.2145357 (cited on pages 77, 79).

[138] D. Wang, C. Lu, A. P. T. Lau, and S. He, "Adaptive chromatic dispersion compensation for coherent communication systems using delay-tap sampling technique", *IEEE Photonics Technol. Lett.*, vol. 23, no. 14, pp. 1016–1018, Jul. 2011. DOI: 10.1109/lpt.2011.2151280 (cited on pages 77, 79, 81).

[139] R. Borkowski, X. Zhang, D. Zibar, R. Younce, and I. T. Monroy, "Experimental demonstration of adaptive digital monitoring and compensation of chromatic dispersion for coherent DP-QPSK receiver", *Opt. Express*, vol. 19, no. 26, B728, Dec. 2011. DOI: 10.1364/oe.19.00b728 (cited on pages 77, 79, 81).

[140] J. C. Diniz, S. Ranzini, V. Ribeiro, E. Magalhães, E. Rosa, V. Parahyba, L. V. Franz, E. E. Ferreira, and J. Oliveira, "Hardware-efficient chromatic dispersion estimator based on parallel gardner timing error detector", in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, CA, USA, Mar. 2012. DOI: 10.1364/ofc.2013.oth3c.6 (cited on pages 77, 79, 81).

[141] F. N. Hauske, N. Stojanovic, C. Xie, and M. Chen, "Impact of optical channel distortions to digital timing recovery in digital coherent transmission systems", in *Proc. Int. Conf. Transparent Opt. Netw.*, Munich, Germany, Jun. 2010, paper We.D1.4. DOI: 10.1109/icton.2010.5549082 (cited on page 79).

[142] E. Ip and J. M. Kahn, "Digital equalization of chromatic dispersion and polarization mode dispersion", *J. Lightw. Technol.*, vol. 25, no. 8, pp. 2033–2043, Aug. 2007. DOI: 10.1109/jlt.2007.900889 (cited on page 85).

[143] I. Fatadin, D. Ives, and S. J. Savory, "Laser linewidth tolerance for 16-QAM coherent optical systems using QPSK partitioning", *IEEE Photonics Technol. Lett.*, vol. 22, no. 9, pp. 631–633, May 2010. DOI: 10.1109/lpt.2010.2043524 (cited on page 85).

[144] X. Zhou, C. Lu, A. P. T. Lau, and K. Long, "Low-complexity carrier phase recovery for square M-QAM based on S-BPS algorithm", *IEEE Photonics Technol. Lett.*, vol. 26, no. 18, pp. 1863–1866, Sep. 2014. DOI: 10.1109/lpt.2014.2332358 (cited on page 85).

[145] J. C. M. Diniz, Q. Fan, S. M. Ranzini, F. N. Khan, F. Da Ros, D. Zibar, and A. P. T. Lau, *Principal component based technique for carrier phase recovery of square M-QAM systems*, to be submitted (cited on pages 85, 90, 91, 93, 96, 97, 102, 106).

[146] S. Hoffmann, C. Wordehoff, A. Al-Bermani, M. El-Darawy, K. Puntsri, U. Ruckert, and R. Noe, "Hardware-efficient phase estimation for digital coherent transmission with star constellation QAM", *IEEE Photon. J.*, vol. 2, no. 2, pp. 174–180, Apr. 2010. DOI: 10.1109/jphot.2010.2044784 (cited on page 86).

[147] H. Abdi and L. J. Williams, "Principal component analysis", "*Wiley Interdiscip. Rev. Comput. Stat.*, vol. 2, no. 4, pp. 433–459, Jun. 2010. DOI: 10.1002/wics.101 (cited on page 89).

[148] R. V. Mises and H. Pollaczek-Geiringer, "Praktische verfahren der gleichungsauflösung .", *ZAMM - Zeitschrift für Angewandte Mathematik und Mechanik*, vol. 9, no. 2, pp. 152–164, 1929. DOI: 10.1002/zamm.19290090206. [Online]. Available: https://doi.org/10.1002/zamm.19290090206 (cited on page 92).

[149] A. Alvarado, E. Agrell, D. Lavery, R. Maher, and P. Bayvel, "Replacing the soft-decision fec limit paradigm in the design of optical communication systems", *J. Lightw. Technol.*, vol. 33, no. 20, pp. 4338–4352, Oct. 2015, ISSN: 0733-8724. DOI: 10.1109/JLT.2015.2450537 (cited on page 100).

[150] E. Agrell and M. Secondini, "Information-theoretic tools for optical communications engineers", in *Proc. Int. Phot. Conf.*, Reston, VA, USA, Sep. 2018, paper MA3.1 (cited on pages 100, 125, 126).

[151] A. Alvarado, D. J. Ives, S. J. Savory, and P. Bayvel, "On the impact of optimal modulation and FEC overhead on future optical networks", *J. Lightw. Technol.*, vol. 34, no. 9, pp. 2339–2352, May 2016 (cited on page 104).

[152] D. A. A. Mello, A. L. N. Souza, J. D. Reis, J. C. M. Diniz, L. H. H. Carvalho, N. G. Gonzalez, J. R. F. Oliveira, D. S. Arantes, and M. H. M. Costa, "Parameter selection in optical networks with variable-code-rate transceivers", in *Int. Conf. Opt. Netw. Design Modeling (ONDM)*, May 2015, pp. 51–56. DOI: 10.1109/ONDM.2015.7127273 (cited on page 104).

[153] G. Colavolpe, T. Foggi, E. Forestieri, and M. Secondini, "Impact of phase noise and compensation techniques in coherent optical systems", *J. Lightw. Technol.*, vol. 29, no. 18, pp. 2790–2800, Sep. 2011. DOI: 10.1109/jlt.2011.2164237 (cited on page 108).

[154] C. B. Czegledi, G. Liga, D. Lavery, M. Karlsson, E. Agrell, S. J. Savory, and P. Bayvel, "Digital backpropagation accounting for polarization-mode dispersion", *Opt. Express*, vol. 25, no. 3, p. 1903, Jan. 2017. DOI: 10.1364/oe.25.001903 (cited on page 108).

[155] J. Renner, T. Fehenberger, M. P. Yankov, F. D. Ros, S. Forchhammer, G. Bocherer, and N. Hanik, "Experimental comparison of probabilistic shaping methods for unrepeated fiber transmission", *J. Lightw. Technol.*, vol. 35, no. 22, pp. 4871–4879, Nov. 2017. DOI: 10.1109/jlt.2017.2752243 (cited on page 110).

[156] R. T. Jones, T. A. Eriksson, M. P. Yankov, and D. Zibar, "Deep learning of geometric constellation shaping including fiber nonlinearities", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Rome, Italy, Sep. 2018, pp. 1–3. DOI: 10.1109/ECOC.2018.8535453 (cited on page 110).

[157] D. A. A. Mello, F. A. Barbosa, and J. D. Reis, "Interplay of probabilistic shaping and the blind phase search algorithm", *J. Lightw. Technol.*, pp. 1–1, 2018. DOI: 10.1109/jlt.2018.2869245 (cited on page 110).

[158] D. Arnold, H.-A. Loeliger, P. Vontobel, A. Kavcic, and W. Zeng, "Simulation-based computation of information rates for channels with memory", *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3498–3508, Aug. 2006. DOI: 10.1109/tit.2006.878110 (cited on pages 125, 126).