

Optical Communications for Long-Haul, Short-Reach, and Chip-Scale Distances Enabled by Digital Signal Processing

Mathieu Chagnon



Department of Electrical & Computer Engineering
McGill University
Montréal, Québec, Canada
December 2015

A thesis submitted to McGill University
in partial fulfilment of the requirements of the degree of
Doctor of Philosophy

© Mathieu Chagnon, 2015

À Geneviève, Sylvie et Pierre: Je ne pourrai jamais assez vous remercier.

À Mina, mon Amour.

No great mind has ever existed without a touch of madness

–Aristotle

Abstract

This thesis covers three topics of optical communications. First, we address long-haul fiber-optic coherent systems and propose a solution to two specific problems: polarization recovery and fiber nonlinearity. Second, we study silicon photonic devices for passive all-optical signal processing and active optical signal modulation enabling dense integration of photonics with advanced microelectronics. Finally, we address the need for faster short-reach transceivers in intra- and inter-data center networks by proposing novel modulation formats, digital signal processing algorithms, and transceiver architectures for Stokes vector modulation and demodulation. The thesis covers a large range of optical transmission distances, from long-haul to chip-scale photonics including short-reach optics for data center networks, all enabled by digital signal processing at both the transmitter and receiver.

In the case of long-haul fiber-optic coherent systems, we propose a blind and fast algorithm allowing recovering the state of polarization of any Dual-Polarization square M-ary Quadrature Amplitude Modulation (DP-MQAM) format. The algorithm allows a significant reduction of the convergence time for any polarization orientation at the receiver compared to the traditional blind derotation method. We also propose a complex modulation format being a power constrained version of the Dual-Polarization-8QAM format and we analytically and experimentally demonstrate greater tolerance to nonlinear effects.

For chip-scale photonic devices, the thesis contains a thorough study of wavelength multicasting using the parametric process of four-wave-mixing in a silicon-on-insulator photonic waveguide where we demonstrate a 1-to-6 multicast of a 16QAM signal and study the impact on performance of the input and output fiber-chip couplers. We also characterize the performance of the first silicon Mach-Zehnder intensity modulator enabling 112 Gb/s using 8-level intensity modulation for CMOS integrated interconnects.

Finally, for optical communications targeting data center applications covering the range of 0.5 to 10 km, we present novel transceivers, modulation formats, and digital signal processing schemes for next generation single wavelength high-speed pluggables. We demonstrate the first

transmission of 112 Gb/s over 10 km of single mode fiber in a silicon photonics modulator using 4-level pulse amplitude modulation (PAM), compliant with IEEE[®]'s new physical specifications of 56 Gbaud signaling and PAM-4 format. The thesis also reports the first demonstration of a 300 Gb/s transceiver enabled by digital to analog and analog to digital converters and novel digital signal processing algorithms where we modulate 6 bits of information per symbol on a single laser source and demodulate with a direct detection Stokes-vector receiver.

We conclude this thesis with a summary of the research presented and possible future research to follow those addressed in this work.

Abrégé

Cette thèse couvre trois aspects des communications optiques. Premièrement, nous abordons les systèmes cohérent de fibre optique à longue portée et nous proposons une solution à deux problèmes spécifiques: le recouvrement de la polarisation au receveur et l'effet nonlinéaire dans une fibre optique. Deuxièmement, nous étudions les dispositifs fabriqués en silicium sur isolant pour le traitement passif de signaux optiques et la modulation active de signaux optiques permettant une intégration compacte de dispositifs photoniques avec des circuits micro-electronique avancés. Finalement, nous abordons le besoin grandissant de transcepteurs plus rapide dans les réseaux interne des centres de données et les réseaux entre centres de données. Nous proposons de nouveaux formats de modulation, des algorithmes de traitement de signaux numériques et des architectures de transcepteurs pour modulation et demodulation en vecteurs Stokes. Cette thèse couvre un large éventail de distances de transmission optique, des systèmes cohérent à longue portée aux dispositif photonique sur puces, en passant par les optiques de courte portée pour les réseaux de centres de données, tous supportés par du traitement de signal numérique à l'émetteur et au récepteur.

Pour les systèmes à fibre optique de longue portée, nous proposons un algorithme aveugle et rapide permettant de recouvrir l'état de polarisation de n'importe quel format carré de modulation en phase et en amplitude multiplexé sur deux polarisations (DP-MQAM). L'algorithme permet une réduction significative du temps de convergence peu importe l'orientation de la polarisation au receveur, lorsque comparé à la méthode traditionnelle de dérotation aveugle de polarisation. Nous proposons également un format de modulation complexe étant une version contrainte en puissance du format de modulation DP-8QAM. Nous démontrons analytiquement et expérimentalement une tolérance accrue aux effets nonlinéaire de la fibre optique.

En ce qui a trait aux dispositifs photonique sur puces, cette thèse contient une étude approfondie de la pluri-diffusion d'une longueur d'onde dans un guide d'onde photonique en silicium sur isolant à l'aide du processus parametric qu'est le mélange à quatre ondes. Nous démontrons une pluri-diffusion de la longueur d'onde de une à six copies, où cette dernière est modulée en format 16QAM. Nous étudions également l'impact sur la performance du couplage d'entrée et de sortie de

la puce. Ensuite, nous caractérisons la performance du premier modulateur Mach-Zehnder sur silicium permettant la transmission de 112 Gb/s pour les prochaines générations de systèmes modulés en intensité avec détection directe.

Finalement, dans la catégorie des communications optiques pour centres de données couvrant une distance entre 0.5 et 10 km, nous présentons de nouveaux transcepteurs, de nouveaux formats de modulations et de nouveaux traitements de signaux numériques pour les prochaines générations de ‘connectables’ à haut débit. Nous démontrons la première transmission à 112 Gb/s sur 10 km de fibre à mode unique utilisant une modulation en amplitude à 4 niveaux à l’aide d’un modulateur Mach-Zehnder en photonique sur silicium. En dernier lieu, cette thèse expose la première démonstration d’un transcepteur à 300 Gb/s propulsé par des algorithmes de traitement de signaux numériques et des convertisseurs numériques-analogiques et analogiques-numériques. Nous modulons 6 bits d’information par symbol utilisant un seul laser source et démodulons à l’aide d’un receveur Stokes en détection directe.

Nous terminons cette thèse par un résumé des travaux présentés et une ouverture sur de futures recherches pour poursuivre celles abordées dans cet ouvrage.

Acknowledgements / Remerciements

First and foremost, I would like to thank my advisor Prof. David V. Plant for everything he gave me in this long adventure that spanned my Master's and Ph.D. in the Photonic Systems Group. David, your social, communication, persuasion, and self-management skills are impressive and exemplary. Every meeting with you turns out to be a highly motivational session. You help me think and organize my thoughts. I immensely learned and matured working with you. You are a striking visionary that provided me with boundless opportunities to propel me as a successful scientist, helping me in conducting sharp, relevant research. You gave me all the tools and counseling to succeed. And most importantly, you believed in me. You are light years away from 'just' being my advisor: you are a model, an inspiration, a mentor. I am extremely privileged to have worked with you for all those years, and I sincerely wish our paths and relationship remain close.

I would also like to thank some colleagues. *Mohamed Osman*, I wouldn't be what I am today if it was not of you. No words can express what I have learned *from* you and *with* you. We formed an inexplicable research duo throughout our Ph.D.'s. Together, we brought back to light this old saying from Aristotle: "*The whole is greater than the sum of its parts.*" In you I discovered a bright scientist, a colleague, a lab mate and a new culture. But most importantly, a discovered a friend. Your reasoning, open-mindedness, dedication and kindness are truly unique and admirable. It was an honor and a pleasure to spend this time with you, and may our paths not wander too far apart.

I cannot skip other colleagues that evolved with me in this long sojourn: Dr. *Qunbi Zhuge*, you are a fast learner and hard worker. It was very impressive watching you evolve in the group. I hope our paths will reunite one day in this small world of research in optical communications. Dr. *Xian Xu*, it was a pleasure working with you. Dr. *Jon Buset*, it's always a pleasure chatting with you, and thanks for helping expediting the writing of this thesis! To the new Crew, *David Patel*, *Alireza Samani*, and others, know that I am always amazed by how bright you guys are. Keep up the good work and I hope we stay connected. To other colleagues in Plant's Group: this was an unforgettable experience! À *Sandrine*: Merci pour tout ce placotage! Du divertissement en boîte!

I would also like to thank specific members of Nortel Networks®/Ciena®. Dr. *Maurice O'Sullivan* for the internship you offered me between my Master's and Ph.D. This sojourn in your team greatly propelled my research skills and changed the course of my Doctoral studies. I am immensely grateful and I sincerely thank you. Dr. *Michael Reimer*, who is simply the brightest man I know! I have to thank you Mike for everything I've learned from you, everything you thought me in such a short period. Your expertise in Mathematics and Fiber Optic Communications are simply astonishing. Dr. *Charles Laperle* and Dr. *Andrzej Borowiec*, I cannot express how much I learned from you in the lab. Your experimental expertise and knowledge are simply a gold mine!

J'aimerais également remercier ma famille proche: Geneviève, Sylvie et Pierre. Vous avez été et serez toujours un catalyseur dans ma vie, un element essentiel à mon bonheur. Vous avez toujours été là pour me supporter et m'encourager. Je vous aime, inconditionnellement! Finalement, j'aimerais terminer en remerciant mes ami(e)s. Vous êtes ma balance, ma soupape de décompression. Vous êtes rires et bonheurs, vous m'avez fait grandir, vous m'avez fait apprendre.

Sincèrement,

Merci

Mat Chagnon

Statement of Originality

The work described in this thesis represents original research by the author, and has not been submitted for a degree or diploma at any other institution. This thesis contains no material previously written or published by any other persons excepted where references are made. As evidence of my original research and contributions to the existing body of knowledge, my Ph.D. work has currently resulted in 8 first-authored peer-reviewed publications, comprising six journal articles [1–6] and two refereed conference proceedings [7, 8]. We note that two of these journals are invited journals [5, 6] and one conference proceeding is a Post-Deadline Paper [7]. A summary of my contributions is as follows:

Formats and DSP for Coherent Optical Transports

- We present both theoretically and experimentally a novel blind and fast method for estimating the State of Polarization (SOP) of a single carrier channel modulated in square Dual Polarization M -QAM format using optical coherent receivers [1]. The method consists of converting the received waveform from Jones to Stokes space and looping over an algorithm until a unitary polarization derotation matrix is estimated. The method is compared against the common blind Constant Modulus Algorithm (CMA) for polarization recovery (PR). We show that the convergence time of this blind PR technique does not depend on the initial SOP as CMA does and allows switching faster to the adaptive stochastic gradient descent filter, by more than an order of magnitude.
- We propose a new modulation format providing 5 bits of information per recovered symbol while maintaining constant the total optical power [2]. The proposed format applies a simple power constraint to the DP-8QAM format. This modulation format provides a passive way to mitigate nonlinear impairments occurring during propagation that are due to the Kerr effect, and most specifically in the first 40 km. The constant power property of the symbol set of the format mitigates the nonlinear effects of self- and cross-phase modulation. We experimentally

validate over a long-haul transmission system in a WDM scenario that the format outperforms its parent DP-8QAM format.

Passive and Active Silicon Photonics

- Using a passive silicon strip waveguide, we present the first high symbol rate, high order QAM 1-to-6 wavelength multicasting of a 22-GBaud 16-QAM single polarization signal in a 20 mm long passive silicon nanowire based on four-wave mixing [3]. The strip waveguide is designed to exhibit a dispersion slope near zero around 1550 nm to minimize the propagation constant difference between the four interplaying waves. The wavelength multicasting performance is compared against the theoretical power transfer during a three-wave mixing process and we show that the average power of degenerate idlers agrees with the predicted 6 dB difference with respect to that of nondegenerate idlers. Finally, we also demonstrate the importance of the output grating coupler's loss and its impacts on the multicasting performance.
- We demonstrate for the first time a 112 Gb/s transmission over 10 km of SMF fiber employing PAM-8 at 37.4 Gbaud using a Silicon Photonic modulator operating on the cost-efficient intensity modulation/direct detection principle [4]. We present the analysis and the performance of an active Silicon Photonic intensity modulator operating at 1.3 μm . The modulator is driven by a digital to analog converter (DAC) and we present the entire DSP stack allowing to maximize the performance of the direct detection system. We also derive the first analytical model for the Q -factor metric applicable for multilevel PAM- N signaling of differing noise power per level that we accurately experimentally verify and validate. We demonstrate the ability of the proposed DAC-powered transmitter to exhibit software-defined transmission for short reach applications by selecting the PAM order, symbol rate and pulse shaping filter.

Optical Communications Targeting Data Center Applications

- We improve the BER performance of the previous work by better equalizing the transmitted waveform, allowing generating 112 Gb/s bit rate using a 4-level PAM format running at 56 Gbaud instead of an 8-level at 37.4 Gbaud [5]. This format and signaling complies with the IEEE 802.3bs standards for 400 GbE, unlike PAM-8. We also show that residual chromatic dispersion and polarization mode dispersion have no impact on the performance up to a transmission distance of 20 km. We study the performance impact of varying the number of transmitter pre-compensation taps and receiver equalization taps, having a direct impact on

system performance. We also describe and explain how the common DC bias voltage applied at the common cathode of the silicon modulator impacts the transmitted waveform. We show that a BER below the FEC limit is obtained for different combinations of transmitter and receiver tap counts, allowing flexibility in the distribution of power consumption.

- We present and validate a transceiver allowing modulating 6 bits/symbol on a single wavelength and employing a self-beating direct detection receiver [6,7]. Throughputs of up to 350 Gb/s are experimentally demonstrated, where the intensity on both polarizations and the inter-polarization phase difference are all modulated. Our 350 Gb/s results are the highest rate ever reported for a direct detection system applicable for data center connectivity using a 4-channel receiver. This work allows a significant increase of the spectral efficiency per wavelength by augmenting the dimensionality of the modulation format while maintaining a cost effective direct detection scheme, increasing the throughput delivered per color. We propose two digital signal processing schemes applied at the receiver to recover the 6 bits/symbol. The two DSP algorithms operate on a two-stage and three-stage stack, respectively. We compared both algorithms using two metrics: the computational complexity and the bit error rate performance. We show that the three stage approach allows significant reduction in the total number of real-multiplications and real-additions required, for the same resulting BER performance. Moreover, tracking the state of polarization using the three-stage proposed method is m -times more efficient, where m is the number of taps in the first filtering stage.

Associated Publications

More specifically, we list here the journal articles and conference proceedings related to this thesis, specifying the contributions of other co-authors.

- [1] **M. Chagnon**, M. Morsy-Osman, X. Xu, Q. Zhuge, and D. V. Plant, “Blind, fast and SOP independent polarization recovery for square dual polarization–M-QAM formats and optical coherent receivers,” *Opt. Express*, vol. 20, no. 25, pp. 27 847–27 865, Dec 2012.

I had the idea, derived the theory, performed the analysis and simulation, and wrote the paper. Exceptionally for this manuscript, I used experimental data collected offline from two colleagues. The other authors contributed in editing the manuscript.

- [2] **M. Chagnon**, M. Morsy-Osman, Q. Zhuge, X. Xu, and D. V. Plant, “Analysis and experimental demonstration of novel 8PolSK-QPSK modulation at 5 bits/symbol for passive mitigation of nonlinear impairments,” *Opt. Express*, vol. 21, no. 25, pp. 30 204–30 220, Dec 2013.

I had the idea, derived the mathematical background and performed the analysis and simulations, collected all experimental data in the laboratory and wrote the paper. The other authors contributed in editing the manuscript.

- [3] **M. Chagnon**, M. Spasojevic, R. Adams, J. Li, D. V. Plant, and L. R. Chen, “Wavelength multicasting at 22-GBaud 16-QAM in a silicon nanowire using four-wave mixing,” *IEEE Photon. Technol. Lett.*, vol. 27, no. 8, pp. 860–863, April 2015.

I performed the experiment, derived the analysis of the data, and I wrote the paper. The other authors helped in performing the experiment and in editing the paper.

- [4] **M. Chagnon**, M. Morsy-Osman, M. Poulin, C. Latrasse, J.-F. Gagné, Y. Painchaud, C. Paquet, S. Lessard, and D. V. Plant, “Experimental study of 112 Gb/s short reach transmission employing PAM formats and SiP intensity modulator at 1.3 μm ,” *Opt. Express*, vol. 22, no. 17, pp. 21 018–21 036, Aug 2014.

I performed all experiments, collected and processed the data, derived all analysis and wrote the manuscript. Authors M. Poulin, C. Latrasse, J.-F. Gagné, Y. Painchaud and C. Paquet from TeraXion[®] provided the Silicon Photonic Mach-Zhender Modulator. The other authors contributed in the discussion and in editing the paper.

- [5] **M. Chagnon**, M. Morsy-Osman, M. Poulin, C. Paquet, S. Lessard, and D. V. Plant, “Experimental parametric study of a silicon photonic modulator enabled 112 Gb/s PAM transmission system with a DAC and ADC,” *J. Lightw. Technol.*, vol. 33, no. 7, pp. 1380–1387, April 2015.

Invited Paper. *I performed all experiments, derived the analysis, and wrote the manuscript. TeraXion[®] provided the SiP MZM. Other authors contributed in editing the manuscript.*

- [6] **M. Chagnon**, M. Morsy-Osman, D. Patel, V. Veerasubramanian, A. Samani, and D. V. Plant, “Digital signal processing for dual-polarization intensity and inter-polarization phase modulation formats using Stokes detection,” *J. Lightw. Technol.*, vol. 34, no. 4, pp. X–X, February 2016.

Invited Paper. *I had the idea, performed all experiments, derived all digital signal processing, performed the analysis and wrote the manuscript. M. Morsy-Osman greatly helped in the laboratory. Other authors contributed in discussing the idea and in editing the manuscript.*

- [7] **M. Chagnon**, M. Morsy-Osman, D. Patel, V. Veerasubramanian, A. Samani, and D. V. Plant, “1 λ , 6 bits/symbol, 280 and 350 Gb/s direct detection transceiver using intensity modulation, polarization multiplexing, and inter-polarization phase modulation,” in Proc. of *Optical Fiber Communications Conference*, March 2015, paper Th5B.2, pp. 1–3.

Post Deadline Paper. *I had the idea, performed all experiments, derived all DSP, performed the analysis and wrote the manuscript. M. Morsy-Osman greatly helped in the laboratory. Other authors contributed in discussing the idea and in editing the manuscript.*

One research publication that I conducted during this Ph.D. is not addressed in this thesis:

- [8] **M. Chagnon**, M. Morsy-Osman, Q. Zhuge, X. Xu, M. Poulin, Y. Painchaud, M. Pelletier, C. Paquet, and D. V. Plant, “Experimental colorless reception of $16\times\text{DP-16QAM}$ at 28 and 30 GBaud using a Si-integrated coherent receiver,” in Proc. of *Optical Fiber Communication Conference*, 2013, paper OTh3C.5, pp. 1–3.

I had the idea, performed the experiment and analysis, and wrote the paper. Other authors contributed in discussing the idea and editing the paper.

Table of Contents

	<u>page</u>
Abstract	iii
Abrégé	v
Acknowledgements / Remerciements	vii
Statement of Originality	ix
List of Tables	xvi
List of Figures	xvii
List of Abbreviations	xx
1 Introduction	1
1.1 Motivation	1
1.1.1 Capacity Needs in Long-Haul Optical Transport Networks	1
1.1.2 Capacity Needs in Intra- and Inter- Data Center Interconnects	2
1.1.3 Integrating Photonics with Electronics	3
1.1.4 Ways to Achieve Needs in Long-Haul and Short-Reach	5
1.2 Review of the State of the Art	6
1.2.1 Polarization Recovery and Fiber Nonlinearity Mitigation	6
1.2.2 Passive and Active Silicon Photonic Devices	8
1.2.3 Optical Short Reach Transceivers	9
1.3 Research Objectives and Thesis Overview	11
2 Optical Communications: From Long-Haul Coherent to Chip-Scale Photonics	15
2.1 Overview	15
2.2 Coherent Optical Communications	17
2.3 The Rise of Silicon Photonics	19
2.4 Connectivity in Data Centers	20
3 DSP in Coherent Long Haul: Polarization Recovery and a 32-Symbol Constant Power Format	22
3.1 Stokes Based Polarization Recovery	22
3.1.1 Introduction	22
3.1.2 DP-MQAM Formats in Stokes Space and Modeling of SOP in Lightwave Systems	24
3.1.2.1 Modeling of SOP in Lightwave Systems	26
3.1.3 New Blind Polarization Rotation Estimation	28
3.1.4 Comparison Metric and Experimental Test Bed	34
3.1.4.1 Comparison Metric	35
3.1.4.2 Experimental Test Bed	36

3.1.5	Results	37
3.1.5.1	Testing Over PMD	40
3.1.6	Conclusion	44
3.2	Constant Power 4D Modulation on a 32-Symbol Set	46
3.2.1	Introduction	46
3.2.2	Proposed 8PolSK-QPSK Modulation Format	48
3.2.2.1	Minimum Distance Detection for 8PolSK-QPSK	50
3.2.3	Nonlinearity Mitigation by Modulating with Constant Power	52
3.2.3.1	Nonlinear Variance Study for 8PolSK-QPSK and DP-8QAM	53
3.2.3.2	Variance of SPM, XPM and XPolM at Launch	57
3.2.4	Experimental Test Bed	60
3.2.4.1	8PolSK-QPSK Signal Generation with a Dual Polarization Emulator	61
3.2.5	Results	62
3.2.6	Conclusion	67
4	Passive and Active Silicon Photonics: Parametric Wave Mixing and Intensity Modulation	68
4.1	Four-Wave Mixing in a Silicon Nanowire	68
4.1.1	Introduction	68
4.1.2	State of the Art	69
4.1.3	Principle of Operation of FWM-based WC and WM	70
4.1.4	Power Transfer in FWM-based Wavelength Multicasting	71
4.1.5	Experimental Setup	72
4.1.6	Results and Discussion	73
4.1.7	Conclusion	76
4.2	Silicon Photonics Intensity Modulator	77
4.2.1	Introduction	77
4.2.2	Fully Packaged SiP Modulator at 1310 nm	78
4.2.3	Experimental Setup	82
4.2.4	Digital Signal Processing at the Transmitter and Receiver	84
4.2.4.1	Transmitter DSP for PAM- N Formats	84
4.2.4.2	Receiver DSP for PAM- N Formats	85
4.2.5	Performance and Results	86
4.2.5.1	Q -factor for PAM- N Formats	86
4.2.5.2	System Performance	88
4.2.6	Conclusion	100
5	Optical Short Reach Transceivers: From Unidimensional to 3D Modulation	101
5.1	Biasing, Pre-Emphasis and Post-Equalization of DAC/ADC Enabled SiP MZM	101
5.1.1	Introduction	102
5.1.2	Further Details on the SiP Modulator	103
5.1.3	Impact of Tx Parameters	104
5.1.4	System Description	106
5.1.5	Experimental Results	108
5.1.5.1	Performance Impact of Modulator Bias Voltage	108
5.1.5.2	Performance Impact from Pre- and Post-Equalizer Lengths	111
5.1.6	Conclusion	115
5.2	Three Dimensional Modulation for Self-Beating Direct Detection Systems	116
5.2.1	Introduction	116
5.2.2	Format Description and Transceiver for Higher Order Modulation, Stokes-Vector Detection	118

5.2.3	Intersymbol Interference Impairments	121
5.2.4	DSP, Filters and Updates	123
5.2.5	Complexity of the Algorithm	125
5.2.6	Experimental Setup and Results	128
5.2.7	Conclusion	131
6	Conclusion	132
6.1	Overview	132
6.2	Summary of Original Contribution	132
6.2.1	DSP for Coherent Transceivers	132
6.2.2	DSP Enhanced Chip-Scale Photonics	133
6.2.3	DAC/ADC Enabled Data Center Interconnects	134
6.3	Future Research	135
	References	138

List of Tables

<u>Table</u>	<u>page</u>
4-1 Power and OSNR in 0.1 nm at points “B” and “C”, with conversion efficiency, required power and power penalty of each channel	74
5-1 Computational Complexity Comparison	127

List of Figures

<u>Figure</u>	<u>page</u>
1-1 Global growth of IP Traffic by Devices: From 2014 to 2019	2
1-2 Projected growth of total data center traffic	3
2-1 Global Submarine Cable Network.	17
2-2 a) DP-IQ transmitter and b) Dual-Polarization Coherent Receiver	18
2-3 Optical network topology	20
3-1 Theoretical constellations of DP-4QAM, DP-16QAM and DP-64QAM in Stokes space.	25
3-2 Visual Inspection of Blind Polarization Derotation Algorithm	30
3-3 Stokes Constellations DP-4, -16, and -64 QAM with noise loading and power screening	31
3-4 Stokes space convergence location of CMA, RDE and MMA algorithm for DP-16QAM	35
3-5 Experimental Test bed	37
3-6 BER vs Distance of first 20K symbols for DP-4QAM format at 28 Gbaud	38
3-7 BER vs Distance of first 20K symbols for DP-16QAM format at 28 Gbaud	39
3-8 BER vs Distance of first 20K symbols for DP-64QAM format at 7 Gbaud	40
3-9 System performance for varying DGDs and Common Group Delays	42
3-10 Required number of symbols to find Urot over varying DGD	43
3-11 Star-8QAM format	49
3-12 Stokes space representation of the 8PolSK-QPSK and DP-8QAM format	51
3-13 SPM, XPM and XPolM variances over distance	55
3-14 Experimental Test Bed to compare 8PolSK-QPSK vs DP-8QAM	60
3-15 Maximum Reach for BER=1.4%, for 8PolSK-QPSK and DP-8QAM	63
3-16 BER versus OSNR in 0.1 nm for 8PolSK-QPSK and DP-8QAM	64
3-17 Required OSNR for BER=1.4% vs Launch Power: 8PolSK-QPSK and DP-8QAM . .	65
3-18 BER versus Distance: 8PolSK-QPSK and DP-8QAM	66
4-1 Principle of operation of Degenerate and non-Degenerate FWM	70
4-2 Experimental setup for Wavelength Multicasting, with cross-section of SNW	72

4-3	Optical spectrum after WM, before and after filtering and amplification	74
4-4	Sensitivity: Q -factor as a function of received power for the 7 cases	75
4-5	Schematic of the series push-pull Mach-Zehnder modulator	79
4-6	Photograph of the 6 mm long series push-pull modulator	79
4-7	Characterization of the SiP series push-pull modulator	81
4-8	EO frequency response of the chip and the packaged SiP modulator	82
4-9	Picture of the fully packaged SiP modulator	82
4-10	Experimental setup to test the packaged SiP modulator	83
4-11	DSP stack at the (a) transmitter and (b) receiver side	85
4-12	Eye diagrams after 10 km for PAM-2, -4, -8, and 16	89
4-13	Histogram and probability density functions of PAM-8 after 10 and 20 km	91
4-14	Histogram and probability density functions of PAM-16 after 10 km	91
4-15	Histogram and PDF of for signals of high Q , for PAM-2 and PAM-4	93
4-16	SNR, Q -factor and BER against varying bit rate, with BER vs Q -factor relation . .	94
4-17	BER and Q -factor for PAM-2, -4 and -8, for varying received signal power	96
4-18	System performance for varying PAM orders, symbol rates, and for roll-off factors . .	98
4-19	Q -factor and BER for PAM-4 and PAM-8, at 112 Gb/s, at varying distances.	100
5-1	EO frequency response of the modulator, for varying reverse biases	103
5-2	Tx DSP stack for DAC-enabled IM-DD transmission	106
5-3	Experimental Setup for DAC-enabled IM-DD transmission	107
5-4	Rx DSP stack for DAC-enabled IM-DD transmission	107
5-5	BER for PAM-4 and PAM-8 at 112 Gb/s, at varying distances and bias voltages . .	109
5-6	BER for a fixed received power after varying distances: PAM-4 and PAM-8	110
5-7	Eye diagrams of PAM-4 and PAM-8 in B2B	111
5-8	BER for PAM-4 at 112 Gb/s, for varying Tx and Rx number of taps	112
5-9	BER for PAM-8 at 112 Gb/s, for varying Tx and Rx number of taps	113
5-10	3D Stokes Vector Transmitter and Receiver architectures.	119
5-11	Visual representation of the three-stage DSP applied at the receiver.	123
5-12	First DSP stage of [7]. In this work, we decouple this stage into two stages.	125
5-13	Experimental setup: transmitter and receiver for the 6 bits/symbol generation and direct detection.	129

5-14 BER vs Bit rate for the 6 bits/symbol DP-PAM4+4PM format.	130
5-15 BER of the two DSP approaches over varying input SOPs.	130
6-1 Summary of the original contributions of the thesis.	133

List of Abbreviations

2D	Two-Dimensional
3D	Three-Dimensional
4D	Four Dimensional
ADC	Analog to Digital Converter
ASE	Amplified Spontaneous Emission
ASIC	Application Specific Integrated Circuit
ASK	Amplitude Shift Keying
AWG	Arbitrary Waveform Generator
AWGN	Additive White Gaussian Noise
B2B	Back-to-Back
BER	Bit Error Rate
BOA	Booster Optical Amplifier
BPF	Band-pass Filter
BPSK	Binary Phase Shift Keying
BW	Bandwidth
CAGR	Compound Annual Growth Rate
CAP	Carrierless Amplitude and Phase
CC	Computational Complexity
CD	Chromatic Dispersion
CMA	Constant Modulus Algorithm
CMOS	Complementary Metal-Oxide-Semiconductor
CW	Continuous Wave
DAC	Digital to Analog Converter
DC	Direct Current

DD	Direct Detection
DD-LMS	Decision-Directed Least-Mean Square
DFB	Distributed Feedback Laser
DFE	Decision Feedback Equalizer
DGD	Differential Group Delay
DML	Directly Modulated Laser
DMT	Discrete Multi-Tone
DOF	Degree Of Freedom
DP	Dual Polarization
DP-IQ	Dual-Polarization In-phase-Quad-phase
DPSK	Differential Phase Shift Keying
DSP	Digital Signal Processing
ECL	External Cavity Laser
EDFA	Erbium-Doped Fibre Amplifier
EML	Electroabsorption-Modulation Laser
ENOB	Effective Number Of Bits
FEC	Forward Error Correction
FFE	Feed-Forward Equalizer
FIR	Finite Impulse Response
FWHM	Full Width at Half Maximum
FWM	Four Wave Mixing
HD	Hard-Decision
HNLF	Highly Nonlinear Fiber
ICA	Independent Component Analysis
IM	Intensity Modulation
IM/DD	Intensity Modulation with Direct Detection
IP	Internet Protocol
IQ	In-phase and Quadrature

ISI	Intersymbol-Interference
LO	Local Oscillator
MIMO	Multiple-Input Multiple-Output
MMA	Multimodulus Algorithm
MMI	Multimode Interference
MZI	Mach-Zehnder Interferometer
MZM	Mach-Zehnder Modulator
NL	Non-Linear
NLS	Non-Linear Schrödinger
OH	Over-Head
OOK	On-Off Keying
OSNR	Optical Signal to Noise Ratio
PAM	Pulse Amplitude Modulation
PAPR	Peak to Average Power Ratio
PC	Polarization Controller
PCA	Principal Component Analysis
PDF	Probability Density Function
PDM	Polarization Division Multiplexing
PIN	P-doped, Intrinsic, N-doped semiconductor
PLL	Phase Locked Loop
PMS	Polarization Maintaining Splitter
PolSK	Polarization Shift Keying
PPLN	Periodically-Poled Lithium Niobate
PR	Polarization Recovery
PS	Polarization Shift
PSK	Phase Shift Keying
PSK	Phase-Shift Keying
QAM	Quadrature Amplitude Modulation

RA	Real Addition
RDE	Radius Directed Equalizer
RM	Real Multiplication
RMS	Root-Mean Square
ROSNR	Required Optical Signal power to Noise power Ratio
RRC	Root-Raised Cosine
Rx	Receiver
RZ	Return-to-Zero
SD	Soft-Decision
SED	Squared Euclidean Distance
Si-EPIC	Silicon Electronic-Photonic Integrated Circuits
Si-Ge	Silicon-Germanium
SiP	Silicon Photonics
SiP-CRx	Silicon Photonics Coherent Receiver
SNR	Signal to Noise Ratio
SNW	Silicon Nanowire
SOA	Semiconductor Optical Amplifier
SOI	Silicon on Insulator
SOP	State of Polarization
SP	Single Polarization
SPM	Self Phase Modulation
SPS	Sample per Symbol
TIA	Transimpedance Amplifier
TPA	Two-Photon Absorption
TPS	Thermal Phase Shifter
Tx	Transmitter
VOA	Variable Optical Attenuator
Vppd	Volts peak-to-peak differential

- WC** Wavelength Conversion
- WM** Wavelength Multicasting
- XPM** Cross-Phase Modulation
- XPolM** Cross-Polarization Modulation

CHAPTER 1

Introduction

1.1 Motivation

The capacity of today's communication networks needs to significantly increase in order to keep up with the growing demand of internet usage by user. This increase comes from the fast growing bandwidth intensive applications and the world-wide proliferation of smart phones and other internet connected hand-held devices. The main drivers for high capacity per-user are services such as video on demand, the explosion of social networks, cloud storage and cloud-based computing and services. The exploitation of new regional markets accentuates the internet protocol (IP) traffic and inflates demands.

1.1.1 Capacity Needs in Long-Haul Optical Tansport Networks

As one of the major providers of networking and communications technology, US-based company Cisco Systems[®] periodically gives forecasts on the development of world-wide IP-traffic to track and forecast global traffic and explore the implications of its growth for service providers through its Visual Networking Index [9]. As per the May 2015 report, annual global IP traffic will pass the zettabyte (10^{21} bytes) threshold by the end of 2016, and will reach 2 zettabytes per year by 2019. Global IP traffic has increased five-fold over the past five years, and will increase three-fold over the next five. Overall, IP traffic will grow at a compound annual growth rate (CAGR) $\left[\text{CAGR} = (\text{End Val.}/\text{Beg. Val.})^{1/(\# \text{ years})} - 1 \right]$ of 23% from 2014 to 2019. Content delivery networks such as Akamai[®] and EdgeCast[®] will carry nearly two-thirds of Internet traffic by 2019, up from 39% in 2014. Metro traffic surpassed long-haul traffic in 2014 and will account for 66% of total IP traffic by 2019. In fact, the higher growth in metro networks is due in part to the increasingly significant role of content delivery networks which bypass long-haul links and deliver traffic to metro and regional backbones via their local Points of Presence. Moreover, the number of devices connected to IP networks will be more than three times the global population by 2019.

As a consequence, mobile traffic is also growing at an impressive rate. By 2019, 67% of all IP traffic worldwide will originate from non-PC devices, consisting mainly of tablets, smartphones, and Machine-to-Machine modules. Service providers and large enterprises need to address the growing needs of ever-increasing demands for high-bandwidth data traffic in metro and data center networks. The following Fig. 1–1 is a visual representation of the global growth of IP traffic, in Exabytes per month, for the years 2014 to 2019, sorted by device type, showing a forecasted compound annual growth rate of 23%.

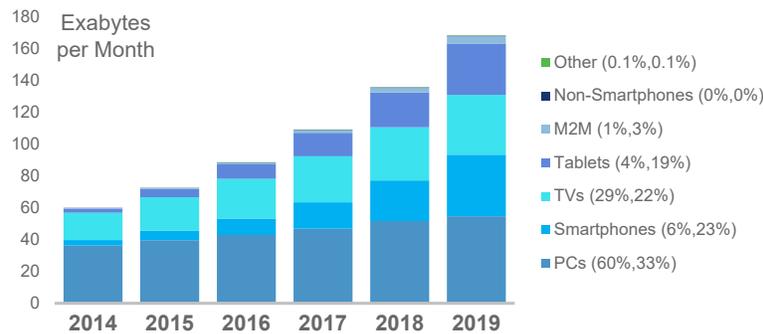


FIGURE 1–1: Global growth of IP Traffic by Devices: From 2014 to 2019.

1.1.2 Capacity Needs in Intra- and Inter- Data Center Interconnects

Data center have also observed their traffic booming in the past years and is no exception to the exponentially increasing bandwidth demand. Data Center traffic includes traffic within a Data Center, between Data Centers, and from a Data Center to users. To visually express the growth of Data Center traffic, we report Cisco[®]'s Global Cloud Index, which forecasts growths specifically for the global Data Center and cloud-based IP traffic. Figure 1–2 shows the total Data Center Traffic Growth forecast until 2019, with respective portions stemming from traditional Data Center traffic and traffic from Cloud-based services. There is a forecast annual growth of 33% of Cloud data center traffic. We see that Cloud-based services already dominate Data Centers traffic, and its portion will increase with time. Global data center traffic will grow nearly three-fold by 2019. Moreover, by 2019, global Data Center traffic will reach 8.6 zettabytes per year and 83% of all data center traffic will come from the cloud.

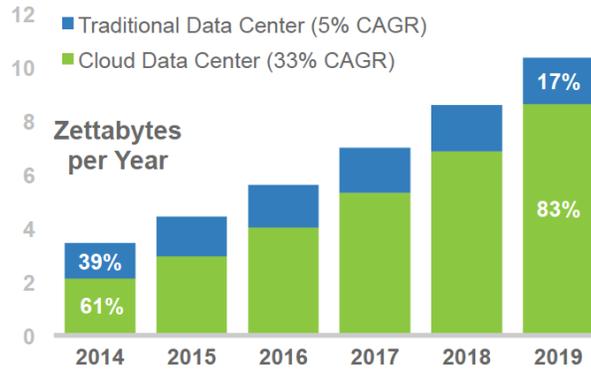


FIGURE 1–2: Projected growth of total data center traffic

Data Center transceivers need to move packets over much shorter distances than what long-haul optical modems are capable of. For Data Centers, reaches of up to ≈ 40 km are required, while transports over distances exceeding 10 000 km are achievable with optical long-haul modems [10].

1.1.3 Integrating Photonics with Electronics

One interesting question one may ask is: Why is silicon photonics so popular today? Silicon photonics is attracting attention not only from academia but it also receives huge interest from the industry. Billions of dollars were invested in the last decade on the development of silicon photonics. One of the major reason is the new massive demand for optics which is coming from the change in the industry of information technology and in our lives and behavior facing the new reality of our internet-based ecosystem, and in particular from the appearance of the ‘Cloud’. Almost all computation from consumer electronics like hand-held devices are directed to data centers. Enterprises are also shifting most of their computing powers outside, outsourcing it to data enters for cloud, distributed computing. Data centers are the backbone of our electronic society. From Cisco’s Global Cloud Index, the total amount of data supporting this Cloud today is 2.6 zettabytes, and it is expected to be growing to 7.7 zettabytes by 2017. However, the data transiting to and from a data center represents only a fraction of these numbers. In fact, only about 25% of the data center traffic leaves the data center. Almost 75% of all those zettabytes are transferred within the data center. In order to have efficient calculations or to provide the answer to a query from a search engine, it is not one or two processors that do the data cranking; its not even done on one rack. The query and process is distributed in a very parallelized system throughout the data center and

collected back. Therefore, communication within the data center is a bottleneck. In order to connect the distant racks together, with distances on the order of several kilometers with recent massive scale data centers, it is important to have very fast communication and large bandwidths. This represent an opportunity for optical communications because at these transfer rates and distances, electrical cables cannot provide the connectivity.

This represents a new reality for the opto-electronic industry. For a long time, the opto-electronic industry was mainly providing products for telecom-types of communication, over long distances. Volumes of production for such optical communication hardware is rather small and margins from equipment manufacturers is large, where each system is sold at around ~\$100k. The market for the emerging datacom is a different one. First, volumes are enormous, and are also growing with time. In a large data center, millions of part are required. Second, the price pressure is much higher. Because of the colossal number of optical transducers required per data center, employing existing products and technology of long-haul/telecom modems for data center optical transceivers would be prohibitive due to their expensive cost. The price per transducer in data centers has to oscillate around hundreds of dollars. The need for faster data shuffling of larger quantities within bigger data centers drives the need to develop new technologies. This is the main reason why there is such a big investment in silicon photonics.

The industry and academia are developing and exploring sets of additional processing modules that allow building structures like modulators, photodetectors and coupling structures right next to transistors in standard CMOS technology. The idea is to build a single die transceiver which consists of both optical components and electrical components integrated on the same chip. The advantage is obvious: it decreases the bill of material. Current transceivers which are based on traditional technology contain about 30-40 components from different vendors that are packaged together with more or less hand-labor involved. The cost of these components, the handling, and their packing constitute a big portion of the total cost of the whole solution. By integrating everything in silicon using the micro-electronics industry approach with mass production, we can significantly decrease the bill of material. Moreover, digital circuitry can also be integrated to augment communication capabilities and restore the signal integrity. Monotonically integrating optics with electronics in a single fabrication run increases yields and greatly reduces the cost per transceiver.

Finally, with photonics in CMOS technology, not only active components like modulators and photodetectors can be built on a die. Passive optical components can also be imprinted by

lithography to perform all-optical signal processing on chip and leverage the capabilities of the silicon photonics technology.

1.1.4 Ways to Achieve Needs in Long-Haul and Short-Reach

One way to meet the increasing traffic demand in long-haul fiber optic transmission systems is to upgrade or to expand the existing fiber network. However, installing new transoceanic fiber links or digging new trenches to upgrade the current optical backbone network comes with tremendous infrastructure costs and is therefore most likely the last option considered. Therefore the target is to increase the capacity while modifying as little as possible the optical infrastructure. Improvements thus have to be realized on the electrical interfaces at the beginning and end of the fiber link, on the modulation formats employed to carry information, and on the digital signal processing applied to generate the waveforms and to process the received waveforms.

The growing demand for faster optical transport is even greater over short-reach distances. Short reach optical modems, called Small Form-factor Pluggables, are employed for much shorter optical transport than that of coherent modems: between 100 m and ≈ 80 km. These pluggables need to be inexpensive. Short reach transceivers have relied on intensity modulation with direct detection (IM/DD) for decades. There are two ways to increase the bit rate delivered by such modems while maintaining the low cost of direct detection (DD): using advanced intensity modulation formats allowing the transfer of more than 1 bit per symbol, and using more dimensions of the optical lightwave to imprint information than the unidimensional single polarization intensity modulation currently deployed.

In recent years, we saw a clear trend to reduce the footprint and cost of transceivers for short reach optical interconnects. One way to reduce the cost and footprint while increasing the speed is by minimizing the number of optical components and increasing the bit rate per component. Said differently, increasing the data rate per wavelength. There has been significant work and progress recently by both research facilities and the industry demonstrating 100 Gb/s on a single wavelength for short reach optics while still employing IM/DD. Moreover, the standards body IEEE Task Force P802.3bs as already ratified in the 400GBASE-DR4 PHY the use of 100 Gb/s/ λ employing PAM-4 at 50 Gsymbols/s over 4 pairs of single mode fibers for the future 400GbE standard, that is expected as early as 2017. With the relentlessly increasing bandwidth demands, wavelengths running at >100 Gb/s are expected in upcoming short reach Terabit Ethernet rates. In data

centers where data needs to be replicated and multicast to different location, such multicasting operation also needs to be performed in a cost effective way.

A second approach to the reduce cost and size of optical interconnects is to use inexpensive technology platforms to fabricate and integrated components. Silicon photonics has proven to be an enabling technology and an inexpensive platform that supports high degrees of integration, where the entire architecture can be put on a substrate using already widely deployed and commercially available CMOS fabrication processes. Silicon photonics can be employed for both passive and active optical signal processing. Passive and active Silicon photonics are differentiated by whether or not electrical signals interplay with the photonic waveguide. Silicon photonics allow integration of both electronics and optics on chip such to reduce footprint, cost and power consumption per Gigabit of high speed pluggables.

In this thesis, we address the growing needs for speed of both long-haul transports and short-reach interconnects.

1.2 Review of the State of the Art

We describe in this section the state of the art in the three fields of optical communication addressed in this thesis. For long-haul optical systems employing coherent receivers, we first present a literature survey of different algorithms for polarization recovery, their respective advantages and pitfalls. We also present different approaches that have been reported for mitigation of fiber nonlinearity. In the realm of passive silicon photonics, we present a wavelength multicasting approach using the parametric process of four-wave-mixing (FWM) in a silicon-on-insulator (SOI) photonic waveguide. For active silicon photonic devices, we study on a Silicon-based Mach-Zehnder Modulator for external laser modulation. Finally, we review the state of the art for optical communications targeting intra- and inter-data center interconnects.

1.2.1 Polarization Recovery and Fiber Nonlinearity Mitigation

Recovering and canceling polarization cross talk is a paramount process in the digital signal processing stack of coherent receivers. Modulation formats employing polarization division multiplexing modulate a complex waveform independently on orthogonal states of polarization of the optical field. After propagation in single mode fiber, the state of polarization has rotated and the receiver's axis do not match that of the principle state of polarization of the lightwave entering the receiver. Therefore, one of the processing required by the circuitry of coherent receivers is to

derotate the received field and untangle polarization cross-talk. Several algorithms can serve this purpose, with varying efficiencies that depend on the algorithm per se but also on the modulation format and the noise level of the received signal. Moreover, some algorithms operate blindly, i.e. without requiring any decision to update their tap values, and some need a decision slicer to update filter coefficients. In such cases, a phase tracking algorithm embedded in the PR algorithm is also required. The constant modulus algorithm [11], a blind PR algorithm, is widely employed for Dual-Polarization-QPSK format, because the QPSK format exhibits a constant modulus at symbol locations. The CMA algorithm applied for polarization multiplexed formats is a 2-by-2 Multiple-Input Multiple-Output (MIMO) filter that minimizes the power variation on each output from the desired constant power. Although simple to implement and operate, CMA has multiple drawbacks: it is slow to converge, its initial performance greatly depends on the input SOP, it has a sub-optimum performance in steady operation, it exhibits a potential singularity problem (converges such that the two outputs are the same), and it is greatly sub-optimum for higher-order Quadrature-Amplitude Modulation (QAM) formats, like 16QAM and higher. Another PR algorithm, the Radius-Directed Equalizer (RDE) [12], relies on the amplitude after decision to update taps. It has been found however that this algorithm does not converge well in practice [13]. A recently proposed cascaded three-modulus blind equalization algorithm [14] has shown worse convergence properties than the CMA and can only be used after channel estimation. The Multi-modulus Algorithm (MMA) is another blind, phase independent algorithm meant to cope with the constellation points of formats of higher QAM orders of varying amplitudes. However, due to its multi-ring nature, it adapts even more slowly than its single ring CMA counterpart. Finally, an algorithm based on independent component analysis (ICA) for blind polarization demultiplexing was presented in [15]. Although ICA converges faster than CMA, it is always more computationally costly than CMA, and even more so for higher order QAM formats.

We also address different approaches recently reported for mitigation of fiber nonlinearities for long-haul optical systems. Several optical and electronics-based (DSP) methods were reported. Amongst these are mid-link optical phase conjugation [16–19], phase conjugated twin-waves [20], digital back-propagation [21–23], Volterra series filters [24] and perturbation-based nonlinear compensation [25–27]. All aforementioned methods have drawbacks and limitations. For methods relying on digital signal processing, the compensation is basically carried out by calculating the

nonlinear noise and cancelling out it from original signal. However, the computational complexity of the nonlinear noise has been an obstacle for implementation with a realistic gate count and power consumption. For instance, digital back-propagation carried by the split-step Fourier method usually requires several concatenation stages of chromatic dispersion compensators and nonlinear compensators to emulate the inverse transfer function of the fiber transmission link and, consequently, its computational complexity becomes quite large. For perturbation-based nonlinear compensation, the nonlinear noise accumulated in all transmission spans can be calculated from the transmitted signal, and thus nonlinear compensator can be implemented with only one stage. The perturbation nonlinear compensation method is thought to be a promising approach to compensate the fiber nonlinearity using DSP resources, in particular intra-channel nonlinearities [28]. Volterra series nonlinear equalizers also fall in the category of digital back-propagation of the signal, with the objective to invert the nonlinear Schrödinger equation. Similar to the Split-step Fourier method, Volterra series filters are also quite computationally hungry. The optical method of mid-link optical phase conjugation requires access and alterations on the fiber-optic links for the insertion of a phase conjugator: a technique that is extremely costly if not impossible for currently deployed submarine links. Finally, the phase conjugated twin-waves method requires to transmit the same information twice on two orthogonal state of polarization. This method therefore inherently loses the dual polarization multiplexing capability of single mode fibers and therefore halves the possible binary throughput. In this thesis, we propose a passive way to mitigate nonlinear impairments where both polarizations contain information and where the method does not require any additional computational resources nor needs access to the middle of the link.

1.2.2 Passive and Active Silicon Photonic Devices

Passive silicon photonics encompass a plethora of different devices where each has its own properties and intended objectives. In this thesis, we focus on only 1 device: a silicon nanowire serving for wavelength multicasting through the parametric process of four-wave-mixing in a silicon-on-insulator photonic waveguide. Silicon-on-insulator can be used to route and convey light on a planar lightwave circuit. Because of the very large refractive index difference between silicon ($n=3.5$) and silica ($n=1.44$, at $\lambda = 1.55 \mu\text{m}$), SOI waveguides efficiently guide light and support bends with radius of curvature as small as $3 \mu\text{m}$ [29]. However, silicon waveguides in silicon integrated photonic circuits exhibit a large waveguide propagation loss of about 1.5 dB/cm [30].

Silicon structures can also perform all-optical signal processing. As the rectangular size of SOI waveguides is quite small, typically of width $500\ \mu\text{m}$ and height $220\ \mu\text{m}$, the area onto which all the optical power propagates is very confined. Consequently, the nonlinear coefficient in SOI waveguides is very large and such waveguides are good medium for nonlinear signal processing [31]. In [32], all-optical wavelength multicasting in silicon was reported using a 40 Gbit/s NZR format and where N -fold multicasting of the high power 40 Gbit/s signal is done by multiplexing N CW laser inside the SOI waveguide. Although this approach allows flexibility in terms of the number of multicast wavelengths, is it power hungry as a total of $N + 1$ laser are required. In [33] a one-to-six wavelength multicast is reported using only 2 CW pumps interplaying with the modulated wavelength. However, the modulation is at a low signaling rate of 5 Gsymbols/s and uses a QAM format of the lowest-complexity: a differential phase-shift-keying format. In this thesis, we report for the first time a one-to-six wavelength multicast of a 16-QAM signal at 22 Gsymbols/s.

Finally, the silicon platform can also serve to modulate light. Although third order nonlinearity is high in silicon, second-order optical nonlinearities are absent in bulk silicon because of its centrosymmetry. However, high-speed modulators can still be fabricated in silicon, based on free-carrier injection, also known as the plasma dispersion effect. These silicon-based modulators fall in the category of ‘Active silicon photonics’. Silicon-on-Insulator waveguides and structures, in conjunction with proper doping levels and metal track depositions, can serve the same modulating purposes as any Lithium Niobate (LiNbO_3) structure. For intensity modulation (IM) only, there is extensive literature on silicon-based IM, which can be done in a Mach-Zehnder Interferometer (MZI) structure [34–38] or in a ring resonator structure [39–42]. Most of these work report NRZ On-Off keying intensity modulation at various bit rates, not exceeding 56 Gbit/s. In this thesis we study a silicon-based MZI and report for the first time an intensity modulation of 112 Gbit/s over 10 km using PAM-4 as a modulation format.

1.2.3 Optical Short Reach Transceivers

The data center landscape is changing rapidly. Broad investments are driving demands for new optical modules to address the needs for faster throughputs, lower power consumption and longer reach. The data center market demands access to fiber optic modules to connect servers, switches and storages. There exist a broad portfolio of commercial products for optical short reach interconnects. These fiber optic solutions are suited for applications within the datacenter,

as well as for longer distances bridging one data center to another. These products span 10G Ethernet, 40G Ethernet and 100G Ethernet over reaches from 2-10km to 80km in form factors such as SFP+, XFP, QSFP+, QSFP28 and CFP/CFP2/CFP4. A good example showing the variety of transceivers for data centers is found in Finisar[®]'s "*Transceivers, Transponders, and Active Optical Cables*" product guide [43]. All currently available optical transceivers for data centers rely on 2-level intensity modulation with direct detection from a single PD, with a maximum signaling rate per lane of 28 Gb/s. To increase the bit rate, the current solution is to multiplex multiple X Gb/s On-Off Keying (OOK) modulated lasers on different fiber lanes. It is well accepted that both the signaling rate and the spectral efficiency of the format have to increase. A 400 Gb/s Physical Layer taskforce was formed in March of 2014, IEEE 802.3bs[™], with the objective of defining a 400Gb/s physical layer specification by March of 2017. The physical layer for 400Gb/s over 500 m has already been ratified in May 2015 as 400GBASE-DR4, specified as 4 optical channels of 100 Gb/s/ λ employing PAM4 at 50 Gbaud. There is a clear trend to improve the spectral efficiency and the bit rate per wavelength by employing more complex modulation formats and faster signaling rates.

A signaling rate increase for the OOK format from 28 Gbit/s to 56 Gbit/s using a distributed feedback laser (DFB) was reported in [44]. Alternatively, the same bit rate was also reported in [45] by maintaining the electrical signaling rate at 28 Gsymbols/s but where the intensity is varied of 4 levels instead of 2. Advanced modulation formats have gained increasing interest from academic research and industry developers as a method to reduce the number of fiber lanes and wavelengths while increasing the total link capacity.

Recent research efforts have shown a number of demonstrations of various advanced intensity modulation formats such as pulse amplitude modulation (PAM) [5], discrete multi-tone modulation (DMT) [46], half-cycle 16-QAM Nyquist subcarrier-modulation [47] and (multiband-) carrier-less amplitude phase modulation (CAP) [48]. Aside from pulse amplitude modulation, all these IM formats for direct detection are enabled by digital signal processing (DSP) at the transmitter. Multi-level PAM- N formats can be generated using $\log_2(N)$ On-Off RF tributaries, from either electrical power combination or from a N -segment, segmented Mach-Zehnder modulator [49]. From current components availability, such unidimensional IM/DD transceivers are limited to around 100G of data payload. In this thesis, we report the first silicon-based 112 Gbit/s intensity modulator relying on 56 Gbaud PAM-4, and also present the first 300 Gbit/s transceiver with direct detection.

1.3 Research Objectives and Thesis Overview

The objectives of this thesis are to demonstrate solutions to specific problems and challenges in all three categories of optical communications: long-haul with coherent detection, chip-scale photonics in silicon, and optical transceivers for data centers. For long-haul with coherent detection, we propose a novel algorithm for polarization recovery. The algorithm converts the received Jones vectors to operate in the Stokes space, blindly finds the state of polarization of the polarization multiplexed signal as it enters the polarization diversity coherent receiver, and applies back the unitary derotation matrix to the input Jones vectors. The algorithm works for any M -QAM polarization multiplexed format, is experimentally validated, and converges orders of magnitude faster than the common blind constant modulus algorithm.

Still in the realm of long-haul optical transport with coherent detection, we propose in the thesis a passive method allowing mitigation of nonlinear effects occurring during propagation in single mode fibers. The passive approach is a novel modulation format for dual-polarization transmission where all symbols exhibit a constant total optical power. The format proposed is a power constrained version of the Dual-Polarization (DP)-8QAM format, that we call ‘8PolSK-QPSK’. The format is experimentally demonstrated and compared against DP-8QAM, where we show that for the same bit rate, the proposed 8PolSK-QPSK format allows greater launch power, extended reach, and smaller penalty of the required optical signal to noise Ratio (OSNR) with increasing launch power.

For photonics in silicon, we present in the thesis one passive and one active optical system working on different principles for different purposes. We present the first high symbol rate, high order QAM 1-to-6 wavelength multicasting in a 20 mm long silicon nanowire of width 500 nm and height 220 nm working on the parametric process of four-wave mixing. For active silicon photonics, we present the first silicon-based Mach-Zehnder intensity modulator operating at 1.3 μm and allowing modulating 112 Gbit/s via an 8-level pulse amplitude modulation running at 37.4 Gsymbols/s with direct detection. Modulation in silicon works via the plasma dispersion effect with carrier depletion in doped pn junctions. We show full characterization and specifications of the Si Mach-Zehnder modulator.

Finally, in the category of short reach optical transceivers, we first present a data center compliant version of the previous 112 Gbit/s transmission, this time employing a PAM-4 format at 56 Gbaud: a physical layer specification that was recently ratified in IEEE 802.3bsTM. The transceiver

is enabled by a DAC and an ADC and we detail all digital signal processing. Instead of covering a propagation length of only 500 m as mentioned in 400GBASE-DR4, we cover distances up to 10 km. We complete the thesis by presenting a single wavelength modulation scheme of high spectral efficiency providing +300 Gbit/s with direct detection via a Stokes receiver. We introduce this novel modulation format and perform experimental validation. We detail two different DSP algorithms to process the received DD signals at the receiver, each of which having different computational complexity.

More specifically, this thesis is organized as follows:

In **Chapter 2**, we briefly introduce a short history of optical communication. We address some milestones in the development of fiber optic systems, up to the revival of coherent optical communications in 2005. We present the optical transmitter and coherent receiver architectures which are at the center of advanced modulation formats and high throughput optical channels in today's high speed coherent links, all enabled by digital signal processing. We also expose the evolution of connectivity in data centers, and the growth of data centers themselves. Finally, we briefly discuss the recent rise of silicon as a host material for dense integration of photonic modules with electronic components.

In **Chapter 3**, we propose and experimentally verify two novel schemes for long-haul digital coherent optical systems. First, we propose a blind and fast algorithm allowing recovering the state of polarization of any square dual polarization (DP)- M -QAM formats. The proposed algorithm allows a significant reduction of the convergence time to estimate the state of polarization of the light with respect to the receiver's axis compared to the traditional blind derotation method. Secondly, we propose a novel modulation format, called 8PolSK-QPSK. The format is a power constrained version of the Dual-Polarization 8PSK format, providing 5 bits/symbol. This power constrained format allows passive mitigation of nonlinear impairments occurring in during propagation in single mode fiber. The greater nonlinear tolerance of this 8PolSK-QPSK format is experimentally demonstrated. Moreover, a thorough analysis of the Manakov-PDM propagation equation applied to both formats is derived and provides analytic explanation of the 8PolSK-QPSK's improved performance.

In **Chapter 4**, we address passive and active silicon photonics. For passive devices, we demonstrate a six-fold multicasting of a wavelength modulated over 16-QAM at 22-GBaud using a dispersion engineered silicon nanowire. The passive silicon photonic waveguide allows all-optical signal

processing where all new converted idlers perform below a bit error rate threshold. The wavelength multicasting performance is compared against the theoretical power transfer during a three-wave mixing process and we demonstrate that the average power of degenerate idlers agrees with the predicted 6 dB difference with respect to that of nondegenerate idlers. Moreover, we also show the importance of the loss of the output grating coupler for passive multicasting using silicon nanowires and how it impacts the multicasting performance.

Subsequently, we study the performance of an active silicon photonics intensity Mach-Zehnder modulator operating at 1.3 μm for pulse amplitude modulation formats with direct detection. Details of the silicon photonic Mach-Zehnder interferometer are presented. We study the system performance at various bit rates, PAM orders and propagation distances. Moreover, we derive an analytical model for the Q -factor metric applicable for multilevel PAM- N signaling that is accurately experimentally verified in the case of Gaussian noise limited detection. We demonstrate that PAM order selection can be optimally chosen as a function of the desired throughput. We report what was the first demonstration of a 112 Gb/s transmission over 10 km of SMF fiber operating below pre-forward error correction bit error rate threshold of 3.8×10^{-3} . Effects of varying the bias voltage on the modulator's V_π , bandwidth, and on chip-insertion loss are studied, and their impact on transmission performance are experimentally investigated for PAM-4 and PAM-8 formats at a throughput of 112 Gb/s over varying distances of 0, 2, 10, and 20 km.

In **Chapter 5**, we improve the performance of the silicon photonics IM presented in **Chapter 4** and report 112 Gbit/s transmission with an electrical signaling rate of 56 Gsymbols/s using PAM-4 modulation. The performance impact for a varying number of transmitter pre-emphasis taps and receiver equalization taps is studied, which has a direct impact on the transceiver's power consumption. We show that residual chromatic dispersion has no impact on performance up to 20 km.

Lastly, we present a novel modulation format and transmitter and receiver architecture allowing to practically triple the bit rate of the intensity modulation/direct detection transceiver presented above. We demonstrate a Stokes Vector modulation using intensity modulation, polarization multiplexing, and inter-polarization phase modulation on a single wavelength. 2 bits are encoded on each intensity of the two orthogonal polarizations, and 2 bits are encoded in the phase difference between the two polarizations, giving a 6 bits per symbol format. The format provides 6

bits/symbol and delivers 300 Gbit/s throughputs. The DP-PAM4+4PM format is demultiplexed at the receiver using DSP algorithms that operate in a 4 dimensional power space after direct detection. We study and compare two digital signal processing approaches to recover the format after rotation at the receiver. We propose a two-stage and a three-stage DSP process and we compare both approaches over their computational complexity and their bit error rate performance.

Chapter 6 concludes the thesis by summarizing the key contributions and provides some potential directions for future work.

CHAPTER 2

Optical Communications: From Long-Haul Coherent to Chip-Scale Photonics

2.1 Overview

Optical communication systems date back to the 1790s, to the optical semaphore telegraph invented by French inventor Claude Chappe [50]. The advent of telegraphy and Morse code in the 1830s replaced the use of light by electricity and began the era of electrical communications. The use of intermediate relay stations allowed communication over long distances (~ 1000 km). The development of worldwide telephone networks during the twentieth century led to many advances in the design of electrical communication systems. The bandwidth of such systems was limited by the frequency-dependent cable losses, which increase rapidly with frequency.

A commonly used figure of merit for communication systems is the bit rate–distance product, BL , where B is the bit rate and L is the repeater spacing [51]. It was realized during the second half of the twentieth century that an increase of several orders of magnitude in the BL product would be possible if optical waves were used as the carrier instead of electrical carriers. However, neither a coherent optical source nor a suitable transmission medium was available during the 1950s. The invention of the laser and its demonstration in 1960 by Theodore Maiman solved the first problem [52]. Attention was then focused on finding transmission medium for using laser light for optical communications over long distances. Many ideas were advanced during the 1960s. It was suggested in 1966 that optical fibers might be the best choice, as they are capable of guiding the light through total internal reflection. Before 1960, all fibers developed were bare and lacked any form of cladding, with total internal reflection occurring at a glass-air interface. Dutch scientist Abraham Van Heel made the crucial innovation of cladding fiber-optic cables by covering a bare fiber with a transparent cladding of lower refractive index. This protected the total reflection surface from contamination and greatly reduced cross talk between fibers. This coating allowed

reduction of the attenuation to about 1000 dB/km, which remained the main problem of optical fibers available during the 1960s. A breakthrough occurred in the summer of 1970 when researchers from Corning Glass began experimenting with fused silica, a material capable of extreme purity with a high melting point and a low refractive index, allowing fiber losses to be reduced to below 20 dB/km. The research team developed a single mode fiber and reduced the losses by doping the fiber core.

The simultaneous availability of compact optical sources and a low-loss optical fibers led to a worldwide effort for developing fiber-optic communication systems. The research phase of fiber-optic communication systems started around 1975. The progress has indeed been rapid with an increase in the bit rate by a factor of 100 000 between 1980 and 2010. Transmission distances have also increased from 10 to 10 000 km over the same time period. Just during the period between 1975 and 2000, the bit rate-distance product doubled every year. As a result, the bit rate-distance product of modern lightwave systems exceeds by a factor greater than 10^7 that of the first-generation lightwave systems.

The invention of the Erbium-Doped Fibre Amplifier (EDFA) in 1987 revolutionized the use of optical communication systems over long distances. EDFAs allowed broad optical amplification of in the entire 1530-1565 nm windows, where attenuation in silica is the lowest. An experiment in 1991 employing optical amplification showed the possibility of data transmission over 21 000 km at 2.5 Gb/s, and over 14 300 km at 5 Gb/s, using a recirculating-loop configuration [53]. This performance indicated that an amplifier-based, all-optical, submarine transmission system was feasible for intercontinental communication.

On 14 December 1988, TAT-8, the first Trans-Atlantic Telephone fiber optic system was completed across the Atlantic. The 5066 km line was capable of handling 40 000 telephone calls, a traffic equivalent to carrying 280 Mbit/s, and was owned by AT&T, France Télécom, and British Telecom. In 1996, the first international, optically amplified, fiber optic submarine system was installed. TAT-12 and -13 systems operating at a line rate 4.8 Gbit/s form the Trans-Atlantic Telephone fiber optic cable network, connecting the U.S. mainland, Britain, and France in a self-restoring loop configuration. Since then, a large number of submarine lightwave systems have been deployed worldwide. The following figure shows the world's submarine telecommunications cable systems as compiled by South-African Greg Mahlknecht [54].

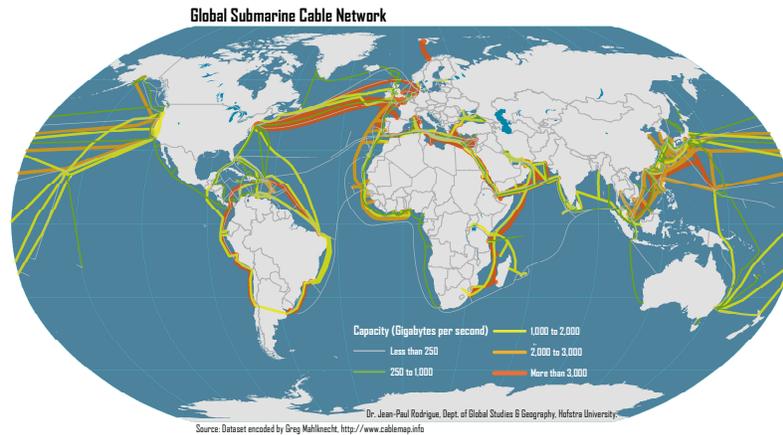


FIGURE 2–1: Global Submarine Cable Network.

2.2 Coherent Optical Communications

Coherent optical fiber communications were studied extensively in the 1980's mainly because the high sensitivity of coherent receivers could elongate the unrepeated transmission distance. However, the research and development on coherent receivers have been interrupted for nearly 20 years behind the rapid progress in high-capacity WDM systems using EDFAs. In 2005, the demonstration of digital carrier phase estimation in coherent receivers has stimulated a widespread interest in coherent optical communications again [55,56]. That same year, Nortel Networks[®] released their first 10G DSP-enabled transceiver, a breakthrough technology called electronic Dynamically Compensating Optics (eDCO) that provided pre-compensation of chromatic dispersion of each wavelength before its transmission, extending wavelength distances while enabling reduced network planning and engineering costs [57]. Digital coherent receiver enables us to employ a variety of spectrally efficient modulation formats such as M-ary phase shift keying (PSK) and quadrature amplitude modulation without relying upon a rather complicated optical phase-locked loop to recover the format at the receiver. In addition, since the phase information is preserved after detection, we can realize electrical post-processing functions such as compensation for chromatic dispersion and polarization-mode dispersion in the digital domain. These advantages of the born-again coherent receiver gave enormous potential for innovating existing optical communication systems. The following Fig. 2–2(a) is a schematic of a generic Dual-Polarization In-phase–Quad-phase (DP-IQ) modulator and Fig. 2–2(b) is a generic Polarization and Phase diversity coherent receiver.

In single mode fibers, an optical field exhibits 4 independent degrees of freedom, being the amplitude and the angle on each orthogonal polarization. A DP-IQ modulator allows full control

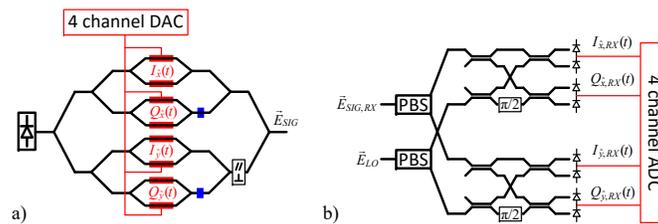


FIGURE 2–2: DP-IQ transmitter and Dual-Polarization Coherent Receiver

on those 4 dimensions. To better understand DP-IQ modulators, one can look at the functionality of a single-polarization IQ modulator. An IQ modulator is a dual-parallel Mach-Zehnder Modulator (MZM) where two MZMs are nested inside a parent MZM. The two ‘child’ MZMs are biased at null while the ‘parent’ MZM is biased such to combine the 2 tributary fields at quadrature. Each of the 2 ‘child’ MZM can be driven in a Push-Pull fashion, where a positive phase shift in one arm is accompanied by an equivalent negative phase shift in the other arm. A DP-IQ contains two of these IQ modulator, where the respective outputs are coupled to orthogonal polarizations in the single mode fiber. The resulting optical field out of a DP-IQ modulator is

$$\vec{E}_{SIG} \propto \begin{bmatrix} (e^{iI_x(t)} - e^{-iI_x(t)}) + i(e^{iQ_x(t)} - e^{-iQ_x(t)}) \\ (e^{iI_y(t)} - e^{-iI_y(t)}) + i(e^{iQ_y(t)} - e^{-iQ_y(t)}) \end{bmatrix} \propto \begin{bmatrix} \sin(I_x(t)) + i \sin(Q_x(t)) \\ \sin(I_y(t)) + i \sin(Q_y(t)) \end{bmatrix} \quad (2.1)$$

Therefore, when the approximation $\sin(x) \approx x$ holds, the 4 electrical signals $I, Q_{x,y}$ are mapped onto 4 orthogonal dimensions of the optical field.

The polarization and phase diversity coherent receiver of Fig. 2–2(b) is comprised of two polarization beam splitters, two 90° optical hybrid and 4 balanced photodetectors. The received optical field $\vec{E}_{SIG,RX}$ differs from the transmitted field \vec{E}_{TX} in that it was affected by polarization rotation, polarization mode dispersion, chromatic dispersion, and additive noise, to name only a few linear impairments out of all linear and nonlinear impairments occurring during propagation. When a local oscillator (LO) laser is tuned to a frequency close to that of the transmitter laser, the 4 output photocurrents allow the recovery of the optical field in the electrical domain, as seen by the receiver’s polarization axis. The benefit of coherent reception lies exactly in this property of having access to the 4 dimensions of the optical field. To the contrary, regular direct detection receivers give access to the total power of the incident field, or said differently to the sum of the squares of the 4 dimension. There is a plethora of digital signal processing algorithms at the receiver after analog to digital conversion that allow recovering and estimating the information transmitted in the

4 dimensions. With digital to analog converter driving the full field modulator of Fig. 2-2(a), there is also an abundance of transmitter side DSP to generate complex optical waveforms. In Chapter 3 of this thesis, we address the recovery of polarization rotation at the receiver and propose a novel modulation format to be generated at the transmitter, using the DP-IQ and the coherent receiver of Fig. 2-2.

2.3 The Rise of Silicon Photonics

In the past decade, significant advances have been made in silicon photonics to address the energy efficiency and cost of traditional optical transceiver using III-V compound materials. Silicon photonics allows the integration of different optical components side-by-side with electrical circuits on a single silicon chip, in standard 90 nm semiconductor fabrication technology. Silicon photonics takes advantage of CMOS foundries with large wafers and high yields. The technology includes a variety of silicon photonics components such as modulators and germanium photodetectors to be integrated with high-performance analog and digital CMOS circuitry. Single-chip optical communications transceivers can now be manufactured in a standard CMOS foundry, rather than assembled from multiple parts. The use of a standard chip manufacturing process alleviates the cost of traditional interconnects. Moreover, high speed digital to analog and analog to digital converters are readily built using CMOS technology [58]. From all the above, we can foresee a natural marriage of silicon photonics, application specific integrated circuits, and digital signal processing, all living on the same silicon-based CMOS platform. Although not the material of choice for semiconductor lasers due to its indirect band gap, silicon has good thermal conductivity, transparency at the traditional telecom wavelengths and low power consumption. Silicon can be employed at all hierarchical levels of optical communications. Figure 2-3 depicts the variety of target distances where optical communication is employed. Silicon photonics can be employed on-chip and from chip-to-chip, but there is a great interest recently in using silicon photonics for inter-rack communication in data centers [59].

High speed modulators are key components of the optical links. Significant progress has been made to advancing silicon photonics optical modulators based on both Mach-Zehnder modulators and electronically tuned ring resonators. In 2004, researchers from Intel Corporation experimentally demonstrated for the first time a silicon-based optical modulator with a bandwidth that exceeded 1 GHz [60]. This was a major milestone in silicon photonics and attracted huge media attention.

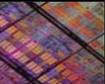
Network type	MAN & WAN	LAN	System	Board	Chip
	Metro & long haul 	Campus, Enterprises 	Intra-rack Inter-rack 	Chip-to-chip 	On-chip 
Distance	Multi-km	10 – 300 m	0.3 – 10 m	0.01 – 0.3 m	<2 cm
Adoption of optical	Since 80s	Since 90s	Since late 00s	After 2012	After 2012
Type of Connectivity	All-optical	Point-to-point and All-optical	Point-to-point	Point-to-point	Point-to-point & all-optical

FIGURE 2–3: Optical networks topology

Today, silicon-based optical modulator attain bandwidths in excess of 40 GHz [34], and a lot of research to further improve bandwidths are still undergoing.

2.4 Connectivity in Data Centers

The exponential growth of Internet traffic continues to push data communication networks to higher data rates. Due to the prevalent use of smart mobile devices and the rising demand for cloud and social media services across the business, educational and entertainment sectors, global data center traffic will grow 3-fold between 2014 and 2019, at a 25% compound annual growth rate [61]. Meeting these huge increases of data center traffic requires architectural and technological advances to the underlying transceivers. These advances are needed in order to enable scalable growth of intra- and inter-data center traffic volumes, whilst simultaneously decreasing transceiver size, cost and power consumption. One promising way to achieve this lies in the reduction of the number of signal lanes, or optical channels, by using advanced modulation formats to increase the transfer rate per wavelength. Modulating multiple dimensions of a lightwave over numerous levels allows increasing spectral efficiencies and binary throughputs of single channel short reach transceivers while maintaining a cost effective direct detection schemes.

Currently, data center are dominated by 10G and 25G IM/DD transceivers, and rely on multiplexing on either multiple fibers or multiple wavelengths these lower speed IM/DD streams to increase the capacity per pluggable [43]. To address the surging demand, the Institute of Electrical and Electronics Engineers has standardized in 2015 in IEEE Std 802.3bs™ the physical layer specifications to transfer of IEEE 802.3 format frames using 100 Gb/s per wavelength, using a slightly more complex modulation format of Pulse Amplitude Modulation over 4 levels instead of the 2

on-off levels currently employed. The standardization body ratified 51.6 Gbaud signaling rates to get to 100G of data payload per wavelength, and specified the use of 4 pairs of single mode fibers to reach the next 400GBASE target, to meet the increasing demand for higher speed interconnects. Data centers themselves are growing in size and quantity around the globe. From the work released by IEEE's Task Force and Study Group and the current research and development effort by the industry and academia, there is a clear trend to increasing the transfer rate per channel and reducing the multiplexing order, itself diminishing the footprint of the transceiver. Modulating wavelengths such to increase their binary throughput in conjunction with the use of silicon photonic modules for optical and electrical integration present great attraction and an immense potential for future short reach optics.

CHAPTER 3

DSP in Coherent Long Haul: Polarization Recovery and a 32-Symbol Constant Power Format

In this Chapter, we propose two solutions for 4 dimensional modulation and detection that are leveraged by digital signal processing at both the transmitter and the receiver. In the first place, we propose a novel algorithm for polarization recovery after the coherent receiver of Fig. 2–2(b) that operates blindly and works for any dual-polarization square– M -QAM modulation format. Secondly, we propose a new modulation format that provides 5 bits of information per recovered symbol, where the 4D constellation points share a common property that allows passive mitigation of nonlinear effects.

3.1 Stokes Based Polarization Recovery

We first begin by presenting a blind, fast and SOP independent polarization recovery for square dual polarization– M -QAM formats and optical coherent receivers.

3.1.1 Introduction

Optical coherent receivers are key to higher transmission rates, where information is imprinted in multiple dimensions of the optical waveform. Polarization is one of those dimensions, where different complex waveforms can be independently imprinted on two orthogonal state of polarizations of the signal at the transmitter. However, at the receiver side, the received SOP is completely unknown. It is mapped onto a different orthogonal basis than the one used at the transmitter and is varying with time and distance.

To blindly untangle the polarization of the incoming waveform in an optical coherent receiver, the constant modulus algorithm is often used [62]. However, the CMA has several disadvantages. First, it suffers from the singularity problem where the algorithm converges to a tap-weight setup

that produces the same transmitted signal at both equalizer outputs [63,64]. Moreover, the convergence speed of the CMA is strongly dependent on the SOP of the received waveform, and consequently varies a lot with the choice of the initial FIR tap values [63,65–67]. This has become a major problem and has recently been addressed to comply with the fast transition requirements of emerging data rate adaptive optical packet networks [66,67], burst mode coherent receivers and agile optical network architectures where the receiver can be dynamically reconfigured to quickly drop a wavelength and switch to another channel [65].

Recently, new blind polarization recovery techniques have been proposed to speed up the convergence time required for blind polarization rotation estimation. A method based on CMA using 25 different initial CMA tap values, all processing in parallel, is presented in [65]. It chooses the test case that provides the lowest CMA error to initialize the central tap of the equalizer. Using such scheme on DP-4QAM format provides a mean convergence time of around 35 ns, with a worst case convergence time of 280 ns [65]. The computational requirement of processing independently and in parallel 25 self-adaptating equalizers is very substantial. In [67], a three-stage CMA enables blind recovery in 200 ns, or 11200 symbols at 56 Gbaud, which is hardly compatible with burst-mode operation [66]. Some proposed techniques to reduce the convergence time of CMA use training symbols or preamble headers. As an example, in [66], a preamble of 30 header symbols is used to estimate to polarization rotation matrix and in [68], a training symbol based algorithm using 20 or 40 symbols is proposed to estimate the Jones channel matrix. Such techniques, although providing quick rotation estimates, not only add overhead to the system but also need synchronization with the preamble header for polarization matrix estimation, and hence are not true blind approaches.

In this section, we present a new method for blindly estimating the SOP of the optical waveform for coherent receivers applicable to any polarization multiplexed square- M -QAM format. The process consists of converting the received waveform into the three dimensional Stokes space and looping over an algorithm looking for desired Stokes states on the Poincaré sphere until a unitary polarization derotation matrix can be estimated. This matrix is then used to initialize the center taps of the subsequent decision directed-LMS polarization equalizer and tracker. The method realizes much shorter convergence time than recent methods using modified version of CMA to improve convergence speed. This section is divided as follows. In Section 3.1.2, we present the mapping of Jones constellations to Stokes constellations for different DP- M -QAM formats. We also summarize a well known model for waveform polarization rotation and polarization mode dispersion

(PMD) in lightwave systems [69]. We will use this model to motivate our approach. In Section 3.1.3 we introduce the new blind PR method and present in detail its operating algorithm. Section 3.1.4 presents the experimental test bed and explains the metric of comparison used to confront the convergence time and initial BER of our Stokes space blind method against that of CMA for 3 different modulation format: DP-4QAM, DP-16QAM and DP-64QAM. Finally, Section 3.1.5 presents experimental performance comparison for the CMA method and the proposed method, as well as numerical simulation results assessing the new method over varying PMD. We conclude this work in Section 3.1.6.

3.1.2 DP-MQAM Formats in Stokes Space and Modeling of SOP in Lightwave Systems

An optical coherent receiver collects simultaneously 4 electrical waveforms that represent the real and imaginary part of an optical waveform mapped into an orthogonal basis formed by the optical coherent front end. Those 4 waveforms of dimension $1 \times N$ can be converted into a single waveform of dimension $2 \times N$ which consists of a concatenation of N Jones vectors as in Eq. (3.1), where $A_{x/y-\text{Re/Im}}$ are the 4 sampled signals and N is the number of samples. The right hand side of Eq. (3.1) is a different representation of the combination of the 4 waveforms, where the magnitude, the absolute phase, the relative amplitude and sign on \hat{x} and \hat{y} and the relative phase of \hat{x} with respect to \hat{y} are all embedded in E_n , δ_n , θ_n and ϕ_n , respectively.

$$|A(nT)\rangle = \begin{bmatrix} A_x \\ A_y \end{bmatrix} = \begin{bmatrix} A_{x-\text{Re}}(nT) + iA_{x-\text{Im}}(nT) \\ A_{y-\text{Re}}(nT) + iA_{y-\text{Im}}(nT) \end{bmatrix} = E_n e^{i\delta_n} \begin{bmatrix} \cos\left(\frac{\theta_n}{2}\right) \\ \sin\left(\frac{\theta_n}{2}\right) e^{i\phi_n} \end{bmatrix} \quad (3.1)$$

The 3-by-1 Stokes representation of Eq. (3.1) is written as

$$\vec{A} = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = E_n^2 \begin{bmatrix} \cos(\theta_n) \\ \sin(\theta_n) \cos(\phi_n) \\ \sin(\theta_n) \sin(\phi_n) \end{bmatrix} \quad (3.2)$$

As can be seen in Eq. (3.2), the absolute phase information $e^{i\delta_n}$ is lost in the conversion from Jones to Stokes space. The new blind Stokes space polarization estimation method we present benefits from this absolute phase information loss as the input Stokes vectors to the algorithm are agnostic to the intermediate frequency offset and the phase variation between the signal and the reference local oscillator, both being common to the two orthogonal polarizations. The reader can refer to [70]

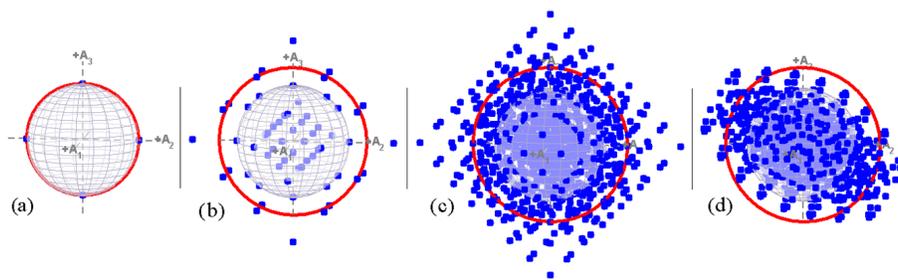


FIGURE 3–1: Mapping of all possible Jones states to Stokes space for theoretical (a) DP-4QAM, (b) DP-16QAM and (c) DP-64QAM. For (d), theoretical DP-64QAM is rotated by a random unitary matrix. The total mean power is always unitary. The gray sphere has a radius of 1.

for details concerning conversions from Jones to Stokes space. Different polarization multiplexed modulation formats can be represented in either the Jones or the Stokes space. Because it exhibits the extra phase information $e^{i\delta_n}$, the Jones space is commonly used to represent constellations. Constellations in Stokes space are rarely shown and we depict here in Fig. 3–1 such constellations for a) DP-4QAM, b) DP-16QAM and c) DP-64QAM, all of unitary mean signal power. As can be observed, much fewer states exist in this space; for instance, the $4^2=16$ states of DP-4QAM map to only 4 states in Stokes space. The red rings in Fig. 3–1 depict the equivalent modulus used in the cost function of the blind CMA polarization derotation algorithm. For every square QAM formats, the modulus has to increase with QAM order [11]. For instance, DP-4QAM, -16QAM and -64QAM formats have moduli of 0.5, 0.66 and ~ 0.69 , respectively. As those moduli apply to both polarizations independently, they are depicted as a single ring of radius double those values. The Stokes space representations of any square DP-MQAM format share a common feature: the most outer Stokes constellation points are all exactly aligned with the Stokes vectors $-\vec{A}_2$, \vec{A}_2 , $-\vec{A}_3$ and \vec{A}_3 at increasingly higher powers of $3(\sqrt{\text{MQAM}} - 1)^2 / (\text{MQAM} - 1)$. Our novel blind Stokes space polarization recovery method uses this theoretical alignment and power increase feature of the most outer Stokes constellation points for estimating the received SOP. As will be detailed below, locating the position of those 4 regions provides an estimate of the unknown unitary rotation imprinted in the received signal. It is out of the scope of this work to explain the generation of specific polarization rotation matrices, or “unitary” matrices. For detailed explanations, we refer the reader to the paper from Gordon and Kogelnik [70]. In summary, the simplest way to rotate a single Stokes vector from one position \hat{A} to another \hat{B} on the Poincaré sphere is done by applying a rotation matrix that

spins the Poincaré sphere on an axis given by the cross product of \hat{A} and \hat{B} , $\hat{A} \times \hat{B} = \hat{p}$, and by an angle given by the inverse cosine of the dot product of \hat{A} and \hat{B} , $\arccos(\hat{A} \cdot \hat{B}) = \phi$. The notation ‘ \hat{A} ’ means the unitary normalized Stokes vector of ‘ \vec{A} ’. There are actually an infinite number of possible unitary rotation matrices that rotate \hat{A} to \hat{B} .

3.1.2.1 Modeling of SOP in Lightwave Systems

As we know, the birefringence in an optical fiber breaks the degeneracy of the 2 eigenmodes that can propagate in a fiber at the same propagation constant. Birefringence gives 2 eigenmodes each having its respective propagation constants $\beta_j(\omega)$. Up to the first order, one can describe those 2 modes in the Jones-matrix and Stokes-matrix forms [69, 71, 72]

$$\frac{\partial |\tilde{A}(z, \omega)\rangle}{\partial z} = -\frac{i}{2} (\Delta\beta_0 + \omega\Delta\beta_1) \mathbf{U}^\dagger \boldsymbol{\sigma}_1 \mathbf{U} |\tilde{A}(z, \omega)\rangle = -\frac{i}{2} (\Delta\beta_0 + \omega\Delta\beta_1) (\hat{b} \cdot \vec{\boldsymbol{\sigma}}) |\tilde{A}(z, \omega)\rangle \quad (3.3)$$

$$\frac{\partial \vec{\tilde{A}}(z, \omega)}{\partial z} = -(\Delta\beta_0 + \omega\Delta\beta_1) \hat{b} \times \vec{\tilde{A}}(z, \omega) \quad (3.4)$$

where $|\tilde{A}(z, \omega)\rangle$ is the Fourier transform of $|A(z, \omega)\rangle$, $\mathbf{U}^\dagger \boldsymbol{\sigma}_1 \mathbf{U} \equiv \hat{b} \cdot \vec{\boldsymbol{\sigma}}$ and \hat{b} is the 3-by-1 birefringence vector. The term $\vec{\boldsymbol{\sigma}} = [\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3]$ is a vector of the Pauli spin matrices $\boldsymbol{\sigma}_k$ and are defined in [69, 70]. The term $\Delta\beta_0$ produces a differential phase shift, while the term $\omega\Delta\beta_1$ leads to a temporal delay between the two orthogonal eigenvectors of \mathbf{U} , which directly leads to PMD. Said differently, $\Delta\beta_0$ engenders simple rotations, or “spinning” of all the states, independent of frequency, whereas $\omega\Delta\beta_1$ leads to spinning of the components of the waveform as a function of their frequency, inherently resulting in differential delays. In the frequency domain, polarization dispersion manifests as a frequency dependent state of polarization at the output of the fiber [73]. One can note that the eigenvectors of $\mathbf{U}^\dagger \boldsymbol{\sigma}_1 \mathbf{U}$ are, in Stokes space, $\pm \hat{b}$, such that for constant birefringence, a SOP aligned on $\pm \hat{b}$ wouldn’t change. However, the birefringence vector \vec{b} has varying orientations and amplitudes along the fiber length and has to be a function of z . For the zeroth order, the magnitude of \vec{b} is $\Delta\beta_0$ and for the first order, its magnitude is $\Delta\beta_1$. The solution of Eq. (3.3) can be written in the form

$$|\tilde{A}(z, \omega)\rangle = \mathbf{W}(z) |\tilde{S}(z, \omega)\rangle \quad (3.5)$$

where $\mathbf{W}(z)$ is a function of solely z . If we apply Eq. (3.5) to Eq. (3.3), we get the 2 following equations

$$\frac{\partial \mathbf{W}(z)}{\partial z} = -\frac{i}{2} \Delta\beta (\hat{\mathbf{b}}(z) \cdot \vec{\sigma}) \mathbf{W}(z) \quad (3.6)$$

$$\frac{\partial |\tilde{\mathcal{S}}(z, \omega)\rangle}{\partial z} = -\frac{i}{2} \omega \Delta\beta_1 \mathbf{W}^{-1}(z) (\hat{\mathbf{b}}(z) \cdot \vec{\sigma}) \mathbf{W}(z) |\tilde{\mathcal{S}}(z, \omega)\rangle \quad (3.7)$$

If $\vec{\mathbf{b}}$ was z -independent, the solution to $\mathbf{W}(z)$ would simply be

$$\mathbf{W}(z) = \exp\left(-\frac{i}{2} \Delta\beta_0 z (\hat{\mathbf{b}} \cdot \vec{\sigma})\right) \quad (3.8)$$

where “exp” is the exponential of the 2-by-2 matrix $-i/2 \Delta\beta_0 z (\hat{\mathbf{b}} \cdot \vec{\sigma})$. To solve for $|\tilde{\mathcal{S}}(z, \omega)\rangle$, one can show that $\mathbf{W}(z)$ and $(\hat{\mathbf{b}}(z) \cdot \vec{\sigma})$ commute, i.e. $[\mathbf{W}(z), (\hat{\mathbf{b}}(z) \cdot \vec{\sigma})] = 0$. Therefore, the differential equation governing $|\tilde{\mathcal{S}}(z, \omega)\rangle$ becomes $\partial |\tilde{\mathcal{S}}(z, \omega)\rangle / \partial z = -i/2 \omega \Delta\beta_1 (\hat{\mathbf{b}} \cdot \vec{\sigma}) |\tilde{\mathcal{S}}(z, \omega)\rangle$, which has the same form as Eq. (3.6) and consequently the same solution form as Eq. (3.8), with $\Delta\beta_0$ changed to $\omega \Delta\beta_1$, adding frequency dependence. $|\tilde{\mathcal{S}}(z, \omega)\rangle$ solve as $|\tilde{\mathcal{S}}(z, \omega)\rangle = \mathbf{T}(z, \omega) |\tilde{\mathcal{S}}(0, \omega)\rangle$, with initial waveform $|\tilde{\mathcal{S}}(0, \omega)\rangle$ at $z=0$. This easily shows mathematically that PMD is simply a polarization rotation that is frequency dependent [73]. If PMD was neglected ($\Delta\beta_1 = 0$) and constant birefringence was assumed, $\vec{A}(z, t)$ would simply precess around $\hat{\mathbf{b}}$ as it propagates, at an angular speed proportional to $\Delta\beta_0$. However, as $\hat{\mathbf{b}}$ is z -dependent, we can assume without any loss of generality that birefringence is piece-wise constant along the fiber, that is, $\hat{\mathbf{b}}(z_j \leq z \leq z_{j+1}) = \hat{\mathbf{b}}_j$ and that for every $z_j \leq z \leq z_{j+1}$, both $\Delta\beta_{0j}$ and $\Delta\beta_{1j}$ vary. The solution of $|\tilde{A}(z, \omega)\rangle$, in the form of Eq. (3.5), is then [69]

$$|\tilde{A}(z, \omega)\rangle = \mathbf{W}(z_N) \mathbf{T}(z_N, \omega) \cdots \mathbf{W}(z_1) \mathbf{T}(z_1, \omega) \mathbf{W}(z_0) \mathbf{T}(z_0, \omega) |\tilde{\mathcal{S}}(0, \omega)\rangle = \mathbf{U}_c(z, \omega) |\tilde{\mathcal{S}}_0\rangle \quad (3.9)$$

where the total fiber length is divided into N shorter pieces. None of the matrices in Eq. (3.9) commute, so the order of application is paramount. $\mathbf{U}_c(z, \omega) |\tilde{\mathcal{S}}_0\rangle$ is the composite Jones matrix of the whole fiber link [71] and is also a unitary matrix at each frequency. In SMF fibers, $\Delta\beta_1$ is normally small. If we neglect PMD, all the $\mathbf{T}(z_j, \omega)$ matrices become the identity matrix \mathbf{I} , $|\tilde{A}(z, \omega)\rangle$ becomes ω -independent and the result of the propagation of $|A(z, t)\rangle$ is simply that its SOP wanders around on the Poincaré sphere as it propagates, adding no impairment on the initial waveform $|\tilde{\mathcal{S}}_0\rangle$. One can observe that for a static link without PMD, the SOP is constant at a fix distance and is only very slowly varying for small first order birefringence. The new polarization

recovery technique allows us to estimate blindly the cumulated polarization rotation $\mathbf{U}_c(z, \omega)$ with the hypothesis that the cumulative PMD after distance z is small. Said differently, it estimates the resulting matrix of Eq. (3.9) with the \mathbf{T} 's assumed to be almost the identity \mathbf{I} .

3.1.3 New Blind Polarization Rotation Estimation

The technique we present here allows estimating blindly the polarization rotation of the received signal after the optical front end of an optical coherent receiver. This technique can be used for any square DP-MQAM modulation format. In this section, we will explain this new method of blind polarization rotation recovery.

This blind polarization rotation estimation operates differently than other blind PR techniques like CMA. The mode of operation of our technique is different and independent of the steady-state mode of operation for polarization tracking. On the contrary, a technique like CMA can be used from system startup and be kept for steady state operation for some modulation formats like DP-4QAM [62]. The CMA algorithm iteratively adapts its coefficients blindly such that they minimize a phase independent cost function [11]. The CMA self-recovering equalizer is more often used as pre-convergence to reduce the effects of channel distortions in polarization before switching to the decision-directed Least-Mean Square (DD-LMS) equalizer which requires an acceptable average of good decisions by the slicer [11, 12].

In the technique presented here, we also share the same end goal of steady state operation using DD-LMS, but the blind technique implemented to get there is different. The proposed SOP estimation is not an iteratively adapting process like CMA. This new blind polarization recovery algorithm can be decomposed into two parts. In the first part, an algorithm is blindly looping through the input vectors until it finds enough vectors satisfying certain criteria. In the second part, those vectors are then employed to estimate the unitary rotation matrix: a 2-by-2 complex matrix \mathbf{U}_{ROT} in the Jones space or a 3-by-3 real valued matrix \mathbf{R}_{ROT} in the Stokes space. In the second part, once the looping algorithm found enough satisfying states, the estimation of the unitary matrix is decomposed into 3 sub-matrices. The first 2 matrices are estimated in the Stokes space and the last is estimated in the Jones space.

The input vectors of the looping algorithm of the first part are not coming directly from the RF waveforms generated by the four Analog to Digital Converters (ADC) of the optical coherent front end. They are obtained after the following operations are applied: I) Convert the 4 RF

waveforms $A_{x-\text{Re}}(nT)$, $A_{x-\text{Im}}(nT)$, $A_{y-\text{Re}}(nT)$ and $A_{y-\text{Im}}(nT)$ in a complex vector form as that depicted in Eq. (3.1); II) Resample the resulting time-domain series of complex two-dimensional (2D) vectors at $T/2$ samples per second, where T is the symbol duration; III) Apply chromatic dispersion (CD) compensation [74]. After operations I, II and III are applied, we assume that the resulting time series of complex 2D vectors are rotated and noisy versions of the Jones vectors that were transmitted. If those resulting vectors are converted to their Stokes space equivalent, the time series represent rotated and noisy versions of the transmitted Stokes vectors. From the theoretical Stokes constellations of DP-4QAM, -16QAM and -64QAM formats in Fig. 3-1 (a), (b) and (c) respectively, the Stokes time series represent rotated and noisy versions of those constellations. Mathematically, the vector time series is assumed to be $|A_{Rx}\rangle = \mathbf{U}_{\text{ROT}} |A_{Tx}\rangle + |n\rangle$ in Jones space, or $\vec{A}_{Rx} = \mathbf{R}_{\text{ROT}} \vec{A}_{Tx} + \vec{n}$ in Stokes space, where \vec{n} ($|n\rangle$) is a (complex) noise source and \mathbf{R}_{ROT} (\mathbf{U}_{ROT}) is the unknown Stokes (Jones) unitary rotation matrix that we are trying to estimate. Once \mathbf{U}_{ROT} is estimated, its 4 complex values are used to initialize the central taps of the subsequent DD-LMS adaptive complex 2-by-2 MIMO filter.

The looping algorithm tries to estimate the location of 3 out of the 4 outer-most states in the received noisy, rotated Stokes constellation. Once 3 distinct regions are obtained, the matrix estimation in the second part tries to realign those estimated corners to their desired location. The looping algorithm iterates through the following three steps:

1. If new input Jones vector has power greater than a certain power threshold P_{th} , retain. Save power of vectors retained. Retained vectors are converted to their Stokes representation, and normalized to unit power. Concatenate vectors in a buffer, $\mathbf{S}_{3,N}$, forming a 3-by- N matrix, where N grows by 1 for each new valid entry.
2. For each new entry to the buffer matrix $\mathbf{S}_{3,N}$, compute a new projection matrix $\mathbf{R}_{\text{proj}} = \mathbf{S}_{3,N+1}^T \mathbf{S}_{3,N+1}$, of size $N+1$ by $N+1$.
3. For each new \mathbf{R}_{proj} , verify if there is a row that satisfies the three following criteria:
 - 3.1 Contains at least C elements greater than $\cos(\gamma_1)$
 - 3.2 Contains at least C elements smaller than $\cos(\pi + \gamma_1)$
 - 3.3 Contains at least C elements smaller than $|\cos(\pi/2 + \gamma_2/2)|$

When all 3 criteria of Item 3 are met for a subset of Stokes vectors in $\mathbf{S}_{3,N}$, the looping algorithm stops. The subset of Stokes vectors saved in the buffer $\mathbf{S}_{3,N}$ that satisfied conditions 3.1 to 3.3 are used to estimate the first 2 matrices that \mathbf{R}_{ROT} is composed of in the following way.

In Step 3, the parameters C , γ_1 and γ_2 are to be chosen such that $C \geq 1$, $2\gamma_1 + \gamma_2 \leq \pi$ and $\gamma_2 \geq 2\gamma_1$. Criterion 3.1 tries to find C elements that are all close enough within a maximum angular distance of $2\gamma_1$. Criterion 3.2 looks for C elements that are also regrouped angularly within $2\gamma_1$ but opposite to those in Criterion 3.2. Finally, criterion 3.3 is looking for C vectors that are perpendicular within $\pi/2 \pm \gamma_2/2$ to those in 3.1 and 3.2 and are likewise gathered within $2\gamma_1$. These operations in Step 3 are well represented in Fig. 3–2, depicting 3 perpendicular and opposite “zones” that the looping algorithm tries to find. In our work, we use $C = 4$, $\gamma_1 = 20^\circ$ and $\gamma_2 = 60^\circ$.

From the 3 clusters of Stokes points, a three-dimensional (3D) plane can be fit in a in least mean square sense, with a plane norm \hat{p}_{hole} of unit-length. The norm of the plane \hat{p}_{hole} is the eigenvector of the smallest eigenvalue of the covariance matrix \mathbf{C}_{COV} of all Stokes vectors forming the 3 clusters, stacked in one 3-by-*time* matrix. This method is also called Principal Component Analysis (PCA) [75]. The first derotation matrix \mathbf{R}_{ROT-1} is the unitary matrix that brings \hat{p}_{hole} back to either \vec{A}_1 or $-\vec{A}_1$ depending on which is the closest to \hat{p}_{hole} . Both Stokes (\mathbf{R}_{ROT-1}) and Jones (\mathbf{U}_{ROT-1}) representation of said rotation matrix are computed. The second derotation matrix \mathbf{R}_{ROT-2} is obtained using the new location of the 3 clusters of Stokes points, after \mathbf{R}_{ROT-1} is applied. The resulting 3 clusters of Stokes vectors now sit somewhere in the $\{\vec{A}_2, \vec{A}_3\}$ plane, but are not necessarily around 3 of the 4 locations they are supposed to be for a noisy, non-rotated Stokes constellation. This second matrix \mathbf{R}_{ROT-2} brings those three clusters around the closest 3 points of $\pm\vec{A}_2, \pm\vec{A}_3$. $n\pi/2$ periodicity is removed by using the 4^{th} power algorithm [76, 77]. Both the Stokes \mathbf{R}_{ROT-2} and the Jones \mathbf{U}_{ROT-2} representation of this unitary matrix are also computed.

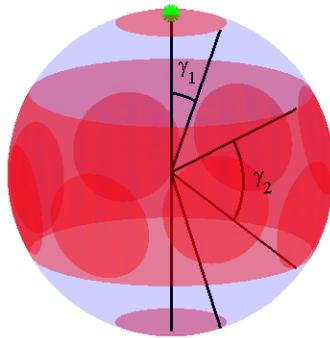


FIGURE 3–2: Explanation of algorithm in Step 3: The n^{th} column of $\mathbf{S}_{3,N}$ is the green dot. By projecting vectors onto each other, we need at least $C - 1$ other vectors close by, C vectors perpendicular and C vectors that are opposite to estimate \mathbf{U}_{ROT} .

Fig. 3-2 is a visual representation of three clusters of Stokes vectors the looping algorithm tries to find blindly in Step 3. Fig. 3-3 is another visual aid, where the result of keeping only Stokes vectors that have power greater than a certain threshold is well represented. The Stokes states \vec{S} that have power $(\vec{S}^T \vec{S})^{1/2} \geq P_{th}$ are marked as green dots and others as blue dots, for formats a) DP-4QAM, b) DP-16QAM and c) DP-64QAM, when the mean signal power is 1. Noise is loaded such that each of the 3 formats yields a theoretical BER of 3.8×10^{-3} . The red sphere has a radius of P_{th} . A 3D plane naturally arises from those high power vectors, of the norm of the plane points in the direction that has the smallest variance.

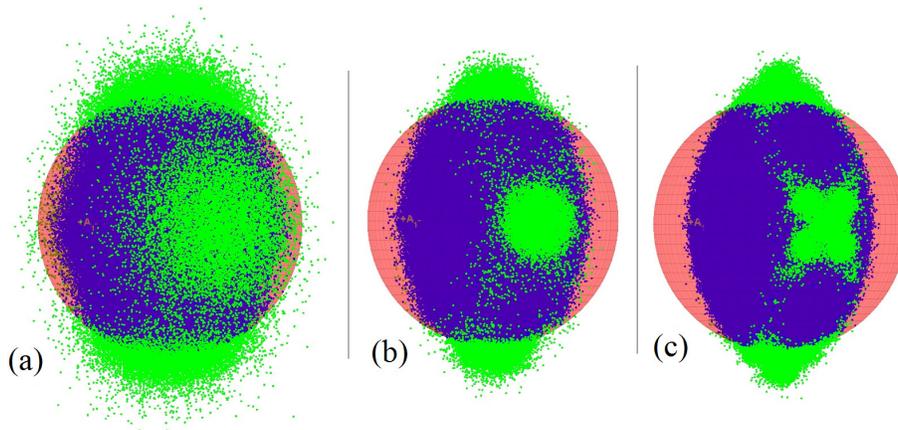


FIGURE 3-3: Power filtering representation in Stokes space for several QAM order. Noise loaded such that BER= 3.8×10^{-3} and P_{th} set such that on average, for every 100 Stokes vectors, 1 has powers exceeding P_{th} . (a) DP-4QAM, SNR of 8.53 dB, $P_{th} = 2.19$ (b) DP-16QAM, SNR of 15.2 dB, $P_{th} = 2.01$, (c) DP-64QAM, SNR of 21.1 dB, $P_{th} = 2.01$.

The last parameter to estimate is the absolute average phase deviation δ , that is accessible only in the Jones space, as demonstrated in Eqs. (3.1) and (3.2). To estimate δ , the intermediate frequency offset Δf needs to be evaluated and suppressed. Δf is evaluated by first applying $\mathbf{U}_{\text{ROT-1}}$ and $\mathbf{U}_{\text{ROT-2}}$ to all Jones vectors since the start of the looping algorithm, even those who had smaller powers. From there, Δf is estimated using the Fast Fourier Transform periodogram method [65]. Once $-\Delta f$ is applied back and the intermediate frequency offset suppressed, the absolute average phase deviation δ can be obtained. To do so, only the Stokes symbols in the last 10 ns before the looping algorithm stopped are employed. As laser linewidths are typically less than 2.5 MHz in coherent fiber optic systems [65], we assume that the received signal's phase noise can be considered constant within those 10 ns. As the phase $e^{i\delta}$ is present in both \hat{x} and \hat{y} polarizations

of each resulting Jones states, the number of phase estimated δ_k to compute the average is twice the number of states. This final absolute phase estimation in \mathbf{U}_{ROT} not essential, but helps the subsequent DD equalizer to start by making more “good” decisions.

The final polarization derotation estimation matrix $\mathbf{U}_{\text{ROT}} = e^{i\delta}\mathbf{U}_{\text{ROT-2}}\mathbf{U}_{\text{ROT-1}}$ is specifically described as

$$\begin{bmatrix} X_{\text{Rx,Derot}} \\ Y_{\text{Rx,Derot}} \end{bmatrix} = \underbrace{e^{i\delta}}_{\text{Phase est.}} \underbrace{\begin{bmatrix} e^{i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{bmatrix}}_{\text{2nd Matrix Estimate}} \underbrace{\begin{bmatrix} \cos(\theta) & e^{-i\chi}\sin(\theta) \\ -e^{i\chi}\sin(\theta) & \cos(\theta) \end{bmatrix}}_{\text{1st Matrix Estimate}} \begin{bmatrix} X_{\text{Rx,Rot}} \\ Y_{\text{Rx,Rot}} \end{bmatrix} \quad (3.10)$$

Computation of \mathbf{R}_{proj} in Step 2 for every new Stokes vector buffered in $\mathbf{S}_{3,N}$ is not a computationally expensive task: One can compute $\mathbf{R}_{\text{proj}}[\mathbf{N}]$ from $\mathbf{R}_{\text{proj}}[\mathbf{N}-1]$ by only computing the projection of all the previous vectors stored in $\mathbf{S}_{3,N-1}$ with the new added vector \vec{S} and append $\mathbf{S}_{3,N-1}^T \vec{S}$ to the previous $\mathbf{R}_{\text{proj}}[\mathbf{N}-1]$. Because of pre-normalization of interplaying Stokes vectors to unity, all \mathbf{R}_{proj} are real valued symmetric matrices with ones on the diagonal and values between 1 and -1 elsewhere. The number of real multiplications and additions required to compute $\mathbf{R}_{\text{proj}}[\mathbf{N}]$ of size N -by- N is $3N(N-1)/2$ and $N(N-1)$, respectively. The data covariance matrix \mathbf{C}_{COV} required to fit a plane in 3D space is expressed as $\mathbf{C}_{\text{COV}}(i,j) = \sum_{n=1}^N (S_{i,n} - \bar{S}_i)(S_{j,n} - \bar{S}_j)$, where \bar{k} is the mean in dimension k . We can show that the total computational complexity of computing \mathbf{C}_{COV} and its smallest eigenvalue and equivalent eigenvector is little. Normally, eigenvalue decomposition are computationally intensive as they are obtained iteratively. However, \mathbf{C}_{COV} has the advantageous property of being a real symmetric 3-by-3 matrix. All eigenvalues of such matrices can be computed directly, as demonstrated in [78]. Moreover, computing the eigenvector once an eigenvalue is known is quite trivial and goes as follows. If \vec{v}_1 is the desired column eigenvector of the eigenvalue λ_1 of matrix \mathbf{C}_{COV} , we can compute $\mathbf{B}\vec{v}_1 = (\mathbf{C}_{\text{COV}} - \lambda_1\mathbf{I}_3)\vec{v}_1 = \begin{bmatrix} \vec{b}_i^T \vec{v}_1, \vec{b}_j^T \vec{v}_1, \vec{b}_k^T \vec{v}_1 \end{bmatrix} = 0$. From the equation we know that the eigenvector \vec{v}_1 is perpendicular to \vec{b}_i , \vec{b}_j and \vec{b}_k . As the determinant of \mathbf{B} is zero, all \vec{b} 's lie in the same plane and consequently \vec{v}_1 is simply the cross product of any two \vec{b} 's, which can subsequently be normalized to unit power such that $\vec{v}_1^T \vec{v}_1 = 1$.

The 4 most power hungry processes required in finding \mathbf{U}_{ROT} are 1) computing the Stokes vectors that satisfy the power criterion P_{th} , 2) projecting each new Stokes vector onto all previously stored vector, 3) computing \mathbf{C}_{COV} and 4) computing λ_1 and \vec{v}_1 . Without detailed steps, we can show that the entire computational complexity of all those processes require only $(3N^2 + 30N/2 + 46)$

real multiplications and $(N^2 + 16N + 16)$ additions, where N is the number of Stokes vectors stacked in $\mathbf{S}_{3,N}$. As N is well below 100 and only 1 symbol every 6 on average are converted to Stokes, those numbers confirm the low complexity.

In the power screening step in Step 1, depicted in Fig. 3–3, one can find a radius P^{th} that leads to a specific power discrimination ratio, defined as the total number of Stokes points over the number of Stokes points outside the red sphere. Fig. 3–3 depicts the results for a ratio of 100. Signals with higher Signal to Noise Ratio (SNR) need a smaller P_{th} to maintain a power discrimination ratio, and vice versa. As propagation and therefore amplification adds noise to the signal, P_{th} has to be slightly increased with distance in order to maintain a ratio and consequently a good estimate of \hat{p}_{hole} . Observing the green “caps” on top of Fig. 3–3 (a), (b) and (c), we understand that signals with higher SNRs exhibit 4 clusters that are more regrouped, providing a better estimate of \hat{p}_{hole} . As higher order QAM inherently require higher SNR for the same BER, this new blind PR technique works very well for higher order QAM modulation formats, which is not true for common blind adaptive PR techniques like CMA, because their inherent cost function are not designed for high order QAM [15].

A major benefit of this blind Stokes space polarization recovery compared to the iteratively adapting CMA algorithm is the avoidance of the singularity problem of CMA [79], where 2 severely rotated inputs can converge to the same output. As we force the rotation to be unitary and as we bring back the average SOP around the $\{\vec{A}_2, \vec{A}_3\}$ plane, the \hat{x} and \hat{y} outputs present independent information. Once \mathbf{U}_{ROT} is estimated, it initializes the subsequent DD-LMS filter, that also contains a built-in digital phase locked loop (PLL). As derived in Eq. (3.9), the proposed algorithm compensates for the “zeroth” order PMD, or said differently for the global polarization rotations occurring on the signal during propagation. The assumption that the cumulative first order PMD in $\mathbf{U}_c(z, \omega)$ at the receiver is small depends on the fiber type and more precisely on its PMD_Q value. Newer fiber types show very little PMD, with maximum PMD_Q of $0.04 \text{ ps}/\sqrt{\text{km}}$. Even after 6400 km, they exhibit a very small mean differential group delay (DGD) of around 3.2 ps . Even at a high symbol rate of 28 GBaud, this value represents less than 10% of the symbol duration. Much greater instantaneous DGD can be tolerated for digital optical coherent receivers. As the fiber used on the experimental setup has a maximum PMD_Q of $0.04 \text{ ps}/\sqrt{\text{km}}$, experimental results do not assess the proposed polarization recovery technique over large PMD with respect to symbol duration. Consequently, we will present numerical simulations in which the algorithm is tested

over varying DGD and sampling offsets in Section 3.1.5.1, after experimental results, and show the accuracy and speed of the algorithm over much larger DGD.

3.1.4 Comparison Metric and Experimental Test Bed

We compare our novel blind polarization recovery algorithm in Stokes space against the CMA [11] scheme. We reason the choice to compare against a single other algorithm because CMA is well known and widely used, but also because most other blind algorithms do not converge as well. For example, the radius-directed algorithm [12] relies on decisions based on the amplitude, and it has been found that this algorithm does not converge well in practice [13]. A recently proposed cascaded three-modulus blind equalization algorithm has shown worse convergence properties than the CMA [14] and can only be used after a channel estimate has been obtained. The same holds true for decision-directed algorithms operating on non-converged data, which depend on most decisions being correct. In [15], the convergence rate of both CMA and the Independent Component Analysis is improved compared to the standard CMA by applying computationally expensive algorithms to the data. One method runs three independent computations of blind polarization recovery using different matrices initializations, similar to [65]. The other method performs the gradient descent based on all the observed symbols from system startup up to cumulative time k , instead of on a single symbol for standard CMA. ICA converges faster but is always more costly than CMA, and even more so for higher order QAM like 16-QAM [80]. We will consequently compare our technique against the easily implementable CMA algorithm with the objective of switching to DD-LMS+PLL as quickly as possible for steady-state operation. We claim that the time saved to get to convergence translates into saves in computational expenses, as an extra computational effort has to be spent during the additional time required for blind adaptive techniques to converge. During this time, 2 complex cost function errors have to be computed for every excessive required symbol, where computing the error of CMA is more computationally demanding than that of LMS.

Fig. 3-4 depicts in Stokes space the convergence location of three blind equalization algorithm schemes for DP-16QAM signals, namely (a) the CMA, (b) the Radius Directed Equalizer [12] and (c) Multi-Modulus Algorithm [81].

Although considered blind, the MMA algorithm is phase dependent as its cost function minimizes the distance between a square instead of a circle(s) as is the case for CMA (RDE). By comparing the figures in Fig. 3-4 against the constellations in Stokes space in Fig. 3-1 (b) for

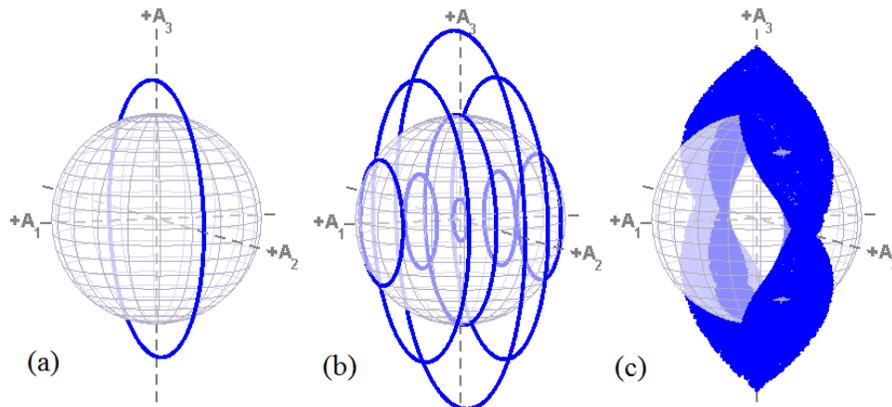


FIGURE 3–4: Representation on the Poincaré sphere of 3 different polarization tracking algorithm: (a) CMA for DP-16QAM ($R = \sqrt{0.66}$), (b) RDE for DP-16QAM with radius to nearest symbols of $\sqrt{0.1}$, $\sqrt{0.5}$ and $\sqrt{0.9}$ (c) MMA with $\sqrt{0.41}$

DP-16QAM, we understand that the “thickness” in the $\pm\hat{A}_1$ dimension as well as the presence of multiple valid rings driven by the cost function have an impact on the speed of convergence of the blind methods for varying initial signal SOP. Even if all those convergence rings align in planes of norm \hat{A} , their multiple presence in several planes, e.g. $A_1 = 0$ (3 rings), $A_1 = \pm 0.4$ (2 rings) and $A_1 = \pm 0.8$ (1 ring), slows the speed at which a rotated SOP is blindly forced back to the desired $A_1 = 0$ plane. On the contrary, the single ring of CMA forces the SOP to fall more quickly to the $A_1 = 0$ plane, while leaving it freely rotating within that plane. A similar conclusion due to the thickness of the MMA convergence region in Fig. 3–4(c) explains its slower blind adaptation speed compared to CMA. Hence, as highlighted in [13], we observed a reduced robustness in filter pre-convergence using the multi-ring CMA and therefore opted for the classic CMA.

3.1.4.1 Comparison Metric

We will use two metrics to assess our method. Firstly, we will compare the convergence time of our technique against that of the CMA having a built-in PLL. We define the convergence time as the number of symbols required by a blind algorithm to switch from blind polarization recovery to DD-LMS+PLL. For our technique, the convergence time is the total number of received symbols before \mathbf{U}_{ROT} can be estimated and used to initialize the taps of the subsequent DD-LMS+PLL. For the blind CMA adaptive technique, the convergence time is the number of symbols required before the average number of bits in error is less than 5% on each polarization, evaluated over a total number of 1000 received bits (500 per polarization). We decided on a BER threshold of

5% as we experimentally observe consistent post-convergence of the DD-LMS filter for any QAM order using this value. Yet this BER is a high bound and is higher than the BER threshold of some soft-decision FEC [82, 83]. As an example, for DP-4QAM, -16QAM and -64QAM formats, the observation window, in symbols, onto which the BER is calculated is 250, 125, and 84, respectively. These short window lengths allow switching to LMS as soon as feasible, preventing to stay in CMA mode for a longer, undesired period. When the BER reaches such a value, the 2-by-2 multiple input-multiple-output (MIMO) filter switches its method of adaptation to DD-LMS. For smooth transition from CMA to DD-LMS, the same PLL is used inside both polarization recovery schemes. To decide if we update the phase of the PLL or not, we use a criterion based on the signal power threshold on each polarization, as presented in [84]. To show in a different way the quickness of our new method, we will also compare the mean number of bits in errors using only the first 20×10^3 symbols. Again, we compare our Stokes space blind polarization recovery technique against that of CMA. This will show that our matrix estimation in Eq. (3.10), including the important initial phase estimate $e^{i\delta}$, yields an initial BER that is much smaller than what obtained using the common CMA algorithm.

3.1.4.2 Experimental Test Bed

We compare performance of our new method against CMA for 3 different modulation formats: DP-4QAM at 28 Gbaud, DP-16QAM at 28 Gbaud and DP-64QAM at 7 Gbaud. This will allow proving experimentally that the technique works for any square M-QAM order. The experimental test bed is depicted in Fig. 3-5. We use an Arbitrary Waveform Generator(AWG) to generate the multi-level signals applied to a Dual-Parallel Mach-Zehnder (DPMZ) modulator that modulates a CW light out of an external cavity laser (ECL) that has a linewidth smaller than 100 kHz. Dual-Polarization is emulated using Polarization Beam Splitters/Combiners and an optical delay line. The single carrier optical signal is then boosted to the desired launch power; we launch at -2.2 dBm for these experiments. The signal then propagates inside an optical recirculating loop which contains 4 spans of 80 km of Corning SMF-28e+ LL[®] and 4 in-line EDFAs. No optical dispersion compensation modules are present. On the receiver side, the signal gets filtered, amplified and re-filtered before hitting the optical coherent receiver front end. The local oscillator laser is the same type as the transmitter. The Dual-Polarization 90° optical hybrid is from Kyria[®], the balanced photodiodes are BPDV2020R's from U²T[®] and the real-time oscilloscope is Agilent[®]'s

DSOX93204A sampling at 80 GSa/s with a 3dB bandwidth of 33 GHz. The recirculating loop allows propagation by steps of 320 km.

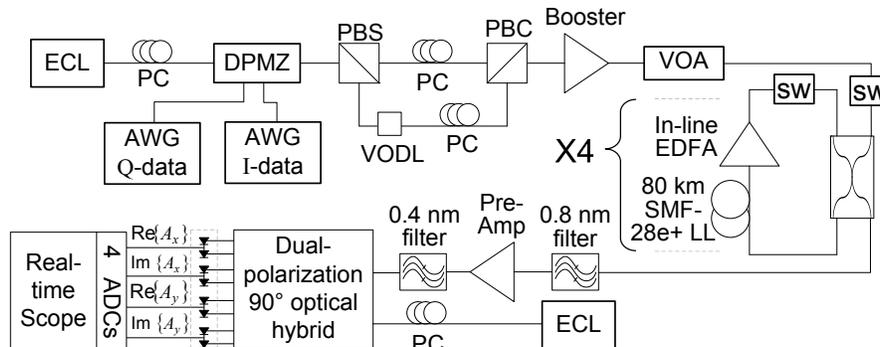


FIGURE 3–5: Experimental Test bed

3.1.5 Results

We present in this section the convergence speed and the initial BER for DP-4QAM in Fig. 3-6, DP-16QAM in Fig. 3-7 and DP-64QAM in Fig. 3-8. We compare the BER of the first 20×10^3 symbols when a blind CMA+PLL filter starts operating on the received data, after resampling at $T/2$ and CD removal, against our Stokes blind recovery method. As our technique has to discard or buffer incoming data before estimating the polarization derotation matrix whereas CMA outputs decisions as soon as it starts processing, we consider the discarded and buffered Jones symbols when computing the BER using the Stokes space method, for a fair comparison of an equal number of samples starting at the same time. Those initial data are polarization entangled, off-phase and freely moving in both polarization and phase, therefore are very bad data for the slicer. For rapid blind convergence of the CMA algorithm, we used a relatively large adaptation coefficient of $\mu = 10^{-3}$ for all QAM orders. Even when considering those discarded/buffered symbols, the initial BER of the Stokes space blind recovery method gives better performance than CMA.

For DP-4QAM format, some traces had their initial SOP at the receiver fortuitously aligned to that of the PBS axis of the coherent receiver. For example, one capture after 1600 km shows that even if the minimum of 250 symbols were needed in CMA+PLL mode before switching to LMS+PLL (from the need to have at least 1000 bits collected to evaluate the BER), the initial BER of the first 20K symbols is still less than that of the Stokes space recovery method that required only 38 symbols to switch to LMS+PLL. The reason is that the first 38 symbols had both

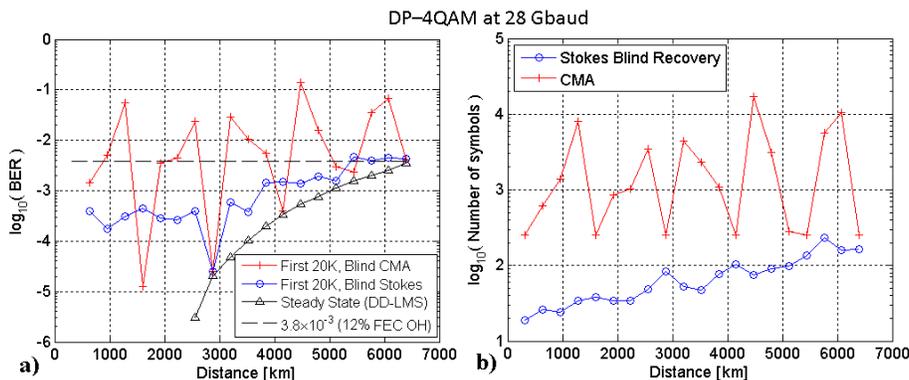


FIGURE 3–6: (a) BER vs Distance of first 20K symbols for DP-4QAM format at 28 Gbaud comparing CMA and Stokes space recovery. (b) Number of symbols required before switching filter from blind to DD-LMS+PLL

their phase and polarization tracked in CMA+PLL whereas they are left wandering in the Stokes space recovery algorithm. It is important to note that if we would assume that buffered incoming symbols could be post-processed by the polarization/phase derotation matrix of Eq. (3.10) after its estimation, before feeding them to the slicer, the initial BER of the Stokes space algorithm would be even smaller, and those rare cases where CMA performed better than Stokes for DP-4QAM would most probably be reversed. The sawtooth look of the required number of symbols for CMA in Fig. 3–6(b) for largely varying noise levels exhibits the strong dependence of the speed of convergence of CMA with respect to the received SOP and no dependence on the SNR. Strongly varying convergence speeds are observed at both large and little noise levels. The Stokes space method, however, shows no dependence on the input SOP, and only a slight dependence on the noise level depicted by the slope. This slope is caused by the voluntary increase of the power discriminator P_{th} with distance in order to keep a power discrimination ratio and consequently a good estimate of \hat{p}_{hole} . For DP-4QAM, we let P_{th} increase as $P_{th} = 1 + \text{LoopNo}/40$, where LoopNo is the loop count. This means that at 6400 km where we reach $BER = 3.8 \times 10^{-3}$, $P_{th} = 1.5$ which is well below $P_{th} = 2.18$ for a ratio of 100 as in Fig. 3–3(a). In fact, $P_{th} = 1.5$ represents a theoretical power discrimination ratio of ~ 5.8 for DP-4QAM at $BER = 3.8 \times 10^{-3}$. As roughly 200 symbols were needed for \mathbf{U}_{ROT} 's estimation after 6400 km, an average of 35 symbols were stacked in $\mathbf{S}_{3,N}$ and used for the 1st matrix estimate, and only $3C = 12$ or a little bit more were kept for the 2nd matrix estimate and the phase estimate of \mathbf{U}_{ROT} . The BER of the first 20×10^3 symbols shows the accuracy of our Stokes space estimate in comparison to the steady state DD-LMS+PLL

curves which represents the lowest bound for the initial BER of any method taken to get to steady state.

As demonstrated in the figures, the average initial BER of the Stokes recovery method can be smaller by an order of magnitude or more compared to that of CMA. This shows the accuracy of our estimation of \mathbf{U}_{ROT} : \hat{p}_{hole} , the angle $\phi/2$ in the 2nd matrices in Eq. (3.10) and the absolute phase $e^{i\delta}$ are well estimated. As both the PLL and the MIMO filters start with good initial conditions, the locking of those processes happens more quickly. For CMA, if the initial polarization entangling is very severe (\hat{p}_{hole} pointing around $\pm\hat{A}_3$ or $\pm\hat{A}_2$) the algorithm will take more time to adaptively blindly untangle the information on both polarizations [65].

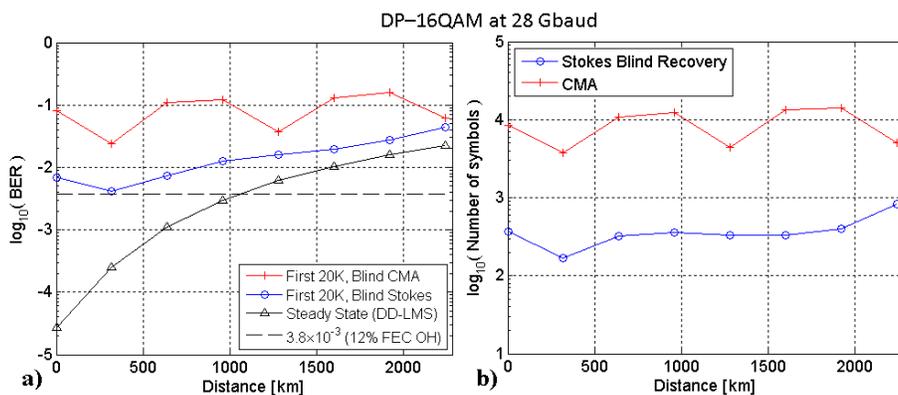


FIGURE 3-7: (a) BER vs Distance of first 20K symbols for DP-16QAM format at 28 Gbaud comparing CMA and Stokes space blind recovery. (b) Number of symbols required before switching from blind to DD-LMS+PLL.

The independent PLL's working on \hat{x} and \hat{y} output polarizations operate on providing an absolute phase to the Jones vectors on a per symbol basis and help locking the SOP on the $\{\vec{A}_2, \vec{A}_3\}$ plane to the \hat{A}_2 and \hat{A}_3 axis, preventing spinning within the plane.

For any square QAM of 4, 16 or 64-ary, after any distance, our experimental results show that the number of required symbols needed to estimate \mathbf{U}_{ROT} before switching to DD-LMS never exceeds 830. For DP-16QAM, at a BER of 2.2% after 2240 km, a total of only 829 symbols impinging on the receiver were needed to estimate \mathbf{U}_{ROT} and start operating in DD-LMS. At 28 Gbaud, this is a very short 29.3 ns required for blind polarization convergence, which is known to be the process taking the most time to converge [65]. For the DP-64QAM format at 7 Gbaud after 640 km at a BER of 1.4%, only 554 symbols were needed, converting into only 79.1 ns. Finally, for DP-4QAM, after 6400 km giving a BER of 0.38%, only 165 symbols were required before successfully

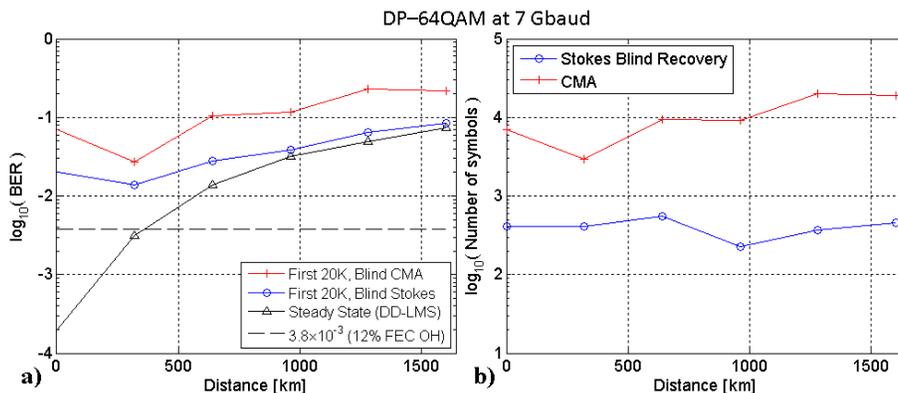


FIGURE 3–8: (a) BER vs Distance of first 20K symbols for DP-64QAM format at 7 Gbaud comparing CMA and Stokes space blind recovery. (b) Number of symbols required before switching from blind to DD-LMS+PLL.

switching to DD-LMS+PLL. This converts to a mere 5.9 ns. We can compare those results with recently reported convergence time of blind CMA algorithm applied on DP-4QAM signals, where modification were applied to the CMA in order to decrease its convergence time. In [65] they report a mean convergence time of 40 ns, with a best case of 20 ns and a worst case of 280 ns. In another work [67], the CMA algorithm achieves blind recovery in 200 ns, or 11200 symbols. By comparing our convergence time results, we demonstrate the rapidity of convergence of our novel blind Stokes space polarization recovery technique.

For the DP-16QAM and -64QAM, the initial BER within the first 20×10^3 samples is always better; around 3 times lower for DP-64QAM and even more for DP-16QAM. In (a) of Figs. 3–6 to 3–8, we can appreciate how close the curves of the Stokes space method are to the lower bound of steady-state operation, proving the accuracy of the method to estimate \mathbf{U}_{ROT} .

3.1.5.1 Testing Over PMD

We present in this subsection numerical simulations where the proposed polarization derotation method is assessed over varying instantaneous DGDs, assumed constant for the short period required by the algorithm to find the derotation matrix \mathbf{U}_{ROT} . The proposed SOP recovery method is also assessed over different DGD when the sampling instances of the ADCs don't align with the center of the symbols' location, for more realistic received waveforms. We call this sampling offset with respect to the symbols' location the Common Group Delay (CGD). This can also be true for two-fold oversampling as is the case for processing the received waveforms in our simulations.

Mathematically, the numerical model is described as follows

$$\begin{bmatrix} \tilde{A}(\omega)_{x,out} \\ \tilde{A}(\omega)_{y,out} \end{bmatrix} = \underbrace{\left[\underbrace{H_{Tx,ISI^{-1}}(\omega)}_{\text{Rx Filter}} \underbrace{e^{i\omega\kappa}}_{\text{CGD}} \underbrace{\mathbf{U}_{rand,2} \mathbf{U}_{rand,1}^\dagger}_{\text{Rotation}} \right]}_{\text{Receiver}} \underbrace{\left[\underbrace{\begin{bmatrix} e^{i\omega\tau/2} & 0 \\ 0 & e^{-i\omega\tau/2} \end{bmatrix}}_{\text{DGD}} \right]}_{\text{Link}} \underbrace{\mathbf{U}_{rand,1}}_{\text{Rotation}} \begin{bmatrix} \tilde{A}(\omega)_{x,in} \\ \tilde{A}(\omega)_{y,in} \end{bmatrix} \quad (3.11)$$

where τ is the DGD and κ is the CGD. $\mathbf{U}_{rand,1}$ and $\mathbf{U}_{rand,2}$ are random rotation matrices as defined in [70]. The two orthogonal eigenvectors of $\mathbf{U}_{rand,1}$ form the basis that will observe a positive and negative DGD of $\tau/2$, and $\mathbf{U}_{rand,2}$ randomly rotates the entire resulting waveform. The waveforms $\tilde{A}(\omega)_{x,in}$ and $\tilde{A}(\omega)_{y,in}$ are respective Fourier transforms of $A(t)_{x,in}$ and $A(t)_{y,in}$. Each $A(t)_{x/y,in}$ is the resulting output of two Pulse Pattern Generators added in quadrature and operating at 28 Gbaud, having a Gaussian impulse response and an output risetime from 20% to 80% of 15 ps. The quadrature addition generates the QPSK signal, on both polarization, hence DP-4QAM format. Due to the risetime, some ISI is present at the transmitter. The required receiver filter to remove transmitter-induced ISI is computed and applied for all captures as the matched filter to apply on the receiver side. Each $A(t)_{x/y,in}$ are initially oversampled at $T/8$ in order to accurately generate $A(t)_{x/y,out}$. Complex white Gaussian noise is added to $A(t)_{x/y,in}$ such that after the transmitter ISI mitigating filter $H_{Tx,ISI^{-1}}$, the bit error rate is at the threshold of 3.8×10^{-3} . Once $A(t)_{x,out}$ and $A(t)_{y,out}$ are generated at $T/8$, perfect ADCs picking 1 sample every 4 mimic two-fold oversampling at $T/2$. Next, the blind PR algorithm presented in Section 3.1.3 starts operating until a unitary rotation matrix is estimated and used to initialize the central tap of a subsequent 25-tap DD-LMS filter.

In the first case, CGD is zero and DGD is varied from 0 to 0.9 T , where T is the symbol duration. This translates to up to 32 ps of DGD at 28 Gbaud; 10 times the mean DGD after 6400km for fibers of $PMD_Q = 0.04 \text{ ps}/\sqrt{\text{km}}$ or 2 times for fibers of $PMD_Q = 0.2 \text{ ps}/\sqrt{\text{km}}$ like the G.652D fiber. In the second study we set the sampling offset to a large value; one third of the symbol duration. For this case, we vary the DGD from 0 to 0.3 T , up to 10.7 ps at 28 GBaud.

Using the Monte Carlo method, for every different combinations of DGD and CGD, 400 random DP-4QAM waveforms with random noise sources and random $\mathbf{U}_{rand,1}$ and $\mathbf{U}_{rand,2}$ were generated to assess the proposed polarization derotation scheme. Then, the proposed blind polarization derotation method starts operating on $|A(t)_{out}\rangle$, estimates \mathbf{U}_{ROT} and uses it as the center tap of the subsequent adaptive DD-LMS filter. The latter then processes data and slowly adapts until

removal of all ISI coming from DGD, CGD and remaining polarization cross-talk due to the slightly imperfect evaluation of the \mathbf{U}_{ROT} . When all ISI is removed, only white Gaussian noise is present and steady state is reached. The unitary matrix \mathbf{U}_{ROT} that our proposed algorithm finds is an estimate of $\mathbf{U}_{\text{rand},2}^\dagger$. The results are presented in the following Fig. 3–9.

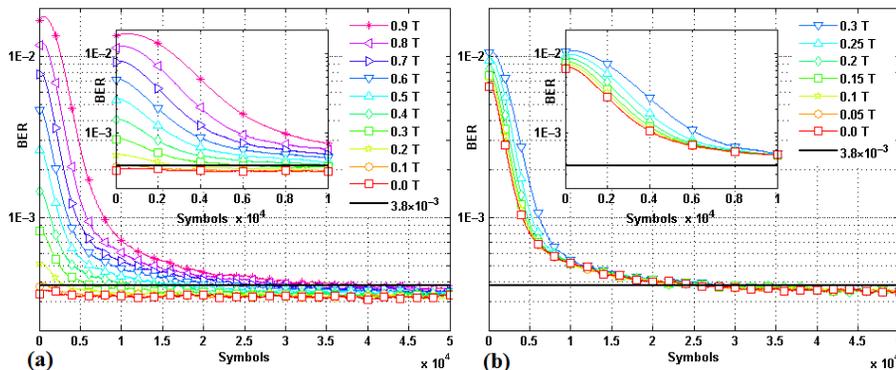


FIGURE 3–9: System performance for varying DGD when the Common Group Delay is (a) 0, and (b) $T/3$. Insets show BER zoom-in for symbols 0 to 10^4 .

We can see the direct impact of DGD on system performance: higher DGD maps as higher bit error rate. For the curves in Fig. 3–9, the BER was filtered by a first order IIR low-pass filter with parameter $\alpha = 0.999$, mimicking a BER computed over a window length of $1/(1 - \alpha) = 1000$ symbols. The initial BER estimate, at index $n = 0$, is the BER if \mathbf{U}_{ROT} was found to be exactly $\mathbf{U}_{\text{rand},2}$ and if no adaptive DD-LMS filter was applied to $A(t)_{x/y,out}$. We can see in the inset of Fig. 3–9(a) that starting at DGD of $0.7 T$, the estimated \mathbf{U}_{ROT} starts to moderately differ from the optimum matrix $\mathbf{U}_{\text{rand},2}$, resulting in an initial BER that is slightly worse than if $\mathbf{U}_{\text{rand},2}$ was applied. This is depicted by the subtle increase of the windowed BER. The accuracy of our method is confirmed by this tenuous different in initial performance from \mathbf{U}_{ROT} to $\mathbf{U}_{\text{rand},2}$ and by convergence attained via a decision directed scheme. Systematic successful \mathbf{U}_{ROT} estimation and post-convergence was observed for DGD as high as $0.9 T$, but higher DGDs failed to converge. The results when $\text{CGD} = T/3$ are shown in Fig. 3–9(b). For this case, the algorithm systematically post-converged to steady state for DGDs upto $0.3 T$.

The maximum DGD thresholds of $0.9 T$ for $\text{CGD} = 0$ and $0.3 T$ for $\text{CGD} = T/3$ can readily be explained. The absolute sum of DGD and CGD applied to one axis is limited to be smaller than $0.5 T$, half the symbol's duration. As an example, for the first case, a $\text{DGD} = 0.9 T$ gives a maximum delay on one axis of $0.45 T$. However when $\text{CGD} = T/3$, the maximum allowed DGD is

$0.3 T$ giving a maximum delay on one axis of $T/3 + 0.3 T/2 = 0.483 T$. As one would imagine, any non data-aided polarization untangling process would require a total delay on any axis with respect to the central position of symbols to be less than half the symbol duration for successful subsequent blind adaptation. After such period, larger DGD is acceptable and tracking will solely depend on the tracking speed and method. Otherwise, if the DGD is too large during the blind process, taps of one output dimension of the adaptive 2-In-2-Out filter are likely to converge to a state that offsets the output symbols by an integer number of symbols with respect to the other output dimension.

The results of Fig. 3-9(b) shows the interesting feature that if the sampling instants are off as much as $\pm T/3$ with respect to the symbols' center position, an instantaneous DGD of $0.3 T$ is still tolerated for our proposed method to successfully find and derotate the SOP and allow direct operation in decision directed mode. Consequently, DGDs up to $0.3 T$ can support CGDs in the wide range of $-T/3$ to $+T/3$, very close to the full range of $-1/2T$ to $+1/2T$.

Finally, the speed of convergence for various DGD is assessed for the proposed blind polarization untangling method. We show in Fig. 3-10 the average required number of symbols for varying DGD for both offset sampling cases.

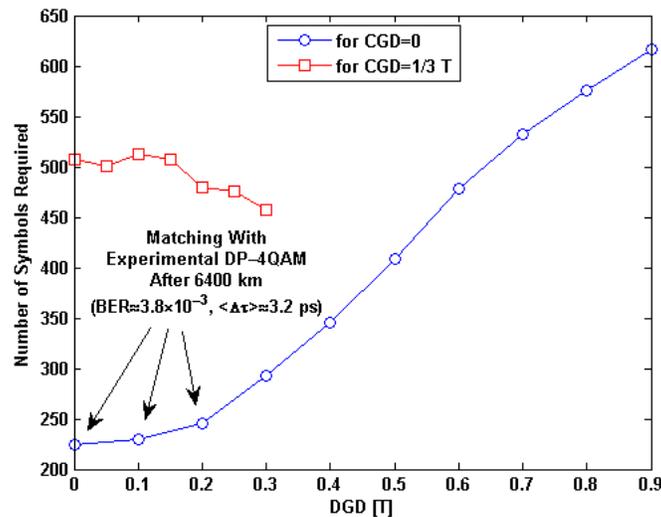


FIGURE 3-10: Required number of symbols to find Urot over varying DGD, for CGD = 0 and $T/3$

We can observe that with increasing DGD, an increased number of symbols are required before successfully switching to DD-LMS. This is explained by the fix power threshold value P_{th} used in

these simulations. To match the P_{th} used in experiments for DP-4QAM after 6400 km where $BER \approx 3.8 \times 10^{-3}$, we used $P_{th} = 1.5$. Using this fix value, we can show that the power discrimination ratio increases with increasing DGD. This explains the monotonically increasing trend of the required number of symbols over DGD. We observe that the required number of symbols is systematically very low even for a large DGD of $0.9 T$ at around 600 symbols, compared to the possible 10^4 symbols required by the CMA process that depends on the SOP alignment as shown in Fig. 3–6(b) when applied to low PMD fiber. Figures 3–9 and 3–10 show the accuracy and speed of the proposed polarization untangling process over largely varying DGD.

Finally, the required number of symbols obtained numerically for instantaneous DGDs smaller than $0.2 T$ match the value obtain for real propagation shown in Fig. 3–6(b). Simulations give between 200 and 245 required symbols on average for DGDs between 0 and 7.1 ps where experiments show around 200 symbols for the same modulation format and noise level with mean DGD of $0.04ps/\sqrt{km} \times \sqrt{6400 km} = 3.2 ps$. Consequently, similarity of the experimental results with simulation over comparable range of PMD confirms the mathematical model used and validates results for larger DGD. The proposed method would successfully operate for fibers with higher PMD_Q like the G.652D. From the Maxwellian distribution of DGD with the sole knowledge of DGD_{MEAN} , we can be compute the probability that the instantaneous DGD be smaller than $0.9 T$ on G.652D fiber after 6400 km with 28 Gbaud signals giving $DGD_{MEAN} = 0.448 T$: a probability of $1 - P(DGD > 0.9/0.448 \times DGD_{MEAN}) = 98.4\%$.

3.1.6 Conclusion

We have presented a novel method for blind estimation of the SOP and polarization rotation for single carrier channels and coherent receivers. This method can be used either on regular system startup or for fast switching burst mode receivers. We have shown experimentally that the algorithm outperforms the convergence time for any DP-*square*-QAM format by about an order of magnitude compared to the standard CMA algorithm. The numbers of symbols required to blindly derotate the received SOP before the self-adapting equalizer can operate in decision-directed LMS mode is used to assess convergence speed. Unlike CMA, this blind Stokes space polarization recovery algorithm is completely independent of the input SOP of the signal. We compared experimental results of convergence time and initial BER when the CMA and when the Stokes space technique are used to blindly untangle the receiver SOP for 3 modulation formats of DP-4QAM, -16QAM

and -64QAM. After propagation over SMF 28e+ LL fiber, we obtain convergence within 5.9 ns at BER=0.38% after 6400 km for DP-4QAM, 29.3 ns at BER=2.2% after 2240 km for DP-16QAM and 79.1 ns at BER=1.4% after 640 km for DP-64QAM. The initial BER of the first 20×10^3 symbols when the Stokes space technique is used is always much closer to the steady-state operation than when CMA is used for tap adaptation, for DP-16QAM and -64QAM formats. For DP-4QAM, some fortuitously well aligned input SOP on system startup can give better initial BER when using CMA. However, the Stokes space method still outperforms CMA even for DP-4QAM as it is robust to any input polarization alignment. Finally, we simulated PMD testing on our Stokes space polarization recovery technique to assess its robustness over larger PMDs than what provided by the experimental test bed. We demonstrated the accuracy and speed in obtaining the derotation matrix over DGDs as large as 90% of the symbol duration for perfect sampling offset and as large as 30% when the initial central tap of the filter is off by one third of a symbol duration with respect to the symbols' center location. The speed of convergence using the blind Stokes method on large PMD is still generally smaller than that of CMA on little PMD.

3.2 Constant Power 4D Modulation on a 32-Symbol Set

In this second section covering the topic of optical communications for coherent long-haul systems, we propose a new modulation format that provides 5 bits of information per recovered symbol while maintaining constant the total optical power at symbol locations. The format allows passive mitigation of nonlinear effects occurring during propagation in single mode fibers.

3.2.1 Introduction

Standard optical fibers exhibit a nonlinear response that introduces a power-dependent phase retardation into the passing wave [69], causing a multitude of effects like self-phase modulation (SPM), cross-phase modulation (XPM) and cross-polarization modulation (XPolM). Undesired phase modulations from nonlinear impairments get converted to amplitude fluctuations during propagation [85], further deteriorating the signal quality. The greater the instantaneous power variation of the optical waveform, the greater the nonlinear impairments.

Modulation formats using only phase and polarization manipulation help reduce the nonlinear impact and increase propagation distance. Polarization Shift (PS)-Quadrature Phase Shift Keying (QPSK) proved to be a good candidate for ultra long-haul transmission allowing greater launch power and increased reach [86, 87]. With the advent of optical coherent receivers, polarization division multiplexing (PDM) rapidly gained interest, as it allows modulation on a 4 dimensional (4D) space and an increase in the binary transmission rate. As an example, a recent publication analyses legacy formats of on-off keying and binary phase shift keying (BPSK) when multiplexed on 4 states of polarization as a novel modulation format employing PDM [88]. Full dual polarization IQ modulators allow the study and use of novel 4D signaling schemes [87]. Formats of higher dimensions were suggested lately, as in [89] where the authors propose a signal constellation design algorithm generating multidimensional signaling schemes that increase the number of dimensions above 4D, where spatial multiplexing in few mode fibers is employed to increase the number of degrees of freedom.

PDM offers drawbacks, amongst which the inherent loss of 3 dB signal power per polarization to maintain the total optical power. Some formats like Polarization Shift Keying (PolSK) rely solely on modulating the state of polarization to imprint information. Such format benefits from the constant power property, but lack bits per symbol from the smaller set size N_{SOP} of possible SOP's, $\log_2(N_{SOP})$.

Higher order modulation formats are widely employed as a way to increase the data throughput by allowing a larger number of different amplitudes and phases. However, such formats inherently exhibit larger power variations from symbol to symbol, accentuating nonlinear effects, especially when polarization multiplexed is used [1]. Higher order modulation formats that do not exhibit larger power variation are of great interest in order to meet the growing need for reach and data throughput.

In this section, we present a new modulation format for single mode fibers relying on a mixture of PolSK and PSK. We entitled this format 8PolSK-QPSK. The format provides 5 bits of information per symbol where each orthogonal polarization exhibits an 8QAM constellation-like Dual Polarization-8QAM format, itself providing 6 bits per symbol. One of the main characteristics differentiating the 2 formats is the waveform independence on orthogonal states: instead of having completely independent 8QAM formats on each polarization as is the case for DP-8QAM, the proposed 8PolSK-QPSK format couples part of the information on each polarization such that each symbol in the dual polarization symbol family share the property of having equal total optical power. We show both analytically and experimentally that modulating only in polarization and phase allowing a symbol set of equal power attenuates the impact of the optical Kerr effects. To tackle both intra- and inter- channel nonlinear effects, we experimentally transmit 7 channels, all modulated in either DP-8QAM or 8PolSK-QPSK, and study the performance of the central channel. For a fair format comparison, the bit rate per channel is kept constant at 129 Gbit/s. To satisfy this constant bit rate the symbol rate of 8PolSK-QPSK has to be higher than that of DP-8QAM by 20%. Besides this modest symbol rate difference, the transmitter architecture and the transmitter implementation penalty are the same for both formats.

The remaining of this section is organized as follows. First, the proposed 8PolSK-QPSK modulation format is presented in Section 3.2.2 and compared to the DP-8QAM format. The advantage of combining PolSK and QPSK as a modulation scheme are highlighted. We also describe a novel minimum distance detection scheme applicable to any constant power format that is less computationally demanding than the conventional complex minimum distance detection. Section 3.2.3 presents an analysis of the nonlinear Manakov-PMD equation [72] and shows how the modified DP-8QAM format can mitigate nonlinear effects. Propagation simulations using both formats demonstrate where and how the 8PolSK-QPSK format suffers less from the Kerr effect. Subsequently, Section 3.2.4 presents the experimental test bed used to compare performance of both

formats. In this section, we also introduce a way to generate 8PolSK-QPSK format with a generic delay-and-add dual polarization emulator and a controlled single polarization (SP) 8QAM signal. Finally, system performance using both formats is presented in Section 3.2.5, where we show performance improvements using the proposed format, including an increased propagation distance of 34%, or 975 km, compared to DP-8QAM, and a reduction of the required OSNR for a BER of 1.4×10^{-2} of 0.95 dB. We conclude in Section 3.2.6.

3.2.2 Proposed 8PolSK-QPSK Modulation Format

The proposed modulation scheme is very similar to the polarization multiplexed 8QAM format. For DP-8QAM, each polarization carries independently an 8QAM modulation format. 8QAM formats can be thought of as a QPSK format (2 bits/symbol) from a symbol family $\{1, i, -1, -i\}$ on top of a binary (1 bit/symbol) format of symbol family $\{a, be^{i\pi/4}\}$, giving a total of 3 bits/symbol/polarization. Of course, both QPSK and binary symbols are independent on each polarization for DP-8QAM. The binary symbol's amplitude corresponds to the 2 possible rings as depicted in Fig. 3-11. For a total mean signal power of 1, the total mean power per polarization is 0.5 and for equidistant constellation points, the 2 radii are $a = 0.4597$ and $b = 0.8881$. Consequently, for DP-8QAM formats, the total optical power jumps from symbol to symbol to any of the 3 possible power levels: $a^2 + a^2 = 0.4226$, $a^2 + b^2 = 1$, or $b^2 + b^2 = 1.5774$.

In the proposed 8PolSK-QPSK format, the binary amplitude bit is not independent on each polarization. This constitutes the main difference between DP-8QAM and the proposed format. In 8PolSK-QPSK, if the \hat{x} -polarization shows a symbol in the inner ring (binary symbol a), the simultaneous symbol on the \hat{y} -polarization has to be a symbol in the outer ring (symbol $be^{i\pi/4}$), and vice versa. This constraint removes one bit of freedom from DP-8QAM's 6 bits, providing 5 bits per symbol for the format. As one can compute, enforcing different amplitude on each polarization restricts the total optical power to always be $a^2 + b^2 = 1$ for any given symbol, relieving DP-8QAM from its large 3-level power fluctuation

In the proposed 8PolSK-QPSK format, we only modulate information on the polarization and phase of an optical waveform. The advantage of combining polarization and phase shift keying is found in the analysis of the nonlinear (NL) Schrödinger (NLS) equation. A well-known, simplistic way to study the Kerr nonlinear effects in a single mode fiber is to neglect all impairments but attenuation and Kerr nonlinearity in the NLS equation and study the equation governing the field

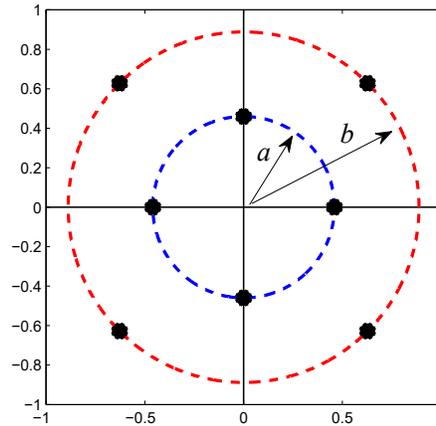


FIGURE 3-11: Star-8QAM format. In DP-8QAM, the 2 polarizations can both independently be at amplitude a or b . For 8PolSK-QPSK format, that independence is removed

evolution during propagation. If we call the optical field at $z = 0$, $|u(z = 0, t)\rangle = |u_o\rangle$, the field $|u(z, t)\rangle$ evolves following $|u(z, t)\rangle = |u_o\rangle e^{i\phi_{\text{NL}}(z, t)}$, where $\phi_{\text{NL}}(z, t) = \gamma(8/9) \int_0^z \langle u_o | u_o \rangle e^{-\alpha z'} dz'$ [85]. The SPM-induced spectral broadening, or the NL impairment, is a consequence of the time dependence of the nonlinear phase $\phi_{\text{NL}}(z, t)$ [85]. If the temporal variation of the total power $\langle u_o | u_o \rangle$ is reduced, the strength of the NL impairments during propagation is also reduced. When modulating in polarization, the total power does not change as PolSK relies on applying a unitary rotation $\mathbf{R}(t)$ to the input field $|u_{\text{CW}}\rangle$ such that $|u_o\rangle = \mathbf{R}(t) |u_{\text{CW}}\rangle$, where $\mathbf{R}^\dagger \mathbf{R} = \mathbf{I}$. Additionally, modulating the phase of a continuous wave (CW) by applying $e^{-i\phi(t)} |u_{\text{CW}}\rangle$ also maintains the power. The combination of both formats, $|u_o\rangle = e^{-i\phi(t)} \mathbf{R}(t) |u_{\text{CW}}\rangle$, give a total power of $P_o(t)$, and the NL phase noise ϕ_{NL} grows as $\gamma(8/9)P_o(t)(1 - e^{-\alpha z})/\alpha$. $P_o(t)$ is the initial launched power waveform, that depends on the pulse shaping at the transmitter. For an intersymbol interference (ISI)-free pulse shape, $P_o(kT)$, $\forall k$. Consequently, formats having a symbol set sharing constant power, as is the case for formats combining polarization and phase shift keying, take advantage of the smaller temporal variation of ϕ_{NL} to reduce signal deterioration from NL effects.

There is a drawback, however, when solely using polarization and phase as a way to imprint information on a waveform. The 8PolSK-QPSK format is a 4D, constant symbol power modulation format that carries $M = 5$ bits/symbol and consequently has a symbol set of size $2^M = 32$. By defining the dimensions as \hat{X}_{REAL} , \hat{X}_{IMAG} , \hat{Y}_{REAL} and \hat{Y}_{IMAG} , the constraint is written as $\hat{X}_{\text{REAL}-i}^2 + \hat{X}_{\text{IMAG}-i}^2 + \hat{Y}_{\text{REAL}-i}^2 + \hat{Y}_{\text{IMAG}-i}^2 = 1$, for all symbol i in the set. Increasing the symbol

set size is equivalent of packing more states in a 4D space, which naturally reduces the minimum squared Euclidean distance (SED) between neighboring states, increasing the required SNR for a target BER [90]. However, by constraining the symbols to satisfy the unit power criterion we lose the independence of all dimensions in the 4D space; all 2^M symbols now have to be sitting on the surface of a 4 dimensional sphere. We end-up with a 4D space of only 3 degrees of freedom. Packing states on a 4D surface normally brings the states closer to each other than when packing them on independent 4D's. Consequently, the drawback of constant total power formats modulating solely polarization and phase is that the SED is reduced, consequently requiring a larger SNR compared to the unconstrained 4D modulation format of equal symbol set size. However, for the specific cases of 64-states DP-star-8QAM and its proposed constrained 32-states 8PolSK-QPSK, the two formats do share the same smallest SED between closest states: $a^2 + a^2 = 0.4226$.

As the Stokes space represents well the power and state of polarization of symbols, we depict in Fig. 3–12 the constellations in the Stokes space of (a) the 8PolSK-QPSK format and (b) the DP-8QAM format, both with added white Gaussian noise giving 16 dB of signal to noise ratio. In Stokes space, the radius of a point represents its power. Figure 3–12(b) clearly depicts the increased power variation of the DP-8QAM format. For DP-8QAM, one fourth of all possible symbols have radii ($r = 0.4226$) below the grey sphere of radius 1, half the symbols lie exactly on the sphere and one fourth reside outside ($r = 1.5774$). For 8PolSK-QPSK format, however, all generated states reside on the sphere. One can observe that the 8 SOP states of DP-8QAM lying on the sphere are exactly the same 8 states for 8PolSK-QPSK. If we distinguish SOPs having the same orientation but of different magnitude, we realize that DP-8QAM has 16 SOPs. For both formats, each SOP can carry 4 different absolute phases. Hence, we can compute that 8PolSK-QPSK supports $8 \times 4 = 32$ symbols and DP-8QAM, $16 \times 4 = 64$. When considered as a pure 3 bit/symbol Stokes space constellation, 8PolSK is referred to as “cubic”-PolSK [91,92], clearly pictured in Fig. 3–12(a).

3.2.2.1 Minimum Distance Detection for 8PolSK-QPSK

The 8PolSK-QPSK format shows a total of $2^5 = 32$ possible symbols. Each symbol can be represented by a Jones vector of the form.

$$|S_i\rangle = \begin{bmatrix} S_{i-x} \\ S_{i-y} \end{bmatrix} \quad (3.12)$$

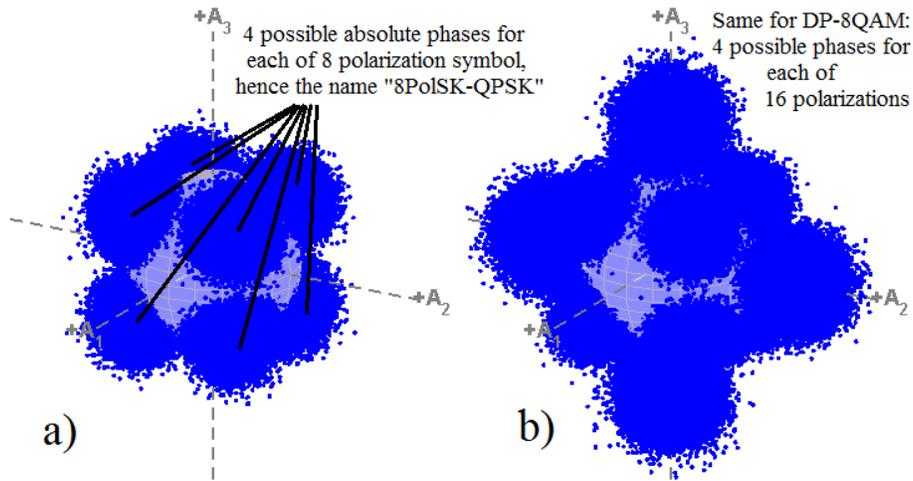


FIGURE 3–12: Stokes space representation of the a) 8PolSK-QPSK and b) DP-8QAM format, both with added noise giving 16 dB of SNR. The axis of both figures are equal. The higher power variations from symbol to symbol is clearly observed for DP-8QAM.

where S_{i-x} and S_{i-y} are the information on \hat{x} and \hat{y} polarization of the i^{th} Jones symbol. As the information on each polarization is coupled, we cannot apply an independent minimum distance detection on each polarization over the 8 constellation points of Fig. 3–11, as is done for DP-8QAM. With coupled polarization information, the minimum distance detector needs to use information on both polarizations simultaneously for symbol detection. If $|R\rangle$ is the received symbol, the minimum distance detector is [90]

$$\min_i \left(\sum_{Pol} \left| |R\rangle - |S_i\rangle \right|^2 \right) \quad (3.13)$$

where we sum the component wise subtraction of each polarization. There is a simpler, less computationally extensive way to derive the minimum distance detection. This technique uses

$$\min_i (\text{Re} \{ \langle R | S_i \rangle \}) \quad (3.14)$$

To show this, we expand the minimization argument of Eq. (3.13) to obtain

$$\min_i (\langle R | R \rangle - 2\text{Re} \{ \langle R | S_i \rangle \} + \langle S_i | S_i \rangle) \quad (3.15)$$

The proposed format exhibits constant total power, so $\langle S_i | S_i \rangle = 1$ for any symbol, and the first term $\langle R | R \rangle$ is independent of different symbols $|S_i\rangle$. The only relevant term becomes the central one. Consequently, instead of computing three terms to obtain Eq. (3.15), the minimum distance detector's complexity is alleviated by computing only one term, as in Eq. (3.14), for the

8PolSK-QPSK format. It is noteworthy to mention that the computation of the minimum distance using Eq. (3.14) is only valid for formats having a symbol set exhibiting constant power.

A different symbol-to-bits mapping is indeed required for the two formats. For DP-8QAM, we use the optimal bit-to-symbol mapping for 8QAM [93] applied independently to both orthogonal polarizations. Optimal mapping means a bit mapping that yields the smallest average number of bits in error per symbol in error. This mapping uses identical Gray-coded QPSK mapping on each ring of Fig. 3–11, where one ring is rotated by 45° , with an extra amplitude bit, for a total of 6 bits. For the 8PolSK-QPSK format, the first bit of each symbol is assigned to power, defining if \hat{x} or \hat{y} was decided to have more power, equivalent of slicing Fig. 3–12(a) on the plane $A_1 = 0$. Then, 2 phase bits on \hat{x} and on \hat{y} are Gray coded independently, complying with the optimal mapping scheme found in [93], for a total of 5 bits.

3.2.3 Nonlinearity Mitigation by Modulating with Constant Power

The idea behind the 8PolSK-QPSK is to reduce the total power variation at the transmitter while keeping a high bit per symbol efficiency. From the wave equation, we know that the nonlinear phase depends on the total optical power in the fiber. The vectorized form of the NLS equation for the modulated optical field, referred to as the Manakov-PMD equation, is described as [69,72,94,95]

$$\frac{\partial |u\rangle}{\partial z} = \underbrace{-\frac{\alpha}{2} |u\rangle}_{\text{Att.}} - \underbrace{i\frac{\Delta\beta_o}{2} (\hat{b} \cdot \vec{\sigma}) |u\rangle}_{\text{Biref.}} - \underbrace{\Delta\beta_1 (\hat{b} \cdot \vec{\sigma}) \frac{\partial |u\rangle}{\partial t}}_{\text{PMD}} + \underbrace{i\frac{\beta_2}{2} \frac{\partial^2 |u\rangle}{\partial t^2}}_{\text{CD}} + \underbrace{i\frac{8}{9}\gamma \langle u|u\rangle |u\rangle}_{\text{NL}} \quad (3.16)$$

where \hat{b} is the unitary birefringence vector, $\vec{\sigma}$ is the vector of the three Pauli spin matrices [85] $|u\rangle = |u(z, t)\rangle = [u_x, u_y]^T$ is the field envelope in the Jones vector notation. The second and third terms on the right-hand side of Eq. (3.16) are the phase-velocity and group-velocity birefringence, the latter being commonly called PMD, with a respective strength of $\Delta\beta_o$ and $\Delta\beta_1$. All \hat{b} , $\Delta\beta_o$ and $\Delta\beta_1$ are z -dependent. All the nonlinear components are lumped in the last term and include SPM, XPM, and XPolM. There is one nonlinear term that is intentionally omitted in Eq. (3.16) and it is the “nonlinear PMD”. This term, thoroughly discussed in [72], is proven to be vanishingly small and to have negligible effects. From Eq. (3.16) it is clear that the nonlinear processes impact the phase of the waveform and that this impact is itself dependent on the power $\langle u|u\rangle$ of the propagating waveform. For single carrier systems, $\langle u|u\rangle = u_x^2(t) + u_y^2(t)$ and consequently the more time varying the power of the waveform, the greater the variation of the nonlinear phase.

If $|u\rangle$ is the sum of modulated envelopes at different frequencies, for instance in the case of a WDM transmission, $|u\rangle = \sum_{k=1}^N |u_k(z, t)\rangle e^{i\omega_k t}$ and one can show that the nonlinear phase term $\langle u|u\rangle |u\rangle$ of Eq. (3.16) becomes

$$\begin{aligned} \langle u|u\rangle |u\rangle &= \langle u_n|u_n\rangle |u_n\rangle + \sum_{m \neq n} \left(\langle u_m|u_m\rangle |u_n\rangle + \langle u_m|u_n\rangle |u_m\rangle \right) \\ &= \underbrace{\langle u_n|u_n\rangle |u_n\rangle}_{\text{SPM}} + \sum_{m \neq n} \left(\underbrace{\frac{3}{2} \langle u_m|u_m\rangle |u_n\rangle}_{\text{XPM}} + \underbrace{\frac{1}{2} (\vec{u}_m \cdot \vec{\sigma}) |u_n\rangle}_{\text{XPoLM}} \right) \end{aligned} \quad (3.17)$$

for the n^{th} carrier, after mathematically retaining only terms oscillating at the n^{th} carrier frequency ω_n . We used the identity $|u_m\rangle \langle u_m| = 1/2 (\langle u_m|u_m\rangle \mathbf{I} + (\vec{u}_m \cdot \vec{\sigma}))$ [70] where \vec{u}_m is the Stokes representation of $|u_m\rangle$ and \mathbf{I} is the 2-by-2 identity matrix. Here, $\omega_k = k\omega_{\text{GRID}}$, where ω_{GRID} is the WDM grid, e.g. 50 GHz.

The third term in Eq. (3.17), representing XPoLM, can be cast as $(\sum_{m \neq n} \vec{u}_m) \cdot \vec{\sigma} = |\sum_{m \neq n} \vec{u}_m| (\hat{v} \cdot \vec{\sigma})$ where $\hat{v} = (\sum_{m \neq n} \vec{u}_m) / |\sum_{m \neq n} \vec{u}_m|$ is a unitary Stokes vector. Put in this form, one can realize that the XPoLM term is expressed similarly as the birefringence term of Eq. (3.16), with a birefringence vector \hat{v} and amplitude $-(8/9)\gamma |\sum_{m \neq n} \vec{u}_m|$. In some literature, XPoLM is in fact called nonlinear birefringence [85,96], as clearly depicted in Eq. (3.17). As a consequence of XPoLM, the state of polarization of the channel of interest is modified by a rotation around the instantaneous sum of the Stokes vectors of the interfering channels, $\sum_{m \neq n} \vec{u}_m$, and this nonlinear birefringence term depends on the statistics of the interfering channels' cumulative power, $|\sum_{m \neq n} \vec{u}_m|$, and that of their cumulative direction [97,98]. This XPoLM term is the only nonlinear term that is polarization dependent, and XPoLM's impact depends on the SOPs of all other co-propagating channels.

3.2.3.1 Nonlinear Variance Study for 8PolSK-QPSK and DP-8QAM

In the following, we study the variance of each nonlinear term appearing in Eq. (3.17) for the two modulation formats: the proposed format and the DP-8QAM format. We study the central channel of an N -channel system using Eq. (3.16). The full split-step Fourier method is employed for propagation. We use, in the simulation, parameters that will be used in a subsequent experimental demonstration presented below, namely $N = 7$ co-propagating channels of identical modulation format and power, launched with a root-raised cosine (RRC) pulse shape of roll-off factor $\alpha = 1/8$, at a symbol rate of 25.8 Gbaud for 8PolSK-QPSK and 21.5 Gbaud for DP-8QAM. The channels sit

in a 50 GHz grid and we study the impact on the central channel. In the simulations, the SOPs of all channels at launch are random. In this simulation, we deliberately neglected amplified spontaneous emission (ASE) noise added by EDFAs, PMD, and birefringence. ASE is neglected because we don't want to interfere the statistics of the signal with external additive white Gaussian noise, as is shown in [99]. PMD is neglected because we do not want to convolve XPolM induced rotations with frequency dependent rotation from PMD, and finally birefringence is neglected as it simply acts as a global unitary rotation equally applied to the entire waveform. Amplification is lumped at every 80 km and the fiber is standard SMF-28 with $\beta_2 = -21.4 \text{ ps}^2/\text{km}$ and an attenuation coefficient α of 0.2 dB/km. The launch power per channel is set at + 1dBm to trigger significant nonlinear impairments. We also use $\gamma=1.3 \text{ W}^{-1}\text{km}^{-1}$.

We are interested in the variance of all nonlinear effects applied to the channel of interest at every distance z , with respect to the mean power of said channel at such distance. As the nonlinear noise is proportional to the signal power [100], one can understand that the NL variance will be proportional to the squared signal power. Consequently, as applied in [101], we normalize the strength of the NL impairments by the squared signal power to alleviate the NL noise from its power dependence. In both our simulations and experiments we set the power per channel to be equal for all channels, allowing normalizing the strength of different NL impairments by the squared power of a single channel. The equations allowing nonlinear monitoring are

$$\text{SPM} : \quad \sigma_{\text{SPM}}^2 = \left(\overline{\langle u_n | u_n \rangle^2} - \langle \overline{u_n | u_n} \rangle^2 \right) / \langle \overline{u_n | u_n} \rangle^2 \quad (3.18)$$

$$\text{XPM} : \quad \sigma_{\text{XPM}}^2 = \left(\overline{\left(\sum_{m \neq n} (3/2) \langle u_m | u_m \rangle \right)^2} - \sum_{m \neq n} \overline{(3/2) \langle u_m | u_m \rangle^2} \right) / \langle \overline{u_n | u_n} \rangle^2 \quad (3.19)$$

$$\text{XPolM} : \quad \sigma_{\text{XPolM}}^2 = \left(\overline{\left| \sum_{m \neq n} \vec{u}_m \right|^2} - \left| \sum_{m \neq n} \overline{\vec{u}_m} \right|^2 \right) / \langle \overline{u_n | u_n} \rangle^2 \quad (3.20)$$

As $|u\rangle = |u(z, t)\rangle$, the nonlinear term $\langle u|u\rangle$ of Eq. (3.16), expended in Eq. (3.17) as $\langle u|u\rangle_{\text{SPM}} + \langle u|u\rangle_{\text{XPM}} + \langle u|u\rangle_{\text{XPolM}}$, is both time and distance dependent. The variance of those three nonlinear terms is applied over the time variable, yielding the z -dependent variances σ_{SPM}^2 , σ_{XPM}^2 and σ_{XPolM}^2 . As the solution of the propagation equation of Eq. (3.16) evolves for every dz , the local temporal variance of the NL processes also evolves with distance. The variance of the three NL processes of Eq. (3.16), namely $x_1 = \langle u_n | u_n \rangle$, $x_2 = \sum_{m \neq n} (3/2) \langle u_m | u_m \rangle$ and $x_3 = \sum_{m \neq n} 1/2 \vec{u}_m \cdot \vec{\sigma}$ are depicted in Fig. 3–13. It is important that the y-axis of Fig. 3–13 be understood as the normalized

variance of x_1 , x_2 and x_3 . By removing the normalization we obtain the observable nonlinear strength variances, $\tilde{\sigma}_{NL}^2 = \sigma_{NL}^2 \times \overline{|u_n|u_n}^2$ where the subscript NL is either SPM, XPM or XPolM. Of course, $\overline{|u_n|u_n}^2$ decays with distance as $e^{-2\alpha z}$, where $\alpha = \alpha_{dB/km}/(10\log_{10}(e))$. The averaging operator $\overline{X(t)}$ is a statistical ensemble average and time average, where the time averaging operator is $\lim_{T \rightarrow \infty} (1/T) \int_{-T/2}^{T/2} X(t) dt$ [102]. Monte Carlo simulations of 100 propagation runs were realized to determine the properties in Eqs. (3.18) to (3.20) with confidence. Each propagation consists of a PRBS sequence of 2^{12} symbols with 8 samples per symbols for each of the 7 channels. The results are presented in Fig. 3–13.

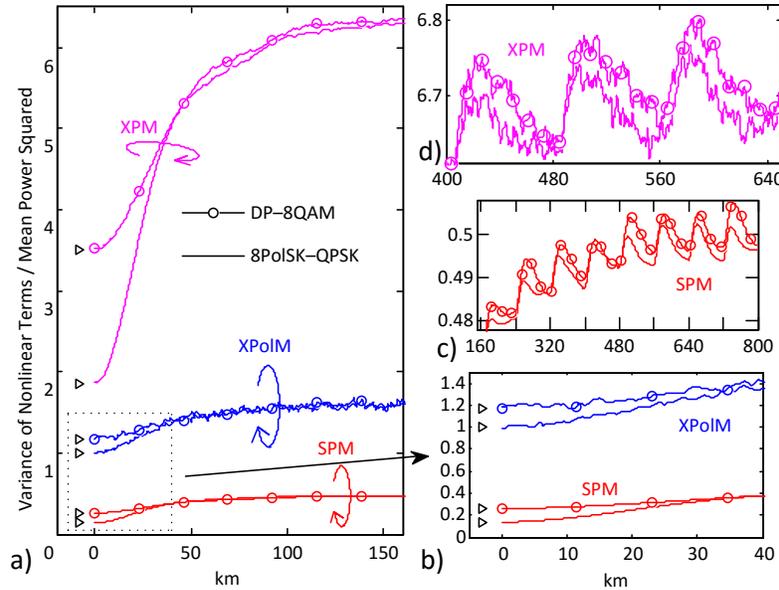


FIGURE 3–13: Temporal variance of SPM, XPM and XPolM, normalized by mean power squared, for 8PolSK-QPSK (solid lines: —) and DP-8QAM (circled solid lines: -o-) a) Main figure showing all NL strengths for first two spans b) Zoom-in from 0 to 40 km showing SPM and XPolM strengths. c) SPM only, from 3rd to 9th span. d) XPM only between 400 and 650 km

We observe that the behaviour of the variance of nonlinear processes at launch, when $z = 0$, is much different than that after 40 km and more, and that this is applicable for both modulation formats. The interplay between chromatic dispersion with channel walk-off and nonlinearity greatly modifies the initial statistics. At launch, we obtain a SPM variance σ_{SPM}^2 of 0.138 for 8PolSK-QPSK and 0.26 for DP-8QAM. Those values show that the constant power format exhibits a power variance for SPM that is roughly 1.9 times less than that of DP-8QAM. The same is applicable to XPM: at launch, XPM's variance σ_{XPM}^2 is roughly 1.86 for 8PolSK-QPSK and 3.52 for DP-8QAM. We

deliberately added markers, \triangleright , at those nonlinear strengths at $z = 0$. We will come back and relate those values to theory in Section 3.2.3.2. As the waveform propagates in the first span, the relative variances σ_{SPM}^2 , σ_{XPM}^2 and σ_{XPoIM}^2 increase very rapidly until they each reach a value roughly maintained for the rest of the fiber length. After about only 40 km, the interplay of CD with nonlinearity renders respective variances seemingly the same for both formats.

We will show in Section 3.2.5 experimental results demonstrating that the constant power 8PolSK-QPSK format performs better than DP-8QAM format. Propagation simulation results in Fig. 3–13 well explain the reasons. This performance improvement is due to the smaller SPM and XPM variances at launch.

It is well known that fiber dispersion plays two roles in the process of XPM induced intensity interference. The first is translating the nonlinear phase modulation into intensity fluctuations and the second is causing signals to propagate with different velocities [69, 103]. It was previously concluded that the nonlinear XPM phase interaction between the signals takes place only in the beginning of the fiber [69, 98, 103], occurring over a short length in which the walk-off between the signals is small. In that part of the fiber, the phase of the channel of interest gets modulated by neighboring channels, as the walk-off between the signals is minimal and pulses are still well defined and contained in their symbol slot, while the intensity modulation is small, as cumulative CD is almost null. After the first walk-off length, the signals then propagate linearly over the rest of the fiber, where the action of dispersion on the phase modulated signal are continuously translated into intensity fluctuations [103], appearing as a continuous additive Gaussian noise [104]. As the bulk of nonlinear XPM phase shift is induced over the first walk-off length only, if power fluctuations can be kept small in this region, the overall XPM penalty will be smaller [98]. Consequently, the better performance of 8PolSK-QPSK format is explicable by the fact that it exhibits a much smaller XPM strength at launch, and SPM for that matter, hence suffering less nonlinear penalty. Results in Fig. 3–13 also concur with those observations from the literature: with a dispersion parameter of $\beta_2 = -21.4 \text{ ps}^2/\text{km}$ and a grid spacing of 50 GHz, the first neighboring channels are delayed at a rate of 1.07 ps/km. With a pulse duration of 38.5 or 46.5 ps, the walk-off length is around 40 km: the required distance before nonlinear variances converge, as we can observe in Fig. 3–13.

Figure 3–13 (c) and (d) depict the behavior of normalized nonlinear strengths after longer propagation distances by providing a closer look at σ_{SPM}^2 and σ_{XPM}^2 , respectively. Although almost identical, we observe that the SPM and XPM variances for DP-8QAM are slightly higher than

that of 8PolSK-QPSK. Albeit not the primary argument justifying improved performance of the proposed format, this subtle reduction of σ_{SPM}^2 and σ_{XPM}^2 occurring at longer distances would also benefit the 8PolSK-QPSK format, as well as the principal and dominant reduction occurring in the first walk-off length. Finally, we also observe an oscillation of the normalized nonlinear strengths of period equal to the EDFA spacing of 80 km, showing a ramp up of the relative power variances for the first 20 to 25 km after each EDFAs, followed by a slow decay for the rest of the span.

3.2.3.2 Variance of SPM, XPM and XPolM at Launch

The relative power variance numbers obtained at launch ($z=0$) describing SPM and XPM can be recovered analytically. In order to analytically compute Eq. (3.18) for SPM, we need to calculate $\overline{\langle u_i | u_i \rangle^2}$. If the launched pulse shape function is $g(t)$ and the polarization multiplexed waveform $|u_i\rangle$ of this central channel i is $|u_i(t)\rangle = \sum_{p=-\infty}^{\infty} |S_p\rangle g(t-pT)$ where $|S_p\rangle$'s are the possible PDM symbol, then $\overline{\langle u_i | u_i \rangle^2}$ equals

$$\overline{\langle u_i | u_i \rangle^2} = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \langle u_m | u_n \rangle \langle u_k | u_l \rangle g^*(t-mT)g(t-nT)g^*(t-kT)g(t-lT) \quad (3.21)$$

After keeping only terms of $\langle u_m | u_n \rangle \langle u_k | u_l \rangle$ that have a non-zero statistical ensemble average, we end up with

$$\begin{aligned} \overline{\langle u_i | u_i \rangle^2} &= \left(\overline{\langle u_k | u_k \rangle \langle u_l | u_l \rangle} + \overline{\langle u_k | u_l \rangle \langle u_l | u_k \rangle} \right) \frac{1}{T} \int_{-\infty}^{\infty} \sum_{m=-\infty}^{\infty} |g(t)|^2 |g(t-mT)|^2 dt + \\ &\quad \left(\overline{\langle u_k | u_k \rangle^2} - \overline{\langle u_k | u_k \rangle \langle u_l | u_l \rangle} - \overline{\langle u_k | u_l \rangle \langle u_l | u_k \rangle} \right) \frac{1}{T} \int_{-\infty}^{\infty} |g(t-mT)|^4 dt \end{aligned} \quad (3.22)$$

This equation is valid for the symbol set $|S_p\rangle$ of any modulation format, and for any pulse shape $g(t)$, even for a pulse shape that includes CD. The respective terms $X_{kkll} = \overline{\langle u_k | u_k \rangle \langle u_l | u_l \rangle}$, $X_{kllk} = \overline{\langle u_k | u_l \rangle \langle u_l | u_k \rangle}$ and $X_{kkkk} = \overline{\langle u_k | u_k \rangle^2}$ are modulation format dependent. If the magnitude of the symbols $|S_p\rangle$ are such that the mean total power is unitary, one can show that $X_{kkll} = 1$, $X_{kllk} = 0.5$ and $X_{kkkk} = 1$ for 8PolSK-QPSK, whereas $X_{kkll} = 1$, $X_{kllk} = 0.5$ and $X_{kkkk} = 7/6$ for DP-8QAM. If the pulse shape is root-raised cosine (RRC) of any roll-off factor $0 < \alpha < 1$, the mean power $\overline{\langle u_i | u_i \rangle}$ can be relatively easily analytically calculated to be $1/T$. However, the analytical expression for $\overline{\langle u_i | u_i \rangle^2}$ is more complicated and it is not our goal to derive it. $\overline{\langle u_i | u_i \rangle^2}$ is both pulse

shape dependent and format dependent. It is easy to obtain the numerical value of Eq. (3.22) using a mathematical software

Now that we have an expression for $\overline{\langle u_i | u_i \rangle^2}$ and its value, we can obtain the relative nonlinear variance of SPM and XPM from Eqs. (3.18) and (3.19) for each format. From the result of σ_{SPM}^2 , σ_{XMP}^2 can easily be obtained: for a N -channel co-propagation, $\sigma_{\text{XMP}}^2 = (N - 1)(3/2)^2 \sigma_{\text{SPM}}^2$. The numerical evaluation of σ_{SPM}^2 for RRC with $\alpha = 1/8$ is 0.13785 and 0.26015 for 8PolSK-QPSK and DP-8QAM respectively. Consequently, with $N = 7$, σ_{XMP}^2 yields 1.861 and 3.512 for the 2 formats, respectively. Those values, obtained from Eqs. (3.18) and (3.19), map exactly to our propagation simulation results at $z = 0$ in Fig. 3–13. Small right-pointing triangle markers “▷” located at those theoretical values of σ_{SPM}^2 and σ_{XMP}^2 for both formats show the match between simulation and theory at launch for σ_{SPM}^2 and σ_{XMP}^2 , when no CD nor nonlinearity is applied on the waveform.

A further discussion on the XPM impact on a channel when an increasingly larger number of copropagating channels are added to the link is of interest. It is known that XPM occurs only if the two pulses from different channels overlap temporally [85]. It is also established that XPM-induced phase shift decreases as the relative group-velocity difference increases, where the latter is proportional to the frequency difference between channels [85]. Consequently, only close neighbours to a channel have a significant XPM impact on said channel. Channels added increasingly farther from a reference channel observe increasingly larger GVD difference, and pulses from the two different channels will traverse at an increasingly different speed, blurring the effect of XPM. As Fig. 3–13 only shows the evolution of the variance of the three nonlinear elements in Eq. (3.16) as it is numerically propagated, the reduced effect of XPM impairment from increasingly farther channels is not directly shown in this Fig. As the number of aggressor channels grows, located farther away from a reference channel, their XPM impact on said reference channel would be negligible compared to XPM’s impact from the closest neighbours.

The variance of XPolM’s rotation strength, on the other hand, is random by nature and increases to the same steady-state average, and becomes rapidly agnostic to the modulation format. The rapid settling of the 3 variances σ_{SPM}^2 , σ_{XMP}^2 and σ_{XPolM}^2 to an approximately steady-state value concurs with a recent conclusion that nonlinear interferences can be modeled as excess additive Gaussian noise [104]. We show here that the variance of that additive Gaussian noise does tend to a coarse steady state value very rapidly.

For XPolM, the statistics of the variable $X = |\sum \vec{u}_m|$ at launch, namely the mean μ_X and variance σ_X^2 , depend on the modulation format, the number of other co-propagating channels, and the pulse shape applied to all channels. The nonlinear XPolM strength X is naturally a random process due to both the instantaneous randomness of the data of every other channel affecting one channel and due to walk-off between channels [98]. As $|\sum \vec{u}_m|$ is the norm of a Stokes space vector, the variable represents time varying power. The mean and variance of $|\sum \vec{u}_m|$ over time can easily be numerically obtained. For a $N = 7$ channel WDM propagation where each channel's power is normalized to 1, with RRC pulse shaped with $\alpha = 1/8$, one can compute that $\mu_X \approx 2.412$ and $\sigma_X^2 \approx 1.01$ for 8PolSK-QPSK where $\mu_X \approx 2.523$ and $\sigma_X^2 \approx 1.19$ for DP-8QAM. Consequently, σ_{XPolM}^2 in Eq. (3.20) equals 1.01 for 8PolSK-QPSK and 1.19 for DP-8QAM. We can observe that the XPolM variance obtained in the propagation simulation in Fig. 3–13, at launch, match exactly those numbers. Again, this match is depicted in Fig. 3–13 by markers \triangleright added at those theoretical values for XpolM. The variance difference of $|\sum \vec{u}_m|$ for both modulation formats can be explained by the fact that the mean squared power $\overline{\langle u|u \rangle^2}$ is larger for DP-8QAM compared to that of 8PolSK-QPSK; $\overline{\langle u|u \rangle^2} = 1$ for 8PolSK-QPSK and $7/6$ for DP-8QAM.

When a large number of channels are launched, all randomly polarized, we can consider that every $u_{i\text{-TOT}}$ in $\sum \vec{u}_m = [u_{1\text{-TOT}}, u_{2\text{-TOT}}, u_{3\text{-TOT}}]^T$ is a Gaussian variable of zero-mean and variance σ^2 . Consequently, the variable $|\sum \vec{u}_m| = X = \sqrt{(u_{1\text{-TOT}}^2 + u_{2\text{-TOT}}^2 + u_{3\text{-TOT}}^2)}$ is a random variable of Chi-distribution with 3 degrees of freedom. It is known that for a Chi-distribution where the generating Gaussian variables all have zero-mean and a variance of σ^2 , the Chi variable X has a mean of $\mu_X = \sigma\sqrt{2} \times \Gamma(2)/\Gamma(3/2)$ where Γ is the Gamma function, and a variance of $\sigma_X^2 = 3\sigma^2 - \mu_X^2 = \sigma^2(3 - 4\Gamma(2)^2/\Gamma(3/2)^2) \approx 0.4535\sigma^2$. One can show numerically that the variance σ^2 of every $u_{i\text{-TOT}}$ is larger by $\approx 10\%$ for the DP-8QAM format compared to that of 8PolSK-QPSK. This explains why both the mean and variance of the strength of XPolM $|\sum \vec{u}_m|$ is smaller for 8PolSK-QPSK at launch. Finally, we can conclude that even the impact of XPolM is lessened for 8PolSK-QPSK format during the early stage of propagation.

It is noteworthy to observe in Eq. (3.17) that the impact of XPolM is 3 times less efficient than that of XPM. Hence the benefit of focusing on reducing the XPM effect over XPolM. The above analysis validates our assumption that having constant power modulation on all wavelengths would reduce nonlinear effects. The experimental test and results presented below confirm our analysis.

3.2.4 Experimental Test Bed

To assess the performance of the two modulation formats, we use the test bed shown in Fig. 3–14. In order to use the same transmitter configuration, and consequently the same transmitter implementation penalty, the only parameter modified when switching from one format to the other is the symbol rate out of the Digital to Analog Converters using a fixed clock rate. The test bed begins with 7 lasers combined to generate the WDM waveform. Six of which are distributed feedback lasers and the central channel of interest is an external cavity laser with laser linewidth <100 khz. Lasers are placed on a 50 GHz grid. The 6 DFBs, of linewidths smaller than 1 MHz, are used as aggressors acting on the central channel. All channels have equal power. This WDM configuration allows assessing the full nonlinear degradation, where all nonlinear impairments are included. A SP-IQ modulator bulk modulates the laser tones using two DACs clocked at a rate of 32 GSa/s for both formats. To modulate the different formats at a their desired symbol rates, a proper root raised cosine pulse shaping filter with a roll-off factor of 1/8 is applied. In order to yield the same bit rate of 129 Gbit/s, the 6 bits/symbol DP-8QAM format runs at a symbol rate of 21.5 GBaud, whereas the 5 bits/symbol 8PolSK-QPSK run at 25.75 GBaud. The reason why this latter symbol rate is not $129/5 = 25.8$ GBaud is twofold. First, the DACs' memory has to be a power of 2 and second, the DAC rate is deliberately fixed for both formats. The difference between a symbol rate of 25.8 and 25.75 is assumed negligible. The DAC's pattern length is set to 2^{19} . The SP-IQ modulator is driven in the linear regime. Following the IQ modulator is the dual polarization emulator with a delay set at 108 ns. This delay translates to an integer number of symbols at both 21.5 GBaud (2322 Symbols) and 25.75 GBaud (2781 Symbols).

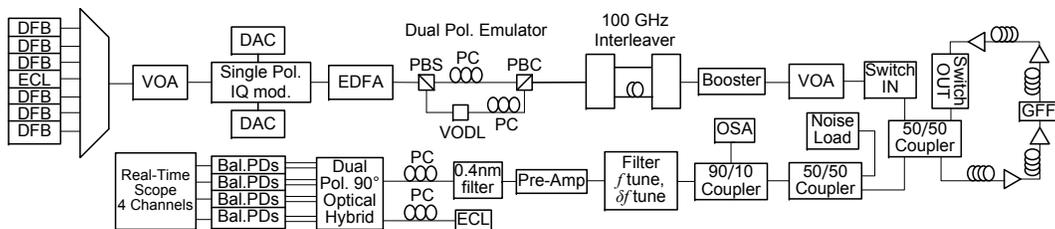


FIGURE 3–14: Experimental Test Bed to compare 8PolSK-QPSK vs DP-8QAM

Then, two 50/100 GHz passive interleavers are used to temporally decorrelate even and odd channels. The WDM waveform is then boosted and attenuated to sweep the launch power before

launching in an optical recirculating loop. The latter is composed of 4 spans of 80 km of Corning[®] SMF-28e+ LL fiber with dual-stage inline EDFAs of noise figure ≈ 5.5 dB. A gain flattening filter is inserted after the second span inside the loop. Chromatic dispersion is uncompensated. A noise loading stage follows the recirculating loop for OSNR sweeping followed by a channel selecting filter, isolating the central channel of interest. Finally, a pre-amp and a 0.4 nm filter are preceding the optical coherent receiver. The receiver consists of a dual polarization 90° optical hybrid, 4 balanced photodiodes (U²T[®] BPDV2020R) and a 4 channel real-time scope sampling at 80 GSa/s.

3.2.4.1 8PolSK-QPSK Signal Generation with a Dual Polarization Emulator

Generating a 8PolSK-QPSK format with a SP-IQ modulator and a dual polarization delay-and-add emulator requires some handling of the signal provided to the SP-MZM. The dual polarization emulator consists of splitting the incoming SP waveform on 2 orthogonal branches, delaying one, and recombining on two orthogonal polarizations, as pictured in Fig. 3–14. For regular Dual Polarization multiplexing where the same format is employed independently on both polarizations, the delay of the DP emulator has no impact on the output format, as long as the delay is an integer number of symbols, is not zero, and is sufficiently large. Generating a 8PolSK-QPSK format out of a SP-8QAM format, however, requires extra manipulation on the SP waveform out of the Mach-Zehnder modulator (MZM). The SP waveform generated has to be related to the DP emulator's delay. As explained in Section 3.2.2, the symbol constellation of a single polarization of the 8PolSK-QPSK format is exactly the same as that of the DP-8QAM format, i.e. the 8QAM constellation of Fig. 3–11. However, the 8PolSK-QPSK signal out of the DP emulator has to exhibit the constant power constraint, not observable before the emulator.

A specific repetitive pattern has to be imprinted on the DACs for the SP-8QAM waveform to get converted to 8PolSK-QPSK after the DP emulator. With a DP emulator delay of 108 ns, the delay is of $N = 2781$ symbols at 25.75 Gsymbols/s. For a DAC sampling rate of 32 GSa/s, a DAC memory of 2^{19} samples and a symbol rate of 25.75 GBaud, a total of $M = 421888$ symbols are repeatedly generated by the DACs. Two different symbols streams are needed to properly mimic both 8PolSK-QPSK and DP-8QAM formats. The first and most simple symbol stream is a random QPSK sequence of length M . The second symbols stream is what will determine if the waveform after the DP emulator is a 8PolSK-QPSK or DP-8QAM format. If the second symbol stream is a random sequence also of length M of symbols $\{S_0 = a, S_1 = be^{i\pi/4}\}$ and both streams are

multiplied symbol by symbol, then the waveform generated is DP-8QAM. To generate an 8PolSK-QPSK format, the second symbol stream is constructed as follows. First, we generate a random sequence of length N from symbol family $\{S_0 = a, S_1 = be^{i\pi/4}\}$. We call this sequence S . Then, a sequence of length $2N$ is generated by concatenating the inverse of S to itself, generating the sequence $S' = [S \bar{S}]$. If $S'(k) = S_0$, then $S'(k+N) = S_1$ and vice versa. This new sequence S' is then replicated and truncated after M symbols are generated. By multiplying this symbol stream with the QPSK stream, we will generate the desired 8PolSK-QPSK format after the DP emulator having a delay of N symbols. One can prove that the generated dual polarization signal $|u(n)\rangle = [u_x(n), u_y(n)]^T = [S'(n), S'(n+N)]^T$ will have the desired constant power $\langle u(n)|u(n)\rangle = 1$ for every symbol n . Of course, this tedious way of generating this constant power signal is avoided with a full Dual Polarization-Dual Parallel Mach-Zehnder Modulator.

3.2.5 Results

In this section we present the performance of both 8PolSK-QPSK and DP-8QAM modulation formats. We assume BER free operation with 29% coding overhead (OH) for hard decision forward error correction (HD-FEC) with a BER threshold of 1.4×10^{-2} , providing slightly more coding overhead than the 25% used in references [105, 106]. First, we present in Fig. 3–15 the maximum reach achievable for a BER of 1.4×10^{-2} as a function of the launch power per channel. As the launch power jumps were relatively large at 2dB/jump, the resolution and hence the optimum powers of Fig. 3–15 are coarse. At such granularity the optimum launch power appears to be the same for both formats at -1 dBm. However, by looking carefully at the trend of the curves, we realize that the optimum launch power for DP-8QAM would sit slightly below -1 dBm had we used a smaller launch power step size. On the other hand, the 8PolSK-QPSK format shows an apparent true optimum launch power at -1 dBm. We can observe, after interpolation of the DP-8QAM curve, that 8PolSK-QPSK allows a slightly greater optimum launch power, by around 0.4 dB. Secondly and most importantly, Fig. 3–15 shows that the 8PolSK-QPSK format allows a much greater propagation distance compared to DP-8QAM: 3800 km compared to 2825 km, a 34% increase, or additional 975 km. In the following Fig. 3–16, we show the experimental BER as a function of the OSNR in 0.1 nm for both formats in back-to-back, along with their theoretical values. The theoretical curves in Fig. 3–16 show a 0.6 dB difference between 8PolSK-QPSK and DP-8QAM at a BER of 1.4×10^{-2} , benefitting 8PolSK-QPSK. Experimentally, we obtain a difference of 0.5 dB.

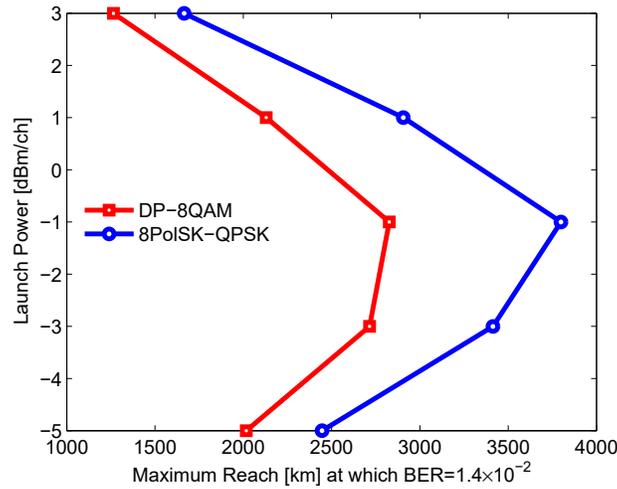


FIGURE 3–15: Launch Power per channel against maximum reach [km] for a BER = 1.4×10^{-2} . The optimum launch power is -1 dBm for 8PolSK-QPSK and around -1.4 dBm for DP-8QAM. 8PolSK-QPSK format propagates 34% more than DP-8QAM, or 975 km more.

The offset from theory for both formats, mainly due to the transmitter implementation penalty, is rather large at 4.37 dB. Our claim that both formats have the same implementation penalty is well proven here. The inset a) in Fig. 3–16 shows the theoretical OSNR (in 0.1 nm) difference between DP-8QAM and 8PolSK-QPSK as a function of the BER, for BERs varying between 2.5×10^{-3} and 2.5×10^{-2} . The independent variable, BER, is put in the y-axis to be consistent with the main outermost axes. This inset demonstrates that as the BER increases, the 8PolSK-QPSK format requires an increasingly smaller OSNR with respect to that for DP-8QAM. This property is beneficial for the proposed format as improving FEC schemes increase BER thresholds; a recent FEC scheme can bring a BER of 2.4×10^{-2} down to 10^{-15} [107]. The theoretical BER against SNR function for DP-8QAM format can be approximated by $\text{BER}(\text{SNR}) = 2/3 \operatorname{erfc}(\sqrt{3 \cdot \text{SNR}/14})$, in the BER range of 10^{-2} to 10^{-4} , with a maximal deviation of ± 0.05 dB [104]. Unfortunately, to the author's knowledge, there is no such theoretical curve in the literature for 8PolSK-QPSK. By using a model

$$\text{BER}(\alpha, \beta, \text{SNR}) = \alpha \cdot \operatorname{erfc}\left(\sqrt{\beta \cdot \text{SNR}}\right) \quad (3.23)$$

as was used for the 4 modulation formats of PDM-BPSK, -QPSK, -8QAM and -16QAM in [104], we found α and β to match the theoretical curve of Fig. 3–16 in a least mean square fashion for the 8PolSK-QPSK format. Within a maximum deviation of ± 0.07 dB, we obtain $\alpha = 0.2817$ and

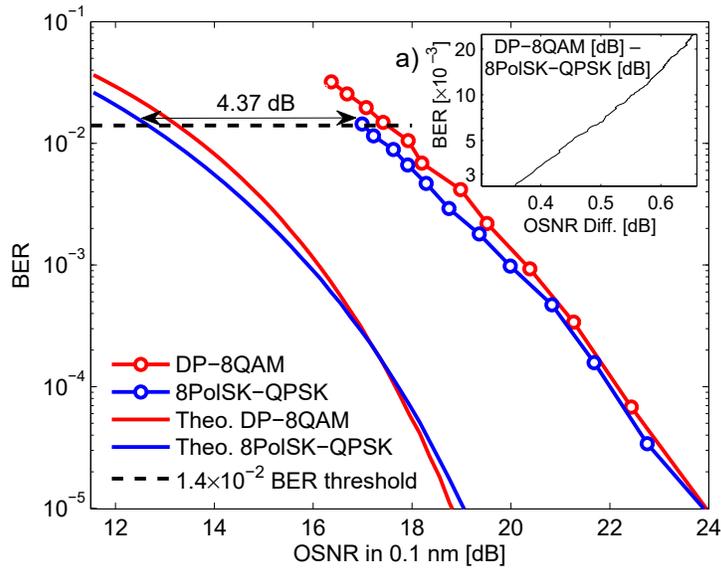


FIGURE 3–16: BER against OSNR in 0.1 nm for both 8PolSK-QPSK and DP-8QAM formats. Inset a): Theoretical OSNR (in 0.1 nm) difference between formats as a function of BER threshold.

$\beta = 0.2225$ over the larger BER range of 10^{-2} to 10^{-5} . We can observe in Fig. 3–16 that the theoretical curves cross around an OSNR in 0.1 nm of 17.2 dB, where 8PolSK-QPSK performs better for worst OSNRs and DP-8QAM starts outperforming for larger OSNRs. Identical formats compared at different symbol rates would simply yield an horizontal shift on the theoretical BER vs OSNR curve. In Fig. 3–16, the two formats exhibit not only a different symbol rate, but also a different number of recovered bits per symbol, and a different symbol-to-bits mapping, where all closest symbols do not always differ by a single bit. Hence, the two formats exhibit a different behavior in their BER against OSNR curve.

Figure 3–17 depicts the required OSNR (ROSNR) in 0.1 nm to obtain a BER of 1.4×10^{-2} after propagation of 1920 km as a function of the launch power, for both formats. We can see that in the fully linear propagation regime, at -5 dBm, the 8PolSK-QPSK format requires 0.5 dB less OSNR compared to DP-8QAM format, as predicted by results in Fig. 3–16. As the launch power increases, the strength of nonlinear impairments increase accordingly and Fig. 3–17 shows the better nonlinear tolerance of the 8PolSK-QPSK format. The smaller nonlinear impairments received by using the proposed format is expressed by the clear decoupling of the 2 curves with growing launch power. At the optimal launch power of -1 dBm where the simultaneous impact of linear and nonlinear effects

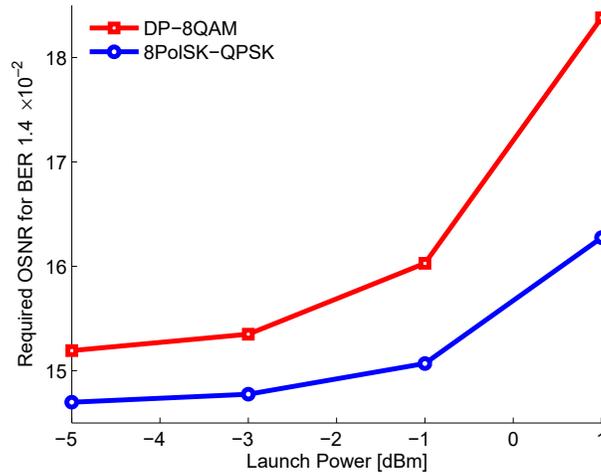


FIGURE 3–17: Required OSNR in 0.1 nm [dB] for $\text{BER}=1.4 \times 10^{-2}$ after 1920 km: 8PolSK-QPSK versus DP-8QAM.

are at minimum, the OSNR requirement is 0.95 dB less for the proposed format. At +1 dBm of launch power, where nonlinearity clearly supersedes linear impairments, 8PolSK-QPSK outperform DP-QAM by 2.1 dB of ROSNR.

As shown in Fig. 3–13 using simulations, the better performance of 8PolSK-QPSK after 1920 km for increasing launch powers comes almost solely from the first 40 km or so, where the proposed format receives significantly less NL phase noise. As the launch power increases, the NL noise applied to both formats also naturally increases, but even more for the DP-8QAM format within those first 40 km. The excess NL phase noise for DP-8QAM that happened in the first 40 km is still observable after longer distances like 1920 km. Figure 3–17 demonstrate very well both the linear and nonlinear advantages of the proposed format.

Finally, as a last experimental result, we present in Fig. 3–18 the BER against Distance curves for the two formats at their optimum launch power of -1 dBm. We can see that at any fixed distance, the BER is always lower when the 8PolS-QPSK format is employed compared to DP-8QAM. Another curve was added to the 2 BER vs Distance curves, showing the ratio of the BERs of the 2 formats as a function of distance, as $\text{BER}_{\text{DP-8QAM}}/\text{BER}_{\text{8PolSK-QPSK}} \times 10^{-6}$ and expressed in log-scale. An offset of 10^{-6} was deliberately applied to the ratio to have it fit at the bottom of the y -axis while using the same y -scale. This curve shows that the ratio of the BERs of the 2 formats is essentially constant with distance. The fact that the BER equally increases with distance for the

2 formats concurs with the simulations observations of Fig. 3–13. It was shown in Fig. 3–13 that the variance of the nonlinear strength quickly saturates to the same value after the second span for both formats. Consequently the nonlinear noise applied to both formats can be assumed equal after a few hundred kilometers. Both NL and linear noise equally increase with distance, explaining the experimental even growth of the BER for both formats.

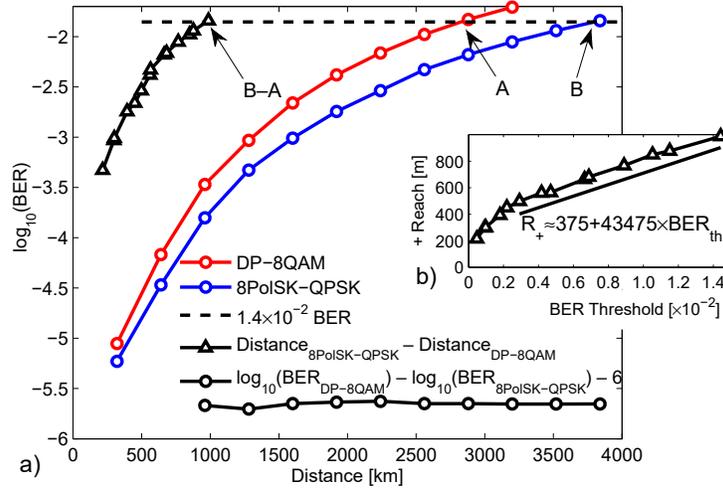


FIGURE 3–18: BER against Distance [km] for both 8PolSK-QPSK and DP-8QAM. Also shown in inset (b) is the reach difference of the two formats as a function of BER in linear scale.

Another informative curve was added to Fig. 3–18 in inset b), showing the excessive reach (R_+) of 8PolSK-QPSK over DP-8QAM at a specific BER, calculated for different BERs. As an example, at a BER of 1.4×10^{-2} , DP-8QAM’s maximum reach is marked by an “A” and 8PolSK-QPSK’s maximum reach is marked by a “B”. B–A is the excessive distance when 8PolSK-QPSK is employed compared to when DP-8QAM is used for a target BER threshold of 1.4×10^{-2} . This reach difference is shown on the left-hand side of Fig. 3–18, with the same x -axis but now representing excessive distance instead of absolute distance. Starting at a BER threshold BER_{th} of 3×10^{-3} , the added reach R_+ when using 8PolSK-QPSK over DP-8QAM can be well approximated by a linear relation with BER_{th} : $R_+ \approx 375 + 43475 \times \text{BER}_{\text{th}}$. The inset b) replicates the reach difference curve with BER in linear scale, put on the x -axis. The trend can be explained from Fig. 3–16, from the difference in the required OSNR between the theoretical curves of the two formats that increases with increasing target BER. The increasingly smaller ROSNR for the proposed 8PolSK-QPSK format compared to

DP-8QAM translates into excessive reach with increasing BER targets. With such supplemental reach monotonically increasing with BER threshold, we can predict a excessive reach by more than 1400 km at the BER threshold of 2.4×10^{-2} ; a threshold currently applicable for soft-decision (SD) FEC [82, 107]. The excessive FEC overhead of 29% used in our experiments would be sufficiently large for such SD-FEC with extra space for protocol overhead.

3.2.6 Conclusion

We have presented a novel modulation format obtained by constraining the well-known DP-8QAM format to exhibit constant symbol power. The proposed format, providing 5 bits of information per symbol and entitled 8PolSK-QPSK, presents a passive way to mitigate nonlinear impairments by significantly reducing the optical power variation from symbol to symbol in the first 40 km of propagation. Using both theoretical analysis, simulations and experimental validations, we confirmed the better nonlinear tolerance of the proposed format. We showed a back-to-back decrease of the required OSNR for a BER of 1.4×10^{-2} by 0.5 dB for 8PolSK-QPSK, showing an improved ASE noise tolerance for 8PolSK-QPSK at such BER. After a distance of 1920 km with a launch power of -1 dBm/channel with 6 other co-propagating channels, the proposed format exhibits a ROSNR reduction of 0.95 dB. We also demonstrated that at the optimum launch power of -1 dBm/channel, the proposed format extends the propagation reach by 34%, or 975 km, for a BER threshold of 1.4×10^{-2} , going from 2825 to 3800 km. We also presented a relation between the supplemental reach provided by the format against increasing BER thresholds. Finally, to bolster the analysis of the format, we presented a detailed study, both analytically and via simulations, of a comparison of the strengths of the three main nonlinear impairments occurring during propagation, namely SPM, XPM and XPolM, where we showed significantly reduced SPM and XPM strengths by using the 8PolSK-QPSK format.

CHAPTER 4

Passive and Active Silicon Photonics: Parametric Wave Mixing and Intensity Modulation

In this Chapter, we address chip-scale optical communications and present experimental studies of passive and active silicon photonic devices. First, we demonstrate the first 1-to-6 wavelength multicasting of a 22-GBaud 16-QAM single polarization signal based on four-wave mixing using a passive silicon strip waveguide. Secondly, we present the analysis and the performance of an active Silicon Photonic intensity modulator operating at 1.3 μm . We demonstrate for the first time a 112 Gb/s transmission over 10 km of single mode fiber employing PAM-8 at 37.4 Gbaud using a silicon-based modulator operating on the cost-efficient principle of intensity modulation/direct detection.

4.1 Four-Wave Mixing in a Silicon Nanowire

We first begin by presenting a passive silicon photonic device. The rectangular silicon waveguide, also called a ‘silicon nanowire’, has a height of 500 nm, a width of 220 nm and a length of 20 mm and serve as a highly nonlinear medium to perform on-chip wavelength multicasting through the parametric process of four-wave-mixing.

4.1.1 Introduction

Wavelength multicasting (WM) differs from wavelength conversion (WC) in that it employs different nonlinear processes. Although based on the same nonlinear effect of four-wave mixing, WM employs multiple degenerate and non-degenerate nonlinear mixing while WC results from a unique degenerate FWM process. The apparent simple addition of a second pump laser engenders a whole new set of nonlinear effects that are not triggered during WC, where a single pump co-propagates with the signal. Considering only first order FWM terms [108], WC generates two new idlers, out of which only one preserves the signal information, while WM engenders nine new idlers,

where five maintain the signal data. With proper pump frequencies, all 5 idlers are clearly separated in the optical spectrum, around the input signal frequency, and allow one-to-six optical spectral broadcasting of the signal.

4.1.2 State of the Art

Multiple approaches for wavelength conversion and multicasting have been demonstrated, including those based on highly nonlinear fibers (HNLF), semiconductor optical amplifiers (SOA), periodically-poled Lithium Niobate (PPLN) and silicon photonics. For instance, FWM-based 1-to-6 WM of a 320 Gb/s Return-to-Zero (RZ)-DPSK signal in HNLF with longitudinal tensioning was demonstrated in [109]. In [110], a two-stage multicasting setup with two separate HNLF sections of precisely engineered dispersion is used to demonstrate the generation of a pump frequency comb using only two input continuous wave light sources, where each generated pump is accompanied by the creation of a phase conjugate and a phase preserving idler on its sides, resulting in more than 60 copies of the input signal. WM of a 40 and 80 Gb/s RZ signal is reported in [111] by exploiting cross-gain modulation and cross-phase modulation in a quantum-dot SOA, where N CW lasers are used to replicate the signal N times. In [112], multicasting of a 20 Gb/s ASK-DPSK signal to 6 equally spaced idlers is realized using FWM in a 32-cm highly nonlinear bismuth oxide fiber employing N pumps for a $2N$ -fold multicast. Finally, high order QAM WM was obtained in [113] at 28 Gbaud with a 16-QAM emulator using a $\chi^{(2)}$ -dominant PPLN waveguide, where $N+1$ pumps are required to wavelength convert a signal N times after employing twice the second order nonlinear effect to realize WM. PPLN waveguides are expensive and not CMOS compatible. In addition to large effective nonlinearity, the interest in exploring SiP approaches is due to its compatibility with CMOS processing. To date however there have been very few demonstrations of WC at high symbol rates and high order modulation formats using passive, pure silicon waveguides. In [114], WC of 40 Gbaud BPSK and QPSK signals was reported using a two compound, 25 mm long waveguide of Si-Ge. In [115], the authors demonstrated FWM-based WC of a 640-Gb/s RZ-DPSK signal in a SiP chip. WM using one signal and two pumps was presented in [116] in a silicon nanowire (SNW), albeit at the low symbol rate and modulation format order of 10 Gb/s DPSK. 16-way WM of a 40 Gb/s OOK signal was shown in [32] using 16 times one-to-one wavelength conversion employing 16 separate CW laser diodes and an 11 mm-long SNW.

In [117] we demonstrated FWM based one-to-one WC of a 28 Gbaud 16-QAM signal in a dispersion engineered SNW. In this section, we present the first high symbol rate, high order QAM 1-to-6 WM in a 20 mm long passive SNW. WM of a 16-QAM, 22 Gbaud signal is experimentally demonstrated where all 6 multicast signals are successfully converted below the bit error rate forward error correction threshold of 3.8×10^{-3} with a worst case power penalty of 8 dB. We also experimentally demonstrate how both coupling losses of the input and output grating couplers of silicon based WM impact respectively the nonlinear conversion efficiency and the final OSNR of all multicast idlers, dictating the BER performance.

4.1.3 Principle of Operation of FWM-based WC and WM

Through a FWM process, three waves of frequencies f_i , f_j , and f_k generate a wave of frequency $f_{ijk} = f_i + f_j - f_k$ ($f_j \neq f_k$). When only two input waves are present, the process of ‘Four’ Wave Mixing can still take place, in which case one of the two input waves is accounted twice in generating the fourth ‘idler’ frequency f_{ijk} . Such process where $f_{iik} = 2f_i - f_k$ is more specifically called a ‘Degenerate’-FWM process. During wavelength conversion, two input waves are interplaying: a pump and the signal to convert. Both f_i and f_k can either be the pump f_1 or the signal f_s , creating in total two idlers at f_{ss1} and f_{11s} . Out of these two, only one maintains the signal information of the source signal: $f_{11s} = 2f_1 - f_s$. The following Fig. 4–1(a) is a schematic illustration of the degenerate-FWM process occurring during WC.

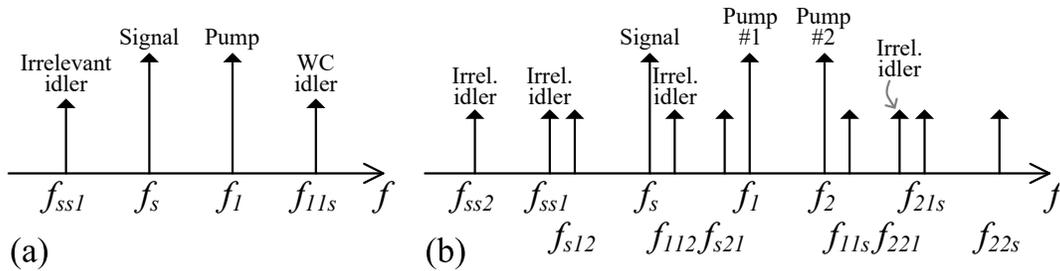


FIGURE 4–1: All idlers created via FWM in (a) a one-pump/one-signal scenario and in (b) a two-pumps/one-signal scenario.

On the other hand, the process of wavelength multicasting via FWM relies on the interaction of three waves, two of which are pumps and one is the signal to multicast. The addition of a second pump of frequency f_2 engenders another pair of degenerate idlers from the interaction of the latter with f_s , where the two new degenerate idlers are at f_{ss2} and f_{22s} . Similarly, only one of these two

new idlers maintains the signal information of the source signal: $f_{22s} = 2f_2 - f_s$. However, the effect of adding a second pump does not stop there. More than 4 idlers are generated in a two-pumps/one-signal scenario. In fact, five extra idlers are generated, giving in total 9 idlers stemming from FWM processes. Of these 5 extra idlers, 3 are conserving the source signal information, all of which are non-degenerate FWM products where the three waves of frequencies f_i, f_j, f_k are employed only once in the creation of a fourth wave $f_{ijk} = f_i + f_j - f_k$. These three waves are located at f_{s12}, f_{s21} and f_{21s} . The final two idlers generated are degenerate-FWM products from the interaction of the two pumps and are located at frequencies f_{112} and f_{221} . Fig. 4-1(b) is a schematic illustrating all 9 idlers generated via FWM in a two-pumps/one-signal scenario.

4.1.4 Power Transfer in FWM-based Wavelength Multicasting

It is known that the time-averaged optical power $P_{ijk}(L, \Delta\beta_{ijk})$ of each generated idler through FWM at the frequency component $f_{ijk} = f_i + f_j - f_k$, with a propagation constant mismatch of $\Delta\beta_{ijk} = \beta_{ijk} - (\beta_i + \beta_j - \beta_k)$ can be written as [118]

$$P_{ijk}(L, \Delta\beta_{ijk}) = \frac{1024\pi^6}{n^4\lambda^2c^2} (\chi_{1111}\Theta)^2 \frac{P_i P_j P_k}{A_{eff}^2} \left| \frac{e^{(i\Delta\beta_{ijk}-\alpha)L} - 1}{i\Delta\beta_{ijk} - \alpha} \right|^2 \quad (4.1)$$

where χ_{1111} is the 3rd order nonlinear susceptibility. For a dispersion slope near zero, the propagation constant difference is $\Delta\beta_{ijk} = (2\pi c/f_k^2)\Delta f_{ik}\Delta f_{jk}D_k$ where D_k is the chromatic dispersion at f_k and $\Delta f_{mn} = |f_m - f_n|$, ($m, n = i, j, k$) [118]. The degeneracy factor $\Theta = 1, 3$ or 6 depending on whether three, two, or none of the frequencies f_i, f_j , and f_k are the same in the generation of the 4th frequency f_{ijk} , respectively. Consequently, non-degenerate FWM idlers have $\Theta = 6$ and degenerate FWM idlers exhibit $\Theta = 3$. By inspection of Eq. (4.1), for equal $\Delta\beta_{ijk}$'s, the power transferred to a non-degenerate idler is $(6/3)^2 = 4$ times that to a degenerate idler, or 6 dB greater. With the conversion efficiency G_c defined as the power of each idler after a SNW of length L with respect to a reference power being that of the data signal [114]

$$G_c \text{ [dB]} = P_{idler}(z = L) \text{ [dBm]} - P_{signal}(z = L) \text{ [dBm]} \quad (4.2)$$

one can expect that the 2 degenerated FWM idlers have a G_c that is approximately 6 dB smaller than non-degenerate idlers.

4.1.5 Experimental Setup

The experimental setup is depicted in Fig. 4-2. It is similar to that reported in [117], with an additional pump laser allowing multicasting. Further details on the SNW can also be found in [117].

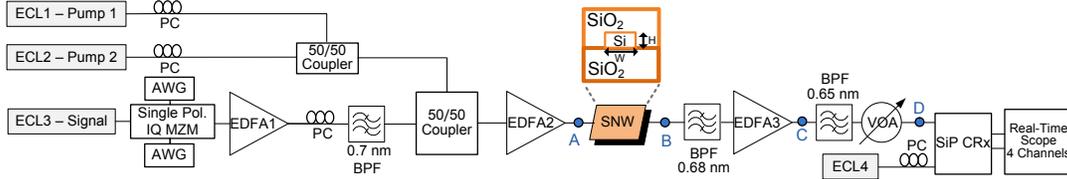


FIGURE 4-2: Experimental setup for Wavelength Multicasting, with cross-section of SNW.

A tunable external cavity laser operating at $f_s=193.60$ THz (1548.5 nm) is modulated by a single polarization IQ modulator driven by two arbitrary waveform generators providing 4-level signals at 22 GBaud, generating an optical 16-QAM signal over a single polarization. The modulated signal is amplified with an erbium-doped fiber amplifier booster (EDFA₁), followed by a polarization controller (PC) and a tunable bandpass filter to eliminate out of band noise. Two pumps, also from ECL's, of frequencies $f_{p1}=194.30$ THz (1542.9 nm) and $f_{p2}=194.00$ THz (1545.3 nm) are individually polarization controlled before coupled together, and further coupled to the signal. The frequencies of f_s , f_{p1} and f_{p2} are such that all idlers are spaced by at least 100 GHz, a spacing much larger than the bandwidth of the 22 Gbaud signal. The combined three waves are amplified by a high power amplifier (EDFA₂). The total optical power launched in the SNW is 24.3 dBm, where the power of the signal, 'pump 1' and 'pump 2' are respectively 12.4, 19.8 and 21.9 dBm at point "A" in Fig. 4-2.

The two-photon absorption (TPA) and free-carrier effect induced propagation loss are measured to be negligible for input optical powers less than 24.5 dBm, while the linear absorption is 2 dB/cm. Polarizations controllers for the signal and for each pump are required to align the state of polarization of all the interplaying signals as they enter the waveguide. After the nonlinear medium, the desired multicast signal is filtered by a tunable wavelength tunable bandwidth filter (BPF), pre-amplified by EDFA₃ and band pass filtered again. Both BPFs exhibit an insertion loss around 4.5 dB and allow to isolate one channel for detection. A variable optical attenuator is inserted to vary the received signal power. The coherent receiver is a silicon photonics coherent receiver (SiP CRx) [8, 119], integrated with transimpedance amplifiers, where the local oscillator is

another ECL at 15.5 dBm at a frequency tuned to the channel being detected. A real-time oscilloscope provides two 8-bit analog to digital converters, sampling the 2 outputs of the SiP CRx at 80 GSa/s. Generic offline digital signal processing [120] for a 16-QAM format is applied to recover the transmitted symbols and compute the average BER. All ECLs employed have linewidths < 100 kHz. In order to maximize the optical power $P_{ijk}(L, \Delta\beta_{ijk})$ and the FWM efficiency of all idlers at the end of the SNW strip waveguide, the latter is engineered to exhibit a flat group velocity dispersion between $\lambda=1535$ and 1565 nm, valued around 140 $ps/nm/km$ as simulated in [117], minimizing $\Delta\beta_{ijk}$'s dependence to third order dispersion $\partial D/\partial\lambda, \forall i, j, k$. The fiber-to-fiber loss is 21 dB, and each vertical grating coupler has a 1-dB bandwidth of 37 nm.

4.1.6 Results and Discussion

Figure 4-3 shows the optical spectrum at the output of the SNW in solid-blue, at point "B" in Fig. 4-2. The figure also shows in dashed-red the optical spectra after each channel is isolated and amplified by EDFA₃, at point "C" in Fig. 4-2. The six red-dashed curves are for the 5 new idlers and the signal. All wavelengths of interest are labelled. The solid-blue curve, obtained using a resolution bandwidth of 0.06 nm, exhibits all the 9 new frequency generated from FWM, the 2 pumps, and the signal. This trace is obtained by redirecting all light from the SNW to the OSA in order to obtain the signal power of each output channel. The 6 dashed-red curves, all superimposed, are independent OSA captures as the central frequency of the BPF preceding EDFA₃ is tuned to isolate a channel.

The Q -factor derived from the measured BER of the five multicast signals, of the source signal, and of the source signal in back to back (B2B, connecting the output of the 50/50 coupler to point "D" in Fig. 4-2), are presented in Fig. 4-4, as a function of received power measured at point "D". The Q -factor and BER are related by $Q = 2^{1/2}erfc^{-1}(2 \times BER)$, with $erfc^{-1}$ being the inverse complementary error function [51]. Conversion below the FEC limit is obtained for all five idlers. A FEC code of 6.69% overhead like the Continuously-Interleaved Bose-Chaudhuri-Hocquenghem brings an input BER of 3.8×10^{-3} to below 10^{-15} [121].

Table 4-1 summarizes in the first four data-columns the signal power and the OSNR of each channel at both points "B" and "C" in Fig. 4-2. The power of all idlers after the SNW is weak, all below -25.7 dBm. With the insertion loss of the following BPF, the input signal power to the final EDFA₃ is always below -30 dBm. Even if the OSNR of each idler varies as much as 7 dB

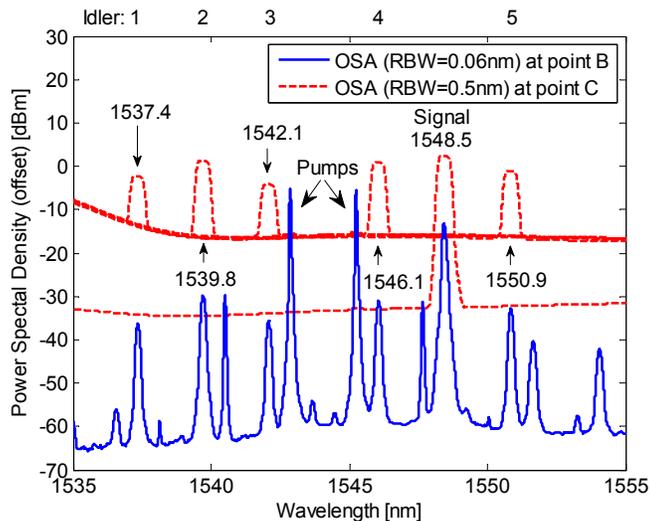


FIGURE 4–3: Optical spectrum at point B (solid-blue) and at point C (dashed-red). Spectra obtained at point “C” are for independently isolated wavelengths.

TABLE 4–1: Power and OSNR in 0.1 nm at points “B” and “C”, with conversion efficiency, required power and power penalty of each channel.

WL [nm]	Pow. at B [dBm]	OSNR at B [dB]	Pow. at C [dBm]	OSNR at C [dB]	Conv. Eff. [dB]	P_{req} [dBm]	P_{pen} [dB]
1537.36	−32.2	28.0	−2.7	17.9	−22.6	−19.8	8.0
1539.74	−25.7	34.9	1.2	23.9	−16.1	−25.2	2.6
1542.14	−31.6	28.3	−4.5	18.7	−22.0	−20.7	7.1
1546.14	−27.1	30.0	0.7	23.6	−17.5	−26.0	1.8
1548.52	−9.6	43.8	2.4	41.5	— — —	−27.5	0.3
1550.92	−28.8	29.0	−1.2	22.4	−19.2	−25.3	2.5

after the SNW between degenerate and non-degenerate idlers, their weak power normalizes their noise floor after EDFA₃. While all idlers receive a gain greater than 31 dB, their OSNR before and after the EDFA decreases by as much as 11 dB, caused by the very small signal power of the FWM idlers seeding the EDFA. However, for the source signal having much higher power seeding EDFA₃ (−14.1 dBm), the latter worsens its OSNR by a mere 2.2 dB, albeit providing a smaller gain of 16.5 dB. From the spectral location of the pumps and signal, the $\Delta\beta_{ijk}$ of the degenerate idler from Pump 1 ($\lambda = 1537.4$ nm) should be $0.7^2/0.4^2 \approx 3$ times that of the degenerate idler from Pump 2 ($\lambda = 1542.1$ nm), giving a theoretical smaller power for the former, as observed in Table 4–1. The signal power after the SNW of the two degenerate idlers is on average 4.9 dB smaller than that of

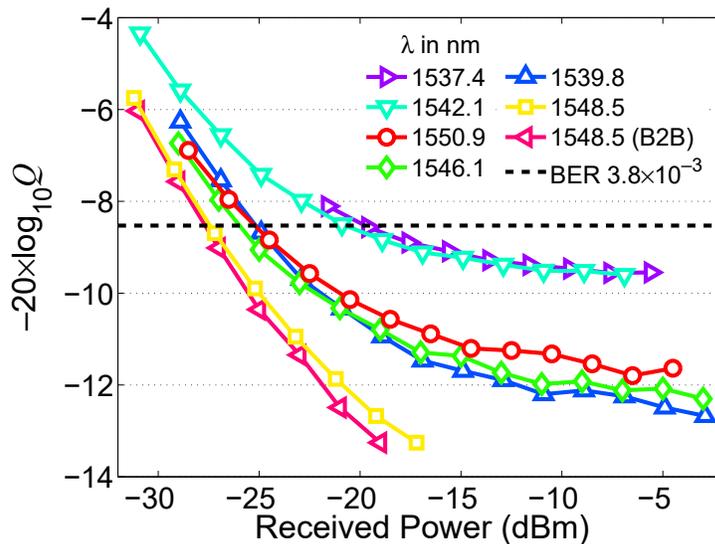


FIGURE 4–4: Sensitivity: Q -factor as a function of received power for the 7 cases.

the three non-degenerates. Moreover, the OSNR of degenerate idlers is on average 5.0 dB smaller than that of non-degenerates. Those numbers approach well Eq. (4.1)’s prediction; the offset from 6 dB can be explained by the non-equal phase mismatch $\Delta\beta_{ijk}$ ’s that is assumed when calculating the power ratio of degenerate and non-degenerate idlers from Eq. (4.1).

Conversion efficiency of each idler is obtained by subtracting the signal power at “B” (–9.6 dBm) from the idlers’ power at “B” in Table 4–1. They are summarized in the fifth data-column of Table 4–1. The sixth data-column shows the required power P_{req} for each channel to perform at the BER FEC limit, as found in Fig. 4–4. The power penalty P_{pen} is obtained by subtracting from that column the required power in B2B (–27.8 dBm). P_{pen} is summarized in the last column. The worst power penalty is 8 dB at 1537.4 nm. The average power penalty is 2.2 dB for non-degenerate idlers and 7.5 dB for degenerates, giving an average excess penalty of 5.3 dB for degenerate idlers.

Table 4–1 also depicts the dependence of the OSNR after EDFA₃ with the signal power before. There is a monotonic relation between the two; the higher the signal power before, the higher the OSNR after. As we are attenuating the signal power before the receiver to values smaller than the ones reported in Table 4–1 at point “C”, the limiting parameter dictating each idler’s BER performance becomes the idler’s OSNR after the last EDFA. From this monotonic relation, improving the idler’s power after the SNW directly translates to improving the performance. For wavelength multicasting using a SNW, the improvement of the idlers power is normally reported

via the nonlinear conversion efficiency metric. The coupling loss out of the nonlinear medium is also a decisive parameter governing the BER performance. Using values in Table 4–1, a decrease of conversion efficiency by 5 dB for all idlers, due to either a larger $\Delta\beta_{ijk}$ mismatch or a worst input coupler loss, would still provide better performance if accompanied by a decrease in the output coupler’s loss by 5 dB.

Higher nonlinear conversion efficiency is achievable by minimizing the input coupler’s loss, engendering an increase of $P_i P_j P_k$ in Eq. (4.1). Higher G_c ’s raise the idler’s power from the optical white noise, increasing its OSNR at the end of the nanowire. This effort is in vain if the output coupler’s loss is high. Minimizing this loss allows seeding the last EDFA with higher input signal power. Both phenomena will dictate the final OSNR, and consequently the BER performance.

Significant reduction in power consumption and footprint for WM in silicon is achievable by removing both EDFAs before the SNW. To maintain the same power for the pump and signal entering the waveguide, the use of lower loss 1D grating couplers (e.g. with an insertion loss below 1 dB [122]) is required. Furthermore, the 3 dB couplers can be replaced by low loss silica-based arrayed-waveguide gratings [123]. These would provide better wavelength multiplexing and system integration for WM.

4.1.7 Conclusion

Using the nonlinear silicon waveguide presented in [117] for one-to-one WC, we have demonstrated 6-fold multicasting of a high order 16-QAM signal at 22 Gbaud, using 2 pumps and exploiting multiple non-degenerate and degenerate FWM processes. All multicast signals perform below the FEC BER threshold of 3.8×10^{-3} . This was the first demonstration of wavelength multicasting of higher QAM signals beyond QPSK in a SNW. We have discussed how the FWM nonlinear process impacts differently the performance of degenerate and non-degenerate multicast idlers and experimentally confirmed their efficiencies with theory. We showed that the output grating coupler’s loss is a critical parameter governing the BER performance of the idlers, along with the nonlinear FWM conversion efficiency. Properties and results using a SNW demonstrate the capability of a silicon based approach to achieve all-optical multicast switching, highlighting the potential for future on-chip optical interconnection and processing systems.

4.2 Silicon Photonics Intensity Modulator

In the previous section, we presented a passive silicon photonics device for on-chip wavelength multicasting. In this section, we present an active silicon photonics device where electronics interplay with the properties of free-carrier in doped silicon to electronically alter the properties of the lightwave traveling through the device. We introduce what was the first Mach-Zehnder modulator made of doped silicon that allows intensity modulation at a rate of 112 Gb/s.

4.2.1 Introduction

The need for faster and inexpensive short reach optics for data centers is growing rapidly due to the incessant spread of cloud services offered by data centers. For an equal bit rate, modulating a single wavelength from a single laser provides a solution that is more cost effective than a hardware multiplexed counterpart which involves the use of multiple wavelengths or fiber lanes to achieve the total desired bit rate. In addition, the single wavelength approach is the more scalable solution to cope with the ever-increasing capacity demand. Although 10 Gb/s transmission of Ethernet frames over single mode fiber has been standardized as 100GBASE-LR4 by using 4×25 Gb/s in a WDM setting employing four colors in the 1300 nm band [124], it is widely agreed that achieving 100 Gb/s on a single wavelength using inexpensive and power efficient components is a mandatory building block in order to realize the envisioned 400 Gb/s bit rate based on 4×100 Gb/s [125].

Recently, 100 Gb/s short-reach transmission experiments have been reported using various advanced modulation formats, such as 4-level pulse amplitude modulation (PAM-4) with polarization division multiplexing [126], half-cycle Nyquist subcarrier modulated 16-ary quadrature amplitude modulation (16-QAM) with polarization division multiplexing [47], multi-band carrierless amplitude phase modulation [127], discrete multi-tone [128] and a performance comparison between some of these formats has been also reported in [129]. More recently, 400 Gb/s was also experimentally demonstrated as 4×100 Gb/s using DMT and 4 directly modulated lasers (DMLs) over 30 km in [46], and using multi-CAP modulation and 4 integrated electroabsorption-modulation laser (EML) over 40 km in [130], both in the O-band.

In addition, Silicon photonics (SiP) is rapidly emerging as an inexpensive solution for a variety of applications, including coherent links [131] and short reach (<10 km) optical interconnects [132]. Silicon is a useful host material for fabricating the external modulator because of its low cost, high yield, large scale integration, and CMOS compatibility, enabling the integration of modulators

with the RF drivers [133]. Recently in [37], a dual-drive SiP modulator operating at 30 GHz bandwidth near 1300 nm was demonstrated and used at a 50 Gb/s rate (using on-off keying). More recently in [36], we reported a single-drive series push-pull SiP modulator at 1310 nm based on our work in [29] and presented some preliminary transmission results using PAM-4 modulation at 107 Gb/s. In fact, externally modulated lasers using SiP modulators such as reported in [36] not only provide a larger bandwidth than directly modulated lasers, they also add the potential of chirp-free transmission.

In this section, we present a detailed system-level experimental study carried out using the series push-pull SiP modulator working at 1310 nm that is presented in [36]. A bit rate of 112 Gb/s is achieved using eight-level pulse amplitude modulation (PAM-8) at 37.4 Gbaud after 10 km of SMF fiber below the hard decision forward error correction threshold of 3.8×10^{-3} , allowing a final bit error rate under 10^{-15} after error correction. To the best of our knowledge, this is the first demonstration of error-free amplitude modulated, 100 Gb/s transmission using a SiP modulator. We also present exhaustive system performance tests where we vary multiple parameters for multi-level PAM- N formats, namely the PAM order, the bit rate, the distance, the pulse shaping roll-off factor, and the received signal power. The DSP algorithms applied at the transmitter and receiver, allowing the high throughput and also a fully software-defined short reach link, are presented. Finally, a model for the Q -factor applicable for multilevel PAM signaling is presented and accurately validated experimentally against the BER performance metric.

The remainder of the section is sub-divided as follows. Section 4.2.2 presents the packaged silicon photonic intensity modulator designed to operate at 1310 nm. Section 4.2.3 details the experimental setup employed to demonstrate the use of the SiP modulator for short reach transmission using PAM- N modulation. Section 4.2.4 presents the digital signal processing applied at the transmitter and receiver for optimum transmission performance of PAM formats. In Section 4.2.5, we present a new model for Q -factor computation applicable to multilevel PAM- N formats, followed by comprehensive experimental system results. We conclude in Section 4.2.6.

4.2.2 Fully Packaged SiP Modulator at 1310 nm

The 1310 nm SiP modulator uses carrier depletion in pn junctions. It is based on a series push-pull configuration [38,134,135] and has been fabricated at A*STAR IME [136] within the framework of the Canadian Silicon Electronic-Photonic Integrated Circuits (Si-EPIC) program [137]. The

schematic and a photograph of the device are shown in Figs. 4–5 and 4–6, respectively. This kind of modulator has already been demonstrated in silicon at 1550 nm [38, 134]. The light is coupled into and out of the chip using surface grating couplers and guided in 220 (height) \times 400 nm (width) strip waveguides. Multimode interference (MMI) couplers are used to split and recombine the light in a Mach-Zehnder configuration. In the active portion of the modulator, the light is guided over two rib waveguides (slab height of 90 nm) used to support the pn junctions. The series push-pull configuration is obtained by connecting in series the two pn junctions, which are driven by a single modulating signal. This series combination reduces the capacitance and increases the speed. Of course, in this configuration, the resistance of the series combination is increased as well. In the center portion of the modulator, the two junctions are thus in contact, through a common n -doped region. A DC bias voltage is applied to this common n -doped region to ensure that the pn junctions are operated in reverse bias at all time. This configuration results in a push-pull action minimizing the chirp, although this is not a critical issue at the selected wavelength of operation.

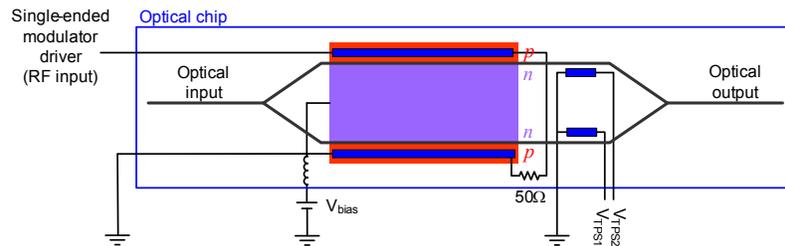


FIGURE 4–5: Schematic of the series push-pull Mach-Zehnder modulator.

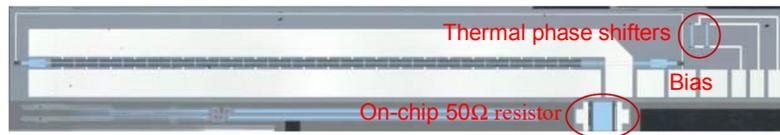


FIGURE 4–6: Photograph of the 6 mm long series push-pull modulator. The portion of the layout supporting the grating couplers is not shown.

The series push-pull configuration was proposed by Walker and first demonstrated in GaAs [135], and, later, in InP [138, 139]. In both GaAs and InP, the two diodes are vertically grown onto a n -type common substrate while in SiP, the pn junctions in series are readily formed onto the silicon substrate, without any epitaxial regrowth, requiring only rib waveguide patterning and doping. Compared to dual-drive modulators, series push-pull modulators require a single driver and

no bias-tee (unless the former is used at null bias [37]). For high-speed operation, a RF traveling-wave coplanar stripline electrode is used, which is carefully designed to have an index close to that of the optical group index of the rib waveguide of 3.8, once loaded with the pn junction structure. The pn junctions are further segmented into 36 sections, for a modulator length of approximately 6 mm. The biasing scheme required by the series push-pull modulator also avoids DC current to be supplied to the line-matching termination as opposed to dual-drive modulators. In our design, the line termination is integrated on-chip (see Fig. 4–6). A value of 50Ω has been measured for the load. A small path mismatch of $150 \mu\text{m}$ is introduced in the interferometer, leading to a 2.6 nm Free-Spectral-Range, to ease the characterization of the modulator. Two thermal phase shifters (TPS) are also incorporated within the interferometer to adjust the operation point at quadrature.

The modulator optical transmission spectrum is shown in Fig. 4–7(a) for different sets of DC voltages applied to the two pn junctions (V_{d1} and V_{d2}). These voltage drops across the junctions are produced by applying properly adjusted DC sources, as illustrated in Fig. 4–7(b). One of these DC sources, V_{bias} , is applied to the bias electrode of the modulator (i.e. biasing the n -doped common region); the other DC source, $V_{\text{RF input}}$, is applied to the RF input electrode of the modulator. One can see that the optical spectrum of the modulator displays a good extinction ratio of more than 20 dB, indicating a good balance of the power from each arm after recombination by the output MMI coupler of the Mach-Zehnder interferometer. The set of curves shown in Fig. 4–7(a) reproduces the push-pull operation around a bias voltage of 4 V ($(V_{d1} + V_{d2})/2 = 4 \text{ V}$) [138]. The blue curve shows the optical spectrum when the two diodes are polarized with 4 V. The purple and light blue curves display the result of decreasing the reverse voltage on one junction (V_{d1}) while increasing the reverse voltage on the other (V_{d2}). The red and green curves are obtained when the role of the two diodes are reversed. The analysis of these results allows for a determination of V_π equal to 6.5 V at a DC bias voltage of 4 V. For this nominal bias voltage, V_π varied from 5.7 V to 6.5 V with a mean of 5.9 V on seven chips characterized. By repeating these measurements for different sets of voltages, the DC V_π can be extracted as a function of the bias voltage. The result is shown in Fig. 4–7(c). It can be noted that a different procedure was used in [38] to obtain V_π . In that case, a single voltage was applied to the pn junctions in series. Using this alternative method, the phase shift obtained is that shown in Fig. 4–7(d), from which a V_π of about 4 V is deduced. It is important to note that this alternative method provides a V_π value corresponding

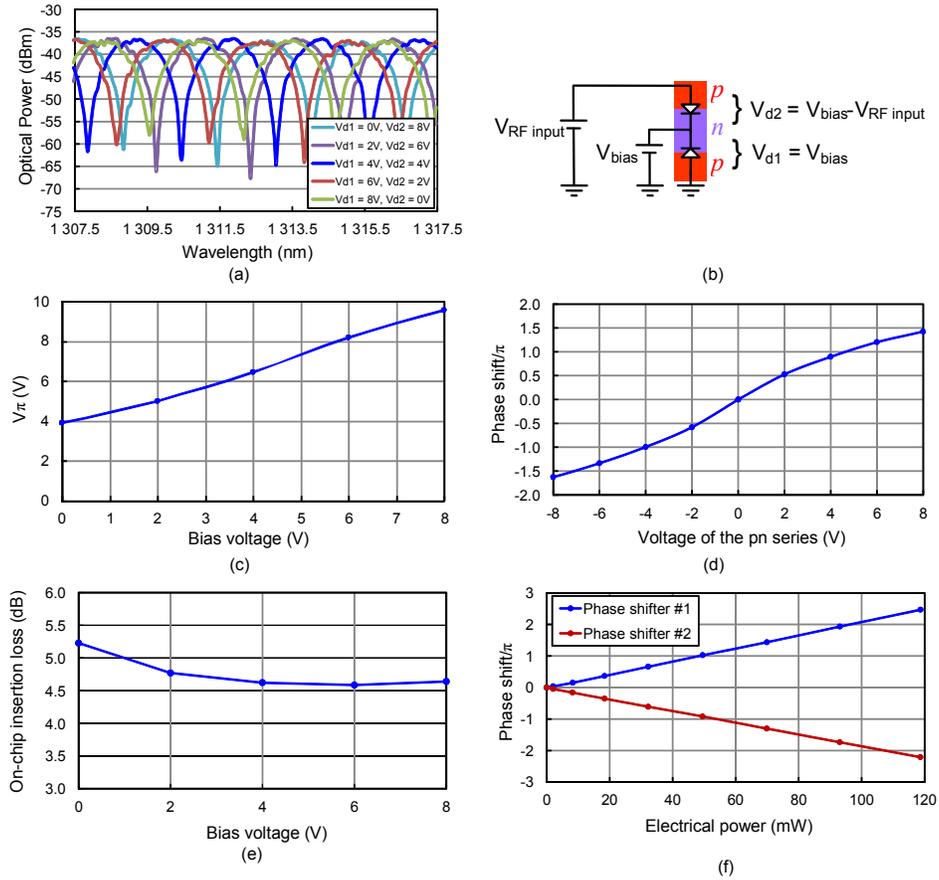


FIGURE 4–7: Characterization of the series push-pull modulator: (a) transmission spectrum at different applied voltages; (b) DC circuit used to mimic the push-pull operation (c) V_{π} as a function of the bias voltage applied to the central electrode; (d) phase shift as a function of the voltage applied to the pn junction series (e) on-chip insertion loss (f) thermal phase shifter efficiency.

to an operation around a bias voltage of 0 V and does not permit a determination of V_{π} at larger bias voltages.

The modulator optical loss is 4.6 dB (at maximum transmission) at a bias of 4 V, excluding the 9 dB fiber coupling loss. Figure 4–7(e) presents the loss dependency to bias voltage. The characterization of the thermal phase shifters, shown in Fig. 4–7(f), results in an efficiency of 50 mW/ π .

The small-signal frequency response of the modulator with the on-chip load wire bonded is presented in Fig. 4–8(a) (blue curve). The curve is normalized to the response at the reference frequency of 1.5 GHz. The bias voltage is 4 V and the modulator is operated at the quadrature

point. A bandwidth slightly above 20 GHz is obtained. The corresponding electrical reflection (S_{11}) spectrum is shown in Fig. 4–8(b) (blue curve), better than -10 dB over a 32 GHz range. After the

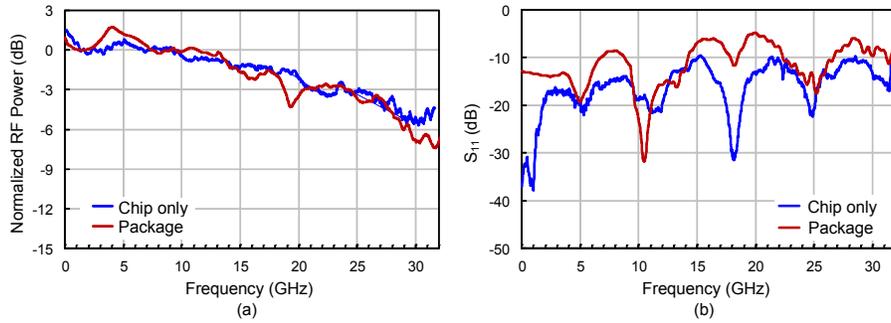


FIGURE 4–8: (a) EO frequency response of the modulator chip (blue) and of the packaged modulator (red). (b) S_{11} of the modulator chip and of the packaged modulator.

initial characterization discussed above, the chip was fully packaged as shown in Fig. 4–9. Optical fibers were attached to the chip using a method outlined in [29], without a noticeable insertion loss penalty. The input fiber is a polarization maintaining fiber. The chip is mounted in a package with access to the modulator electrode through a GPPO connector and also to the thermal phase shifters and DC bias connection points. The frequency response and S_{11} characteristics of the packaged modulator are slightly degraded, as illustrated by the red curves in Fig. 4–8 (a) and (b), respectively.



FIGURE 4–9: Picture of the fully packaged SiP modulator.

4.2.3 Experimental Setup

We present in Fig. 4–10 the experimental setup employed to demonstrate the use of the SiP modulator presented in the previous section for short reach transmission using PAM- N modulation. A 20 mW 1310 nm DFB laser with a measured linewidth of 440 kHz is modulated by the SiP

intensity modulator. The RF signals driving the modulator are first generated from a Digital to Analog Converter. This AC-coupled 8-bit DAC runs at 70 Gsamples/s and feeds the differential input of an Inphi IN3214SZ linear driver. The amplified RF signal is then applied to the RF input of the SiP modulator which was constantly operated at the quadrature point of the optical transfer function by appropriately tuning the thermal phase shifters. The modulated signal is launched into various lengths of single mode fiber: either 0, 2, 10 or 20 km of Corning SMF-28e+ fiber. An AC-coupled Picometrix PT-40D PIN+TIA receiver is employed before an Agilent DSO-X 96204Q real-time oscilloscope serving as an 8-bit Analog to Digital Converter sampling at 80 GSamples/s. The DAC has a 3 dB bandwidth of 14 GHz and a maximum output differential swing of 1.2 Vppd.

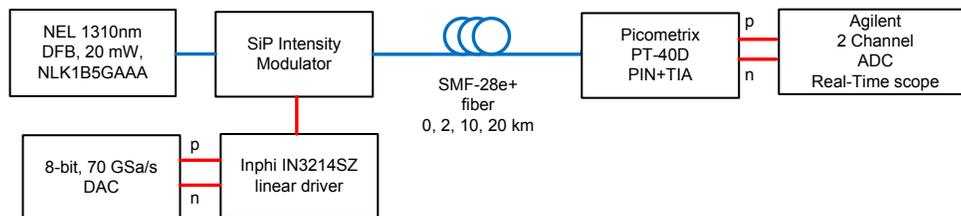


FIGURE 4–10: Experimental setup to test the packaged SiP modulator.

The linear driver has a gain of 17 dB, maximal single-ended output swing of 6 Vpp, and a 3 dB bandwidth of 34 GHz. The 3 spools of different fiber lengths exhibit a loss of 0.34 dB/km at 1310 nm. The PIN+TIA receiver has a 3 dB bandwidth of 35 GHz and a conversion gain of 1.3 V/mW, and the real-time scope has a 3 dB bandwidth of 33 GHz.

In the forthcoming sections, we compare the performance of different PAM orders over variable fiber lengths in order to reach a 100G Ethernet transport rate. To achieve this information transfer rate, the transmitter needs to provide 112 Gb/s. From the OTU4 standard, a client payload of 99.5328 Gb/s is transmitted at a line rate of $255/227 \times 99.5328 = 111.80997$ Gb/s [140], including a 6.7% overhead for the forward error correction. In this work, we assume FEC encoding and decoding at the transmitter and receiver, respectively. FEC is required in currently deployed 100 Gb/s metro and long-haul fiber optic transmission systems because it enables a significant reduction of the acceptable raw BER at the expense of signal overhead. The 6.7% FEC overhead in the OTU4 standard produces a net coding gain of 9.19 dB at a corrected BER of 10^{-15} , translating to a pre-FEC BER threshold of 3.8×10^{-3} . Measured BER values below this threshold are regarded as ‘error-free’ in the context of optical transport networks [105, 141].

In this work we focus on a comparison of PAM-4 and PAM-8 as a 100G candidate, but also collect PAM-2 (On-Off Keying) and PAM-16 data for completeness. The 8-bit DAC, with its maximum memory of 2^{18} samples, supports the generation of these differing PAM orders at varying symbol rates.

4.2.4 Digital Signal Processing at the Transmitter and Receiver

As the system employs a DAC and an ADC, digital signal processing can be applied at both the transmitter and the receiver. This section describes the different DSP schemes applied at both ends of the transmission system, from symbol generation to symbol detection, in order to best generate and recover the desired waveforms.

4.2.4.1 Transmitter DSP for PAM- N Formats

We first introduce the required processing to apply at the transmitter on the multilevel data symbols. There are 4 processes to apply on the data symbols before the 8-bit waveform quantization: 3 linear processes and one nonlinear. The processes are applied sequentially on the symbols generated, where we assume a fixed sampling rate of 70 GSa/s. The symbols for PAM- N format come from an alphabet of integers from $-N+1$ to $N-1$ by jumps of 2. The first process, for a desired symbol rate of R_B in Gbaud, is to upsample the symbol stream from 1 sample per symbol (SPS) to $70/R_B$ SPS. This upsampling is readily done in the frequency domain. Secondly, the desired pulse shaping filter is applied. This pulse shaping filter $h_{RRC}[n]$ is typically a root-raised cosine filter. From the RRC pulse theory, the frequency (in GHz) at which the pulse shaping filter's power transfer is reduced by half is $1/2R_B$, and the maximum spectral content is located at $1/2R_B(1-\alpha)$, where α is called the roll-off factor of the pulse shaping filter [90]. Consequently, for a desired symbol rate of R_B , the roll-off factor can only be within the range $0 < \alpha < \min(1, 70/R_B - 1)$ in order to avoid aliasing and respect the Nyquist sampling theorem [142]. Thirdly, the inherent nonlinear raised cosine power transfer function of the SiP Mach-Zehnder intensity modulator is compensated for. This process equalizes the spacing between the optical power levels after modulation by applying an *arcsin* function to the waveform. This nonlinear transform will be cancelled and linearized after the MZI with direct detection. Finally, the last step is to pre-compensate the frequency response of all the components at the transmitter that precede the SiP modulator. The concatenated linear analog response of the DAC, RF driver, and SiP modulator can all be compensated using one lumped

inverse response function, $h_{\text{inv}}[n]$. It is important to mention that the main frequency constraining component on the transmitter side is the DAC with a 3 dB bandwidth of only 14 GHz. From the Nyquist sampling theory, the content out of the DAC can be manipulated anywhere from DC to 35 GHz, when sampling at 70 GSa/s. It can be difficult and unnecessary to fully equalize the frequency response all the way to 35 GHz, as 1) this frequency exhibits very low output swing and 2) the pulse shaping filter can be designed such that its maximum frequency $1/2R_B(1 - \alpha)$ is less than 35 GHz. Naturally, one drawback of this analog response compensation is a reduction of the output peak-to-peak voltage, leading to smaller driving signals for the SiP modulator, itself leading to reduced level separation of the optical PAM- N signal. Figure 4–11(a) shows the sequence of DSP operations applied at the transmitter.

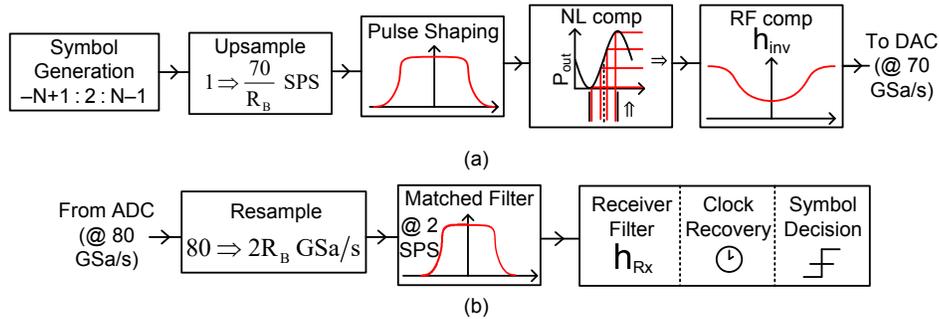


FIGURE 4–11: DSP stack at the (a) transmitter and (b) receiver side.

4.2.4.2 Receiver DSP for PAM- N Formats

At the receiver, the required signal processing to be performed off-line is straightforward, where the only major process is linear filtering. The processes are detailed as follows. First, we resample from the fixed ADC rate of 80 GSa/s to twice the symbol rate $2R_B$. Secondly, we apply the matched filter, defined at 2 SPS. The stream of samples is then filtered by a short linear FIR filter h_{Rx} , with constant coefficients stored in the receiver. During first system startup, the coefficients of this receiver filter can be unknown, in which case all coefficients are set to “0” except the central tap which is set to “1”. The optimum coefficients that will later remain constant and be stored can be obtained in a blind decision-directed scheme using a short period of a few hundred symbols. Once the coefficients are determined, h_{Rx} is subsequently linearly applied to the input samples. The receiver also has to implement a digital clock recovery algorithm to recover the transmitter clock and apply symbol decision at proper sampling instants. Clock recovery at the receiver is used

to compensate any sampling phase and frequency offset that may exist between the transmitter and receiver clocks. As both clocks are free running, a digital clock recovery is needed for residual clock offset removal. For this task, we employ the blind feedforward timing error estimator in [143] which operates at 2 samples/symbol and is transparent to modulation format, operating for any PAM- N format. The timing error estimate provided by the estimator, updated on a block-by-block basis, is then used to control a piecewise parabolic interpolator in a recursive fashion similar to the one explained in [144]. The final output symbols are used to compute the desired performance metric, including the mean signal to noise ratio, the mean bit error rate or the equivalent Q -factor. Figure 4–11(b) shows the sequential DSP blocks applied at the receiver.

4.2.5 Performance and Results

In this section, we present the metrics employed to assess the performance of the system. We first start by defining a Q -factor for multilevel PAM signal, followed by system performance tests. We use both this Q -factor and the BER from error counting metrics to assess performance.

4.2.5.1 Q -factor for PAM- N Formats

System performance is often reported using the Q -factor metric, where the noise distribution is assumed to follow a Gaussian distribution. The metric is well suited for binary On-Off signaling, i.e. PAM-2, where the received sampled value I fluctuates from symbol to symbol around an average value of μ_1 or μ_2 depending on whether the bit corresponds to 0 or 1 in the binary symbol stream, respectively, with respective noise variance for each level of σ_1^2 and σ_2^2 . For this format, the BER is found to be [69]

$$BER_Q = \frac{1}{2} \left(\frac{1}{2} \operatorname{erfc} \left(\frac{\mu_2 - I^{th}}{\sigma_2 \sqrt{2}} \right) + \frac{1}{2} \operatorname{erfc} \left(\frac{I^{th} - \mu_1}{\sigma_1 \sqrt{2}} \right) \right) \quad (4.3)$$

where the optimum decision threshold I^{th} is

$$I^{th} = \frac{\mu_1 \sigma_2^2 - \mu_2 \sigma_1^2 + \sigma_2 \sigma_1 \sqrt{(\mu_2 - \mu_1)^2 + 2(\sigma_2^2 - \sigma_1^2)(\ln(\sigma_2/\sigma_1))}}{\sigma_2^2 - \sigma_1^2} \quad (4.4)$$

and where erfc is the complementary error function and \ln is the natural logarithm. With the assumption $(\mu_2 - \mu_1)^2 \gg (\sigma_2^2 - \sigma_1^2)$ in Eq. (4.4), I^{th} simplifies to $I^{th} = (\mu_1 \sigma_2 + \mu_2 \sigma_1) / (\sigma_1 + \sigma_2)$ and the argument of each erfc of Eq. (4.3) become identical: $(\mu_2 - I^{th}) / \sigma_2 = (I^{th} - \mu_1) / \sigma_1 \stackrel{\text{def}}{=} Q$. This Q can be rewritten as $Q = (\mu_2 - \mu_1) / (\sigma_2 + \sigma_1)$ and is also called Q -factor. By inverting Eq. (4.3),

one can compute the Q -factor from the knowledge of the BER_Q with

$$Q = \sqrt{2} \operatorname{erfc}^{-1}(2 BER_Q) \quad (4.5)$$

It is important to point out that BER_Q in Eq. (4.3) with the subscript Q is used to denote the theoretical BER when two levels μ_1 and μ_2 are corrupted by Gaussian noise with variances of σ_1^2 and σ_2^2 , respectively. This BER_Q is minimized when I^{th} equals Eq. (4.4). The reason we introduce the subscript Q in BER_Q is to distinguish this BER from the BER calculated from bit error counting denoted by BER , while BER_Q and Q -factor represent the same metric, cast differently and related via Eq. (4.5). In the forthcoming sections, we will use two performance metrics namely BER and Q -factor. BER represents the BER obtained from counting the number of bit errors. This counting is performed offline after all receiver DSP processes and with knowledge of the transmitted bit stream. BER_Q , and therefore Q -factor, will be derived using a new definition of BER_Q for multilevel PAM- N formats that considers the statistics μ and σ^2 of all levels, instead of just the 2 levels in Eq. (4.3). Finally, the Q -factors reported are always cast in dB scale, i.e. $10 \times \log_{10}(Q)$.

We define here a BER_Q model that considers the statistics μ and σ^2 of all levels of a multilevel format. Moreover, instead of the single optimum threshold I^{th} given in Eq. (4.4) for ON-OFF keying, $N - 1$ optimum thresholds have to be defined for a PAM- N format, where each optimum threshold uses the statistics μ and σ of its two closest levels. Using a similar approach as that for PAM-2 in Eq. (4.3), the analytical BER_Q for PAM- N formats is defined as

$$BER_Q = \frac{1}{\log_2(N)} \sum_{i=1}^N p(I_i) (P(I_{i-1}|I_i) + P(I_{i+1}|I_i)) \quad (4.6)$$

where N is the PAM order, I_i is symbol level i out of N possible levels, $p(I_i)$ is the probability of transmitting symbol level i , and finally $P(I_{i-1}|I_i)$ or $P(I_{i+1}|I_i)$ are the probabilities of deciding I_{i-1} or I_{i+1} when I_i was transmitted, respectively. $1/\log_2(N)$ is the number of erroneous bits when a wrong symbol decision is made, assuming Gray coding and that erroneous symbol decisions are made on either of the two closest neighboring symbols. Of course, for the first and last symbol levels I_1 and I_N , erroneous symbol level decision can only be made on one side and $P(I_0|I_1) = P(I_{N+1}|I_N) = 0$. As $p(I_i)$ is equiprobable for all levels i , $p(I_i) = 1/N$, $\forall i$. From the theory relating $P(I_{i\pm 1}|I_i)$ to

the error function *erfc*, one can show that

$$BER_Q = \frac{1}{N \log_2(N)} \frac{1}{2} \sum_{i=1}^N \left[\operatorname{erfc} \left(\frac{|\hat{\mu}_i - I_{i-low}^{th}|}{\hat{\sigma}_i \sqrt{2}} \right) + \operatorname{erfc} \left(\frac{|\hat{\mu}_i - I_{i-high}^{th}|}{\hat{\sigma}_i \sqrt{2}} \right) \right] \quad (4.7)$$

where $\hat{\mu}_i$ is the average value of the received symbols that were transmitted as I_i and $\hat{\sigma}_i^2$ is their variance. Both $\hat{\mu}_i$ and $\hat{\sigma}_i^2$ are computed and obtained at the receiver before the hard decision. The quantities I_{i-low}^{th} and I_{i-high}^{th} are the optimum decision level thresholds between I_{i-1} and I_i , and between I_{i+1} and I_i , respectively. From this analytical BER_Q obtained using solely the first 2 statistical moments of all levels, $\hat{\mu}_i$ and $\hat{\sigma}_i^2$, we compute the equivalent Q -factor using Eq. (4.5). This gives the equivalent Q -factor for multilevel PAM signaling from analytical BER_Q derivation.

4.2.5.2 System Performance

In this subsection, we present the system-level performance of the short-reach optical link that was depicted in Fig. 4–10. The system performance is assessed qualitatively with the aid of eye diagrams and quantitatively in terms of three performance metrics: Q -factor according to its definition in Section 4.2.5.1, BER from error counting, and signal-to-noise ratio which is defined, for a PAM- N signal, as the ratio of the average signal power to the average power of the noise riding on top of the signal. For a PAM- N signal of possible values $-N + 1$ to $N - 1$ by jumps of 2, the mean signal power is $(N^2 - 1)/3$. For the average BER computation, a total of 10^7 bits were captured for each trace. Therefore, we assume the minimum BER that is reported with confidence is at around $100/10^7 \approx 10^{-5}$.

We first present in Fig. 4–12 a few sample eye diagrams for PAM orders of 2, 4, 8 and 16 at both low and high symbol rate, after 10 km of propagation. Eye diagrams are a good qualitative way to observe the quality of the received signal. The eyes are obtained after the receiver DSP presented in Fig. 4–11(b) is applied, that is after resampling the ADC output at $2R_B$ and applying the matched filter and linear receiver filter h_{RX} with clock recovery. To obtain multiple points within a symbol duration in the eye diagrams, we upsample the 2 SPS signal obtained after h_{RX} , without creating spectral content, by padding zeros in the frequency domain.

Above each eye diagram presented in Fig. 4–12, we give the PAM order, the bit rate, the BER (from error counting), the Q -factor and the roll-off factor α of the Tx and Rx pulse shaping and matched filters, respectively. Figures 4–12(a) and (b) show PAM-2 eyes at 30 and 60 Gb/s,

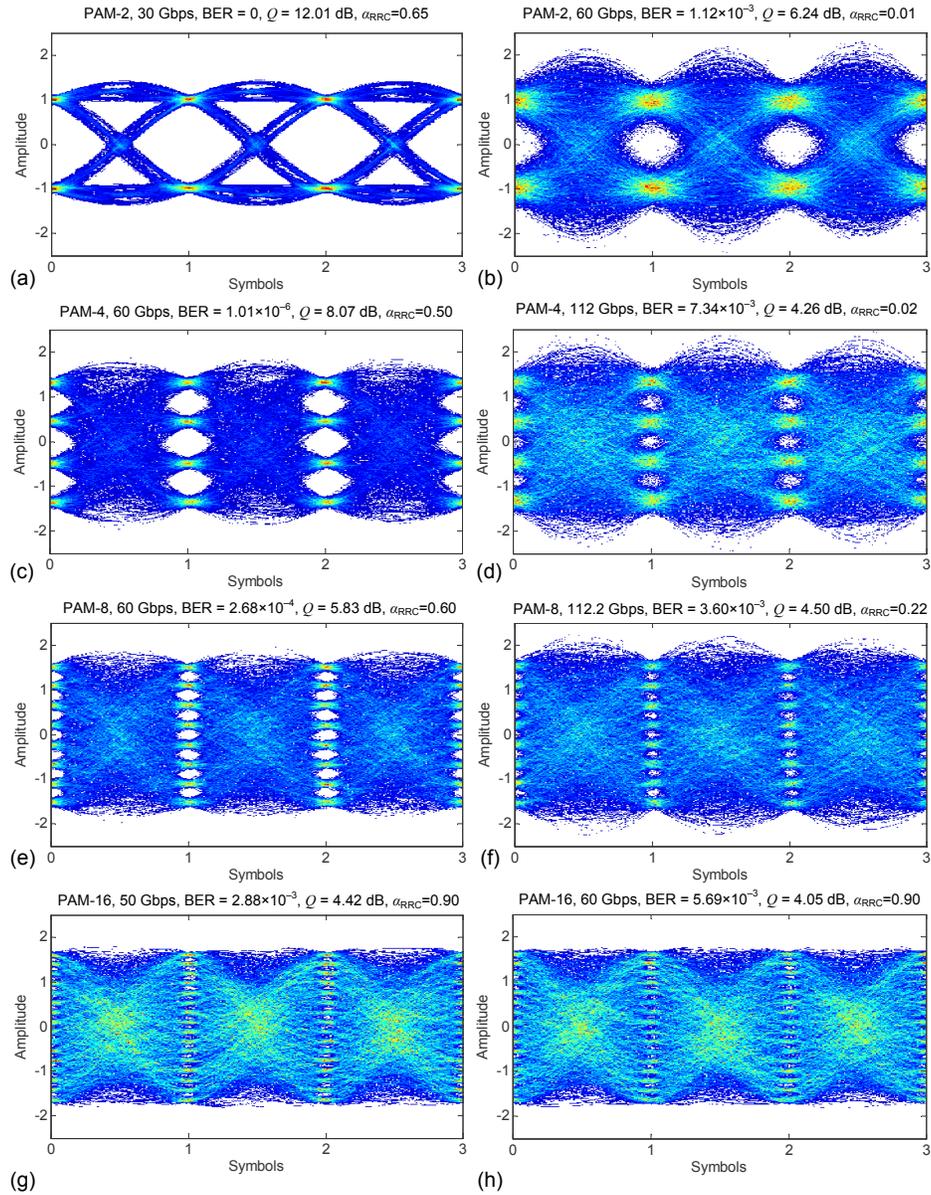


FIGURE 4–12: Eye diagrams, after 10 km, for (a) PAM-2 at 30 Gb/s, (b) PAM-2 at 60 Gb/s, (c) PAM-4 at 60 Gb/s, (d) PAM-4 at 112 Gb/s, (e) PAM-8 at 60 Gb/s, (f) PAM-8 at 112 Gb/s, (g) PAM-16 at 50 Gb/s and (h) PAM-16 at 60 Gb/s. All eyes obtained after receiver DSP.

respectively, Figs. 4–12(c) and (d) show PAM-4 eyes at 60 and 112 Gb/s, Figs. 4–12(e) and (f) show PAM-8 eyes also at 60 and 112 Gb/s, respectively, and finally Figs. 4–12(g) and (h) show PAM-16 eyes at 50 and 60 Gb/s, respectively. It is interesting to compare the eyes for 30 Gb/s PAM-2 and 60 Gb/s PAM-4, as they are generated at the same symbol rate. One can see that

PAM-2 at 30 Gb/s provides a signal of very high quality, with a BER that is most probably already below the desired error rate of 10^{-15} : the limited memory length of the ADC prevents accurate BER measurements from error counting at such high signal quality. Keeping the same signaling rate but doubling the number of generated levels from 2 to 4 still gives a rather good signal quality, but with an increase in the BER. It is also interesting to observe the eyes for PAM-16: even if eyes look very similar, an increase in bit rate from 50 Gb/s to 60 Gb/s doubles the BER from an ‘error-free’ rate of 2.88×10^{-3} to a BER falling above the BER FEC threshold. This shows the limited qualitative information provided by eye diagrams.

For a fixed symbol rate, jumping from PAM- N_k to a higher PAM- N_i naturally decreases the Q -factor by

$$5 \log_{10} \left(\frac{N_i^2 - 1}{N_k^2 - 1} \right) + \text{Excess loss [dB]} \quad (4.8)$$

where the mean signal power is assumed equal for both PAM formats, and where the standard deviation of the noise around each level is assumed constant notwithstanding the size of the input alphabet (N_k or N_i). From PAM-2 to PAM-4, $N_k = 2$ and $N_i = 4$, and the Q -factor drops by at least $5 \times \log_{10}((4^2 - 1)/(2^2 - 1)) = 3.5$ dB. This decrease arises due to neighboring levels being closer to each other as the PAM order increases. In our case, Q diminished by 3.94 dB, showing a small excess loss of signal quality of 0.44 dB by going from 2 to 4 level waveform generation at 30 Gbaud.

In the following, we test the accuracy of the BER_Q model of Eq. (4.7) presented in Section 4.2.5.1 and its equivalent Q -factor. The accuracy of the Q -factor, related to the BER_Q via Eq. (4.5), is tested by comparing the BER_Q of Eq. (4.7) with the BER obtained by error counting. Figure 4-13 shows good agreement between BER_Q and BER . The figure also demonstrates the accuracy of the Gaussian model, by showing the normalized histogram of the received symbols (blue curves), overlaid with the aggregate probability density functions (PDF) of each level i of the multilevel signal assuming Gaussian noise statistics. The accumulated PDF shown by the red curves in Fig. 4-13 is $\sum_{i=1}^N \mathcal{N}(\hat{\mu}_i, \hat{\sigma}_i^2)$, where $\mathcal{N}(\mu_i, \sigma_i^2)$ is a Gaussian distribution of mean μ and variance σ^2 . Results are shown for PAM-8 at 112 Gb/s after propagation of 10 and 20 km.

At both distances, BER from error counting and BER_Q from Eq. (4.7) show very good agreement. After 20 km, agreement is also high because the receiver is solely Gaussian noise

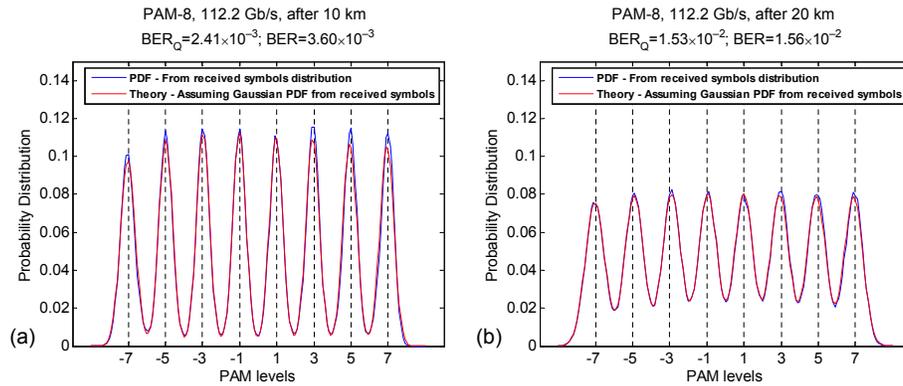


FIGURE 4–13: Histogram of received symbols and probability density function (PDF) of PAM-8 after (a) 10 km, giving a BER of 3.6×10^{-3} and after (b) 20 km, giving a BER of 1.56×10^{-2} .

limited. To show that the model fits any PAM order, Fig. 4–14 shows its accuracy for PAM-16 at 50 Gb/s, after 10 km.

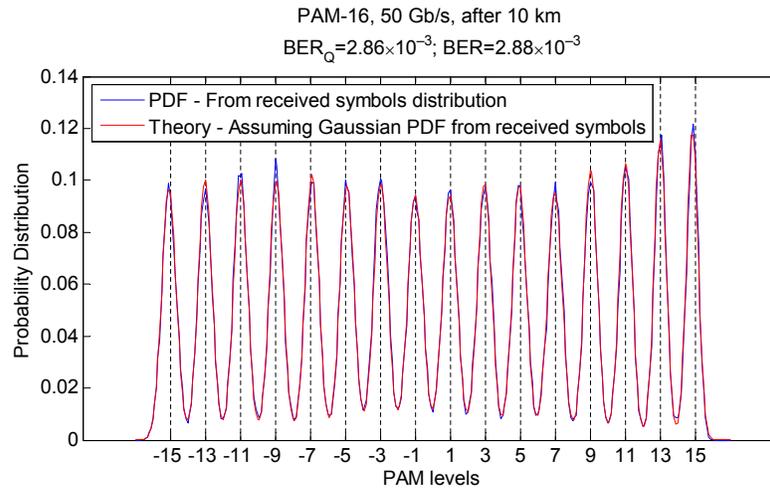


FIGURE 4–14: Probability distribution function of PAM-16 at 50 Gb/s, after 10 km. Blue: received symbol distribution. Red: PDF computed from of mean μ_i and variance σ_i^2 of each 16 level.

The match between both performance metrics, BER , and BER_Q , is clear, demonstrating the strong accuracy of the BER_Q model of Eq. (4.7) compared to the BER from error counting for multilevel PAM- N formats. Faithful matching of the histogram curves of PAM-8 and PAM-16 signals with the accumulated PDF curve further validates the model. The accuracy is high for received signals with low Q -factors (large BER) as shown in the previous Figs. 4–13 and 4–14

where BERs were greater than 10^{-3} . Low Q 's are observed for two cases: 1) for low received signal power (after 20 km of propagation) where the receiver is Gaussian thermal noise limited and 2) for high bit rates, where the transmitter's symbol rate is large, requiring heavy DAC compensation and consequently generating a waveform of smaller swing, of lesser signal-to-noise ratio, and moreover where the receiver integrates more inband noise from a larger symbol rate.

The match between the BER and the BER_Q (Q -factor) diverges for two cases: for received signals of high Q (low BER), and for decreasing PAM orders. The former is explained by the following. High Q signals at the receiver are attainable when both a high Q signal is generated at the transmitter and when sufficient power is presented to the receiver. For signals of high enough power, the highest PAM level of a multilevel PAM- N format typically has a noise distribution that does not exactly match a Gaussian distribution. The cause is twofold. First, the two edge levels generated at the transmitter can be readily compressed by the raised cosine power transfer function of the Mach-Zehnder intensity modulator, while the central levels remain in the linear transfer region. Second, and most importantly, the PIN+TIA receiver can start saturating when receiving the highest PAM level of a higher power signal. Those cumulative effects at the transmitter and receiver will modify the additive white Gaussian noise (AWGN) assumption employed in the Q -factor computation. This effect builds up as PAM order decreases, from 16 to 8, 4 and 2, as the ratio of non-AWGN-only levels over AWGN levels increases, explaining the divergence of the two metrics as the PAM order decreases. The non-AWGN-only distribution of the highest level is naturally more clearly observed for signals of high quality, when the non-AWGN portion of the distribution is not buried under other AWGN noise sources, like thermal noise at low received signal power. To demonstrate this, we show in Fig. 4-15 the histogram of received symbols and the equivalent aggregate Gaussian PDF for (a) PAM-2 and (b) PAM-4 signals of high Q . In Fig. 4-15(a) we see how both inner tails of the histogram spread out slightly more than what is predicted by a Gaussian-only distribution. The excess spreading is moderately larger for the higher level. In Fig. 4-15(b), for high Q PAM-4, we clearly observe the non-AWGN-only distribution of the higher level (level from alphabet label '3') and its larger histogram inner tail, significantly increasing the BER from error counting from the predicted BER_Q using a AWGN-only model. At such low BERs, a slightly larger tail does significantly impact the BER from error counting.

Figure 4-16 shows the system performance as we vary the bit rate from 25 Gb/s to 120 Gb/s for different PAM orders of 2, 4, 8 and 16, over different propagation distances of 0, 2, 10 and 20

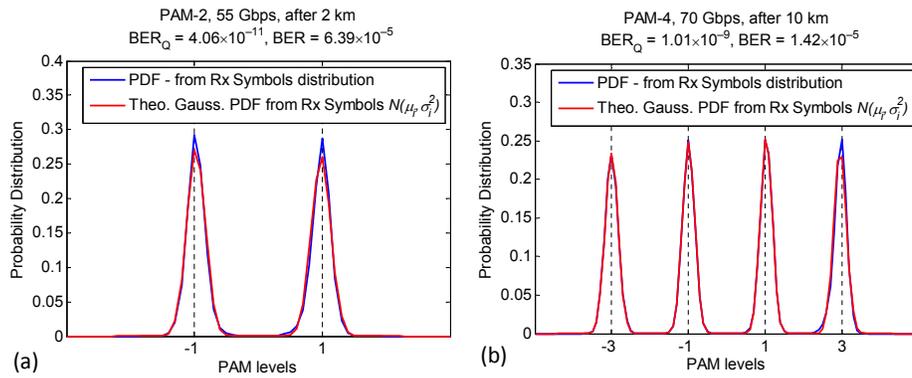


FIGURE 4–15: Histogram of received symbols and aggregate PDFs for signals of high Q (low BER), for (a) PAM-2 and (b) PAM-4.

km. For each bit rate of each PAM order, the roll-off factor of the pulse shaping filter is optimized to give the best system performance. We first present the transmission performance results using the Q -factor and the BER metrics in Fig. 4–16(b) and (c), respectively. Those results exhibit important characteristics for PAM order selection. First, for a fixed PAM order, it is observed that increasing the bit rate degrades the BER and the Q -factor. This is a normal behavior as a bit rate increase translates into increasing the signaling rate (symbol rate), itself increasing the bandwidth of the signal, where the larger matched filter at the receiver will integrate more in-band noise. Integrating more in-band noise power while maintaining the average signal power degrades both the BER and Q -factor. A more revealing observation drawn from both Fig. 4–16(b) and (c) is the following: looking at a fixed distance, we observe that as the bit rate increases, the highest Q and the lowest BER are attained using increasingly higher PAM orders, where transitions happen at specific bit rates. From the Q curves in Fig. 4–16(b), bit rates below 53 Gb/s have the highest Q using PAM-2, bit rates between 53 Gb/s and 109 Gb/s have the highest Q using PAM-4, and bit rates above 109 Gb/s exhibit higher Q using PAM-8. The same trend is observed using the BER metric as shown in Fig. 4–16(c), at different transitional bit rates. For instance, the transitional bit rate at which PAM-8 gives a smaller BER than PAM-4 as the bit rate is increased is 104 Gb/s instead of 109 Gb/s when the Q -factor metric is employed, showing a good match of the transitional bitrate using two very different performance metrics. At lower bit rates, the transition where PAM-2 gives a smaller BER than PAM-4 as the bit rate is reduced occurs at a bit rate that exhibits a BER that cannot be accurately measured from the bit error counting method due to the limited memory of the receiver. This PAM-2 to -4 transitional bit rate would also be smaller using the

BER technique than the one obtained using the *Q*-factor metric. Notwithstanding this off-line measurement limitation, the *BER* slopes of PAM-2 and PAM-4 in Fig. 4–16(c) clearly indicate that as the bit rate is lowered a transition point will be reached at which PAM-2 will outperform PAM-4. This greater deviation of the transitional bit rate between PAM-2 and -4 for the two metrics can be explained from Fig. 4–16(d), where we investigate the relation between each *y*-axis variable in Fig. 4–16(b) and (c), i.e. where we plot the error counting *BER* as a function of the *Q*-factor.

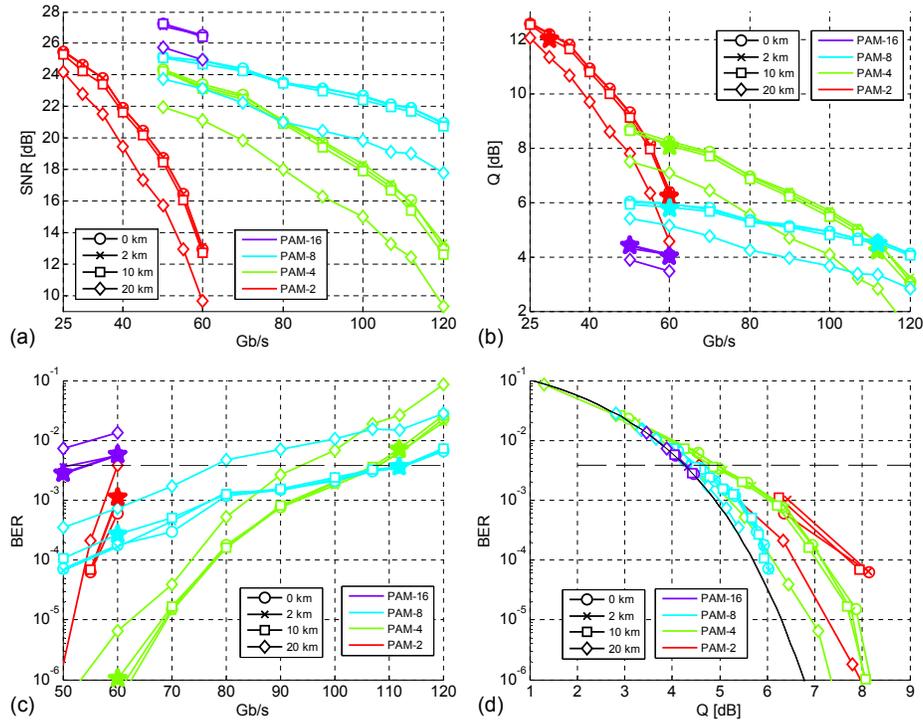


FIGURE 4–16: (a) SNR, (b) *Q*-factor [dB] and (c) *BER* for PAM orders 2, 4, 8 and 16, after propagation distances of 0, 2, 10 and 20 km, for varying bit rates. Figure (d) shows the *BER* from error counting against $Q(BER_Q)$. Black solid curve in (d) represents the theoretical $BER_Q(Q)$ relation of Eq. (4.5). Hashed line in (c) and (d) is $BER=3.8 \times 10^{-3}$ threshold

The waterfall curves in Fig. 4–16(d) corroborate the histograms, PDFs and remarks drawn from Figs. 4–13 to 4–15. Figure 4–16(d) shows that the two metrics of *Q*-factor and *BER* match very well for signals of low *Q* (high noise floor) and signals of large PAM order, and start diverging as the signal quality is increased and as PAM order is decreased. As the transitional bit rate between PAM-2 and -4 occurs at both high *Q* and low PAM order, the two metrics exhibit a different transitional value.

For completion, SNR versus bit rate curves are also added in Fig. 4-16(a) as a third performance metric in order to compare it with the two main metrics of *BER* and *Q*-factor. An important observation is that for any desired bit rate, the SNR is always greater at larger PAM orders, when monitored after a fixed distance. This is misleading and contradicts the observations of the *BER* and *Q*-factor metrics. It shows the inaptitude of the SNR metric for multilevel PAM signaling to provide meaningful and representative qualitative figure of merit, like the *Q*-factor delivers. This being said, we can explain this SNR behavior by the lower required symbol rate at any desired bit rate as the PAM order increases. Smaller symbol rate translates to a further accurate waveform generation with more samples per symbols, and a matched filter of smaller bandwidth. A second observation from the SNR curves shows that as the bit rate increases, the SNR of PAM-4 decreases faster than that of PAM-8. The rationale is closely related to the previous observation. For an increase in bit rate by 30 Gb/s, PAM-4 requires an increase in signaling rate by 15 Gsymbols/s, where PAM-8 only requires an increase by 10 Gsymbols/s. Rapid increase of signaling rate translates into rapid decrease of the signal to noise ratio of the transmitter generated waveform. This trend rapidly accumulates as the bit rate is increased from 50 Gb/s to 120 Gb/s.

In light of our observations in Fig. 4-16, it is interesting to look back at the eye diagrams of Fig. 4-12. First, the eyes of Fig. 4-12 are represented by bold stars in Fig. 4-16(b) and (c). The optimum PAM order selection as a function of bit rate observed in the performance curves of Fig. 4-16 is in fact visually well depicted in the eyes of Fig. 4-12. For low bit rates in the region of 30 Gb/s, it is clear that the PAM-2 format offers better performance than PAM-4: the eye is wide open, levels are largely separated and the raw *BER* is potentially already below 10^{-15} . However, for a target *BER* in the region of 60 Gb/s, Fig. 4-12(b), (c), and (e) show that PAM-4 provides the cleanest signal at the receiver, with the most well defined, distinctly separated levels. Indeed, both the error-counting *BER* and the *Q*-factor are better when using PAM-4 at 60 Gb/s compared to when using PAM-2 or PAM-8. On the other side, if the desired bit rate is in the 112 Gb/s region, Fig. 4-12(d) and (f) show that the *BER* obtained using PAM-4 can be halved by reducing the signaling rate by a third of its value and using 8 levels instead of 4. The *Q*-factor also improves at 112 Gb/s with PAM-8. Trends in Fig. 4-16(c) show that the *BER* improvement using PAM-8 over PAM-4 further increases with increasing bit rates starting at around 104 Gb/s.

In the following Fig. 4-17 we present the system performance as we vary the signal power presented to the PIN+TIA receiver, for different PAM orders. To collect this data the experimental

test bed of Fig. 4–10 is slightly modified to add power measurements: the system is in back-to-back (0 km) and the receiver is preceded by a variable optical attenuator (VOA), followed by 99/1 coupler with characterized port ratio and a power meter on the 1% port. For PAM-4 and PAM-8, the bit rate is set to 107 Gb/s and for PAM-2, 60 Gb/s. Both the *BER* and *Q*-factor metrics are presented to monitor the performance, respectively in Fig. 4–17(a) and (b). The figure shows that the optimum received signal power is dependent on the PAM order. For PAM-8, the received signal power yielding the lowest *BER* is around -4.3 dBm. For PAM-4 and PAM-2 the maximum received power was lower than the optimum operational point, allowing room for improved performance. The coupling loss of the grating couplers of the integrated SiP modulator provides the bulk fiber to fiber loss of the modulator. Better light coupling into the optical chip would significantly improve the available power to present to the PIN+TIA receiver and allow operation at optimum Rx power for both lower PAM orders. Nonetheless, looking at the trends around the highest received signal power of approximately -3 dBm for PAM-4 and PAM-2, we observe that PAM-4 is optimum just above -3 dBm, and PAM-2 requires even higher power before it reaches optimum operation. This optimum power reduction with increasing PAM order shows the receiver linearity requirement as we increase the number of signaling levels. A received optical power that is too large will compress the higher levels of a PAM-*N* format: a compression that is more detrimental as *N* increases.

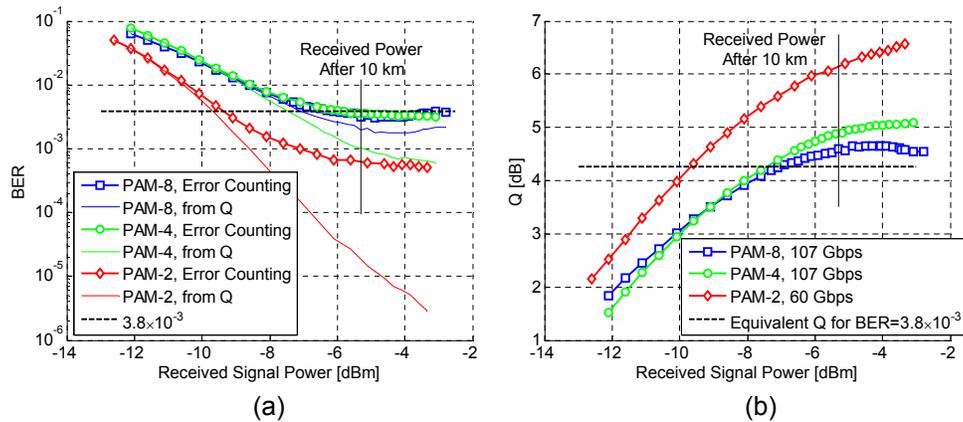


FIGURE 4–17: *BER* and (b) *Q*-factor for PAM-2, -4 and -8, for varying received signal power. For both (a) and (b), PAM-4 and -8 are running at 107 Gb/s and PAM-2 at 60 Gb/s.

Figure 4–17(a) shows both the *BER* obtained from error counting, as well as the equivalent BER_Q from the computed *Q*-factor of Fig. 4–17(b). For low signal power, where the receiver

operates fully linearly and is limited by Gaussian thermal noise, both BERs (BER and BER_Q) match perfectly, meaning that in this regime, the AWGN assumption of Eq. (4.7) is met. From the initial slope of Fig. 4–17(b), an increase of 1 dB in received power is linearly translated to a Q increase of 0.76 dB. As the power is further increased, solely by changing the VOA attenuation, the receiver starts behaving non-linearly and the dB-for-dB relation vanishes. The higher PAM levels start compressing down and the Gaussian noise statistics are modified. The Q -factor from the computed mean $\hat{\mu}_i$ and variance $\hat{\sigma}_i^2$ of all levels give an equivalent BER_Q that is much smaller than the real BER from error-counting. This is because the AWGN assumption on all PAM levels does not hold well at increasing powers, where the highest levels are compressed by the PIN+TIA and exhibit non-AWGN noise distribution, as mentioned in the discussion of Fig. 4–15. Computing the first two statistical moments of all levels and applying the Gaussian model to each level fails to provide an accurate BER estimate. A sum of two different noise processes, with one being power dependent, would better model the noise statistics at the receiver and provide a better BER estimation. Figure 4–17(a) and (b) show a vertical line at -5.3 dBm, representing the received optical signal after 10 km of SMF-28e+.

In the following Fig. 4–18, we present the signal quality, by showing the BER and the Q -factor, for PAM orders of 2, 4 and 8 at different symbol rates, against the roll-off factor α of the root raised cosine pulse shaping filter applied at the transmitter and receiver. Results are all captured after 10 km of propagation. BER 's log-scale is deliberately limited to 10^{-6} from the limited resolution of BER from error counting. For all PAM orders, we observe that both Q and BER have an optimum roll-off factor that decreases with increasing symbol rates. It is important to reiterate here that the DAC's sampling rate is always fixed at 70 GSamples/s. This first limits the range of roll-off factors as a function of the symbol rate (Section 4.2.4.1) and second, changes the number of SPS as the symbol rate varies.

Smaller SPS translates to closer spectral images and less accurate analog waveform generation, to the extent that at 1 SPS, the DAC acts as a multilevel PPG with no pulse shaping capabilities. Moreover, at an equal roll-off factor, increasing the symbol rate by 5 Gbaud means increasing the maximum spectral content of the signal that needs to be generated by the DAC by $5(1+\alpha)/2$ GHz. From this, the experimental decrease of α becomes intuitive. Three co-acting effects worsen the quality of the generated signal with larger spectral content. The first effect is the reduced output swing. The analog response of all the RF components at the transmitter side rapidly decreases in

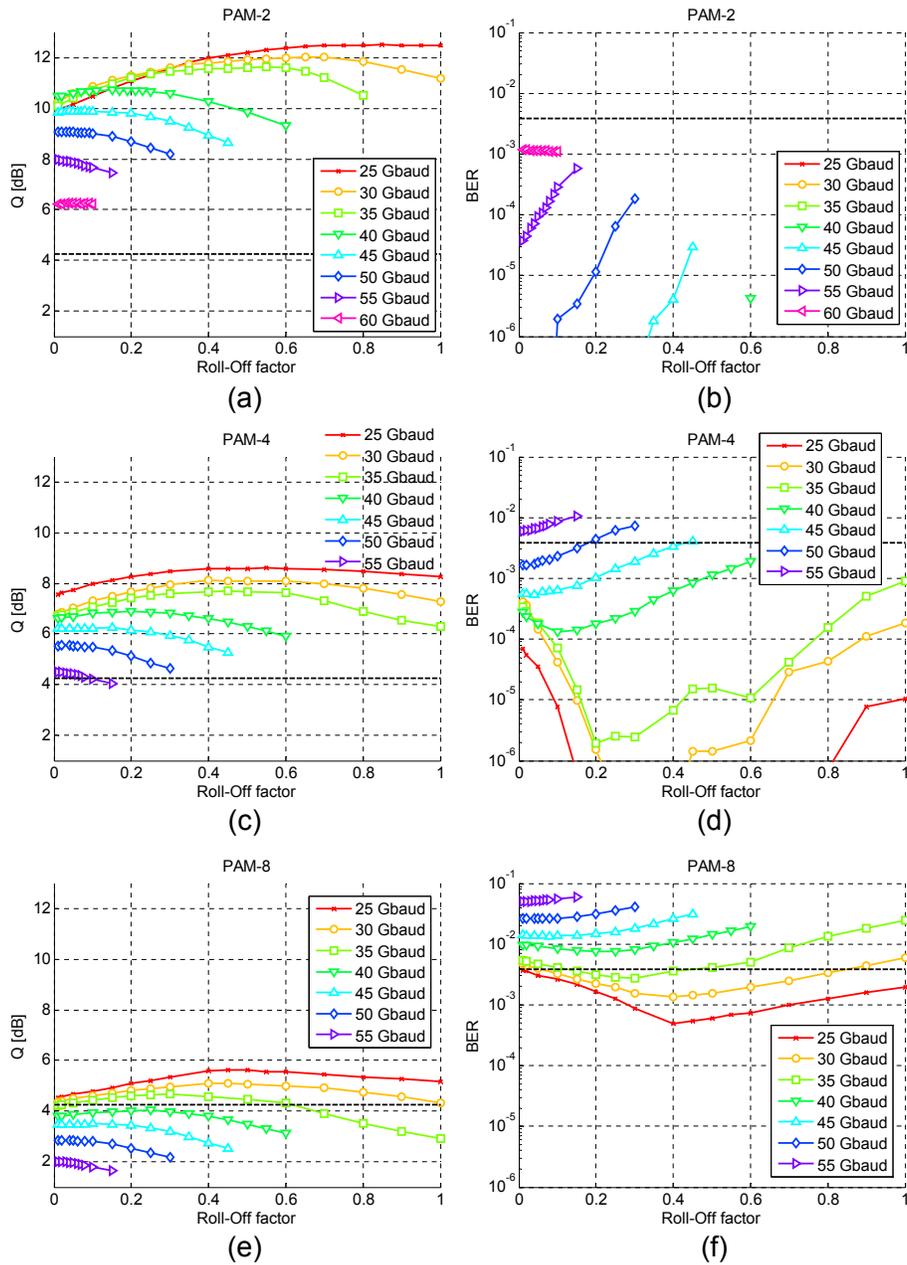


FIGURE 4–18: System performance assessment for varying PAM orders, at different symbol rates, for varying roll-off factors. (a) Q for PAM-2, (b) BER for PAM-2, (c) Q for PAM-4, (d) BER for PAM-4, (e) Q for PAM-8, (f) BER for PAM-8. Hashed black line is BER of 3.8×10^{-3} or the Q -factor equivalent.

amplitude with increasing frequency. The response can still be equalized after the addition of the extra bandwidth to generate the desired pulse shape, but this is done at the expense of reduced swing out of the DAC. As all the PAM- N levels have to fit in this reduced electrical swing, the

electrical signal to noise ratio of the generated RF waveform is inevitably reduced. This will get translated to the optical waveform generation, and observed at the receiver. The second effect worsening the quality of the generated signal after a symbol rate increase is the effective number of bits (ENOB) of the DAC, which decreases with increasing frequency. Consequently, the signal quality within the additional bandwidth, required for the increased symbol rate, will inherently be of worse quality. Thirdly, increasing the symbol rate widens the full width at half max (FWHM) of the receiver matched filter, which translates into integrating more noise power with the signal as the matched filter is applied. Consequently, even if equal signal power was assumed after a symbol rate increment, the SNR at the receiver would intrinsically worsen. The sum of those three effects explains 1) the worse performance at a fix roll-off factor and PAM order for increasing symbol rates, 2) the reduced optimum roll-off factor at a fixed PAM order for increasing symbol rates, and 3) the reduced performance for increasing PAM order at a fixed symbol rate and fixed roll-off factor.

Finally, for low symbol rates of 25, 30 and 35 Gbaud where 2 SPS or more are employed to accurately generate the desired waveform, one can observe that the optimal roll-off factor decreases for increasing PAM order. It is noteworthy to reiterate that all data of different bit rates and PAM orders presented in Figs. 4–16 and 4–17 employ the optimum roll-off factor presented in Fig. 4–18.

The following Fig. 4–19 shows (a) the *BER* and (b) the *Q*-factor for PAM-4 and PAM-8 when running at 112 Gb/s, over varying distances of 0, 2, 10 and 20 km. We can observe that PAM-8 is able to yield a BER that is below 3.8×10^{-3} after 10 km. At a greater distance of 20 km, the received signal power is too low and the BER increases with decreasing signal power. The received power after 20 km was -9 dBm. For 0 and 2 km, the received signal power had to be slightly attenuated in order to operate at the optimum powers, as shown in Fig. 4–17. PAM-4, however, could not deliver the 112 Gb/s throughput while providing a BER below the 3.8×10^{-3} FEC threshold at any distance. This format also shoots to smaller *Q*'s and higher BER's as the received signal power decreases from 10 km to 20 km of SMF-28e+. Figure 4–19 presents both metric performance, however, only the BER from error counting defines whether or not a modulation format performs below the BER FEC threshold after a specific distance, consequently confirming or refuting its error free operation.

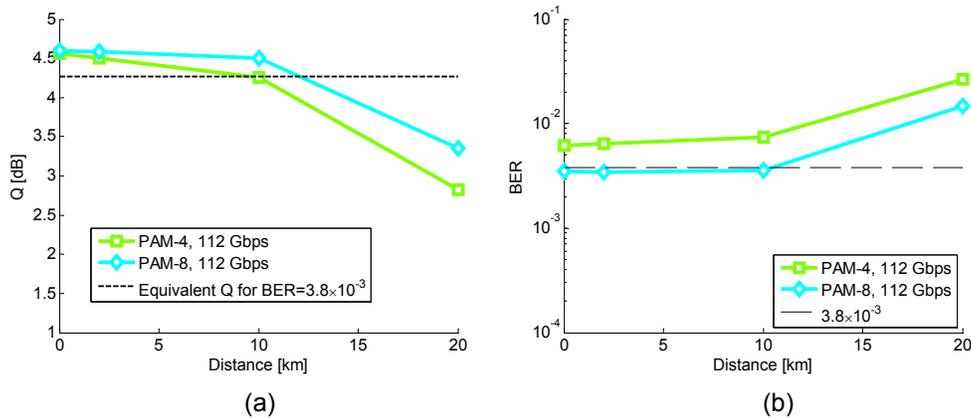


FIGURE 4–19: (a) Q -factor and (b) BER for PAM-4 and PAM-8, at 112 Gb/s, at varying distances.

4.2.6 Conclusion

The performance of a SiP intensity modulator at 1310 nm for short reach transmission using PAM modulation format was experimentally studied. Transmission of 112 Gb/s over 10 km of SMF fiber below the pre-FEC BER of 3.8×10^{-3} using PAM-8 at 37.4 Gsymbols/s was experimentally validated. A large parameter space was swept for both the transmitter and receiver, namely the PAM order, symbol rate and bit rate, pulse shaping's roll-off factor, propagation distance and received signal power. We showed that to optimize the system performance and minimize the BER, there exists an optimum PAM order at a given desired bit rate, with PAM of an increasingly higher order being preferable as the bit rate is increased. We also presented a new Q -factor model for multilevel PAM signaling using Gaussian statistics. We experimentally demonstrated the accuracy of the model by first comparing its equivalent BER_Q to the BER from error counting, and second by comparing the aggregate probability density function of the model with the histogram of the received data. The two approaches show an accurate match for Gaussian noise limited transmission, validating the Q -factor model.

CHAPTER 5

Optical Short Reach Transceivers: From Unidimensional to 3D Modulation

After long-haul optical coherent and chip-scale photonics addressed earlier in this thesis, the last level in the optical network topology to address is that of short reach optical transceivers. Such optical transceivers target distances below 40 km, typically between 500 m and 10 km. They are intended for connectivity within and between data centers and are commonly called ‘Data center interconnects’. We have presented in Section 4.2 in the previous chapter a modulator that is intended to be employed for data center applications, capable of transmissions of up to 10 km. The novelty about this modulator was that it is entirely made in silicon and is capable of modulating 112 Gb/s via the PAM-8 format, which was a first when publicly reported. The previous section focused on detailing the specifications of this SiP modulator and presented a plethora of parameters to be optimized when dealing with DSP+DAC based IM/DD transmission applied to a SiP modulator. In this chapter, we focus on data centers transceivers per se. In the first section, we study the same device but this time in the context of data center applications instead of as the first 112 Gb/s SiP modulator to be reported. We focus on the impact of the tap length of the pre-emphasis and equalizer filters for DSP enabled, DAC and ADC-based transceivers and present a digitally improved version of the SiP MZ modulator presented in Section 4.2. In the second part of this chapter, we present another DSP enabled data center interconnect working on a completely different principal of operation that allows almost tripling the transmission speed of the previous transceiver by modulating 3 degrees of freedom of a single wavelength while maintaining direct detection.

5.1 Biasing, Pre-Emphasis and Post-Equalization of DAC/ADC Enabled SiP MZM

In Section 4.2, we demonstrated 112 Gb/s over 8-level pulse amplitude modulation at 37.4 Gsymbols/s using a silicon photonics Mach-Zehnder Modulator. This modulation format and signaling rate are not compliant to standardization bodies for next generation single wavelength

IM/DD transceiver. The physical layer specification that was recently ratified in IEEE 802.3bsTM as 400GBASE-DR4 specifies PAM-4 as the preferred modulation format and 56 Gsymbols/s as the electronic signaling rate. In this section, we demonstrate how better signal equalization at the transmitter (Tx) applied to the same hardware can greatly improve the final bit error rate of the DSP enabled IM/DD transceiver. We demonstrate again 112 Gb/s, this time at the more aggressive and IEEE 802.3bsTM compliant symbol rate of 56 Gbaud using the same SiP modulator, which has a 3-dB bandwidth of only ~20 GHz.

5.1.1 Introduction

Silicon photonics is a rapidly evolving technology that is fabricated using CMOS manufacturing processes and enables device and systems designers to build dense photonic circuits on silicon. Fabricating external Mach-Zehnder modulators on a silicon platform allows design of highly complex and integrated systems.

Higher order intensity modulation with direct detection on a single wavelength is an attractive approach to efficiently increase the bit rate throughput without the expensive capital expenditure required by multi-wavelength multiplexing. PAM-4 was recently chosen by IEEE's 400Gb/s Physical Layer Taskforce as the desired modulation format for the next physical layer specifications of 400GBASE-DR4, -FR8, and -LR8 as the intensity modulation format for optical short reach transceivers.

The generation of PAM- N formats can be done via passive RF combination of $\log_2(N)$ single binary On-Off streams, or equivalently from a $\log_2(N)$ -bit DAC. However, the quality of the optical intensity signal generated at the transmitter can be greatly leveraged if a multi-bit DAC ($>\log_2(N)$) is employed. Moreover, the performance and throughput of a direct detection receiver (Rx) can also be leveraged by employing an analog-to-digital converter after the PIN+TIA receiver followed with filtering from a feed forward equalization (FFE) filter [145]. The equalizer can initially be adaptive until tap convergence and switch to a static finite impulse response (FIR) filter in steady state operation.

We study here the transmission performance of PAM-4 and PAM-8 formats running at 112 Gb/s using an 8-bit DAC and show how DACs can greatly benefit PAM formats by employing more output levels to generate the PAM waveform. With an 8-bit ADC at the Rx, we show that transmission performance can be greatly improved with feed-forward equalization. We study the

impact of varying the number of Tx pre-compensation and Rx post-compensation taps on bit error rate performance. When implemented in an application specific integrated circuit (ASIC), the number of Tx and Rx taps has a direct impact on chip power consumption. The Mach-Zehnder silicon photonic modulator employed in this study is the same as that presented in Section 4.2, i.e. a series push-pull device constructed from two series connected pn junctions having a common n-type region. The device therefore requires a DC bias voltage applied at the common cathode in order to reverse bias the diodes. The bias voltage has an important effect on the BER depending on the order of the PAM format and on the strength of the Tx pre-compensation filter.

In this section we further the description of the SiP modulator in Section 5.1.2 and explain how the common DC bias voltage impacts the transmitted waveform. We examine the impact of some Tx parameters in Section 5.1.3, where BER is used as the performance metric. Section 5.1.4 presents the system test bed and experimental results are found in Section 5.1.5. We conclude in Section 5.1.6.

5.1.2 Further Details on the SiP Modulator

The SiP modulator is the same as that presented in Section 4.2.2. We are adding here the electro-optic frequency response of the modulator for biases of 0, 2, 4 and 6 V in Fig. 5–1, where all curves are normalized at 1.5 GHz.

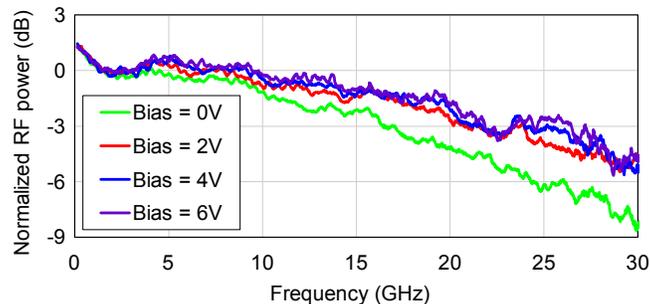


FIGURE 5–1: EO frequency response of the modulator, for $V_{\text{bias}} = 0, 2, 4,$ and 6 V.

In Fig. 4–7(c) we saw that the bias voltage has an important effect on the efficiency of the modulator. At a small bias voltage, the region of depleted charges, located at the center of the pn junction, is narrow and coincides with the region where the optical mode field intensity is high. A change of the voltage will thus have a large impact on the phase of light, hence the smaller V_{π} .

In the opposite direction, when a large DC bias voltage is applied, the depletion region widens and the RF signal modulating its width around that point will have a lower impact on the phase of light because the charge distribution variation then occurs where the mode field possesses less energy, hence the greater V_π . However, operating the pn junctions around a low bias voltage increases the optical losses, as seen in Fig. 4-7(e), and reduces the modulator speed from the larger junction capacitance, as seen in Fig. 5-1.

When the bias is very low, the AC swing could even send the pn junction into a carrier injection mode during part of the modulation cycle (although the forward conduction mode appears at measured forward voltages of around 0.7–1 V). In carrier injection mode, the speed is reduced dramatically, this time from charge accumulation and removal through recombination during the RF modulation cycle. The antisymmetry of the index/phase change produced in the two interferometer arms around the nominal mean voltage can be broken, reducing the efficiency of the push-pull action.

There are thus tradeoffs in the optimization of the bias point V_{bias} : a small bias voltage leads to increased insertion loss, reduced speed, deficient push-pull action, but increased efficiency (i.e., reduced V_π). A large bias voltage reduces insertion loss and increases speed but also V_π . The bias point at which optimal performance is obtained also depends on the strength of the applied RF modulating signal. Tx spectral compensation using the DAC lowers the low frequency content with respect to the higher frequency content of the modulating signal spectrum, leading to a reduced RMS voltage applied to the modulator. This opens the way to reducing the bias voltage while maintaining depletion mode operation of the pn junctions at all times. There is thus an interest in studying the impact of this parameter on system performance, as an optimized bias point yields additional system performance margin.

5.1.3 Impact of Tx Parameters

For PAM formats generated from a DAC, several parameters can be varied and modified to improve the system performance. We define system performance via the recovered BER at the receiver. The following is a non-exhaustive list of Tx parameters having an impact on performance for PAM formats: the DACs nominal and effective number of bits, the amount of DAC clipping, the strength of the Tx pre-compensation filter, the pulse shaping filter's roll-off factor, the strength of the nonlinear power transfer function compensation of the MZM, and the bias voltage. All previously mentioned parameters, except the bias voltage V_{bias} , are controlled digitally in the Tx

digital signal processing stack [4]. Each parameter modifies in its own way the signal quality generated at the Tx and the performance monitored at the Rx.

The DAC nominal number of bits N_b sets the number of quantization levels (from -2^{N_b-1} to $2^{N_b-1} - 1$). The signal to quantization noise ratio (in decibel) equals $\text{SNR}_Q = 6.02N_b + X$, where X depends on the mean signal power before full-scale quantization, for a uniform quantization error. A pure sinusoidal signal has $X = 1.76$, for instance. The ENOB can be viewed as the base-2 logarithm of the effective number of resolvable levels as a function of frequency [145].

The Tx symbol rate R_B can be any value less than or equal to the DAC's sampling rate F_s . The symbols to transmit are upsampled and pulse shaped at F_s/R_B samples per symbol. Typical pulse shaping filters are root-raised cosine filters, where a roll-off factor, denoted by α , defines how steep the pulse shaping filter's amplitude drops in frequency around $f = 1/2R_B$. The roll-off factor ranges between $0 \leq \alpha \leq \alpha_{\max}$, where $\alpha_{\max} = F_s/R_B - 1$. A rapid spectral descent, due to a small α , results in a temporal pulse that lasts longer. For a fixed DAC sampling rate, higher symbol rates are constrained to a smaller range of possible α , and consequently of longer pulse shaping impulse responses [4].

Independent symbols are successively pulse shaped and transmitted. The waveform value at a given time instant is a sum of multiple shifted and weighted pulse shapes, where weights are from the upsampled F_s/R_B SPS symbol input. As the roll-off factor decreases, more neighboring symbols interfere with the waveform value at one time instant and the sum includes more elements. As the interference can be randomly destructive or constructive, smaller roll-off factors give rise to greater peaks, increasing the peak to average power ratio (PAPR) of the waveform.

Quantizing the full range of a peaky signal can be detrimental to the output signal quality, as the signal's standard deviation can greatly differ from its peak to mean value when uploaded as full scale on the uniform quantizer, reducing SNR_Q . Clipping of the waveform before quantization is an irreversible operation that can improve the resulting SNR. However, aggressive clipping can greatly distort the signal and become detrimental to the output signal quality.

A pre-compensation filter $h_{\text{inv}}[n]$ can be applied before DAC quantization. This FIR filter can partially compensate for the analog frequency response of the successive components following the DAC, including the DAC itself. Partial compensation of the Tx RF elements can be very favorable to the signal quality. However, compensation inherently increases the signal's PAPR, and aggressive compensation becomes detrimental. The strength and the accuracy of this compensation can be

varied by varying the number of taps of the pre-compensation FIR filter, $h_{inv}[n]$. Tap values of h_{inv} are found experimentally, stored in memory, and employed as a static Tx DSP FIR block.

In addition, MZM have a nonlinear, sinusoidal-like power transfer function with respect to the applied voltage. For IM formats like PAM, the Rx expects equally spaced levels at proper sampling instances. Large RF swings on the MZM will distort the symbols' amplitude spacing. Nonlinear compensation can be readily done by applying an inverse sine function to a prescaled waveform. The scaling factor g in the interval $0 \leq g \leq 1$ determines the nonlinear compensation strength. Larger scaling factors will increase the waveform's PAPR. Proper nonlinear compensation will be canceled by the MZM itself. Figure 5-2 presents the DSP stack applied at the Tx, repeated from Fig. 4-11(a) for completeness.

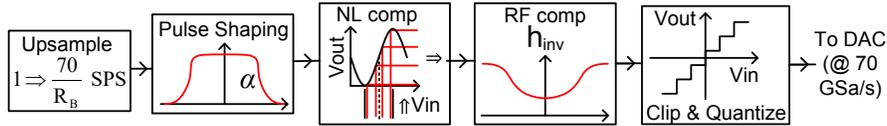


FIGURE 5-2: Tx DSP stack for DAC-enabled IM-DD transmission.

Finally, the modulator bias voltage V_{bias} , the sole parameter not related to digital processing, also has an impact on system performance. Figure 4-7(c), Fig. 4-7(e), and Fig. 5-1 inform us on the interplay between the modulator's V_{π} , its insertion loss, and its electro-optic response as a function of DC bias.

We experimentally investigate two of these Tx parameters, namely, the performance variation as a function of the applied bias voltage V_{bias} , and the transmission performance for varying tap numbers of the pre- and postcompensation filters, h_{inv} and h_{Rx} , respectively.

5.1.4 System Description

The following Fig. 5-3 presents the experimental test bed deployed to assess the BER variations as V_{bias} is varied from 4 to 2 V and as the Tx and Rx number of taps are changed. The setup is moderately modified from that employed in the tests of Section 4.2 and depicted in Fig. 4-10.

A 30 mW 1310 nm DFB laser with a measured linewidth of 440 kHz is modulated by the SiP intensity modulator. The RF signal driving the modulator is first generated from a balanced AC-coupled 8-bit DAC running at 70 Gsamples/s before being amplified by an Inphi IN3214SZ linear driver providing 17 dB of differential to single ended gain. The SiP modulator is constantly

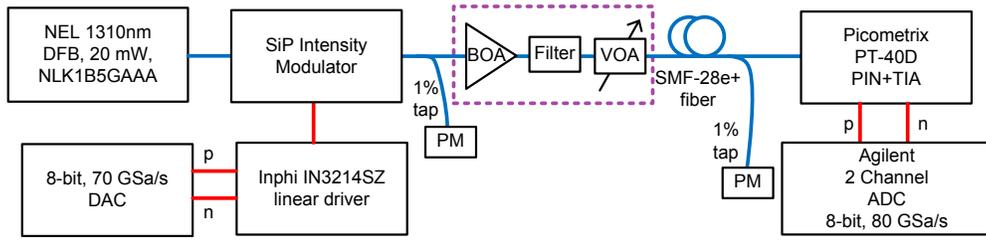


FIGURE 5-3: Experimental Setup for DAC-enabled IM-DD transmission.

operated at quadrature by appropriately tuning the thermal phase shifters (see Figs. 4-5 and 4-6). For experiments where needed, a booster optical amplifier (BOA) with an optical filter and a variable optical attenuator are added. Otherwise, the intensity modulated light from the MZM is launched into various lengths of single mode fiber: either 0, 2, 10 or 20 km of Corning SMF-28e+. A balanced AC-coupled PIN+TIA Rx is employed before each balanced RF ports are real time sampled at 80 Gsamples/s by an Agilent DSO-X 96204Q oscilloscope serving as an 8-bit ADC. The DAC has a 3 dB bandwidth of 15 GHz and a maximum output differential swing of 1.2 Vppd. The linear driver has maximal single-ended output voltage swing of 6 Vpp and a 3 dB bandwidth of 34 GHz. The PIN+TIA Rx has a bandwidth of 35 GHz and a conversion gain of 1.3 V/mW, and the real-time scope's bandwidth is 33 GHz. Finally, the SiP modulator exhibits a 3 dB bandwidth slightly above 20 GHz.

The two-channel real-time oscilloscope digitizes and stores the Rx waveforms for offline processing. Figure 5-4 shows the Rx DSP stack, also repeated from Fig. 4-11(b) for completeness. The two ports of the balanced receiver are first subtracted digitally to remove common noise terms. The DSP consists of resampling at 2 SPS, matched filtering, and Rx equalization h_{Rx} that is embedded with symbol decision and clock recovery [145].

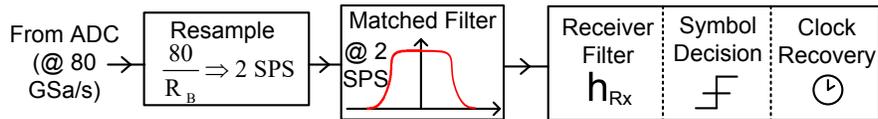


FIGURE 5-4: Rx DSP stack for DAC-enabled IM-DD transmission.

The Rx compensation filter h_{Rx} compensates for the Rx front end's frequency response and helps mitigate residual ISI and the remaining spectral response that the Tx precompensation filter did not equalize. h_{Rx} can be an adaptive Rx equalizer on system startup until tap convergence. A

typical Rx DSP-based adaptive equalizer employs a FFE and a decision feedback equalizer (DFE) block. An FFE is characterized by the number of taps and their spacing, and typically equalizes the precursor portion of the residual ISI. A DFE is also characterized by its number of taps and typically cancels the post-cursor portion of the ISI [145].

The adaptive equalization scheme employed for the Rx DSP, before switching to a static FIR, is a delayed FFE where both pre- and post-cursor taps are used for a two sided equalization. FFE's tap spacing is T/k , $k \in \mathbb{N}^+ = \{1, 2, \dots\}$. This "quantization" of the FFE tap spacing is required due to the "filter and decimate" action of the Rx equalization block. Decimating a symbol every k samples is more complex when k is not an integer as it requires resampling within the equalization+decision block. We employ $T/2$ tap spacing for low complexity and accurate frequency representation.

Tap adaptation of the $T/2$ spaced FFE uses the least mean square update algorithm with the error being the difference between the current symbol after equalization, before and after decision.

5.1.5 Experimental Results

The impacts of V_{bias} and of Tx and Rx tap count are presented. First, we study the variation of the modulator bias voltage for 56 Gbaud PAM-4 and 37.4 Gbaud PAM-8 over propagated distances of 0, 2, 10, and 20 km. Second, BER degradation is investigated as the length of pre- and post-equalizing filters are varied, for both formats. The optimum pulse shaping roll-off factors used for 56 Gbaud PAM-4 and 37.4 Gbaud PAM-8 are 0.03 and 0.4, respectively, as found in [4]. The $V_{\text{peak}} - V_{\text{sto}} - V_{\text{peak}}$ swing driving the MZM is 2.85 V for PAM-4 and 3.20 V for PAM-8. For both formats and for V_{bias} ranging between 2 to 4 V, the optical power out of the SiP MZM always remains between -1.6 and -2 dBm.

5.1.5.1 Performance Impact of Modulator Bias Voltage

Figure 5-5 presents the BER for PAM-4 and PAM-8 at 112 Gb/s for increasing distances as the bias voltages is varied from 4 to 2 V. With a fixed DAC rate of 70 Gsamples/s, a signal of lower symbol rate is inherently of greater quality: it is generated with more SPS, and has a larger roll-off factor, providing shorter temporal pulse duration, better tolerance to jitter [146], and a larger RF swing.

However, higher SNR does not directly translate to lower BER when the number of signal level increases. In back-to-back, PAM-8 offers a SNR of 22.7 dB while PAM-4 provides a SNR of 18.4

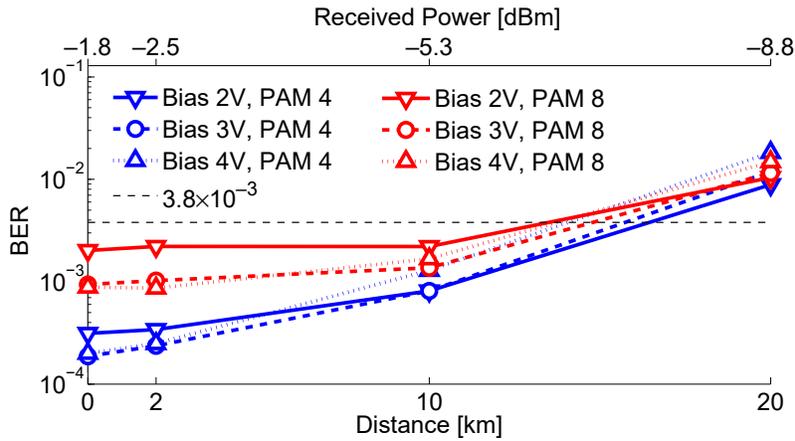


FIGURE 5–5: BER for PAM-4 and PAM-8 at 112 Gb/s, at varying distances, for bias voltages of 4, 3 and 2 V. 83 FIR taps are employed at the Tx, and 19 taps at the Rx.

dB. However, as seen in Fig. 5–5, PAM-4 still outperforms PAM-8 with a BER almost five times lower. We observe in B2B that a bias of 4 V performs better than at 2 V. The reason for this can be understood from Fig. 4–7(c). As V_{bias} decreases, the RF driving signal amplitude with respect to V_{π} increases. At a bias of 4 V, the RF amplitude swings mostly in the linear portion of the sine power transfer function. Consequently, little nonlinear compensation is required. This lower PAPR, high quality RF waveform is linearly converted from the electrical to the optical domain, and back to the electrical domain to produce a high quality received signal of large AC swing from the high power received after 0 km.

As V_{bias} is reduced to 2 V, the RF signal amplitude significantly swings in the compressed portions of the power transfer function, requiring compensation. This compensation is slightly detrimental in the B2B configuration because the Rx does not lack power and the TIA already receives a large AC swing. As the propagation distance increases, the received signal power decreases and both the mean and variance of the optical power decrease. In this case, increasing the AC swing of the optical waveform by swinging wider over the V_{π} range becomes beneficial as it gets translated directed to the AC-coupled TIA, increasing the peak to peak voltage swing on the ADC. The small mean signal power drop from increased on-chip insertion loss that occurs when lowering V_{bias} from 4 to 2 V is negligible and even compensated by the increased optical signal power variance. This relationship between optimum V_{bias} as a function of distance is observed for both PAM-4 and PAM-8 formats.

Figure 5–5 shows that the performance of the 56 GBaud signal degrades more rapidly with distance (received power) than the 37.4 GBaud signal. As the received signal power is reduced, thermal noise becomes the dominant Rx noise source. As thermal noise is white Gaussian noise [147], a larger symbol rate inherently integrates more in-band noise power, explaining the faster deterioration of PAM-4 with distance.

Figure 5–6 shows the BER when we maintain the power presented to the Rx, after different distances. To fix the received power, the components inside the dashed box in Fig. 5–3 are inserted in the test bed. A BOA at 1310 nm amplifies the signal out of the MZM to more than 10 dBm. To reduce the out-of-band ASE noise added by amplification that would show after direct detection, a tuneable optical filter is inserted. A VOA sets the launch power such that the power at the Rx after propagation is always the optimum power of the format being transmitted. The optimum power for PAM-4 is -1.5 dBm and for PAM-8, -2.9 dBm. Fiber spools of lengths 10 and 20 km are added for different transmission distances. The quasi-independence of the BER with increasing propagation distance shows that group velocity dispersion is negligible and has no tangible impact on the performance. Moreover, it also shows that polarization mode dispersion after 20 km for a 56 Gbaud signal ($T = 17.9$ ps) is insignificant and does not spread the pulse.

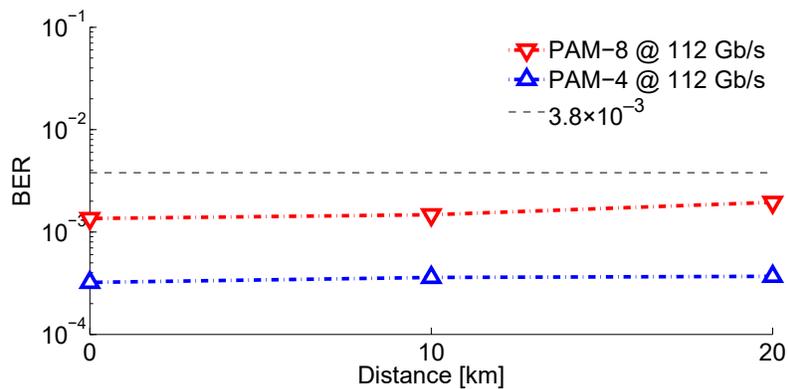


FIGURE 5–6: BER for a fixed received power after varying distances. PAM-4: -1.5 dBm; PAM-8: -2.9 dBm.

In light of Fig. 5–6, the performance of the system over varying distances presented in Fig. 5–5 is purely limited by the available power after propagation. The “distance” axis could be replaced by “received signal power” where the received signal power P_K after K km is about $P_K = P_0 - 0.35 \text{ dB} \times K$, where P_0 is the power out of the MZM. An additional top x -axis in Fig. 5–5 shows

the received power at each distance. Differing BERs in B2B for PAM-4 and -8 become similar after 20 km in Fig. 5-5 as the Rx becomes thermal noise limited. We observe that both PAM-4 and PAM-8 at 112 Gb/s perform below the FEC threshold of 3.8×10^{-3} after 10 km, and both exceed the FEC limit after 20 km.

The SiP MZM employed has a fiber coupling loss close to 9 dB at 1309 nm and approximately 4.7 dB on-chip insertion loss at $V_{\text{bias}} = 3$ V at maximum power transfer, or 7.7 dB at quadrature. Reducing this rather large total insertion loss of ≈ 16.5 dB by 3 dB can be done by improving the coupling efficiency of each surface grating coupler that couples light in and out of the chip. For example, a fiber to chip grating coupler of 2 dB insertion loss was recently demonstrated in [148]; better than the present ≈ 4.5 dB per grating coupler.

To provide visual inspection of the signal quality, Fig. 5-7 shows the eye diagrams for both PAM-4 and PAM-8 in B2B.

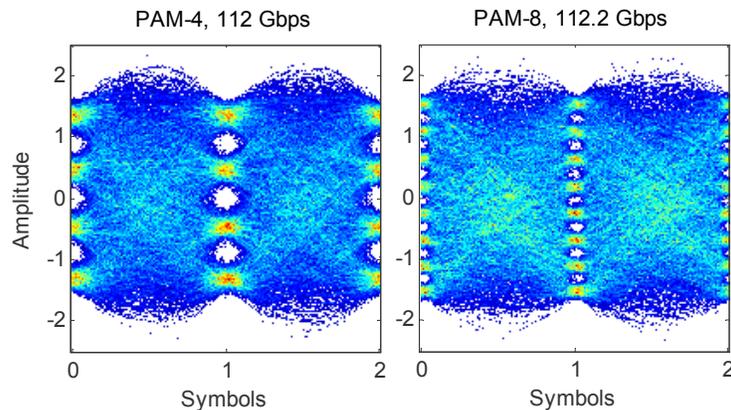


FIGURE 5-7: Eye diagrams: PAM-4 ($\text{BER} = 3.2 \times 10^{-4}$); PAM-8 ($\text{BER} = 1.4 \times 10^{-3}$).

5.1.5.2 Performance Impact from Pre- and Post-Equalizer Lengths

In the following, we investigate the impact on the transmission performance for varying number of equalization taps used at the Tx and Rx. The number of taps in h_{inv} and h_{Rx} directly impacts the power consumption of the Tx and Rx, respectively.

We study the taps' impact after 10 km of SMF fiber, for both PAM-4 and PAM-8 at 112 Gb/s. It is noteworthy to mention that we used different tap spacings for the Tx compensation and the Rx equalization. At the Tx, the pre-compensation filter is applied at the DAC's clock rate, i.e.,

70 Gsamples/s, giving a tap spacing of ≈ 14.3 ps. However, at the Rx, the adaptive equalizer on startup is always applied at 2 SPS, giving a tap spacing that is symbol rate dependant and equals $T/2$, where $T \approx 26.7$ ps for PAM-8 and ≈ 17.9 ps for PAM-4. Applying the Rx equalization at the same rate as the Tx sampling speed would better mimic a real ASIC where both Tx and Rx blocks operate on the same clock rate. However, filter adaptation and sample decimation performed at the Rx is more complex for non-integer SPS. We experimentally observed that using more than 83 taps at the Tx did not provide better transmission performance, notwithstanding the number of the Rx taps. Hence, the variation of the number of Tx taps is upper bounded at 83.

First we present in Fig. 5-8 the BER as a function of the number of pre-compensation taps employed at the Tx. The DC bias applied to the SiP-MZM is set to 3 V. Each curve represents a different number of equalizer taps employed at the Rx. The best BER for 112 Gb/s PAM-4 after 10 km is 4.3×10^{-4} and is obtained by using 75 Tx taps and 35 Rx taps.

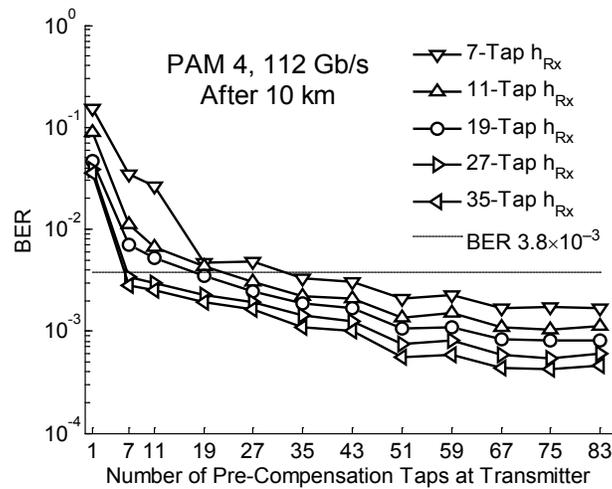


FIGURE 5-8: PAM-4, 112 Gb/s BER variation after 10 km as both the Tx and Rx number of pre- and post-compensation are varied. $V_{\text{bias}} = 3$ V.

It is interesting to note the interplay between the number of FIR taps employed at the Tx and Rx and their impact on the BER performance. A decreasing number of Tx taps increases the BER, but can greatly be compensated by an increase of the number of Rx taps. For PAM-4, one can reach a BER below the FEC limit by using either a combination of 19 taps at the Tx and 19 at the Rx, or reduce the complexity of the Tx by employing only 7 Tx taps and offload on the Rx and use

27 Rx taps. One can also decide to load the Tx and use 35 Tx taps with an end goal of reducing the number of Rx FFE taps to only 7. All cases providing a BER below FEC limit.

Figure 5-9 provides a similar study, applied to 112 Gb/s PAM-8, where we maintain the 3 Volts DC bias. A similar trend of increased BER with decreasing Tx tap length is observed. However, we can observe that decreasing h_{inv} from 51 taps to 7 taps has a greater impact for the PAM-4 format than for PAM-8. This can be understood by noting their different symbol rates and pulse shaping roll-off factors. Pre-compensation of a 56 GBaud signal with small roll-off factor requires quasi-perfect spectral compensation and flattening up to 28 GHz. On the other side, pre-compensating a 37.4 GBaud signal of larger roll-off factor requires only spectral compensation up to about 18.7 GHz, with a more relaxed compensation constrain from the slower pulse shaping filter’s amplitude decay around $f = 1/2R_B$.

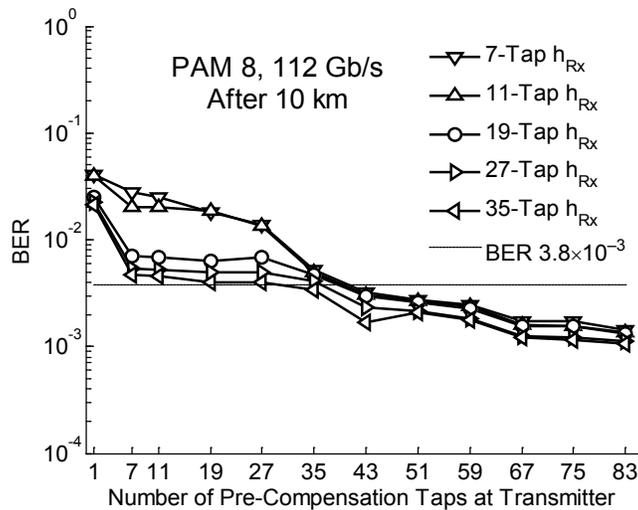


FIGURE 5-9: PAM-8, 112 Gb/s BER variation after 10 km as both the Tx and Rx number of pre- and post-compensation are varied. $V_{bias} = 3$ V.

For a large number of Tx taps between 51 and 83, increasing the Rx tap count only marginally improves the BER for PAM-8, as opposed to a more substantial improvement with PAM-4. This can be explained by the fact that starting at 51 taps, h_{inv} is able to correct for almost all the required spectral compensation when PAM-8 is employed. PAM-8’s optimum BER is about 2.5 times that of PAM-4 when 83 Tx and 35 Rx taps are employed. Even if PAM-4’s BER increases

about 1.7 times faster than PAM-8's when the Tx taps length is reduced from 75 to 7, PAM-4 always outperforms PAM-8 for any Rx taps length greater than 11.

Although the number of equalization taps directly impacts the power consumption, it is difficult and not our intent to evaluate a finite “Joules/second” cost as the energy depends on a plethora of parameters, including CMOS manufacturing technology, DAC/ADC architectures, filtering/processing implementation and the average bit resolution of each DSP module [149]. Nonetheless a study on the relative power consumption would be appropriate in the context of our investigation on pre- and post-equalizers' tap length. As power consumption is proportional to the filter's complexity [150], we use the latter for evaluating power consumption.

The complexity of a filter can be evaluated by its required number of real multiplication [149, 150]. In this DAC/ADC based IM/DD system, signals and equalizers are all real valued. Complexity can be assessed over a finite duration of a nanosecond instead of “per bit” or “per symbol” as the Tx and Rx operate at different speeds. As specified in [150], a digital signal processor operating at f_s Gsamples/s and applying a filter of length N on f_s samples (there are f_s samples in a nanosecond) has a complexity proportional to $f_s N$ for a time-domain implementation and to $f_s N_{FFT} \log_2(N_{FFT}) / N_{Data}$ for a frequency-domain implementation, where $N_{FFT} = N_{Data} + N - 1$ is the FFT size and N_{Data} is the number of data points equalized at a time. The Tx operates at a sampling rate of $f_s = f_{s-DAC} > R_B$ and the Rx at $f_s = 2R_B$, or twice the symbol rate. There is a maximum number of taps after which a frequency domain implementation yields a smaller complexity. From [150], this tap number threshold can be as little as 14 and suggests that the pre- and post-equalizers at Tx and Rx may be deployed in different implementation domains. Notwithstanding the implementation domain of the filter, there is always a monotonic relation between complexity and filter length. Signal quality and equalizer complexity becomes a tradeoff as seen in Figs. 5–8 and 5–9.

For lower Rx post-compensation complexity and power consumption, the feed-forward equalization could be done at the symbol rate, using a T -spaced equalizer. In this case, with a realistic assumption that the Rx ADC samples at the same clock rate as the DAC's rate of 70 Gsamples/s, the DSP stack shown in Fig. 5–4 would change: no resampling required, bulk matched filtering done at $70/R_B$ SPS, followed by bulk downsampling from $70/R_B$ to 1 SPS. At which point, the FFE equalization is performed at a T -spacing, and half the Rx tap counts mentioned below are required

to cover the same temporal duration of the filter. The tradeoff is a smaller spectral content covered and equalized by the FFE: only half that compared to when equalizing at $T/2$.

Finally, to further reduce the Rx complexity, pulse shaped signals with a roll-off factor very close to zero or equating zero can be sampled and processed at the Rx directly at the symbol rate. In this case, the Rx DSP stack does not contain any resampling nor matched filtering blocks, and only T -spaced FFE with clock recovery is performed. This allows suppression of two Rx DSP blocks, significantly lowering the Rx power consumption, at the expense of worsening the Rx performance.

5.1.6 Conclusion

We studied the impact on system performance of the bias voltage applied to the common n -doped region of a series push-pull traveling wave silicon MZM for IM/DD transmission over short distances between 0 and 20 km, at 1.31 μm , for PAM-4 and PAM-8 format at a bit throughput of 112 Gb/s. We show that in B2B a low V_{bias} value is detrimental as the “AC” portion of the optical power presented to the Rx is large and stronger nonlinear compensation is unnecessary when reducing V_{bias} as it increases the PAPR and slightly impedes the modulator’s speed. We show that after propagation when the system is thermal noise limited, as both the “DC” and “AC” portions of the optical power decrease, increasing the AC swing with respect to V_{π} at the Tx becomes beneficial. Those observations apply for both PAM-4 and PAM-8 formats. We observe error-free operation for biases between 2 and 4 V for both PAM orders up to 10 km. In this section, compared to what we reported in Fig. 4–19 in Section 4.2, we realize 112 Gb/s using the physical layer specification that was recently ratified in IEEE 802.3bsTM: PAM-4 at 56 Gbaud below the hard decision FEC threshold of 3.8×10^{-3} . We show that the transmission system is simply power limited by demonstrating equal BER performance over propagation distances ranging from 0 to 20 km when the power presented to the Rx is maintained. We studied the BER impact when varying Tx and Rx pre-compensation and equalization filter lengths are employed in this DSP-based IM/DD link. We show that a decreasing number of taps at the Tx increases the BER, but can be compensated by an increase of the number of taps at the Rx. For PAM-4, one can reach a BER below the FEC limit by using either of the following combination of (Tx,Rx) taps: (19,19), (7,27), or (35,7). As the number of taps has an immediate impact on the ASIC power consumption, it provides room to decide if the load of compensation and equalization is mostly applied at the Tx, Rx, or equally split.

5.2 Three Dimensional Modulation for Self-Beating Direct Detection Systems

After the unidimensional modulation employing IM/DD in Section 5.1, we propose a new format to greatly improve the bit rate delivered for short reach optical transceivers targeting data center interconnects. We propose to modulate the three degrees of freedom of a lightwave in single mode fiber while maintaining a direct detection scheme. The modulation format allows to almost triple to binary throughput of single wavelength modulation with direct detection compared to the IM/DD transceiver of Section 5.1. In this section, we present the transmitter, the Stokes receiver and we study two digital signal processing schemes to recover information on all three dimensions.

5.2.1 Introduction

Data center traffic is increasing incessantly, from the unabated growth of cloud-based applications, server virtualization, cloud computing, web services, and content delivery. This growth is driving the demand for faster inexpensive short reach optics for intra- and inter-data center traffic. Amongst the various solutions, externally modulating a laser to produce 100 Gb/s capacities on a single wavelength is desirable not only because this approach scales well to next generation targets of 400 Gb/s and 1 Tb/s, but also because this approach has a lower component count compared to multiplexing approaches that use multiple wavelengths and/or fibers. Intensity modulation with direct detection also proves to be more economical than coherent approaches [151]. Several 100 Gb/s short-reach IM/DD transmission experiments using various higher order modulation formats operating on a single wavelength modulated on a single polarization have recently been reported. These results include pulse amplitude modulation over 4 levels, discrete multi-tone, and single- or multi-band carrier-less amplitude and phase modulation [5, 129, 152, 153].

Self-coherent approaches have also recently been reported [151]. In this approach the LO is transmitted alongside the modulated data on the orthogonal state of polarization. In addition to increased receiver sensitivity, self-coherent systems allow transmission over greater distances because of the availability at the receiver to a linear mapping of the optical field from the received electrical signals. Transmission over 160 km at 160 Gb/s was reported in [154] using self-coherent detection. For applications requiring shorter distance transmission (e.g. IEEE standard 802.3bm-2015 which includes a 10 km Long-Range specification [155]) polarization division multiplexing provides a means of increasing data throughput. As an example, we reported in [156] transmission at 224 Gb/s over 10 km using PDM with Stokes-vector direct reception, where each SOP is modulated using PAM4

at 56 GSymbols/s. Other experiments reporting IM/DD on two polarizations perform polarization demultiplexing optically and manually at 112 Gb/s [47].

More recently, we reported in [7] 336 Gb/s transmission using a single 1550 nm laser which was intensity modulated on two polarizations, with an additional modulation of the phase difference between the two polarizations. Multi-level signaling was generated on all three degrees of freedom (DOF) of the Stokes representation of an optical field. Specifically, 56 Gbaud PAM4 intensity modulation on two polarizations was followed by 56 Gbaud, 4 phase level inter-polarization phase modulation to achieve 6 bits/symbol DP-PAM4+4PM signaling. The recovery of the 3 DOFs of this format was achieved using direct detection and a two stage DSP algorithm that operated on 4D power vectors after direct detection.

In this work, we propose a new DSP approach to process such 4D vectors and recover the IM and PM information. The DSP is comprised of three stages instead of the two stages employed in [7]. We compare the two DSP approaches using two metrics: the computational complexity (CC) and the performance of the algorithms. CC is calculated using the total required number of real multiplications (RM) and real additions (RA) applied to the 4D input signals to recover the two intensity and phase symbols, and are calculated per symbol. The performance of the two DSP approaches is compared using the bit error ratios from error counting. We show that the three stage DSP requires much less RM and RA per symbol than the two stage approach used in [7], for the same system performance. In the case of a static system with a stable SOP and equal filter lengths, the three stage algorithm is approximately 2.85 times less computationally intensive than the two stage algorithm, and in the more realistic case where SOP is slowly varying and requires tracking, the proposed algorithm is m -times more efficient, where m is the number of the taps in the first filtering stage.

This section is sub-divided as follows: we present in Section 5.2.2 the 6 bits/symbol format including a description of the transmitter and receiver used to generate and recover the format. We also introduce the concept of 4D power vectors and the impact of polarization rotation on these vectors. In Section 5.2.3, we address the impact of intersymbol interference from components employed at the transmitter and receiver and how ISI can be mitigated using DSP in digital to analog converters and analog to digital converters. Section 5.2.4 describes the three stage DSP stack and explains how SOP is efficiently tracked by updating a square matrix. In Section 5.2.5, we assess the CC of the two DSP approaches under study. Section 5.2.6 describes the experimental

test bed and the performance results obtained when we vary both the symbol duration and the SOP of the optical signal. Finally, we conclude in Section 5.2.7.

5.2.2 Format Description and Transceiver for Higher Order Modulation, Stokes-Vector Detection

There are two principal ways to vectorially represent an optical field: the Jones and Stokes representations. Both are based on a reference polarization coordinate system. The Jones space is a two dimensional complex space where each dimension represents an orthogonal SOP and where the argument of each dimension is the complex optical field in said dimension. The Stokes space, on the other hand, is a three dimensional real-valued space, where each dimension represents the difference in optical power after the field passes through different pairs of orthogonal polarizers. The first dimension is the difference in optical power after linear polarizers in the \hat{x} - and \hat{y} -directions, the second dimension is that after polarizers aligned at $+45^\circ$ and -45° , and the third dimension is after a right hand circular polarizer and a left hand circular polarizer. The Jones representation of an optical field is

$$\vec{E} = \begin{bmatrix} E_{\hat{x}} \\ E_{\hat{y}} \end{bmatrix} = \begin{bmatrix} E_x e^{i\phi_x} \\ E_y e^{i\phi_y} \end{bmatrix} = e^{i\phi_x} \begin{bmatrix} E_x \\ E_y e^{i(\phi_y - \phi_x)} \end{bmatrix} \quad (5.1)$$

and the Stokes representation of the same field is

$$\vec{S} = \begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} E_{\hat{x}} E_{\hat{x}}^* - E_{\hat{y}} E_{\hat{y}}^* \\ 2Re\{E_{\hat{x}} E_{\hat{y}}^*\} \\ -2Im\{E_{\hat{x}} E_{\hat{y}}^*\} \end{bmatrix} \quad (5.2)$$

$E_{\hat{x}}$ and $E_{\hat{y}}$ are complex valued, E_x and E_y are signed real values and * represents complex conjugation. Another Stokes term, $S_0 = (S_1^2 + S_2^2 + S_3^2)^{1/2} = E_{\hat{x}} E_{\hat{x}}^* + E_{\hat{y}} E_{\hat{y}}^*$ which is the sum of the total powers on \hat{x} and \hat{y} is often employed for Stokes-space field visualization, in which case S_0 is appended to the Stokes-vector of Eq. (5.2), giving a 4D Stokes-vector $[S_0, S_1, S_2, S_3]^T$. The appearance of * in all multiplied fields in Eq. (5.2) cancels any absolute phase information present in Eq. (5.1), i.e. cancels ϕ_x . Therefore, Jones representations have 4 degrees of freedom, being the powers on \hat{x} and \hat{y} ($|E_x|^2$ and $|E_y|^2$) and the absolute phase on \hat{x} and \hat{y} (ϕ_x and ϕ_y), while the Stokes representations only have 3 DOFs: the powers on \hat{x} and \hat{y} and the phase difference between \hat{x} and \hat{y} ($\phi_y - \phi_x$). In this work, all 3 Stokes-space DOFs are modulated and detected, as we did

in [7]. Figure 5–10 depicts one type of transmitter architecture for modulating all Stokes space’s 3 DOFs.

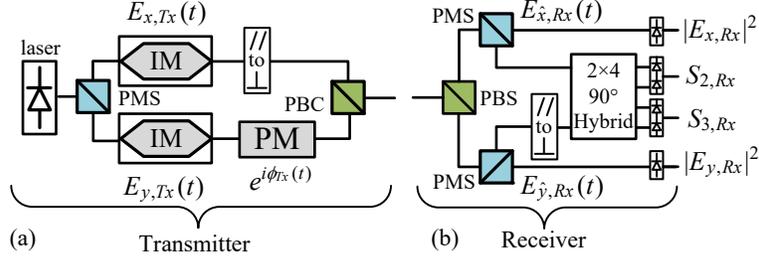


FIGURE 5–10: 3D Stokes Vector (a) Transmitter and (b) Receiver architectures

The transmitter externally modulates a CW laser. The laser is split using a polarization maintaining splitter (PMS). The \hat{x} -branch (top) is intensity modulated using a Mach-Zehnder Interferometer while the \hat{y} -branch (bottom) is also intensity modulated by another MZI, and subsequently phase modulated. The top branch receives a 90° polarization rotation before both tributaries are orthogonally combined by a polarization beam combiner. Because only ϕ_y is modulated, $\phi_x = 0$ and the third DOF ($\phi_y - \phi_x$) is directly proportional to the RF waveform applied to the phase modulator. The axis of the polarization beam combiner of the transmitter defines the Jones and Stokes representation reference axes. When MZIs are employed to perform IM, each MZI is biased at quadrature. From the theoretical field and power transfer function of a MZI, the AC-coupled RF waveform $RF_{\hat{x},\hat{y}}$ applied to each MZI yields an optical field of amplitude $E_{\hat{x},\hat{y},Tx}(t) = \sin(1/4\pi + 1/2RF_{\hat{x},\hat{y}})$. In the presence of a single polarization, a square-law detector would then generate a photocurrent proportional to $|E_{\hat{x},\hat{y},Tx}(t)|^2 = 1/2(1 + \sin(RF_{\hat{x},\hat{y}}))$. If $RF_{\hat{x},\hat{y}}$ is small with respect to the MZI’s V_π or if the RF waveform was compensated by the arcsin function before being applied to the interferometer, the photocurrent is then directly proportional to $RF_{\hat{x},\hat{y}}$.

Figure 5–10(b) shows one type of receiver architecture capable of recovering information on all three dimensions for a Stokes-space representation of the optical field, as seen by the receiver’s polarization axes. A polarization beam splitter first splits the polarization multiplexed incoming light into two orthogonal SOPs. Each branch is further split by a polarization maintaining splitter where one of the two outputs are directed to PIN-TIA photo-receivers, forming the electrical fields $|E_{\hat{x},\hat{y},Rx}(t)|^2$. The other outputs of the two PMS are combined in a 90° optical hybrid, where one is first rotated by 90° to insure that field interference in the hybrid is on the same polarization. Each

hybrid output pair is detected by balance photodiodes to remove the common RF terms per pair. The two outputs after balanced detection give RF waveforms $S_{2,Rx}$ and $S_{3,Rx}$. From Eq. (5.2) and using the relation $S_0 = |E_{\hat{x},Rx}|^2 + |E_{\hat{y},Rx}|^2$, the receiver structure of Fig. 5–10(b) gives access to $|E_{\hat{x},Rx}|^2 = 1/2(S_{0,Rx} + S_{1,Rx})$, $|E_{\hat{y},Rx}|^2 = 1/2(S_{0,Rx} - S_{1,Rx})$, $S_{2,Rx}$ and $S_{3,Rx}$. In this work, the 4D-power vector that we collect and process digitally is $\vec{V} = [|E_{\hat{x},Rx}|^2, |E_{\hat{y},Rx}|^2, S_{2,Rx}, S_{3,Rx}]^T$ instead of the common Stokes representation $[|E_{\hat{x},Rx}|^2 + |E_{\hat{y},Rx}|^2, |E_{\hat{x},Rx}|^2 - |E_{\hat{y},Rx}|^2, S_{2,Rx}, S_{3,Rx}]^T$.

It is known that single mode fibers, like the one connecting the transmitter to the receiver, do not maintain the state of polarization of light. The SOP of the light undergoes a unitary rotation with respect to the transmitter polarization axis. Rotations in SMF fiber are represented by a unitary matrix applied to the complex Jones vector. Applying such a matrix to the transmitted Jones vector of Eq. (5.1) gives the resulting Jones vector

$$\underbrace{\begin{bmatrix} a & -b \\ b^* & a^* \end{bmatrix}}_U \begin{bmatrix} E_{\hat{x},Tx} \\ E_{\hat{y},Tx} \end{bmatrix} = \begin{bmatrix} E_{\hat{x},Rx} \\ E_{\hat{y},Rx} \end{bmatrix} = \begin{bmatrix} aE_{\hat{x},Tx} - bE_{\hat{y},Tx} \\ b^*E_{\hat{x},Tx} + a^*E_{\hat{y},Tx} \end{bmatrix} \quad (5.3)$$

where U is the Cayley-Klein representation of a unitary matrix and where a and b are complex entries related by $|a|^2 + |b|^2 = 1$ [157]. Calculating the 4 receiver-side components $|E_{\hat{x},Rx}|^2$, $|E_{\hat{y},Rx}|^2$, $S_{2,Rx}$ and $S_{3,Rx}$ from Eq. (5.3) and relating them the 4 transmitter-side components having a ‘Tx’ subscript reveals the 4D power vectors relation via the following matrix equation

$$\begin{bmatrix} |E_{\hat{x},Rx}|^2 \\ |E_{\hat{y},Rx}|^2 \\ S_{2,Rx} \\ S_{3,Rx} \end{bmatrix} = \underbrace{\begin{bmatrix} |a|^2 & |b|^2 & -Re\{ab^*\} & -Im\{ab^*\} \\ |b|^2 & |a|^2 & Re\{ab^*\} & Im\{ab^*\} \\ 2Re\{ab\} & -2Re\{ab\} & Re\{a^2\} - Re\{b^2\} & Im\{a^2\} + Im\{b^2\} \\ -2Im\{ab\} & 2Im\{ab\} & Im\{b^2\} - Im\{a^2\} & Re\{a^2\} + Re\{b^2\} \end{bmatrix}}_M \begin{bmatrix} |E_{\hat{x},Tx}|^2 \\ |E_{\hat{y},Tx}|^2 \\ S_{2,Tx} \\ S_{3,Tx} \end{bmatrix} \quad (5.4)$$

where Re and Im stand for real and imaginary. As one can see, SOP rotation generates linear cross-talk on the received 4D vector. Depending on the rotation, each component of the received 4D vector is a weighted, linear combination of those sent at the transmitter. The M matrix in Eq. (5.4) is quite similar to the Mueller matrix [157]: the Mueller matrix is obtained from M by adding and subtracting the first 2 rows and the first 2 columns. From M , one can consequently also find the equivalent Stokes 3×3 rotation matrix, this time by simply subtracting the first two rows and first two columns.

Polarization Mode Dispersion is another phenomenon occurring in single mode fibers that can be detrimental for long-reach IM/DD systems of large symbol rates. First order PMD leads to a differential group delay between two orthogonal polarization states. The mean DGD of a fiber link of length L can be calculated from the PMD coefficient as $DGD_{\text{MEAN}} = PMD_{\text{COEF}} \times \sqrt{L}$. The PMD coefficient, also known as the PMD link value, of ‘older’ single mode fibers is specified to be ≤ 0.1 ps/ $\sqrt{\text{km}}$, while newer SMF fibers exhibit $PMD_{\text{COEF}} = 0.04$ ps/ $\sqrt{\text{km}}$. For a short reach distance of 10 km, the mean DGD is below 0.32 ps. Even at 56 GSymbols/s, such a differential group delay has negligible impact on the performance. An experimental validation of the negligible impact of DGD on a single-polarization 56 Gbaud PAM-4 IM/DD system was demonstrated in [5], where performances remain unchanged between 0 and 20 km when the received signal power is kept constant.

5.2.3 Intersymbol Interference Impairments

There are three main signal impairments that the receiver DSP has to compensate for. Two of which are related to the limited bandwidth of the electronic and/or optoelectronic components at *i*) the transmitter and *ii*) the receiver, and one impairment is related to *iii*) polarization rotation in SMF fiber. We have explained the impact of SOP rotation on the received 4D power vector in Section 5.2.2. In this section, we address the impact of the limited bandwidths of the transmitter-side and the receiver-side components, both of which induce waveform ISI that DSP from the DACs and the ADCs can mitigate.

The interference between neighboring symbols is due to long and imperfect impulse responses of RF components. The non-flatness of the spectral response magnitude, the 3-dB and the 10-dB bandwidths define the amount of ISI induced by each component. The amount of ISI induced also naturally depends on the symbol rate and the pulse shape of the waveform propagating through the component. DACs typically have smaller 3-dB bandwidths (BW) than their ADC counterparts. For instance, ADCs with 3-dB BWs greater than 30 GHz have been available since 2009 [158], while DACs with bandwidths of 24 GHz have only recently become available [159]. Moreover, the flat spectral magnitude response from DC to 32 GHz of ADCs as reported in Fig. 3 of [158], exhibiting a variation within solely ± 0.25 dB (between 0.97 and 1.03 in linear scale) up to 29.5 GHz, expresses how little waveform distortion analog to digital converters impose on the waveform. Generating

signals from a DAC of high symbol rate with high spectral integrity, however, requires waveform pre-compensation, also called transmitter pre-emphasis, where a finite impulse response filter is applied to the waveform before it is transferred to the DAC. There are a plethora of documentations on transmitter-side signal pre-emphasis allowing opening the eye diagram out of a digital to analog converter. For example, optimization of such FIR filter under LMS criterion is presented in [160], while transmitter pre-emphasis and receiver equalization behaviors are addressed in [161].

Reducing the ISI and opening the eye is beneficial in terms of signal integrity, but reduces significantly the effective output voltage swing, itself reducing the amplitude of the driving voltage applied to the electro-optic transducers (either the MZI or the PM). DACs can compensate not only their own spectral response but also that of subsequent components. Fortunately, electro-optic transducers and linear modulator drivers offer much larger 3-dB BWs (≥ 30 GHz) than DACs and therefore these components are not bottle necks at the transmitter.

On the receiver side, ADCs offer larger BWs, higher effective number of bits over increasing frequency and larger voltage conversion range than DACs. Moreover, large bandwidth photo-receivers and photo-detectors are also readily available. Consequently, there is generally less ISI imprinted on the signal by receiver-side components, including the ADC, than the DAC induced ISI imprinted at the transmitter. As the Rx components are less frequency constraining and because ADCs offer better conversion, it becomes interesting to split and share the load of the DAC compensation towards the ADC. There are pros and cons of waveform compensation via DSP using either a DAC or an ADC and in this context we address the following two cases: one where the DAC fully compensates for all ISI, and one where the DAC does not apply any FIR pre-compensation filtering and therefore the DSP applied after the ADC mitigates all ISI.

In the first case, DAC ISI compensation mitigates ISI but reduces signal swing, or signal power. As a consequence, a multi-level signal will experience a decrease in the spacing between the levels as the compensation increases. The increased signal quality achieved by reducing ISI is accompanied by a decrease of the signal to noise ratio, where transmitter-side noise is dominated by both DAC output noise and modulator driver noise. In the second case, not compensating for the limited DAC plus RF component frequency response yields a signal of larger swing and power, and therefore of larger SNR, but with very poor signal quality. Thus, there is an interplay between the signal quality and SNR with transmitters employing DACs.

If no DAC compensation is applied, ISI mitigation of the high power signal is required after the ADC in order to improve signal quality and recover distinguishable levels. The predominant detrimental effect of receiver side signal equalization is that it also ‘equalizes’ and boosts the noise coming with the signal. This noise can be ASE noise from a preceding optical amplifier, thermal and shot noise from photodiodes, or noise coming from TIA amplification. While this equalization reduces ISI, it also increases the noise variance and therefore reduces the SNR. There is also an interplay between signal quality and SNR with equalization via DSP from an ADC, and we note that while the SNR metric quantifies the signal power with respect to surrounding noise, it does not quantify overall signal quality. Therefore, for a given system using specific components, the best splitting ratio of signal compensation between the DAC and the ADC is found empirically.

5.2.4 DSP, Filters and Updates

In addition to residual transmitter-side and receiver-side ISI, the digital signal processing applied at the receiver also needs to derotate the signal after rotation in SMF fiber. The receiver side DSP for higher order modulation formats with Stokes-vector detection consists of a three stage process: Stage 1) receiver-side ISI compensation, Stage 2) derotation, and Stage 3) residual transmitter side ISI mitigation, which is applied in reverse order relative to impairments experienced by the transmitted signal along the channel. Figure 5–11 is a visual representation of this DSP processing.

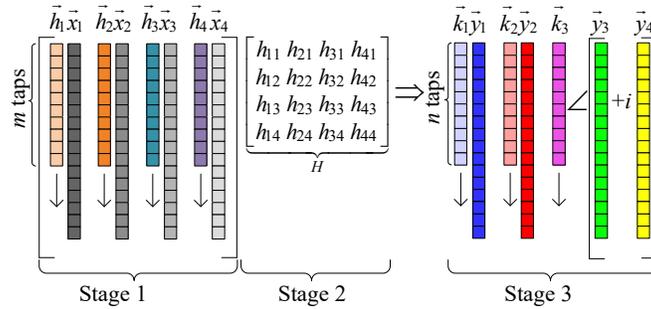


FIGURE 5–11: Visual representation of the three-stage DSP applied at the receiver

Stage 1 filters all 4 received waveforms independently and serves to mitigate ISI that each Rx channel of Fig. 5–10(b) may have imprinted on their respective waveforms. The inputs $\vec{x}_{1,2,3,4}$ are respectively $|E_{\hat{x},Rx}(t)|^2$, $|E_{\hat{y},Rx}(t)|^2$, $S_{2,Rx}(t)$ and $S_{3,Rx}(t)$ taken from the ADCs and each filter \vec{h}_i filters one input \vec{x}_i . The second stage, Stage 2, outputs four weighted, linear combinations of the

4 filtered outputs of Stage 1. The first and second combinations give an estimate of $\langle |E_{\hat{x},Tx}|^2 \rangle$ and $\langle |E_{\hat{y},Tx}|^2 \rangle$ which is the intensity modulation on each polarization generated at the transmitter. The third and fourth combinations, when added in a complex form and after taking the angle of the result, give an estimate of the phase modulation $\langle \phi_y \rangle$ generated at the transmitter. In Fig. 5–11, we identify those 4 combinations as $\vec{y}_{1,2,3,4}$ for $\langle |E_{\hat{x},Tx}|^2 \rangle$, $\langle |E_{\hat{y},Tx}|^2 \rangle$, $\langle S_{2,Tx} \rangle$ and $\langle S_{3,Tx} \rangle$, respectively. Equation (5.5) shows the mathematical representation of the combined first two stages.

$$\begin{aligned}
h_{11}\vec{h}_1^T \vec{x}_1 + h_{12}\vec{h}_2^T \vec{x}_2 + h_{13}\vec{h}_3^T \vec{x}_3 + h_{14}\vec{h}_4^T \vec{x}_4 &= \vec{y}_1 \\
h_{21}\vec{h}_1^T \vec{x}_1 + h_{22}\vec{h}_2^T \vec{x}_2 + h_{23}\vec{h}_3^T \vec{x}_3 + h_{24}\vec{h}_4^T \vec{x}_4 &= \vec{y}_2 \\
h_{31}\vec{h}_1^T \vec{x}_1 + h_{32}\vec{h}_2^T \vec{x}_2 + h_{33}\vec{h}_3^T \vec{x}_3 + h_{34}\vec{h}_4^T \vec{x}_4 &= \vec{y}_3 \\
h_{41}\vec{h}_1^T \vec{x}_1 + h_{42}\vec{h}_2^T \vec{x}_2 + h_{43}\vec{h}_3^T \vec{x}_3 + h_{44}\vec{h}_4^T \vec{x}_4 &= \vec{y}_4
\end{aligned} \tag{5.5}$$

If not all transmitter-side ISI is compensated by the DACs, the intensity and phase modulation $\langle |E_{\hat{x},Tx}|^2 \rangle$, $\langle |E_{\hat{y},Tx}|^2 \rangle$, $\langle \phi_y \rangle$ generated at the Tx will have residual ISI. Stage 3 mitigates this residual ISI by applying an independent FIR filter to the phase and the two intensity waveforms. In Fig. 5–11, the two intensity filters are represented by $\vec{k}_{1,2}$ and the phase filter by \vec{k}_3 . Equation (5.6) is a mathematical description of the final Stage 3 filtering.

$$\begin{aligned}
\vec{k}_1^T \vec{y}_1 &= \vec{z}_1 = \overline{|E_{x,Tx}|^2} \\
\vec{k}_2^T \vec{y}_2 &= \vec{z}_2 = \overline{|E_{y,Tx}|^2} \\
\vec{k}_3^T (\vec{y}_3 + i\vec{y}_4) &= \vec{z}_3 = \overline{\phi_y}
\end{aligned} \tag{5.6}$$

In practice, neither $\vec{h}_{1,2,3,4}$ nor $\vec{k}_{1,2,3}$ evolves in time: once those filters are characterized, they remain invariant. However, the SOP inside the SMF fiber varies with temperature, stress, and vibration. Consequently, the state of polarization of the light as it enters the receiver is unknown and slowly varying. As the SOP varies, the 4-by-4 matrix H in the second stage needs to adapt its coefficients to track the SOP. Coefficients adaptation can be done using the least mean square algorithm, also known as the stochastic gradient descent. When updating using LMS, an error signal is minimized in a least mean square sense. This error signal is the difference between the *desired* symbol and the current output value. To compute the errors for outputs $y_{1,2,3,4}$, we need to find the 4 desired values to compare y_i against, $\forall i$. In theory, all final symbol s_i of the last stage would need to be back-filtered by the inverse (\vec{k}_i^{-1}) of the filter \vec{k}_i . In practice, this is very computationally

intensive and, as the analog BW of the components employed at the receiver is large, little Rx-side ISI is imprinted. Therefore, directly using s_i 's to update the second stage is a sufficient approximation. The target outputs for y_3 and y_3 need be reconstructed from the desired power and phase symbols $s_{1,2,3}$ as $d_3 = A\cos(s_3)$ and $d_4 = A\sin(s_3)$, where $A = 2(1/2 + s_1)^{1/2}(1/2 + s_2)^{1/2}$.

Updating the 16 coefficients of the H matrix of Fig. 5–11 using the LMS adaptation is performed as follows

$$h_{ij} = h_{ij} + \mu (d_i - y_i) x'_j \quad (5.7)$$

where x'_i are the 4 outputs of the first stage, $x'_i = \vec{h}_i^T \vec{x}_i$ and where $d_{1,2,3,4}$ are respectively s_1 , s_2 , $A\cos(s_3)$ and $A\sin(s_3)$.

5.2.5 Complexity of the Algorithm

In this section we compare the computational complexity of the receiver-side DSP approach presented in our previous work [7] to the new approach proposed here. In order to visualize the differences between the current approach and the approach found in [7], Fig. 5–12 depicts how Stage 1 and Stage 2 shown in Fig. 5–11 were merged into a single stage in [7].

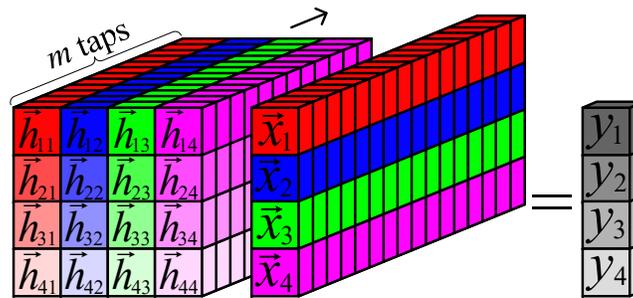


FIGURE 5–12: First DSP stage of [7]. In this work, we decouple this stage into two stages.

Computational complexity will be assessed using the total number of real additions and real multiplications needed to recover one symbol, i.e. the CC per symbol. From there, the CC per bit can be derived by dividing the latter by 6.

In [7], the DSP stack is based on a two-stage process. The first stage is a combination of the first two stages of this work, where receiver-side ISI and polarization derotation are combined

in a single stage. The third stage also differs because in [7] outputs \bar{y}_3 and \bar{y}_4 of Eq. (5.5) are independently post-filtered before being combined in a complex fashion to recover the angle.

We compute the CC based on the temporal length of the first and third stage filters, assuming for simplicity that all filters per stage have the same length. This may not be necessary in practice because, for instance, the spectral response of the PDs collecting S_2 and S_3 could be different from those collecting $|E_{\hat{x}}|^2$ and $|E_{\hat{y}}|^2$. We also assume that the filters having a temporal dimension, i.e. those of Stage 1 and Stage 3, are processing data at a sampling rate of 2 samples per symbol. In the complexity comparison, the temporal dimension of the first and of the second stage in [7] will be the same as that in first and third stage of this work.

We define the length of the first and third stages as m and n , respectively. If m and n are even, they cover $m/2$ and $n/2$ symbols, respectively. The CC's of both approaches are obtained by summing the RM and RA required per final recovered intensity and phase. We first evaluate the complexity of computing all 4 $y_{1,2,3,4}$, followed by what is required to compute the final three outputs $z_{1,2,3}$. In this three stage approach, each $\vec{h}_i^T \vec{x}_i$ needs to be computed only once for all outputs $y_{1,2,3,4}$, that are obtained from a weighted sum of the $\vec{h}_i^T \vec{x}_i$'s. In [7], 16 $\vec{h}_{ji}^T \vec{x}_i$ are required to obtain all 4 $y_{1,2,3,4}$. Compared to Eq. (5.5), each y_i in [7] are $y_i = \vec{h}_{i1}^T \vec{x}_1 + \vec{h}_{i2}^T \vec{x}_2 + \vec{h}_{i3}^T \vec{x}_3 + \vec{h}_{i4}^T \vec{x}_4$. Knowing that performing one filter $\vec{h}^T \vec{x}$ requires m RM and $m - 1$ RA per output, we conclude that obtaining the 4 $y_{1,2,3,4}$ requires $4m + 16$ RM and $4m + 8$ RA using this decoupled approach and $16m$ RM and $16m - 4$ RA using the merged approach of [7]. z_1 and z_2 are obtained the same way in both approaches and each require n RM and $n - 1$ RA.

For the phase recovery used in both approaches, we assume that generating a complex number $(a + ib)$ does not require an effort larger than what is required to compute a and b individually. Moreover, we assume that obtaining the angle of a complex number $(a + ib)$ can be performed via a look-up table where the starting quadrant is obtained from the signs of a and b . We refrain from assessing the complexity of computing the angle of a complex number and therefore this operation is not taken into account when evaluating CC. Obtaining the final phase output z_3 is different in the two approaches. In [7], y_3 and y_4 are independently filtered and then added in a complex fashion to obtain the final angle. In the present approach, y_3 and y_4 are first added in a complex fashion to obtain the angle that is subsequently filtered thus producing a different CC. $2n$ RM and $2n - 2$ RA are required in [7] while only half of those numbers are required in the present approach. As the

sampling rate is 2 samples per symbol, 2 final output data points need to be computed for every symbol. Table I summarizes the total RM and RA for both approaches.

TABLE 5–1: Computational Complexity Comparison

Process	RM	RA
Three-Stage	$8m + 6n + 32$	$8m + 6n + 10$
Two-Stage [7]	$32m + 8n$	$32m + 8n - 16$

In the new DSP version, the complexity with respect to m is 4 times less due to the splitting of the first merged stage in [7] into Stages 1 and 2. Moreover, the approach to process the phase is also less demanding because only the phase is filtered in the final stage instead of its real and imaginary components, reducing the complexity as a function of n by a factor of $8/6 = 1.33$. For the special case where $m = n$, the proposed approach reduces the CC by a factor of approximately $80/28 = 2.85$. The complexities reported in Table 5–1 are for steady state operation when no updating or tracking are required.

As mentioned in Section 5.2.4, polarization rotations need to be tracked. It is important to mention that tracking the SOP using the proposed method with decoupled stages is much less computationally intensive than tracking SOP using the merged stage in [7]. Using the present method, only the 16 coefficients of the second stage need to be adapted in order to fully track SOP, as demonstrated in Eq. (5.7). Change of input SOP has no impact on the receiver-side ISI removal filters $\vec{h}_{1,2,3,4}$ nor on the residual transmitter-side ISI removal filters $\vec{k}_{1,2,3}$. However, in [7], all 16 \vec{h}_{ij} filters, each of length m , need to adapt to track SOP. Tracking SOP in [7] is therefore m -times more intensive than the tracking using the 3-stage approach.

It is not our intent to evaluate the cost per bit of the proposed format and receiver digital signal processing. Defining a finite “\$/bit” number depends on a plethora of parameter, including CMOS manufacturing technology, DAC/ADC architectures, filtering/processing implementation and the average bit resolution of each DSP module [149], to name a few. However, the complexity of a given process node determines its power dissipation [150]. Consequently, the computational complexity assessment of the proposed process and that of the previous approach simultaneously provides an assessment of the cost in “Joule/second” of the processes that can be further converted to “\$/bit”. The lessening in CC of the proposed process will be reflected in a reduction of power consumption.

5.2.6 Experimental Setup and Results

In this section we present the experimental test bed employed to compare the performance of the two DSP approaches used to process the four waveforms $\hat{x}_{1,2,3,4}$ in order to recover the 6 bits/symbol encoded on the two intensities and the inter-polarization phase. The performance comparison in this section complements the CC comparison of Section 5.2.5 by providing a different metric to compare the two approaches, namely by BER counting.

A schematic representation of the transmitter and receiver test bed is presented in Fig. 5–13. A 14 dBm TUNICS T100S-HP tunable laser operating at 1550 nm is intensity modulated over 4 levels by a silicon photonics intensity Mach-Zehnder modulator fabricated at IME A*STAR, as presented in [34]. The intensity modulator is driven by a DAC running at 70 GSamples/s and amplified by a SHF 806E amplifier. The intensity modulated optical signal is amplified by an Erbium Doped Fiber Amplifier and further filtered by a 0.8 nm optical filter to remove out of band ASE noise. A polarization controller aligns the SOP to the slow axis of a subsequent polarization maintaining splitter which splits the optical signal entering the dual polarization emulator. One branch of the emulator receives an additional phase modulation, also over 4 levels, while the other is delayed by a variable delay line and rotated by 90°. The optical delay serves to decorrelate the intensity modulation on the two orthogonal SOPs. Both tributaries are next combined by a polarization beam combiner. The low- V_π LiNbO₃ phase modulator is driven by another amplified DAC channel. As one polarization receives phase modulation and the other does not, inter-polarization phase variations are proportional to the RF signal applied to the PM.

The receiver shown in Fig. 5–13 is composed of a polarization beam splitter that splits the received signal into two orthogonal polarization axes. Each output port is further split using a 90/10 coupler, where the 90% ports are fed to the two inputs of a 90° optical hybrid and the 10% ports are directed to two Picometrix PT-40D PIN-TIA photo-receivers. The 4 outputs of the 90° hybrid are connected to two U2T BPDV2020R balanced photodetectors. Uneven coupling ratios were necessary because the two BPDs did not have TIAs. In the case where all photo-detectors are connected to trans-impedance amplifiers, the coupling ratios should be 66/33 [156]. Temporal and polarization alignment of the two optical signals prior to entering the hybrid is performed by a VODL and polarization controllers, respectively. We note that the transmitter and receiver shown in Fig. 5–10(a) and Fig. 5–10(b), respectively, can be built in a silicon photonic platform, and

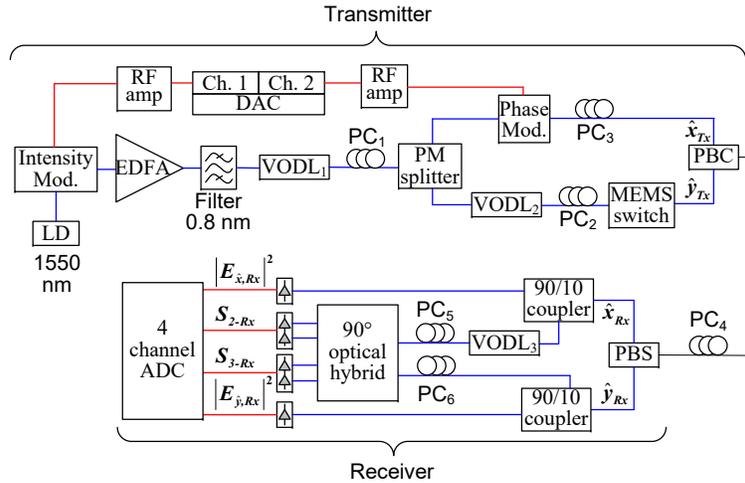


FIGURE 5–13: Experimental setup: transmitter and receiver for the 6 bits/symbol generation and direct detection

that several components shown in Fig. 5–13 would not be required when realizing an integrated transceiver solution.

The 4 RF waveforms generated by the two photo-receivers and the two BPDs are real-time sampled by four 8-bit Analog to Digital Converters sampling at 80 GSamples/s from an Agilent DSO-X 96204Q real-time oscilloscope and stored for offline processing. For both DSP approaches under study, resampling of the 4 waveforms from 80 GSa/s to 2 samples per symbol is initially performed. The 4 resampled waveforms are then fed to either the DSP stack of [7] or the one presented in this work.

For the performance comparison of the two DSP approaches, we first vary the symbol rate of the DP-PAM4+4PM format from 28 Gbaud to 58.4 Gbaud and compute the bit error ratio using error counting. The results are shown in Fig. 5–14. Very similar performance is obtained using both techniques. Thus, the proposed DSP stack requires much less complexity for the same BER performance.

In Fig. 5–15, we compare the performance of the two approaches when we vary the input SOP of the signal incident on the receiver. Results show that the bit error rates yielded by the two DSP approaches are, again, almost identical. We reiterate that tracking the SOP using the proposed method is much less computationally extensive. In [7], $16m$ taps need to change as the SOP changes, as depicted in Fig. 5–12, even if the ISI caused by Rx components is agnostic to

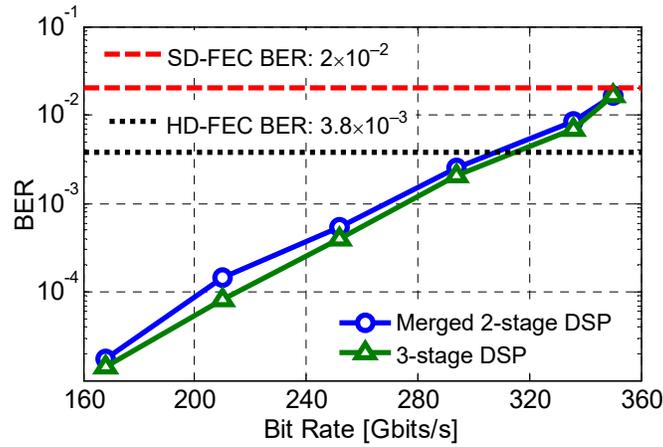


FIGURE 5–14: BER vs Bit rate for the 6 bits/symbol DP-PAM4+4PM format using: the merged two stage (circles) and the proposed three stage (triangle) DSP

SOP orientation. The splitting of this first merged stage into two stages reduces the complexity of tracking SOP by a factor m , without affecting performances, as reported in Fig. 5–15.

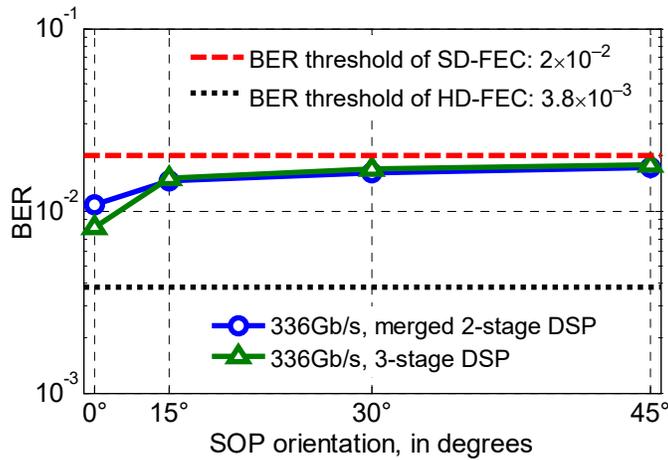


FIGURE 5–15: BER of the two DSP approaches over varying input SOPs. Using: the merged two stage (circles) and the proposed three stage (triangle)

The DP-PAM4+4PM modulation format provides a large modulation efficiency of 6 bits/symbol. Such efficiency can be leveraged in different ways. It can allow reducing the signaling rate while maintaining a target bit rate. For instance, delivering 112 Gbits/s requires operating the system at only 18.7 Gsymbols/s compared to 56 GSymbols/s for PAM4 on a single polarization. Reducing

the signaling rate allows relaxing the bandwidth requirements of analog devices and using less expensive RF components. Conversely, such large modulation efficiency also allows for a significant increase of the data throughput by operating at large signaling rates as demonstrated in this work where 300 Gbits/s are modulated and demodulated below HD-FEC. As DACs are often spectrally constraining components in the transmission system, a thorough analysis of DAC requirements for next-generation high symbol rate systems was presented in [162].

5.2.7 Conclusion

We studied two DSP approaches enabling recovery of all three degrees of freedom of the Stokes representation of an optical field using a Stokes-vector direct detector. We focus on higher order modulation per DOF; specifically, we modulate PAM4 signals on two orthogonal polarizations with additional inter-polarization phase modulation also over 4 levels, giving 6 bits/symbol. The novel DSP proposed in this work to process the three DOFs is compared to the one presented in our previous work [7]. The metrics employed for comparison are the computational complexity of the multi-stage DSP stacks and the BER performance over varying scenarios. Performance is evaluated using the bit error ratio computed for varying symbol rates of the 6 bits/symbol format and over varying receiver input SOPs. We show that the two approaches perform equally versus varying symbol rates and varying SOPs. However, the DSP presented in this work requires a significantly smaller number of RMs and RAs. The computational complexity reduction primarily comes from splitting the first merged stage in [7] into two independent stages, where receiver-side ISI mitigation and polarization derotation are performed separately. For the same filter temporal duration, in the case where the input polarization is stable, the proposed three-stage DSP requires approximately 4 times less RM and RA for the Rx ISI mitigating filters, and approximately 1.33 times less RM and RA for the Tx ISI mitigating filters. In the more likely case where SOP is not stable and slowly varying, tracking of the SOP is m -times more efficient, where m is temporal duration in tap count of the first filtering stage.

CHAPTER 6

Conclusion

6.1 Overview

We demonstrated in this thesis that digital signal processing plays a paramount role in all hierarchical levels of the optical network topology. The rather recent revival of the coherent transceiver allowed a plethora of DSP to be applied at both the transmitter and receiver which gave rise to a disruptive change in optical communications. When applied to Data Center transceivers, be it on the unidimensional IM/DD system or on more complex multi-dimensional transceivers with self-beating detection, DSP has proven to play a key role in the drastic increase of bit rate transfers per wavelength. Finally, for chip-scale photonics, we have show that DSP also leverages the possibilities of passive components through our experiment on 16QAM wavelength multicasting in a silicon nanowire. We have also shown the immense potential of DSP as a system equalizer to greatly ‘open the eye’ and improvement performance when applied to a DAC-driven Silicon Photonic intensity Mach-Zehnder modulator running at very large symbol rates.

6.2 Summary of Original Contribution

In Fig. 6–1, we summarize the original contributions achieved in the thesis on the three aspects of optical communication: Tx and Rx DSP for coherent transceivers, DSP for data center interconnects and DSP enabled chip-scale photonics.

6.2.1 DSP for Coherent Transceivers

In Chapter 3, we tackled optical transmissions using advanced modulation formats and coherent reception and proposed a novel DSP algorithm at the receiver and a novel modulation format for the transmitter. At the receiver, we proposed a new algorithm to recover the state of polarization of the polarization multiplexed signal presented to the receiver. The algorithm performs blindly and allows switching to the steady state DD-LMS polarization tracking mode orders of magnitude faster than the conventional blind Constant Modulus Algorithm. The proposed algorithm works for any input

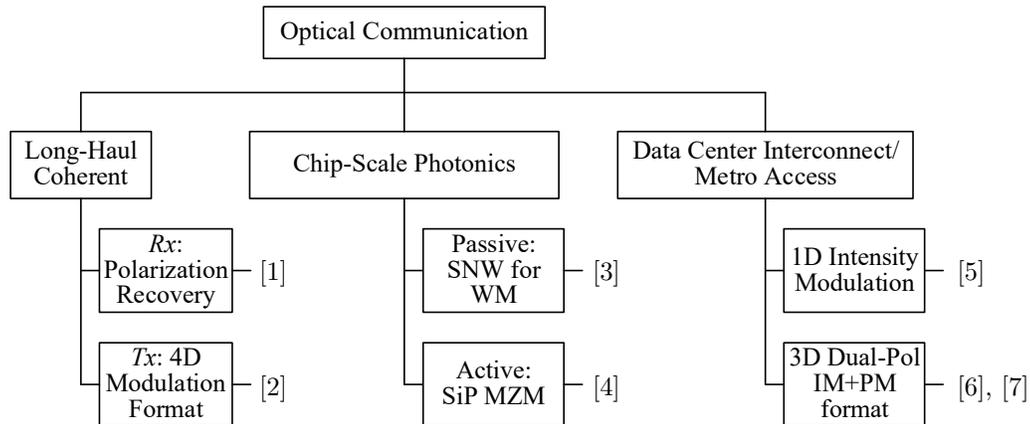


FIGURE 6–1: Summary of the original contributions of the thesis

SOP and any polarization multiplexed square M -QAM format, as experimentally validated. At the transmitter, we proposed and developed a new constant power modulation format. The format is a power constrained version of the Dual-Polarization–8QAM format, having 32 possible symbols and giving in total 5 bits of information per recovered symbol. We experimentally demonstrated that the format is more tolerant to the Kerr nonlinear effects in single mode fiber. We performed an analytical study demonstrating the increased tolerance of the format compared to the power varying format DP–8QAM, and performed experimental validation.

6.2.2 DSP Enhanced Chip-Scale Photonics

In Chapter 4, we studied a passive and an active silicon photonics device. The first device under study was a passive silicon nanowire that, with the help of DSP at the receiver, allowed one-to-six wavelength multicasting of a high order 16QAM signal. This was the first high symbol rate, high order QAM wavelength multicasting in a 20 mm long silicon nanowire. We related the experimental results to the theory and explain the different in the signal to noise ratios of degenerate and non-degenerate idlers from the four-wave mixing process. We also showed the importance of the output grating coupler’s loss and how it impacts the multicasting performance, in conjunction with the four-wave mixing efficiency.

Moreover, we demonstrated the first DSP enhanced active silicon photonics Mach-Zehnder Modulator operating at $1.3 \mu\text{m}$ on a pulse amplitude modulation format for short reach transmission employing a DAC for the transmitter RF signal generation. We presented all the details of the 6 mm

long series push-pull SiP Mach-Zehnder modulator and studied the system performance at various symbol rates, PAM orders, propagation distances, pulse shaping roll-off factor and received signal power. We derived the first analytical model for the Q -factor performance metric applicable for multi-level PAM signaling that takes into account different noise power per level, and we accurately experimentally verified the model in the case of Gaussian noise limited detection. By employing PAM-8 at 37.4 Gbaud, we reported the first demonstration of a 112 Gb/s transmission over 10 km of SMF fiber operating below pre-FEC BER threshold of 3.8×10^{-3} .

6.2.3 DAC/ADC Enabled Data Center Interconnects

In Chapter 5, we improved the BER performance of the previous work by better equalizing the transmitted waveform and by generating the 112 Gb/s bit rate using a 4-level PAM format running at 56 Gbaud: a signaling rate and modulation format recently ratified by the IEEE Standardization body in the 400GBASE-DR4 Physical layer specification. We showed that residual chromatic dispersion and polarization mode dispersion have no impact on the performance up to a transmission distance of 20 km. We also studied the impact on system performance of the bias voltage applied to the common n -doped region of the silicon MZM. Finally and most importantly, we investigated the impact of varying the number of transmitter pre-emphasis taps and receiver equalization taps. We showed that for a target BER, different combinations of transmitter and receiver tap lengths can be used, allowing flexibility in power consumption allocation.

Finally, we reported the first demonstration of a transceiver capable of up to 300 Gb/s using a single 1550 nm laser which is intensity modulated over 4-levels on 2-polarizations with an additional 4-level inter-polarization phase modulation and that uses direct detection. The format provides a spectral efficiency of 6 bits per symbol. The bit error rate is below the hard-decision thresholds of 3.8×10^{-3} . The DP-PAM-4+4PM format is demultiplexed at the receiver using a Stokes receiver. We also proposed and compare two digital signal processing approaches to recover the intensity modulation on the two orthogonal polarizations and the inter-polarization phase modulation. The first approach proposed relies on a 2-stage DSP stack and the second relies on a 3-stage DSP stack. We show that the two approaches exhibit equal BER performance, but the three-stage DSP approach significantly reduces the total required number of real additions and real multiplications by recovered symbol.

6.3 Future Research

Despite all the accomplishments reported in this thesis, there still exists a lot of opportunities for future work to build upon the research conducted. We list below some of the prospective research avenues derived from each of the topics proposed in the thesis.

DSP for Coherent Transceivers

There is obviously a plethora of new digital signal processing to investigate for digital coherent transceivers. All linear impairments occurring in single mode fibers have been addressed and their compensation and mitigation is already implemented in one way or another in current coherent cards. Of course, research on algorithm optimization and on newer, better approaches is always welcome and still presents interests by the research community. However, it is well accepted that the remaining technologically challenging problem to address with DSP, DAC/ADC-leveraged optical coherent transceiver is the compensation of nonlinear effects occurring during signal propagation in SMF. The nonlinear Kerr effects impair the signal quality and greatly reduce both system reach and the modulation order from the limited signal launch power. Fiber nonlinearity can be compensated digitally at either ends of the fiber link. However, the computational complexity of the nonlinear noise compensation is an obstacle for real-time implementation on an ASIC with realistic gate count and power consumption. Several approaches using digital signal processing have been proposed, such as digital back-propagation, Volterra series filtering and perturbation-based nonlinear compensation. Although some have legitimate merits, the accuracy and efficiency of all DSP-based approaches are limited by their required area and power consumption on an ASIC. Research needs to continue in finding better, less demanding approaches to mitigate nonlinear effects. Continuing investigating DSP-based approaches is one solution, but other methods can also serve the means of mitigating nonlinear effects. One such method relies on novel modulation formats and coding. Constant power formats, like the one presented in this thesis [2], is one such approach to deal with nonlinearity. The industry has recently shown great interest in the development of constant power modulation formats for ultra-long systems [163]. However, it would be interesting to develop constant power formats of large alphabet sizes of 256 and more to address simultaneously the problem of nonlinearity and the need for faster wavelengths.

DSP-enhanced Chip-scale Photonics

It is still unclear how much of a role will DSP take in next-generation silicon photonics chips. It is important to distinguish between digital signal processing to ‘open the eye’ and digital signal processing that performs signal processing, like the digital removal of signal cross-talk due to polarization rotation in Stokes-based direct detection of polarization multiplexed IM signals. DSP intended to improve the quality of the signal and to ‘open the eye’ of the received signal is time-invariant: it can be done with a N -tap FIR filter where the tap values do not change in time. The same functionality of this FIR filter can be implemented using a N -tap adjustable Feed-Forward Equalizer imprinted in the CMOS driving circuit [164]. Tap values would be adjusted only once. In fact, a digital signal processing chip is not required in such a case as there is no need to perform a numerical pre-treatment of the signal and push the resulting treated, ‘weird’ signal to a DAC.

On the other hand, digital signal processing can also be used to perform different operations that are not to ‘open the eye’. For instance, it can be used to remove polarization cross-talk after a Stokes receiver, like we did in [6] and [7]. To do this, one can indeed use 4 analog to digital converters followed by a 4-input DSP chip that performs numerically all required operations. If we refer back to the landscape of optical interconnects of Fig. 2–3, for chip-to-chip and On-chip distances, whether electrons or photons are used to convey information, it is very difficult to foresee and justify the use of a digital signal processing chip and a DAC from the extremely short transmission length. A simple eye opening, FFE approach would suffice.

There are multiple possible future venues to study for chip-scale interconnects. First, as signaling rates are increasing to meet the demand, research should be conducted in the development of Feed-Forward Equalizer designs of multiple tap for such high symbol rates. Also, it would be of great interest to investigate FFEs for multilevel formats to follow the current trend of improving the transmission efficiency using higher order signaling. For instance, what would be the better approach to generate and equalize via FFEs a 2^m -level electrical signal using m on-off keying signals? Finally, doing a comparison of the power consumption against system performance of equalization using a N -tap analogous FFE versus using a digital FIR filter and a digital to analog converter would be greatly relevant [165]. One could find if there is a number of taps after which a DSP+DAC approach is more efficient than an analogous FFE, with a possible dependency on the symbol rate.

DAC/ADC-driven Data Center Interconnects

In this Doctoral degree, we have studied multiple different modulation formats and transceiver architectures for self-beating reception over the last year of this doctoral research. We believe that we have not exhausted the research area and that there is still newer, better ways to increase the spectral efficiency of single wavelength modulation with self-beating direct detection. One important aspect of DSP-based Stokes receivers is the power consumption of the DSP functions. Especially, one interesting path to study to lower power is to modify the receiver front end such to do polarization derotation in the optical domain, when the lightwave is still propagating through the substrate, instead of digital, as we did in this thesis. There are well known designs of ‘polarization rotators’ on all material platform (InP, LiNbO₃, SiP). One of the challenges would be to find an efficient, low speed feed-back system to track the wanderings of the state of polarization.

The question to answer is how much power can be allocated to digital signal processing functions is those CMOS integrated all-in-one chips. Minimizing power consumption is always a priority for optical transceivers. Searching for simpler, less computationally demanding processes like our efficient MIMO presented in [6] is one approach to reducing the Watt/Gigabit number of next-generation self-beating transceivers. Another very interesting avenue to lower power is to completely remove the 4 ADCs in the Stokes 4D receiver. In this thesis, we use the 4 ADCs to 1) equalizes the electrical response of the 4 RF tracks of the receiver front-end and to 2) digitally remove the cross-talk amongst the 4 received signals via a linear, bipolar (positive or negative) weighted sum of the 4 tributaries. We adapted the digital weights numerically using the DD-LMS algorithm.

It would be a milestone to be able to remove all 4 ADCs and perform all processing analogically. A proposed receiver front-end could be as follows. The 4 photocurrents of the Stokes receiver could be followed by 4 independent N -tap feed-forward equalizer circuits to equalize the spectral response of the 4 RF tracks. The values of the N taps are set only once to ‘open the eye’. The 4 analog outputs could be summed with different weights, where each ‘weight’ can simply be an adjustable 1-tap FFE that can be controlled analogously. The value of the controlling voltage for each summing tap-weight could be governed via a certain metric to be defined. The metric could be as simple and blind as to minimize the root-mean-square of the output currents. It could also be an analogous version of a DD-LMS adaptation, similar to Proportional-Integral-Derivative controllers.

Doors are always open for newer, better, and innovative designs in the boundless world of Optical Communications

References

- [1] **M. Chagnon**, M. Morsy-Osman, X. Xu, Q. Zhuge, and D. V. Plant, “Blind, fast and SOP independent polarization recovery for square dual polarization–M-QAM formats and optical coherent receivers,” *Opt. Express*, vol. 20, no. 25, pp. 27 847–27 865, Dec 2012.
- [2] **M. Chagnon**, M. Morsy-Osman, Q. Zhuge, X. Xu, and D. V. Plant, “Analysis and experimental demonstration of novel 8PolSK-QPSK modulation at 5 bits/symbol for passive mitigation of nonlinear impairments,” *Opt. Express*, vol. 21, no. 25, pp. 30 204–30 220, Dec 2013.
- [3] **M. Chagnon**, M. Spasojevic, R. Adams, J. Li, D. V. Plant, and L. R. Chen, “Wavelength multicasting at 22-Gbaud 16-QAM in a silicon nanowire using four-wave mixing,” *IEEE Photon. Technol. Lett.*, vol. 27, no. 8, pp. 860–863, April 2015.
- [4] **M. Chagnon**, M. Morsy-Osman, M. Poulin, C. Latrasse, J.-F. Gagné, Y. Painchaud, C. Paquet, S. Lessard, and D. V. Plant, “Experimental study of 112 Gb/s short reach transmission employing PAM formats and SiP intensity modulator at 1.3 μm ,” *Opt. Express*, vol. 22, no. 17, pp. 21 018–21 036, Aug 2014.
- [5] **M. Chagnon**, M. Morsy-Osman, M. Poulin, C. Paquet, S. Lessard, and D. V. Plant, “Experimental parametric study of a silicon photonic modulator enabled 112 Gb/s PAM transmission system with a DAC and ADC,” *J. Lightw. Technol.*, vol. 33, no. 7, pp. 1380–1387, April 2015.
- [6] **M. Chagnon**, M. Morsy-Osman, D. Patel, V. Veerasubramanian, A. Samani, and D. V. Plant, “Digital signal processing for dual-polarization intensity and inter-polarization phase modulation formats using Stokes detection,” *J. Lightw. Technol.*, vol. 34, no. 4, pp. X–X, February 2016.
- [7] **M. Chagnon**, M. Morsy-Osman, D. Patel, V. Veerasubramanian, A. Samani, and D. V. Plant, “1 λ , 6 bits/symbol, 280 and 350 Gb/s direct detection transceiver using intensity modulation, polarization multiplexing, and inter-polarization phase modulation,” in Proc. of *Optical Fiber Communications Conference*, March 2015, paper Th5B.2, pp. 1–3.
- [8] **M. Chagnon**, M. Morsy-Osman, Q. Zhuge, X. Xu, M. Poulin, Y. Painchaud, M. Pelletier, C. Paquet, and D. V. Plant, “Experimental colorless reception of 16 \times DP-16QAM at 28 and 30 Gbaud using a Si-integrated coherent receiver,” in Proc. of *Optical Fiber Communication Conference*, 2013, paper OTh3C.5, pp. 1–3.
- [9] *The zettabyte era: Trends and analysis*, Cisco Systems, Inc., May 2015.
- [10] *WaveLogic 3 Extreme and WaveLogic 3 Nano: New coherent chipsets for a Web-scale World*, Ciena, 2015.
- [11] 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.

- [12] I. Fatadin, D. Ives, and S. J. Savory, "Blind equalization and carrier phase recovery in a 16-QAM optical coherent system," *J. Lightw. Technol.*, vol. 27, no. 15, pp. 3042–3049, Aug 2009.
- [13] P. Winzer, A. H. Gnauck, C. R. Doerr, M. Magarini, and L. L. Buhl, "Spectrally efficient long-haul optical networking using 112 Gb/s polarization-multiplexed 16-QAM," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 547–556, Feb 2010.
- [14] X. Zhou and J. Yu, "200 Gb/s PDM-16QAM generation using a new synthesizing method," in Proc. of *European Conference on Optical Communication*, Sept 2009, paper 10.3.5, pp. 1–2.
- [15] P. Johannisson, H. Wymeersch, M. Sjödin, A. S. Tan, E. Agrell, P. A. Andrekson, and M. Karlsson, "Convergence comparison of the CMA and ICA for blind polarization demultiplexing," *IEEE J. Opt. Commun. Netw.*, vol. 3, no. 6, pp. 493–501, June 2011.
- [16] R. A. Fisher, B. R. Snydam, and D. Yevick, "Optical phase conjugation for time-domain undoing of dispersive self-phase-modulation effects," *Opt. Lett.*, vol. 8, no. 12, pp. 611–613, Dec 1983.
- [17] A. H. Gnauck, R. M. Jopson, and R. M. Derosier, "10 Gb/s 360-km transmission over dispersive fiber using midsystem spectral inversion," *IEEE Photon. Technol. Lett.*, vol. 5, no. 6, pp. 663–666, June 1993.
- [18] S. Watanabe, T. Chikama, G. Ishikawa, T. Terahara, and H. Kuwahara, "Compensation of pulse shape distortion due to chromatic dispersion and Kerr effect by optical phase conjugation," *IEEE Photon. Technol. Lett.*, vol. 5, no. 10, pp. 1241–1243, Oct 1993.
- [19] X. Chen, X. Liu, S. Chandrasekhar, B. Zhu, and R. W. Tkach, "Experimental demonstration of fiber nonlinearity mitigation using digital phase conjugation," in Proc. of *Optical Fiber Communication Conference*, March 2012, pp. 1–3.
- [20] X. Liu, A. R. Chraplyvy, P. J. Winzer, R. W. Tkach, and S. Chandrasekhar, "Phase-conjugated twin waves for communication beyond the Kerr nonlinearity limit," *Nat. Photon.*, vol. 7, pp. 560–568, May 2013.
- [21] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *J. Lightw. Technol.*, vol. 26, no. 20, pp. 3416–3425, Oct 2008.
- [22] W. Yan, Z. Tao, L. Dou, L. Li, S. Oda, T. Tanimura, T. Hoshida, and J. C. Rasmussen, "Low complexity digital perturbation back-propagation," in Proc. of *European Conference on Optical Communication*, Sept 2011, paper Tu.3.A.2, pp. 1–3.
- [23] E. Ip, N. Bai, and T. Wang, "Complexity versus performance tradeoff for fiber nonlinearity compensation using frequency-shaped, multi-subband backpropagation," in Proc. of *Optical Fiber Communication Conference*, March 2011, paper OThF4, pp. 1–3.
- [24] L. Liu, L. Li, Y. Huang, K. Cui, Q. Xiong, F. N. Hauske, C. Xie, and Y. Cai, "Intrachannel nonlinearity compensation by inverse volterra series transfer function," *J. Lightw. Technol.*, vol. 30, no. 3, pp. 310–316, Feb 2012.
- [25] Z. Tao, L. Dou, W. Yan, L. Li, T. Hoshida, and J. C. Rasmussen, "Multiplier-free intrachannel nonlinearity compensating algorithm operating at symbol rate," *J. Lightw. Technol.*, vol. 29, no. 17, pp. 2570–2576, Sept 2011.

- [26] L. Dou, Z. Tao, W. Yan, L. Li, T. Hoshida, and J. C. Rasmussen, "Pre-distortion method for intra-channel nonlinearity compensation with phase-rotated perturbation term," in Proc. of *Optical Fiber Communication Conference*, March 2012, paper OTh3C.2, pp. 1–3.
- [27] T. Oyama, H. Nakashima, S. Oda, T. Yamauchi, Z. Tao, T. Hoshida, and J. C. Rasmussen, "Robust and efficient receiver-side compensation method for intra-channel nonlinear effects," in Proc. of *Optical Fiber Communications Conference*, March 2014, paper Tu3A.3, pp. 1–3.
- [28] T. Oyama, T. Hoshida, H. Nakashima, S. Oda, T. Yamauchi, T. Tanimura, L. Dou, Y. Zhao, Z. Tao, and J. C. Rasmussen, "Nonlinear compensation technologies for future optical communication systems," *Proc. of SPIE*, vol. 9389, pp. 93 890I–93 890I–11, 2015.
- [29] Y. Painchaud, M. Poulin, F. Pelletier, C. Latrasse, J.-F. Gagné, S. Savard, G. Robidoux, M.-J. Picard, S. Paquet, C.-A. Davidson, M. Pelletier, M. Cyr, C. Paquet, M. Guy, M. Morsy-Osman, **M. Chagnon**, and D. V. Plant, "Silicon-based products and solutions," vol. 8988, 2014, pp. 89 880L–89 880L–14.
- [30] C. E. Doerr, "Integrated photonic platforms for telecommunications: InP and Si," *IEICE Transactions on Electronics*, vol. 96, no. 7, pp. 950–957, Jul 2013.
- [31] C. Koos, P. Vorreau, T. Vallaitis, P. Dumon, W. Bogaerts, R. Baets, B. Esembeson, I. Biaggio, T. Michinobu, F. Diederich, W. Freude, and J. Leuthold, "All-optical high-speed signal processing with silicon-organic hybrid slot waveguides," *Nat. Photon.*, vol. 3, pp. 216–219, March 2009.
- [32] A. Biberman, B. G. Lee, A. C. Turner-Foster, M. A. Foster, M. Lipson, A. L. Gaeta, and K. Bergman, "Wavelength multicasting in silicon photonic nanowires," *Opt. Express*, vol. 18, no. 17, pp. 18 047–18 055, Aug 2010.
- [33] M. Pu, H. Hu, H. Ji, M. Galili, L. K. Oxenlowe, K. Yvind, P. Jeppesen, and J. M. Hvam, "Non-degenerate four-wave mixing in a silicon nanowire and its application for one-to-six WDM multicasting," in Proc. of *European Conference on Optical Communication*, Sept 2011, paper Th.11.LeSaleve.2, pp. 1–3.
- [34] D. Patel, S. Ghosh, **M. Chagnon**, A. Samani, V. Veerasubramanian, M. Morsy-Osman, and D. V. Plant, "Design, analysis, and transmission system performance of a 41 GHz silicon photonic modulator," *Opt. Express*, vol. 23, no. 11, pp. 14 263–14 287, Jun 2015.
- [35] A. Samani, **M. Chagnon**, D. Patel, V. Veerasubramanian, S. Ghosh, M. Morsy-Osman, Q. Zhong, and D. V. Plant, "A low-voltage 35-GHz Silicon photonic modulator-enabled 112 Gb/s transmission system," *IEEE Photon. J.*, vol. 7, no. 3, pp. 1–13, June 2015.
- [36] M. Poulin, C. Latrasse, J.-F. Gagné, Y. Painchaud, M. Cyr, C. Paquet, M. Morsy-Osman, **M. Chagnon**, S. Lessard, and D. V. Plant, "107 Gb/s PAM-4 transmission over 10 km using a SiP series push-pull modulator at 1310 nm," in Proc. of *European Conference on Optical Communication*, Sept 2014, paper Mo.4.5.3, pp. 1–3.
- [37] M. Streshinsky, R. Ding, Y. Liu, A. Novack, Y. Yang, Y. Ma, X. Tu, E. K. S. Chee, A. E.-J. Lim, P. G.-Q. Lo, T. Baehr-Jones, and M. Hochberg, "Low power 50 Gb/s silicon traveling wave Mach-Zehnder modulator near 1300 nm," *Opt. Express*, vol. 21, no. 25, pp. 30 350–30 357, Dec 2013.
- [38] P. Dong, L. Chen, and Y.-K. Chen, "High-speed low-voltage single-drive push-pull silicon Mach-Zehnder modulators," *Opt. Express*, vol. 20, no. 6, pp. 6163–6169, Mar 2012.

- [39] W. Bogaerts, P. D. Heyn, T. V. Vaerenbergh, K. D. Vos, S. K. Selvaraja, T. Claes, P. Dumon, P. Bienstman, D. V. Thourhout, and R. Baets, "Silicon microring resonators," *Laser & Photonics Reviews*, vol. 6, no. 1, pp. 47–73, 2012.
- [40] T. Baba, S. Akiyama, M. Imai, N. Hirayama, H. Takahashi, Y. Noguchi, T. Horikawa, and T. Usuki, "50 Gb/s ring-resonator-based silicon modulator," *Opt. Express*, vol. 21, no. 10, pp. 11 869–11 876, May 2013.
- [41] G. Li, A. V. Krishnamoorthy, I. Shubin, J. Yao, Y. Luo, H. Thacker, X. Zheng, K. Raj, and J. E. Cunningham, "Ring resonator modulators in Silicon for interchip photonic links," *IEEE J. Sel. Topics Quantum Electron.*, vol. 19, no. 6, pp. 95–113, Nov 2013.
- [42] S. Akiyama and T. Usuki, "High-speed and efficient silicon modulator based on forward-biased pin diodes," *Frontiers in Physics*, vol. 2, no. 65, 2014.
- [43] Finisar, *Transceivers, Transponders, and Active Optical Cables*, 2015.
- [44] Y. Matsui, T. Pham, T. Sudo, G. Carey, B. Young, and C. Roxlo, "112 Gb/s WDM link using two directly modulated Al-MQW BH DFB lasers at 56 Gb/s," in Proc. of *Optical Fiber Communication Conference*, 2015, p. Th5B.6.
- [45] F. Karinou, C. Prodaniuc, N. Stojanovic, M. Ortsiefer, A. Daly, R. Hohenleitner, B. Kogel, and C. Neumeyr, "Directly PAM-4 modulated 1530-nm VCSEL enabling 56 Gb/s/ λ data-center interconnects," *IEEE Photon. Technol. Lett.*, vol. 27, no. 17, pp. 1872–1875, Sept 2015.
- [46] T. Tanaka, M. Nishihara, T. Takahara, W. Yan, L. Li, Z. Tao, M. Matsuda, K. Takabayashi, and J. Rasmussen, "Experimental demonstration of 448 Gbps+ DMT transmission over 30-km SMF," in Proc. of *Optical Fiber Communication Conference*, 2014, paper M2I.5.
- [47] A. S. Karar and J. C. Cartledge, "Generation and detection of a 112 Gb/s dual polarization signal using a directly modulated laser and half-cycle 16-QAM Nyquist-subcarrier-modulation," in Proc. of *European Conference on Optical Communication*, 2012, paper Th.3.A.4.
- [48] M. I. Olmedo, Z. Tianjian, J. B. Jensen, Z. Qiwen, X. Xiaogeng, and I. T. Monroy, "Towards 400GBASE 4-lane solution using direct detection of MultiCAP signal in 14 GHz bandwidth per lane," in Proc. of *Optical Fiber Communication Conference*, March 2013, paper PDP5C.10, pp. 1–3.
- [49] D. Patel, A. Samani, V. Veerasubramanian, S. Ghosh, and D. V. Plant, "Silicon photonic segmented modulator-based electro-optic DAC for 100 Gb/s PAM-4 generation," *IEEE Photon. Technol. Lett.*, vol. 27, no. 23, pp. 2433–2436, Dec 2015.
- [50] V. Alwayn, *Fiber-Optic Technologies*, Cisco Press, 2004.
- [51] G. P. Agrawal, *Fiber-optic communication systems*, 3rd ed. Wiley-Interscience, 2002.
- [52] T. H. Maiman, "Stimulated optical radiation in ruby," *Nature*, vol. 187, pp. 493–494, Aug 1960.
- [53] N. S. Bergano, J. Aspell, C. R. Davidson, P. R. Trischitta, B. M. Nyman, and F. W. Kerfoot, "A 9000 km 5 Gb/s and 21,000 km 2.4 Gb/s Feasibility Demonstration of Transoceanic EDFA Systems Using a Circulating Loop," in Proc. of *Optical Fiber Communication Conference*, 1991, paper PD13.
- [54] G. Mahlknecht, *Greg's Cable Map*, <http://www.cablemap.info/>.

- [55] D.-S. Ly-Gagnon, K. Katoh, and K. Kikuchi, "Unrepeated 210-km transmission with coherent detection and digital signal processing of 20 Gb/s QPSK signal," in Proc. of *Optical Fiber Communication Conference*, 2005, p. OTuL4.
- [56] S. Tsukamoto, D.-S. Ly-Gagnon, K. Katoh, and K. Kikuchi, "Coherent demodulation of 40 Gbit/s polarization-multiplexed QPSK signals with 16-GHz spacing after 200 km transmission," in Proc. of *Optical Fiber Communication Conference*, 2005, p. PDP29.
- [57] *Breaking the physical barriers with electronic Dynamically Compensating Optics (eDCO)*, Nortel Networks, 2006.
- [58] Fujitsu, *Fujitsu Factsheet LEIA 55-65 GSa/s CMOS 8-bit DAC*, 2011.
- [59] C. Kachris, K. Bergman, and I. Tomkos, *Optical interconnects for future data center networks*, 1st ed. Springer-Verlag, 2013.
- [60] A. Liu, R. Jones, L. Liao, D. Samara-Rubio, D. Rubin, O. Cohen, R. Nicolaescu, and M. Paniccia, "A high-speed silicon optical modulator based on a metal-oxide-semiconductor capacitor," *Nature*, Feb. 2004.
- [61] *Cisco Global Cloud Index: Forecast and Methodology, 2014-2019*, Cisco Press, 2015.
- [62] S. J. Savory, "Digital coherent optical receivers: Algorithms and subsystems," *IEEE J. Sel. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1164–1179, Sept 2010.
- [63] L. Liu, Z. Tao, W. Yan, S. Oda, T. Hoshida, and J. C. Rasmussen, "Initial tap setup of constant modulus algorithm for polarization de-multiplexing in optical coherent receivers," in Proc. of *Optical Fiber Communication Conference*, March 2009, paper OMT2, pp. 1–3.
- [64] A. Vgenis, C. S. Petrou, C. B. Papadias, I. Roudas, and L. Raptis, "Nonsingular constant modulus equalizer for PDM-QPSK coherent optical receivers," *IEEE Photon. Technol. Lett.*, vol. 22, no. 1, pp. 45–47, Jan 2010.
- [65] R. Maher, D. S. Millar, S. J. Savory, and B. C. Thomsen, "Widely tunable burst mode digital coherent receiver with fast reconfiguration time for 112 Gb/s DP-QPSK WDM networks," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3924–3930, Dec 2012.
- [66] F. Vacondio, O. Rival, Y. Pointurier, C. Simonneau, L. Lorcy, J.-C. Antona, and S. Bigo, "Coherent receiver enabling data rate adaptive optical packet networks," in Proc. of *European Conference on Optical Communication*, Sept 2011, paper Mo.2.A.4, pp. 1–3.
- [67] J. E. Simsarian, J. Gripp, A. H. Gnauck, G. Raybon, and P. J. Winzer, "Fast-tuning 224 Gb/s intradyne receiver for optical packet networks," in Proc. of *Optical Fiber Communication Conference*, March 2010, paper PDPB5, pp. 1–3.
- [68] M. Morsy-Osman, **M. Chagnon**, Q. Zhuge, X. Xu, M. E. M. Pasandi, Z. A. El-Sahn, and D. V. Plant, "Training symbol based channel estimation for ultrafast polarization demultiplexing in coherent single-carrier transmission systems with M-QAM constellations," in Proc. of *European Conference on Optical Communication*, Sept 2012, paper Mo.1.A.4, pp. 1–3.
- [69] G. P. Agrawal, *Lightwave technology: Telecommunication systems*, 1st ed. Wiley-Interscience, 2005.
- [70] J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," *Proceedings of the National Academy of Sciences*, vol. 97, no. 9, pp. 4541–4550, 2000.

- [71] S. J. Savory and F. P. Payne, "Pulse propagation in fibers with polarization-mode dispersion," *J. Lightw. Technol.*, vol. 19, no. 3, pp. 350–357, Mar 2001.
- [72] D. Marcuse, C. R. Menyuk, and P. K. A. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightw. Technol.*, vol. 15, no. 9, pp. 1735–1746, Sep 1997.
- [73] C. D. Poole and J. Nagel, *Polarization effects in lightwave systems*, 1st ed. Academic Press, 1997, vol. A, Ch. 6.
- [74] 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.
- [75] I. T. Jolliffe, *Principal component analysis*, 1st ed. Springer, 2002.
- [76] D.-S. Ly-Gagnon, S. Tsukamoto, K. Katoh, and K. Kikuchi, "Coherent detection of optical quadrature phase-shift keying signals with carrier phase estimation," *J. Lightw. Technol.*, vol. 24, no. 1, pp. 12–21, Jan 2006.
- [77] 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.
- [78] O. K. Smith, "Eigenvalues of a symmetric 3×3 Matrix," *Communications of the ACM*, vol. 4, no. 4, p. 168, Apr 1961.
- [79] V. N. Rozental, T. F. Portela, D. V. Souto, H. B. Ferreira, and D. A. A. Mello, "Experimental analysis of singularity-avoidance techniques for CMA equalization in DP-QPSK 112 Gb/s optical systems," *Opt. Express*, vol. 19, no. 19, pp. 18 655–18 664, Sep 2011.
- [80] P. Johannisson, H. Wymeersch, M. Sjodin, A. S. Tan, E. Agrell, P. Andrekson, and M. Karlsson, "Convergence comparison of CMA and ICA for blind polarization demultiplexing of QPSK and 16-QAM signals," in Proc. of *European Conference on Optical Communication*, Sept 2010, paper Th.9.A.3, pp. 1–3.
- [81] J. Yang, J.-J. Werner, and G. A. Dumont, "The multimodulus blind equalization and its generalized algorithms," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 5, pp. 997–1015, Jun 2002.
- [82] X. Zhou, L. E. Nelson, R. Isaac, P. Magill, B. Zhu, D. W. Peckham, P. Borel, and K. Carlson, "1200 km transmission of 50 GHz spaced, 5×504 Gb/s PDM-32-64 hybrid QAM using electrical and optical spectral shaping," in Proc. of *Optical Fiber Communication Conference*, March 2012, paper OM2A.2, pp. 1–3.
- [83] K. Roberts, S. H. Foo, M. Hubbard, A. Sinclair, J. Gaudette, and C. Laperle, "High Capacity Transport - 100G and Beyond," *Journal of Lightwave Technology*, vol. 33, no. 3, pp. 563–578, Feb 2015.
- [84] M. J. Ready and R. P. Gooch, "Blind equalization based on radius directed adaptation," in Proc. of *International Conference on Acoustics, Speech, and Signal Processing*, Apr 1990, paper D11.16, pp. 1699–1702.
- [85] G. P. Agrawal, *Nonlinear fiber optics*, 3rd ed. Academic Press, 2001.

- [86] D. S. Millar, D. Lavery, S. Makovejs, C. Behrens, B. C. Thomsen, P. Bayvel, and S. J. Savory, "Generation and long-haul transmission of polarization-switched QPSK at 42.9 Gb/s," *Opt. Express*, vol. 19, no. 10, pp. 9296–9302, May 2011.
- [87] E. Agrell and M. Karlsson, "Power-efficient modulation formats in coherent transmission systems," *J. Lightw. Technol.*, vol. 27, no. 22, pp. 5115–5126, Nov 2009.
- [88] Z. Chen, L. Yan, W. Pan, B. Luo, Y. Guo, H. Jiang, A. Yi, Y. Sun, and X. Wu, "Transmission of multi-polarization-multiplexed signals: another freedom to explore?" *Opt. Express*, vol. 21, no. 9, pp. 11 590–11 605, May 2013.
- [89] I. B. Djordjevic, A. Z. Jovanovic, M. Cvijetic, and Z. Peric, "Multidimensional vector quantization-based signal constellation design enabling beyond 1 Pb/s serial optical transport networks," *IEEE Photon. J.*, vol. 5, no. 4, pp. 7 901 312–7 901 312, Aug 2013.
- [90] J. G. Proakis, *Digital communications*, 4th ed. McGraw-Hill Science, 2000.
- [91] S. Benedetto, R. Gaudino, and P. Poggiolini, "Direct detection of optical digital transmission based on polarization shift keying modulation," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 3, pp. 531–542, Apr 1995.
- [92] S. Benedetto and P. Poggiolini, "Theory of polarization shift keying modulation," *IEEE Trans. Commun.*, vol. 40, no. 4, pp. 708–721, Apr 1992.
- [93] E. Ip and J. M. Kahn, "Carrier synchronization for 3- and 4-bit-per-symbol optical transmission," *J. Lightw. Technol.*, vol. 23, no. 12, pp. 4110–4124, Dec 2005.
- [94] M. Winter, C.-A. Bunge, D. Setti, and K. Petermann, "A statistical treatment of cross-polarization modulation in DWDM systems," *J. Lightw. Technol.*, vol. 27, no. 17, pp. 3739–3751, Sept 2009.
- [95] P. K. A. Wai and C. R. Menyuk, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence," *J. Lightw. Technol.*, vol. 14, no. 2, pp. 148–157, Feb 1996.
- [96] J. H. Lee, K. J. Park, C. H. Kim, and Y. C. Chung, "Effects of nonlinear crosstalk in optical PMD compensation," *IEEE Photon. Technol. Lett.*, vol. 14, no. 8, pp. 1082–1084, Aug 2002.
- [97] M. Winter, "A statistical treatment of cross-polarization modulation in DWDM systems & its application," Ph.D. Dissertation, Technical University of Berlin, Berlin, 2010.
- [98] M. Karlsson and H. Sunnerud, "Effects of nonlinearities on PMD-induced system impairments," *J. Lightw. Technol.*, vol. 24, no. 11, pp. 4127–4137, Nov 2006.
- [99] K.-P. Ho and J. M. Kahn, "Electronic compensation technique to mitigate nonlinear phase noise," *J. Lightw. Technol.*, vol. 22, no. 3, pp. 779–783, March 2004.
- [100] S. Kumar, "Effect of dispersion on nonlinear phase noise in optical transmission systems," *Opt. Lett.*, vol. 30, no. 24, pp. 3278–3280, Dec 2005.
- [101] K.-P. Ho, "Asymptotic probability density of nonlinear phase noise," *Opt. Lett.*, vol. 28, no. 15, pp. 1350–1352, Aug 2003.
- [102] M. Nazarathy, W. V. Sorin, D. M. Baney, and S. A. Newton, "Spectral analysis of optical mixing measurements," *J. Lightw. Technol.*, vol. 7, no. 7, pp. 1083–1096, Jul 1989.

- [103] M. Shtaif, "Analytical description of cross-phase modulation in dispersive optical fibers," *Opt. Lett.*, vol. 23, no. 15, pp. 1191–1193, Aug 1998.
- [104] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, "Modeling of the impact of nonlinear propagation effects in uncompensated optical coherent transmission links," *J. Lightw. Technol.*, vol. 30, no. 10, pp. 1524–1539, May 2012.
- [105] *Forward error correction for high bit-rate DWDM submarine systems*, International Telecommunication Union - Telecommunication Standardization Sector (ITU-T) G.975.1, 2004.
- [106] D. Lavery, C. Behrens, and S. J. Savory, "A comparison of modulation formats for passive optical networks," *Opt. Express*, vol. 19, no. 26, pp. B836–B841, Dec 2011.
- [107] D. Chang, F. Yu, Z. Xiao, Y. Li, N. Stojanovic, C. Xie, X. Shi, X. Xu, and Q. Xiong, "FPGA verification of a single QC-LDPC code for 100 Gb/s optical systems without error floor down to BER of 10^{-15} ," in Proc. of *Optical Fiber Communication Conference*, March 2011, paper OTuN2, pp. 1–3.
- [108] S. Song, "Higher-order four-wave mixing and its effect in WDM systems," *Opt. Express*, vol. 7, no. 4, pp. 166–171, Aug 2000.
- [109] C. S. Bres, A. O. J. Wiberg, B. P. Kuo, E. Myslivets, and S. Radic, "320 Gb/s RZ-DPSK data multicasting in self seeded parametric mixer," in Proc. of *Optical Fiber Communication Conference*, March 2011, paper OThC7, pp. 1–3.
- [110] B. P.-P. Kuo, E. Myslivets, N. Alic, and S. Radic, "Wavelength multicasting via frequency comb generation in a bandwidth-enhanced fiber optical parametric mixer," *J. Lightw. Technol.*, vol. 29, no. 23, pp. 3515–3522, Dec 2011.
- [111] G. Contestabile, A. Maruta, S. Sekiguchi, K. Morito, M. Sugawara, and K. ichi Kitayama, "All-optical wavelength multicasting in a QD-SOA," *IEEE J. Quantum Electron.*, vol. 47, no. 4, pp. 541–547, April 2011.
- [112] M. P. Fok and C. Shu, "Performance investigation of one-to-six wavelength multicasting of ASK-DPSK signal in a highly nonlinear bismuth oxide fiber," *J. Lightw. Technol.*, vol. 27, no. 15, pp. 2953–2957, Aug 2009.
- [113] A. Malacarne, G. Meloni, G. Berrettini, L. Poti, and A. Bogoni, "Optical multicasting of a 224 Gb/s PM-16QAM signal in a periodically-poled Lithium Niobate waveguide," in Proc. of *Optical Fiber Communication Conference*, March 2013, paper OM2G.2, pp. 1–3.
- [114] M. A. Ettabib, K. Hammani, F. Parmigiani, L. Jones, A. Kapsalis, A. Bogris, D. Syvridis, M. Brun, P. Labeye, S. Nicoletti, and P. Petropoulos, "FWM-based wavelength conversion of 40 Gbaud PSK signals in a silicon germanium waveguide," *Opt. Express*, vol. 21, no. 14, pp. 16 683–16 689, Jul 2013.
- [115] H. Hu, H. Ji, M. Galili, M. Pu, C. Peucheret, H. C. H. Mulvad, K. Yvind, J. M. Hvam, P. Jeppesen, and L. K. Oxenløwe, "Ultra-high-speed wavelength conversion in a silicon photonic chip," *Opt. Express*, vol. 19, no. 21, pp. 19 886–19 894, Oct 2011.
- [116] M. Pu, H. Hu, H. Ji, M. Galili, L. K. Oxenløwe, P. Jeppesen, J. M. Hvam, and K. Yvind, "One-to-six WDM multicasting of DPSK signals based on dual-pump four-wave mixing in a silicon waveguide," *Opt. Express*, vol. 19, no. 24, pp. 24 448–24 453, Nov 2011.

- [117] R. Adams, M. Spasojevic, **M. Chagnon**, M. Malekiha, J. Li, D. V. Plant, and L. R. Chen, “Wavelength conversion of 28 GBaud 16-QAM signals based on four-wave mixing in a silicon nanowire,” *Opt. Express*, vol. 22, no. 4, pp. 4083–4090, Feb 2014.
- [118] N. Shibata, R. P. Braun, and R. G. Waarts, “Phase-mismatch dependence of efficiency of wave generation through four-wave mixing in a single-mode optical fiber,” *IEEE J. Quantum Electron.*, vol. 23, no. 7, pp. 1205–1210, Jul 1987.
- [119] Y. Painchaud, M. Pelletier, M. Poulin, F. Pelletier, C. Latrasse, G. Robidoux, S. Savard, J.-F. Gagné, V. Trudel, M.-J. Picard, P. Poulin, P. Sirois, F. D’Amours, D. Asselin, S. Paquet, C. Paquet, M. Cyr, M. Guy, M. Morsy-Osman, Q. Zhuge, X. Xu, **M. Chagnon**, and D. V. Plant, “Ultra-compact coherent receiver based on hybrid integration on silicon,” in Proc. of *Optical Fiber Communication Conference*, March 2013, paper OM2J.2, pp. 1–3.
- [120] M. Kuschnerov, F. N. Hauske, K. Piyawanno, B. Spinnler, M. S. Alfiad, A. Napoli, and B. Lankl, “DSP for coherent single-carrier Receivers,” *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3614–3622, Aug 2009.
- [121] M. Scholten, T. Coe, and J. Dillard, “Continuously-interleaved BCH (CI-BCH) FEC delivers best in class NECG for 40G and 100G metro applications,” in Proc. of *Optical Fiber Communication Conference*, March 2010, paper NTuB3.
- [122] L. Chrostowski and M. Hochberg, *Silicon photonics design: From devices to systems*. Cambridge University Press, May 2015, Ch. 1 - Fabless silicon photonics.
- [123] A. Sugita, A. Kaneko, K. Okamoto, M. Itoh, A. Himeno, and Y. Ohmori, “Very low insertion loss arrayed-waveguide grating with vertically tapered waveguides,” *IEEE Photon. Technol. Lett.*, vol. 12, no. 9, pp. 1180–1182, Sept 2000.
- [124] *IEEE Standard for Information Technology—Local and metropolitan area networks—Specific requirements— Part 3: CSMA/CD access method and physical layer specifications amendment 4: Media access control parameters, physical layers, and management parameters for 40 Gb/s and 100 Gb/s operation*, IEEE Standard for Information Technology, June 2010.
- [125] *IEEE P802.3bs 400 GbE Task Force*, IEEE Standard for Information Technology, May 2014.
- [126] R. Rodes, J. Estaran, B. Li, M. Muller, J. B. Jensen, T. Gruendl, M. Ortsiefer, C. Neumeyr, J. Rosskopf, K. J. Larsen, M.-C. Amann, and I. T. Monroy, “100 Gb/s single VCSEL data transmission link,” in Proc. of *Optical Fiber Communication Conference*, March 2012, paper PDP5D.10, pp. 1–3.
- [127] M. I. Olmedo, T. Zuo, J. B. Jensen, Q. Zhong, X. Xu, S. Popov, and I. T. Monroy, “Multi-band carrierless amplitude phase modulation for high capacity optical data links,” *J. Lightw. Technol.*, vol. 32, no. 4, pp. 798–804, Feb 2014.
- [128] W. Yan, T. Tanaka, B. Liu, M. Nishihara, L. Li, T. Takahara, Z. Tao, J. C. Rasmussen, and T. Drenski, “100 Gb/s optical IM-DD transmission with 10G-Class devices enabled by 65 GSamples/s CMOS DAC core,” in Proc. of *Optical Fiber Communication Conference*, 2013, paper OM3H.1.
- [129] Y. Kai, M. Nishihara, T. Tanaka, T. Takahara, L. Li, Z. Tao, B. Liu, J. C. Rasmussen, and T. Drenski, “Experimental comparison of pulse amplitude modulation (PAM) and discrete multi-tone (DMT) for short-reach 400 Gbps data communication,” in Proc. of *European Conference on Optical Communication*, Sept 2013, paper Th.1.F.3, pp. 1–3.

- [130] T. Zuo, A. Tatarczak, M. I. Olmedo, J. Estaran, J. B. Jensen, Q. Zhong, X. Xu, and I. T. Monroy, "O-band 400 Gbit/s client side optical transmission link," in Proc. of *Optical Fiber Communications Conference*, March 2014, paper M2E.4, pp. 1–3.
- [131] M. Morsy-Osman, **M. Chagnon**, X. Xu, Q. Zhuge, M. Poulin, Y. Painchaud, M. Pelletier, C. Paquet, and D. V. Plant, "Analytical and experimental performance evaluation of an integrated Si-photonic balanced coherent receiver in a colorless scenario," *Opt. Express*, vol. 22, no. 5, pp. 5693–5730, Mar 2014.
- [132] P. de Dobbelaere, A. Narasimha, A. Mekis, B. Welch, C. Bradbury, C. Sohn, D. Song, D. Foltz, D. Guckenberger, G. Masini, J. Schramm, J. White, J. Redman, K. Yokoyama, M. Harrison, M. Peterson, M. Mack, M. Sharp, R. LeBlanc, S. Abdalla, S. Gloeckner, S. Hovey, S. Jackson, S. Sahni, S. Yu, T. Pinguet, and Y. Liang, "Silicon photonics for high data rate optical interconnect," in Proc. of *Optical Interconnects Conference*, May 2012, paper WA2, pp. 113–114.
- [133] D. Liang and J. E. Bowers, "Photonic integration: Si or InP substrates?" *Electronics Lett.*, vol. 45, no. 12, pp. 578–581, June 2009.
- [134] L. Chen, C. R. Doerr, P. Dong, and Y. kai Chen, "Monolithic silicon chip with 10 modulator channels at 25 Gbps and 100-GHz spacing," *Opt. Express*, vol. 19, no. 26, pp. B946–B951, Dec 2011.
- [135] R. G. Walker, "High-speed III-V semiconductor intensity modulators," *IEEE J. Quantum Electron.*, vol. 27, no. 3, pp. 654–667, Mar 1991.
- [136] *Agency for Science, Technology and Research, (A*STAR)*, Institute of Microelectronics (IME), Singapore.
- [137] *CREATE Silicon Electronic-Photonic Integrated Circuits (Si-EPIC)*, Natural Sciences and Engineering Research Council (NSERC), Canada.
- [138] S. Akiyama, H. Itoh, T. Takeuchi, A. Kuramata, and T. Yamamoto, "Wide-wavelength-band (30 nm) 10 Gb/s operation of InP-based Mach-Zehnder modulator with constant driving voltage of 2 Vpp," *IEEE Photon. Technol. Lett.*, vol. 17, no. 7, pp. 1408–1410, July 2005.
- [139] K. Prosyk, A. Ait-Ouali, C. Bornholdt, T. Brast, M. Gruner, M. Hamacher, D. Hoffmann, R. Kaiser, R. Millett, K.-O. Velthaus, and I. Woods, "High performance 40 GHz InP Mach-Zehnder modulator," in Proc. of *Optical Fiber Communication Conference*, March 2012, paper OW4F.7, pp. 1–3.
- [140] Optical Internetworking Forum Workshop, Dallas, TX, October 2008.
- [141] J. Cho, C. Xie, and P. J. Winzer, "Analysis of soft-decision FEC on non-AWGN channels," *Opt. Express*, vol. 20, no. 7, pp. 7915–7928, Mar 2012.
- [142] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-time signal processing*, 2nd ed. Prentice Hall, 1999.
- [143] Y. Wang, E. Serpedin, and P. Ciblat, "An alternative blind feedforward symbol timing estimator using two samples per symbol," *IEEE Trans. Commun.*, vol. 51, no. 9, pp. 1451–1455, Sept 2003.

- [144] M. Morsy-Osman, **M. Chagnon**, Q. Zhuge, X. Xu, and D. V. Plant, “Non-data-aided feed-forward timing recovery for flexible transceivers employing PDM-MQAM modulations,” in Proc. of *Optical Fiber Communications Conference*, March 2014, paper W3B.4, pp. 1–3.
- [145] C. Cole, I. Lyubomirsky, A. Ghiasi, and V. Telang, “Higher-order modulation for client optics,” *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 50–57, March 2013.
- [146] K. Roberts, “1 Tb/s coherent transceiver,” in Proc. of *Advanced Photonics Congress*, 2012, paper SpW2B.5.
- [147] H. A. Haus, *Electromagnetic noise and quantum optical measurements*, 1st ed. Springer, 2000, Ch. 4 - Shot Noise and Thermal Noise.
- [148] R. Shi, H. Guan, A. Novack, M. Streshinsky, A. E.-J. Lim, G.-Q. Lo, T. Baehr-Jones, and M. Hochberg, “High-efficiency grating couplers near 1310 nm fabricated by 248-nm DUV lithography,” *IEEE Photon. Technol. Lett.*, vol. 26, no. 15, pp. 1569–1572, Aug 2014.
- [149] B. S. G. Pillai, B. Sedighi, K. Guan, N. P. Anthapadmanabhan, W. Shieh, K. J. Hinton, and R. S. Tucker, “End-to-end energy modeling and analysis of long-haul coherent transmission systems,” *J. Lightw. Technol.*, vol. 32, no. 18, pp. 3093–3111, Sept 2014.
- [150] C. Laperle and M. O’Sullivan, “Advances in high-speed DACs, ADCs, and DSP for optical coherent transceivers,” *J. Lightw. Technol.*, vol. 32, no. 4, pp. 629–643, Feb 2014.
- [151] Q. Hu, D. Che, Y. Wang, and W. Shieh, “Advanced modulation formats for high-performance short-reach optical interconnects,” *Opt. Express*, vol. 23, no. 3, pp. 3245–3259, Feb 2015.
- [152] K. Zhong, X. Zhou, T. Gui, L. Tao, Y. Gao, W. Chen, J. Man, L. Zeng, A. P. T. Lau, and C. Lu, “Experimental study of PAM-4, CAP-16, and DMT for 100 Gb/s short reach optical transmission systems,” *Opt. Express*, vol. 23, no. 2, pp. 1176–1189, Jan 2015.
- [153] K. Zhong, X. Zhou, Y. Gao, W. Chen, J. Man, L. Zeng, A. P. T. Lau, and C. Lu, “140-Gb/s 20-km transmission of PAM-4 signal at 1.3 μ m for short reach communications,” *IEEE Photon. Technol. Lett.*, vol. 27, no. 16, pp. 1757–1760, Aug 2015.
- [154] D. Che, A. Li, X. Chen, Q. Hu, Y. Wang, and W. Shieh, “160 Gb/s Stokes vector direct detection for short reach optical communication,” in Proc. of *Optical Fiber Communication Conference*, 2014, paper Th5C.7.
- [155] *IEEE Standard for Ethernet - Amendment 3: Physical layer specifications and management parameters for 40 Gb/s and 100 Gb/s operation over fiber optic cables*, March 2015.
- [156] M. Morsy-Osman, **M. Chagnon**, M. Poulin, S. Lessard, and D. V. Plant, “224 Gb/s 10-km transmission of PDM PAM-4 at 1.3 μ m using a single intensity-modulated laser and a direct-detection MIMO DSP-based receiver,” *J. Lightw. Technol.*, vol. 33, no. 7, pp. 1417–1424, April 2015.
- [157] J. N. Damask, *Polarization optics in telecommunications*. Springer, 2005.
- [158] *Techniques to achieve oscilloscope bandwidths of greater than 16 GHz*, Agilent Technologies Application Note, 2010.
- [159] S. Randel, S. Corteselli, P. J. Winzer, A. Adamiecki, A. Gnauck, S. Chandrasekhar, A. Bielik, L. Altenhain, T. Ellermeyer, U. Dümmler, H. Langenhagen, and R. Schmid, “Generation of a digitally shaped 55-GBd 64-QAM single-carrier signal using novel high-speed DACs,” in Proc. of *Optical Fiber Communication Conference*, 2014, paper M2A.3.

- [160] M. Li, T. Kwasniewski, S. Wang, and Y. Tao, "FIR filter optimization as preemphasis of high-speed backplane data transmission," in Proc. of *International Conference on Communications, Circuits and Systems*, vol. 2, June 2004, pp. 773–776.
- [161] *Understanding the pre-emphasis and linear equalization features in Stratix IV GX devices*, ALTERA Application Note, 2010.
- [162] D. Rafique, A. Napoli, S. Calabro, and B. Spinnler, "Digital preemphasis in optical communication systems: On the DAC requirements for terabit transmission applications," *J. Lightw. Technol.*, vol. 32, no. 19, pp. 3247–3256, Oct 2014.
- [163] A. D. Shiner, M. Reimer, A. Borowiec, S. O. Gharan, J. Gaudette, P. Mehta, D. Charlton, K. Roberts, and M. O'Sullivan, "Demonstration of an 8-dimensional modulation format with reduced inter-channel nonlinearities in a polarization multiplexed coherent system," *Opt. Express*, vol. 22, no. 17, pp. 20 366–20 374, Aug 2014.
- [164] A. V. Rylyakov, C. L. Schow, B. G. Lee, F. E. Doany, C. W. Baks, and J. A. Kash, "Transmitter predistortion for simultaneous improvements in bit rate, sensitivity, jitter, and power efficiency in 20 Gb/s CMOS-driven VCSEL links," *J. Lightw. Technol.*, vol. 30, no. 4, pp. 399–405, Feb 2012.
- [165] F. Doany, "Power-efficient, high-bandwidth optical interconnects for high performance computing," in Proc. of *High-Performance Interconnects Conference*, Aug 2012, paper Keynote III.