# ADVANCED MODULATION FORMATS FOR OPTICAL INTERCONNECTS

by

Alexander Michael Brisson

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Optics and Photonics

MONTANA STATE UNIVERSITY
Bozeman, Montana

May 2025

©COPYRIGHT

by

Alexander Michael Brisson

2025

All Rights Reserved

# ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Dr. Ioannis Roudas, for his invaluable mentorship. His insights, support, and readiness to offer guidance whenever needed were instrumental in the completion of this work. I am also grateful to Dr. Eric Fink for his assistance in setting up the SVM simplex optimization experiment and for his willingness to answer my many questions on algorithms, optimized constellations, and other fun mathematical topics. Additionally, I extend my appreciation to Dr. Fabio Barbosa and Dr. Filipe Ferreira for their productive collaboration, insightful discussions, and for providing the numerical framework that enabled the advanced studies of Mode Vector Modulation presented in this thesis. I would also like to acknowledge the support of the Air Force Research Laboratory (AFRL) through grant #FA8750-20-1-1004, which made this research possible.

Finally, I offer my heartfelt thanks to my parents, Carl and Lynn, for their unwavering love and support throughout my graduate studies. Your encouragement has been, and always will be, one of my greatest assets.

TABLE OF CONTENTS

1. INTRODUCTION ..... 1

Overview of Intra-Data Center Networks ..... 1
Modulation Format Considerations for Intra-DCNs ..... 2
Scaling Pulse Amplitude Modulation ..... 4
Coherent Quadrature Amplitude Modulation ..... 5
Motivation for Advanced Modulation Formats ..... 6
Thesis Organization ..... 6

2. THEORY ..... 7

Overview ..... 7
Jones and Stokes Representations of Polarization ..... 7
Mode Vector Modulation ..... 10
The MVM Concept ..... 10
Optimized Constellations of SVM/MVM ..... 12
Analytical Performance of MVM ..... 14

3. EXPERIMENTAL DESCRIPTION ..... 18

Overview ..... 18
Experimental Devices ..... 18
Experimental Setup ..... 20
Optical Modulation Analysis and MATLAB Interaction ..... 22
Transmitter Calibrations ..... 23
DC Biasing of the Mach-Zehnder Interferometers ..... 23
Receiver Calibrations ..... 25
90-degree Optical Hybrid Imbalance ..... 25
Channel Skew ..... 26

4. EXPERIMENTAL OPTIMIZATION OF THE SVM SIMPLEX CONSTELLATION ..... 28

Overview ..... 28
SVM/MVM Generation using Quadrature Modulators ..... 29
Conversions between Jones and Stokes Spaces ..... 29
Relating MZI Transfer Functions and Modulator Driving Voltages ..... 32
Transmitter Characterization using Tone FM ..... 34
Evaluating the AWG Waveforms ..... 42
Experimental Generation of the SVM Simplex Constellation ..... 46
Matched Filter ..... 47

TABLE OF CONTENTS – CONTINUED

Jones to Stokes Conversion ..... 48
Symbol Sequence Resolution ..... 48
Pseudo-Inverse ..... 51
Bit Error Rate Evaluation and Error Vector Magnitude ..... 53
SVM Simplex Optimization using the Steepest Ascent Procedure ..... 56
Numerical Simulation ..... 62
Experimental Implementation ..... 65
Future Work ..... 70

5. IMPACT OF MODAL DISPERSION AND INTERMODAL CROSSTALK ON MVM ..... 72

Overview ..... 72
Semi-Analytical Model ..... 73
Description of the Simulation Parameters ..... 75
Simulation Diagram and Considerations ..... 75
Simulation Parameters for 3-mode Propagation ..... 78
Numerical Simulation of (6,64)-MVM ..... 79
MVM Transmitter ..... 79
Crosstalk Accumulation and Modal Dispersion ..... 81
Noise Loading and the Ideal MVM Receiver ..... 82
Polarization Derotation using the Least-Mean-Square Algorithm ..... 84
Simulation Results ..... 89
Future Work ..... 94

6. CONCLUSION ..... 95

REFERENCES CITED ..... 97

APPENDICES ..... 102

APPENDIX A : Acronyms ..... 103
APPENDIX B : MATLAB Code ..... 106

# LIST OF TABLES

Table Page

1. Table 1 Scenarios for achieving an aggregate bit rate of 1.6 Tbps for future data centers. ..... 16

2. Table 2 Device models. ..... 21

3. Table 3 Modulator driving voltages for fixed $\beta = \frac{\pi}{2}$. ..... 42

4. Table 4 Normalized voltage mapping between the $\gamma$ coefficients used to drive each MZI and the SVM simplex Jones symbol vectors. ..... 43

5. Table 5 SVM simplex bit-to-symbol mapping. ..... 54

6. Table 6 Simulation parameters. ..... 78

# LIST OF FIGURES

Figure Page

1. Figure 1 Google's DCN optical technology evolution [41] ..... 3

2. Figure 2 Propagating TEM wave with $x$ and $y$ polarization components. ..... 8

3. Figure 3 Arbitrary unit Stokes vector $\hat{s}$ plotted on the Poincaré sphere. ..... 9

4. Figure 4 The SVM concept. ..... 10

5. Figure 5 (a) The MVM concept; (b) multicore fiber (MCF) and few-mode fiber (FMF) with $k$ single-mode cores or modes. In reality, the modes of the FMF overlap; the depiction is purely qualitative. ..... 11

6. Figure 6 Optimization of 8-SVM. (a) Reduction of the potential energy $\xi$ after $\approx$ 900 iterations; (b) starting (left) and optimized (right) constellations of 8-SVM. The constellation on the right represents the square antiprism (twisted cubic constellation) ..... 13

7. Figure 7 MVM spectral efficiency per spatial degree of freedom (SDoF) vs required bit SNR per SDoF to achieve a bit error probability of $10^{-4}$. The blue, red, and black curves represent dimensionality $N = 2$, 4, and 8, respectively. Each data point represents the number of points $M$ in the constellation [19] ..... 14

8. Figure 8 MVM simplex spectral efficiency per spatial degree of freedom (SDoF) vs required bit SNR per SDoF to achieve a bit error probability $= 10^{-4}$. Other modulation format results, including $M$-ary PAM and QAM for comparison [19]. ..... 15

9. Figure 9 Experimental devices placed on the server rack inside the optical communications lab at Montana State University. ..... 19

10. Figure 10 Back-to-back experimental block diagram. ..... 20

11. Figure 11 OMA and MATLAB interaction. ..... 23

LIST OF FIGURES – CONTINUED

Figure Page

12. Figure 12 (a) Quantifi Photonics IQTX block diagram; (b) one of the two MZMs inside the polarization diverse IQ modulator. ..... 23

13. Figure 13 DC bias sweeps performed by the ABC before external modulation for (a) $x$-polarization I, Q, and phase; (b) $y$-polarization I, Q, and phase. ..... 25

14. Figure 14 (a) The SVM simplex Jones vectors plotted in the complex plane; (b) the formed regular tetrahedron in Stokes space after using the quadratic form with the Pauli spin matrices. ..... 31

15. Figure 15 The MZI intensity (dotted) and amplitude (solid) transfer functions. ..... 32

16. Figure 16 Transmitter characterization experiment block diagram. ..... 35

17. Figure 17 (a) Computed power spectrum with the corresponding odd-order Bessel functions occurring at multiples of $f_o \pm (2n+1)f_m$; (b) Normalized $\beta/\pi$ plotted as a function of $r$. ..... 37

18. Figure 18 Qualitative BPSK scheme with $R_s = 2f_m$. ..... 38

19. Figure 19 Transmitter characterization DSP. ..... 39

20. Figure 20 Experimental constellations of BPSK processing through various DSP functions. Received symbols plotted (a) after clock recovery and symbol downsampling; (b) after performing SOP estimation; (c) after removing the IF offset and performing phase estimation. ..... 40

21. Figure 21 Final steps for recovering the modulation index $\beta$. (a) Interpolated BPSK signal representing the received sinusoid; (b) RF spectrum after taking an FFT of the signal. ..... 40

22. Figure 22 Normalized modulation index $\beta/\pi$ plotted against AWG peak-to-peak voltage setting $2v_0$ for each MZI in the transmitter. ..... 41

23. Figure 23 Example of a de Bruijn sequence with $k = n = 2$. ..... 44

LIST OF FIGURES – CONTINUED

Figure Page

24. Figure 24 Plots of waveforms to be uploaded to the AWG for generation of the simplex SVM constellation. Figures include waveform plots up to 2000 samples for modulators controlling the (a) $x$-pol in-phase; (b) $x$-pol quadrature; (c) $y$-pol in-phase; (d) $y$-pol quadrature. ..... 45

25. Figure 25 The SVM simplex experimental block diagram. ..... 46

26. Figure 26 DSP functions for the SVM simplex including hardware impairment compensation, matched filter, clock recovery, Jones to Stokes conversion, symbol sequence resolution, pseudo-inverse, and evaluations of the BER and error vector magnitude (EVM). ..... 47

27. Figure 27 Simplex SVM matched filter implemented on the $x$-pol in-phase channel. ..... 48

28. Figure 28 (a) Received Stokes vectors rotated away from the nominal Stokes symbols forming the regular tetrahedron; (b) tetrahedron formed by the cluster averages of received vectors with unknown symbol vertices. ..... 49

29. Figure 29 (a) Clustered groups of received Stokes vectors with labels {a, b, c, d}; (b) permutations of {a, b, c, d} corresponding to symbol vertices. ..... 50

30. Figure 30 (a) Peak correlations of $\rho_n$ plotted against $n$; (b) Cross-correlation between the permuted sequence $\mathbf{y_{13}}$ and the known repeated transmitted sequence $\mathbf{x}'$. ..... 51

31. Figure 31 (a) Derotated SVM simplex constellation; (b) histogram of inner products of received Stokes vectors. ..... 52

32. Figure 32 BER versus average received optical power for the initial generation of the SVM simplex constellation. ..... 55

33. Figure 33 The SVM simplex plotted in the $yz$-plane with computed EVM $\approx 0.36$. ..... 56

34. Figure 34 Parallelepiped spanned by $\vec{a}, \vec{b}, \vec{c}$. ..... 57

LIST OF FIGURES – CONTINUED

Figure Page

35. Figure 35 Tetrahedron inscribed in the unit Poincaré sphere (not shown). Notation: Position vectors of the four vertices, $\hat{s}_i$, $i = 1, \ldots, 4$ (in green); Three edges $\vec{d_i}$, $i = 1, \ldots, 3$ (in red)..... 58

36. Figure 36 (a) cost function $\xi$ plotted against the iteration index $n$; (b) The formed tetrahedron after 17 iterations. The black points represent the resulting Stokes vectors that maximize the volume and the red points represent successive iterations..... 64

37. Figure 37 Experimental block diagram incorporating the steepest ascent optimization procedure to update MZI scaling coefficients $\gamma$. ..... 66

38. Figure 38 Normalized volume $\hat{\xi}$ plotted against number of iterations $n$. ..... 67

39. Figure 39 The SVM simplex plotted in the $yz$-plane with EVM $\approx 0.16$. ..... 68

40. Figure 40 Figures display experimental results of the SVM simplex after 23 iterations of steepest ascent optimization. (a) Received Stokes vectors plotted on the Poincaré sphere; (b) The histogram of inner products. ..... 69

41. Figure 41 BER vs average receive power showing an $\approx 5.5$ dB improvement in sensitivity after optimization..... 70

42. Figure 42 Future experimental generations of high-cardinality SVM constellations eventually paving the way toward the generation of promising $(N, M)$-MVM constellations. ..... 71

43. Figure 43 MVM simulation block diagram denoting two simulation scenarios. The top fiber represents a standard 10 km FMF link with positive DMD, where the bottom fiber represents propagation through 5 km of FMF with positive DMD and another 5 km of FMF with negative DMD. ..... 76

44. Figure 44 RRC impulse response. ..... 79

LIST OF FIGURES – CONTINUED

Figure Page

45. Figure 45 Simulation transmission waveforms for (6,64)-MVM. Each figure shows the real and imaginary parts of the 6-dimensional transmitted Jones vectors for modes (a) $LP_{01}$ $x$-pol; (b) $LP_{01}$ $y$-pol; (c) $LP_{11a}$ $x$-pol; (d) $LP_{11a}$ $y$-pol; (e) $LP_{11b}$ $x$-pol; (f) $LP_{11b}$ $y$-pol. ..... 80

46. Figure 46 Coupling matrices between mode indices for the FMF supporting 3 spatial modes. Each figure shows the magnitude of each coupling matrix element in dB for the (a) first step in the fiber where $z = 10$ m; (b) the accumulation of coupling matrices where $z = 10^4$ m. ..... 81

47. Figure 47 The real part of the first dimension of $R$, corresponding to the $x$-polarization of the first polarization mode ($LP_{01}$), is shown for the first 1000 samples, assuming a signal-to-noise ratio $\gamma_b = 10$ dB. ..... 83

48. Figure 48 Filtered spectrum of the real part of the $x$-polarization of the first polarization mode ($LP_{01}$). ..... 84

49. Figure 49 Time-domain view of transmitted Stokes symbols partitioned into pilot and payload packets. ..... 85

50. Figure 50 Least-mean squares pilot-aided adaptive channel equalizer. ..... 86

51. Figure 51 Least-mean squares decision-directed adaptive channel equalizer. ..... 87

52. Figure 52 Simulation of the PA+DD LMS equalizer for (6,64)-MVM. ..... 88

53. Figure 53 Simulation results of (6,64)-MVM through a 10 km FMF link assuming a crosstalk parameter XT = -30 dB/km. Each figure depicts the BER vs bit SNR with and without optical fiber DMD compensation. The figures are segmented into the following: (a) no fiber management at $R_s = 56$ GBd; (b) no fiber management at $R_s = 112$ GBd; (c) with fiber management at $R_s = 56$ GBd; (d) with fiber management at $R_s = 112$ GBd. ..... 91

LIST OF FIGURES – CONTINUED

Figure Page

54. Figure 54 Simulation results of (6,64)-MVM through a 10 km FMF link assuming various levels of crosstalk and fixed MD. Each figure depicts the BER vs bit SNR without optical fiber DMD compensation. The figures are segmented into the following: (a) MD = 0.1 ps/km and $R_s$ = 56 GBd; (b) MD = 0.1 ps/km and $R_s$ = 112 GBd; (c) MD = 0.4 ps/km and $R_s$ = 56 GBd; (d) MD = 0.4 ps/km and $R_s$ = 112 GBd. ..... 92

55. Figure 55 Simulation results of (6,64)-MVM through a 10 km FMF link assuming various levels of crosstalk and fixed MD. Each figure depicts the BER vs bit SNR without optical fiber DMD compensation. The figures are segmented into the following: (a) MD = 12 ps/km and $R_s$ = 56 GBd; (b) MD = 12 ps/km and $R_s$ = 112 GBd; (c) MD = 24 ps/km and $R_s$ = 56 GBd; (d) MD = 24 ps/km and $R_s$ = 112 GBd. ..... 93

# ABSTRACT

The rapid growth of artificial intelligence, machine learning, cloud computing, and the Internet of Things has dramatically increased global internet traffic, driving an unprecedented demand for computing resources. As data centers scale to accommodate this surge, their energy consumption has become a pressing concern, necessitating the development of efficiency solutions across various aspects of their design, construction, and implementation. Geometrically optimized constellations of Stokes Vector Modulation (SVM) and Mode Vector Modulation (MVM) are encouraging energy-efficient advanced modulation formats suitable for data centers, offering trade-offs between spectral efficiency and energy consumption under various optical link constraints.

At present, being largely theoretical, research into practical realizations of these constellations, due in part to their challenging symmetries, combined with their performance under real-world transmission effects such as modal dispersion and intermodal crosstalk, has remained limited. This thesis, which aims to address these questions, finds that with adequate optimization frameworks such as steepest ascent, fine-tuning of Mach-Zehnder Modulator driving voltages can significantly enhance the constellation symmetry of SVM and MVM, leading to the first experimental realization of the theoretically ideal four-point SVM simplex en route toward the generation of SVM/MVM constellations exhibiting higher cardinalities. Additionally, numerical simulations of (6,64)-MVM under realistic transmission conditions reveal that, despite severe impairment, when combined with adequate optical compensation techniques and digital signal processing, MVM exhibits substantial gains in terms of receiver sensitivity compared to conventional 4-PAM, achieving an aggregate bit rate of up to 672 Gbps when transmitted over three spatial lanes.

# 1

# INTRODUCTION

## Overview of Intra-Data Center Networks

Data centers form the backbone of modern digital infrastructure, supporting applications such as cloud computing, artificial intelligence (AI), machine learning (ML), and large-scale data analytics. These facilities house thousands to millions of interconnected servers that exchange vast amounts of data. The internal network, known as the Intra-Data Center Network (DCN), is crucial for ensuring efficient, high-speed communication among switching nodes, storage units, and computing resources.

Intra-DCNs must meet high bandwidth demands, low-latency requirements, and energy efficiency constraints. They utilize various network topologies, including fat-tree, Clos, mesh, torus, and dragonfly architectures, to optimize performance and scalability [29, 38]. The growing need for faster data processing and reduced latency has driven significant advancements in networking technologies, particularly through high-speed optical communication solutions.

Since the late 1990s and continuing into the 2010s, silicon-based technology has played a key role in advancing data center infrastructure, enabling high-speed, high-performance computing (HPC) essential for AI- and ML-driven workloads. However, after the 2010s, the exponential growth of HPC computational speed, as measured by the LINPACK benchmark, began to plateau. To address increasing workload demands, the production of integrated circuits has increased hardware parallelization, pushing transistor footprint constraints and increasing DCN power consumption [29]. This has also led to challenges related to high power densities and chip aging [30, 32]. Since 2011, Google's aggregate energy consumption has grown exponentially, surpassing 25.9 terawatt-hours in 2023 [31].

As AI and ML drive unprecedented data demands, innovations within Intra-DCNs are essential not only for meeting data capacity requirements but also for addressing the increasingly unsustainable energy consumption trends. Strategies such as redesigning switching layers and implementing liquid cooling can help reduce DCN energy impact [10, 29]; however, the choice of modulation format also plays a critical role in determining both the data-carrying capacity and energy efficiency of the network.

## Modulation Format Considerations for Intra-DCNs

A modulation format encodes digital data (bits) onto an optical carrier signal for transmission by modulating the amplitude, phase, or polarization of laser light. The choice of modulation format directly impacts key performance metrics, including data rate, spectral efficiency, power consumption, and transmission reach. As Intra-DCNs face increasing capacity demands, selecting an optimal modulation format becomes essential for balancing performance, efficiency, and cost-effectiveness.

Traditional modulation formats, such as Non-Return-to-Zero On-Off Keying (NRZ-OOK), have long served as the foundation of optical communication systems. OOK, a form of intensity modulation, represents binary data by toggling the presence or absence of light at the receiver, enabling direct detection with a single photodiode. While simple and widely used in short-haul optical networks ($\approx 100\,\mathrm{m} - 10\,\mathrm{km}$) due to its ease of implementation [11], OOK suffers from poor spectral efficiency. Since it only utilizes two symbol states—on or off—its bandwidth requirements increase significantly at higher data rates compared to more spectrally-efficient modulation formats.

Spectrally-efficient formats manipulate amplitude levels, polarization states, phase, or a combination of these properties. These approaches enable multi-symbol constellations with $M = 2^k$ states, where each state represents $k$ bits of information. Modern Intra-DCNs primarily use $M$-ary Pulse Amplitude Modulation ($M$-PAM), which divides laser

amplitude into $M$ levels, each encoding $k$ bits. This format enhances spectral efficiency while maintaining compatibility with direct detection using single-photodiode receivers.

Google's data center architecture began evolving in 2010 with fiber lane parallelization of NRZ-OOK to meet rising aggregate bit rate demands. This transition led to the adoption of Quad-lane and later Octal-lane Small Form Factor Pluggable (QSFP/OSFP) modules, implementing 100 Gbps 4-PAM, ultimately achieving aggregate bit rates of up to 800 Gbps by 2020 (Fig. 1).

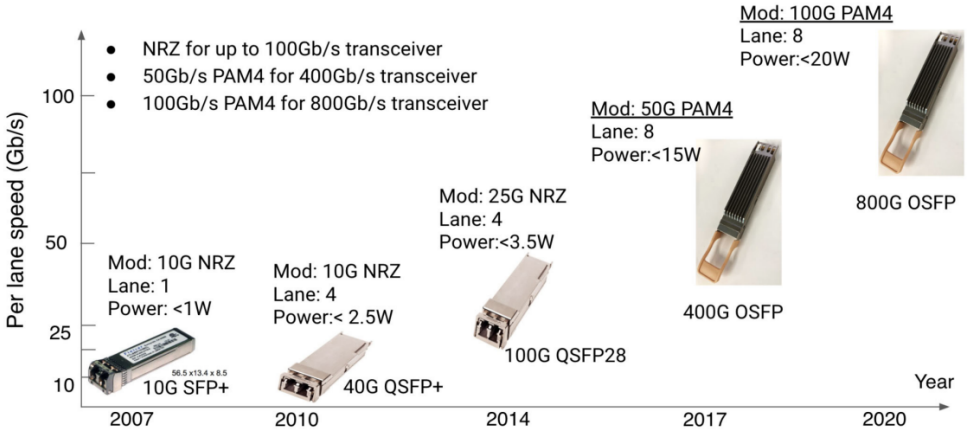



Figure 1: Google's DCN optical technology evolution [41].

Future Intra-DCNs will require data rates exceeding 1.6 Tbps to support next-generation data traffic. Meeting these demands may involve further parallelization of 4-PAM technology or the adoption of more advanced, spectrally efficient modulation formats. Notable candidates include higher-order PAM formats such as 8-PAM and FlexPAM, as well as hardware-simplified or DSP-free coherent Quadrature Amplitude Modulation (QAM).

The following sections provide brief descriptions of the aforementioned PAM and QAM formats, highlighting recent research on optimizing spectral efficiency and reducing complexity. While these approaches are promising, they also inherit several inherent drawbacks from their respective working principles.

Scaling Pulse Amplitude Modulation

8-PAM increases the number of amplitude levels to eight, allowing for higher bits per symbol than 4-PAM. For example, a study demonstrated a record-setting 402 Gbps 8-PAM intensity modulated direct detection transmission using a single digital-to-analog converter (DAC) to drive an O-band electro-absorption modulated laser (EML). The setup achieved a net bit rate of 348.62 Gbps over 2 km of standard single-mode fiber (SSMF) [14], marking a significant leap in high-speed data transmission.

However, the use of multi-level PAM formats like 8-PAM introduces sensitivity challenges. To mitigate these, researchers have explored pre-equalization techniques employing FIR filters to compensate for channel distortions. One study optimized linear pre-equalization for short-reach 225 Gbps 8-PAM systems, showing improvements in bit error rate (BER) by an order of magnitude after tuning the filter parameters [36].

In addition to 8-PAM, FlexPAM presents an alternative by using non-standard formats such as 5-PAM and 6-PAM, offering finer granularity between amplitude levels. This flexibility makes FlexPAM particularly valuable for intra-DCNs, where varying transmission requirements can benefit from more adaptable modulation schemes. Experimental results have demonstrated successful transmission up to 16-PAM at 20 GBd, showcasing the potential of FlexPAM for high-speed data communication [35].

While both 8-PAM and FlexPAM improve spectral efficiency, they introduce challenges related to energy consumption. The closer spacing of amplitude levels in these formats reduces signal-to-noise ratio (SNR) tolerance, making them less energy-efficient than 4-PAM. As seen in Fig. 1, energy consumption increases as the square of the number of constellation states ($M^2$), causing significant power consumption penalties as $M$ increases.

Coherent Quadrature Amplitude Modulation

Coherent QAM is an amplitude and phase-diverse modulation format known for its high spectral efficiency and receiver sensitivity. It has traditionally been used in long-haul and metro optical networks due to its ability to compensate for linear impairments, like chromatic dispersion (CD), using DSP. While it offers superior spectral efficiency compared to 4-PAM, making it a promising option for short-reach intra-DCN applications, it also presents several challenges.

Implementing coherent QAM requires precise control of both the amplitude and phase components, demanding sophisticated hardware such as a reference local oscillator (LO) for phase tracking, 90-degree hybrids, balanced photodiodes, and high-speed analog-to-digital converters (ADCs). Additionally, complex DSP algorithms are needed for carrier recovery, phase recovery, and polarization control in dual-polarization configurations.

Hardware and processing demands make coherent QAM adoption in Intra-DCNs challenging in terms of cost and energy. Research has focused on simplifying DSP or receiver hardware. Kramers-Kronig (KK) receivers recover phase information digitally from intensity measurements, eliminating the need for an LO and 90-degree hybrids [23, 27], but suffer from reduced LO gain and sensitivity [23]. Other efforts modify QAM constellation shaping to enhance phase-noise tolerance, reducing reliance on costly narrow-linewidth lasers [41].

To further simplify the DSP requirements, recent proposals like the coherent-lite transceiver use a copy of the transmit laser for self-homodyne detection, eliminating the need for carrier recovery and managing polarization and phase derotation with phase shifters [24].

Despite its superior spectral efficiency, the complexity, cost, and energy consumption of coherent QAM pose significant challenges for short-haul Intra-DCNs. While alternative approaches simplify certain aspects, they often sacrifice efficiency or performance. Direct-detection-based formats, though less resource-intensive, still offer lower data capacity. As a result, the practical deployment of QAM in intra-DCNs remains uncertain.

## Motivation for Advanced Modulation Formats

To meet the capacity demands of data centers while curbing the unsustainable trend of energy consumption, the Optical Communications Group at Montana State University has explored novel modulation formats as energy-saving alternatives to PAM and QAM for Intra-DCNs, relying exclusively on polarization-based techniques. Specifically, the approach involves using optimally shaped Stokes Vector Modulation (SVM) constellations, which utilize pure polarization states, along with its promising extension, Mode Vector Modulation (MVM), which leverages multiple spatial degrees of freedom simultaneously, allowing multi-dimensional signal spaces for constellation optimization containing a large number of symbol states $M$, while remaining compatible with direct-detection receivers [19]. This thesis focuses on the experimental generation and numerical evaluation of these emerging formats, laying the foundation for realistic evaluations of these constellations, some of which have been analytically shown to offer an optimal balance between energy and spectral efficiency suitable for Intra-DCN applications.

## Thesis Organization

This thesis is organized as follows: Chapter 2 establishes the theoretical foundations of SVM and MVM, providing a clear understanding of their operation and principles. Chapter 3 details the experimental setup for generating SVM/MVM constellations. Building on this, Chapter 4 presents the first experimental realization and optimization of a specialized constellation of SVM, the simplex SVM constellation. Finally, in Chapter 5, simulated investigations of MVM are conducted, exploring various propagation degradations and comparing the (6,64)-MVM constellation with 4-SVM and 4-PAM to achieve an aggregate data rate of 672 Gbps.

# 7

# THEORY

## Overview

Stokes Vector Modulation (SVM) and Mode Vector Modulation (MVM) are modulation formats that leverage the state of polarization (SOP) of an electromagnetic (EM) wave for data encoding. These formats are particularly advantageous for their compatibility with direct-detection receivers, offering a simpler, more power-efficient alternative to coherent detection formats. Despite their amenability toward direct detection (DD), SVM and MVM can also be generated and detected using coherent transceivers, providing additional flexibility for testing their performance. SVM operates within the three-dimensional Stokes space, while MVM extends its concept to higher dimensionalities by utilizing multiple spatial degrees of freedom (DoFs), such as those available in multicore or few-mode fibers (MCF/FMF).

This chapter briefly explores the theoretical foundations of both SVM and MVM, detailing their representations both mathematically and qualitatively so that one may better understand the potential advantages of these constellations for data center optical interconnects, particularly in terms of their balance between spectral efficiency, energy efficiency, and implementation complexity.

## Jones and Stokes Representations of Polarization

SVM is a modulation format that encodes binary data into the various SOPs of a propagating EM wave. The properties of any polarized EM wave can be expressed in terms of its orthogonal polarization components, $E_x$ and $E_y$, which represent the complex components of the electric field's orientation along the cardinal $x$- and $y$-axes in space (Fig. 2).

These components are superimposed with differential phase and amplitude differences

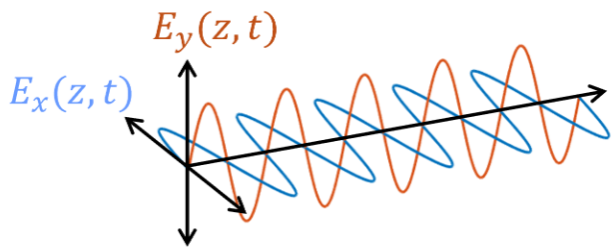



Figure 2: Propagating TEM wave with $x$ and $y$ polarization components.

to generate any desired SOP [3, 18]. In general, the components are functions of both space and time, ($z$, $t$); however, assuming $z$=0 at the fiber input, can be expressed by the following equations:

$$E_x(t) := a_x(t)e^{i(\omega_0 t+\phi_x(t))}\,|x\rangle\,, \tag{2.1a}$$

$$E_y(t) := a_y(t)e^{i(\omega_0 t+\phi_y(t))}\,|y\rangle\,, \tag{2.1b}$$

where $\omega_0$ denotes the carrier angular frequency, and $a(t)$ and $\phi(t)$ represent the time-varying amplitude and phase of each component, respectively. The basis vectors, represented using Dirac's ket notation $|\cdot\rangle$, align the components along the $x$- or $y$-axis of the Cartesian plane. Placing (2.1) in vector notation, the Jones vector $|s\rangle$ is defined and describes the SOP of the EM wave:

$$|s\rangle := [E_x, E_y]^T. \tag{2.2}$$

If a geometric representation using real vectors is of primary interest, the Stokes vector representation of the EM wave is used:

$$\vec{s} := [S_1, S_2, S_3]^T. \tag{2.3}$$

The Stokes vector components are given by the following equations, where the differential phase, $\delta$, is defined as the argument of the ratio of polarization components: $\delta := \arg(E_x/E_y)$.

$$S_1 := |E_x|^2 - |E_y|^2, \tag{2.4a}$$

$$S_2 := 2|E_x|^2|E_y|^2\cos(\delta) = 2\mathrm{Re}\{E_x^* E_y\}, \tag{2.4b}$$

$$S_3 := 2|E_x|^2|E_y|^2\sin(\delta) = 2\mathrm{Im}\{E_x^* E_y\}. \tag{2.4c}$$

The amplitude of the Stokes vector $S_0$ is given by the vector norm of $\vec{s}$:

$$S_0 = \|\vec{s}\| = \sqrt{S_1^2 + S_2^2 + S_3^2}, \tag{2.5}$$

where $\vec{s}$ may be scaled by its norm, e.g., $\hat{s} = \vec{s}/\|\vec{s}\|$ such that the amplitude $S_0 = 1$. The SOP can then be visualized on the Poincaré sphere, as depicted in the figure below.

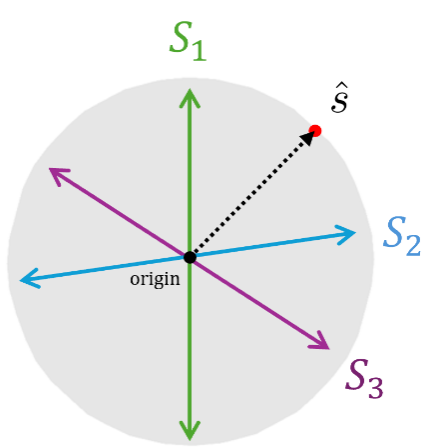



Figure 3: Arbitrary unit Stokes vector $\hat{s}$ plotted on the Poincaré sphere.

## Mode Vector Modulation

MVM extends the concept of SVM beyond the conventional three-dimensional Stokes space into a generalized $(N^2 - 1)$-dimensional Stokes space, where $N$ refers to the number of polarization DoFs. This expansion leverages all spatial and polarization DoFs in multicore or few-mode fiber infrastructure, significantly enhancing the signal space for constellation optimization and enabling greater separation between MVM Stokes symbols.

As a result, MVM provides resilience to transmission impairments and noise, improving receiver sensitivity. Additionally, with its higher-dimensional signal space, MVM constellations offer optimal trade-offs between energy and spectral efficiency, making them a promising alternative for short-haul optical links.

This section introduces the fundamental concepts of MVM and explores preliminary findings from previous theoretical studies on this emerging modulation format.

### The MVM Concept

Conventional SVM uses single-mode fiber (SMF) infrastructure, providing a single spatial DoF (mode) with two polarization components, $x$ and $y$, as depicted in Fig. 4.

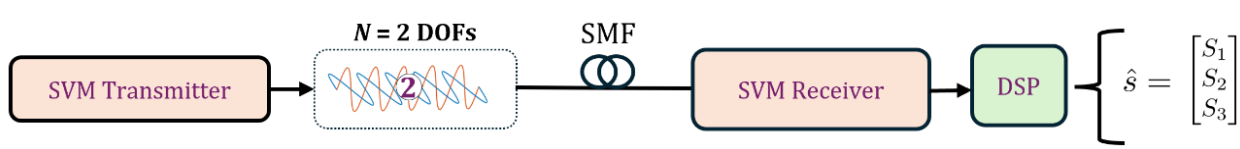



Figure 4: The SVM concept.

The SVM transmitter modulates the amplitude and phase between the two polarization DoFs to generate various SOPs of the outgoing EM wave [18]. Since the SMF has generalized dimensionality $N^2 - 1 = 3$, the SVM transmitter can generate numerous constellation arrangements supporting $M$ points that may be visualized using the Poincaré sphere. The ideal and noiseless SVM/DD receiver measures the intensity components of the received SOP

using photodiodes, whereas digital signal processing (DSP) compensates for polarization rotations caused by fiber birefringence. Ultimately, the transmitted Stokes vector symbol is recovered, and the bit error rate (BER) is computed [18].

The MVM concept is a Multiple-Input Multiple-Output (MIMO) extension of SVM that utilizes all $k$ single-mode cores or few-mode fiber modes at the same time with accumulated polarization DoFs $N > 2$ (Fig. 5). Rather than treating each spatial DoF as a separate SVM channel, MVM modulates the amplitude and phase of each polarization component simultaneously to uniquely represent an $(N^2 - 1)$-dimensional Stokes vector. This vector can only be qualitatively imagined, as direct visualization in dimensions greater than three is not possible. Instead, MVM constellations are conceptually depicted using the surface of an $(N^2 - 1)$-dimensional Poincaré hypersphere [19].

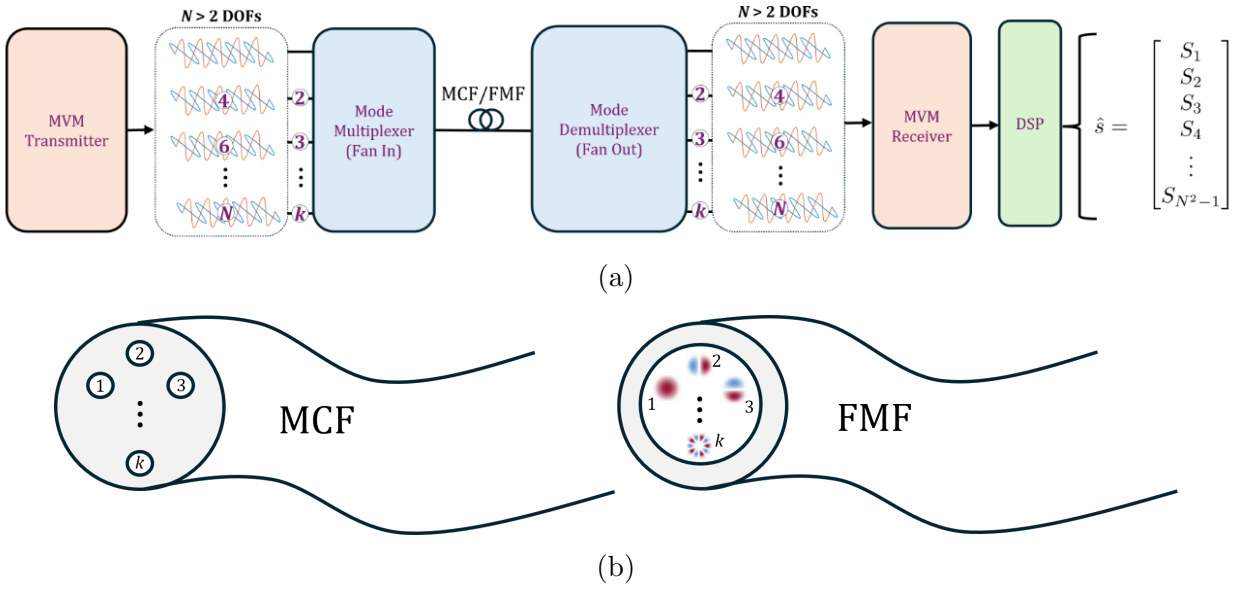



Figure 5: (a) The MVM concept; (b) multicore fiber (MCF) and few-mode fiber (FMF) with $k$ single-mode cores or modes. In reality, the modes of the FMF overlap; the depiction is purely qualitative.

As shown in Fig. 5, the MVM transmitter modulates the amplitude and phase of all $N$ polarization components, selecting from an optimally shaped constellation including an

$M$-number of $N$-dimensional Jones vectors [19]. Each pair of polarization components is launched into a single core or mode of the MCF/FMF fiber using a fan-in topology or mode multiplexer.

At the fiber end, all $N$ polarization DoFs are demultiplexed. The ideal and noiseless MVM/DD receiver detects the polarization components of each mode and, in essence, functions as a high-speed multimode polarimeter. The optical communications group at MSU has devoted significant efforts to designing the MVM receiver architecture, which currently exhibits linear hardware complexity scaling with the number of DoFs $N$ [19]. Finally, the DSP block compensates for polarization rotations using an adaptive algorithm such as Least Mean Squares (LMS) to eventually recover the transmitted Stokes vector [37].

Optimized Constellations of SVM/MVM

The constellations of MVM are denoted as ($N$, $M$)-MVM, where $N$ represents the number of polarization degrees of freedom and $M$ is the constellation cardinality, indicating the number of points distributed along the surface of the $(N^2 - 1)$-dimensional Poincaré hypersphere. Constellations for SVM ($N = 2$) were numerically optimized in [4] to maximize the Euclidean distance between Stokes symbol vectors. Later, further optimizations based on minimizing the symbol error probability were performed, achieving theoretical optima for constellations with up to 32 points [3]. Additionally, practical implementations of SVM have been developed using theoretically suboptimal cubic constellations, incorporating variations of intensity-modulated SVM [18, 25].

To extend this analysis into the generalized Stokes space and enhance existing SVM constellations, the optical communications group at MSU developed a method for generating optimally separated Stokes symbol vectors in any dimensionality [19]. This method is inspired by the physical analogy of electrostatics and Coulomb forces [39].

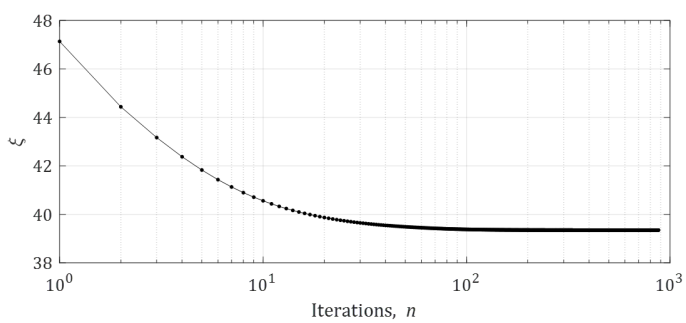
In this approach, each Stokes symbol in an $M$-point SVM constellation behaves like a

point charge on a superconducting Poincaré sphere, repelling others under Coulomb forces to minimize potential energy. This principle extends naturally to the generalized $(N^2-1)$-dimensional Stokes space, where $M$ points will also equilibrate in higher dimensions.

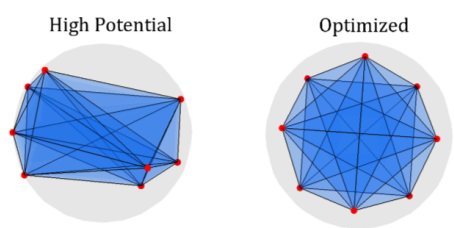
The method involves minimizing an objective function $\xi$, which represents the system's potential energy, using gradient descent [5]:

$$\xi := \sum_{i=1}^{M} \sum_{j=i+1}^{M} \Omega(d_{ij}). \tag{2.6}$$

The function $\Omega(d_{ij}) \approx d_{ij}^{-1}$ is inversely proportional to the Euclidean distance between all pairs of $M$ Stokes vertices. As the cumulative distance between vertices increases, the overall potential energy $\xi$ decreases. Figure 6 illustrates this process, where an 8-SVM constellation is optimized from an initially random collection of 8 Stokes vectors.




(a)




(b)

Figure 6: Optimization of 8-SVM. (a) Reduction of the potential energy $\xi$ after $\approx 900$ iterations; (b) starting (left) and optimized (right) constellations of 8-SVM. The constellation on the right represents the square antiprism (twisted cubic constellation).

Analytical Performance of MVM

The performance of various $M$-point optimized SVM and MVM constellations was analytically studied in terms of their receiver sensitivity under an increasing number of polarization DoFs $N$ (Fig. 7). As shown in the figure, the bit signal-to-noise ratio (SNR) required to achieve a bit error probability of $10^{-4}$ significantly decreases, approaching the Shannon capacity limit, when $N$ increases from the single-mode case ($N = 2$) to the multicore/multimode case carrying 2 and then 4 spatial DoFs ($N = 4, 8$). While increasing the number of constellation points improves the spectral efficiency of the optical link, the energy gains begin to diminish beyond a certain point.

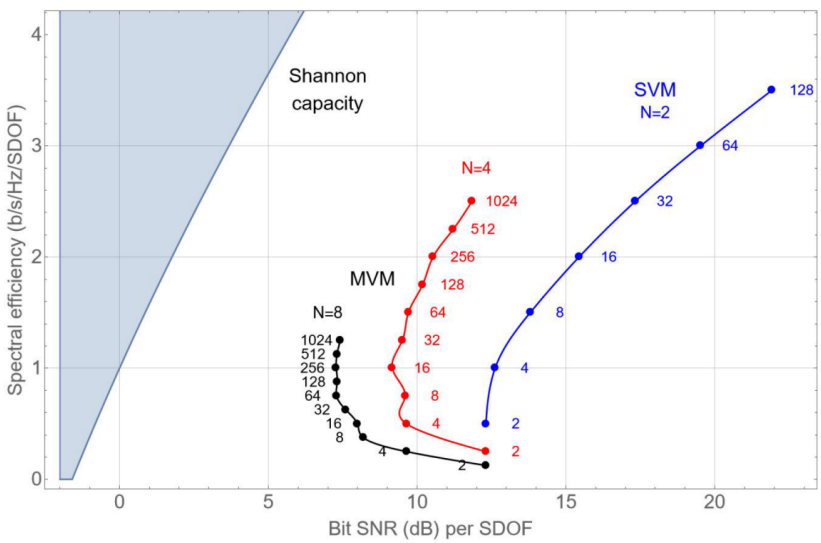



Figure 7: MVM spectral efficiency per spatial degree of freedom (SDoF) vs required bit SNR per SDoF to achieve a bit error probability of $10^{-4}$. The blue, red, and black curves represent dimensionality $N = 2$, 4, and 8, respectively. Each data point represents the number of points $M$ in the constellation [19].

These points—belonging to a subset of MVM constellations—offer the optimal trade-off between spectral and energy efficiency. They have been theoretically identified using symmetric, informationally complete, positive operator-valued measures (SIC-POVM) [12].

These constellations, known as the simplex constellations of SVM and MVM, occur when $M = N^2$ for any dimensionality $N$. The experimental generation and optimization of the SVM simplex (4-SVM), which portrays a regular tetrahedron inscribed within the Poincaré sphere, is discussed later in this thesis.

Considering just the simplex constellations of MVM, the receiver sensitivity can be compared to that of $M$-ary Pulse Amplitude Modulation ($M$-PAM) and $M$-ary Quadrature Amplitude Modulation ($M$-QAM) (Fig. 8). As shown in the figure, coherent QAM provides the highest spectral efficiency; however, when compared to DD formats, it is hardware complex and incurs significant energy penalties with an increasing number of constellation points.

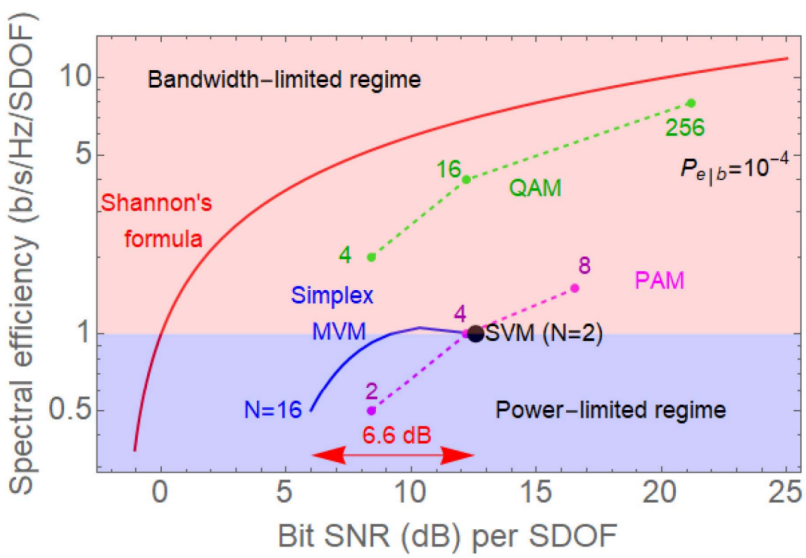



Figure 8: MVM simplex spectral efficiency per spatial degree of freedom (SDoF) vs required bit SNR per SDoF to achieve a bit error probability = $10^{-4}$. Other modulation format results, including $M$-ary PAM and QAM for comparison [19].

Simpler, direct-detection-based formats such as PAM and MVM do not match QAM in terms of spectral efficiency but offer significantly lower receiver complexity. This is the primary reason why 4-PAM is widely used in short-haul optical interconnects. However,

as illustrated in the figure, MVM can improve the energy efficiency by up to 6.6 dB over 4-PAM when utilizing 16 polarization DoFs simultaneously, with only a moderate reduction in spectral efficiency. Certain MVM constellations may even surpass 4-PAM in terms of spectral efficiency while still achieving a few dBs of energy gain.

Future data center optical interconnects aim to support aggregate bit rates of 1.6 Tbps and beyond. However, scaling to higher bit rates presents additional design challenges, including the hardware complexity of advanced modulation formats, footprint constraints, energy efficiency, and overall cost. Table 1 presents various scenarios for future optical interconnects striving to meet the 1.6 Tbps benchmark, assuming direct detection remains the preferred choice for short-haul links due to its simplicity and cost advantages.

| **Modulation Format** | **Spatial Lanes** | **Fiber** | **Symbol Rate Per Lane (GBd)** |
|---|---|---|---|
| 4-PAM | 16 | SMF | 50 |
| 4-SVM | 16 | SMF | 50 |
| (4,16)-MVM | 8 | dual-core MCF | 50 |
| (8,256)-MVM | 4 | quad-core MCF | 50 |
| 4-PAM | 8 | SMF | 100 |
| 4-SVM | 8 | SMF | 100 |
| (4,16)-MVM | 4 | dual-core MCF | 100 |
| (8,256)-MVM | 2 | quad-core MCF | 100 |

Table 1: Scenarios for achieving an aggregate bit rate of 1.6 Tbps for future data centers.

Achieving 1.6 Tbps is possible with 4-PAM, assuming greater parallelization utilizing additional spatial lanes. However, adopting MVM with 8 polarization DoFs enables larger

constellations, reaching up to, and potentially beyond, 256 points, as shown in the table above.

Due to their superior energy efficiency over 4-PAM and coherent QAM, their compatibility with direct detection, and advanced constellation designs offering a broad range of performance capabilities, SVM and MVM provide a compelling solution for achieving the data rate demands of future data center optical interconnects while minimizing energy consumption.

# EXPERIMENTAL DESCRIPTION

## Overview

This chapter provides a detailed description of the setup for Stokes Vector Modulation (SVM) and Mode Vector Modulation (MVM) experiments, including the devices used and the calibration procedures implemented. The calibrations are crucial for validating and testing the proposed modulation formats. The chapter is organized as follows: first, a brief description of the key experimental devices and their roles in the setup; second, an explanation of the communication between the Optical Modulation Analysis (OMA) software, MATLAB, and the personal computer (PC); and finally, an outline of the calibration steps required to ensure accurate measurements.

## Experimental Devices

SVM/MVM experiments involve a range of optical and electrical devices designed for signal generation, transmission, reception, and measurement. Located within the optical communications lab at Montana State University, a server rack houses several key components essential for the advanced testing of optical communication systems and various modulation formats (Fig. 9). This server rack supports fiber transmission experiments within a recirculating loop architecture (RCL), where a propagating signal can recirculate through the same fiber spool up to $N$ times, emulating $N$ spans in long-haul optical communication systems.

The primary components of the RCL include an arbitrary waveform generator (AWG) for generating transmission signals at gigahertz frequencies, a dual-polarization In-phase and Quadrature (IQ) transmitter, an RCL switch for transitioning between loop and load states of the RCL, erbium-doped fiber amplifiers (EDFAs), optical fibers, a variable optical attenuator

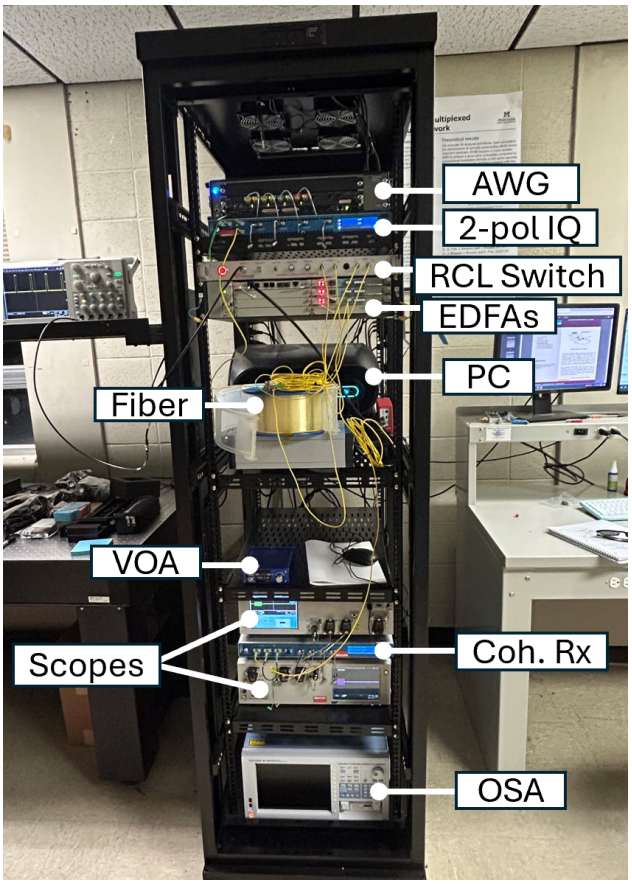



Figure 9: Experimental devices placed on the server rack inside the optical communications lab at Montana State University.

(VOA), a coherent intradyne receiver, real-time digital oscilloscopes, and an optical spectrum analyzer (OSA).

Each experimental device can be programmed using MATLAB through a central PC. However, signal reception—specifically, sample data from the oscilloscopes—is managed by Tektronix's OMA software. This software interacts with MATLAB to perform offline digital signal processing (DSP). Details of the interaction between the OMA software and MATLAB are discussed in later sections.

Experimental Setup

For the experiments conducted in this thesis, the RCL is bypassed during back-to-back experiments aimed at initial SVM constellation generation. These experiments use only a subset of the components from the server rack. As shown in Fig. 10, the quad-channel AWG drives each quadrature port of both the $x$ and $y$ polarization components of the dual-polarization IQ transmitter. This transmitter is responsible for generating the optical signals that are launched into the optical fiber. By utilizing both polarization components of a continuous wave (CW) laser, the transmitter modulates the in-phase and quadrature parts of each component, enabling phase and amplitude diverse modulation formats. The modulation is performed by two Mach-Zehnder modulators (MZMs), each containing a pair of Mach-Zehnder interferometers (MZIs). The MZIs convert electrical signals from the baseband AWG source into an optical output.

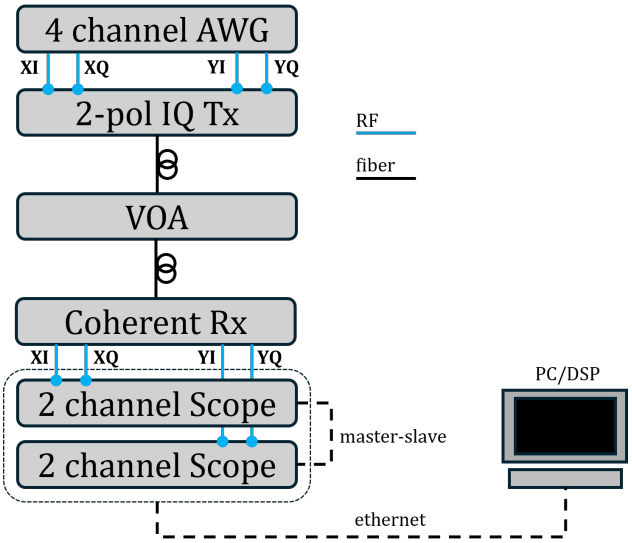



Figure 10: Back-to-back experimental block diagram.

A short fiber cable (approximately 3 m) connects the transmitter's optical output to the VOA, which adjusts the optical signal power to a desired level, simulating real-world transmission conditions by emulating fiber attenuation or signal degradation. This allows

for real-time control of the optical signal's power during experimentation.

A second fiber, of the same length, connects the VOA to the coherent intradyne receiver. The receiver detects the incoming optical signal and converts it back into a series of electrical signals for analysis. The receiver architecture includes a local oscillator (LO), a pair of 90-degree optical hybrids, and balanced photodiodes, enabling the extraction of both the in-phase and quadrature parts from each polarization component of the received signal.

Finally, real-time digital oscilloscopes sample and display each electrical signal coming from the receiver. The sampled data is then sent to the host PC, where offline DSP is performed.

The experimental setup includes the devices summarized in Table 2. It consists of the 92 GSa/s Keysight M8196A arbitrary waveform generator, a 40 GHz Quantifi Photonics IQTX-1203 dual-polarization IQ transmitter, an Optilab VOA-C-M variable optical attenuator, a Quantifi Photonics IQRX-1004 coherent receiver, and two synchronized 100 GSa/s Tektronix DPO75002SX oscilloscopes.

| **Device** | **Model** |
|---|---|
| Arbitrary Waveform Generator | Keysight M8196A |
| 2-pol IQ Transmitter | Quantifi Photonics IQTX-1203 |
| Variable Optical Attenuator | Optilab VOA-C-M |
| Coherent Receiver | Quantifi Photonics IQRX-1004 |
| Oscilloscopes | Tektronix DPO75002SX |

Table 2: Device models.

## Optical Modulation Analysis and MATLAB Interaction

The Tektronix OMA software is a powerful tool designed to analyze and process electrical signals sampled by Tektronix oscilloscopes. It enables the demodulation of optical signals, extracting the underlying properties from modulated optical carriers, and performs in-depth analysis of modulation formats such as quadrature amplitude modulation (QAM) and phase-shift keying (PSK). The software generates constellation plots, which visually represent the signal's quality in terms of key metrics like signal-to-noise ratio (SNR) and symbol error rate (SER). It also calculates the Error Vector Magnitude (EVM), a statistical measure of constellation noise and degradation, and produces eye diagrams to evaluate the timing and shape of transmitted pulses. Additionally, the OMA software supports offline DSP using MATLAB, allowing detailed step-by-step analysis of received signals before visualizations and error rates are produced and computed. With its support for arbitrary complex modulation formats and its ability to support user-defined DSP, the OMA software is a key tool for studying advanced modulation formats using a coherent receiver.

The interaction between MATLAB and the OMA software is illustrated in Fig. 11. Before a record of the received signal is sampled by the oscilloscopes, the OMA software writes relevant analysis parameters to the MATLAB workspace, such as the modulation format, symbol rate, and calibration data. The OMA software then queries a record acquisition by the oscilloscopes where sampled data is transferred to the PC and written to the MATLAB workspace. With access to experimental data, MATLAB performs various user-defined DSP functions to process the data according to custom modulation formats. The results of these computations are then sent back to the OMA software for visualizations, such as eye diagrams, constellation plots, or representations of the state of polarization (SOP) using the Poincaré sphere. Additionally, the OMA software computes important performance metrics, including bit error rate (BER) and other signal statistical measures like the EVM.

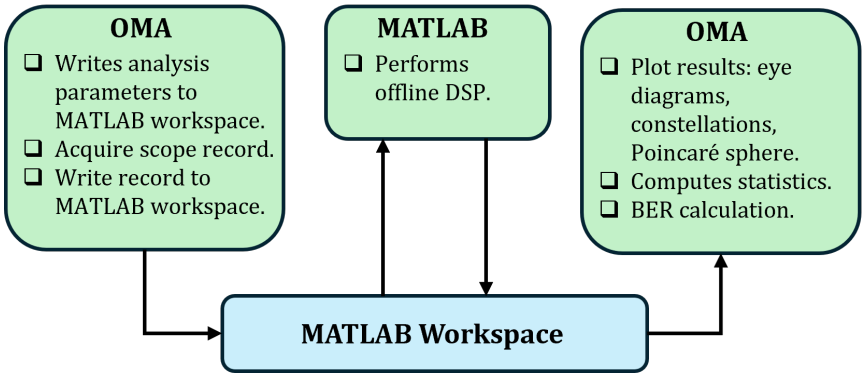
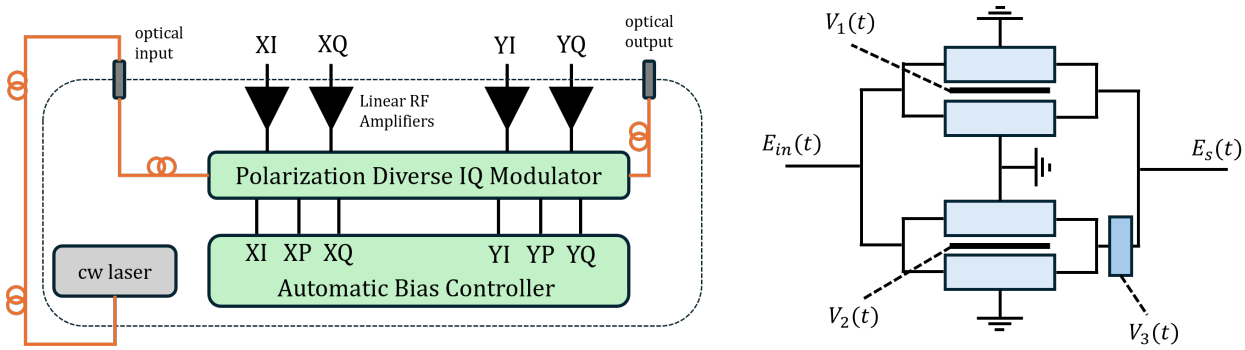



Figure 11: OMA and MATLAB interaction.

## Transmitter Calibrations

### DC Biasing of the Mach-Zehnder Interferometers

Calibrating the IQ transmitter is crucial for accurate signal generation and measurement. DC bias calibration of the MZIs ensures proper modulation by maintaining the desired operating point. Figure 12a shows a block diagram of the Quantifi Photonics IQTX, where each MZI is automatically biased via an internal bias controller (ABC). The XP/YP phase component introduces a phase offset between the arms of each MZM (Fig. 12b).




(a) (b)

Figure 12: (a) Quantifi Photonics IQTX block diagram; (b) one of the two MZMs inside the polarization diverse IQ modulator.

The MZM consists of two MZIs that control the real and imaginary parts of a single polarization component. The polarization-diverse IQ modulator inside the IQTX contains two MZMs—one for each polarization component of the signal. The electric field $E_s(t)$ that emerges from each MZM can be expressed in terms of the input electric field $E_{in}(t)$ and the driving voltages $V(t)$ as follows:

$$E_s(t) = \frac{i}{2} e^{\frac{-i\pi V_3(t)}{2V_{\pi 3}}} \left[ \sin\left(\frac{\pi}{2}\frac{V_1(t)}{V_{\pi 1}}\right) - e^{\frac{-i\pi V_3(t)}{2V_{\pi 3}}} \sin\left(\frac{\pi}{2}\frac{V_2(t)}{V_{\pi 2}}\right) \right] E_{in}(t), \tag{3.1}$$

where $V_\pi$ represents the half-wave voltage for each MZI and the additional phase-shifter. The DC bias voltages set by the ABC are locked to ensure null intensity, e.g., destructive interference between MZI arms, at the modulator output in the absence of an externally applied AC voltage. The AC-coupled voltages, $V_1(t)$ and $V_2(t)$, are then used to modulate the signal around the operating bias.

The voltage $V_3(t)$ is fixed to a DC value $V_3(t) = V_{\pi 3}$, which results in a $\pi/2$ phase difference between the interferometric arms of the MZM. This allows the expression to be simplified further as:

$$E_s(t) = \frac{E_{in}(t)}{2} \left[ \sin\left(\frac{\pi}{2}\frac{V_1(t)}{V_{\pi 1}}\right) + i \sin\left(\frac{\pi}{2}\frac{V_2(t)}{V_{\pi 2}}\right) \right]. \tag{3.2}$$

The terms inside the parentheses represent one polarization component's in-phase and quadrature parts. For the polarization-diverse IQ transmitter, these terms, when combined, can generate any desired Jones vector symbol $|s\rangle$ for use in SVM and MVM experiments.

To perform bias calibration, the IQTX ABC conducts a DC voltage sweep, ranging from -14 to 14 volts, on each MZI and phase-shifter while measuring the output power in milliwatts, as shown in Fig. 13. After the sweep, the ABC locks to the voltage corresponding to the minimum intensity for each bias curve, both for the in-phase and quadrature components, as well as for the quadrature point ($\pi/2$) of the two phase-shifters. However,

external factors such as temperature fluctuations can cause the DC bias corresponding to these points to shift over time. To compensate for these variations, the ABC applies a small dither signal to the bias (a fraction of $V_\pi$) to track and adjust the biases, correcting any lateral shifts in their respective transfer functions.

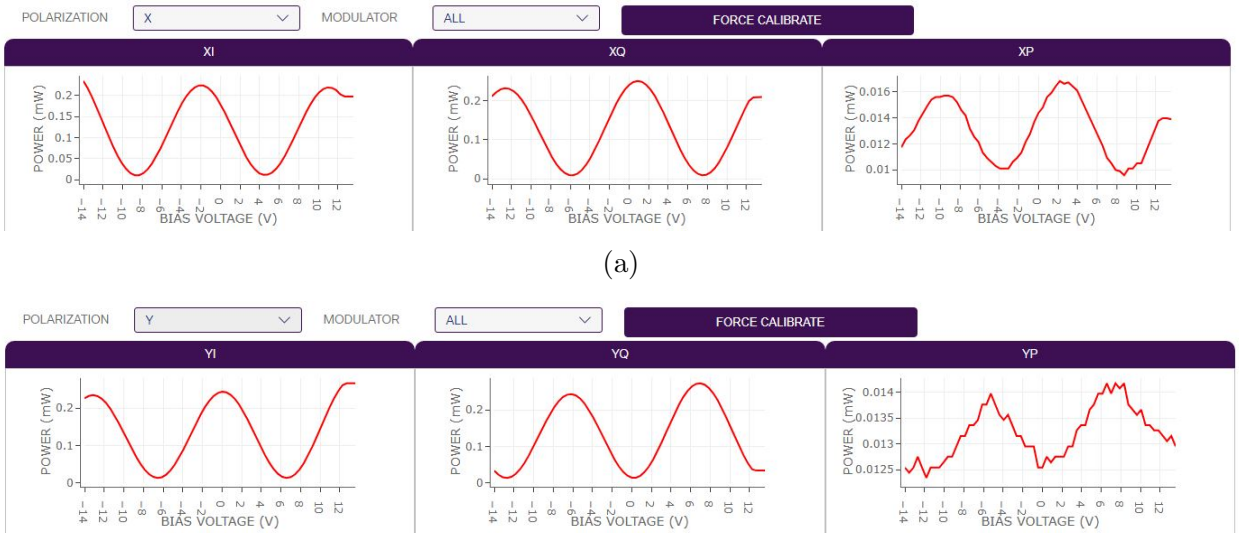



Figure 13: DC bias sweeps performed by the ABC before external modulation for (a) $x$-polarization I, Q, and phase; (b) $y$-polarization I, Q, and phase.

## Receiver Calibrations

### 90-degree Optical Hybrid Imbalance

Receiver calibration is essential for ensuring the accuracy and reliability of the received signal in coherent optical systems. One of the most common issues encountered in coherent receivers is 90-degree hybrid imbalance. This occurs when the hybrid, which is designed to split the received optical signal into its respective polarization IQ components, does not perfectly divide the signal. As a result, phase and amplitude imbalances can occur between the I and Q channels, meaning that they are no longer exactly in quadrature, or one channel

may receive more power than the other.

Ideally, the I and Q channels should have equal amplitude and a phase offset of exactly 90 degrees. However, practical hybrid components may introduce discrepancies that deviate from this ideal. Any discrepancy caused by the hybrid can introduce significant errors in the received electrical signals, negatively impacting constellation uniformity. Therefore, a calibration matrix is used to adjust the phase and amplitude of the received signals to address these issues.

The hybrid imbalance is corrected by mixing the signal and LO lasers with a frequency offset on the order of megahertz. In the ideal case, the sampled signals, when plotted in the complex plane, should rotate at this intermediate frequency (IF) offset, forming a uniform circular shape. Any deviation from rotational circular symmetry, such as ellipticity, indicates a phase imbalance in the hybrid. This imbalance is corrected by applying a compensating matrix $\mathbf{H}$. Similarly, amplitude discrepancies are corrected within $\mathbf{H}$ by adjusting the relative gains of either the I or Q channel to equalize their powers. The calibrated hybrid imbalance matrix for experimental implementations is a $2 \times 4$ matrix and is given by the following:

$$\mathbf{H} := \begin{bmatrix} 0.6900 + 0.0000i, -0.0052 - 0.6901i, 0.0017 - 0.0025i, -0.0028 - 0.0019i \\ 0.0000 + 0.0000i, -0.0002 - 0.0004i, 0.6469 - 0.0000i, -0.0009 - 0.0713i \end{bmatrix}. \quad (3.3)$$

<u>Channel Skew</u>

Another critical calibration aspect is channel skew, which occurs when the signals corresponding to different polarization states of the transmitted signal are received with temporal delays. These relative delays arise from the RF connectors and unequal cable lengths, typically found to be on the order of picoseconds. Timing offsets between channels may cause signal distortion, as each IQ channel is sampled at sub-optimal instances.

To address this, phase slope measurements are performed. By measuring the phase difference between the signal and LO lasers, the phase slope is calculated, revealing the timing offset between the IQ channels. If discrepancies are detected, the timing can be adjusted by delaying or advancing one of the channels until the phase slope becomes linear, indicating proper synchronization.

The channel skew measurements are represented by a $1 \times 4$ vector $\mathbf{s}$, with channel 1 as the reference. The elements of $\mathbf{s}$ are provided in units of picoseconds:

$$\mathbf{s} := \left[00.0, 32.0, -12.0, 31.2\right]. \tag{3.4}$$

# EXPERIMENTAL OPTIMIZATION OF THE SVM SIMPLEX CONSTELLATION

## Overview

While the simplex Stokes Vector Modulation (SVM) constellation has been identified to be theoretically optimal for $N = 2, M = 4$ [1, 2, 4], its practical implementation has been non-existent. The primary challenge in generating the SVM simplex and other Mode Vector Modulation (MVM) simplexes in the lab lies in the ability to form them with high symmetry. Consequently, most experimental research involving quaternary SVM constellations has opted for simpler, yet sub-optimal, square constellations within Stokes space [15, 21, 22]. While these alternatives may be easier to implement, they fall short of capturing the full theoretical advantages of the simplex SVM constellation.

This inaugural study, focused on the experimental realization of the SVM simplex constellation, marks a pivotal step in advancing polarization-based optical communications systems. By overcoming the challenge of experimentally generating, optimizing, and testing the constellation, this study sets the stage for future research involving the experimental generation of the highly promising simplex MVM constellations across different dimensionalities.

This chapter is organized as follows: first, a brief outline of the equivalence mapping between the mathematical descriptions of generic SVM/MVM symbols and the IQ modulators is given; then, using the aforementioned mathematical equivalencies, the first experimental generation and performance evaluation of the SVM simplex is presented including several necessary steps involving transmitter characterizations and digital signal processing (DSP). Finally, an optimization procedure is investigated for reducing the implementation penalty of the format. It is found that, with adequate feedback control of modulator driving voltages, SVM/MVM constellations can be generated with high symmetry in the lab, potentially unlocking their full energy efficiency advantage over other modulation formats considered

for optical interconnects.

## SVM/MVM Generation using Quadrature Modulators

### Conversions between Jones and Stokes Spaces

Both $M$-ary SVM and ($N$, $M$)-ary MVM encode information into various states of polarization (SOP) of a propagating electromagnetic (EM) wave. Once the encoded signal reaches the receiver, the Jones vector components are measured and analyzed to determine the transmitted symbol. Since IQ modulators, operating in Jones space, are used to generate the SVM/MVM symbols, conversions between Jones space and Stokes space are commonly performed during constellation visualization and inside the DSP functions, assuming an intradyne coherent receiver is used to detect the received optical signal.

SVM symbol coordinates are encoded using a Jones vector with dimensionality $N = 2$. For SVM, any corresponding Jones vector symbol $|s\rangle$ belonging to its constellation contains two complex-valued elements, $E := a + ib$, each representing a single polarization component of the propagating EM wave (4.1). Similarly, for MVM, a generalized Jones vector containing $N$ polarization components is used for representation of its symbols, where $k = N/2$ represents the number of fiber modes or cores allocated for multi-dimensional modulations (4.2).

$$|s\rangle = \Big[E_x, E_y\Big]^T \tag{4.1}$$

$$|s\rangle = \Big[E_{x1}, E_{y1}, \cdots, E_{xk}, E_{yk}\Big]^T \tag{4.2}$$

For this experimental example, consider the SVM simplex constellation where $M = 4$. This 4-point SVM constellation offers the theoretical optimal trade-off between energy and spectral efficiency for $N = 2$. The vertices of this constellation form a regular tetrahedron

in Stokes space and are created through the superposition of two complex orthogonal polarization components, making it suitable for implementation inside the dual-polarization nested Mach-Zehnder Interferometer (MZI) structure. The mathematical description of the vectors is given in Jones space (4.3), where the basis vectors $|x\rangle$ and $|y\rangle$ represent the $x$ and $y$ polarization components, respectively.

$$|s_1\rangle = |x\rangle\,, \tag{4.3a}$$

$$|s_2\rangle = \frac{1}{\sqrt{3}}|x\rangle + \sqrt{\frac{2}{3}}|y\rangle\,, \tag{4.3b}$$

$$|s_3\rangle = \frac{1}{\sqrt{3}}|x\rangle + \sqrt{\frac{2}{3}}e^{i\frac{2\pi}{3}}|y\rangle\,, \tag{4.3c}$$

$$|s_4\rangle = \frac{1}{\sqrt{3}}|x\rangle + \sqrt{\frac{2}{3}}e^{i\frac{4\pi}{3}}|y\rangle\,. \tag{4.3d}$$

The SVM simplex constellation can be rotated by any unitary matrix to generate a new set of Jones vectors; however, selective rotations can simplify implementation. Considering the vectors above, the polarization basis $|x\rangle$ is purely real. Therefore, it is possible to produce the SVM simplex using only three of the four available modulators in the coherent transmitter. To view the SVM simplex constellation in Stokes space, one must first convert the Jones vector symbols into Stokes vectors that can be plotted on the Poincaré sphere. For $N = 2$, the conversion is done using the unitary transformations of the Pauli spin matrices $\sigma_i$:

$$\sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \sigma_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma_3 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \tag{4.4}$$

where the elements $S_i$ of the corresponding Stokes vector $\hat{s}$ can be found using the following

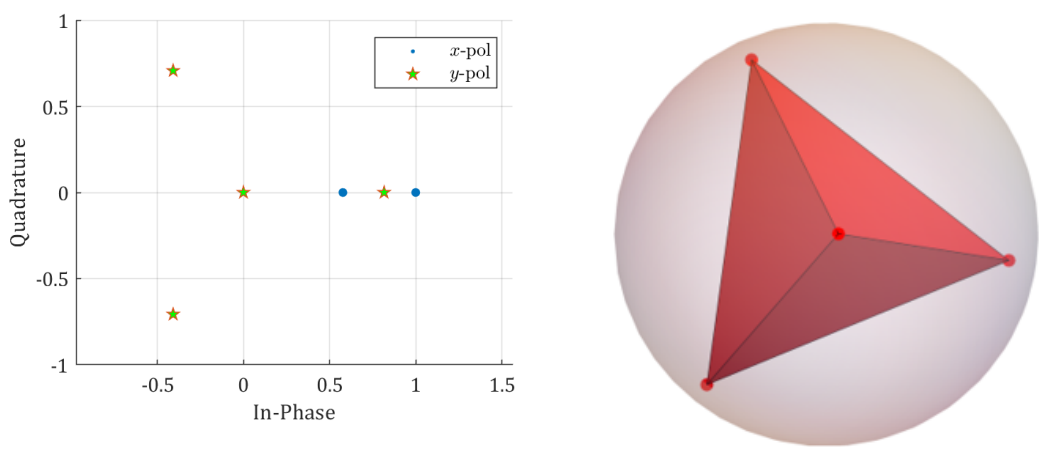
quadratic form [13]:

$$S_i = \langle s| \sigma_i |s\rangle \,, \tag{4.5}$$

where the vector $\langle s|$ represents the Jones vector's complex conjugate transpose.

Upon conversion into Stokes space, the quadratic form computation results in the loss of the absolute phase of the EM wave. Thus, any transformation back into Jones space would lead to absolute phase ambiguity. The Stokes vector receiver is a direct-detection receiver and does not require knowledge of the absolute phase. Therefore, when coherent receivers are used to detect the Jones vector components, conversion into Stokes space is performed in the DSP to emulate true Stokes vector receivers.

The Jones vectors from (4.3) can be used to plot the simplex SVM constellation in the complex plane. In Fig. 14a, the real (in-phase) and imaginary (quadrature) parts of each polarization component are shown.




(a) (b)

Figure 14: (a) The SVM simplex Jones vectors plotted in the complex plane; (b) the formed regular tetrahedron in Stokes space after using the quadratic form with the Pauli spin matrices.

Performing the conversion given by (4.5), the corresponding simplex SVM constellation visualized in the 3-dimensional Stokes space is shown in Fig. 14b. To form the constellation shape in Stokes space, each Stokes symbol vertex is connected by a line, forming a regular triangle with three adjacent Stokes vertices. After connecting all four facets, it can be seen that the Stokes vectors corresponding to the simplex SVM constellation form a regular tetrahedron in the 3-dimensional Stokes space.

Relating MZI Transfer Functions and Modulator Driving Voltages

To relate the set of Jones vectors to modulator driving voltages, special consideration must be given to each MZI transfer function. In general, each transfer function represents a sinusoidal variation in output optical amplitude with applied voltage. However, the sinusoidal variation can be described either in terms of amplitude or intensity. The $2V_\pi$ range corresponds to a peak-to-peak swing in intensity or a minimum-to-peak swing in amplitude and is illustrated for both functions in Fig. 15.

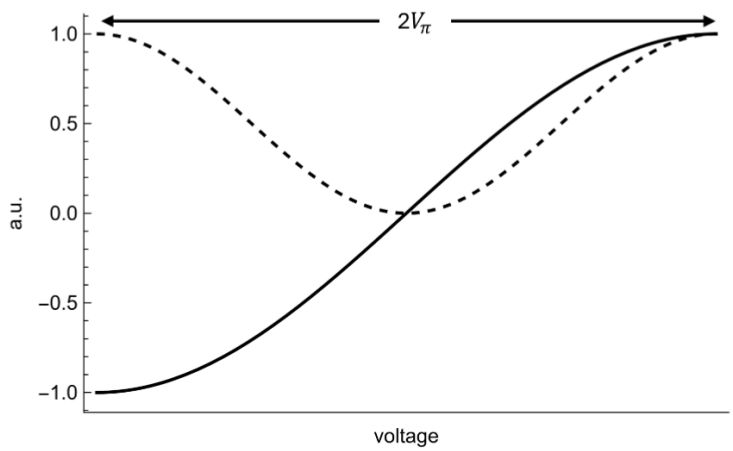



Figure 15: The MZI intensity (dotted) and amplitude (solid) transfer functions.

Assuming a null bias is applied, each MZI operates along the minimum point of the intensity curve, or equivalently, the quadrature point of the amplitude curve. The DC bias

voltage can be set experimentally using an automatic bias controller (ABC), which holds the null point for each MZI. Although both transfer functions vary sinusoidally, mathematically, they will produce different results when equated to the SVM simplex symbol Jones vectors. For SVM/MVM and coherent modulations, the amplitude transfer function is considered when applying an external voltage to the modulator.

The complete amplitude transfer function used to describe each MZI is given by the following equation:

$$\sin\left(\frac{\pi}{2}\frac{\alpha}{V_\pi}v(t)\right). \tag{4.6}$$

Since each MZI half-wave voltage is on the order of volts, internal RF amplifiers are required to boost the arbitrary waveform generator (AWG) small signal voltage $v(t)$, on the order of millivolts, by a gain coefficient $\alpha$.

With an inclusive description of the MZI amplitude transfer function established, several important relations must be satisfied to evaluate the corresponding voltages used to generate the SVM simplex. In compact notation, this requires equating the real and imaginary parts of the Jones vector $x$ and $y$ components with the amplitude transfer function of each MZI, as described in the following equations, where $v_0$ is a fixed voltage corresponding to the amplitude of $v(t)$ and $\gamma$ is a unit-less scalar representing the pulse shape and temporal fluctuation of $v_0$.

$$\mathrm{Re}\{E_x\} = \sin\left(\frac{\pi}{2}\frac{\alpha_1}{V_{\pi 1}}\gamma_1 v_{01}\right), \tag{4.7a}$$

$$\mathrm{Im}\{E_x\} = \sin\left(\frac{\pi}{2}\frac{\alpha_2}{V_{\pi 2}}\gamma_2 v_{02}\right), \tag{4.7b}$$

$$\mathrm{Re}\{E_y\} = \sin\left(\frac{\pi}{2}\frac{\alpha_3}{V_{\pi 3}}\gamma_3 v_{03}\right), \tag{4.7c}$$

$$\mathrm{Im}\{E_y\} = \sin\left(\frac{\pi}{2}\frac{\alpha_4}{V_{\pi 4}}\gamma_4 v_{04}\right). \tag{4.7d}$$

The goal is to solve for the normalized voltage amplitudes $\gamma$, which are defined to be limited by the dynamic range of the AWG digital-to-analog converter (DAC), e.g., [-1, +1]. Within the AWG soft panel software, peak-to-peak amplitude settings, $2v_0$, for each channel must be configured before a user-defined waveform may be uploaded. The normalized voltages, $\gamma$, serve as scaling coefficients to the fixed peak voltage amplitude setting $v_0$, ultimately determining the output voltage level for each DAC clock instance.

Transmitter Characterization using Tone FM

Solving the equations in 4.7 for $\gamma$ is a non-trivial task, primarily due to the unknown factor $\frac{\alpha}{V_\pi}$, which varies for each MZI. As shown in Fig. 16, internal linear RF amplifiers with unknown gain transfer functions amplify the small signal outputs $v(t)$ from the AWG to higher voltages for modulation.

Moreover, the MZI transfer functions are not perfectly sinusoidal as assumed, introducing frequency-dependent variations in $V_\pi$ across each MZI. Additionally, the frequency response of each AWG channel output is structured, varying moderately between channels. The combination of these factors can result in poor constellation symmetry of SVM/MVM modulations. Assuming the factor $\frac{\alpha}{V_\pi}$ a priori for each MZI is insufficient for generating

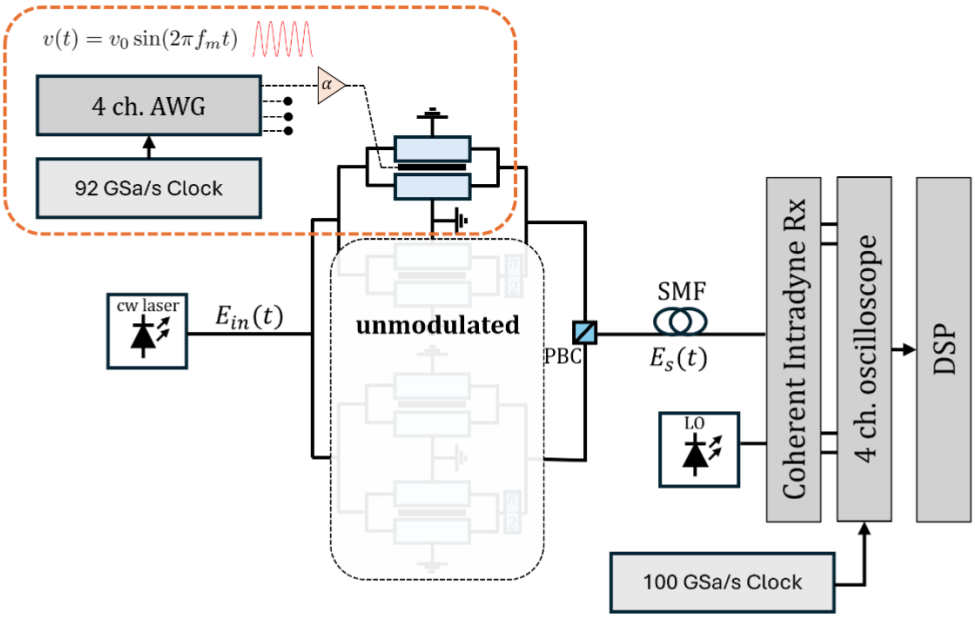



Figure 16: Transmitter characterization experiment block diagram.

highly symmetric simplex SVM/MVM constellations. Therefore, it is necessary to perform coarse characterizations of the transmitter to resolve the arguments of (4.7).

One way to evaluate the ratio $\frac{\alpha}{V_\pi}$ for each of the four MZIs involves independent sinusoidal modulation at a particular frequency of interest, enabling a coarse characterization of the transmitter, as shown within the dotted orange line in Fig. 16. For this purpose, a sinusoidal voltage $v(t)$ is generated using the AWG output corresponding to one of the four MZIs. The voltage $v(t)$ is given by the following equation:

$$v(t) = v_0 \sin(2\pi f_m t), \tag{4.8}$$

where $v_0$ is the amplitude of the sinusoid and $f_m$ is the modulation frequency. The other three MZIs are biased at null and receive no modulation.

A continuous wave (CW) laser with optical frequency $f_0$ and electric field amplitude $E_0$ is used for modulation of the MZI under test. Substituting (4.8) into (4.6), the output

electric field $E_{out}(t)$ is found to be given by the equation:

$$E_s(t) = \frac{E_0}{2} \sin\left[\beta \sin(2\pi f_m t)\right] \cos(2\pi f_0 t), \tag{4.9}$$

which represents a double-sideband suppressed-carrier (DSB-SC) signal with a message signal $m(t) = \sin\left[\beta \sin(2\pi f_m t)\right]$. This type of sinusoidal modulation is analogous to tone frequency modulation (tone FM) [20], where the argument of the MZI transfer function contains an FM index $\beta$, characteristic of the argument in equation (4.6), and defined by the following equation:

$$\beta := \frac{\pi}{2} \frac{\alpha}{V_\pi} v_0. \tag{4.10}$$

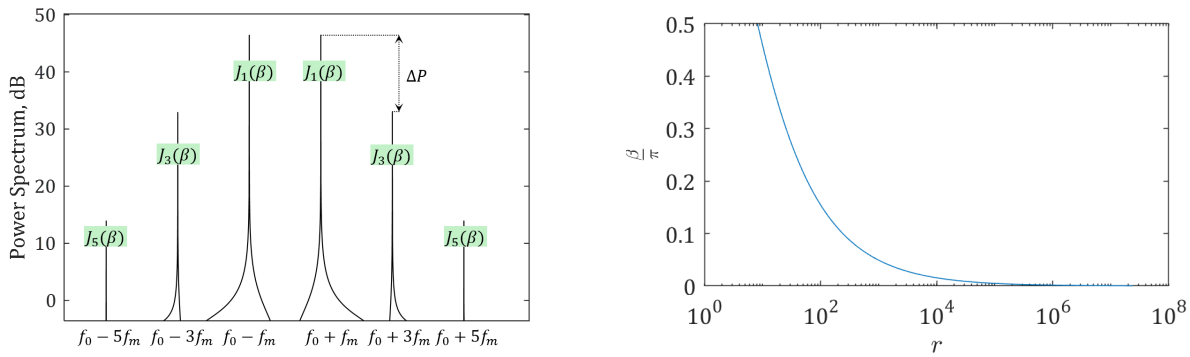
By expanding the message signal $m(t)$ using its Fourier series, it can be shown that $E_s(t)$ may be represented as a sum of sinusoids, corresponding to both the upper and lower sidebands relative to the optical carrier frequency $f_0$ (4.11). Each sinusoid has a unique amplitude determined by the Bessel function of the first kind with argument $\beta$, $J_n(\beta)$, where $n$ is the order of the function. At baseband, the frequency of each sinusoid is an odd multiple of $\pm f_m$. However, when expressed around the frequency $f_0$, the sinusoids are spaced by $f_0 \pm (2n+1) f_m$, where $n$ is an integer.

$$\begin{aligned} E_s(t) &= \frac{E_0}{2} \sum_{n=0}^{\infty} J_{2n+1}(\beta) \sin\left[2\pi f_0 t + 2\pi(2n+1) f_m t\right] \\ &- \sum_{n=0}^{\infty} J_{2n+1}(\beta) \sin\left[2\pi f_0 t - 2\pi(2n+1) f_m t\right]. \end{aligned} \tag{4.11}$$

An optical spectrum analyzer can be used to measure the optical spectrum of the DSB-SC signal, as shown qualitatively in Fig. 17a. The ratio $r$ of the first spectral line to the

second spectral line in the upper sideband is measured from the optical spectrum and used to calculate the value of $\beta$ by numerically solving the equation below:

$$J_1(\beta) = rJ_3(\beta). \tag{4.12}$$




(a) (b)

Figure 17: (a) Computed power spectrum with the corresponding odd-order Bessel functions occurring at multiples of $f_o \pm (2n+1)f_m$; (b) Normalized $\beta/\pi$ plotted as a function of $r$.

The ratio $r$ is estimated using the optical spectrum by first measuring the magnitude difference $\Delta P = P_1 - P_3$ between the frequency components $f_0 + f_m$ and $f_0 + 3f_m$. The ratio is then linearized and given by $r = 10^{\frac{\Delta P}{10}}$. This process can be repeated for each MZI in the transmitter. With a measured FM index $\beta$ for each MZI and prior knowledge of the AWG peak-to-peak voltage setting $2v_0$, the ratio of the RF amplifier gain coefficient to the half-wave voltage, $\frac{\alpha}{V_\pi}$, of each MZI can be inferred.

The goal is to equate both sides of each equation in (4.7), solving for all $\gamma$ coefficients, which will yield the scaled waveforms suitable for uploading to the AWG. For characterization, the modulation frequency $f_m$ is chosen to be 4.6 GHz to ensure an even number of samples per period for generating the sinusoid. Although using an optical spectrum analyzer to detect the optical spectrum and determine the modulation index $\beta$

may be adequate with high-resolution sampling, the Yokogawa AQ6370D analyzer deployed in the lab is limited to a resolution of only 0.02 nm, which, with a laser center wavelength $\lambda_0 = 1550$ nm, corresponds to a frequency resolution of approximately 2.5 GHz. Frequency sampling at the order of $f_m$ may lead to inaccurate power measurements in the optical spectrum. Therefore, the coherent receiver is used to bring the signal to baseband, where it can be sampled at a rate much higher than $f_m$. While detecting the signal with the coherent receiver introduces more complexity in terms of detection impairments and noise, several DSP algorithms can be employed to compensate for these impairments and provide a more accurate measurement than the optical spectrum analyzer.

As shown in the experimental setup in Fig. 16, a single MZI is driven by a sinusoidal tone with frequency $f_m = 4.6\,\text{GHz}$ and peak-to-peak amplitude $2v_0$. The internal RF amplifier increases the amplitude of the small-signal AWG voltage by the gain factor $\alpha$. The signal is transmitted via a short fiber jumper and detected by the coherent receiver with additional blocks representing analog-to-digital conversion and subsequent DSP processing.

For the received signal to be processed by DSP functions, including hardware impairment compensation, clock recovery, SOP estimation, and phase estimation, it must be sampled on a symbol-by-symbol basis (Fig. 18).

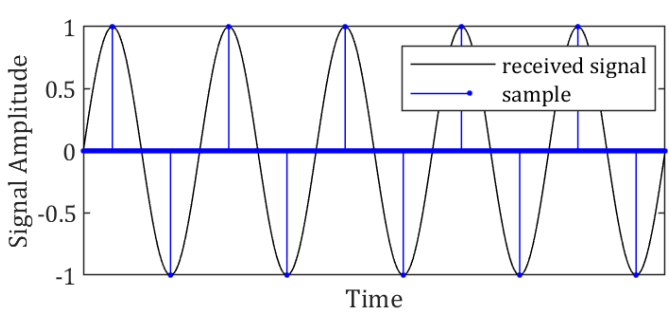



Figure 18: Qualitative BPSK scheme with $R_s = 2f_m$.

After reception, the signal components are sampled by the oscilloscopes at a rate of

$f_s = 100\,\text{GSa/s}$. The received sinusoid is treated as a coherent binary phase shift keying (BPSK) signal with symbol rate $R_s = 2f_m$, and is interpreted by the DSP as a binary waveform with $+1$ or $-1$ amplitude states. In this setup, the DSP processes the signal by taking 2 samples per period of the sinusoid, using these samples for all downstream processing.

The DSP block in Fig. 16 can be expanded into its constituent functions, including hardware impairment compensation, clock recovery, SOP estimation, phase estimation, interpolation, the Fast Fourier Transform, and, ultimately, resolving the modulation index $\beta$ (Fig. 19).

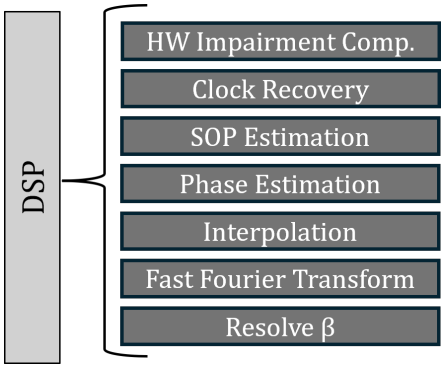



Figure 19: Transmitter characterization DSP.

The hardware impairment compensation corrects for 90° hybrid imbalances and inter-channel skew [34]. The clock recovery function determines the symbol clock frequency and phase, generating a time-series Jones vector spaced by the symbol period $T_s$. Following clock recovery, the sampled Jones vectors may be visualized using the complex plane (Fig. 20a).

Fiber birefringence causes some power in $x$ to leak into $y$, while temporal decoherence between the signal and LO lasers leads to an intermediate frequency (IF) offset, causing the points to rotate. The SOP estimation restores all of the signal power into the $x$-polarization [34] (Fig. 20b). Thereafter, the phase estimator removes the IF offset and subtracts laser phase noise, producing the desired BPSK symbols [33] (Fig. 20c).

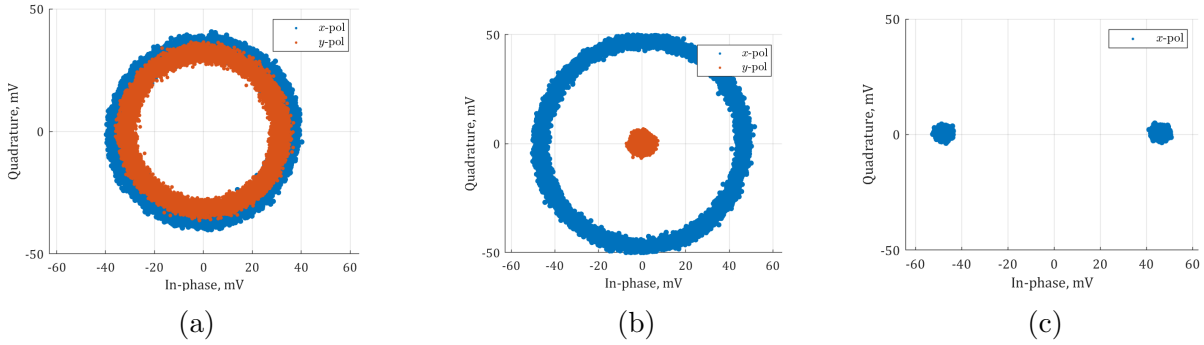



Figure 20: Experimental constellations of BPSK processing through various DSP functions. Received symbols plotted (a) after clock recovery and symbol downsampling; (b) after performing SOP estimation; (c) after removing the IF offset and performing phase estimation.

BPSK symbols aligned on the real axis are interpolated to reconstruct the received waveform with an effective sample rate significantly higher than $f_m$ [34]. An FFT then reveals the modulation index $\beta$ (Fig. 21).

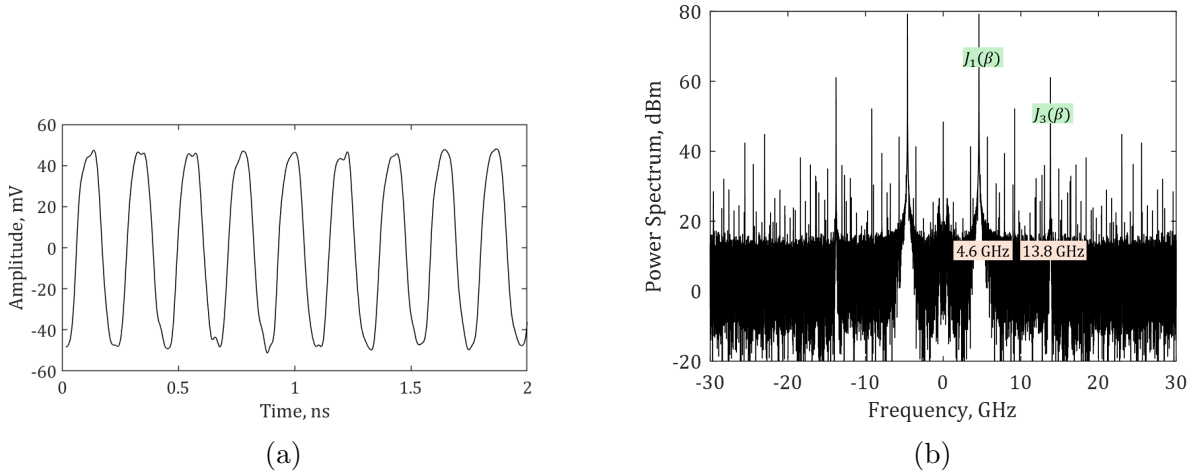



Figure 21: Final steps for recovering the modulation index $\beta$. (a) Interpolated BPSK signal representing the received sinusoid; (b) RF spectrum after taking an FFT of the signal.

Once the RF spectrum is obtained, the modulation index $\beta$ can be calculated using equation (4.12). The experimental results are shown in Fig. 22, where the modulation

indices $\beta$ are plotted against the AWG driving peak-to-peak voltage setting, $2v_0$, for each MZI modulator (1-4).

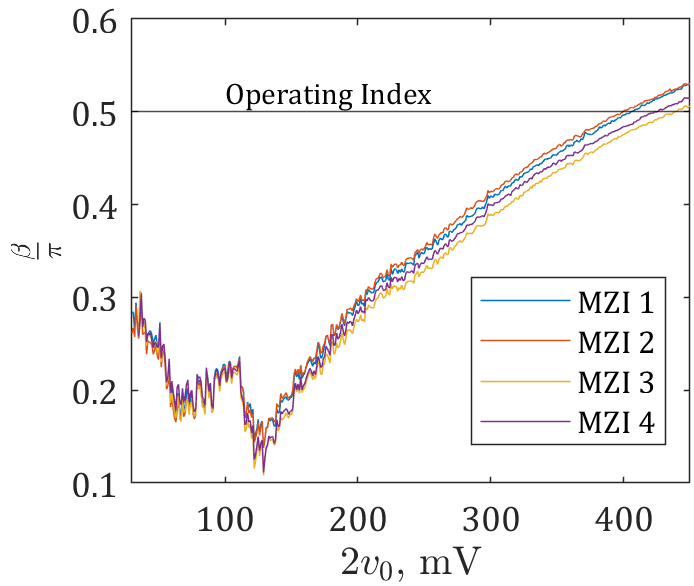



Figure 22: Normalized modulation index $\beta/\pi$ plotted against AWG peak-to-peak voltage setting $2v_0$ for each MZI in the transmitter.

Once the modulation index $\beta$ is determined, an operating index can be selected from Fig. 22 and used to compute the scalar coefficients $\gamma$. As stated in (4.7) and (4.10), $\beta$ represents the full argument of each MZI transfer function, leaving only $\gamma$ to be solved for. Rewriting (4.7), the following set of equations links the MZI transfer functions to the real and imaginary parts of each Jones vector component in the simplex SVM symbol set:

$$\mathrm{Re}\{E_x\} = \sin(\beta\gamma_1), \tag{4.13a}$$

$$\mathrm{Im}\{E_x\} = \sin(\beta\gamma_2), \tag{4.13b}$$

$$\mathrm{Re}\{E_y\} = \sin(\beta\gamma_3), \tag{4.13c}$$

$$\mathrm{Im}\{E_y\} = \sin(\beta\gamma_4). \tag{4.13d}$$

Choosing an operating value for $\beta$ is somewhat arbitrary; however, larger values of $2v_0$ are preferred due to their ability to improve the signal-to-noise ratio (SNR). Therefore, the operating index is set to $\beta = \frac{\pi}{2}$, which approximates $\alpha v_0 \approx V_\pi$. With $\beta = \frac{\pi}{2}$, each MZI must be driven by a different value of $2v_0$, as illustrated in Fig. 22. These voltage settings are summarized in the table below.

| **Modulator** | $\mathbf{2v_0}$ |
|---|---|
| MZI 1 | 406 mV |
| MZI 2 | 400 mV |
| MZI 3 | 439 mV |
| MZI 4 | 424 mV |

Table 3: Modulator driving voltages for fixed $\beta = \frac{\pi}{2}$.

Evaluating the AWG Waveforms

Waveform generation follows the coarse characterization of the transmitter, assuming a fixed modulation index $\beta = \frac{\pi}{2}$ for each MZI. The generation of the AWG waveforms begins by evaluating the equations (4.13) for the scaling coefficients $\gamma$ for each MZI and constellation symbol. Using the symbol Jones vectors from (4.3), Table 4 depicts the corresponding normalized voltages $\gamma$ for each Jones symbol vector $|s\rangle$ of the SVM simplex.

After the scaling coefficients $\gamma$ are computed, attention can be turned to the type of symbol sequencing that is used in the experiment. Many sequencing algorithms apply to various scenarios; however, for direct-detection systems, impairments caused by chromatic dispersion (CD) are of primary interest since these systems cannot electronically compensate for CD.

| **Scaling Coefficient** | $\|\mathbf{s_1}\rangle$ | $\|\mathbf{s_2}\rangle$ | $\|\mathbf{s_3}\rangle$ | $\|\mathbf{s_4}\rangle$ |
|---|---|---|---|---|
| $\gamma_1$ | 1.0000 | 0.0000 | 0.0000 | 0.0000 |
| $\gamma_2$ | 0.3918 | 0.0000 | 0.6082 | 0.0000 |
| $\gamma_3$ | 0.3918 | 0.0000 | -0.2677 | 0.5000 |
| $\gamma_4$ | 0.3918 | 0.0000 | -0.2677 | -0.5000 |

Table 4: Normalized voltage mapping between the $\gamma$ coefficients used to drive each MZI and the SVM simplex Jones symbol vectors.

Due to the CW laser's moderately polychromatic spectrum, CD results in a temporal spreading of the optical pulse during propagation. As the pulse spreads, interference occurs between adjacent pulses, leading to Inter-symbol Interference (ISI) [16].

Certain symbol sequencing techniques can help mitigate the impact of ISI by ensuring all combinations of adjacent symbols in the constellation appear in the sequence. One such sequence commonly used in communications is the de Bruijn sequence [6]. By carefully choosing adjacent symbols, de Bruijn sequences ensure a balanced distribution of symbol combinations, thereby reducing the likelihood of interference between neighboring symbols and minimizing ISI-induced degradation. Each de Bruijn sequence is defined by an alphabet $A$ of length $k$ and order $n$. The order $n$ refers to a subsequence in which every possible string of length $n$ from the alphabet $A$ appears at least once. For example, as shown in Fig. 23, the sequence {0,0,1,1} is generated from all possible subsequences of length 2 taken from an alphabet consisting of symbols 0 and 1. This sequence has an alphabet length of $k = 2$ and order $n = 2$.

Since the simplex SVM constellation alphabet consists of four symbols, the de Bruijn sequence will exhibit an alphabet length $k = 4$. In this experiment, since a short fiber jumper is used, CD is assumed to be negligible. Therefore, an order of $n = 5$ is chosen, which is sufficient for this initial back-to-back experiment. Using a de Bruijn generator in

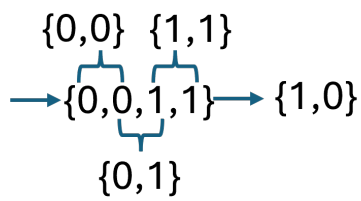



Figure 23: Example of a de Bruijn sequence with $k = n = 2$.

MATLAB [17], the following de Bruijn symbol sequence $\mathbf{x}$ of length $4^5 = 1024$ is generated as the following:

$$\mathbf{x} = [4, 3, 2, 2, 2, 3, 1, 1, 4, ...]. \tag{4.14}$$

Including the symbol sequence and symbol vectors $|s\rangle$ from the simplex SVM constellation, along with the corresponding scaling coefficients $\gamma$ from Table 4, the waveforms are generated and adapted to the desired symbol rate $R_s$. The symbol rate is given as the ratio of the DAC clock rate of the AWG, $f_{\text{clk}}$, to the number of samples per symbol, $N_{\text{sps}}$.

For this experiment, the transmitter characterization was performed with modulation frequency $f_m = 4.6$ GHz, which coarsely characterizes the transmitter for a symbol rate $R_s = 4.6$ GBd. With a DAC clock rate of 92 GHz, each coefficient $\gamma$ is repeated 20 times ($N_{\text{sps}} = 20$). The de Bruijn sequence is used to generate plots of the simplex SVM waveforms for each MZI, shown in Fig. 24, up to the first 2000 samples. The symbols are shaped by square pulses, with amplitudes corresponding to the scaling coefficients $\gamma$. The amplitudes of each waveform are expressed as factors of $V_\pi$, since the modulation index is taken as $\beta = \frac{\pi}{2}$.

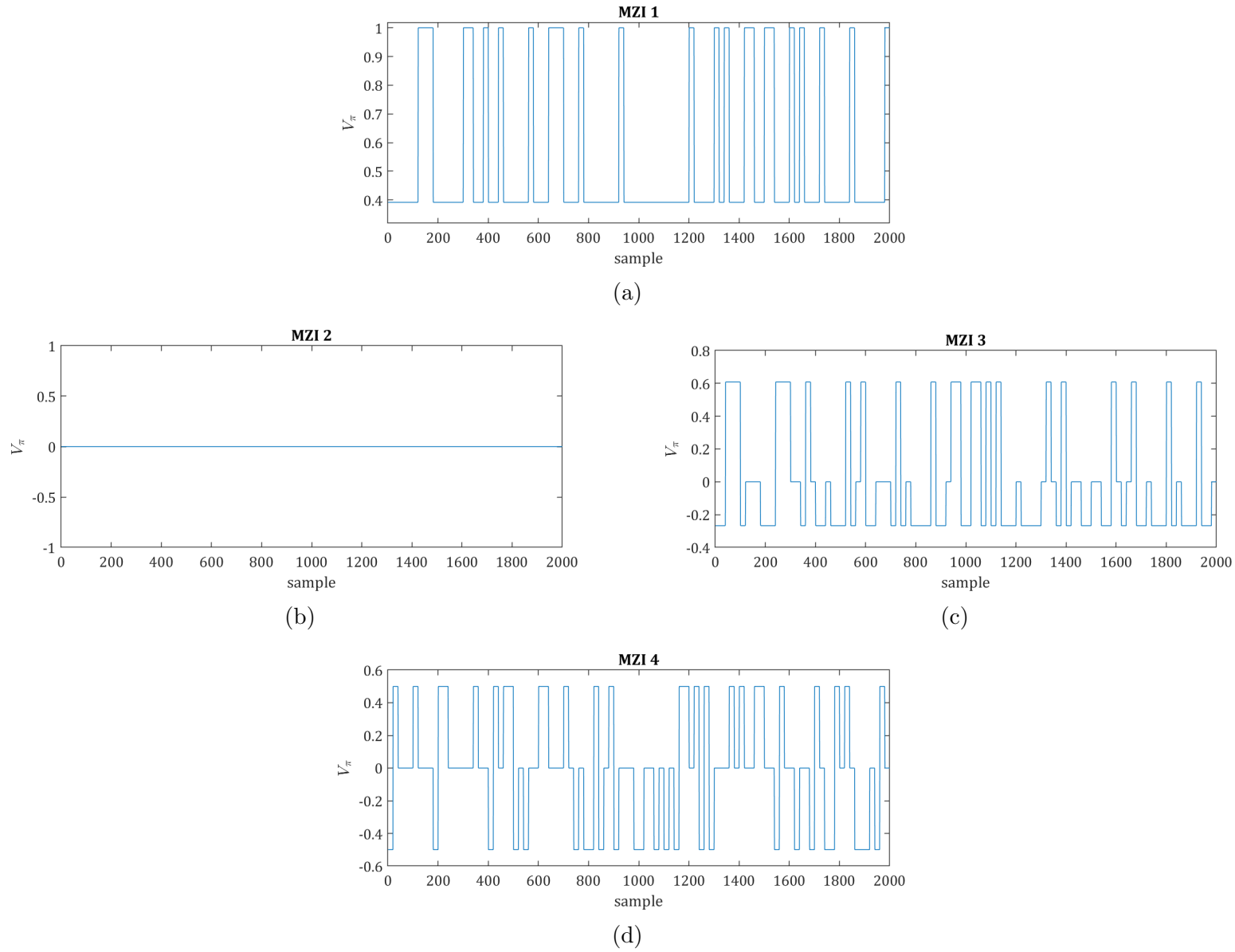



Figure 24: Plots of waveforms to be uploaded to the AWG for generation of the simplex SVM constellation. Figures include waveform plots up to 2000 samples for modulators controlling the (a) $x$-pol in-phase; (b) $x$-pol quadrature; (c) $y$-pol in-phase; (d) $y$-pol quadrature.

Experimental Generation of the SVM Simplex Constellation

The signal waveforms, as shown in Fig. 24, are uploaded to the AWG. The peak-to-peak voltage for each AWG channel, $2v_0$, is set according to Table 3 to achieve a fixed modulation index of $\beta = \frac{\pi}{2}$ across all MZIs. As depicted in Fig. 25, each AWG channel voltage modulates the MZIs corresponding to the real and imaginary parts of the Jones vector polarization components, thereby modulating the SOP of the incoming CW laser's electric field $E_{in}(t)$.

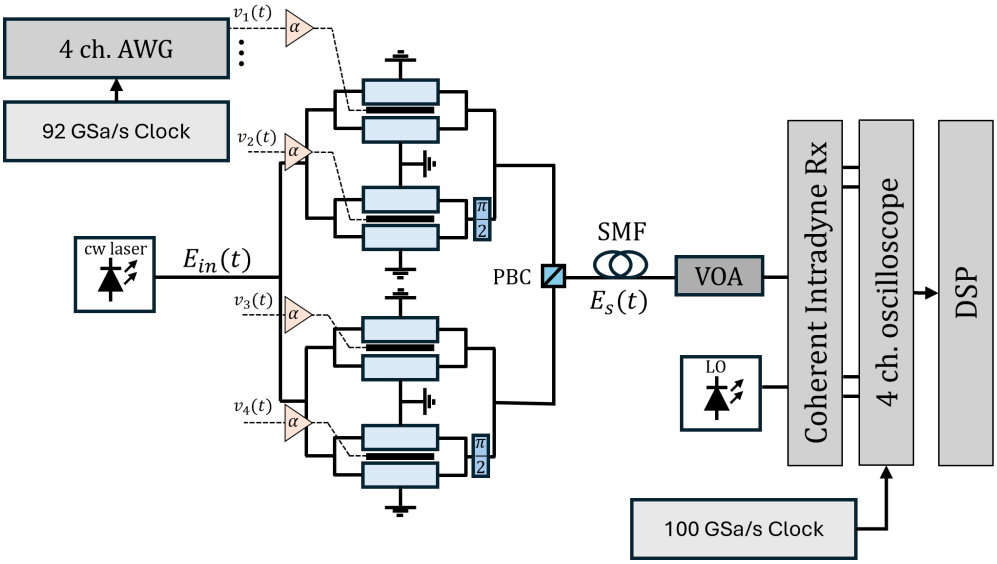



Figure 25: The SVM simplex experimental block diagram.

The modulated signal then propagates through a short fiber jumper, where it acquires random and time-varying birefringence, as well as controlled attenuation from the variable optical attenuator (VOA). It is then detected by the intradyne coherent receiver, where several offline DSP functions are performed to assess the constellation symmetry and evaluate the BER. The DSP functions are outlined in Fig. 26 and are discussed in detail in this section.

Identical to several DSP functions used in the previous section for transmitter characterizations, the hardware impairment compensation and clock recovery functions are performed as part of the DSP chain. These compensations correct 90° hybrid amplitude

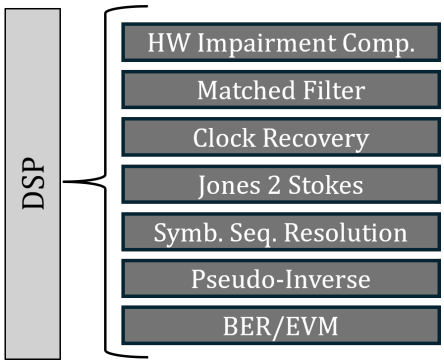



Figure 26: DSP functions for the SVM simplex including hardware impairment compensation, matched filter, clock recovery, Jones to Stokes conversion, symbol sequence resolution, pseudo-inverse, and evaluations of the BER and error vector magnitude (EVM).

and phase imbalances and remove inter-channel skew. The clock recovery function uses the baseband spectrum to determine the symbol clock frequency and phase, outputting a time-series Jones vector sampled at $R_s$. The other functions listed in Fig. 26 are internally developed.

Matched Filter

After applying hardware impairment compensations to the received waveforms from the four-channel oscilloscope, a matched filter is used to remove excess noise from the received signal. Since the driving waveforms are shaped by square pulses, the matched filter impulse response is the moving average filter with a duration equal to the number of samples per symbol $N_{\text{sps}}$.

Since the symbol rate is $R_s = 4.6$ GBd and the oscilloscope sample rate is configured to 100 GSa/s, the moving average filter is approximated with $N_{\text{sps}} \approx 20$. The power spectrum of both the raw and filtered waveforms for the $x$-polarization in-phase channel is shown in Fig. 27. Each channel is filtered using the same impulse response.

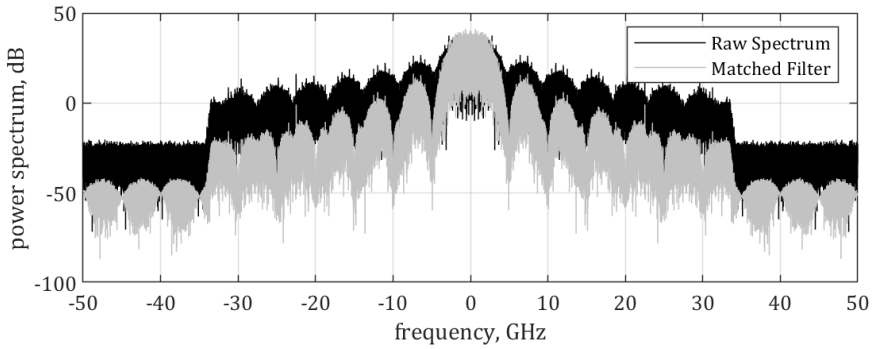



Figure 27: Simplex SVM matched filter implemented on the $x$-pol in-phase channel.

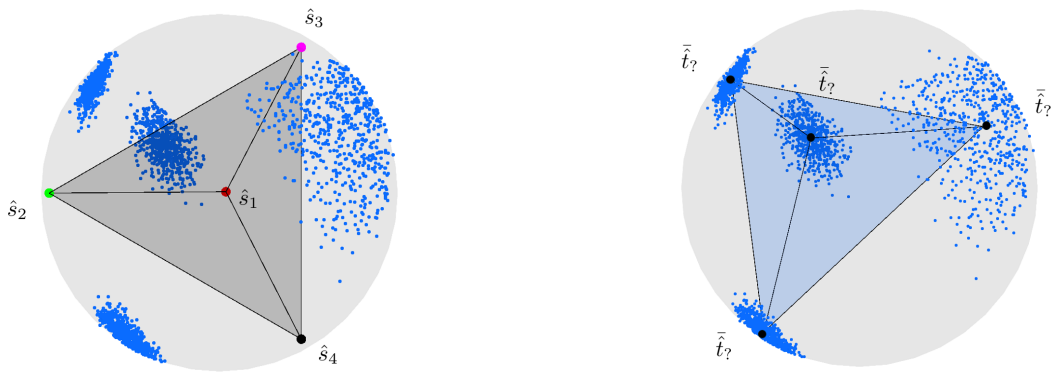
Jones to Stokes Conversion

Once the clock recovery algorithm determines the symbol clock and phase, the time-series Jones vectors, $\mathbf{J} = [|s(t_1)\rangle, |s(t_2)\rangle, |s(t_3)\rangle, ..., |s(t_N)\rangle]$, sampled at $R_s$, are converted into their corresponding Stokes vector representations to emulate true direct detection receiver processing. At this point, the absolute phase information provided by the coherent receiver is lost to the DSP. The conversion from time-series Jones vectors to time-series Stokes vectors is performed using (4.4) and (4.5).

Symbol Sequence Resolution

Once the received time-series Stokes vectors, $\mathbf{T} = \left[\hat{t}(t_1), \hat{t}(t_2), \hat{t}(t_3), ..., \hat{t}(t_N)\right]$, are computed using the quadratic form, they are normalized to unit magnitude and can be plotted on the surface of the Poincaré sphere, shown as the blue points in Fig. 28a. Due to the birefringence of the short fiber cable, the received vector clusters in blue form a tetrahedron that is rigidly rotated away from the nominal transmitted Stokes vector positions, forming the vertices of the shaded black regular tetrahedron. The primary task of the SVM DSP is to counter-rotate the constellation back to the nominal positions, allowing symbol decisions to be made and, eventually, the BER to be evaluated.

Since the DSP is performed offline, the real-time oscilloscopes must capture a single



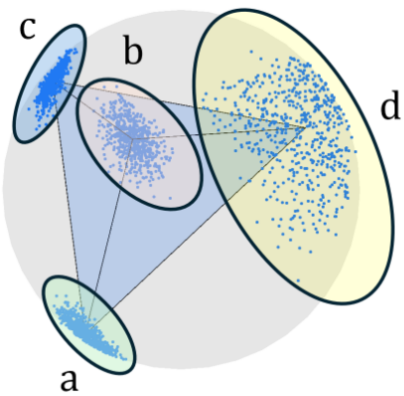

(a) (b)

Figure 28: (a) Received Stokes vectors rotated away from the nominal Stokes symbols forming the regular tetrahedron; (b) tetrahedron formed by the cluster averages of received vectors with unknown symbol vertices.

acquisition of samples from the continuously incoming signal. The acquisition record length is kept sufficiently long to capture at least two full $4^5$ de Bruijn sequence lengths (in practice, $N \approx 2{,}300$ symbols). Because the timing of the acquisitions is asynchronous with the transmitter, there is a symbolic delay to the known transmitted sequence. Therefore, symbol sequence resolution is required, as the future derotation step based on the pseudo-inverse depends on knowledge of the received symbol sequence, which is inherently unknown to the DSP (Fig. 28b).

Without knowledge of the fiber Mueller matrix $\mathbf{R}$ and due to the symmetry of the constellation, each symbol cluster average vector $\bar{\hat{t}}_i$ could correspond to any one of the four symbol vertices $\hat{s}_i$. Thus, both the received symbol sequence and symbol ambiguity must be resolved to perform a derotation of the constellation. Using the MATLAB functions `linkage()` and `cluster()`, received Stokes vectors are clustered together and assigned indices {a, b, c, d} arbitrarily (Fig. 29a). Since any cluster may correspond to any Stokes symbol in the SVM simplex, the cross-correlation $\rho$ between a repeated transmitted sequence

$\mathbf{x}'$ and all permutations of the received sequence $\mathbf{y}_n$, where $n$ is the permutation index including 4! permutations, is evaluated (Fig. 29b).




(a)

| **permutation (n)** | **a** | **b** | **c** | **d** |
|---|---|---|---|---|
| 1 | 4 | 3 | 2 | 1 |
| 2 | 3 | 4 | 1 | 2 |
| 3 | 1 | 3 | 2 | 4 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 4! | 1 | 2 | 3 | 4 |

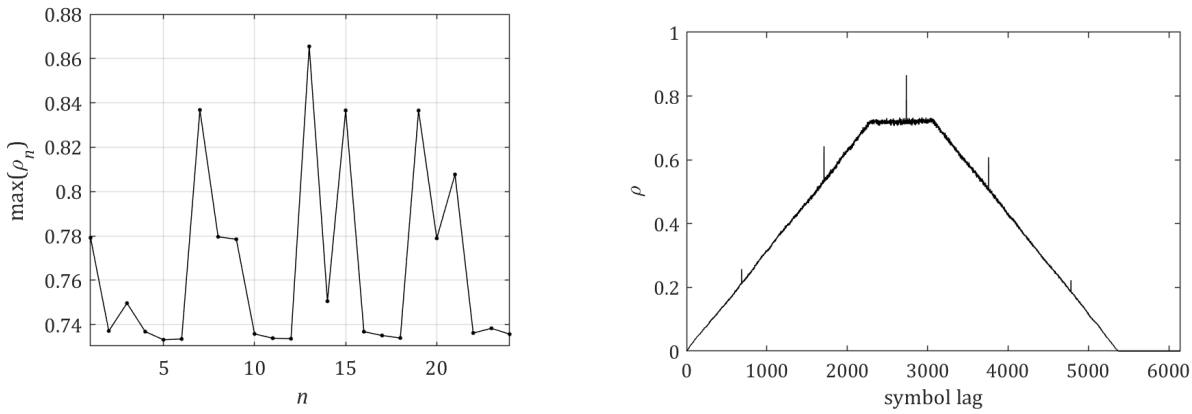
(b)

Figure 29: (a) Clustered groups of received Stokes vectors with labels {a, b, c, d}; (b) permutations of {a, b, c, d} corresponding to symbol vertices.

The length of the cross-correlation vector $\rho$ is $2K - 1$, where $K$ represents the length of the repeated transmitted sequence vector $\mathbf{x}'$. This ensures that the correlation sequence is long enough to correlate all $N$ symbols in the acquisition record. For example, assuming a $4^5$ de Bruijn sequence, the length $K$ is given by $K = 1024 \times \epsilon$, where $\epsilon$ is the repetition factor.

For the case where $N = 2300$, the repetition factor $\epsilon = 3$. Thus, the transmitted sequence $\mathbf{x}'$ is formed by repeating the array $\mathbf{x}$ three times: $\mathbf{x}' = [\mathbf{x}, \mathbf{x}, \mathbf{x}]$. The correct permutation sequence $\mathbf{y_n}$ will produce the highest correlation peak with $\mathbf{x}'$.

As shown in Fig. 30a, the peak correlation of $\rho_n$ occurs when $n = 13$, though it is less than 1 due to excess symbols in $K > \mathbf{y_n}$, as seen in Fig. 30b. Once the correct permutation sequence $\mathbf{y}_{13}$ is identified, the transmitted time-series Stokes symbols $\mathbf{S} = [\hat{s}(t_1), \hat{s}(t_2), \hat{s}(t_3), ..., \hat{s}(t_N)]$ are defined using the symbol set from (4.3), converted to Stokes vectors, and $\mathbf{y}_{13}$ as the symbol indices.




(a) (b)

Figure 30: (a) Peak correlations of $\rho_n$ plotted against $n$; (b) Cross-correlation between the permuted sequence $\mathbf{y_{13}}$ and the known repeated transmitted sequence $\mathbf{x}'$.

Pseudo-Inverse

Naturally, one might assume that the Mueller matrix $\mathbf{R}$ could be determined through a straightforward matrix inversion operation. However, the dimensionality of each transmitted Stokes vector $\hat{s}$ implies that the inverse does not exist. As a result, the resolved time-series transmitted and received Stokes vectors are used to form the following matrices:

$$\mathbf{T} := \begin{bmatrix} | & & | \\ \hat{t}(t_1) & \dots & \hat{t}(t_N) \\ | & & | \end{bmatrix}, \quad \mathbf{S} := \begin{bmatrix} | & & | \\ \hat{s}(t_1) & \dots & \hat{s}(t_N) \\ | & & | \end{bmatrix}. \tag{4.15}$$

Combining the matrices, the fiber rotation $\hat{t}(t) = \mathbf{R}\hat{s}(t)$ is rewritten in a more compact form:

$$\mathbf{T} = \mathbf{RS}, \tag{4.16}$$

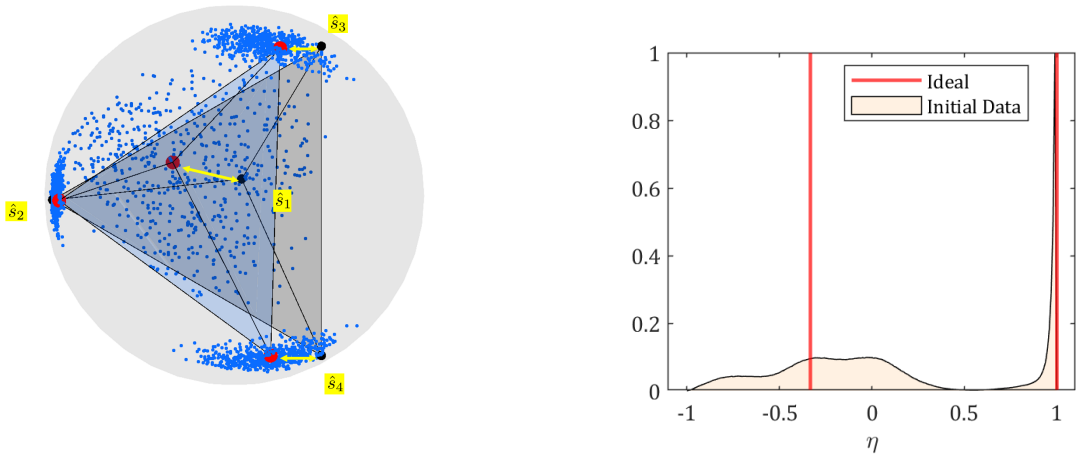
where the dimensions of matrices are $3 \times N$, $3 \times 3$, and $3 \times N$, respectively. Then, an

estimation of the Mueller matrix $\mathbf{R}$ is made using the pseudo-inverse:

$$\mathbf{R} = \mathbf{T}\mathbf{S}^T(\mathbf{S}\mathbf{S}^T)^{-1}. \tag{4.17}$$

Note that $\mathbf{S}\mathbf{S}^T$ is invertible provided that $\mathbf{S}$ is full rank. Since $N > 3$, this condition is easily satisfied. Assuming the fiber rotation is subsequently characterized via the estimated Mueller matrix $\mathbf{R}$, the received symbols $\hat{t}(t)$ at the receiver can be derotated by evaluating $\mathbf{R}^{-1}\hat{t}(t) \approx \hat{s}(t)$ in the DSP.

By implementing the pseudo-inverse on the received symbols shown in Fig. 28a, the Mueller matrix $\mathbf{R}$ is recovered. Evaluating $\mathbf{R}^{-1}\hat{t}(t) \approx \hat{s}(t)$ and re-normalizing the Stokes vectors to unit magnitude, the rotated constellation is found and depicted in Fig. 31a.




(a) (b)

Figure 31: (a) Derotated SVM simplex constellation; (b) histogram of inner products of received Stokes vectors.

The pseudo-inverse derotates the received Stokes vectors $\hat{t}(t)$ to positions that minimize the separation between each cluster average $\bar{\hat{t}}_i$ and the corresponding transmitted Stokes symbols $\hat{s}_i$. By computing the averages of each cluster and plotting them in red, the deviation

between the cluster averages and the nominal Stokes symbol vertices can be observed, emphasized by the yellow arrows. This deviation is due to the sub-optimal shape of the received tetrahedron formed by the cluster averages. Furthermore, to quantify constellation symmetry, the histogram of inner products, shown in Fig. 31b, is used. The inner product $\eta$ is defined as:

$$\eta = \hat{t}_i \cdot \hat{t}_j. \tag{4.18}$$

For an ideal (regular) tetrahedron corresponding to the SVM simplex, this is evaluated to $\eta = 1$ $(i = j)$ or $\eta = -\frac{1}{3}$ $(i \neq j)$. As observed in the histogram, there is a strong distribution around 1, indicating the large number of received Stokes vectors that underwent an inner product with themselves. However, the relatively broader distribution around $-\frac{1}{3}$ indicates poor constellation symmetry. This broadened distribution results from the non-ideal shape of the received tetrahedron and noise.

Bit Error Rate Evaluation and Error Vector Magnitude

After estimating $\mathbf{R}$ using the pseudo-inverse, the derotated received Stokes vectors are denoted as $\tilde{\mathbf{T}} = \left[\tilde{\hat{t}}(t_1), \tilde{\hat{t}}(t_2), \tilde{\hat{t}}(t_3), ..., \tilde{\hat{t}}(t_N)\right]$. The final step in the DSP involves calculating the BER. To make symbol decisions, maximum likelihood estimation (MLE) is applied, defined by:

$$\hat{\hat{s}} := \text{argmax}\left(\tilde{\hat{t}} \cdot \hat{s}\right), \hat{s} \in \hat{\mathbf{S}}, \tag{4.19}$$

where $\hat{\mathbf{S}} = [\hat{s}_1, \hat{s}_2, \hat{s}_3, \hat{s}_4]$ is the nominal Stokes vector matrix containing the transmitted Stokes vectors. The value $\hat{\hat{s}}$ corresponds to the estimated symbol decision Stokes vector that produced the greatest inner product with $\tilde{\hat{t}}(t_i)$. Following MLE for each received symbol, the symbol decision Stokes vector matrix $\tilde{\mathbf{T}}' = \left[\hat{\hat{s}}(t_1), \hat{\hat{s}}(t_2), \hat{\hat{s}}(t_3), ..., \hat{\hat{s}}(t_N)\right]$ is formed.

Since the symbol vertices forming the SVM simplex are equidistant, the bit encoding can be arbitrarily assigned without introducing additional penalty. The bit encoding used in the experiment is summarized in Table 5.

| **Symbol** | **Bit Sequence** |
|---|---|
| $\hat{s}_1$ | 00 |
| $\hat{s}_2$ | 01 |
| $\hat{s}_3$ | 10 |
| $\hat{s}_4$ | 11 |

Table 5: SVM simplex bit-to-symbol mapping.

Both the transmitted and received symbol matrices, $\mathbf{S}$ and $\tilde{\mathbf{T}}'$, respectively, are converted to their corresponding binary sequences and flattened into one-dimensional vectors. The number of errors $N_{\text{errs}}$ is used to calculate the number of bit positions that differ between the transmitted and received bit sequences. Finally, the BER is computed as the ratio of the number of errors $N_{\text{errs}}$ to the total number of bits $N_{\text{bits}}$:

$$\text{BER} := \frac{N_{\text{errs}}}{N_{\text{bits}}}. \tag{4.20}$$

Using an in-line VOA, the BER is measured for varying levels of average received optical power, expressed in units of dBm (Fig. 32). Adopting the KP4 forward error correction (FEC) threshold of $2.4 \times 10^{-4}$, it is found that the initial generation of the simplex SVM constellation requires approximately -13.5 dBm of received power to reach the KP4 FEC limit.

As a final figure of merit to quantify constellation noise, the EVM is defined based on the distance of the Stokes vectors in each cluster from the average normalized vector $\bar{\hat{t}}_i$ within the cluster:

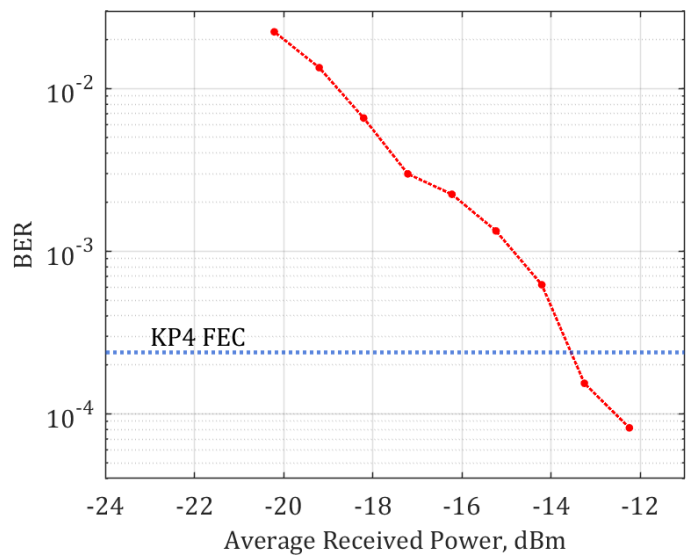



Figure 32: BER versus average received optical power for the initial generation of the SVM simplex constellation.

$$\text{EVM} := \sqrt{\left\langle \left\| \hat{t} - \bar{\hat{t}}_i \right\|^2 \right\rangle}, \tag{4.21}$$

where the angled brackets indicate averaging across the cluster. Cluster noise can be qualitatively visualized by projecting the 3-dimensional Poincaré sphere onto the $yz$-plane, as shown in Fig. 33. From the figure, it is observed that the computed EVM is approximately 0.36. In the following section, an optimization procedure is proposed to improve both the constellation symmetry and EVM of the simplex SVM constellation.

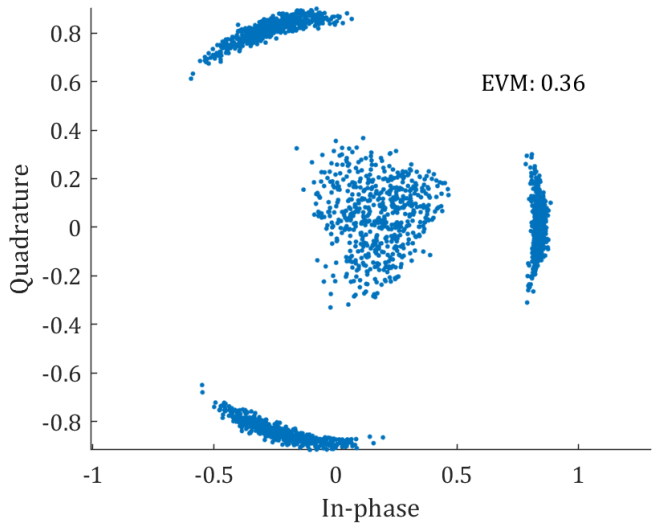



Figure 33: The SVM simplex plotted in the $yz$-plane with computed EVM $\approx 0.36$.

SVM Simplex Optimization using the Steepest Ascent Procedure

Sub-optimal generations of the SVM simplex constellation are attributed to the high sensitivity of modulator driving voltages to each MZI inside the coherent transmitter. The received constellation can be leveraged to optimize these driving voltages, thereby generating SVM/MVM constellations with enhanced symmetry and reduced noise.

To improve the modulator voltages for the SVM simplex, the volume of the received tetrahedron in Stokes space is used as an optimization metric. This is based on the principle that the largest-volume tetrahedron inscribed in a sphere is a regular one. By applying a recursive process using the steepest ascent procedure, the volume of the tetrahedron is maximized, leading to improved constellation characteristics.

Consider three non-collinear vectors, $\vec{a}$, $\vec{b}$, and $\vec{c}$, in 3D Stokes space, as shown in Fig. 34. The figure depicts a parallelepiped spanned by the vectors $\vec{a}$, $\vec{b}$, and $\vec{c}$, where the three adjacent sides of the parallelepiped, denoted as $P$, are given by these vectors. The volume of the parallelepiped is determined by the absolute value of the scalar triple product of the

vectors:

$$\text{volume}(P) = \left|(\vec{a} \times \vec{b}) \cdot \vec{c}\right|. \tag{4.22}$$

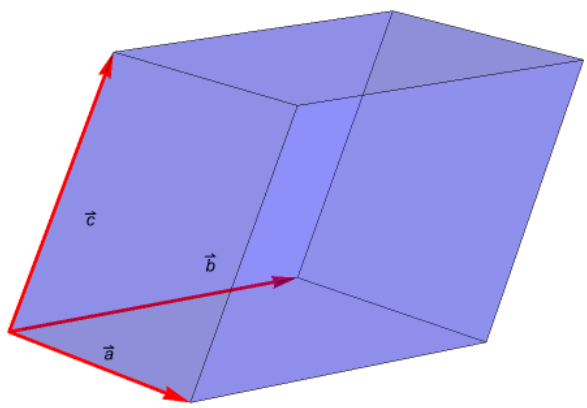



Figure 34: Parallelepiped spanned by $\vec{a}, \vec{b}, \vec{c}$.

Using the formula for the cross product in component form, which is given by the determinant, the evaluation of $\vec{a} \times \vec{b}$ is given by the following:

$$\vec{a} \times \vec{b} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}, \tag{4.23}$$

where $\hat{i}, \hat{j}, \hat{k}$ denote the unit vectors in 3D Stokes space. Using (4.23), the scalar triple

product in (4.22) is given below in component form:

$$(\vec{a} \times \vec{b}) \cdot \vec{c} = \begin{vmatrix} c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}. \tag{4.24}$$

Now, consider a tetrahedron with edges $\vec{a}$, $\vec{b}$, and $\vec{c}$. The volume of this tetrahedron is given by one-sixth of the volume of the parallelepiped spanned by the same vectors. By imagining a tetrahedron inscribed within the unit Poincaré sphere, as illustrated in Fig. 35, let the position vectors of the vertices of the tetrahedron be denoted by $\hat{s}_i$ for $i = 1, \ldots, 4$ in green. The edges of the tetrahedron, denoted $\vec{d}_1$, $\vec{d}_2$, and $\vec{d}_3$, are defined as $\vec{d}_i := \hat{s}_i - \hat{s}_4$ for $i = 1, 2, 3$ depicted in red.

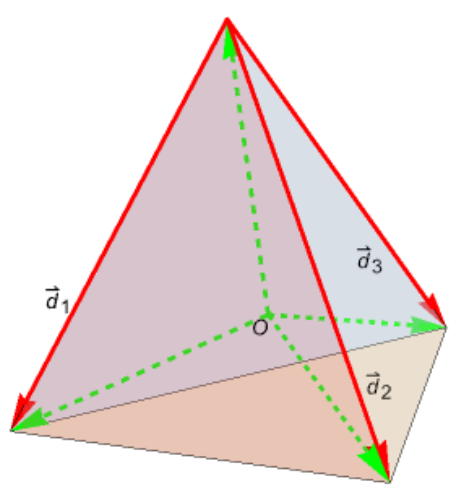



Figure 35: Tetrahedron inscribed in the unit Poincaré sphere (not shown). Notation: Position vectors of the four vertices, $\hat{s}_i$, $i = 1, \ldots, 4$ (in green); Three edges $\vec{d}_i$, $i = 1, \ldots, 3$ (in red).

The three edge vectors $\vec{d}_1, \vec{d}_2, \vec{d}_3$ of the tetrahedron can be arranged into the matrix $\mathbf{D} := [\vec{d}_1, \vec{d}_2, \vec{d}_3]^T$. Based on the determinant expression in (4.24), the volume of the tetrahedron

is given by:

$$V = \frac{1}{6}|\det(\mathbf{D})|. \tag{4.25}$$

Since the leading coefficient $\frac{1}{6}$ does not impact the maximization of the volume, it can be neglected. Consequently, the volume expression serves as the cost function for the optimization process, which is based on the square of the determinant:

$$\xi := \det(\mathbf{D})^2. \tag{4.26}$$

Using properties of determinants, the cost function can be rewritten as:

$$\xi = \det(\mathbf{G}), \tag{4.27}$$

where $\mathbf{G} := \mathbf{D}\mathbf{D}^T$ is the Gram matrix. This Gram matrix can be expanded element-wise as the dot product between the edge vectors $\vec{d}$:

$$\begin{aligned} \mathbf{G}_{ij} &= \vec{d_i} \cdot \vec{d_j} = (\hat{s}_i - \hat{s}_4) \cdot (\hat{s}_j - \hat{s}_4) \\ &= (\hat{s}_i \cdot \hat{s}_j) - (\hat{s}_i \cdot \hat{s}_4) - (\hat{s}_j \cdot \hat{s}_4) + 1. \end{aligned} \tag{4.28}$$

For the generation of the SVM symbols, two quadrature modulators are employed in parallel, one controlling the $x$ SOP and the other controlling the $y$ SOP. Ultimately, the cost function needs to be expressed in terms of Jones vectors, as the modulator voltages directly control the components of these vectors. Therefore, it is advantageous to operate solely within Jones space.

The dot product of Stokes vectors in terms of the dot product of Jones vectors can be

expressed by the following:

$$\hat{s}_i \cdot \hat{s}_j = 2|\langle s_i|s_j\rangle|^2 - 1. \tag{4.29}$$

By substituting (4.29) into (4.28), an alternative form of the elements of the Gram matrix expressed exclusively in terms of Jones vectors is obtained and expressed below.

$$\mathbf{G}_{ij} = 2\left(|\langle s_i|s_j\rangle|^2 - |\langle s_i|s_4\rangle|^2 - |\langle s_j|s_4\rangle|^2 + 1\right). \tag{4.30}$$

For the diagonal elements of the Gram matrix, i.e., where $i = j$, equation (4.30) yields the simplified expression:

$$\mathbf{G}_{i=j} = 4\left(1 - |\langle s_i|s_4\rangle|^2\right). \tag{4.31}$$

The next task is to analytically evaluate the gradient of the cost function with respect to the real and imaginary parts of the polarization components of the Jones vector symbols. Letting the variable $E := a + ib$ represent the real and imaginary parts of the Jones vector, serving as a placeholder, allows the cost function $\xi$ to be expressed in terms of the real and imaginary parts of a Jones vector component. To find the gradient of $\xi$, the derivative of $\xi$ with respect to $E$ must be computed. This derivative can be determined using Jacobi's formula, which provides the derivative of the determinant of a matrix:

$$\frac{d\xi}{dE} = \frac{d}{dE}\det(\mathbf{G}) = \det(\mathbf{G})\,\mathrm{Tr}\left[\mathbf{G}^{-1}\frac{d\mathbf{G}}{dE}\right]. \tag{4.32}$$

Notice from equation (4.32), the gradient only applies to the Gram matrix $\mathbf{G}$. The

Gram matrix has been shown to include constants, which are easily differentiable, and inner products of Jones vectors, which are less easily differentiable. Nevertheless, the expressions for the partial derivatives in terms of the real and imaginary parts of the Jones vector components are found to be given as follows:

$$\frac{\partial|\langle s_i|s_j\rangle|^2}{\partial a_{km}} = 2\,\mathrm{Re}\left\{\langle s_j|s_i\rangle\, s_{jm}\delta_{ik} + \langle s_i|s_j\rangle\, s_{im}\delta_{jk}\right\}, \tag{4.33a}$$

$$\frac{\partial|\langle s_i|s_j\rangle|^2}{\partial b_{km}} = 2\,\mathrm{Im}\left\{\langle s_j|s_i\rangle\, s_{jm}\delta_{ik} + \langle s_i|s_j\rangle\, s_{im}\delta_{jk}\right\}, \tag{4.33b}$$

where $\delta_{ik}$ denotes the Kronecker delta, $m$ is the polarization index (1, 2), and $k$ is an integer spanning 1, ..., 4. The above relationships can be combined into a single expression for the partial derivative with respect to the complex variable $E_{km}$:

$$\frac{\partial|\langle s_i|s_j\rangle|^2}{\partial E_{km}} := \frac{\partial|\langle s_i|s_j\rangle|^2}{\partial a_{km}} + i\frac{\partial|\langle s_i|s_j\rangle|^2}{\partial b_{km}} = 2\left(\langle s_j|s_i\rangle\, s_{jm}\delta_{ik} + \langle s_i|s_j\rangle\, s_{im}\delta_{jk}\right). \tag{4.34}$$

In the experimental setup, one iteration of the steepest ascent method consists of collecting $N$ Stokes vectors. These vectors are subdivided by the MATLAB clustering algorithm into four clusters of points, each corresponding to one of the four SVM symbol vectors in the simplex constellation. The average vector for each cluster $\bar{\hat{t}}_k$ is computed and normalized to unit length. The normalized average Stokes vectors correspond to four Jones vectors, $|s_k\rangle$, each with two complex components $E_{km} = a_{km} + ib_{km}$. The quadrature components $a_{km}$ and $b_{km}$ are used to control the MZI voltages, or more specifically, the scaling coefficients $\gamma$. The derivation of the final expression of the partial derivatives in equation (4.33) in terms of MZI scaling coefficients $\gamma$ are left to future sections.

For simplification, one of the position Jones vectors, such as $|s_4\rangle$, can remain fixed during the optimization process, while the remaining three vectors, $|s_k\rangle$ for $k = 1, \ldots, 3$,

are adjusted to maximize the volume of the tetrahedron. To facilitate this, the six complex components of $|s_k\rangle$ are grouped into a single matrix $\mathbf{J}$. The steepest ascent method can then be written as a recursion in the 6-dimensional complex space:

$$\mathbf{J}_{n+1} = \mathbf{J}_n + \mu \nabla \xi(\mathbf{J}_n), \tag{4.35}$$

where $\mu$ is the step size.

It is important to note that the recursion defined by equation (4.35) does not impose any constraints on the components of $\mathbf{J}_n$. As a result, there is no guarantee that the Jones vectors $|s_k(n+1)\rangle$ corresponding to the $(n+1)$-th iteration will be normalized to unity. This means that the corresponding Stokes vectors may not lie on the surface of the Poincaré sphere. Therefore, after computing $\mathbf{J}_{n+1}$ from equation (4.35), the Jones vectors $|s_k(n+1)\rangle$ must be rescaled to unit length:

$$\frac{|s_k(n+1)\rangle}{\| \, |s_k(n+1)\rangle \, \|}. \tag{4.36}$$

Subsequently, the rescaled components are used to update the vector $\mathbf{J}_{n+1}$.

Numerical Simulation

As a verification step for the steepest ascent procedure, a numerical simulation is performed assuming that the receiver obtains four Stokes vectors, with their vertices forming a tetrahedron in Stokes space with sub-optimal volume. The components of each initial Stokes vector are drawn from a normal distribution, where the magnitudes of each vector are normalized to ensure they lie on the Poincaré sphere. Typically, the step size is experimentally determined; however, for demonstration purposes, a step size of 0.01 is sufficient. The termination criterion is set to stop the optimization once the algorithm converges to a value that satisfies the procedure's requirements. For demonstration, a

convergence threshold of $10^{-4}$ is selected, representing the difference between successive iterations, $\xi_n - \xi_{n-1}$, where $n$ is the iteration index.

It is important to note that, to evaluate the partial derivatives in equation (4.33), it is necessary to convert the received Stokes vectors back into Jones vectors. As mentioned earlier, this conversion introduces an absolute phase ambiguity. However, in the context of equation (4.33), the sign of the Jones vectors at the current iteration, $|s\rangle$, is insignificant so long as the differential phase between polarization components is retained. The Stokes to Jones conversion is described by the following set of equations, where $S_i$ represents the $i$-th Stokes component:

$$|s\rangle := \begin{bmatrix} \cos\left(\frac{\theta}{2}\right) \\ \cos(\phi)\sqrt{1-\cos^2\left(\frac{\theta}{2}\right)} + i\sin(\phi)\sqrt{1-\cos^2\left(\frac{\theta}{2}\right)} \end{bmatrix}, \tag{4.37a}$$

$$\theta := \cos^{-1}(S_1), \tag{4.37b}$$

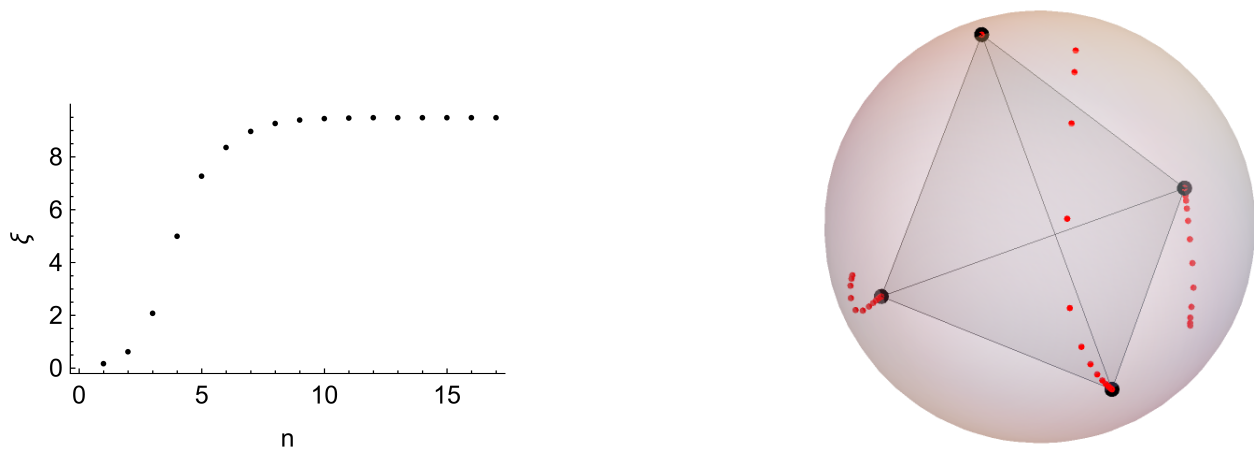
$$\cos(\phi) := \frac{S_2}{\sqrt{1-\cos^2(\theta)}}, \tag{4.37c}$$

$$\sin(\phi) := \frac{S_3}{\sqrt{1-\cos^2(\theta)}}. \tag{4.37d}$$

After applying the steepest ascent procedure to maximize the volume, the optimization terminates after only 17 iterations, as shown in Fig. 36a. From a qualitative perspective, Fig. 36b illustrates the rotation of Stokes vectors along the surface of the Poincaré sphere. Successive iterations, indicated by the red points, eventually stop at positions defined by the black points. These positions form a tetrahedron with a maximized volume ($\xi \approx 9.48$).

The vector near the top of the sphere remains fixed during the procedure. This provides illustrative clarity, showing that only three out of the four vectors change between iterations. To further demonstrate that the volume is indeed maximized, a quality matrix can be used, calculated by leveraging the unique property of the regular tetrahedron formed by four unit

vectors. In this case, each Stokes vector undergoes an inner product with every other Stokes vector, as defined by the relationship in equation (4.18).




(a) (b)

Figure 36: (a) cost function $\xi$ plotted against the iteration index $n$; (b) The formed tetrahedron after 17 iterations. The black points represent the resulting Stokes vectors that maximize the volume and the red points represent successive iterations.

A newly defined Gram matrix is formed by taking the inner products of all permutations of indices $i$ and $j$. This is done by computing the matrix product of the optimized Stokes vectors, $\mathbf{G}^T\mathbf{G}$, where the Gram matrix $\mathbf{G}$ is the column matrix containing the components of the optimized Stokes vectors. Using the results illustrated above, the following Gram matrix is obtained:

$$\mathbf{G}^T\mathbf{G} = \begin{bmatrix} 1.0000 & -0.3326 & -0.3335 & -0.3315 \\ -0.3326 & 1.0000 & -0.3339 & -0.3329 \\ -0.3335 & -0.3339 & 1.0000 & -0.3356 \\ -0.3315 & -0.3329 & -0.3356 & 1.0000 \end{bmatrix}. \tag{4.38}$$

From the matrix above, the diagonal elements are equal to one, representing the permutations of dot products where $i = j$. The off-diagonal elements correspond to

permutations of dot products where $i \neq j$. It is evident that the off-diagonal elements exhibit error in their trailing significant digits and do not exactly equal $-\frac{1}{3}$. This discrepancy arises due to the finite termination criterion employed in the numerical optimization scheme. To achieve a more precise result, one could lower the termination criterion, though this would come at the cost of a greater number of iterations.

Experimental Implementation

The experimental implementation of the steepest ascent optimization procedure to improve the generated simplex SVM constellation can be achieved by recursively updating the AWG waveform scaling coefficients $\gamma$ as given in Table 4, effectively optimizing the AWG driving voltages. Since the Jones vectors themselves cannot be updated recursively with the AWG in the experiment, an additional step is required to express the partial derivatives from (4.33) in terms of the scaling coefficients $\gamma$. This can be accomplished using the equations from (4.13) and the chain rule, where (4.33) becomes:

$$\frac{\partial |\langle s_i|s_j\rangle|^2}{\partial \gamma_{km}} = 2\,\mathrm{Re}\left(\langle s_j|s_i\rangle\, s_{jm}\delta_{ik} + \langle s_i|s_j\rangle\, s_{im}\delta_{jk}\right)\beta\cos\left(\beta\gamma_{km}\right), \tag{4.39a}$$

$$\frac{\partial |\langle s_i|s_j\rangle|^2}{\partial \gamma_{km}} = 2\,\mathrm{Im}\left(\langle s_j|s_i\rangle\, s_{jm}\delta_{ik} + \langle s_i|s_j\rangle\, s_{im}\delta_{jk}\right)\beta\cos\left(\beta\gamma_{km}\right). \tag{4.39b}$$

From this point, the recursion occurs on the scaling coefficients listed in Table 4. These coefficients are redefined in terms of $\mathbf{\Gamma}$, which represents the 16-element matrix containing the updated $\gamma$ values:

$$\mathbf{\Gamma}_{k+1} = \mathbf{\Gamma}_k + \mu \boldsymbol{\nabla} \xi\left(\mathbf{\Gamma}_k\right). \tag{4.40}$$

The matrix $\mathbf{\Gamma}$ contains all the $\gamma$ coefficients for each MZI, enabling each one to recursively update to find the optimal MZI driving configuration.

The experimental block diagram, which includes the steepest ascent optimization procedure and the re-computation and upload of MZI waveforms, is shown in Fig. 37.

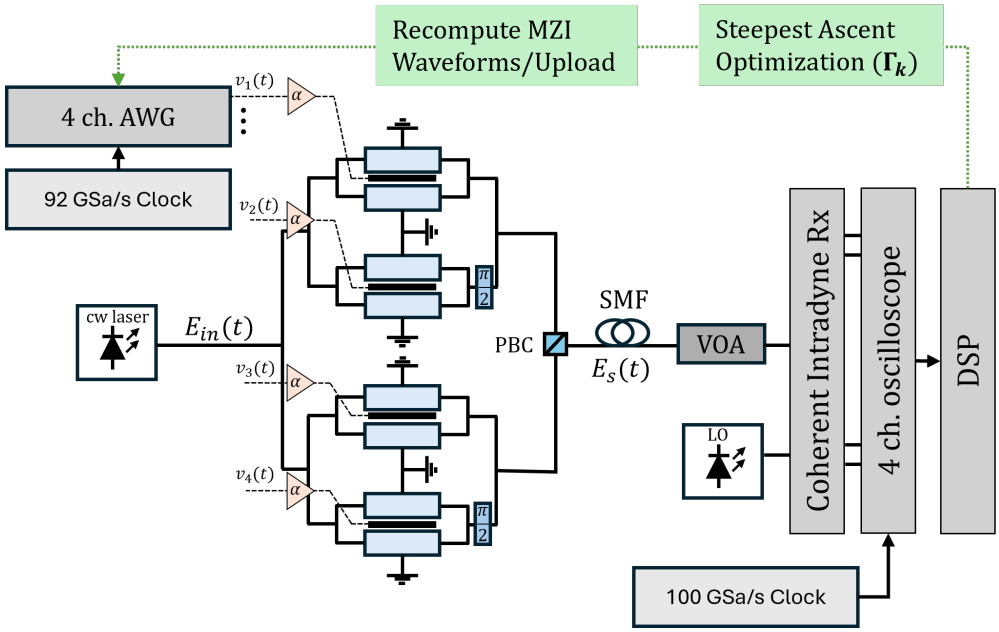



Figure 37: Experimental block diagram incorporating the steepest ascent optimization procedure to update MZI scaling coefficients $\gamma$.

After each iteration of steepest ascent, a new set of $\gamma$ coefficients is computed and used to generate a new set of MZI waveforms, maintaining the same de Bruijn sequence as shown in Fig. 24. With the partial derivatives now redefined in terms of the scaling coefficients, the previous normalization step is no longer necessary. However, it is important to keep the $\gamma$ values within the dynamic range of the AWG DAC, which is [-1, +1]. Conveniently, if any $\gamma$ coefficient is updated outside of this range, the AWG will automatically rescale the coefficient to the corresponding maximum or minimum value. In the DSP, all functions are carried out through the symbol sequence resolution step, as shown in Fig. 26. Once the symbol sequence is determined, the symbol vector averages for all clusters $\bar{\hat{t}}_k$, $k = 1, ..., 4$ are

computed and used to calculate the next iteration of $\gamma$ coefficients, denoted as $\mathbf{\Gamma}_{k+1}$. Using the updated set of $\gamma$ coefficients, new MZI waveforms are uploaded to each channel of the AWG, and the process repeats.

Following each acquisition of a new set of $N \approx 2,300$ symbols, the steepest ascent procedure iterates with step size $\mu = 0.01$ and progressively increases the volume of the tetrahedron formed by the newly acquired Stokes vectors. The improvements between successive iterations are illustrated in Fig. 38.

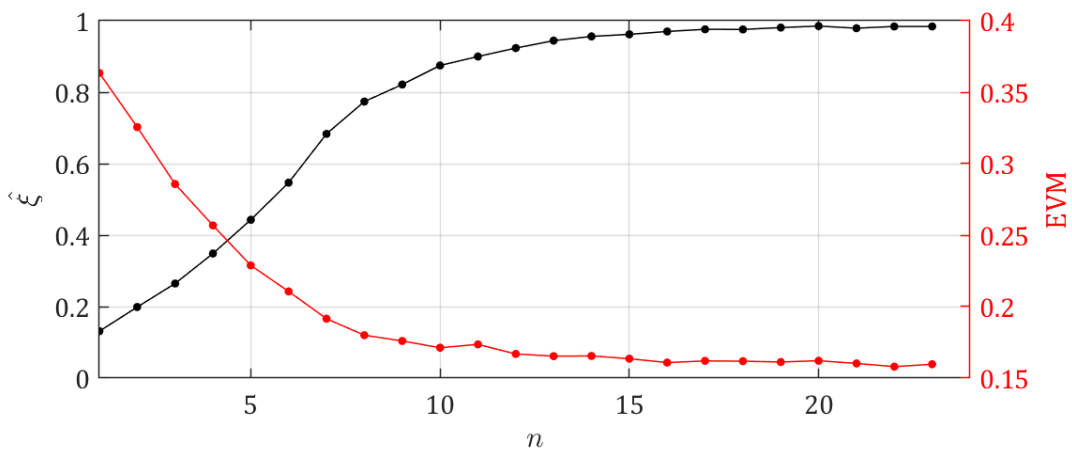



Figure 38: Normalized volume $\hat{\xi}$ plotted against number of iterations $n$.

Starting from the initial set of Stokes vectors shown in Fig. 28a, the normalized cost function $\hat{\xi}$ is initially evaluated at around 15% of the maximal achievable volume. However, as demonstrated in the figure, after just 23 iterations of the steepest ascent procedure, the volume of the received tetrahedron increased to approximately 98% of the maximal achievable volume. In addition to this volume improvement, there was a notable reduction in cluster noise, as quantified by the EVM. Initially, the Stokes vectors produced an EVM of approximately 0.36, which is reduced to approximately 0.16 after only 23 iterations, indicating a substantial decrease in the aggregate cluster noise.

The projection of the received Stokes constellation onto the $yz$-plane is shown in Fig. 39.

A qualitative comparison with the initial set of Stokes vectors in Fig. 33 reveals significant noise reduction, highlighting the effectiveness of the optimization procedure.

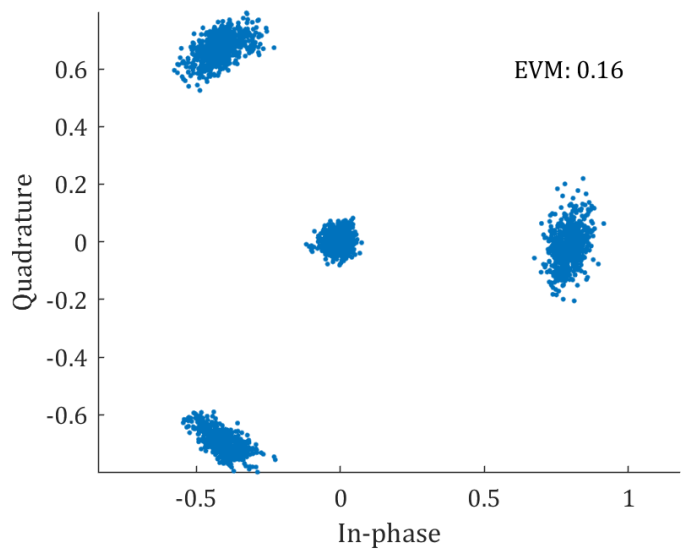
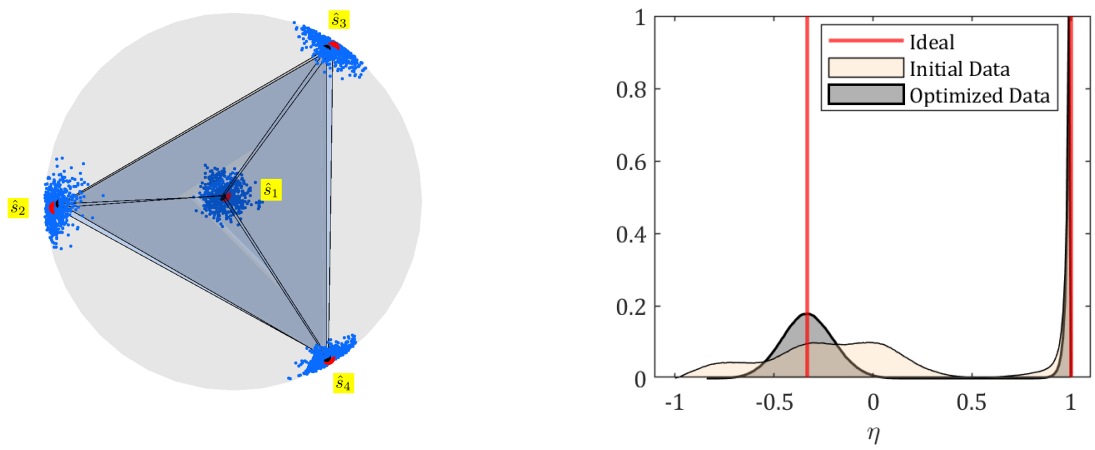



Figure 39: The SVM simplex plotted in the $yz$-plane with EVM $\approx 0.16$.

The final iteration of the optimized Stokes vectors, after derotation using the pseudo-inverse, is visualized on the Poincaré sphere and compared to the ideal regular tetrahedron and Stokes symbol vertices (Fig. 40a). The received tetrahedron closely resembles the ideal shape, and the cluster averages are nearly coincident with the ideal Stokes symbol vertices. Although the overlap is not perfect, it can be considered sufficient for performance evaluations of Stokes vector constellations in practical applications. Additionally, the symmetry of the constellation shows significant improvement, as illustrated in Fig. 40b. Compared to the initial distribution of Stokes vector inner products (shown in light orange), the optimized vectors exhibit a much stronger concentration around the ideal inner product values of $-\frac{1}{3}$ and 1.




(a) (b)

Figure 40: Figures display experimental results of the SVM simplex after 23 iterations of steepest ascent optimization. (a) Received Stokes vectors plotted on the Poincaré sphere; (b) The histogram of inner products.

The quantitative improvement of the optimization procedure is demonstrated by experimentally measuring the BER of the newly optimized constellation at decreasing levels of receive power. As shown in Fig. 41, at the KP4 FEC threshold, the optimized Stokes vectors produce a curve that is shifted significantly to the left (by approximately 5.5 dB) compared to the initial set of received Stokes vectors, indicated by the dotted red curve. This improvement in BER is substantial and leads to a dramatic reduction in the overall implementation penalty of the simplex SVM constellation.

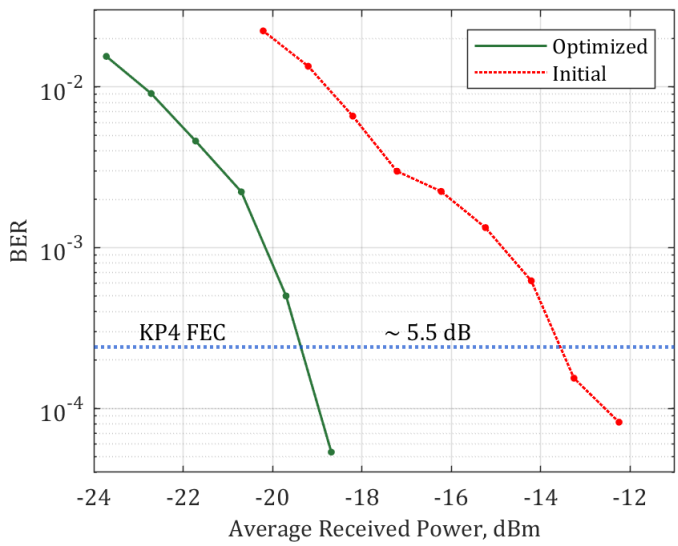



Figure 41: BER vs average receive power showing an $\approx$ 5.5 dB improvement in sensitivity after optimization.

## Future Work

This chapter introduces a novel approach for generating Stokes Vector Modulation and Mode Vector Modulation constellations that exhibit high symmetry and low noise. In the context of modern short-haul optical interconnects, maximizing spectral efficiency for SVM/MVM constellations is highly valuable, especially given their superior energy efficiency compared to conventional $M$-ary pulse amplitude modulation. By building upon prior work in optimizing the constellation shaping of points on the Poincaré sphere based on steepest ascent, there is significant potential to realize highly dense SVM and MVM constellations (Fig. 42).

Looking forward, high-cardinality and high-dimensional polarization-based constellations may emerge as viable alternatives to other direct detection formats. This can be achieved by optimizing modulator driving voltages via feedback to generate constellations with high symmetry. The scalability of this optimization method to higher-cardinality constellations, combined with the use of more advanced cost functions, makes it possible

to design SVM/MVM constellations with minimal implementation penalties. This progress paves the way for the experimental realization of their energy efficiency advantages in practical settings.

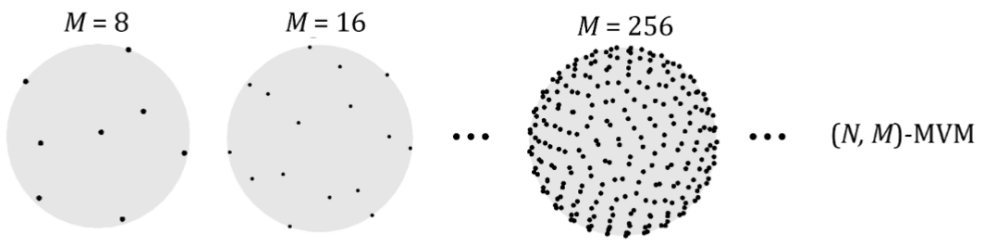



Figure 42: Future experimental generations of high-cardinality SVM constellations eventually paving the way toward the generation of promising $(N, M)$-MVM constellations.

# IMPACT OF MODAL DISPERSION AND INTERMODAL CROSSTALK ON MVM

## Overview

Modal dispersion and intermodal crosstalk are significant challenges in few-mode fiber (FMF) systems that can substantially impact the performance of Mode Vector Modulation (MVM). In a FMF, light propagates through several distinct paths (modes), each characterized by different propagation constants, expressed as $\beta_k := 2\pi n_k/\lambda$, where $n_k$ is the effective refractive index of the $k$-th guided mode. The propagation constants of these modes are different and depend on their respective spatial distributions in the fiber. Modal dispersion occurs when the modes of the FMF travel at varying speeds, resulting in MVM pulse broadening. In an MVM system, where simultaneous sampling of each mode is required, modal dispersion complicates the precise timing and recovery of the $N$-dimensional Jones vector components at the receiver. The resulting timing errors between sampled modes reduce signal integrity, negatively impacting the overall performance of MVM.

Intermodal crosstalk, on the other hand, refers to the coupling of light between the different modes within the FMF. This coupling arises from fiber imperfections, such as variations in core shape or refractive index, or from non-linearities induced by high-power signals [40]. In the context of MVM, intermodal crosstalk distorts the power ratios between the polarization components of each mode, which affects the received Stokes vector and, ultimately, the symmetry of the MVM constellation. This mixing of power between modes complicates the maintenance of the high symmetry and low noise characteristics that are essential for MVM.

Collaborators in the following study from University College London (UCL) have developed a semi-analytical model of linear mode coupling due to waveguide imperfections, which is suitable for simulating multimode propagation in FMFs [9]. They have also designed a trench-assisted graded-index core profile optimized for low differential mode

delay (DMD) in the C-band wavelength range [8]. Using these approaches, numerical simulations of MVM over a $k$-mode FMF channel can be performed. These simulations allow for the evaluation of MVM performance under varying conditions of modal dispersion and intermodal crosstalk, offering a clearer understanding of MVM's advantages over other single-mode direct-detection formats, such as $M$-ary Pulse Amplitude Modulation ($M$-PAM) and $M$-ary Stokes Vector Modulation ($M$-SVM).

## Semi-Analytical Model

In the simulations presented in this study, a multi-section fiber model, as introduced in [9], is employed to emulate the realistic propagation of modes through a FMF. This model is based on a set of coupled-mode equations that describe the evolution of the mode amplitudes along the fiber's length. These equations, expressed in matrix form, enable a detailed analysis of how the optical modes interact and propagate within the fiber, and their general form is given below:

$$\partial_z \tilde{\mathbf{A}}(z,\omega) = -\mathrm{j}[\boldsymbol{\beta}(z,\omega) + \mathbf{K}(z) + \mathbf{R}(z)]\tilde{\mathbf{A}}(z,\omega). \tag{5.1}$$

In the above equation, $\tilde{\mathbf{A}}(z,\omega)$ is a vector that represents the complex envelope of the electric field for each mode at a specific position $z$ along the fiber and a given optical frequency $\omega$. The elements of this vector correspond to the amplitudes of each mode, with the spatial distribution of the field within the fiber being influenced by these amplitudes. The matrix $\boldsymbol{\beta}(z,\omega)$ contains the propagation constants for each mode, which depend on the position $z$ and the frequency $\omega$, encapsulating the modal dispersion characteristics of the fiber. The matrix $\mathbf{K}(z)$ represents the power coupling between the $N$ polarization modes, with each element $\mathbf{K}_{mn}(z)$ capturing the interaction strength between modes $m$ and $n$ through the overlap of their respective electric fields. Finally, $\mathbf{R}(z)$ is a unitary block-diagonal matrix

accounting for polarization rotations in the fiber.

To make the solution of these coupled-mode equations more tractable, it is assumed that modal dispersion and linear mode coupling can be treated as independent over small segments of the fiber, each of length $\Delta z$. Within each segment, the modal properties, including propagation constants and coupling matrices, are assumed to be approximately constant, simplifying the overall model. The transfer matrix for each small segment of fiber is thus given by:

$$\tilde{\mathbf{M}}(\omega) \cong \exp\{-\mathrm{j}\,[\boldsymbol{\beta} + \mathbf{K} + \mathbf{R}]\,\Delta z\}. \tag{5.2}$$

This approximation assumes that over the short distance $\Delta z$, the effects of modal dispersion, e.g., the frequency dependence of the propagation constants, do not change significantly. As a result, for each fiber segment, the propagation constant matrix $\boldsymbol{\beta}$ can be expanded in a Taylor series around a central angular frequency $\omega_0$:

$$\boldsymbol{\beta} = \sum_l \boldsymbol{\beta}_l (\omega - \omega_0)^l. \tag{5.3}$$

In this expansion, $\boldsymbol{\beta}_l$ represents the $l$-th derivative of the propagation constant matrix with respect to the frequency $\omega$. The elements $\boldsymbol{\beta}_m$ and the coupling matrix elements $\mathbf{K}_{mn}$ are determined from the refractive index profiles of the fiber, which are designed to minimize modal group delay. These refractive index profiles assume a graded-index fiber with a cladding trench, a design that improves the fiber's performance in reducing DMD. As described in [8], the fiber's parameters for the simulation include a graded-index exponent of 1.95 for the core, a relative refractive index difference of 0.01 between the core and cladding, a core radius of 6.8 $\mu$m, a trench width of 4 $\mu$m, and a trench depth of $-0.0034$.

In the simulation model, the propagation constant elements $\boldsymbol{\beta}_m$ are assumed to be constant across all fiber sections, while the mode coupling elements $\mathbf{K}_{mn}$ are allowed to vary

in each section due to random offsets in the radial and azimuthal positions of the modes. These random offsets introduce variability in the mode coupling across the fiber, simulating the effects of fiber imperfections that can lead to mode mixing and crosstalk.

Finally, the overall transfer matrix for the entire fiber is obtained by multiplying the transfer matrices for each fiber section. This end-to-end transfer matrix, denoted as $\mathbf{H}(\omega)$, encapsulates the cumulative effects of modal dispersion, mode coupling, and polarization rotation across the entire fiber:

$$\mathbf{H}(\omega) = \tilde{\mathbf{M}}_K(\omega) \cdot \tilde{\mathbf{M}}_{K-1}(\omega) \cdot \tilde{\mathbf{M}}_{K-2}(\omega) \ldots \cdot \tilde{\mathbf{M}}_1(\omega), \tag{5.4}$$

where $K$ represents the total number of fiber sections considered in the simulation. By computing this transfer matrix, it is possible to model and analyze the few-mode propagation effects within the fiber, such as modal dispersion, intermodal coupling, and polarization-dependent behaviors, which are essential for evaluating the performance of MVM systems in practical few-mode fiber channels.

## Description of the Simulation Parameters

### Simulation Diagram and Considerations

The block diagram of the optical transmission system using Mode Vector Modulation with direct-detection (MVM/DD) for point-to-point communication is shown in Fig. 43. At the transmission end, MVM signals with varying degrees of freedom (DoFs) $N$ and different constellation sizes $M$ are generated and launched into the optical fiber. These signals are formed through the use of geometrically optimized constellations with efficient bit-to-symbol mappings, as explained in [19]. The generated MVM pulses are shaped using a root-raised cosine (RRC) filter before being sent through the fiber.

In the simulation setup presented in Fig. 43, two different fiber configurations are

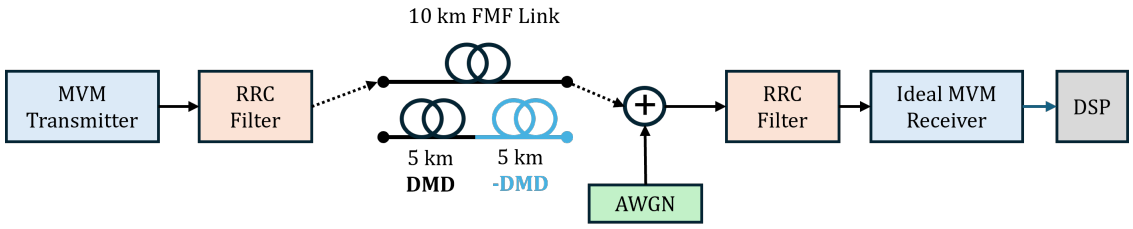



Figure 43: MVM simulation block diagram denoting two simulation scenarios. The top fiber represents a standard 10 km FMF link with positive DMD, where the bottom fiber represents propagation through 5 km of FMF with positive DMD and another 5 km of FMF with negative DMD.

considered. The first option uses a single type of FMF throughout the entire fiber span, which is referred to as the modal dispersion unmanaged link "MD unmanaged link." In the second setup, the fiber span is split into two segments: the first segment uses a FMF transmission fiber, while the second segment uses a FMF compensating fiber. This second configuration, known as the "MD managed link", is designed to reduce modal dispersion by using opposite-sign fiber DMD parameters. When the intermodal coupling is weak, cascading two FMFs of the same length but with opposite DMDs can effectively reduce DMD, as described in [28].

Pulse propagation through the fiber is modeled using the multi-section approach, where short segments of fiber are treated separately, and the fiber characteristics, including the distorted core-cladding boundary, are accounted for using matrix concatenation. This model takes into account all relevant linear impairments such as Rayleigh scattering, macro-bend loss (MBL), mode coupling, and DMD.

To simulate the intermodal crosstalk (XT) typically encountered in FMFs, Monte Carlo simulations are used. These simulations are configured to achieve a specific average intermodal XT per unit length, as commonly seen in FMFs [9]. The radial displacement between the core and cladding is adjusted to target the desired level of intermodal XT. Additionally, the uncoupled DMD of the fiber is artificially scaled to match a desired value.

Alternatively, a redesign of the fiber can be done to meet the exact DMD specifications, as outlined in [8].

In these simulations, chromatic dispersion (CD) is not considered, assuming that the system operates within the O-band [7], where CD is negligible. The noise effect from amplified spontaneous emission (ASE) is incorporated using an additive white Gaussian noise (AWGN) model at the receiver. After noise loading, the received signal is processed through an optical matched RRC filter [19]. It is assumed that the receiver is ideal, noiseless, and capable of performing MVM direct-detection. The ideal receiver converts the received symbols from the generalized Jones space to the generalized Stokes space [19].

Following reception, the digital signal processing (DSP) unit compensates for polarization rotations introduced during fiber propagation. Operating in Stokes space, the DSP multiplies the received Stokes vectors by a compensating Mueller matrix, with the matrix elements adjusted iteratively through a least-mean-square (LMS) algorithm [19]. Initially, the LMS algorithm uses a pilot-aided approach for the first 3,000 received symbols, then switches to a data-aided decision-directed approach for the remaining symbols with a fixed step size of 0.01. After polarization derotation, maximum-likelihood estimation (MLE) is used to estimate the transmitted data, and the bit error rate (BER) is computed.

To allow for a fair comparison, all modulation formats used in the simulations are designed to have the same spectral efficiency per spatial degree of freedom. In the following section, the performance of a (6,64)-MVM format over a 6-polarization-mode FMF is compared with that of three separate 4-PAM or 4-SVM tributaries transmitted over three single-mode fiber (SMF) lanes. The results for 4-PAM are derived under the assumption of signal-ASE noise-limited operation, as discussed in [26], with a one-sided noise bandwidth equal to half the symbol rate. For all the modulation formats, an average of 2 bits per symbol per spatial degree of freedom is transmitted. Standard optical interface (OI) symbol rates of 56 and 112 GBd are considered, with 12% Ethernet and KP4 forward error correction

(FEC) overhead.

Simulation Parameters for 3-mode Propagation

The simulations are conducted using a 3-mode FMF that supports the fundamental $LP_{01}$ mode as well as the degenerate $LP_{11a}$ and $LP_{11b}$ modes. For each simulation, the parameters governing the FMF transmission are provided in Table 6. In these simulations, the effects of CD and nonlinearity are neglected, meaning that the dispersion coefficient $D$ and the nonlinear parameter $n_l$ are set to zero. All other parameters are fixed throughout the simulation. Since the performance of the MVM format will be compared under varying levels of DMD and XT, these parameters are kept general and are expressed per unit length.

| **Simulation Parameter** | **Value** |
|---|---|
| Differential Mode Dispersion (DMD) | Set by MD parameter [ps/km] |
| Intermodal crosstalk (XT) | Set by XT parameter [dB/km] |
| Mode Dispersion Parameter ($D$) | 0 |
| Mode Dispersion Slope ($S$) | $LP_{01}$: 67.0994; $LP_{11}$: 68.5156 [ps/nm·km] |
| Nonlinear Coefficient ($n_l$) | 0 |
| Effective Index of Refraction ($n_{eff}$) | $LP_{01}$: 1.4490; $LP_{11}$: 1.4466 |
| Attenuation Coefficient ($\alpha$) | $LP_{01}$: 0.1640; $LP_{11}$: 0.1589 [dB/km] |
| Fiber Length ($L$) | 10 [km] |
| Step Size ($\Delta z$) | 10 [m] |

Table 6: Simulation parameters.

## Numerical Simulation of (6,64)-MVM

### MVM Transmitter

The MVM transmitter block in Fig. 43 takes the MVM constellation DOF $N$ and cardinality $M$ as inputs. Based on these, the $N$-dimensional MVM Jones vectors are imported. For (6,64)-MVM, there are 64 Jones vectors $|s_i\rangle$, $i = 1, ..., 64$ with 6 complex elements each, corresponding to the $x$ and $y$ polarization components of the 3 modes:

$$|s\rangle = \left[E_{x1}, E_{y1}, E_{x2}, E_{y2}, E_{x3}, E_{y3}\right]^T. \tag{5.5}$$

For each simulation, the number of simulated bits is fixed to $N_{\text{bits}} = 2^{16}$. With $M = 64$ and 6 bits per symbol, the number of symbols is $N_{\text{sym}} = 10,922$. Using `randi()`, a random symbol sequence is generated in MATLAB, and the pulses are shaped using the RRC filter impulse response shown in Fig. 44, with 8 samples per symbol and a roll-off factor of 0.1. The transmission waveforms for the first 1000 samples are shown in Fig. 45.

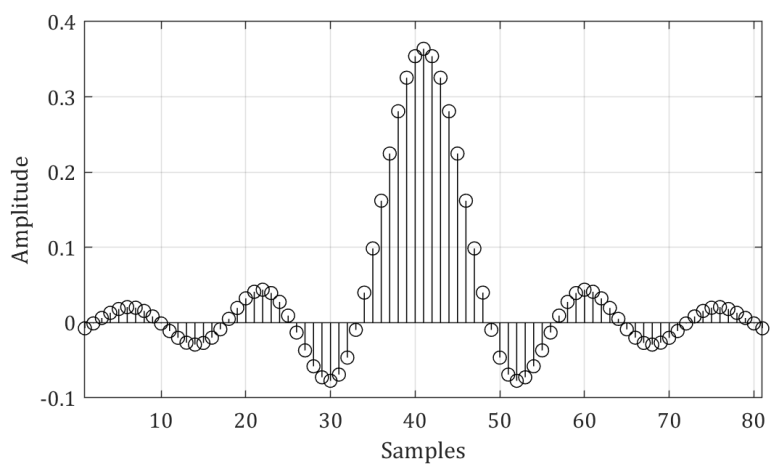



Figure 44: RRC impulse response.

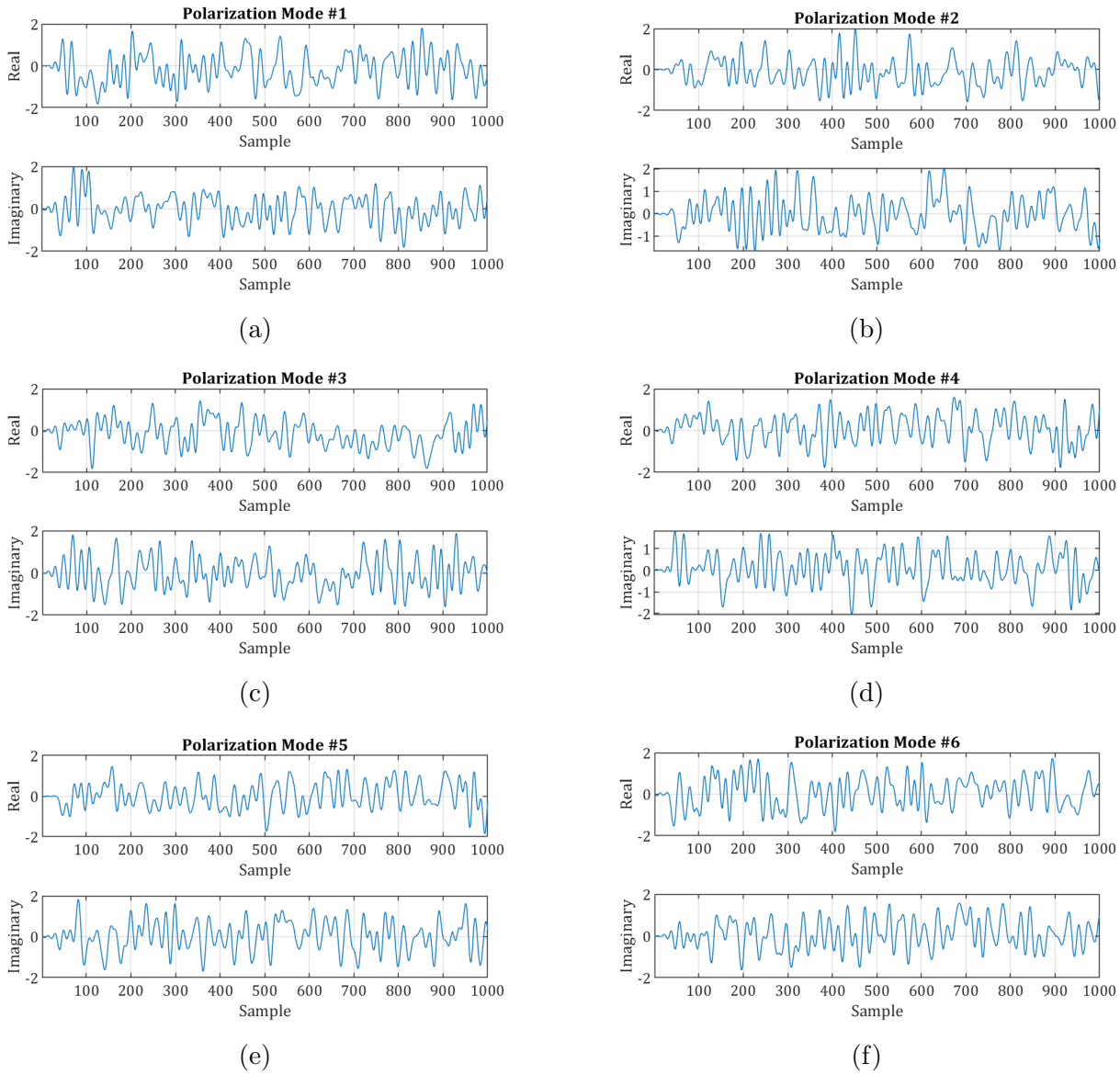



Figure 45: Simulation transmission waveforms for (6,64)-MVM. Each figure shows the real and imaginary parts of the 6-dimensional transmitted Jones vectors for modes (a) $LP_{01}$ $x$-pol; (b) $LP_{01}$ $y$-pol; (c) $LP_{11a}$ $x$-pol; (d) $LP_{11a}$ $y$-pol; (e) $LP_{11b}$ $x$-pol; (f) $LP_{11b}$ $y$-pol.

Crosstalk Accumulation and Modal Dispersion

Following the generation of each polarization component, crosstalk coupling matrices are calculated for each fiber segment of length $\Delta z$ in the simulation [9]. These coupling matrices are designed to accumulate the desired average crosstalk level over the length of the fiber. For example, suppose the target average crosstalk is set to -30 dB/km. In that case, each crosstalk matrix for a segment of size $\Delta z$ is adjusted to ensure that the cumulative crosstalk over the entire fiber approaches this target value. The coupling matrices are visualized by plotting each mode index as $10 \log |\mathbf{K}_{mn}|^2$. Figures illustrating these coupling matrices are provided below.

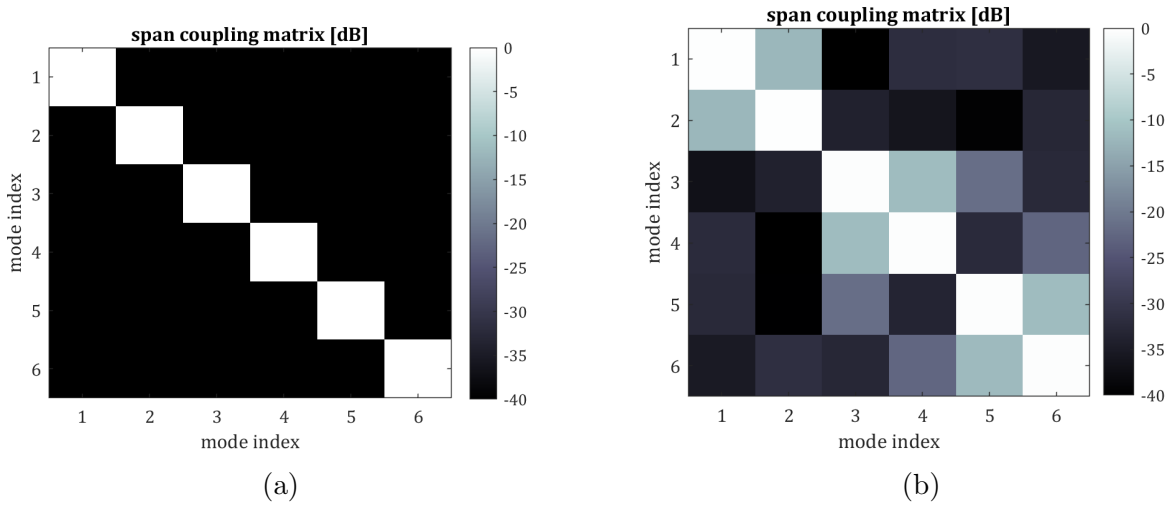



Figure 46: Coupling matrices between mode indices for the FMF supporting 3 spatial modes. Each figure shows the magnitude of each coupling matrix element in dB for the (a) first step in the fiber where $z = 10$ m; (b) the accumulation of coupling matrices where $z = 10^4$ m.

Each mode propagates through the fiber in discrete steps, accumulating crosstalk between polarization modes at each stage. In addition to this, the propagation of each mode is influenced by its specific group velocity, which depends on the MD parameter. The MD parameter, set in the simulation and measured in ps/km, defines the 6-dimensional uncoupled DMD vector, which represents the differential delay for each mode in units of picoseconds.

This vector is directly related to the first derivative of the propagation constant with respect to frequency, $\beta_1$, for each mode.

To determine the higher-order derivatives of the propagation constant $\beta$, the dispersion and dispersion slope parameters ($D$ and $S$) of the optical fiber are used. These are given by the following equations, where $c$ and $\lambda_0$ represent the vacuum velocity of light and center wavelength, respectively:

$$\beta_2 := -\frac{D\lambda_0^2}{2\pi c}, \tag{5.6}$$

$$\beta_3 := \frac{\lambda_0^2}{(2\pi c)^2(\lambda_0^2 S + 2\lambda_0 D)}. \tag{5.7}$$

If optical compensation for DMD is assumed, the first derivative of the fiber's propagation constant, $\beta_1$, is negated for the final 5 km of the 10 km fiber length.

Noise Loading and the Ideal MVM Receiver

After fiber propagation, a noise vector is generated and added to the received waveforms. This noise vector is produced using the `normrnd()` function in MATLAB, which draws random numbers from a normal distribution with a mean $\mu = 0$ and standard deviation $\sigma$. The standard deviation $\sigma$ depends on the average symbol energy $\tilde{E}$:

$$\sigma := \frac{\sqrt{\tilde{E}}}{2L_b\gamma_b}, \tag{5.8}$$

where $L_b$ is the bit word length, $\gamma_b$ represents the bit signal-to-noise ratio. The noise vector $n$ is complex-valued with dimensions $T \times 6$, where $T$ denotes the length of the received samples. The received noisy waveforms, denoted as $R$, are obtained by the following equation:

$$R := T + n. \tag{5.9}$$

Figure 47 shows a plot of the real part of the $x$-polarization component of the $LP_{01}$ mode's resulting waveform after fiber propagation and AWGN, contained in $R$, assuming a signal-to-noise ratio $\gamma_b = 10$ dB.

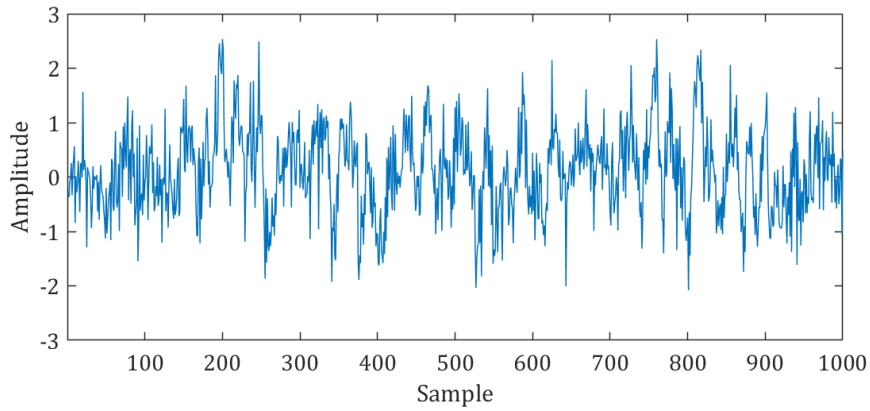



Figure 47: The real part of the first dimension of $R$, corresponding to the $x$-polarization of the first polarization mode ($LP_{01}$), is shown for the first 1000 samples, assuming a signal-to-noise ratio $\gamma_b = 10$ dB.

The noisy received waveforms $R$ are then filtered using the RRC impulse response, shown in Fig. 44, and downsampled to a single sample per symbol corresponding to the symbol rate $R_s$. The matched filter helps reduce excess noise, as illustrated in Fig. 48.

Inside the ideal MVM receiver, each sampled 6-dimensional Jones vector $|s\rangle$ is converted into a 35-dimensional Stokes vector $|\hat{s}\rangle$ using the Gell-Mann matrix in the generalized Stokes space [19]. The resulting Stokes vectors are then normalized to unit magnitude. Before making symbol decisions in the generalized Stokes space, polarization derotation is applied using the LMS algorithm, which is discussed in the following section.

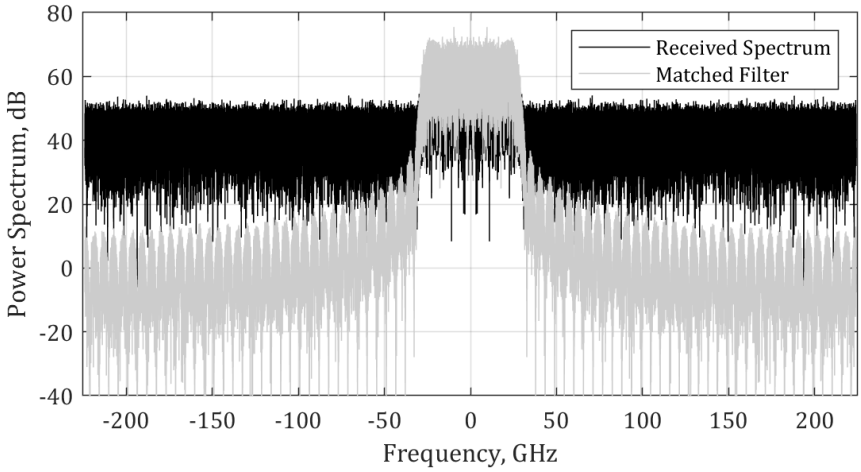



Figure 48: Filtered spectrum of the real part of the $x$-polarization of the first polarization mode ($LP_{01}$).

Polarization Derotation using the Least-Mean-Square Algorithm

One of the most critical steps in digital signal processing for SVM/MVM transmission is the estimation of the time-varying and random fiber Mueller matrix $\mathbf{R}$. Several methods are available for accurately estimating $\mathbf{R}$. As previously discussed in this thesis, the pseudo-inverse method requires knowledge of the transmitted symbol sequence. However, in the absence of access to the transmitted sequence, pilot-aided methods can be used. This approach assumes that the time-varying birefringence of the fiber remains approximately constant over the timescale of the fiber-optic baud rate. With this assumption, a packet of known pilot symbols of length $k$ can be periodically transmitted to estimate $\mathbf{R}$. After the pilot symbols are transmitted, a payload packet of length $L$ and duration $\tau$ follows. The duration $\tau$ represents the maximum period before $\mathbf{R}$ must be re-estimated, as shown in Fig. 49.

With known pilot symbols, the Mueller matrix $\mathbf{R}$ can be estimated using the pseudo-inverse; however, careful attention must be given to the overhead caused by the pilot symbols, which affects the rate $k/L$. A higher overhead reduces the effective data-carrying capacity

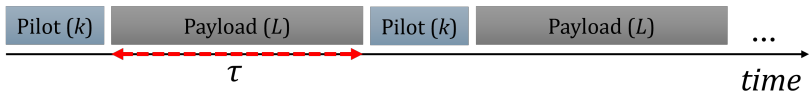



Figure 49: Time-domain view of transmitted Stokes symbols partitioned into pilot and payload packets.

of fiber transmissions.

In dynamic environments, determining the maximum period $\tau$ can be challenging, which may require adaptive equalization for continuous estimation of the Mueller matrix. The LMS algorithm is a suitable choice due to its ability to track relatively slowly varying changes, such as polarization rotations, by monitoring an error signal $\vec{e}(t)$.

The LMS algorithm can be applied in both pilot-aided and data-aided decision-directed modes. Initially, using a sequence of known pilot symbols, the LMS algorithm derotates the received MVM constellation in the generalized Stokes space. To reduce transmission overhead, the pilot symbols are then turned off, and the algorithm transitions to a decision-directed MLE approach for adaptive control of future fiber rotations. This section first discusses the use of pilot symbols for polarization derotation, before moving on to the decision-directed mode.

In Fig. 50, the received pilot Stokes vector $\hat{t}_k$ at the $k$-th time instant is multiplied by a compensating generalized Mueller matrix $\mathbf{M}(k)$. The resulting generalized Stokes vector $\hat{z}_k$ is then subtracted from the $k$-th pilot symbol $\hat{s}_k$ to form the error signal $\vec{e}_k$:

$$\vec{e}_k = \hat{s}_k - \hat{z}_k. \tag{5.10}$$

The error signal $\vec{e}_k$ is used to compute the cost function $\xi$ for the LMS equalizer. The LMS algorithm minimizes the mean square error (MSE) of the error signal with $E\{\cdot\}$ denoting the expectation operator and is expressed as:

$$\xi(k) = E\left\{\|\vec{e}_k\|^2\right\}. \tag{5.11}$$

In the blind (decision-directed) implementation of the algorithm, the LMS iteratively adjusts the matrix $\mathbf{M}(k)$ such that $\xi(k)$ is kept below a predefined threshold. For pilot-aided symbols, the recursion is terminated after $k$ pilot symbols, which are used to determine $\mathbf{M}(k)$.

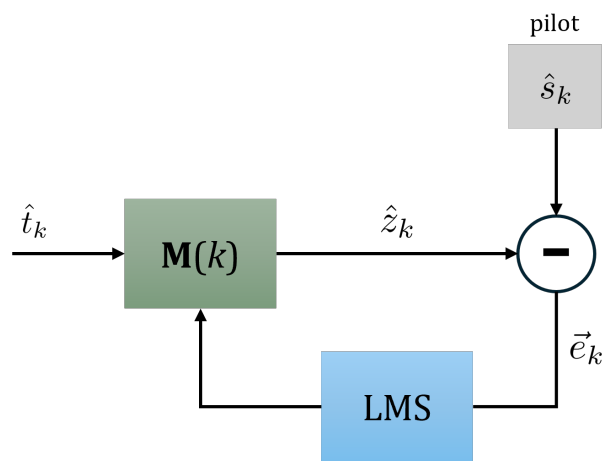



Figure 50: Least-mean squares pilot-aided adaptive channel equalizer.

The LMS algorithm minimizes the cost function by recursively adjusting the elements of the Mueller matrix $\mathbf{M}(k)$ as shown in Fig. 50. Considering expectation's instantaneous value $\|\vec{e}_k\|^2$, and expanding the norm, the following expression is obtained:

$$\xi(k) = \|\hat{s}_k - \hat{z}_k\|^2 = 1 + \|\hat{z}_k\|^2 - 2(\hat{s}_k \cdot \hat{z}_k). \tag{5.12}$$

Minimizing this instantaneous cost function involves calculating the partial derivatives of $\xi(k)$ with respect to the elements of $\mathbf{M}_{ij}$. In the pilot-aided approach, the known pilot symbols $\hat{s}_k$ are unaffected by small changes in $\mathbf{M}_{ij}$. Therefore, the partial derivatives of the terms involving $\hat{z}_k$ with respect to $\mathbf{M}_{ij}$ are evaluated. If blind equalization is used, the same

assumption can be made to first order. The gradient is then computed as:

$$\nabla\xi = \left[\frac{\partial\xi}{\partial M_{ij}}\right], \tag{5.13}$$

$$\frac{\partial\xi}{\partial M_{ij}} \simeq 2\left(\vec{z}_k - \hat{s}_k\right)\hat{t}_k. \tag{5.14}$$

The gradient descent update for the Mueller matrix $\mathbf{M}(k)$ is given by:

$$\mathbf{M}(k+1) = \mathbf{M}(k) - \mu\nabla\xi(k) = \mathbf{M}(k) - 2\mu\left[\hat{z}_k - \hat{s}_k\right] \otimes \hat{t}_k, \tag{5.15}$$

where $\otimes$ denotes the outer product.

In the decision-directed mode, the LMS algorithm is augmented with a MLE block to estimate the transmitted symbol $\hat{\hat{s}}_k$, as shown in Fig. 51.

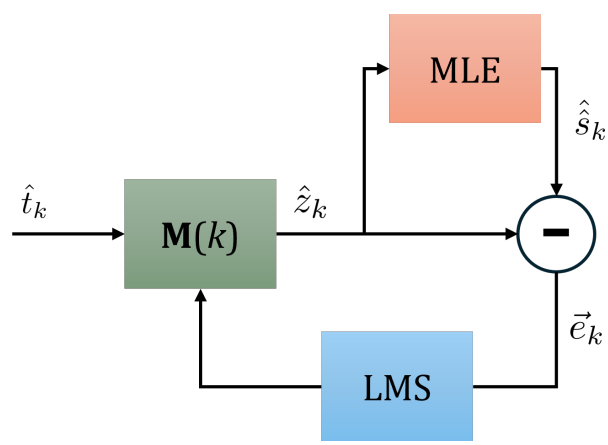



Figure 51: Least-mean squares decision-directed adaptive channel equalizer.

The same mathematical framework applies, but the error vector is redefined as:

$$\vec{e}_k = \hat{\hat{s}}_k - \hat{z}_k. \tag{5.16}$$

Maximum likelihood estimation involves making symbol decisions $\hat{\hat{s}}_k$ by comparing the inner product between the received Stokes vector $\hat{z}_k$ and the set of nominal Stokes vectors from the (6,64)-MVM constellation, represented by the $64 \times 35$ matrix $\hat{\mathbf{S}}$. The symbol decision $\hat{\hat{s}}_k$ corresponding to the largest inner product is selected using the equation below:

$$\hat{\hat{s}}_k \quad := \quad \text{argmax}\left(\hat{z}_k \cdot \hat{s}\right), \hat{s} \in \hat{\mathbf{S}}. \tag{5.17}$$

In the DSP of the (6,64)-MVM simulation, the LMS algorithm is employed to perform polarization derotation of the constellation in the generalized Stokes space. In this implementation, all received Stokes vectors are utilized, with the first 3,000 symbols assumed to be known to the receiver in a pilot-aided (PA) fashion. After this, the decision-directed (DD) approach is used for the remaining symbols. The simulation, conducted with a signal-to-noise ratio of $\gamma_b = 10\,\text{dB}$, is shown in Fig. 52.

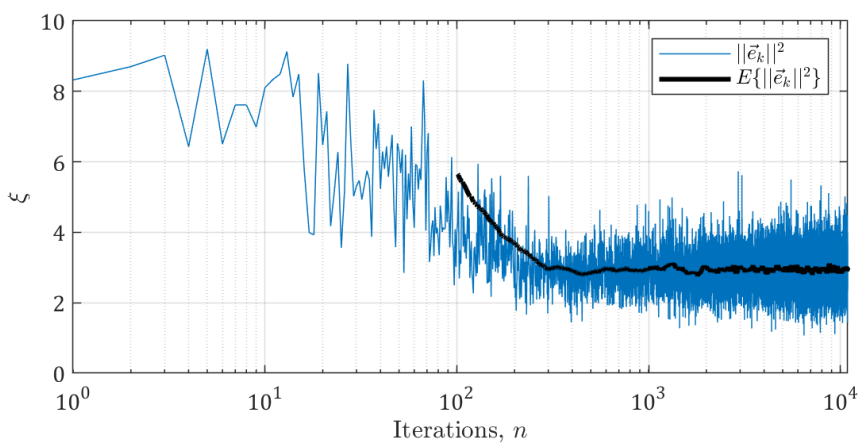



Figure 52: Simulation of the PA+DD LMS equalizer for (6,64)-MVM.

As illustrated in the figure, the cost function $\xi$ decreases after approximately 500 known pilot symbols are used for equalization in the PA phase. Due to fundamental noise

limitations, $\xi$ will not reach zero. After 3,000 symbols, when the pilot symbols are turned off and the DD approach is adopted, $\xi$ remains stable, demonstrating the LMS algorithm's effectiveness. After each symbol is processed by the LMS, the pilot symbols are discarded and not used for determining the BER.

## Simulation Results

After polarization derotation using the LMS algorithm, a final iteration of MLE is performed to map the symbols to bits, utilizing the predefined coding for the constellation; thereafter, the BER is evaluated. All BER calculations are averaged over 10 channel realizations. For the initial simulations, the target intermodal crosstalk per unit length is set to -30 dB/km. The simulation is conducted under two conditions: with or without optical fiber DMD compensation. These conditions are referred to as "MD managed" and "MD unmanaged", respectively.

In the "MD unmanaged" scenario, the MD parameter must remain relatively low, typically on the order of fractions of 1 ps/km, to prevent significant performance penalties. This is especially crucial as the symbol rate increases from 56 GBd to 112 GBd. The results for this case are presented in Fig. 53. Additionally, the figure includes a scenario with much higher DMD (on the order of tens of ps/km), but with DMD fiber management. The results show that optical fibers with opposite-sign DMD can fully compensate for modal dispersion up to 48 ps/km for both symbol rates.

The next study examined the effects of varying levels of crosstalk under fixed, low MD levels (on the order of fractions of 1 ps/km) and no DMD fiber management, as illustrated in Fig. 54. In the absence of DMD fiber management, the results align with those shown in Fig. 53, where the penalty from DMD becomes more significant as both DMD and symbol rate increase. However, when DMD is kept low, the impact of crosstalk remains negligible up to 0 dB/km. This is likely because the effects of uncompensated DMD appear to overshadow

any potential penalty from crosstalk. One possible explanation of this has to do with the LMS algorithm's ability to compensate polarization coupling as long as the power coupling between modes is on the order of the symbol period.

When MD increases, e.g., up to 24 ps/km, and DMD fiber management is employed, the influence of crosstalk becomes more pronounced. In this scenario, strong crosstalk around 0 dB/km can contribute up to a 1 dB penalty at the KP4 FEC threshold, as shown in Fig. 55. This effect is likely due to the accumulation of crosstalk before DMD compensation, which contains significant DMD (on the order of tens of ps/km). The combined effects of large DMD and crosstalk result in power coupling between larger arrays of adjacent symbol components of the three propagating modes, leading to greater penalties at the receiver.

The initial simulation results show that optical fiber compensation for DMD eliminates the performance penalty associated with DMD in the (6,64)-MVM format, with the resulting curves matching the back-to-back performance. In contrast, without DMD compensation, MVM becomes highly sensitive to DMD, experiencing significant penalties starting at 0.4 ps/km. This increased sensitivity negates the advantage of (6,64)-MVM over 4-SVM. As DMD becomes more significant, the role of crosstalk also intensifies, likely because the LMS algorithm struggles to effectively equalize under its effects.

Alternatively, with DMD compensation, MVM appears to offer significant sensitivity benefits over both 4-SVM and 4-PAM, amounting to several dBs, even in the presence of substantial amounts of DMD and crosstalk.

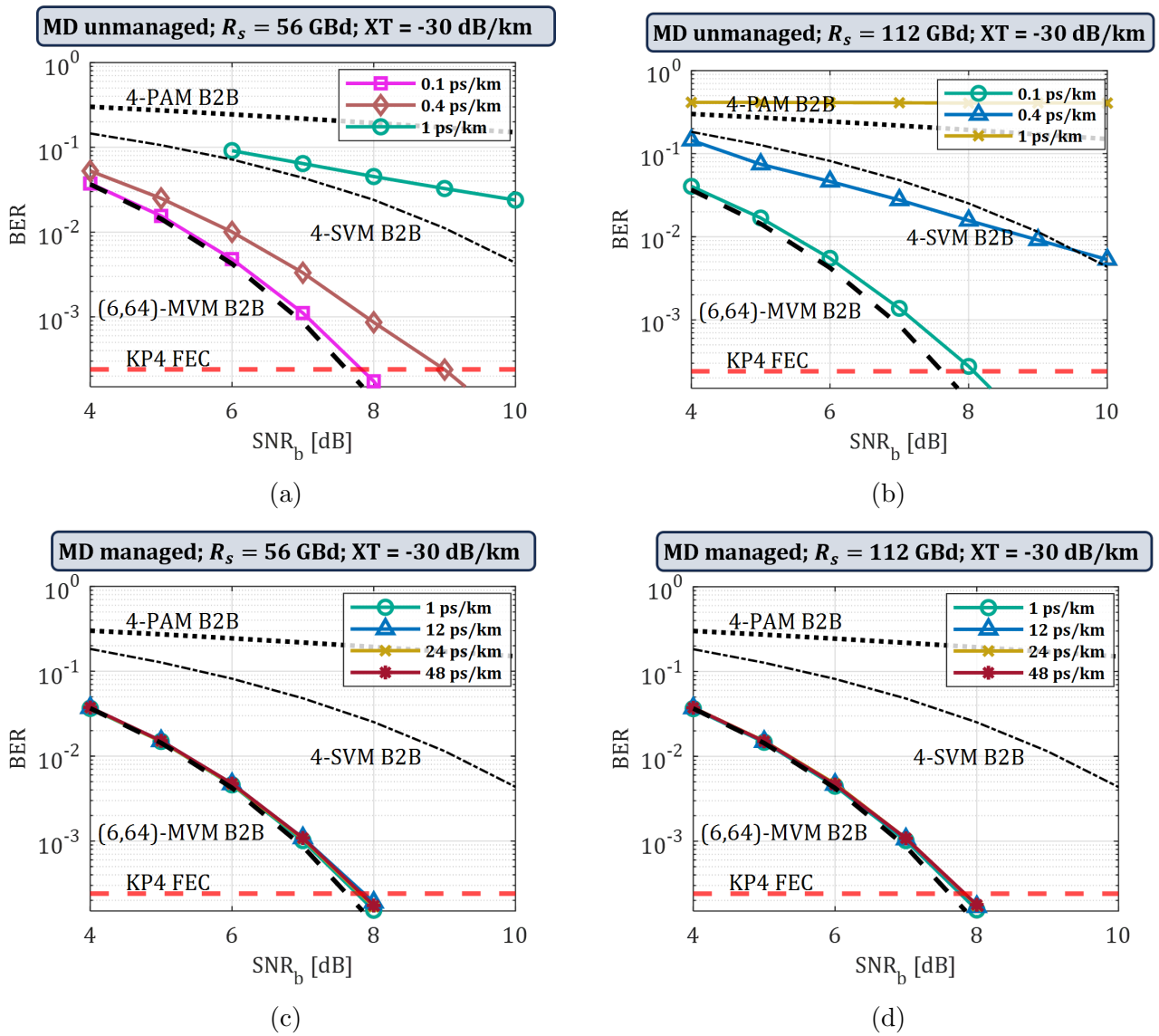



Figure 53: Simulation results of (6,64)-MVM through a 10 km FMF link assuming a crosstalk parameter XT = -30 dB/km. Each figure depicts the BER vs bit SNR with and without optical fiber DMD compensation. The figures are segmented into the following: (a) no fiber management at $R_s$ = 56 GBd; (b) no fiber management at $R_s$ = 112 GBd; (c) with fiber management at $R_s$ = 56 GBd; (d) with fiber management at $R_s$ = 112 GBd.

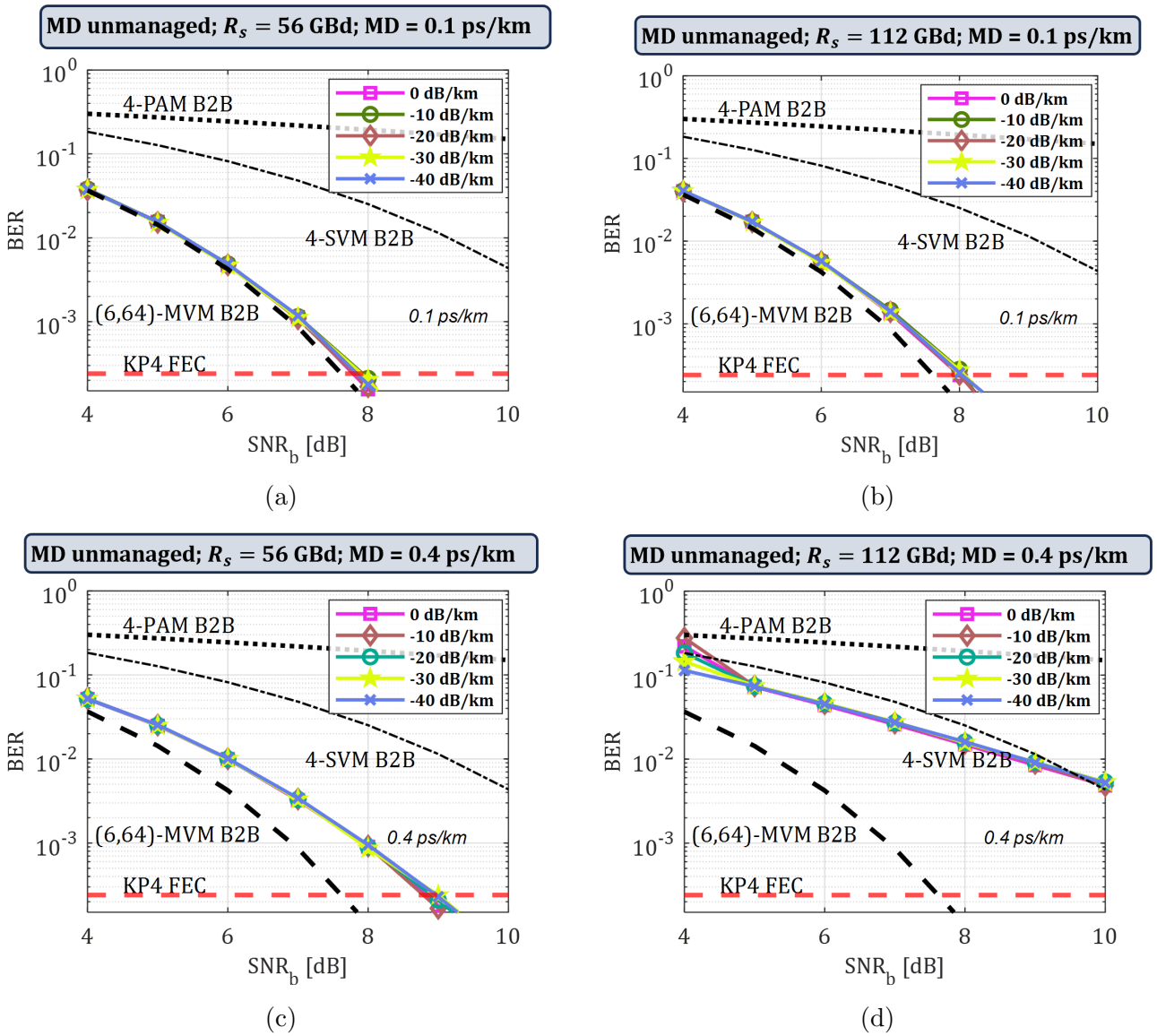



Figure 54: Simulation results of (6,64)-MVM through a 10 km FMF link assuming various levels of crosstalk and fixed MD. Each figure depicts the BER vs bit SNR without optical fiber DMD compensation. The figures are segmented into the following: (a) MD = 0.1 ps/km and $R_s$ = 56 GBd; (b) MD = 0.1 ps/km and $R_s$ = 112 GBd; (c) MD = 0.4 ps/km and $R_s$ = 56 GBd; (d) MD = 0.4 ps/km and $R_s$ = 112 GBd.

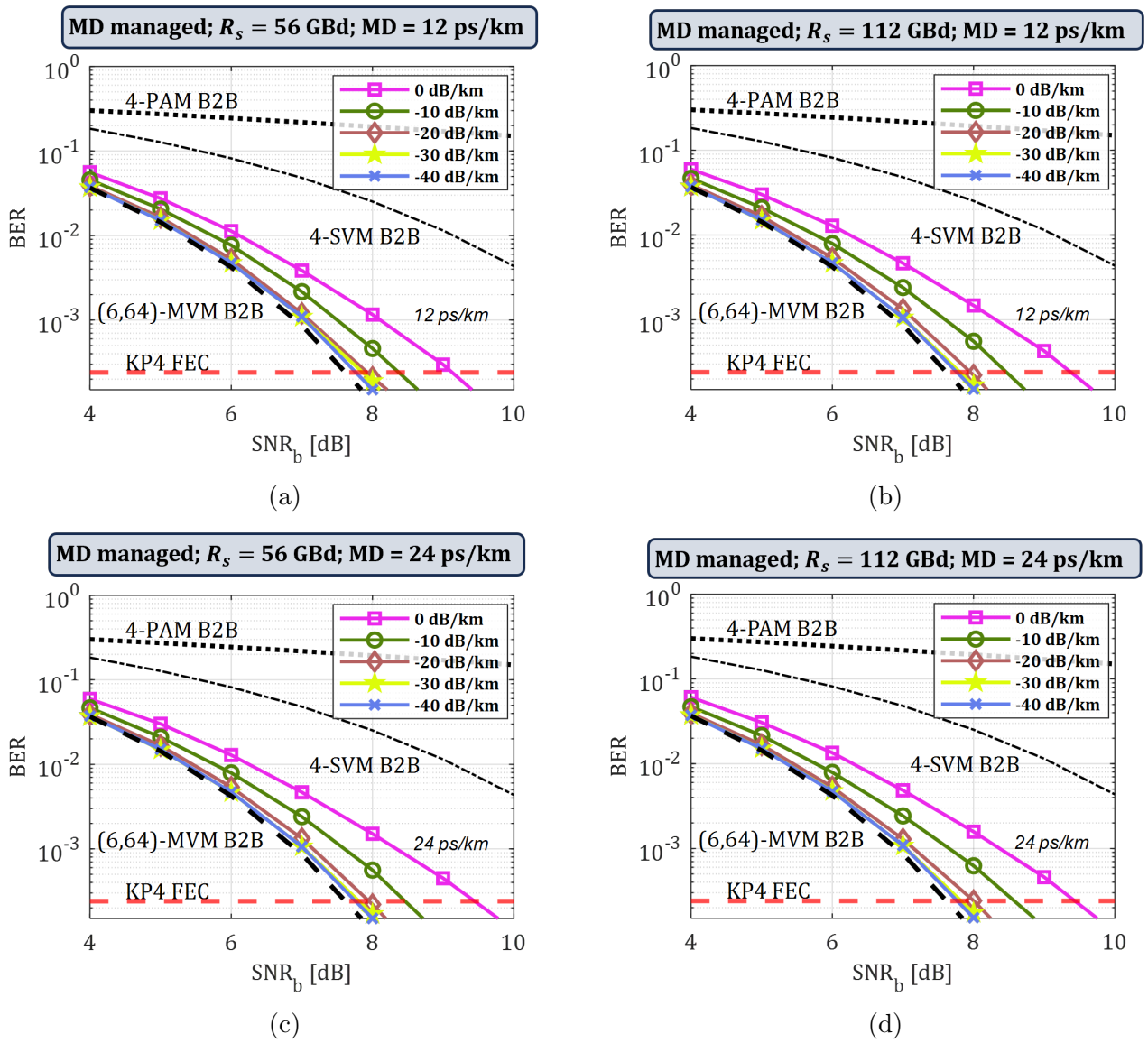



Figure 55: Simulation results of (6,64)-MVM through a 10 km FMF link assuming various levels of crosstalk and fixed MD. Each figure depicts the BER vs bit SNR without optical fiber DMD compensation. The figures are segmented into the following: (a) MD = 12 ps/km and $R_s$ = 56 GBd; (b) MD = 12 ps/km and $R_s$ = 112 GBd; (c) MD = 24 ps/km and $R_s$ = 56 GBd; (d) MD = 24 ps/km and $R_s$ = 112 GBd.

## Future Work

The MVM simulation environment provides a powerful platform for exploring complex signal distortions commonly encountered in real-world optical transmission systems. These distortions include chromatic dispersion, polarization mode dispersion, modal crosstalk, differential mode delay, and non-linearities, all of which significantly affect signal integrity and system performance. The simulation's flexibility also enables the exploration of more advanced MVM constellations, incorporating higher degrees of freedom and larger signal spaces with greater cardinalities. This capability is crucial for expanding the range and adaptability of MVM, allowing more sophisticated modulation formats to be tested, optimized, and compared against other direct-detection formats to meet the high-capacity and energy efficiency demands of the future.

Looking ahead, future efforts will focus on developing more robust simulations that not only account for all fiber propagation effects but also incorporate realistic MVM transmitter and receiver designs, considering the added complexities of hardware and signal processing.

# CONCLUSION

This thesis has explored the theoretical, experimental, and numerical aspects of Stokes Vector Modulation (SVM) and Mode Vector Modulation (MVM), advancing their potential as energy-efficient alternatives to conventional modulation formats for short-haul optical links. A key contribution was the successful mapping of SVM and MVM symbols to modulator driving voltages within a dual-polarization Mach-Zehnder Interferometer structure, leading to the first experimental demonstration of the theoretically optimal simplex SVM constellation. Additionally, an optimization procedure for modulator driving voltages was introduced, enabling the generation of the SVM simplex with a reduced implementation penalty. This optimization framework, which can be extended into higher dimensions for MVM, provides a systematic approach for improving the performance of these modulation formats and realizing their full theoretical potential in practical systems.

In parallel, a numerical study of (6,64)-MVM was conducted to evaluate the impact of modal dispersion and crosstalk, key impairments in multicore and few-mode fibers (MCF/FMF). The results numerically verified MVM's significant sensitivity advantages over 4-PAM, even in the presence of substantial transmission impairments. Additionally, the simulation framework developed in this work provides a foundation for assessing MVM under a broader set of transmission effects, including chromatic dispersion, polarization mode dispersion, and nonlinearities. These findings are essential for determining MVM's feasibility in intra-data center interconnects (DCNs), where balancing energy efficiency and spectral efficiency is critical.

The contributions of this research extend beyond theoretical and experimental advancements; they highlight the potential of optimally shaped SVM and MVM constellations to address critical challenges in modern optical communication systems. As data centers continue to experience exponential traffic growth, the demand for energy-efficient modulation

schemes becomes increasingly urgent. If SVM and MVM can be successfully implemented in short-haul optical networks, they could provide an optimal trade-off between energy consumption and spectral efficiency, mitigating the rising energy demands of high-capacity data centers.

Future research will build upon these results through ongoing experimental efforts, particularly with the development of a 2×2 multiple-input multiple-output (MIMO) experimental setup to validate MVM for $N$=4. The combination of simulation-based analysis and experimental verification will be essential for refining MVM's implementation and evaluating its practical advantages over existing direct-detection modulation schemes. By accurately modeling real-world fiber distortions and leveraging experimental feedback, this research aims to establish a robust framework for assessing MVM's viability in next-generation optical networks. Ultimately, the insights gained from this work will contribute to the development of high-performance, energy-efficient optical communication systems capable of meeting the growing demands of modern data centers.

# REFERENCES CITED

[1] S. Benedetto and P.T. Poggiolini. Multilevel polarization shift keying: optimum receiver structure and performance evaluation. *IEEE Transactions on Communications*, 42(234): 1174–1186, 1994. doi: 10.1109/TCOMM.1994.580226.

[2] S. Benedetto, R. Gaudino, and P. Poggiolini. Direct detection of optical digital transmission based on polarization shift keying modulation. *IEEE Journal on Selected Areas in Communications*, 13(3):531–542, 1995. doi: 10.1109/49.372412.

[3] Sergio Benedetto and PT Poggiolini. Multilevel polarization shift keying: Optimum receiver structure and performance evaluation. *IEEE Transactions on communications*, 42(234):1174–1186, 1994.

[4] S. Betti, F. Curti, G. De Marchis, and E. Iannone. Multilevel coherent optical system based on stokes parameters modulation. *Journal of Lightwave Technology*, 8(7):1127–1136, 1990. doi: 10.1109/50.56417.

[5] Stephen P Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[6] N.G. De Bruijn. A combinatorial problem. *Mathematics of Computation*, 3(19):97–98, 1946.

[7] Filipe Ferreira, Daniel Fonseca, and Henrique Silva. Design of few-mode fibers with arbitrary and flattened differential mode delay. *IEEE Photonics Technology Letters*, 25 (5):438–441, 2013. doi: 10.1109/LPT.2013.2241047.

[8] Filipe M Ferreira and Fabio A Barbosa. Maximizing the capacity of graded-index multimode fibers in the linear regime. *Journal of Lightwave Technology*, 2023.

[9] Filipe Marques Ferreira, Christian S. Costa, Stylianos Sygletos, and Andrew D. Ellis. Semi-analytical modelling of linear mode coupling in few-mode fibers. *Journal of Lightwave Technology*, 35(18):4011–4022, 2017. doi: 10.1109/JLT.2017.2727441.

[10] Flexential. Data Center Liquid Cooling: Deep Dive, 2024. URL `https://www.flexential.com/resources/blog/data-centers-liquid-cooling`.

[11] Lou Frenzel. Understanding modern digital modulation techniques. *Electronic Design*, July 2021. URL `https://www.electronicdesign.com/technologies/communications/article/21798737/electronic-design-understanding-modern-digital-modulation-techniques`.

[12] Christopher A Fuchs, Michael C Hoang, and Blake C Stacey. The sic question: History and state of play. *Axioms*, 6(3):21, 2017.

[13] James P. Gordon and Herwig Kogelnik. PMD fundamentals: polarization mode dispersion in optical fibers. *Proceedings of the National Academy of Sciences (PNAS)*, 97 9:4541–50, 2000.

[14] Md Sabbir-Bin Hossain, Jinlong Wei, Fabio Pittalà, Nebojša Stojanović, Stefano Calabrò, Talha Rahman, Tom Wettlin, Changsong Xie, Maxim Kuschnerov, and Stephan Pachnicke. Single-lane 402 gb/s pam-8 im/dd transmission based on a single dac and an o-band commercial eml. In *Optoelectronics and Communications Conference*, pages T5A–6. Optica Publishing Group, 2021.

[15] E. Hu, Y. Hsueh, K. Wong, M. Marhic, L. Kazovsky, K. Shimizu, and N. Kikuchi. 4-level direct-detection polarization shift-keying (dd-polsk) system with phase modulators. In *OFC 2003 Optical Fiber Communications Conference, 2003.*, pages 647–649 vol.2, 2003. doi: 10.1109/OFC.2003.316199.

[16] EXFO Inc. Chromatic dispersion and its impact on optical networks, 2023. URL `https://www.exfo.com/en/`.

[17] P.M. Williams K.E. Strom. de bruijn sequence generator, 2010. URL `https://www.mathworks.com/matlabcentral/fileexchange/28015-de-bruijn-sequence-generator`.

[18] Kazuro Kikuchi and Shojiro Kawakami. Multi-level signaling in the stokes space and its application to large-capacity optical communications. *Optics express*, 22(7):7374–7387, 2014.

[19] Jaroslaw Kwapisz, Ioannis Roudas, and Eric Fink. Mode vector modulation: Extending stokes vector modulation to higher dimensions. *Journal of Lightwave Technology*, 2023.

[20] B. P. Lathi and Zhi Ding, 1962. *Modern digital and analog communication systems*. Oxford series in electrical and computer engineering. Oxford University Press, New York, 4th ed edition, 2009. ISBN 978-0-19-533145-5 0-19-533145-1 978-0-19-538493-2 0-19-538493-8.

[21] Yiming Lin, Jinjing Tao, Xueguang Yuan, Yangan Zhang, and Jinnan Zhang. A novel structure of 4-polsk system using on free-space optical communications. In *Asia Communications and Photonics Conference 2013*, page AF2F.44. Optica Publishing Group, 2013.

[22] Weijie Liu, Zhangwei Ye, Nuo Huang, Shangbin Li, and Zhengyuan Xu. Multilevel polarization shift keying modulation for turbulence-robust underwater optical wireless communication. *Optics Express*, 31(5):8400–8413, 2 2023. doi: 10.1364/OE.483397.

[23] Antonio Mecozzi, Cristian Antonelli, and Mark Shtaif. Kramers–kronig coherent receiver. *Optica*, 3(11):1220–1227, 2016.

[24] Mohamed Morsy-Osman, Mohammed Sowailem, Eslam El-Fiky, Tristan Goodwill, Thang Hoang, Stephane Lessard, and David V Plant. Dsp-free 'coherent-lite'transceiver for next generation single wavelength optical intra-datacenter interconnects. *Optics Express*, 26(7):8890–8903, 2018.

[25] Mohamed Morsy-Osman, Md Samiul Alam, Kh Arif Shahriar, Stephane Lessard, and David V Plant. Optimum three-dimensional constellations for stokes vector direct detect receivers. *IEEE Photonics Technology Letters*, 31(8):587–590, 2019.

[26] Jose Krause Perin. *Spectrally and Power Efficient Optical Communication Systems*. PhD thesis, Stanford Unversity, CA, 2018.

[27] Jose Krause Perin, Anujit Shastri, and Joseph M Kahn. Coherent data center links. *Journal of Lightwave Technology*, 39(3):730–741, 2020.

[28] Roi Rath, Dennis Clausen, Simon Ohlendorf, Stephan Pachnicke, and Werner Rosenkranz. Tomlinson–Harashima precoding for dispersion uncompensated PAM-4 transmission with direct-detection. *Journal of Lightwave Technology*, 35(18):3909–3917, 2017. doi: 10.1109/JLT.2017.2724032.

[29] Ken-ichi Sato. Optical switching will innovate intra data center networks [invited tutorial]. *Journal of Optical Communications and Networking*, 16(1):A1–A23, 2023.

[30] Semiengineering. Chip aging becoming key factor in data center economics. *Semiengineering*, 2021. URL `https://semiengineering.com/chip-aging-becoming-key-factor-in-data-center-economics/`.

[31] Statista. Energy consumption of Google from 2011 to 2021, 2022. URL `https://www.statista.com/statistics/788540/energy-consumption-of-google/`.

[32] Cadence Design Systems. Future facilities: The future of integrated circuit design and data center architecture, 2021. URL `https://www.cadence.com/content/dam/cadence-www/global/en_US/documents/tools/cimdata-commentary-cadence-future-facilities.pdf`.

[33] Michael G. Taylor. Phase estimation methods for optical coherent detection using digital signal processing. *J. Lightwave Technol.*, 27(7):901–914, Apr 2009.

[34] Inc. Tektronix. *OM1106 Optical Modulation Analysis Software*. Tektronix, Inc.

[35] Marcos Troncoso-Costas, Devika Dass, Colm Browning, Francisco J Diaz-Otero, Chris GH Roeloffzen, and Liam P Barry. Intra-data centre flexible pam transmission system using an integrated inp-si 3 n 4 dual laser module. *IEEE Photonics Journal*, 14 (1):1–6, 2021.

[36] Jinlong Wei, Liang Zhang, Cristian Prodaniuc, Nebojša Stojanović, and Changsong Xie. Linear pre-equalization techniques for short reach single lambda 225 gb/s pam imdd systems. In *2018 European Conference on Optical Communication (ECOC)*, pages 1–3. IEEE, 2018.

[37] Bernard Widrow and Marcian E. Hoff. Adaptive switching circuits. *1960 IRE WESCON Convention Record*, pages 96–104, 1960.

[38] Wikipedia contributors. Data center network architectures, 2024. URL `https://en.wikipedia.org/wiki/Data_center_network_architectures`.

[39] Wikipedia contributors. Thomson problem — Wikipedia, The Free Encyclopedia, 2025. URL `https://en.wikipedia.org/wiki/Thomson_problem`.

[40] Junhe Zhou and Haoqian Pu. Generalized analytical study on the random crosstalk in multicorex002f;multimode fibers. *Journal of Lightwave Technology*, 40(4):1112–1120, 2022. doi: 10.1109/JLT.2021.3127865.

[41] Xiang Zhou, Ryohei Urata, and Hong Liu. Beyond 1 tb/s intra-data center interconnect technology: Im-dd or coherent? *Journal of Lightwave Technology*, 38(2):475–484, 2019.

# APPENDICES

# APPENDIX A

## ACRONYMS

| Acronym | Description |
|---|---|
| ABC | Automatic Bias Controller |
| ADC | Analog-to-Digital Converter |
| AI | Artificial Intelligence |
| ASE | Amplified Spontaneous Emission |
| AWG | Arbitrary Waveform Generator |
| AWGN | Additive White Gaussian Noise |
| BER | Bit Error Rate |
| BPSK | Binary Phase Shift Keying |
| CD | Chromatic Dispersion |
| CW | Continuous Wave |
| DAC | Digital-to-Analog Converter |
| DCN | Data Center Network |
| DD | Direct-Detection/Decision-Directed |
| DMD | Differential Mode Delay |
| DOF | Degree of Freedom |
| DSB-SC | Double Sideband Suppressed Carrier |
| DSP | Digital Signal Processing |
| EDFA | Erbium Doped Fiber Amplifier |
| EM | Electromagnetic |
| EML | Electro-Absorption Modulated Laser |
| EVM | Error Vector Magnitude |
| FEC | Forward Error Correction |
| FFT | Fast Fourier Transform |
| FIR | Finite Impulse Response |
| FM | Frequency Modulation |
| FMF | Few-Mode Fiber |
| HPC | High Performance Computing |
| IF | Intermediate Frequency |
| IQ | In-phase and Quadrature |
| ISI | Intersymbol Interference |
| KK | Kramers-Kronig |
| LMS | Least Mean Squares |
| LO | Local Oscillator |
| MBL | Macro-Bend Loss |
| MCF | Multicore Fiber |
| MD | Modal Dispersion |
| MIMO | Multiple Input Multiple Output |
| ML | Machine Learning |
| MLE | Maximum Likelihood Estimation |
| MSU | Montana State University |

| | |
|---|---|
| MVM | Mode Vector Modulation |
| MZI | Mach-Zehnder Interferometer |
| MZM | Mach-Zehnder Modulator |
| NRZ-OOK | Non-Return-to-Zero On-Off Keying |
| OI | Optical Interface |
| OMA | Optical Modulation Analysis |
| OSA | Optical Spectrum Analyzer |
| OSFP | Octal-Small Form Factor Pluggable |
| PA | Pilot-Aided |
| PAM | Pulse Amplitude Modulation |
| PC | Personal Computer |
| PSK | Phase Shift Keying |
| QAM | Quadrature Amplitude Modulation |
| QSFP | Quad-Small Form Factor Pluggable |
| RCL | Recirculating Loop |
| RRC | Root Raised Cosine |
| SDOF | Spatial Degree of Freedom |
| SER | Symbol Error Rate |
| SIC-POVM | Symmetric, Informationally Complete, Positive Operator-Valued Measures |
| SMF | Single Mode Fiber |
| SNR | Signal-to-Noise Ratio |
| SOP | State of Polarization |
| SVM | Stokes Vector Modulation |
| UCL | University College London |
| VOA | Variable Optical Attenuator |
| XT | Crosstalk |

# APPENDIX B

## MATLAB CODE

Various MATLAB scripts were developed to optimize the SVM simplex constellation and are available in my personal GitHub repository:
https://github.com/alexbrisson/MS-Thesis-4SVM. The primary scripts, `GradientAscent_SVM` and `GradientAscent_SVM_OUI`, handle the optimization and SVM digital signal processing, respectively. Additional functions support various processing tasks, including calculations, conversions, and signal generation.