

UNIVERSITY OF OTTAWA
MASTERS' THESIS

Photonic Integration with III-V Semiconductor Technologies

Author:

Tuhin PAUL

Supervisor:

Dr. Ksenia DOLGALEVA

*A thesis submitted in fulfillment of the requirements
for the degree of Masters of Science in Physics*



uOttawa

Faculté des sciences
Faculty of Science

Department of Physics
University of Ottawa

© Tuhin Paul, Ottawa, Canada, 2022

Declaration of Authorship

I, Tuhin PAUL, declare that this thesis titled, “Photonic Integration with III-V Semiconductor Technologies” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

“My whole life had been designing computers which I could never build.”

Steve Wozniak

Abstract

This dissertation documents works on two projects, which are broadly related to photonic integration using III-V semiconductor platform for fiber-based optical communication. Our principal project aims to demonstrate continuous variable quantum key distribution (CV-QKD) with InP-based photonic integrated circuit at the 1550 nanometer of optical wavelength. CV QKD protocols, in which the key is encoded in the quadrature variables of light, has generated immense interest over the years because of its compatibility with the existing telecom infrastructure. In this thesis, we have proposed a design of a photonic integrated circuit potentially capable of realizing this protocol with coherent states of light. From the practical perspective, we have basically designed an optical transmitter and an optical receiver capable of carrying out coherent communication via the optical fiber. Initially, we established a mathematical model of the transceiver system based on the optical transfer matrix of the foundry specific (Fraunhofer Heinrich Hertz Institute-Germany) building blocks. We have shown that our chip design is versatile in the sense that it can support multiple modulation schemes. Based on the mathematical model, we estimated the link budget to assess the feasibility of on-chip implementation of our protocol. Then we ran a circuit level simulation using the process design kit provided by our foundry to put our analysis on a better footing. The encouraging result from this step prompted us to generate the mask layout for our transceiver chips, which we eventually submitted to the foundry. The other project in the thesis grew out of a collaboration with one of our industry partners. The goal of the project is to enhance the performance of a distributed feedback laser emitting at the 1310 nanometer of optical wavelength by optimizing its design. To that end, we first derived the expression for transmission and reflection spectrum for the laser cavity. Those expressions contained parameters which needed to be obtained from the transverse and the longitudinal mode analysis of the laser. We performed the transverse mode analysis and the longitudinal mode analysis with commercially available numerical solvers. Those mode profiles critically depend on the grating physical parameters. Therefore by tweaking grating dimensions one can control the transmission characteristics of the laser.

Keywords: Gaussian Quantum Information, Continuous Variable Quantum Key Distribution, Phase Space Quantum Optics, Private Communication, Photonic Integrated Circuit, Indium Phosphide, Distributed Feedback Laser, Coupled Mode Theory, Diffraction Grating, Coherent Optical Communication

Acknowledgements

I feel exceedingly grateful to work as a graduate student under the supervision of Professor Ksenia Dolgaleva. First of all, I thank her profusely for letting me work in the CV QKD project, which has been very close to my heart, as part of masters' thesis. From the very beginning, we both knew the amount of hurdles we have to get over in order to bring this project into fruition. Nevertheless she has encouraged me in every possible way. As a mentor she has been amazing. I have grown as a researcher under her guidance. As a human being, she is wonderful. She is very empathetic toward students' needs and struggles. This very attribute has been very reassuring for me in my early days while I was still adjusting to the new environment.

I am particularly thankful to Professor Jeff Lundeen who virtually co-supervised me on the CV QKD project. He has also provided me constant support throughout my research. Specially his deep knowledge on the quantum aspects of the project has helped me in my design analysis.

I would like to express my profound gratitude to Valery Tolstikhin from Inten-gent Inc. for formulating the project on distributed feedback laser. Working for this project has been a quite a learning experience for me. I also garnered valuable industry insights from our occasional meeting.

Special thanks to our post doctoral fellow Boris Rosenstein Levin who has mentored me on the DFB laser project. He had always been available every time I needed to help. I cannot thank enough Sina Aghili, PhD student in the group, whose work on the same project has been a source of reference for me.

I am deeply indebted to Fabian Laudenbach (Xanadu) with whom I have been in touch on numerous occasions. We shared the common research interest. During his time in Austrian Institute of Technology he was involved in experimental CV QKD. His publication on the practical implementation of CV QKD is the source which I can single out for furnishing me with the most background details on the subject matter. I would like to thank Ronald Broeke from Bright Integrated photonics for mentoring me in producing the mask layout of my photonic integrated circuit design. Working with him has been the most rewarding experience for me.

At the various stages of my thesis writing I have communicated privately with a number of experts either to clarify or to solidify my understanding in their area of expertise. I would like to convey my gratitude to the following list of people: Fred Kish (North Carolina State University) for Chapter 1; Mark Wilde (Louisiana State University), Andreas Winter (Autonomous University of Barcelona, Spain), Alessio Serafini (University College London, UK), Carlos Navarrete-Benlloch (Shanghai Jiao Tong University, China), and Anthony Leverrier (INRS, France) for Chapter 2; Shiva Kumar (McMaster University), Moritz Baier, and Durvasa Das Gupta (Fraunhofer HHI, Berlin) for Chapter 3.

I fondly acknowledge the contribution of my fellow colleagues from Advanced Research Complex. Without their presence research would have not been as

interesting. I would specially mention my research group colleague Kaustubh Vyas with whom I also shared my office space. He was instrumental in maintaining a lighthearted environment in the office. Without his help it would have been impossible for me to work remotely at the later stage of my thesis writing. I would also like to mention Dilip Paneru, Jean-Luc Begin, Jacob Manalo, and Yasser Salem for the many stimulating conversations we had on physics over a cup of tea. Last but not the least, I thank my family in Bangladesh. Without their unwavering support, I would have been clueless, specially in the trying times of the global pandemic. I owe them for everything.

Contents

Declaration of Authorship	ii
Acknowledgements	v
1 Photonic Integration Technology	1
1.1 Introduction	1
1.2 Motivation for Photonic Integration	1
1.3 Conceptual Placement of this Thesis	6
1.4 Generic Model for Photonic Integration	8
1.4.1 Generic Integration Model	9
1.4.2 Generic Foundry Model	11
1.4.3 Multi-project Wafer Run	11
1.4.4 Generic InP-based PIC technology	12
1.5 ASPIC Design Environment	12
1.5.1 Component Design & Optimization	13
1.5.2 Circuit Level Simulation	14
Simulating in the frequency domain	14
Simulating in the time domain	14
1.5.3 Circuit Layout Generation	15
1.5.4 Verification	15
1.5.5 PDK	16
1.6 Conclusion	16
2 Continuous Variable Quantum Key Distribution	17
2.1 Introduction	17
2.2 Quantum Key Distribution	18
2.3 Gaussian Quantum Optics	20
2.3.1 Wigner Distribution	20
2.3.2 Examples of Gaussian States	25
2.3.3 Gaussian Operations	27
Gaussian Unitary Transformation	27
Examples of Gaussian Unitaries	28
2.3.4 Gaussian Measurements	29
Partial Measurements	29
2.3.5 Entropy of a Gaussian State	30
2.4 Basics of a coherent state CV QKD protocol	32
2.5 Gaussian Modulation	34
2.6 Entanglement based version of CV QKD	35
2.6.1 Discrete Modulation	40

2.7	Transmission and Noise	43
2.8	State Measurement	44
2.9	Estimating the secret key rate	45
2.10	Signal to Noise Ratio and Mutual Information	45
2.11	Calculation of Holevo Information	46
2.11.1	Universal Analysis of Purification Attacks	47
	Homodyne Detection	48
	Heterodyne Detection	48
2.12	Practical Parameter Estimation	49
2.13	Security Aspect of CV QKD	50
2.14	Conclusion	52
3	On Chip Implementation of CV QKD	54
3.1	Introduction	54
3.2	High Level Overview	54
3.3	Description of the Building Blocks	56
3.4	Mathematical description of the whole system	77
3.5	Circuit Level Schematics	81
3.5.1	Power Budget Estimation	100
3.6	Mask Layout Generation	104
3.7	Wafer Validation	106
3.8	Conclusion	107
4	Electromagnetic Analysis of a DFB Laser	109
4.1	Introduction	109
4.2	Distributed Feedback Laser	109
4.3	Transverse Mode Analysis	112
4.4	Longitudinal Mode Analysis	119
	Making use of boundary conditions:	121
4.4.1	Transmission and Reflection Spectrum	124
4.5	Conclusion	125
5	Numerical Analysis of a DFB Laser	126
5.1	Introduction	126
5.2	Fundamental Mode Profile	127
5.3	Partial Wave Analysis	129
5.4	Longitudinal Mode Profile	130
5.5	Transmission & Reflection Spectrum	131
5.6	Conclusion	133
6	Final Remarks	134
A	Private Classical Communication	137
A.1	Classical Shannon Theory	137
A.1.1	Notions from discrete probability	138
A.1.2	Classical Typicality	139
	Weak Typicality	139
	Strong Typicality	141

A.1.3	Source Coding	143
A.1.4	Channel Coding	146
A.2	Notions from Quantum Mechanics	149
A.3	Quantum Information-theoretic Measures	151
A.4	Quantum Typicality	152
A.5	The Classical Capacity Theorem	156
A.6	Private Classical Capacity Theorem	160
A.7	Entropy with continuous variables	162
	Gaussian Distribution	162
B	Phase Space Quantum Mechanics	166
B.1	Wigner Characteristic Function	166
B.2	Properties of Wigner Function	168
B.3	Admissible condition for a covariance matrix	169
B.4	Gaussian Unitaries	171
B.5	Symplectic Analysis of Gaussian States	171
	B.5.1 Two mode Gaussian States:	173
B.6	Hilbert Space vs Phase Space	174
B.7	Conditional Gaussian Dynamics	174
C	Discrete Modulation	177
C.1	Discrete modulation equations derivation	177
D	Complementary to Chapter 3	181
D.1	Working Principle of a 2×2 3-dB MMI	181
D.2	Derivation of equation (3.21)	184
D.3	Gaussian Modulation	185
D.4	Half Wave Plate	187
D.5	8-PSK with MZM:	188
D.6	Eye Diagram	189
E	Complementary to Chapter 4	192
	Bibliography	206

List of Figures

1.1	An artist's rendition of a photonic integrated circuit showcasing different optical components reproduced from [2].	2
1.2	Scale comparison between an optical table set up and an optical chip[3]	3
1.3	Decomposing an N port unitary $U(N)$ into $SU(2)$ group elements: beam splitters and phase shifters [5]	3
1.4	Block diagram representation of a coherent communication system. The diagram also captures the essence of a CV-QKD system realization. Symbols: BS=Beam Splitter, MZM=Mach Zehnder Modulator,PBC=Polarization Beam Combiner, DSP= Digital Signal Processing, DAC=Digital to Analog Converter, LO = Local Oscillator, PBS=Polarization Beam Splitter, TIA= Trans Impedance Amplifier, ADC= Analog to Digital Converter)	4
1.5	Waveguide structures for the InP, SOI (Silicon-on-insulator), and $Si_3N_4 - SiO_2$ material platforms [9]	5
1.6	Illustration of taper assisted multi-guide vertical integration principle.	7
1.7	Examples of the functionalities that can be realized with (a) passive waveguide devices alone, or in combination with (b) optical amplifiers and (c) phase modulators [28]. The second component in (c) shows a Mach Zehnder Interferometer (MZI). This composite building block (CBB) consists of simpler building blocks such as MMI's, waveguide sections, and phase shifters. The MMI is another CBB made out of single mode and multi mode waveguide sections (BBBs).	9
1.8	Schematic diagram of the hierarchical structure for the different building blocks [28].	10
1.9	Example of a multi-project wafer (MPW) realised in the COBRA process. The wafer has 52 cells in the usable wafer area. Each ASPIC is copied at four different locations on the wafer with room for 12 users and one test cell for wafer validation. The figure shows one test cell and its position on the wafer. In this way, each user will receive four copies of his chips per wafer in the MPW batch. [28]	11
1.10	Flow involved in an ASPIC design environment [29]	13

- 2.1 The figure illustrates the idea behind the BB84 protocol. Alice's key consists of random sequences of classical bits $\{0, 1\}$. She has two non orthogonal basis sets $\{|0\rangle, |1\rangle\}, \{|+\rangle, |-\rangle\}$ at her disposal which she can employ to encode her key bit. She also picks her basis states randomly. For the choice of $\{|0\rangle, |1\rangle\}$, the mapping corresponds to $0 \mapsto |0\rangle$ and $1 \mapsto |1\rangle$ while for the choice of $\{|+\rangle, |-\rangle\}$, it is $0 \mapsto |+\rangle$ and $1 \mapsto |-\rangle$. Bob knows about the basis sets Alice uses for encoding. But he does not have any knowledge about which random basis Alice has picked. He performs a measurement by randomly choosing a basis on the transmitted state to decode the key. If by chance, Alice's and Bob's choice of basis is the same, Bob will be able to read out the correct key bit (columns 1,3,6,8). But if Eve performs an intermediate measurement, it will randomize Bob's measurement result even though he is matched with Alice for the choice of basis 19
- 2.2 (a) Phase space representation of Gaussian modulation. The red fuzzy circle denotes one of the many coherent states prepared by Alice. The fuzziness is due to the fluctuation associated with the quadrature operator measurements on a *single* coherent state. The mean values of both the quadratures for *all* the coherent states follow a Gaussian distribution. The standard deviation of this distribution is $\sqrt{V_{\text{mod}}}$ [54] (b) The state ensemble effectively looks like a thermal state (the outer circle) with variance V_{mod} 34
- 2.3 Entanglement based scheme of CV QKD. The 50 : 50 beam splitters characterize the measurements on Alice's and Bob's sides; while the transmittance T , and noise ξ are controlled by Eve. The QM box represents Eve's quantum memory [59]. 36
- 2.4 (a) (b) (c) Encoding schemes used for the two-state, four-state, and eight-state protocol respectively. 40
- 2.5 Comparison between the correlations: Z_G for Gaussian modulation (solid red), Z_2 for 2-state modulation (dotted blue), Z_4 for 4-state modulation (dotted green), Z_8 for 8-state modulation (dotted black). It is to be noted that for variance $V_{\text{mod}} \leq 1$, Z_G and Z_8 are almost indistinguishable, meaning that in this regime, one has the Holevo bound $S_8(B, E) \approx S_G(B, E)$. Therefore, the security bounds for the Gaussian modulation can be applied here as well. 42
- 3.1 Block diagram of the transmitter architecture. The acronyms used in the diagram refer to DFBL = Distributed Feedback Laser; TMZI = Tunable Mach Zehnder Interferometer; MZM = Mach Zehnder Modulator; PC = Polarization Converter; TPBS = Tunable Polarization Beam Splitter; SSC = Spot Size Converter. 55
- 3.2 Block diagram of the receiver architecture. The acronyms used in the diagram refer to MMI = Multimode Interferometer; PM = Phase Modulator; PD = Photo Detector; EPU = Electronic Processing Unit. The Σ symbol refers to the electronic summer/subtractor. 55

3.3	Schematic of generic InP PIC platform as developed at Fraunhofer HHI (BG = Bragg grating; PD = photodiode; DFB = DFB laser; EAM = electro-absorption modulator; SOA = semiconductor optical amplifier)[73], [74]	57
3.4	(a) Transfer matrix of a 2×2 component with (a_1, a_2) input amplitudes and (b_1, b_2) output amplitudes. Two components arranged in cascaded configuration (b) have an overall transfer matrix which is the product of the transfer matrices of both the components (c)[9]	58
3.5	In a slab waveguide the core and the claddings have rectangular geometry. The cross section is assumed to be in the $x - y$ plane. The step index change is along the $\hat{x}$ axis. The mode propagates along $\hat{z}$ axis. Throughout this document we stick to this convention. The mode picks up a phase factor as it travels from $z = 0$ to $z = L$	58
3.6	Different waveguide sections available from Fraunhofer-HHI. The waveguides consist of InP/InGaAsP layers. The rib width w is $2 \mu m$. The rib height h determines the index contrast and correspondingly the minimum bend radius R_{min} [73]	59
3.7	A tunable phase shifter consists of a phase actuator of length d mounted on top of a waveguide section.	60
3.8	(a) An MMI coupler with N inputs and M outputs with a multimode region in between. (b) Schematic of the 2×2 MMI from HHI which we have used in our design. This building block is equivalent of a symmetric power splitter with a $\pi/2$ relative phase between the two outputs.	61
3.9	Schematic of a tunable MZI consisting of two MMIs at the input and the output. The phase shifters can be adjusted to produce outputs with any splitting ratio between 0 to 1.	62
3.10	(a) A diagrammatic view of an MZM showing the control voltages and the outputs. (b) Procedure to find out $\phi_{i,1}, \phi_{i,2}$ for $s_i = r_1 e^{j\theta_i}$	66
3.11	A diagrammatic view of an MZM with the control voltages and the outputs for binary phase shift keying.	67
3.12	A diagrammatic view of an MZM with the control voltages and the outputs for quadrature phase shift keying.	69
3.13	A diagrammatic view of implementing quadrature phase shift keying with cascaded MZMs.	70
3.14	A diagrammatic view of implementing 8-PSK with cascaded MZMs.	71
3.15	(a) SEM image of the slanted waveguide cross section from HHI. The waveguide width is defined at the top. $\varepsilon = 52^\circ$ in this work [87]. (b) Birefringence and modal tilt as a function of waveguide width. The slope of these curves become smaller for larger widths, making them more favorable for fabrication tolerance [88]. (c) Schematic presentation of the two-section slanted sidewall PRs with relevant parameters[87].	73

3.16	(a) 3D cartoon of the overall structure of the PR design including input-output waveguides in different colors. The asymmetry is flipped between the two sections. All transitions are tapered to minimize losses and reflection [89]. (c) Performance of a polarization converter building block consisting of a 2-section slanted waveguide structure. Simulated polarization extinction ratio (PER) as a function of waveguide top widths of the two sections. The dashed magenta lines are iso-PER contour lines. The first one for device $\alpha_1 = -\alpha_2 = \pi/4$, $\delta_1 = 0.25$, $\delta_2 = 0.75$ with waveguide width tolerance of 140 nm. The second one for device $\alpha_1 = -\alpha_2 = 0.21\pi$, $\delta_1 = 0.29$, $\delta_2 = 0.74$ with waveguide width tolerance of 240 nm. These parameters are obtained from numerical optimization of the Jones matrix tolerance equation (??). It is assumed that a deviation in width will be equal over both sections [88]. (b) Measurement results of PER as a function of deviation from the nominal width = $1.2 \mu\text{m}$ [87]	74
3.17	(a) Schematic view of the MZI-PBS [91] (c) Simulated dependence of birefringence on waveguide widths with etch depth of 1700 nm. At the waveguide width of $1.7 \mu\text{m}$ the group indices of the two polarization states are equal, and hence no birefringence at this point [92]. (b) PER as a function of heating current for a PBS with arm length of $200 \mu\text{m}$ [92].	75
3.18	(a) Optical gain plotted against pump current of a $500 \mu\text{m}$ long SOA. The SOA had an input and an output waveguide both terminating to single mode fiber via spot size converters (b) Small signal response vs bias current curve for a $200 \mu\text{m}$ long DFB laser [73]	76
3.19	Schematic set-up for the characterization of an on-chip DFB laser	81
3.20	The output waveform from the on-chip DFBL at the target wavelength of 1550 nanometer at 100 mA bias current at 300K temperature.	82
3.21	Output power vs bias current (LI) characteristics of the on-chip DFBL at the target wavelength of 1550 nanometer at 300K temperature.	82
3.22	Power spectral density of the on-chip DFBL at the target wavelength of 1550 nanometer at two different bias currents at 300K temperature.	82
3.23	Schematic set-up for the characterization of a photonic integrated MMI	83
3.24	The two output waveforms of an MMI with respect to its input.	83
3.25	Schematic set-up for the characterization of two sequential MMIs.	84
3.26	Input vs output waveforms when two MMI's are placed sequentially	84
3.27	Schematic view of the tunable beam splitter consisting of an input MMI, an output MMI, and two thermo-optic phase shifters on both the connecting arms.	85

3.28	Schematic set-up for the characterization of a tunable beam splitter. The expanded form of the Tunable BS component is shown in the previous figure.	86
3.29	Interference pattern obtained at the two output ports of the TBS subject to the variation in the control current for the phase shifter on the upper arm.	86
3.30	Schematic set-up for the characterization of a polarization converter. POL stands for optical polarization analyzer.	87
3.31	Polarization rotation by the polarization converter within a certain wavelength range.	87
3.32	Schematic set-up to analyze the behavior of a polarization beam splitter. Operation of only one heater suffices to produce the necessary phase differences.	88
3.33	Interference pattern obtained at the output ports of the PBS for the TE wave subject to the variation in the phase shifter control current.	88
3.34	Interference pattern obtained at the output ports of the PBS for the TM wave subject to the variation in the phase shifter control current.	89
3.35	Relative distribution of the TE and TM waves at Port 2 of the PBS subject to the variation in the phase shifter control current.	89
3.36	Relative distribution of the TE and TM waves at Port 4 of the PBS subject to the variation in the phase shifter control current.	89
3.37	Schematic showing a PBS acting both as a polarization multiplexer (the first one from the left) and a demultiplexer (the second one).	90
3.38	Schematic showing the reverse situation to the one shown in figure (3.37)	90
3.39	Relative distribution of the TE and TM waves at at Port 2 and Port 4 of the PBS subject to the variation in the phase shifter control current.	91
3.40	A Mach Zehnder modulator is composed of an input MMI, an output MMI, and two current injection phase modulators (CIPM) on both the connecting arms.	91
3.41	Schematic set-up for the characterization of an MZM.	92
3.42	Interference pattern obtained at the two output ports of the MZM subject to the variation in the control current for CIPM on the upper arm.	93
3.43	Schematic showing implementation of intensity modulation and direct detection with an MZM.	93
3.44	Non return to zero electrical modulating signal at the rate of 25 Gbps.	94
3.45	Optical waveforms observed at the output ports of the MZM corresponding to on off keying operation.	94
3.46	Photocurrents generated by the photodiodes as observed with an electrical oscilloscope in IMDD.	94

3.47	Schematic showing a current injection phase modulator in conjunction with an on chip distributed feedback laser.	95
3.48	Input and output waveforms of the CIPM.	95
3.49	Output waveforms from the CIPM at I_0 , $I_{\frac{\pi}{2}}$, and I_{π} bias currents.	96
3.50	Schematic illustrating implementation of binary phase shift keying with a single MZM in dual drive operation.	96
3.51	Non return to zero electrical modulating signal at the rate of 25 Gbps.	97
3.52	Photocurrents corresponding to the in-phase and quadrature-phase quadratures of the modulated signal.	97
3.53	Eye diagram results for P and Q quadrature in BPSK.	98
3.54	The signal statistic with respect to the decision boundaries.	98
3.55	Schematic illustrating implementation of quadrature phase shift keying with a single MZM in dual drive operation.	99
3.56	Photocurrents corresponding to the in-phase and quadrature-phase quadratures of the modulated signal in QPSK.	99
3.57	Eye diagram results for P and Q quadrature in QPSK.	100
3.58	The signal statistic with respect to the decision boundaries in QPSK.	100
3.59	The circuit level simulation for our transceiver without the interconnects.	102
3.60	The circuit level simulation for our transceiver with all the connecting waveguides.	103
3.61	The signal statistics, against the backdrop of the decision boundaries, of our transceiver system aiming to implement QPSK.	104
3.62	Mask layout of a photonic integrated transmitter for CV QKD.	105
3.63	Mask layout of a photonic integrated receiver for CV QKD. The receiver chip contains coherent receiver components for balanced detection as well as some test structures for characterization.	106
3.64	Spectrum of a DFB laser with a target wavelength of $1.54 \mu\text{m}$, measured at various currents and at (a) 20°C (b) 50°C . (c) DFB series resistance above threshold for various DFB wavelengths. DFB LIV curves obtained using a CW current (red) and a pulsed current (blue) for (d) 20°C (e) 50°C	107
4.1	(a),(b) Transmission spectrum of a Fabry-Perot laser and a DFB laser [24]. (c) Phase matching condition illustrated in a Fabry Perot cavity [94]. (d) Surface corrugation to define periodic grating in a DFB laser [95]. (e) Phase matching condition illustrated in a DFB cavity [95].	110
4.2	(a) The contra-propagating Bragg modes in a 1st order DFB laser feed off each other. The grating induces the coupling. (b) Plot of the relative amplitudes of left travelling wave S and right travelling wave R vs distance.	111
4.3	Side view of a four layer DFB laser structure. The grating is described by its period Λ , and the tooth shape functions $w_1(x)$, and $w_2(x)$	112

4.4	A generic DFB longitudinal section with additional reflector on one end.	120
5.1	Workflow related with the numerical analysis of a generic DFB Laser.	126
5.2	The cross-section of the TAVI DFBL. For confidential reasons, the layers are not labeled.	127
5.3	The grating profile of the TAVI DFBL. The geometric parameters of the trapezoid-shaped grating are illustrated with respect to its y-z cut plane. Image courtesy Sina Aghili	127
5.4	Fundamental mode of the laser under consideration	128
5.5	$m = \pm 1, \pm 2, \pm 3$ partial wave for duty cycle $\gamma = 0.655$. Image courtesy Sina Aghili	130
5.6	Pair of solution for the modal gain and the Bragg detuning cavity length = 1 mm, duty cycle = 0.75 and mirror reflectivity = -0.565. .	131
5.7	Transmission spectrum of our TAVI distributed feedback laser. . .	132
5.8	Reflection spectrum of our TAVI DFB laser.	132
6.1	The workflow involved in a research project which seeks to demonstrate validation of an idea with respect to a photonic integrated circuit.	134
6.2	(a) Standard package from Technobis (Tipps) for photonic integrated circuits. (b) A customized package from the same company. Their PDK is available through Phoenix design kit	135
A.1	Shannon's model for the transmission of a message	137
A.2	Schematic showing Shannon's idea for a source code. Alice has an encoding channel such that $\mathcal{E} : x^n \mapsto c^{\tilde{n}}$. $\tilde{n} < n$, as indicated by fewer lines following the encoder. The transmission channel is noiseless indicated by "id" for identity. Bob's decoding channel undoes Alice's encoding $\mathcal{D} : c^{\tilde{n}} \mapsto x^n$. A good code will allow Bob to decode Alice's message with arbitrarily low probability of error [101]	143
A.3	Schematic showing Shannon's idea for a channel code. Alice chooses a message $m \in [M]$. The encoder channel outputs according to $\mathcal{E} : m \mapsto x^n$. The noisy channel induces $\mathcal{N} : x^n \mapsto y^n$. Bob's decoder finds $\mathcal{D} : y^n \mapsto \hat{m}$. A reliable channel code will allow Alice and Bob to communicate at a finite rate with vanishing probability of error $\Pr[m \neq \hat{m}] \rightarrow 0$. [101]	146
A.4	Schematic of information transmission through an AWGN channel. The input signal is power limited since the channel cannot withstand arbitrary large power. The challenge is to enhance correlation between the input and the output in the presence of Gaussian noise while also taking into account the power constraint.	164

D.1	(a) General layout of a MMI coupler implemented by means of a 2D slab waveguide (access waveguides are not shown) [9]. (b) Supported modes in the multimode region [106]. (c) Multimode waveguide showing the input field $\psi(x, 0)$, a mirrored single image at $3L_\pi$, a direct single image at $23L_\pi$, and two-fold images at $3L_\pi/2$ and $3 \cdot 3L_\pi/2$ [106].	181
D.2	Diagrammatic view of a Mach Zehnder Modulator with only one input.	184
D.3	The green arrow and the blue arrow indicate electric field parallel and perpendicular to the optic axis respectively. The red arrow is the resultant field. Linearly polarized light entering a half-wave plate can be resolved into two waves, parallel and perpendicular to the optic axis of the waveplate. In the plate, the parallel wave propagates slightly slower than the perpendicular one. At the far side of the plate, the parallel wave is exactly half of a wavelength delayed relative to the perpendicular wave, and the resulting combination is a mirror-image of the entry polarization state (relative to the optic axis)[107].	188
D.4	Schematic illustrating implementation of 8 phase shift keying with a single MZM in dual drive operation.	189
D.5	Photocurrents corresponding to the in-phase and quadrature-phase quadratures of the modulated signal in 8PSK.	190
D.6	Eye diagram results for P and Q quadrature in 8PSK.	190
D.7	The signal statistic with respect to the decision boundaries in 8PSK.	190
D.8	Construction of an eye diagram. [108]	191
D.9	Anatomy of an eye diagram [108]	191

List of Tables

3.1	BPSK implementation parameters for a single dual drive MZM	67
3.2	QPSK implementation parameters for a single dual drive MZM	68
3.3	Overview of the validation results reported by the foundry.	106
5.1	Streifer coefficients obtained from the partial wave analysis.	129
B.1	Hilbert Space versus phase space representation of Gaussian states	174

I dedicate this work to my parents.

Organization of the thesis

The thesis is divided into six chapters reflecting my involvement in two different projects as a masters research student. The underlying thread connecting both the projects is the development of photonic integrated technology based on Indium Phosphide (InP). Therefore, our first chapter, Chapter 1, gives an overview of the area of photonic integration implemented with Indium Phosphide material system so that the reader can put both the projects in a larger context. Chapter 2, and Chapter 3 contain discussion on the continuous variable quantum key distribution project. In particular, Chapter 2 details the theoretical analysis of continuous variable quantum key distribution, which build on the background of quantum Shannon theory and phase space quantum optics. The design analysis for the photonic integrated transceiver chips with foundry specific components are included in Chapter 3. Similar approach is taken for the distributed feedback laser project. Chapter 4 discusses, at length, the electromagnetic theory of an emblematic distributed feedback (DFB) laser. The formulas mentioned in this chapter are then employed in the numerical analysis of the transverse and longitudinal mode profile of a DFB laser with a certain cross-section and a grating profile in Chapter 5. Lastly, Chapter 6 wraps up the thesis by providing some future perspective for both the projects. To complement the discussions presented in those chapters mentioned, five appendices have been attached. Appendix A tries to explain the information theoretic reasoning behind the formula for the secret key rate of a quantum key distribution protocol. Appendix B deals with some of the notions from phase space quantum optics. Appendix C details the derivation of the covariance matrix of a two-state protocol. Appendix D complements Chapter 3. Finally, Appendix E shows the in-between steps of the mathematical results presented in Chapter 4.

Chapter 1

Photonic Integration Technology

1.1 Introduction

In this thesis, we pursue two projects associated with Indium Phosphide (InP) based photonic integration. The first project is related to photonic integration at the system level: realization of the continuous variable quantum key distribution on-a-chip. The second project is related to photonic integration at the component level: performance enhancement of an on-chip distributed feedback laser by optimizing its physical parameters. Therefore, we begin our discussion with a chapter which provides an overview of the current landscape of integrated photonics with a focus on InP-based material system. First, we motivate the case for integrated photonics to replace systems based on bulk, discrete optical components citing three application areas. The most relevant application area in the context of this thesis is that of coherent optical communication. As we shall see in Chapter 3, when Alice and Bob are trying to generate a secret key with a continuous variable quantum key distribution protocol, they are effectively performing coherent optical communication. Then, we describe the specific motivations which prompted us to undertake our own research projects. We lay out precisely how our effort tries to bring novelty in the relevant field of research. Both of our projects involve generic photonic integration technology. On that account, we briefly talk about the adoption of generic model for integrated photonics, which aims at emulating the success of integrated electronics. Finally, we discuss, in detail, the work flow associated with the design of a photonic integrated circuit. This discussion will be helpful in following the design analysis presented in Chapter 3.

1.2 Motivation for Photonic Integration

The idea behind photonic integration is to integrate a number of photonic components (building blocks) with well-defined functionalities on a common planar substrate so that it forms a photonic integrated circuit/chip (PIC) to perform a certain application or a set of applications [1] (Figure 1.1). The components are connected via optical wires, or *waveguides*, which ensure controlled propagation of light throughout the circuit.

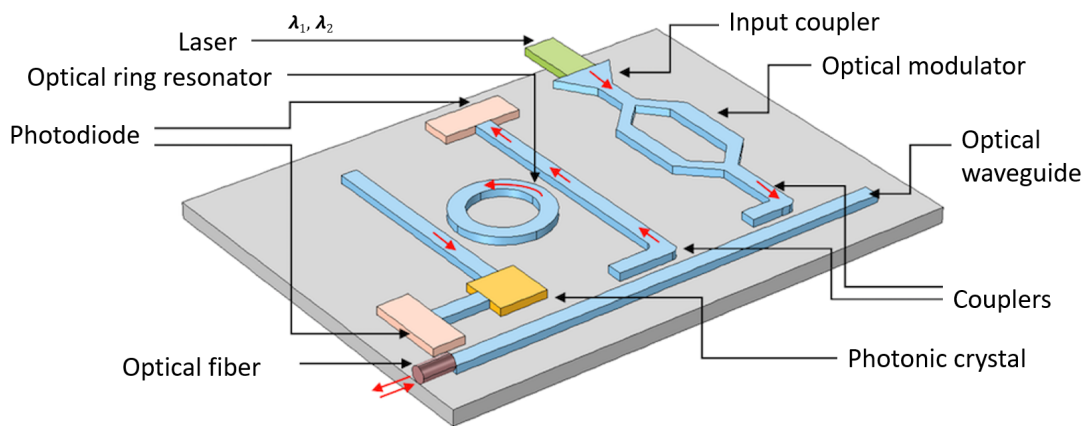


FIGURE 1.1: An artist's rendition of a photonic integrated circuit showcasing different optical components reproduced from [2].

Photonic integration has largely been motivated by microelectronic integration. The miniaturization of electronic circuits from the bulk-scale to the chip-scale has increased scalability, performance, and reliability of those circuits. At the same time, it has reduced their size, power, and costs. The motivation behind photonic integration is to replicate the same impact in the field of photonics based technology. In the following, we talk about three application areas where photonic integration can bring dividends.

The traditional way of performing an optical experiment is to set up bulk optical components on an optical table (8-12 feet long). The components require careful and stable alignment (protection against vibration, dust, moisture, and temperature drift). With photonic integration, it is possible to downscale the same experiment to a compact size of $1 \times 1 \text{ cm}^2$ on a semiconductor chip with all the functionalities retained. Moreover, one does not need to worry about the misalignment of components, and opto-mechanic instability since optical waveguides ensure confined and well-directed propagation of light. For example, with a Mach-Zehnder Interferometer fabricated on a photonic chip, interferometric experiments can be done more reliably, as one can adjust the phase mismatch between the interacting beams in a more continuous fashion with proper voltage bias instead of mechanical movement of mirror(s).

The quantum unit of light in a particular mode is a photon. With respect to some observable space (such as polarization), a single photon can be considered as the physical embodiment of a qubit, the irreducible unit for quantum information processing. A quantum gate (unitary) operation on a qubit can be physically implemented with a linear optical element acting on a photon. Moreover, an arbitrary quantum gate is basically a $N \times N$ unitary operation with N qubits as inputs and N qubits as outputs. The famous KLM protocol [4] demonstrated that any $SU(N)$ operation can be split down to $SU(2)$ operations, which means any quantum gate operation can be constructed out of mirrors, phase shifters, beam splitters, or their combinations like Mach-Zehnder Interferometers, in different numbers and topologies (Figure 1.3) together with sources and

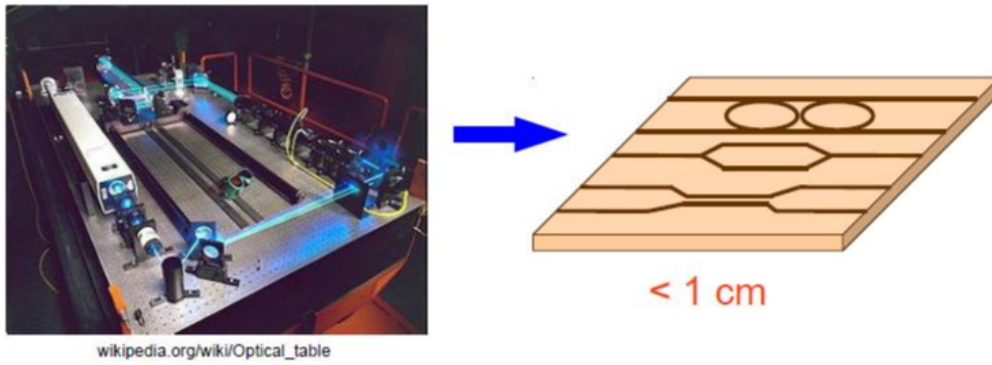


FIGURE 1.2: Scale comparison between an optical table set up and an optical chip[3]

detectors. As the complexity of the quantum operation increases, the number

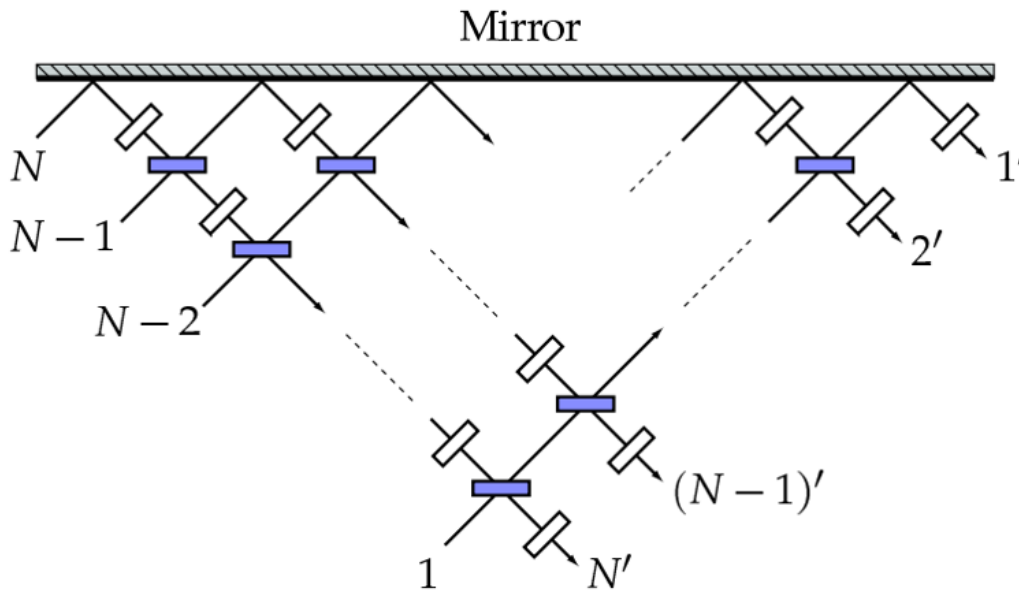


FIGURE 1.3: Decomposing an N port unitary $U(N)$ into $SU(2)$ group elements: beam splitters and phase shifters [5]

for integrating optical elements would scale as well. For this scenario, photonic integrated circuits are perfectly suited.

Arguably the greatest potential for photonic integration lies in modern day optical communication network. In fact, the area of long-haul and metropolitan optical communication network has been identified as the so called "killer application" for PICs.

Light is an electromagnetic wave capable of carrying information encoded in its amplitude, phase, and polarization. With wavelength division multiplexing (WDM) and advanced modulation formats, modern day optical transmitters and receivers can communicate Terabit/second (Tb/s) data via fiber optic

link. To process this high volume of data, these transmitters and receivers may contain hundreds of components and interconnections. Figure (1.4) shows the block diagram of a point-to-point coherent communication system.¹ On the transmitter side, a laser emits narrowband light centered around some wavelength (typically 1550 nm) with a specific polarization. A beam splitter (typically 50:50) splits the light into two branches. On each branch, the light can be modulated both in amplitude and phase by actual data stream (in the form of electronic signals driving the modulators). A polarization converter on one branch induces a 90° rotation on the polarization state of light. Now, the two orthogonally polarized signals are combined with a polarization beam combiner so that they can be sent over the same fiber channel. On the receiver side, there is another laser which works as a local oscillator for phase referencing. The transmitted signal from the transmitter is again split into two branches depending on their polarization. Light from the local oscillator is split with a 50:50 beam splitter. The x-polarized light (y-polarized light) is mixed with one branch of LO output (another branch of LO output) in a 90° hybrid. The outputs of the 90° hybrid impinge on an array of photodetectors. The information encoded on the amplitude and the phase of the signal is recovered by the so called balanced detection technique followed by electronic post processing.

In a WDM system, there may be hundreds of transceiver pairs, like the one described above, per link operating at different center wavelengths. A multiplexer multiplexes signals from different transmitters and sends them over a single fiber. A demultiplexer on the other end of the fiber demultiplexes the transmitted signals and route them to the receiver circuits for information decoding.

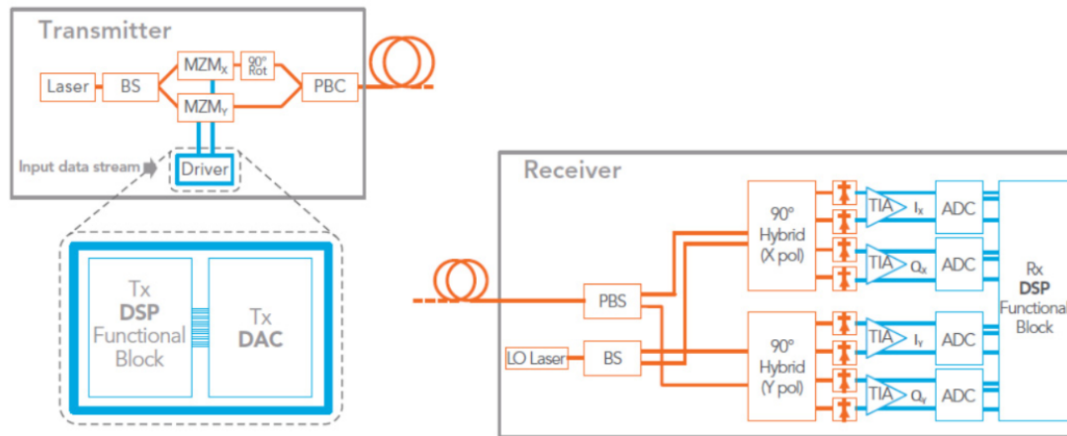


FIGURE 1.4: Block diagram representation of a coherent communication system. The diagram also captures the essence of a CV-QKD system realization. Symbols: BS=Beam Splitter, MZM=Mach Zehnder Modulator, PBC=Polarization Beam Combiner, DSP=Digital Signal Processing, DAC=Digital to Analog Converter, LO = Local Oscillator, PBS=Polarization Beam Splitter, TIA= Trans Impedance Amplifier, ADC= Analog to Digital Converter)

¹In Chapter 3 we shall see the set up is essentially the same with the circuit diagram of our photonic integrated circuit intended to implement CV-QKD.

Modern day WDM systems are mostly equipped with assembly of discretely packaged photonic devices. There is little room for scaling. Besides, there is a significant amount of cost associated with assembling and individual testing of components. With photonic integration it is possible to lower the packaging costs by reducing the number of fiber ports as the components are connected through a common waveguide structure. Furthermore, it enables simultaneous on-wafer testing as the testing sources and receivers can be integrated on chip. The reliability also increases by reducing the optical coupling loss.

At present three material platforms have reached a reasonable degree of maturity in terms of technology (physical structure of the components and their fabrication processes). Figure (1.5) illustrates these three platforms namely Indium Phosphide (InP), Silicon-on-insulator (SOI), and Silicon Nitride (Si_3N_4). Each of them has its own advantages and limitations. A comparative analysis can be found on [6], [7], [8]. For the purpose of this dissertation, the relevant material system is Indium Phosphide (InP). InP-based PIC technology has the unique capability of accommodating both active and passive components on the same chip without the need for any hybrid/heterogeneous integration. At present, there are two foundries in the world which provide dedicated InP-based PIC services to the outside customers: Smart Photonics in the Netherlands, and Fraunhofer-Heinrich Hertz Institute in Germany.

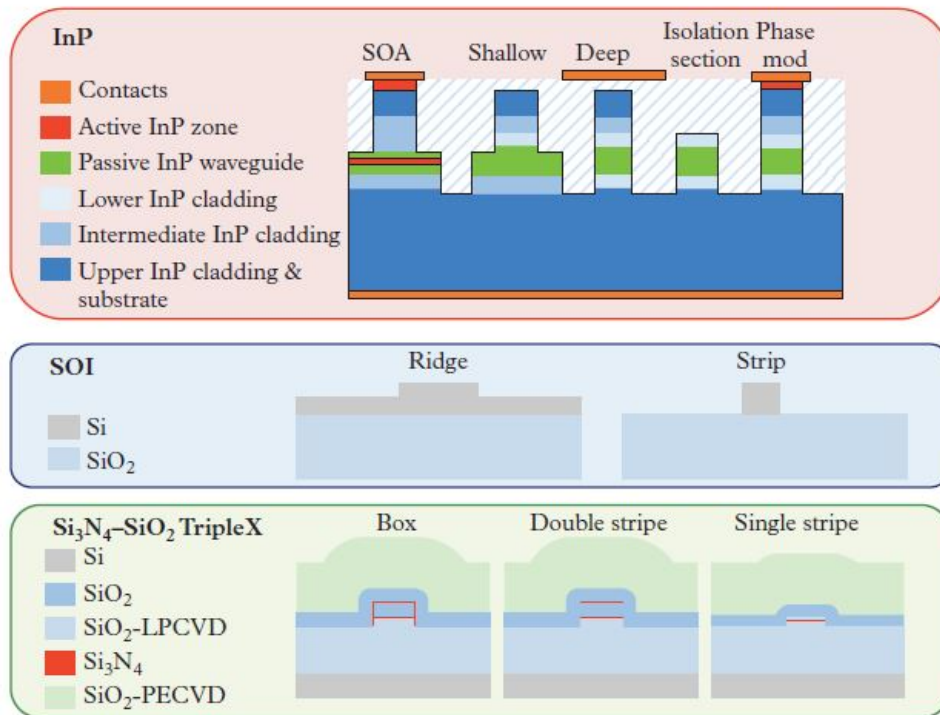


FIGURE 1.5: Waveguide structures for the InP, SOI (Silicon-on-insulator), and Si_3N_4 – SiO_2 material platforms [9]

1.3 Conceptual Placement of this Thesis

In recent times, there has been a massive surge in interest in developing quantum technologies with integrated photonics [10], [11]. Apart from the reduction in size and cost for practical deployment and commercialization, as mentioned earlier, these integrated quantum photonic (IQP) circuits are also expected to bring improvements in terms of robustness and reproducibility. Current research effort focuses not only on developing advanced photonic integrated components optimized for quantum applications, but also on shoehorning entire quantum optics systems into system-on-chip realizations leading to advanced solutions for large scale system and network level integration. The two research projects we address here attempt to bolster this ongoing research effort.

The realization of a fully integrated version of quantum key distribution system is an emerging field of research and development in quantum communication. Quantum key distribution is a cryptographic precondition which is immune to undetected eavesdropping. Since its first appearance in the scientific literature in 1984, quantum key distribution is now an established technology, and even commercially available from companies like ID Quantique, Toshiba, NEC. However, these commercial QKD systems are primarily based on single photons implemented with bulk optical/optoelectronic components. In fact, integrating single photon sources and detectors at the chip scale is inherently difficult.

With the advent of continuous variable quantum key distribution (CV QKD), one can implement a QKD protocol with the modulation of the amplitude and the phase of light coming from a laser followed by demodulation with balanced homodyne/heterodyne detection schemes. From the implementation point of view, this is nothing but coherent optical communication. Therefore, to implement on-chip CV QKD, one would require an integrated transmitter, and an integrated receiver capable of performing coherent optical communication. An added advantage with CV QKD is that it is perfectly compatible with the existing optical fiber based telecom infrastructure operating at 1550 nanometer of optical wavelength. In fact, a successful field test for CV QKD within the metropolitan range in Madrid, Spain has already been carried out which involved important players (Huawei, Telefonica) in the fiber optic industry [12]. In Europe, flagship projects such as CiViQ [13], and UNIQORN [14] are underway, attempting to realize stable, miniaturized, and cost-effective CV QKD systems so that they can be installed in the European fiber optic infrastructure for future commercial application.

To our knowledge, the first realization of a fully integrated CV QKD protocol, at the laboratory research level, has been reported with silicon photonics (SiPh) [15], [16]. The transmitters and receivers, mentioned in the relevant publications, have all the components integrated on chip except for the sources. In both the reported cases, an off-chip laser acts as a source which is integrated with other optical components on a Silicon chip. The estimated key rates are 0.14 Kbps and 2.3 Kbps, respectively over a distance of 100 km in optical fiber.

In this thesis, we undertake an attempt to realize on-chip CV QKD with InP material platform. The novelty of our approach resides in the fact that we design a monolithically integrated transmitter and receiver based on InP technology, without resorting to hybrid integration based on Si-III-V platforms. Moreover, we claim that our design is versatile, in the sense that the transceiver architecture should support multiple modulation formats.

The performance of a system-on-chip largely depends on the performances of the individual components. While using single mode fiber for point to point communication, it is imperative to have an on-chip source with narrowband optical transmission spectrum. Hence, research endeavor targeting the development of integrated distributed feedback laser is necessary. DFB lasers quintessentially have sharp peak spectrum with large side-mode suppression ratio, making them ideal for single mode fiber optic communication. DFB lasers with first order grating have the lowest radiation loss and the lowest threshold current. But, often times, lithographic limitations dictate that we fabricate laser with higher order gratings at the expense of some performance trade-offs. In this thesis, we present thorough numerical analysis showing how to relate performance parameters such as coupling coefficient with the physical dimensions of a third-order DFB laser emitting at 1310 nanometer of optical wavelength. The laser will be fabricated for taper assisted vertical integration (TAVI) platform, a regrowth-free photonic integration technology in Indium Phosphide [17].

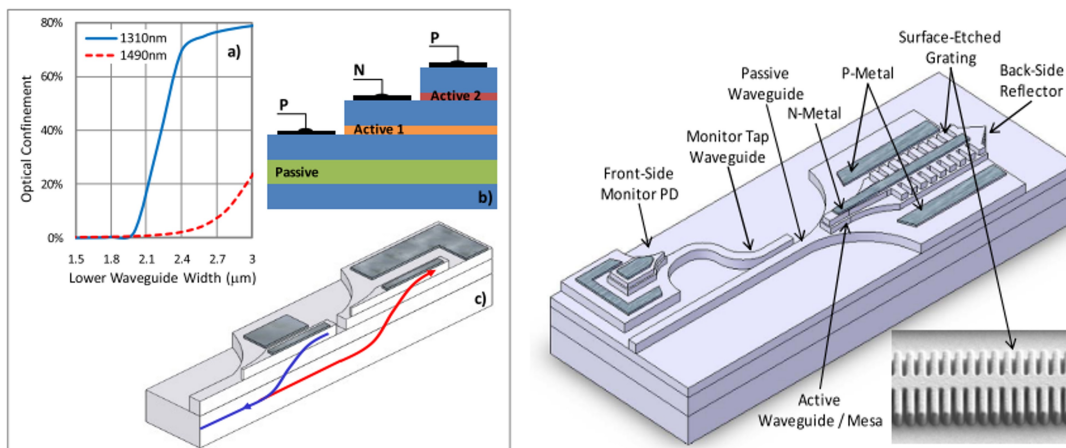


FIGURE 1.6: Illustration of taper assisted multi-guide vertical integration principle.

In TAVI, instead of planar integration, all the required materials are grown in one epitaxial growth step and vertically integrated in a form of dedicated functional waveguides, stacked one above another in a multi-guide structure [18], [19]. Vertically stacked functional waveguides are evanescently coupled, such that their vertical optical modes overlap at their tails, and providing a vertical transition between the functional waveguides thereby. This mechanism is based on lateral tapering of the vertically evanescent-field coupled functional waveguides, such that tapering down the ridge waveguide results into pushing its optical mode down into the functional waveguide below it, until the optical mode of the tapered functional waveguide is eventually cut-off. The coupled

light can be laterally routed at this functional waveguide level and eventually, can be vertically transferred to the functional waveguide of another waveguide device. In this way, functional integration is achieved without either epitaxial regrowth or butt coupling. Figure (1.6) tries to illustrate the concept of TAVI integration technology. The left figure shows schematics of the multi-guide vertical integration principle on an example of the 1310 nm laser – 1490 nm detector optical duplex [18].

However, for all its benefits, the TAVI platform requires an unconventional on-chip laser source design [20]. A TAVI distributed feedback laser differs from a conventional ridge waveguide DFB laser in a number of ways. For example, in a TAVI-compatible DFB laser, waveguide Bragg gratings cannot be defined at the interface between laser guide layers since it requires an epitaxial regrowth. Rather the DFB cavity is either a laterally coupled [21], or vertically-coupled [22] surface etched grating (SEG). In either case, the SEG is etched into the N-InP layer, but does not penetrate the laser SCH (separate confinement heterostructure) guide layers. In optical communication windows at 1.3 μm and 1.5 μm , the Bragg wavelength, defined as $\lambda_B = 2\Lambda n_L(\lambda_B)/m$, where Λ is the grating period, $n_L(\lambda)$ is the wavelength dependent effective index of the laser mode and m is grating order, is too short for the I-line stepper to define a 1st order SEG. Typically, a 3rd order SEG is used instead. While this reduces the coupling efficiency, 3rd order SEG also has an advantage in that it allows engineering of the imaginary part of coupling coefficient to achieve gain-like or loss-like coupling, without the SEG actually penetrating the gain region or having the absorption loss. The image on the right in Figure (1.6) shows a transmitter building block featuring laterally coupled surface etched grating distributed feedback laser (LC-SEG-DFBL) with front-side passive waveguide / power monitor PD and back-side retro-reflector [20]

The project we discuss in this thesis is about a TAVI-compatible, 3rd order, vertically-coupled SEG distributed feedback laser, which employs a specific cross-section profile. Our goal is to investigate the impact of the grating parameters on the output performances for this laser. The effects of the 3rd order grating on DFBL are related to: radiation loss, complex coupling, mechanism of the longitudinal mode selection, and spatial hole burning. Since TAVI DFBL shares the same basic physics and operating principle with any other DFBL, we take advantage of the already established theoretical analysis of higher order DFBL by Striefer [23], and the subsequently developed numerical analysis techniques based on this theory by Millett et al [24].

1.4 Generic Model for Photonic Integration

Since its inception, the technology development for photonic integration has mostly been application specific. This is because to justify commercial success of any integrated circuit, the economic value derived from integrating components must outweigh the cost of integration itself. Therefore, although numerous proof of principle demonstrations of chip-scale photonic systems or subsystems have been shown at the research level, only a few have justified investment

into their commercial implementation. That is why most often photonic integrated circuits have remained order of magnitude more expensive compared to their electronic counterparts. The reason why electronic ICs are cheap and ubiquitous is largely because of the adoption of generic model for microelectronic integration introduced by Conway and Mead [25]. The generic approach has greatly reduced the prototyping costs and the throughput time of ICs, and paved the way for very-large-scale integration (VLSI) [26]. There are two aspects to the generic model

- **Generic Integration Model (GIM):** Develop a few generic integration techniques by realizing a highly standardized set of building blocks with a broad range of functionalities.
- **Generic Foundry Model (GFM):** Develop a foundry infrastructure for providing low-cost open access to these generic technologies.

For photonic integration a similar approach has been suggested [27], and is now actively being pursued.

1.4.1 Generic Integration Model

A generic integration model refers to the monolithic integration of a small set of accurately characterized building blocks. These building blocks offer the basic

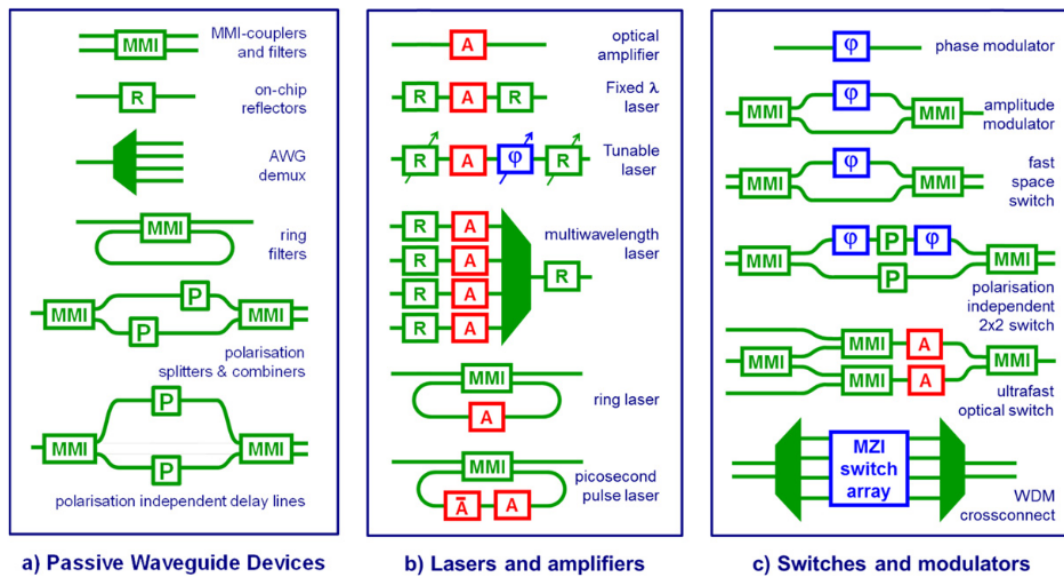


FIGURE 1.7: Examples of the functionalities that can be realized with (a) passive waveguide devices alone, or in combination with (b) optical amplifiers and (c) phase modulators [28]. The second component in (c) shows a Mach Zehnder Interferometer (MZI). This composite building block (CBB) consists of simpler building blocks such as MMI's, waveguide sections, and phase shifters. The MMI is another CBB made out of single mode and multi mode waveguide sections (BBBs).

functionalities required to realize PICs for a broad range of different applications. To that end, a set of basic building blocks (BBB) is identified first. **BBB:**

BBBs are the irreducible set of building blocks required to realize the full functionality provided by the generic technology. With such a set of building block one can construct complex building blocks (CBB), with well-defined functionality of their own. **CBB:** CBBs are building blocks composed out of two or more BBBs. They do not require additional process steps for the combination. The property of a CBB can always be derived from the properties of its constituting basic building blocks. A Mach Zehnder Modulator is an example of a CBB. It has one power splitter (MMI), two waveguide sections (BBB) with phase modulators on top, and one power combiner (MMI). The MMIs themselves are building blocks made out of passive waveguide sections (BBBs). The BBBs and the CBBs can be put together in different numbers and configurations to design, and eventually fabricate any *ASPIC* (Application specific photonic integrated circuit). A particular ASPIC can in turn be used as a composite building block (sub-circuit) for another ASPIC. Figure (1.7) illustrates the idea described so far.

For photonic integration applications there is a wide variety of functional building blocks. Their physical structures would usually differ from each other in order to reflect this functional variation. Each basic building block may have a different semiconductor layer stack with varying band-gap, doping, and/or heterostructures. But sometimes it is possible to derive multiple functions from a single structure.

BTB: BTBs are physical structures that can have different functions and that can act as different BBBs under different operational conditions. For example, an SOA (Semiconductor optical amplifier) can act as the gain section for a laser (under forward bias), or as a photodetector (under reverse bias). An advantage of generic

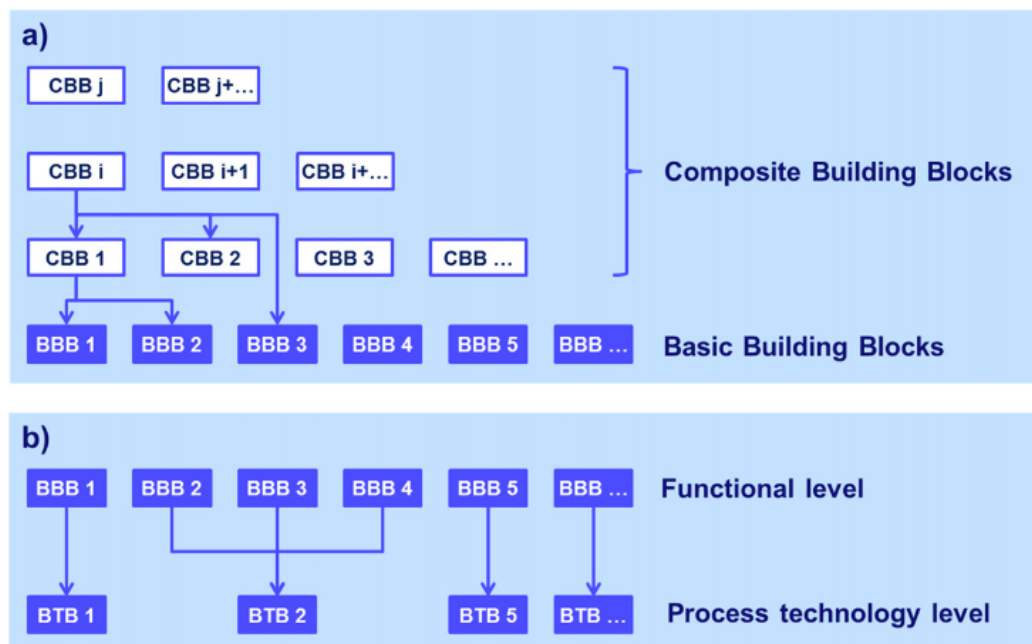


FIGURE 1.8: Schematic diagram of the hierarchical structure for the different building blocks [28].

integration model is that rather than having fragmented technologies for optimizing different applications, a single generic process (like CMOS in microelectronics), or a few can suit ranges of applications. Therefore it justifies investments in developing the technology for improved performance and reliability at the level of building blocks. Figure (1.8) illustrates the flow of hierarchical order from the process technology level to the functional level in a generic integration model.

1.4.2 Generic Foundry Model

Once the generic integration technology is standard and mature, the chip manufacturer provides open access to it to external users. The access is provided in terms of a design manual, and a process design kit (PDK) The design manual contains all the experimentally characterized data about the building blocks for a user to design his ASPIC. The Process Design Kit is a licensed plug-in library containing all the detailed information about the foundry-specific building blocks, which can be integrated with the PIC design software environment.

In this way the task of designing a chip, and manufacturing is decoupled, and can be distributed to two different entities: *pure play company*: companies which actually own fabrication facilities with standardized integration technologies, and *fabless company*: companies which do not have chip manufacturing capability, but can provide design services with foundry specific design considerations. The chip manufacturers can run their facilities at the fullest capacity by finding work from a world wide pool of fabless companies. The operational expenditure of cleanroom facilities can be distributed over multiple users.

1.4.3 Multi-project Wafer Run

Since in the generic model PICs are fabricated using the same standardized technology, a single wafer can accommodate designs from different customers.

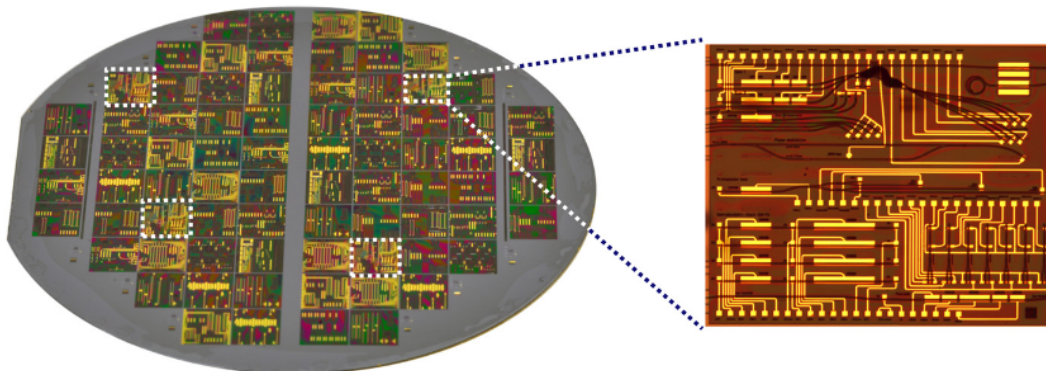


FIGURE 1.9: Example of a multi-project wafer (MPW) realised in the COBRA process. The wafer has 52 cells in the usable wafer area. Each ASPIC is copied at four different locations on the wafer with room for 12 users and one test cell for wafer validation. The figure shows one test cell and its position on the wafer. In this way, each user will receive four copies of his chips per wafer in the MPW batch. [28]

Such a wafer is called a Multi-Project Wafer (MPW). Since the cost of a single run can be shared by a number of customers, multi-project wafer runs provide a cost advantage for small and medium enterprises (SMEs) and universities. The participants in an MPW-run submit their designs in GDS files to foundries where the final wafer-layout is made using dedicated software tools. Figure 1.9 shows a multi-project wafer of InP from The COBRA Institute at Technical University of Eindhoven, Netherlands.

1.4.4 Generic InP-based PIC technology

In this thesis, we choose to work with InP technology as it is the material platform which provides the most comprehensive list of optical components with diverse functionalities, and offers the unique advantage of monolithic integration of the active and the passive devices. The thrust for developing InP-based generic technology initially came from Europe. Notably, the COBRA research institute of TU Eindhoven (now Institute of Photonic Integration) did numerous pioneering works of developing a generic process for integration of InP based optical elements eventually opening its process to external users in 2008. Two European R&D projects EuroPIC and PARADIGM aimed at not only developing the technology of generic InP based integration platform [28]. These projects resulted in transferring the foundry model from the university environment to an industrial environment. Open access MPW services were launched on a pre-commercial basis by three European entities- UK based Oclaro (now owned by Lumentum), Fraunhofer HHI of Germany, Smart Photonics in the Netherlands, the latter commercializing the technology of the COBRA institute. SMART photonics and HHI provide pre-specified building blocks in the form of process design kits (PDKs). For our project, we booked a spot for HHI's MPW run HHI17. The performance parameters of the building blocks available from HHI can be found on the website of JePPIX [6].

1.5 ASPIC Design Environment

To translate a functional idea into a working integrated photonic chip involves a number of stages: design, fabrication, packaging, and characterization (testing). In this section we focus on the first stage namely the design of a photonic integrated circuit (PIC). This is because the principal project presented in this thesis (Chapter 5) concerns with the design of a PIC for quantum key distribution. The workflow associated with the design of a photonic integrated circuit can be seen in the context of the following sequence of tasks:

1. Component level analysis;
2. Circuit level analysis;
3. Mask layout generation;
4. Verification

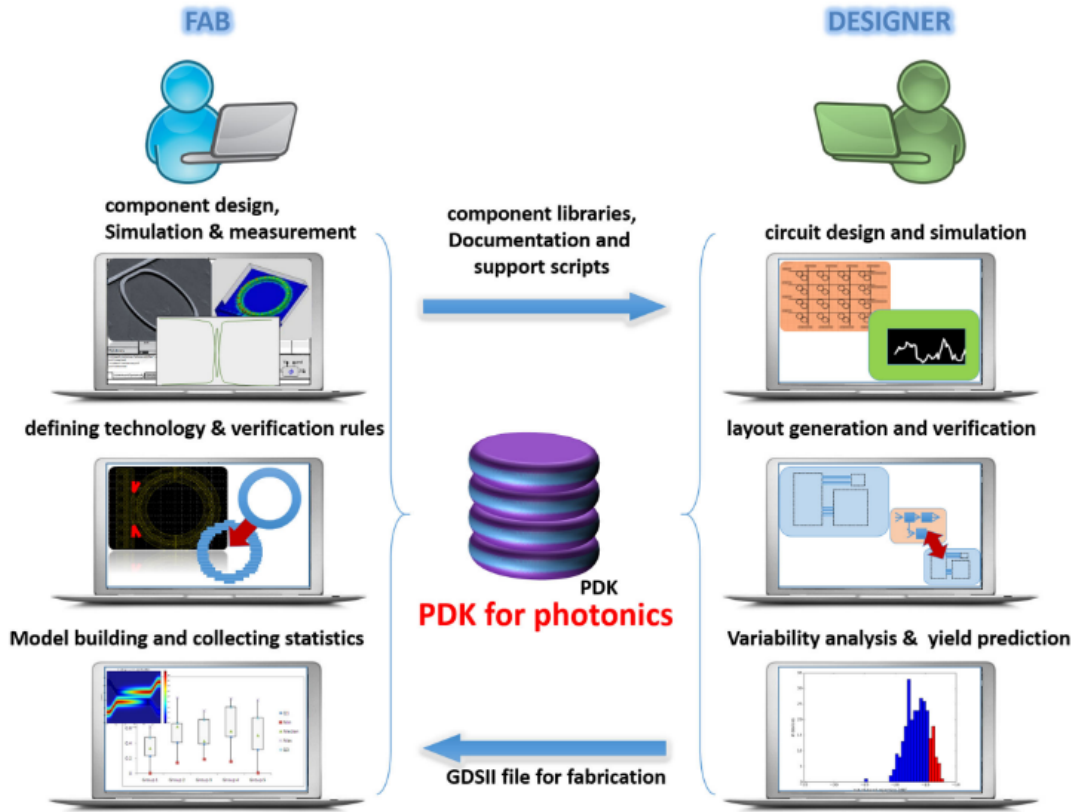


FIGURE 1.10: Flow involved in an ASPIC design environment [29]

1.5.1 Component Design & Optimization

The goal of the component level analysis is to create a compact model for a building block so that it can be subsequently used for circuit level simulation. Physical simulations of optical and optoelectronic components are carried out by the foundries to develop their compact model library. The paradigm of component level analysis can be explained with a specific example.

Let us consider a single-mode, straight waveguide section available from a particular material platform. The straight waveguide is optimized in terms of its core and cladding materials, and geometry. A compact model of this waveguide should provide information on how the effective index, group index (group delay), and propagation loss of the guided mode varies with wavelength. The compact model can be generated with an eigenmode solver (for example Lumerical FDE solver). A mode solver solves time-independent Maxwell's equation numerically for a given structure. The guided modes are the eigenmodes and the effective indices are the eigenvalues. By sweeping over a range of wavelength, one can obtain the effective index vs wavelength relation. This in turn will lead to the evaluation of the group index dispersion according to

$$n_g = n_{\text{eff}} - \lambda \frac{dn_{\text{eff}}}{d\lambda}. \quad (1.1)$$

To calculate the propagation loss, one needs to propagate the mode over the length of the waveguide. Therefore, it would require time dependent solution of Maxwell's equation (for example with Lumerical FDTD solver). The propagation/insertion loss can be given in terms of S-parameter values. A straight waveguide transmits all of the input light. There is no back reflection. But for a bent waveguide section, there might a non-negligible reflection. In this case, all the entries in the scattering matrix need to be specified to define the relevant compact model.

1.5.2 Circuit Level Simulation

The compact models derived from the physical simulation stage for the building blocks can be used for running simulations at the circuit and system level. An industry standard circuit solver software package contains a library of optical components described by their behavioral models. It allows the designer to construct a photonic integrated circuit in schematic form by drawing components from the library and linking them together in a custom configuration. The basic assumption is that the light is flowing in clearly defined waveguide modes throughout the circuit. The simulation analysis can be carried out both in the time domain and the frequency domain.

Simulating in the frequency domain

Frequency domain circuit simulation is useful in studying a photonic integrated circuit in steady state. Each component in the circuit is characterized by a complex scattering matrix or S-matrix. The matrix entries describe the relationships between its bi-directional ports as well as relationships between the modes supported. Running a frequency domain simulation would calculate the overall scattering matrix of the circuit from the S-matrices of the individual elements at each frequency. Some common results extracted from the calculated overall matrix include transmission and reflection spectra, group delay, and resonator properties such as free spectral range (FSR) and quality factor (Q).

Simulating in the time domain

Time domain simulation is used to calculate the response of a photonic circuit when it is triggered by a time dependent source. The response is solved by time-stepping the signals between all the nodes in the circuit, and updating all the compact models at each step. Time domain models for the passive linear components are derived from their corresponding frequency response in the form of a digital filter either with finite impulse response (FIR), or infinite impulse response (IIR). Since the signal is modeled as a modulation on top of a carrier wavelength, these models work in a limited bandwidth. Generating compact models for the active building blocks is not straightforward since they require multi-physics simulation. Therefore, often times, the compact models for modulators, semiconductor optical amplifiers, or photodetectors are essentially curve fits based on a mix of physical understanding derived from the physical simulations, and optical performance measurements. In this way, the resource

cost of time domain simulation of these devices is minimized without compromising in the prediction of the circuit performance too much.

1.5.3 Circuit Layout Generation

The schematic diagram of an intended photonic integrated circuit needs to be translated into a circuit layout. This is the standard way of communicating one's design to a foundry for fabrication. The layout contains different mask layers which correspond to different fabrication process steps, or to logical operations. The layout is usually generated in GDSII (graphic data system) file format. The layout of a circuit design consists of hierarchical cells, where some parts of the geometry can be parameterized (hence called parametric cells or PCells). The PCells are connected via optical waveguides or electrical wiring. Waveguides should respect a minimum bend radius and spacing in order to prevent excess waveguide losses and parasitic coupling. PCells are usually defined in a scripting language which is integrated into a design software. Instead of drawing the shape, a designer sets some parameters by writing codes. The software then translates this design intent into a set of geometric primitives on different mask layers. Some examples of scripting platforms integrated into a design environment include SKILL in Cadence, Ample in Mentor Graphics Pyxis, or just Python in IPKISS, and KLayout.

1.5.4 Verification

Before submitting it to the foundry, the circuit layout should be checked against potential errors and violations of the design rules prescribed by the fab. The design rules from the fab include specification on minimum line/space widths, sharp angles, or overlap of layers that might cause unpredictable results during processing. The physical verification process, known as *design rule check* (DRC) is usually performed with software tools designed for electronics such as Calibre by Mentor, IC Validator by Synopsis, or Physical Verification System by Cadence. The foundries provide a design rule checking deck which can be integrated with these tools. However, the automated DRC for electronics cannot handle curvilinear geometries often utilized in photonics design, and hence may produce false positives. Research is ongoing at present to address this issue. Another important difference between photonics DRC and electronics DRC at present is that layout versus schematic (LVS) verification is nonexistent in photonics. LVS is a functional verification stage which checks if the layout mirrors the original design intent of the abstract circuit in electronics. In photonics design flow, historically, the circuit layout process developed separately from the circuit simulation process. Hence it is the designers' responsibility to make sure the correct components along with their parameters are used, and they are properly connected for the desired circuit performance.

1.5.5 PDK

In a generic foundry approach, the process design kit (PDK) bridges the gap between the foundry and the designers. A PDK is a piece of software code(s) that can be plugged into a design environment, and depending on its maturity may contain all the relevant information to facilitate design activities. For instance, the PDK may contain component geometry, library of compact models, circuit layout information (different mask layers), design rule deck, and so on. Therefore, to realize a photonic integrated circuit with a foundry specific technology, a designer must integrate the relevant PDK to all the different proprietary software used at different stages of design flow.

1.6 Conclusion

In this chapter, we have emphasized on detailing the sequence of tasks associated with the design of a photonic integrated circuit. The discussion may help the reader in better appreciating the design aspects of our photonic integrated transceiver, intended to realize a continuous variable quantum key distribution protocol, mentioned in Chapter 3. We shall conclude this chapter with a remark. The photonic integrated circuit design process is still at the stage of development. Although it shares many similarities with the microelectronic design flow such as hierarchical layouts and PDKs, there is still no seamless transition from the functional circuit level simulation to the physical layout generation. For this very reason, it is difficult to scale up the complexity of photonic circuits and verify their functionality.

Chapter 2

Continuous Variable Quantum Key Distribution

2.1 Introduction

Quantum key distribution enables two parties to communicate in an information-theoretic secure [30] way even in the presence of a malevolent eavesdropper [30], (Appendix A). This provable security arises from the fundamental laws of quantum mechanics. The first quantum key distribution protocol, the BB84 protocol, was put forward by Bennett and Brassard in a seminal publication in 1984 [31]. Since its inception, the BB84 protocol has matured over the years as a technology and is now commercially available [32]. One way to implement this protocol is to encode the key into the different polarization states of light/photon. Since the polarization states of a quantum state of light form a discrete eigenbasis, the BB84 protocol, along with host of other protocols, is categorized under discrete variable protocols. There exists, however a different type of protocol called the continuous variable protocols. The basic idea behind continuous variable quantum key distribution (CV QKD) is that the sender encodes the "key" in those observables of light which have continuous eigenspectra [33]. The receiver performs some relevant measurements to decode the key. There are numerous different ways to implement such a protocol. For example, in terms of states of light used - coherent, squeezed, or thermal; in terms of modulation schemes - discrete or (quasi)continuous; in terms of detection schemes - homodyne, or heterodyne; in terms of reconciliation strategies - forward, or reverse; and so on. For each different case the security analysis shapes up differently. In this chapter, the CV QKD protocol using coherent states will be discussed at length, since our on-chip light source emits coherent states. The first reported coherent state CV QKD protocol was the GG02 protocol [34]. All other protocols explored subsequently [35], [36], [37], [38], [39], [40], [41] can be seen as variants around this protocol. The proof of principle demonstration of CV QKD that we aim to achieve on-chip also closely resembles GG02. It essentially consists of two stages like every QKD protocol: (a) quantum state preparation, transmission, and detection; (b) classical post-processing. In this chapter, we mainly concern ourselves with the theoretical understanding of the quantum stage of the protocol. In that regard, first, we discuss quantum key distribution (QKD) as a cryptographic primitive in section (2.2). Then, in section (2.3), we

familiarize the reader with the concepts and notions from Gaussian quantum optics. These are the essential tools for appreciating the theoretical formulation of CV QKD with coherent states of light presented in the subsequent sections.

2.2 Quantum Key Distribution

Quantum key distribution is used as a part of a larger protocol such as quantum cryptography. The art of cryptography has long been in practice for safeguarding sensitive information while transmission. The working procedure of a generic cryptographic protocol can be understood in the following setting: Alice wants to communicate a message to Bob. In order to prevent interception by a potential adversary named Eve, Alice encrypts the message with a key. The encrypted message can only be recovered by Bob who has the corresponding decrypting key. In symmetric cryptography, to which we limit ourselves, the encryption and the decryption keys are identical. For anyone who is not in possession of the key such as Eve, the encrypted message (ciphertext) appears gibberish. Symbolically, one can summarise this concept as

$$\text{Enc}(m, k) = c, \quad \text{Dec}(c, k) = m, \quad (2.1)$$

where m, c, k stands for the message, the ciphertext and the key, respectively.

The "One-Time Pad" cryptographic scheme has been proved to be information-theoretically secure. In this scheme, modular addition (bit-wise XOR) constitutes both the encryption and the decryption function. This operation can successfully fulfill the desiderata for a cryptographic protocol: (a) The protocol is correct; Bob can recover Alice's intended message with the application of the key on the ciphertext

$$\text{Enc}(m, k) = m \oplus k = c, \quad \text{Dec}(c, k) = c \oplus k = m. \quad (2.2)$$

(b) The protocol is perfectly secret; any possible inspection of the ciphertext by Eve does not reveal her any information about the original message

$$\Pr(M = m | C = c) = \Pr(M = m), \quad (2.3)$$

where M and C denote any random message bit and ciphertext bit, respectively. Let us choose the message space to be $\mathcal{M} \in \{0, 1\}^n$, set of all possible n bit strings. The key is chosen uniformly from the key space $\mathcal{K} \in \{0, 1\}^n$. For arbitrary n -bit sequences $M = m$ and $K = k$, the ciphertext is $C = c = m \oplus k$. From equation (2.3)

$$\begin{aligned} \Pr(M = m | C = c) &\stackrel{(a)}{=} \Pr(C = c | M = m) \cdot \frac{\Pr[M = m]}{\Pr[C = c]} \\ &\stackrel{(b)}{=} \Pr[K = m \oplus c] \frac{\Pr[M = m]}{\Pr[C = c]} \stackrel{(c)}{=} 2^{-n} \frac{\Pr[M = m]}{\Pr[C = c]}. \end{aligned} \quad (2.4)$$

(a) follows from the Bayes' rule, (b) follows from modular addition, (c) follows from the uniformity of k . Now

$$\begin{aligned} \Pr[C = c] &= \sum_{m'} \Pr[C = c | M = m'] \Pr[M = m'] \\ &= \sum_{m'} \Pr[K = m' \oplus c] \Pr[M = m'] = 2^{-n} \sum_{m'} \Pr[M = m'] = 2^{-n}. \end{aligned} \quad (2.5)$$

Combining equations (2.4) and (A.5), one arrives at the equation (2.3).

The proof of unconditional-security for the one time pad requires that the length of the key should be at least as long as the message. Therefore, every time Alice and Bob want to exchange a new message, they need to generate a new key. Moreover, a basic assumption behind the proof is that "Alice and Bob already share a key". Therefore, the fundamental problem to address is how to distribute the key securely assuming Alice and Bob cannot meet in person. With classical means Alice and Bob cannot distribute the key in a provably secure way, if the channel between them is subject to eavesdropping. Whereas with quantum strategies, the laws of quantum physics will come into play to render their task of key distribution inherently secure.

Alice's random bit	0	1	1	0	1	0	0	1
Alice's random sending basis	+	+	×	+	×	×	×	+
Photon polarization Alice sends	↑	→	↘	↑	↘	↗	↗	→
Eve's random measuring basis	+	×	+	+	×	+	×	+
Polarization Eve measures and sends	↑	↗	→	↑	↘	→	↗	→
Bob's random measuring basis	+	×	×	×	+	×	+	+
Photon polarization Bob measures	↑	↗	↗	↘	→	↗	↑	→
PUBLIC DISCUSSION OF BASIS								
Shared secret key	0		0			0		1
Errors in key	✓		×			✓		✓

FIGURE 2.1: The figure illustrates the idea behind the BB84 protocol. Alice's key consists of random sequences of classical bits $\{0, 1\}$. She has two non orthogonal basis sets $\{|0\rangle, |1\rangle\}, \{|+\rangle, |-\rangle\}$ at her disposal which she can employ to encode her key bit. She also picks her basis states randomly. For the choice of $\{|0\rangle, |1\rangle\}$, the mapping corresponds to $0 \mapsto |0\rangle$ and $1 \mapsto |1\rangle$ while for the choice of $\{|+\rangle, |-\rangle\}$, it is $0 \mapsto |+\rangle$ and $1 \mapsto |-\rangle$. Bob knows about the basis sets Alice uses for encoding. But he does not have any knowledge about which random basis Alice has picked. He performs a measurement by randomly choosing a basis on the transmitted state to decode the key. If by chance, Alice's and Bob's choice of basis is the same, Bob will be able to read out the correct key bit (columns 1,3,6,8). But if Eve performs an intermediate measurement, it will randomize Bob's measurement result even though he is matched with Alice for the choice of basis

With quantum key distribution, Alice encodes the key in a quantum state, and Bob needs to recover the key with an appropriate measurement on that state. If Alice chooses a set of non-orthogonal states for her encoding scheme, which is sufficient to realize quantum key distribution [42], any attempt by Eve to

extract information about the key from the transmitted quantum state might leave a footprint on that state. Therefore, Alice and Bob can detect the presence of Eve with a very high degree of probability. Figure (2.1) illustrates the basic idea behind quantum key distribution protocol with an exemplary BB84 protocol [31]. Alice encodes her key using the set of four states consisting of $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$. In reality she can prepare single photons with polarizations pointing at vertical, horizontal, diagonal(45°), or antidiagonal directions (-45°). Bob measures the transmitted state without knowing which basis was chosen by Alice. When their choices of basis coincide by chance, they should extract the same key bit (Column 1,3,6,8 in the Figure (2.1)). On the contrary, if Eve intercepts in the middle, her intermediate measurement might randomize Bob's measurement results even if he is matched with Alice's basis (Column 3 in the Figure (2.1)). Moreover, Eve cannot copy/clone any intercepted state for subsequent measurement. Any such operation is fundamentally prohibited by the *no cloning theorem*, which is a consequence of linearity in quantum mechanics. Therefore, Alice and Bob can be aware of an eavesdropper by performing basis reconciliation. The basis reconciliation task requires a public authenticated channel, which guarantees that Eve cannot pose as anyone of them. If Alice and Bob agree that Eve's interception was below a certain limit, they can use the partially secure key (row 9 in the Figure (2.1) to distill a perfectly secure key by *privacy amplification*.

The above discussion suggests that as an information-processing task quantum key distribution requires local quantum operation followed by classical communication (LOCC). Therefore, any quantum key distribution protocol has two phases: quantum phase (preparation, transmission, and measurement of quantum states) and classical phase (information reconciliation, error correction, privacy amplification). However, the above example of QKD differs fundamentally from the continuous variable quantum key distribution (CV QKD) protocol with which we are concerned here. The photon polarizations in BB84 have discrete eigenspectra, whereas in CV QKD the key is encoded in the continuous observables of light. Nevertheless, the sequences of action for implementing QKD will remain the same as we shall see in Section (2.4).

2.3 Gaussian Quantum Optics

In this section we look at the essentials of Gaussian quantum information pertinent to the theoretical analysis of CV QKD. We start our discussion with the Wigner function. It is known that a quantum state, at any time, is represented by a vector in a complex Hilbert space. Wigner function is an equivalent way of representing a quantum optical state in the phase space.

2.3.1 Wigner Distribution

A quantum system composed of N quantized radiation modes, i.e. N quantum harmonic oscillators is described by a tensor-product Hilbert space namely $\mathcal{H}^{\otimes N} = \otimes_{k=1}^N \mathcal{H}_k$, where $\mathcal{H}_k$ corresponds to the state-space for the k -th mode

of radiation [43], [44]. The Hamiltonian of the composite system is given by $\hat{H} = \sum_{k=1}^N \hat{H}_k$, with $\hat{H}_k = \hbar\omega_k(\hat{a}_k^\dagger\hat{a}_k + \frac{1}{2})$. The creation and annihilation operators $\{\hat{a}, \hat{a}^\dagger\}$ for different modes are related via the bosonic commutation relation

$$\begin{aligned} [\hat{a}_j, \hat{a}_k^\dagger] &= \delta_{jk}, \\ [\hat{a}_j, \hat{a}_k] &= 0, \quad [\hat{a}_j^\dagger, \hat{a}_k^\dagger] = 0. \end{aligned} \quad (2.6)$$

If we group N -pairs of field operators $\{\hat{a}_k, \hat{a}_k^\dagger\}_{k=1}^N$ in a vector $\hat{\mathbf{b}} := (\hat{a}_1, \hat{a}_1^\dagger, \dots, \hat{a}_N, \hat{a}_N^\dagger)^T$, the bosonic commutation relation can succinctly be written in terms of the elements of $\hat{\mathbf{b}}$ as [45]

$$[\hat{b}_i, \hat{b}_j] = \Omega_{ij} \quad (i, j = 1, \dots, 2N), \quad (2.7)$$

where Ω_{ij} is the generic element of the following $2N \times 2N$ matrix

$$\Omega := \bigoplus_{k=1}^N \omega = \begin{pmatrix} \omega & & \\ & \ddots & \\ & & \omega \end{pmatrix} \quad \omega := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (2.8)$$

The above structure is known as the symplectic form. An equivalent description for the N -mode system can be provided in terms of the quadrature operators¹ $\{\hat{q}_k, \hat{p}_k\}_{k=1}^N$ with

$$\hat{q}_k = \frac{\hat{a}_k + \hat{a}_k^\dagger}{\sqrt{2}}, \quad \hat{p}_k = \frac{\hat{a}_k - \hat{a}_k^\dagger}{j\sqrt{2}}, \quad (2.9)$$

so that,

$$\begin{aligned} [\hat{q}_j, \hat{p}_k] &= j\delta_{jk}, \\ [\hat{q}_j, \hat{q}_k] &= 0, \quad [\hat{p}_j, \hat{p}_k] = 0. \end{aligned} \quad (2.10)$$

Grouped together in a vector

$$\hat{\mathbf{x}} = \{\hat{q}_1, \hat{p}_1, \hat{q}_2, \hat{p}_2, \dots, \hat{q}_N, \hat{p}_N\}^T, \quad (2.11)$$

the canonical commutation relation (CCR) in terms of the vector elements of $\hat{\mathbf{x}}$ is

$$[\hat{x}_i, \hat{x}_j] = j\Omega_{ij}. \quad (2.12)$$

For the N -mode Hilbert space one can write

$$\hat{\mathbf{x}}^T |\mathbf{x}\rangle = \mathbf{x}^T |\mathbf{x}\rangle \quad (2.13)$$

with $\mathbf{x} \in \mathbb{R}^{2N}$ and $|\mathbf{x}\rangle = (|x\rangle_1, \dots, |x\rangle_{2N})^T$. The vector $\mathbf{x}$ containing continuous quadrature eigenvalues can describe the entire bosonic system.

Associated with each mode k , there is a vacuum state $|0\rangle_k \in \mathcal{H}_k$. The vacuum state of the complete N -mode system is $|0\rangle = \bigotimes_{k=1}^N |0\rangle_k \in \mathcal{H}$. Coherent states

¹equivalent to position and momentum operators for a mechanical harmonic oscillator

$|\alpha\rangle_k$, as described earlier, are mathematically generated by applying the single mode Weyl displacement operator $\hat{D}_k$ to the vacuum $|0\rangle_k$: $|\alpha\rangle_k = \hat{D}_k(\alpha) |0\rangle_k$ where $\hat{D}_k(\alpha) = e^{\alpha \hat{a}_k^\dagger - \alpha^* \hat{a}_k}$, and $\alpha \in \mathbb{C}$ is the coherent amplitude. In a similar fashion, N -mode coherent states are obtained by applying the N -mode Weyl operator $\hat{D}(\boldsymbol{\eta})$ to the global vacuum $|0\rangle$ defined above. The expression for $\hat{D}(\boldsymbol{\eta})$ in terms of the canonical quadrature operator vector $\hat{\mathbf{x}}$ is

$$\hat{D}(\boldsymbol{\eta}) = e^{j\hat{\mathbf{x}}^T \boldsymbol{\Omega} \boldsymbol{\eta}} \quad (2.14)$$

with $\boldsymbol{\eta} \in \mathbb{R}^{2N}$ [46]. $\boldsymbol{\eta}$ is the generalization of α in the N -mode case. Subsequently, one has $|\boldsymbol{\eta}\rangle = \hat{D}(\boldsymbol{\eta}) |0\rangle$. The multimode Weyl operator allows us to express quantum (Wigner) characteristic function for any state ρ in the Hilbert space (Appendix B)

$$\chi(\boldsymbol{\eta}) = \text{Tr}[\hat{\rho} \hat{D}(\boldsymbol{\eta})], \quad (2.15)$$

which, in turn, is related to the *Wigner function* [47] via the Fourier transform

$$W(\mathbf{x}) = \int_{\mathbb{R}^{2N}} \frac{d^{2N}\boldsymbol{\eta}}{(2\pi)^{2N}} \exp(-j\mathbf{x}^T \boldsymbol{\Omega} \boldsymbol{\eta}) \chi(\boldsymbol{\eta}). \quad (2.16)$$

$W(\mathbf{x})$ is normalized to one but generally non-positive. In the above equation, one can consider the continuous variables $\mathbf{x} \in \mathbb{R}^{2N}$ as the eigenvalues of the quadrature operators $\hat{\mathbf{x}}$. These variables span a real symplectic space $\Gamma := (\mathbb{R}^{2N}, \boldsymbol{\Omega})$, called the phase space. The symplectic condition means that any operational transformation on the quadrature operators should preserve the CCRs, or in mathematical notation, $S\boldsymbol{\Omega}S^T = \boldsymbol{\Omega}$. Therefore, one can assert, an arbitrary quantum state $\hat{\rho}$ of an N -mode bosonic system is equivalent to a Wigner function $W(\mathbf{x})$ defined over a $2N$ -dimensional phase space Γ [48]. From the definition of the characteristic function, one can see the tensor product structure for the composite system in Hilbert space is replaced by a direct sum structure in phase space: the N -mode phase space decomposes as the tensor sum $\Gamma = \bigoplus_k \Gamma_k$, where Γ_k is the local phase space associated with mode k .

The Wigner function *almost* qualifies as a joint distribution function for the quadrature coordinates [49], since

(i) The individual quadrature distributions appear as the marginals of the Wigner function. For example, for a single mode field

$$\int_{\mathbb{R}} dp W(\mathbf{x}) = \langle q | \rho | q \rangle, \quad (2.17a)$$

$$\int_{\mathbb{R}} dq W(\mathbf{x}) = \langle p | \rho | p \rangle, \quad (2.17b)$$

where $\mathbf{x} = \begin{pmatrix} q \\ p \end{pmatrix}$ is the coordinate vector [43].

(ii) The Wigner function is real: $W_\rho(\mathbf{x}) \in \mathbb{R} \forall \mathbf{x}, \rho$.

(iii) It is normalized

$$\int d^2\mathbf{x} W_\rho(\mathbf{x}) = 1. \quad (2.18)$$

(iv) The expectation value of any observable $\hat{O}$, which is expressed as a symmetrically (Weyl) ordered function of quadrature operators, can be calculated from $W(\mathbf{x})$ [50]

$$\text{Tr}(\rho\hat{O}) = \langle\hat{O}\rangle = \left\langle f^{(s)}(\hat{q}^m\hat{p}^n) \right\rangle = (2\pi) \int_{\mathbb{R}^2} d^2\mathbf{x} W(\mathbf{x}) \underbrace{f(q^m, p^n)}_{W_{\hat{O}}(\mathbf{x})}, \quad (2.19)$$

where the superscript "s" means symmetrically ordered function, and where $W_O(\mathbf{x})$ is the Wigner function (also called the Weyl-expression) for the operator $\hat{O}$ defined over the phase space. All these properties are very desirable as far as a joint probability distribution is concerned. But, $W(\mathbf{x})$ can be negative as can be seen from a particular example of equation (2.19)

$$\text{Tr}(\rho_1\rho_2) = 2\pi \int_{\mathbb{R}^2} d^2\mathbf{x} W_{\rho_1}(\mathbf{x})W_{\rho_2}(\mathbf{x}) = |\langle\psi_1|\psi_2\rangle|^2; \quad \rho_j = |\psi_j\rangle\langle\psi_j|. \quad (2.20)$$

If $|\psi_1\rangle \perp |\psi_2\rangle$, then $\langle\psi_1|\psi_2\rangle = 0$. It suggests that $W_\rho(\mathbf{x})$ is negative for some values of $(\rho, \mathbf{x})$. That is why $W_\rho(\mathbf{x})$ is not a true distribution, and instead is called a quasi-probability distribution. Like any classical probability distribution, the Wigner (quasi-probability) distribution is completely characterized by its statistical moments. The first moment is called the displacement vector, or simply the mean vector with elements

$$\bar{x}_i = \int_{\mathbb{R}^{2N}} d^{2N}\mathbf{x} W_\rho(\mathbf{x}) x_i = \langle\hat{x}_i\rangle. \quad (2.21)$$

Alternatively, in compact matrix notation

$$\bar{\mathbf{x}} := \langle\hat{\mathbf{x}}\rangle = \text{Tr}(\hat{\mathbf{x}}\rho) = \text{col}(\langle\hat{q}\rangle_1, \langle\hat{p}\rangle_1, \dots, \langle\hat{q}\rangle_N, \langle\hat{p}\rangle_N). \quad (2.22)$$

The second order moment is called the covariance matrix Σ , which is a positive semi-definite $2N \times 2N$ real symmetric matrix with elements

$$\Sigma_{j,k} = \int_{\mathbb{R}^{2N}} d^{2N}\mathbf{x} W_\rho(\mathbf{x}) \delta x_j \delta x_k = \frac{1}{2} \langle \delta \hat{x}_j \delta \hat{x}_k + \delta \hat{x}_k \delta \hat{x}_j \rangle = \frac{1}{2} \langle \{ \delta \hat{x}_j, \delta \hat{x}_k \} \rangle. \quad (2.23)$$

Here $\delta x_j = x_j - \bar{x}_j$ denotes the fluctuation of the quadrature from its mean, while $\delta \hat{x}_j = \hat{x}_j - \langle\hat{x}_j\rangle$ is the fluctuation operator for the j^{th} quadrature. In equation (2.23), the fluctuation operators are put in symmetric order. From this definition, it further follows that

$$\begin{aligned} \Sigma_{jk} &= \frac{1}{2} \langle \{ \hat{x}_j - \langle\hat{x}_j\rangle, \hat{x}_k - \langle\hat{x}_k\rangle \} \rangle = \frac{1}{2} \text{Tr}[\rho \{ \hat{x}_j - \bar{x}_j, \hat{x}_k - \bar{x}_k \}] \\ &= \frac{1}{2} \left(\langle \hat{x}_j \hat{x}_k \rangle + \langle \hat{x}_k \hat{x}_j \rangle \right) - \langle \hat{x}_j \rangle \langle \hat{x}_k \rangle \\ &=: \mathbb{E}(\hat{x}_j, \hat{x}_k), \end{aligned} \quad (2.24)$$

where $\{\}$ denotes the anticommutator. Along the diagonal, one has

$$\Sigma_{jj} = \langle (\hat{x}_j)^2 \rangle - \langle \hat{x}_j \rangle^2 = V(\hat{x}_j) = \langle \delta \hat{x}_j^2 \rangle. \quad (2.25)$$

In compact matrix notation, one can write

$$\Sigma = \langle \delta \hat{\mathbf{x}} \delta \hat{\mathbf{x}}^T \rangle - \mathbf{J} \Omega, \quad (2.26)$$

which in the expanded form looks as the following

$$\Sigma = \begin{pmatrix} V(\hat{q}_1) & \mathbb{E}(\hat{q}_1, \hat{p}_1) & \mathbb{E}(\hat{q}_1, \hat{q}_2) & \mathbb{E}(\hat{q}_1, \hat{p}_2) & \dots & \mathbb{E}(\hat{q}_1, \hat{q}_N) & \mathbb{E}(\hat{q}_1, \hat{p}_N) \\ & V(\hat{p}_1) & \mathbb{E}(\hat{p}_1, \hat{q}_2) & \mathbb{E}(\hat{p}_1, \hat{p}_2) & \dots & \mathbb{E}(\hat{p}_1, \hat{q}_N) & \mathbb{E}(\hat{p}_1, \hat{p}_N) \\ & & V(\hat{q}_2) & \mathbb{E}(\hat{q}_2, \hat{p}_2) & \dots & \mathbb{E}(\hat{q}_2, \hat{q}_N) & \mathbb{E}(\hat{q}_2, \hat{p}_N) \\ & & & V(\hat{p}_2) & \dots & \mathbb{E}(\hat{p}_2, \hat{q}_N) & \mathbb{E}(\hat{p}_2, \hat{p}_N) \\ & & & & \ddots & \vdots & \vdots \\ & \text{symmetric} & & & & V(\hat{q}_N) & \mathbb{E}(\hat{q}_N, \hat{p}_N) \\ & & & & & & V(\hat{p}_N) \end{pmatrix}. \quad (2.27)$$

The covariance matrix describes the mutual correlations among the mode quadratures. If the cross terms $\mathbb{E}(\hat{x}_j, \hat{x}_k)$ are zero for the quadratures of two modes, then these two modes are uncorrelated to each other, that is, separable. The admissible condition on a real symmetric positive-definite $2N \times 2N$ matrix Σ to be a bona fide (quantum) variance matrix is that the Hermitian matrix $\Sigma + \mathbf{J} \Omega$ must be positive semi-definite [51]

$$\Sigma + \mathbf{J} \Omega \geq 0. \quad (2.28)$$

A particular class of states, called the *Gaussian states*, are completely characterized only by their first and second order moments. Any higher order moment of these states can be expressed in terms of their first two moments (Gaussian moment theorem). By definition, Gaussian states have Gaussian Wigner representation [52]

$$W(\mathbf{x}) = \frac{\exp\left\{-\frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}})^T \Sigma^{-1}(\mathbf{x} - \bar{\mathbf{x}})\right\}}{(2\pi)^N \sqrt{\det \Sigma}} \text{ with } \mathbf{x} \in \mathbb{R}^{2N}, \Sigma \in \mathbb{R}^{2N} \times \mathbb{R}^{2N}, \quad (2.29)$$

which is nothing but the $2N$ dimensional generalization of the Gaussian distribution. Together with the definition of Wigner distribution in terms of the characteristic function, and employing the multidimensional Gaussian integral [53]

$$\int_{\mathbb{R}^{2N}} d^{2N} \boldsymbol{\eta} \exp\left[-\frac{1}{2} \boldsymbol{\eta}^T \Sigma \boldsymbol{\eta} + \mathbf{x}^T \boldsymbol{\eta}\right] = \frac{(2\pi)^{2N}}{\sqrt{\det \Sigma}} \exp\left[\frac{1}{2} \mathbf{x}^T \Sigma^{-1} \mathbf{x}\right], \quad (2.30)$$

one can deduce the expression for χ working backwards

$$\chi(\boldsymbol{\eta}) := \exp\left\{-\frac{1}{2} \boldsymbol{\eta}^T (\Omega \Sigma \Omega^T) \boldsymbol{\eta} - \mathbf{J} (\Omega \bar{\mathbf{x}})^T \boldsymbol{\eta}\right\}. \quad (2.31)$$

This subsection attempts to provide background information on how the concept of Gaussian state arises. In summary, quantum states can be represented in terms of phase space Wigner distribution. This Wigner distribution is a Fourier transform of the Wigner characteristic function, which in turn can be shown to be the expectation value of the displacement operator. Quantum states with a Gaussian Wigner distribution are called Gaussian states such as the coherent states and the squeezed states. Gaussian states are fully characterized by their first and second order moments.

2.3.2 Examples of Gaussian States

One mode Gaussian states

A generic one mode Gaussian state has a mean vector $\bar{\mathbf{x}} = \text{col}(\langle \hat{q} \rangle, \langle \hat{p} \rangle)$ containing two elements, and 2×2 covariance matrix of the form

$$\Sigma = \begin{pmatrix} a & c \\ c & b \end{pmatrix}, \quad (2.32)$$

where $a = \text{Var}(\hat{q})$, $b = \text{Var}(\hat{p})$, $c = \mathbb{E}(\hat{q}, \hat{p})$.

Vacuum and Coherent States: The vacuum state is characterized by its displacement vector, $\bar{\mathbf{x}} = (0, 0)$ and covariance matrix $\Sigma = \mathbb{1}_2$. Both the quadratures have a variance of 1 in shot-noise unit. Since coherent states are just displaced vacuum, they have the same covariance matrix as the vacuum, but non-zero displacement vector $\bar{\mathbf{x}} = (q, p)$, as they can be centered anywhere in the phase space.

Squeezed States: Squeezed states differ from coherent states in having squeezed variance (smaller than one shot-noise unit) along one quadrature at the expense of elongated covariance (greater than one shot-noise unit) along the other quadrature. Therefore, they have the same first order moment as the coherent states, but their second order moment is modified as follows

$$\Sigma = \begin{pmatrix} e^{-2r} & 0 \\ 0 & e^{2r} \end{pmatrix}, \quad (2.33)$$

where r is a squeezing parameter. For $r > 0$, the quadrature $\hat{q}$ is squeezed, while the quadrature $\hat{p}$ is anti-squeezed.

Thermal State: The thermal state has null mean value and covariance matrix

$$\Sigma = \begin{pmatrix} V & 0 \\ 0 & V \end{pmatrix}. \quad (2.34)$$

The quantity V can be expressed in terms of the average number of photons $\bar{n}$ contained in the thermal state as $V = 2\bar{n} + 1$. The vacuum can be seen as a special case of a thermal state with $\bar{n} = 0$.

Noisy Coherent State: A thermal state can be displaced by amount $\mathbf{x}$ on the phase space, and its variance can be interpreted as $V = 1 + \varepsilon$, where ε is due

to excess noise on top of shot-noise. Therefore, a displaced thermal state can be considered as a noisy coherent state.

Two mode Gaussian States: A generic two mode Gaussian state is characterized by a displacement vector $\bar{\mathbf{x}} = \bar{\mathbf{x}}_1 \otimes \bar{\mathbf{x}}_2 = (q_1, p_1, q_2, p_2)$, and a covariance matrix

$$\Sigma_{\text{tot}} = \Sigma_1 \oplus \Sigma_2 = \begin{pmatrix} \Sigma_1 & C \\ C^T & \Sigma_2 \end{pmatrix}. \quad (2.35)$$

$\Sigma_{1(2)}$ is the covariance matrix of the first (second) mode after tracing out the second (first) mode from the global state.

Partial Trace Let us consider a system made up of two Gaussian modes A and B . In the phase space picture, tracing out mode B is equivalent to calculating the marginal of the Wigner function of the system with respect to the quadrature variables (eigenvalues) of B namely (q_B, p_B) [53]

$$\text{Tr}_B[\rho_{AB}] = (2\pi)^2 \int_{\mathbb{R}^2} W(q_A, p_A, q_B, p_B) dq_B dp_B \quad (2.36)$$

which can be shown equivalent to setting $\xi_B = 0$ in the characteristic function

$$\chi_A(\xi_A) = \text{Tr}_A[\rho_A \hat{D}_{\xi_A}] = \text{Tr}_{AB}[\rho_{AB} \hat{D}_{\xi_A} \otimes \mathbb{1}_B] = \chi_{AB}(\xi_A, \xi_B = 0) \quad (2.37)$$

In terms of statistical operation, the tracing out operation can be summarized as follows

$$\Sigma_{AB} = \begin{pmatrix} \Sigma_A & C \\ C^T & \Sigma_B \end{pmatrix} \xrightarrow{\text{P.T.}} \Sigma_A \quad (2.38)$$

$$\bar{\mathbf{x}}_{AB} = (\bar{\mathbf{x}}_A, \bar{\mathbf{x}}_B) \xrightarrow{\text{P.T.}} \bar{\mathbf{x}}_A \quad (2.39)$$

C is the matrix that gives the correlation (classical, or quantum) between the two modes. The case $C = 0$ corresponds to an absence of correlation between two Gaussian modes. Generally speaking, an N -mode Gaussian state with a product-state structure $\rho = \rho_1 \otimes \rho_2 \otimes \dots \otimes \rho_N$ has the block diagonal covariance matrix structure $\Sigma = \Sigma_1 \oplus \Sigma_2 \oplus \dots \oplus \Sigma_N$.

Two Mode Squeezed Vacuum States (TMSVS): For our purpose, the most interesting two mode Gaussian state is the *two mode squeezed vacuum states* with null mean $\bar{\mathbf{x}} = (0, 0, 0, 0)$ and covariance matrix

$$\Sigma_{\text{TMSVS}} = \begin{pmatrix} \cosh 2r \mathbb{1}_2 & \sinh 2r \sigma_z \\ \sinh 2r \sigma_z & \cosh 2r \mathbb{1}_2 \end{pmatrix} \quad (2.40)$$

where $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ [54]. Tracing out one mode of a TMSVS gives a thermal state with variance $\cosh 2r = 2\bar{n} + 1$. Thus the TMSVS can be thought of as the purification of a thermal state. In this respect, in the security analysis of CV QKD, the TMSVS plays the same role as the Bell state $|\phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ in discrete variable QKD.

This subsection introduces some example Gaussian states in terms of their first and second order moments. Coherent states are the enablers of our QKD technology. For coherent states, the dispersion in both the quadratures equal to one shot noise unit. Squeezed states show sub-shot-noise variance in one quadrature. The two mode squeezed vacuum state is an important theoretical construct to understand the entanglement based version of the CV QKD protocol.

2.3.3 Gaussian Operations

In general, a quantum operation $\mathcal{O}$ is called Gaussian if it maps one Gaussian state to another $\mathcal{O} : \rho_G(\bar{\mathbf{x}}, \Sigma) \rightarrow \rho_G(\bar{\mathbf{x}}', \Sigma')$. The class of Gaussian operations which are completely-positive and trace preserving i.e. $\text{Tr}[\rho] = \text{Tr}[\mathcal{O}(\rho)]$ are called Gaussian channels. Reversible Gaussian channels are called Gaussian unitaries $\hat{U}_G$, which transform a state according to the rule $\rho \rightarrow \hat{U}_G \rho \hat{U}_G^\dagger$.

Gaussian Unitary Transformation

It can be shown that the Hamiltonian $\hat{H}$ generating a Gaussian unitary ($\hat{U} = \exp[-j\hat{H}]$) can at best be quadratic in the quadrature operators ($\hat{q}, \hat{p}$) [46]. Consequently, it is possible to implement Gaussian unitaries with commonly accessible laboratory apparatus (phase shifter, beam splitters, squeezers etc). A Gaussian unitary is completely characterized by its action on the statistical moments of the Gaussian states. This action can be deduced from the consequence of Stone-von Neumann theorem: *any unitary transformation $\hat{U}$ on the Hilbert space uniquely corresponds to a real symplectic operation $\mathbf{S} \in \text{Sp}(2N, \mathbb{R})$ on the phase space* [55]. In particular, a Gaussian unitary $\hat{U}_{\mathbf{S}, \mathbf{d}}$ acting on a state $\rho(\bar{\mathbf{x}}, \Sigma)$ induces the following transformations on the state's statistical moments

$$\bar{\mathbf{x}} \longrightarrow \mathbf{S}\bar{\mathbf{x}} + \mathbf{d}, \quad (2.41a)$$

$$\Sigma \longrightarrow \mathbf{S}\Sigma\mathbf{S}^T \quad (2.41b)$$

with $\mathbf{d} \in \mathbb{R}^N$, $\mathbf{S} \in \mathbb{R}^{2N \times 2N}$. There is no restriction on $\mathbf{d}$, but $\mathbf{S}$ must preserve the canonical commutation relation

$$\mathbf{S}\Omega\mathbf{S}^T = \Omega. \quad (2.42)$$

Equations (2.41a), (2.41b) have the following implication on the quadrature operator vector (Appendix B)

$$\hat{U}_{\mathbf{S}, \mathbf{d}} : \hat{\mathbf{x}} \rightarrow \mathbf{S}\hat{\mathbf{x}} + \mathbf{d}. \quad (2.43)$$

Gaussian unitaries can be classified into *passive transformations* and *active transformations*. A passive unitary transformation conserves the mean photon of the Gaussian state. For an N -mode Gaussian state, the mean photon number is given by ²

$$\sum_{l=1}^N \langle \hat{N}_l \rangle = \frac{1}{4} \text{Tr}[\Sigma] + \frac{1}{4} \bar{\mathbf{x}}^2 - \frac{N}{2}. \quad (2.44)$$

²extends from the specific case of a single mode, $\langle \hat{n} \rangle = \frac{1}{4} (\langle \hat{q}^2 \rangle + \langle \hat{p}^2 \rangle) - \frac{1}{2}$

Therefore, a Gaussian unitary will be a passive transformation if and only if

$$\mathbf{d} = 0 \quad \text{and} \quad \mathbf{S}^T \mathbf{S} = \mathbf{1}_{2N \times 2N}. \quad (2.45)$$

The condition on $\mathbf{S}$ means it has to be orthogonal $\mathbf{S}^T = \mathbf{S}^{-1}$. Putting everything together, a passive Gaussian unitary is not only symplectic, but also orthogonal forming a group $P(N) = \text{Sp}(2N, \mathbb{R}) \cap O(2N)$. Passive transformations include phase shift, beam splitter, or combination of them in different configurations (multiport interferometer). In contrast, active transformations $A(N) = \text{Sp}(2N, \mathbb{R}) \setminus P(N)$ (such as squeezing, displacement) do not conserve the mean photon number.

Examples of Gaussian Unitaries

Displacement: The effect of displacement operator on a Gaussian state is only to translate its mean, leaving the covariance matrix invariant. For the single mode case, the displacement operator is $\hat{D}(\alpha) := \exp(\alpha \hat{a}^\dagger - \alpha^* \hat{a})$, where $\alpha = (q + jp)/\sqrt{2}$, and for the N -mode case it is the same as equation (2.14). The quadrature operators $\hat{\mathbf{x}} = (\hat{q}, \hat{p})^T$ get translated as $\hat{\mathbf{x}} \rightarrow \hat{\mathbf{x}} + \mathbf{d}$, where $\mathbf{d} = (\Delta q, \Delta p)^T$. Displacement is an active unitary ($\mathbf{d} \neq 0$).

Phase Shift: A phase shift operation rotates the mean value of a Gaussian state. For the single mode case it corresponds to a symplectic map $S_{\text{PS}}(\theta) \in \text{Sp}(2, \mathbb{R})$

$$S_{\text{PS}}(\theta) = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}. \quad (2.46)$$

Squeezing: A single mode squeezing operation corresponds to the symplectic transformation $S_{\text{SQ}}^1(r) \in \text{Sp}(2, \mathbb{R})$

$$S_{\text{SQ}}^1(r) = \begin{pmatrix} e^{-r} & 0 \\ 0 & e^r \end{pmatrix}, \quad (2.47)$$

where r is the squeezing factor. For the two-mode squeezing operation, the symplectic transformation $S_{\text{SQ}}^2(r) \in \text{Sp}(4, \mathbb{R})$ is

$$S_{\text{SQ}}^2(r) = \begin{pmatrix} \cosh r \mathbb{1}_2 & \sinh r \sigma_z \\ \sinh r \sigma_z & \cosh r \mathbb{1}_2 \end{pmatrix}. \quad (2.48)$$

Here σ_z is the Pauli- z matrix.

Beam splitter: A beam splitter with characteristic transmittance T performs unitary operation on two Gaussian modes and is described by the symplectic transformation $S_{\text{BS}}(T) \in \text{Sp}(4, \mathbb{R})$ as

$$S_{\text{BS}}(T) = \begin{pmatrix} \sqrt{T} \mathbb{1}_2 & \sqrt{1-T} \mathbb{1}_2 \\ -\sqrt{1-T} \mathbb{1}_2 & \sqrt{T} \mathbb{1}_2 \end{pmatrix}. \quad (2.49)$$

This is the most important symplectic transformation in the context of this dissertation.

Generalized Unitary Transformation: Any real symplectic matrix $S \in \text{Sp}(2N, \mathbb{R})$ can be decomposed in the following form [56]

$$S = K \bigoplus_{k=1}^N \begin{pmatrix} e^{-r_k} & 0 \\ 0 & e^{r_k} \end{pmatrix} L, \quad (2.50)$$

which is known as the Euler decomposition, or the Bloch-Messiah reduction of a symplectic matrix. Here $K, L \in P(N)$ and $r_k \in \mathbb{R}$. Since (tensor) addition in the phase space corresponds to (tensor) product in the Hilbert space, the above result is equivalent to

$$\hat{U}_S = \hat{U}_K \underbrace{\left[\bigotimes_{k=1}^N S(r_k) \right]}_{\text{Euler decomposition}} \hat{U}_L, \quad (2.51)$$

which can be interpreted as: any Gaussian unitary map on N modes can be implemented by successive application of a linear interferometer K (N inputs $\times$ N outputs), followed by a set of N single mode squeezers, followed by another interferometer L (N inputs $\times$ N outputs).

In summary, Gaussian operations or unitaries do not alter the Gaussianity of the corresponding states. A unitary transformation on the states in the Hilbert space translates to a symplectic transformation of their moments in the phase space. The symplectic transformation must preserve canonical commutation relations. The most important Gaussian unitary going forward is the beam splitter.

2.3.4 Gaussian Measurements

Partial Measurements

Consider an N mode Gaussian state described by a covariance matrix of the dimension $2N \times 2N$ (two quadratures per mode). A measurement on one mode will modify the Gaussian state of the remaining modes, depending on their correlations to the measured one. Let B be the mode on which a measurement is performed. A be the $N - 1$ remaining modes and C the correlations between B and A . Then the covariance matrix of the total state before the measurement can be written as

$$\Sigma = \begin{pmatrix} \Sigma_A & \Sigma_C \\ \Sigma_C^T & \Sigma_B \end{pmatrix}, \quad (2.52)$$

where $\Sigma_A \in \mathbb{R}^{2(N-1) \times 2(N-1)}$, $\Sigma_B \in \mathbb{R}^{2 \times 2}$ and $\Sigma_C \in \mathbb{R}^{2(N-1) \times 2}$. The impact of a partial measurement on mode B on the covariance matrix of the other modes, $\Sigma_A \rightarrow \Sigma_{A|B}$, depends on whether homodyne or heterodyne measurement is performed, that is, on whether only one or both of B 's quadratures are measured. In case of homodyne detection of q or p , by a partial measurement on

B , the state ρ_A will be projected onto the Gaussian state ρ'_A characterized by the covariance matrix (please refer to Appendix B)

$$\Sigma_{A|B} = \Sigma_A - \Sigma_C(\Pi_{q,p}\Sigma_B\Pi_{q,p})^{-1}\Sigma_C^T, \quad (2.53)$$

where $\Pi_q = \text{diag}(1, 0)$ (in case q is measured) and $\Pi_p = \text{diag}(0, 1)$ (in case p is measured). Therefore, one can express the transformation as

$$\text{q measurement : } \Pi_q\Sigma_B\Pi_q = \begin{pmatrix} \Sigma_B^{11} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} V(\hat{q}_B) & 0 \\ 0 & 0 \end{pmatrix} \quad (2.54a)$$

$$\text{p measurement : } \Pi_p\Sigma_B\Pi_p = \begin{pmatrix} 0 & 0 \\ 0 & \Sigma_B^{22} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & V(\hat{p}_B) \end{pmatrix} \quad (2.54b)$$

The diagonal matrix keeps track of which quadrature was measured. The inverse expression indicates a pseudo-inverse. However if a matrix is diagonal, as it is the case for $\Pi\Sigma_B\Pi$ (for any Σ_B) the pseudo-inverse is just an inversion of the elements: $\text{diag}(a_1, a_2, \dots, a_n)^{-1} = \text{diag}(a_1^{-1}, a_2^{-1}, \dots, a_n^{-1})$. Re-expressing (2.53) in convenient form, one obtains [57]

$$\Sigma_{A|B} = \Sigma_A - \frac{1}{V(\hat{q}_B, \hat{p}_B)}\Sigma_C\Pi_{q,p}\Sigma_C^T. \quad (2.55)$$

A heterodyne detection of the mode B will affect the other modes as follows

$$\Sigma_{A|B} = \Sigma_A - \Sigma_C(\Sigma_B + \mathbb{1}_2)^{-1}\Sigma_C^T. \quad (2.56)$$

When B is a coherent state with $V(\hat{q}_B) = V(\hat{p}_B) =: V_B$, equation (2.55) simplifies to [57]

$$\Sigma_{A|B} = \Sigma_A - \frac{1}{V_B + 1}\Sigma_C\Sigma_C^T. \quad (2.57)$$

The above concept of partial measurement will be helpful to understand the entanglement based version of the CV QKD protocol. Firstly, Alice's heterodyne measurement of her mode, derived from the two mode squeezed vacuum states, with a 50:50 beam splitter will project Bob's state into a coherent state. Secondly, the upper bound on Eve's knowledge about the communication between Alice and Bob will depend on whether Bob performs homodyne or heterodyne detection on his coherent states.

2.3.5 Entropy of a Gaussian State

A generic N -mode Gaussian state of a continuous variable system is described by its density operator in the Fock basis

$$\rho = \sum_{m,n=0}^{\infty} \rho_{m_1,m_2,\dots,m_N,n_1,n_2,\dots,n_N} |m_1, m_2, \dots, m_N\rangle \langle n_1, n_2, \dots, n_N|. \quad (2.58)$$

The Von Neuman entropy of this state is

$$S(\rho) = -\text{Tr} \rho \log \rho. \quad (2.59)$$

The entropy of a Gaussian state is entirely determined by its covariance matrix. According to Williamson's theorem for every positive definite real matrix $V \in \mathbb{R}^{2N \times 2N}$, there exists a symplectic matrix S and diagonal matrix D such that [58]

$$V = SDS^T \Rightarrow \Sigma = S\Sigma^\oplus S^T, \quad \Sigma^\oplus := \bigoplus_{k=1}^N \nu_k \mathbb{1}_2, \quad (2.60)$$

where ν_k 's are called the symplectic eigenvalues of Σ , or equivalently the eigenvalues of $|J\Omega\Sigma|$ where $||$ represents element-wise absolute value. The symplectic transformation on the covariance matrix suggests there exists a corresponding unitary mapping on the Hilbert space which transforms the Gaussian state ρ_G to a product of N single mode thermal states with $\bar{n}_k = \frac{1}{2}(\nu_k - 1)$ photons in the mode k

$$\hat{\rho}_G = \bigotimes_{k=1}^N \hat{\rho}_{\text{th}}\left(\frac{\nu_k - 1}{2}\right). \quad (2.61)$$

If $\rho_{\text{th}}(\bar{n})$ is the single mode thermal state with a mean photon number $\bar{n}$, then

$$S(\rho_G) = \sum_{k=1}^N S(\rho_{\text{th}}(\bar{n}_k)). \quad (2.62)$$

Entropy of a thermal state

The density operator of a single mode thermal state ρ_{th} is given by [43]

$$\rho_{\text{th}} = \sum_{n=0}^{\infty} \frac{\bar{n}^n}{(\bar{n} + 1)^{n+1}} |n\rangle \langle n|, \quad (2.63)$$

where $\bar{n}$ is the mean number of photon thermal photons. The corresponding Von Neumann entropy is

$$\begin{aligned} S(\rho_{\text{th}}) &= -\text{Tr}[\rho_{\text{th}} \log_2 \rho_{\text{th}}] \\ &= -\frac{1}{\bar{n} + 1} \sum_{k=0}^{\infty} \left(\frac{\bar{n}}{\bar{n} + 1}\right)^k \log_2 \left[\frac{1}{\bar{n} + 1} \left(\frac{\bar{n}}{\bar{n} + 1}\right)^k \right] \\ &= -\frac{1}{\bar{n} + 1} \sum_{k=0}^{\infty} \left(\frac{\bar{n}}{\bar{n} + 1}\right)^k \left[-\log_2(\bar{n} + 1) + k \log_2 \frac{\bar{n}}{\bar{n} + 1} \right]. \end{aligned} \quad (2.64)$$

By making use of the following identity

$$\sum_{k=0}^{\infty} kx^k = \frac{x}{(1-x)^2}, \quad (2.65)$$

one finally arrives at

$$\begin{aligned} S(\rho_{\text{th}}) &= (\bar{n} + 1) \log_2(\bar{n} + 1) - \bar{n} \log_2 \bar{n} \\ &= \left(\frac{\nu + 1}{2}\right) \log_2\left(\frac{\nu + 1}{2}\right) - \left(\frac{\nu - 1}{2}\right) \log_2\left(\frac{\nu - 1}{2}\right) := g(\nu). \end{aligned} \quad (2.66)$$

Therefore, from equation (2.62), $S(\rho_G) = \sum_k g(\nu_k)$. The concept of Von Neumann entropy of a Gaussian state, or a thermal state will be required in section (2.11) to find the upper bound on Eve's knowledge about the shared states between Alice and Bob.

2.4 Basics of a coherent state CV QKD protocol

Early quantum optical implementation of quantum key distribution protocols involved encoding the key into those observables of light which have discrete eigenspectra such as polarization, phase, or time bin of a single photon state of light (hence the name *discrete variable* (DV) QKD protocols). The key is recovered upon detection of single photons. On the other hand, *continuous variable* (CV) quantum key distribution, first introduced with squeezed states and subsequently implemented with coherent states, employs continuous observables of Gaussian states of light (such as field quadratures/amplitude and phase) for encoding the secret key. The key is recovered via balanced coherent detection (homodyne or heterodyne) of light. The CV QKD protocols with coherent states have many practical advantages. For example: (i) the sources for generating, and the equipment for manipulating coherent states are readily available in the laboratory; (ii) single photon detectors are replaced by homodyne detection which is much faster and more efficient; (iii) compatible with the already established fiber-based telecommunication infrastructure, hence more suited toward commercial deployment.

The essential steps involved in a coherent state based CV QKD protocol are outlined below. Like for any QKD protocol, the steps can be divided into a quantum phase (steps 1 to 3) involving generation, transmission, and measurement of coherent states, followed by classical post-processing (steps 4 to 7). The classical post processing (basically error correction) enables Alice and Bob to extract a secure key from the raw key distributed during the quantum phase. The quantum states can be transmitted via a quantum channel (optical fiber) which Eve can tap in. On the other hand, the error correcting steps are performed via a classical authenticated channel. With such a channel, Alice and Bob can be certain that they are communicating among themselves, and Eve is not posing as either of them (ruling out the possibility for man-in-the-middle attack).

1. **Preparation:** Alice prepares a sequence of N coherent states $|\alpha_1\rangle, |\alpha_2\rangle, \dots, |\alpha_j\rangle, \dots, |\alpha_N\rangle$ according to some distribution (for example, Gaussian or discrete) where

$$|\alpha_j\rangle = |q_j + jp_j\rangle \quad (2.67)$$

(q_j, p_j) is the coordinate on the complex $q - p$ plane around which $|\alpha_j\rangle$ is centered.

2. **Transmission:** She sends the N coherent states to Bob through the quantum channel.
3. **Measurement:** Bob uses either homodyne detection, or heterodyne detection to measure the eigenvalues of one or both the quadrature operators. In the former case, he obtains N classical variables (randomly measured q_i or p_i for any $|\alpha\rangle_i$) while in the later case he obtains $2N$ variables (q_i 's and p_i 's for any $|\alpha\rangle_i$).
4. **Sifting:** In the homodyne detection case, Bob needs to inform Alice about the choice of quadratures. Alice keeps only the relevant quadrature for each state. This step is known as *sifting*. At the end of the process, Alice and Bob share N pairs of correlated classical variables $(x_1, y_1), \dots, (x_N, y_N)$. In the heterodyne detection case, no sifting is required.
5. **Parameter Estimation:** Alice and Bob randomly choose a subset (m or $2m$ number) of their data $(x_{i_1}, y_{i_1}), \dots, (x_{i_m}, y_{i_m})$, and reveal them publicly to compare the quadrature values which were used for state preparation and the corresponding measurements. This comparison allows them to estimate the transmission T and the excess noise ξ of the quantum channel. This *parameter estimation* step allows Alice and Bob to compute their mutual information I_{AB} , and bound Eve's information χ . If $\beta I_{AB} < \chi$, where β is the reconciliation efficiency, they abort the protocol. Otherwise they carry on with the protocol with n (or $2n$) pairs of classical correlated variables where $n = N - m$ (or $2n = 2N - 2m$).
6. **Information Reconciliation:** The next task for Alice and Bob is to perform information reconciliation on their correlated set of data. There are two variations: *direct reconciliation*- Bob corrects his bits according to Alice's data, or *reverse reconciliation*-Alice corrects her bit according to Bob's data. In the direct reconciliation scenario, if $T < 0.5$ ($\approx 3\text{dB}$), which means that half of Alice's states were lost through the channel, hence no secret key can be distilled assuming Eve is responsible for the entire loss. On the other hand, this problem can be overcome with the reverse reconciliation. Since Alice's information on Bob's measurement results is always greater than Eve's $I_{AB} > \chi$ for any amount of transmission T (in this case the excess noise, ξ plays a more important role). In reality, this reconciliation is achieved thanks to the error correcting codes (such as low density parity check codes) widely used in the telecommunication industry.
7. **Privacy Amplification:** Following the reconciliation step, Alice and Bob should share a common bit string. That bit string is not completely secret, since some of the information was leaked to Eve during the reconciliation procedure. To extract a key, Alice and Bob will perform *privacy amplification*. In this step Alice and Bob choose randomly a hashing-function from a family of universal functions which takes their common binary string as input and outputs a secure key of size l .

Moving forward, we shall be particularly interested in the theoretical implication of the quantum stage of the protocol (preparation, transmission, and

measurement of coherent states) and how one can estimate the useful channel parameters, the transmittance and the channel noise, from the experimentally observed data.

2.5 Gaussian Modulation

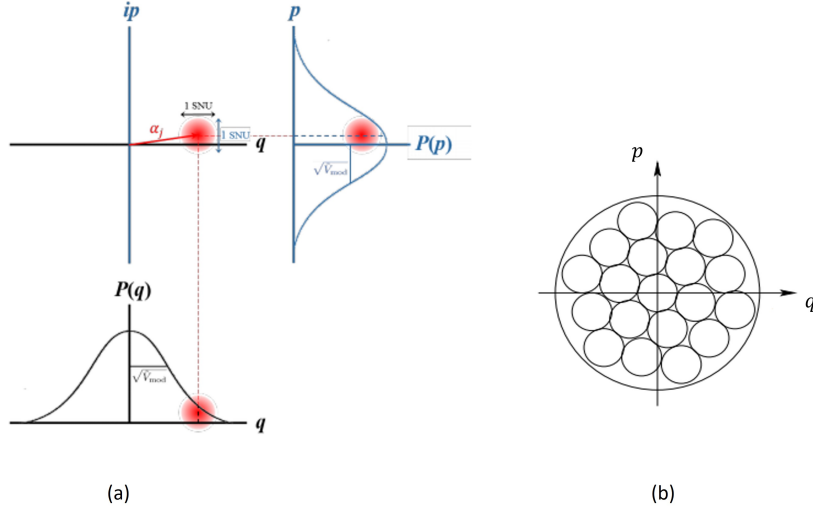


FIGURE 2.2: (a) Phase space representation of Gaussian modulation. The red fuzzy circle denotes one of the many coherent states prepared by Alice. The fuzziness is due to the fluctuation associated with the quadrature operator measurements on a *single* coherent state. The mean values of both the quadratures for *all* the coherent states follow a Gaussian distribution. The standard deviation of this distribution is $\sqrt{\tilde{V}_{\text{mod}}}$ [54] (b) The state ensemble effectively looks like a thermal state (the outer circle) with variance V_{mod}

In the Gaussian modulation scheme (Figure (2.2)), for the sequence of states prepared by Alice in equation (2.67), the quadrature means q_j 's and p_j 's ($2N$ in total) can be thought of as the realizations of two independent and identically distributed (i.i.d.) random variables $\mathcal{Q}$ and $\mathcal{P}$ with zero-mean Gaussian distribution

$$\mathcal{Q} \sim \mathcal{P} \sim \mathcal{N}(0, \tilde{V}_{\text{mod}}), \quad (2.68)$$

where $\tilde{V}_{\text{mod}}$ is the modulation variance of the means. If we denote the state ensemble by

$$\rho = \sum_j P_j |\alpha_j\rangle \langle \alpha_j|, \quad (2.69)$$

where P_j 's are the weights appropriate for a Gaussian modulation, then it follows that

$$\langle \mathcal{Q} \rangle = \sum_j P_j q_j = \langle \mathcal{P} \rangle = \sum_j P_j p_j = 0, \quad (2.70a)$$

$$\langle \mathcal{Q}^2 \rangle = \sum_j P_j q_j^2 = \langle \mathcal{P}^2 \rangle = \sum_j P_j p_j^2 = \tilde{V}_{\text{mod}}. \quad (2.70b)$$

The mean photon number of each individual state $|\alpha_j\rangle$ is given by

$$\langle \hat{n}_j \rangle = \langle \alpha_j | \hat{n} | \alpha_j \rangle = |\alpha_j|^2 = q_j^2 + p_j^2, \quad (2.71)$$

from which the mean photon number of the state ensemble can be deduced as

$$\langle \hat{n} \rangle = \langle \mathcal{Q}^2 \rangle + \langle \mathcal{P}^2 \rangle = 2\tilde{V}_{\text{mod}}, \quad (2.72)$$

where q_j , and p_j are considered as the realization of the random variables $\mathcal{Q}$ and $\mathcal{P}$. Now the variance of the quadrature operators applied to the state ensemble is

$$V(\hat{q}) = \langle \hat{q}^2 \rangle - \underbrace{\langle \hat{q} \rangle^2}_{=0} = \langle \hat{q}^2 \rangle \quad (2.73)$$

with

$$\langle \hat{q}^2 \rangle = \text{Tr}(\rho \hat{q}^2) = \text{Tr}\left(\sum_j P_j |\alpha_j\rangle \langle \alpha_j| \hat{q}^2\right) = 4 \sum_j P_j q_j^2 + 1 = 4\tilde{V}_{\text{mod}} + 1. \quad (2.74)$$

Let us define $4\tilde{V}_{\text{mod}} + 1 := V_{\text{mod}} + 1$. To distinguish, $\tilde{V}_{\text{mod}}$ represents the variance of the classical random variables $(\mathcal{Q}, \mathcal{P})$ whose realizations are the quadrature means, whereas $V := V_{\text{mod}} + 1$ represents the modulation variance of quantum mechanical observables (the quadrature operators $\hat{q}, \hat{p}$) of the state ensemble. The particular case $\tilde{V}_{\text{mod}} = 0$ represents the vacuum, and the relevant variance for the quadrature operators in this case is $V_0 = 1$. V_0 is referred to as shot noise, which arises independently, statistically speaking, from the modulation variance. Finally, the mean photon number of the ensemble expressed in terms of the variance of the quadrature operators is

$$\langle \hat{n} \rangle = \frac{1}{2} V_{\text{mod}} = \frac{1}{2} (V - 1). \quad (2.75)$$

After the preparation of each coherent state, Alice transmits $|\alpha_j\rangle$ to Bob per use of a Gaussian quantum channel. Upon reception, Bob uses either homodyne or heterodyne detection to measure the eigenvalue of one or both the quadrature operators, and he obtains N (for homodyne) or $2N$ (for heterodyne) correlated values. This is the so-called prepare-and-measure (P&M) way of implementing a CV QKD protocol.

2.6 Entanglement based version of CV QKD

There is an equivalent way of describing the quantum phase of the Gaussian modulated (GM) CV-QKD protocol (shown in Figure 2.3), which turns out to be more conducive to security analysis. In this entanglement based version of the protocol, there is a source (there is no restriction on where the source is located) which prepares two-mode squeezed vacuum states (TMSVS) and distributes one mode to Alice and another mode to Bob. The TMSVS in ket notation can be expressed as follows

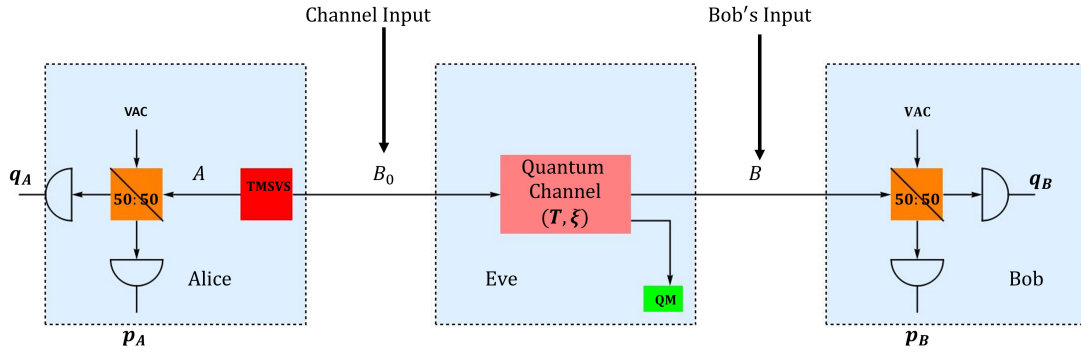


FIGURE 2.3: Entanglement based scheme of CV QKD. The 50 : 50 beam splitters characterize the measurements on Alice's and Bob's sides; while the transmittance T , and noise ξ are controlled by Eve. The QM box represents Eve's quantum memory [59].

$$|\text{TMSVS}\rangle = \left(\frac{2}{V+1}\right)^{1/2} \sum_{n=0}^{\infty} \left(\frac{V-1}{V+1}\right)^{n/2} |n, n\rangle \quad (2.76)$$

with $V = \cosh 2r$ quantifying the noise variance in both of the quadratures. The covariance matrix of this TMSVS can be derived according to [54] as

$$\Sigma_0^{\text{EB}} = \Sigma_{\text{TMSVS}} = \begin{pmatrix} V\mathbb{1}_2 & \sqrt{V^2-1}\sigma_z \\ \sqrt{V^2-1}\sigma_z & V\mathbb{1}_2 \end{pmatrix} := \begin{pmatrix} v & w \\ w & v \end{pmatrix} \quad (2.77)$$

with $v = V\mathbb{1}_2$ and $w = \sqrt{V^2-1}$. Equation (2.77) describes the covariance matrix for the joint state of Alice and Bob at the start of the protocol. In order to collapse Bob's mode into a coherent state, Alice performs a balanced heterodyne measurement on her mode. She splits her mode in half with a balanced beam splitter, and performs simultaneous measurement on both the quadratures [60]. The input covariance matrix to the beamsplitter includes an additional vacuum mode

$$\Sigma_1^{\text{EB}} = \Sigma_{\text{vac}} \oplus \Sigma_{\text{TMSVS}} = \begin{pmatrix} \mathbb{1}_2 & 0 & 0 \\ 0 & v & w \\ 0 & w & v \end{pmatrix}. \quad (2.78)$$

The symplectic expression for the beamsplitter matrix as per equation (2.49) is

$$\text{BS} = \begin{pmatrix} \sqrt{T}\mathbb{1}_2 & \sqrt{1-T}\mathbb{1}_2 \\ -\sqrt{1-T}\mathbb{1}_2 & \sqrt{T}\mathbb{1}_2 \end{pmatrix} := \begin{pmatrix} t & r \\ -r & t \end{pmatrix}. \quad (2.79)$$

The 50 : 50 beam splitter operation does nothing to the Bob's mode, but splits Alice's mode in half. Therefore, effectively,

$$\text{BS}_{\text{eff}} = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1}_2 & \mathbb{1}_2 \\ -\mathbb{1}_2 & \mathbb{1}_2 \end{pmatrix}_{\text{A, vac}} \oplus \mathbb{1}_B = \begin{pmatrix} \mathbb{1}_2/\sqrt{2} & \mathbb{1}_2/\sqrt{2} & 0 \\ -\mathbb{1}_2/\sqrt{2} & \mathbb{1}_2/\sqrt{2} & 0 \\ 0 & 0 & \mathbb{1}_2 \end{pmatrix}. \quad (2.80)$$

The transformed covariance matrix is then

$$\begin{aligned}
 \Sigma_1^{\text{EB}}(\rho^{A_1 A_2 B}) &= \text{BS}_{\text{eff}} \Sigma^{\text{EB}} \text{BS}_{\text{eff}}^T \\
 &= \begin{pmatrix} \mathbb{1}_2/\sqrt{2} & \mathbb{1}_2/\sqrt{2} & 0 \\ -\mathbb{1}_2/\sqrt{2} & \mathbb{1}_2/\sqrt{2} & 0 \\ 0 & 0 & \mathbb{1}_2 \end{pmatrix} \begin{pmatrix} \mathbb{1}_2 & 0 & 0 \\ 0 & v & w \\ 0 & w & v \end{pmatrix} \begin{pmatrix} \mathbb{1}_2/\sqrt{2} & -\mathbb{1}_2/\sqrt{2} & 0 \\ \mathbb{1}_2/\sqrt{2} & \mathbb{1}_2/\sqrt{2} & 0 \\ 0 & 0 & \mathbb{1}_2 \end{pmatrix} \\
 &= \begin{pmatrix} \frac{V+1}{2} \mathbb{1}_2 & \frac{V-1}{2} \mathbb{1}_2 & \sqrt{\frac{1}{2}(V^2-1)} \sigma_z \\ \frac{V-1}{2} \mathbb{1}_2 & \frac{V+1}{2} \mathbb{1}_2 & \sqrt{\frac{1}{2}(V^2-1)} \sigma_z \\ \sqrt{\frac{1}{2}(V^2-1)} \sigma_z & \sqrt{\frac{1}{2}(V^2-1)} \sigma_z & V \mathbb{1}_2 \end{pmatrix}. \quad (2.81)
 \end{aligned}$$

The above covariance matrix refers to the joint state of Alice and Bob just before Alice's photodetector has registered a current. The effect of Alice's balanced heterodyne detection on Bob's mode is succinctly described according to the equation (2.57)

$$\begin{aligned}
 \Sigma_{B|A}^{\text{EB}} &= \Sigma_B - \frac{1}{V_A + 1} \Sigma_C \Sigma_C^T \quad \text{where } V_A = \text{variance of Alice} \\
 &= \begin{pmatrix} V & 0 \\ 0 & V \end{pmatrix} - \frac{1}{V+1} \begin{pmatrix} \sqrt{V^2-1} & 0 \\ 0 & -\sqrt{V^2-1} \end{pmatrix} \begin{pmatrix} \sqrt{V^2-1} & 0 \\ 0 & -\sqrt{V^2-1} \end{pmatrix} \\
 &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (2.82)
 \end{aligned}$$

which is nothing but the representative covariance matrix for a coherent state. Moreover, Bob's mean vector is modified as [53]

$$\bar{\mathbf{x}}_B \mapsto \bar{\mathbf{x}}_B + \Sigma_C \frac{1}{\Sigma_A + \mathbb{1}_2} (\bar{\mathbf{x}}_m - \bar{\mathbf{x}}_A) = \begin{pmatrix} q_m \sqrt{\frac{V-1}{V+1}} \\ p_m \sqrt{\frac{V-1}{V+1}} \end{pmatrix}, \quad (2.83)$$

meaning that Bob's projected coherent state will be centered at $(q_m \sqrt{\frac{V-1}{V+1}}, p_m \sqrt{\frac{V-1}{V+1}})$ on the phase space. While arriving at the above equation, we have used the fact that for a TMSVS, $\bar{\mathbf{x}}_A = \bar{\mathbf{x}}_B = (0, 0)^T$. The measurement outcomes for Alice's heterodyne detection measurements are listed in $\bar{\mathbf{x}}_m = (q_m, p_m)^T$. Therefore one can conclude that with N copies of TMSVS, it is possible to simulate a Gaussian modulated CV QKD. Since the TMSVS's are Gaussian states, Alice's heterodyne measurements will have Gaussian statistics for q_m and p_m , resulting in the same statistics for q_B and p_B . One can write the following equivalence

$$|\text{TMSVS}\rangle^{\otimes N} \equiv \sum_{i=1}^N p_{Q,P} |q_i + jp_i\rangle \langle q_i + jp_i| \otimes |\alpha\rangle_i, \quad (2.84)$$

where $p_{Q,P}$ is the joint bi-variate Gaussian distribution from which Alice draws her q 's and p 's at the same time. $|q_i + jp_i\rangle$ denotes two classical registers, and $|\alpha\rangle_i$ is the quantum coherent state prepared by Alice according to the values from her registers.

The transmission of one of the TMSVS modes to Bob through a noisy, lossy channel transforms Alice and Bob's covariance matrix in the following way. The total covariance matrix taking channel noise into account is

$$\Sigma_2^{\text{EB}} = \Sigma_1^{\text{EB}} \oplus \Sigma_{\text{noise}} = \begin{pmatrix} \frac{V+1}{2}\mathbb{1}_2 & \frac{V-1}{2}\mathbb{1}_2 & \sqrt{\frac{1}{2}(V^2-1)}\sigma_z & 0 \\ \frac{V-1}{2}\mathbb{1}_2 & \frac{V+1}{2}\mathbb{1}_2 & \sqrt{\frac{1}{2}(V^2-1)}\sigma_z & 0 \\ \sqrt{\frac{1}{2}(V^2-1)}\sigma_z & \sqrt{\frac{1}{2}(V^2-1)}\sigma_z & V\mathbb{1}_2 & 0 \\ 0 & 0 & 0 & N\mathbb{1}_2 \end{pmatrix}, \quad (2.85)$$

where we have assumed the noisy input has equal variance N in both the quadratures. The lossy channel can be modeled with a beamsplitter of transmittance T , and since its effect is only on Bob's mode,

$$\text{BS}_{\text{eff}} = \mathbb{1}_A \oplus \begin{pmatrix} \sqrt{T}\mathbb{1}_2 & \sqrt{1-T}\mathbb{1}_2 \\ -\sqrt{1-T}\mathbb{1}_2 & \sqrt{T}\mathbb{1}_2 \end{pmatrix} = \begin{pmatrix} \mathbb{1}_2 & 0 & 0 & 0 \\ 0 & \mathbb{1}_2 & 0 & 0 \\ 0 & 0 & \sqrt{T}\mathbb{1}_2 & \sqrt{1-T}\mathbb{1}_2 \\ 0 & 0 & -\sqrt{1-T}\mathbb{1}_2 & \sqrt{T}\mathbb{1}_2 \end{pmatrix}. \quad (2.86)$$

Therefore, the covariance matrix prior to Bob's measurement stands as

$$\Sigma_3^{\text{EB}}(\rho^{A_1 A_2 B N}) = \text{BS}_{\text{eff}} \Sigma_2^{\text{EB}} \text{BS}_{\text{eff}}^T = \begin{pmatrix} \frac{V+1}{2}\mathbb{1}_2 & \frac{V-1}{2}\mathbb{1}_2 & \sqrt{\frac{T}{2}(V^2-1)}\sigma_z & -\sqrt{\frac{1-T}{2}(V^2-1)}\sigma_z \\ \frac{V-1}{2}\mathbb{1}_2 & \frac{V+1}{2}\mathbb{1}_2 & \sqrt{\frac{T}{2}(V^2-1)}\sigma_z & -\sqrt{\frac{1-T}{2}(V^2-1)}\sigma_z \\ \sqrt{\frac{T}{2}(V^2-1)}\sigma_z & \sqrt{\frac{T}{2}(V^2-1)}\sigma_z & ([1-T]N + TV)\mathbb{1}_2 & \sqrt{T(1-T)}[N-V]\mathbb{1}_2 \\ -\sqrt{\frac{1-T}{2}(V^2-1)}\sigma_z & -\sqrt{\frac{1-T}{2}(V^2-1)}\sigma_z & \sqrt{T(1-T)}[N-V]\mathbb{1}_2 & ([1-T]N + TV)\mathbb{1}_2 \end{pmatrix}. \quad (2.87)$$

Considering only Alice and Bob's modes (i.e. $\text{Tr}_N(\rho^{A_1 A_2 B N})$),

$$\Sigma_3^{\text{EB}}(\rho^{A_1 A_2 B}) = \begin{pmatrix} \frac{V+1}{2}\mathbb{1}_2 & \frac{V-1}{2}\mathbb{1}_2 & \sqrt{\frac{T}{2}(V^2-1)}\sigma_z \\ \frac{V-1}{2}\mathbb{1}_2 & \frac{V+1}{2}\mathbb{1}_2 & \sqrt{\frac{T}{2}(V^2-1)}\sigma_z \\ \sqrt{\frac{T}{2}(V^2-1)}\sigma_z & \sqrt{\frac{T}{2}(V^2-1)}\sigma_z & ([1-T]N + TV)\mathbb{1}_2 \end{pmatrix}. \quad (2.88)$$

At this point, it is instructive to see the equivalence between the prepare and measure (P & M) version and the entanglement based (EB) version of the protocol. In practice, Alice and Bob do not share a two-mode squeezed vacuum states (TMSVS). Rather, Alice prepares an ensemble of coherent states

$$\rho = \sum_{i=1}^N p_{Q,P} |q_i + jp_i\rangle \langle q_i + jp_i| \otimes |\alpha\rangle_i. \quad (2.89)$$

with quadrature components q and p according to a bi-variate Gaussian distribution with zero-mean and variance $\tilde{V}_{\text{mod}}$. Since Alice knows her distribution in arbitrary detail, she is not limited by the shot-noise. Therefore, the variance of the quadrature operators $\hat{q}_A$ and $\hat{p}_A$ associated with Alice's state ensemble is

$$V_A = V(\hat{q}_A) = V(\hat{p}_A) = V_{\text{mod}} = 4\tilde{V}_{\text{mod}}. \quad (2.90)$$

Taking into account the inherent fundamental noise associated with the quadrature measurements of a coherent state, as well as excess noise and imperfect transmission of the channel, Bob's measurement statistics will have the variance

$$\hat{q}_B \sim \sqrt{T}\hat{q}_A + \mathcal{N}(0, 1 + \xi) \Rightarrow V(\hat{q}_B) = TV(\hat{q}_A) + 1 + \xi = TV_{\text{mod}} + 1 + \xi = V(\hat{p}_B), \quad (2.91)$$

³ where the variance of the statistically independent normal distribution $\mathcal{N}(0, 1 + \xi)$ corresponds to 1-shot noise unit and excess noise ξ . The covariance of Alice's and Bob's quadrature operators reads

$$\langle \hat{q}_A \hat{q}_B \rangle = \langle (\sqrt{T}\hat{q}_A + \mathcal{N})\hat{q}_A \rangle = \sqrt{T} \langle \hat{q}_A^2 \rangle + \underbrace{\langle \mathcal{N}\hat{q}_A \rangle}_{=0} = \sqrt{T}V_{\text{mod}} = \langle \hat{p}_A \hat{p}_B \rangle. \quad (2.92)$$

Therefore, the final covariance matrix, including the channel transmission T and excess noise ξ for the P&M version, is

$$\Sigma^{\text{PM}} = \begin{pmatrix} V_A \mathbb{1}_2 & \text{Cov}(A, B) \mathbb{1}_2 \\ \text{Cov}(A, B) \mathbb{1}_2 & V_B \mathbb{1}_2 \end{pmatrix} = \begin{pmatrix} V_{\text{mod}} \mathbb{1}_2 & \sqrt{T}V_{\text{mod}} \mathbb{1}_2 \\ \sqrt{T}V_{\text{mod}} \mathbb{1}_2 & (TV_{\text{mod}} + 1 + \xi) \mathbb{1}_2 \end{pmatrix}. \quad (2.93)$$

If Bob decides to measure both the quadratures of his mode, he applies heterodyne detection. In this case the covariance matrix, immediately before his detectors click, reads

$$\Sigma^{\text{PM}} = \begin{pmatrix} V_{\text{mod}} \mathbb{1}_2 & \sqrt{\frac{T}{2}} V_{\text{mod}} \mathbb{1}_2 & -\sqrt{\frac{T}{2}} V_{\text{mod}} \mathbb{1}_2 \\ \sqrt{\frac{T}{2}} V_{\text{mod}} \mathbb{1}_2 & [\frac{TV_{\text{mod}}}{2} + 1 + \frac{\xi}{2}] \mathbb{1}_2 & -\frac{1}{2} [TV_{\text{mod}} + \xi] \mathbb{1}_2 \\ -\sqrt{\frac{T}{2}} V_{\text{mod}} \mathbb{1}_2 & -\frac{1}{2} [TV_{\text{mod}} + \xi] \mathbb{1}_2 & [\frac{TV_{\text{mod}}}{2} + 1 + \frac{\xi}{2}] \mathbb{1}_2 \end{pmatrix}. \quad (2.94)$$

This P&M covariance matrix contains the correlation between Alice's known *classical* probability distribution of the quadrature components, and *quantum* experimental statistics generated by Bob upon measuring the quadrature components of the state ensemble (of which he does not have any knowledge) sent by Alice.

It is possible to establish an equivalence between the covariance matrices derived from the EB picture with the ones from the P&M picture if Alice rescales

³If $Z \mapsto aX + Y$, then $V(Z) = a^2V(X) + V(Y)$, where X, Y, Z are random variables.

her measurement operators in the following way

$$\hat{q}_A^{\text{EB}} \rightarrow \sqrt{\frac{V+1}{2(V-1)}} \hat{q}_A^{\text{PM}}, \quad (2.95a)$$

$$\hat{p}_A^{\text{EB}} \rightarrow -\sqrt{\frac{V+1}{2(V-1)}} \hat{p}_A^{\text{PM}}. \quad (2.95b)$$

Considering the case for heterodyne measurement by Bob, one can easily verify that

$$V(\hat{q}_A^{\text{EB}}) = \frac{V+1}{2(V-1)} V(\hat{q}_A^{\text{PM}}) = \frac{V+1}{2(V-1)} (V-1) = \frac{V+1}{2} = V(\hat{p}_A^{\text{EB}}), \quad (2.96a)$$

$$\begin{aligned} \langle \hat{q}_A^{\text{EB}} \hat{q}_B^{\text{EB}} \rangle &= \sqrt{\frac{V+1}{2(V-1)}} \langle \hat{q}_A^{\text{PM}} \hat{q}_B^{\text{PM}} \rangle = \sqrt{\frac{V+1}{2(V-1)}} \cdot \sqrt{\frac{T}{2}} (V-1) = \frac{1}{2} \sqrt{T(V^2-1)} = \\ &= -\langle \hat{p}_A^{\text{EB}} \hat{p}_B^{\text{EB}} \rangle, \end{aligned} \quad (2.96b)$$

where we have used $V_{\text{mod}} = V-1$. Therefore, with the aforementioned scaling factors Alice can convert her P&M covariance matrix to EB covariance matrix. The equivalence in the homodyne case can be similarly shown.

The P&M covariance matrix corresponds to the experimentally realistic scenario. On the other hand, the EB covariance matrix is more useful in the security analysis (for example calculating the Holevo bound- the maximum information Eve can extract by intercepting Alice's bound). Therefore switching back and forth between these two matrices is a crucial procedure.

2.6.1 Discrete Modulation

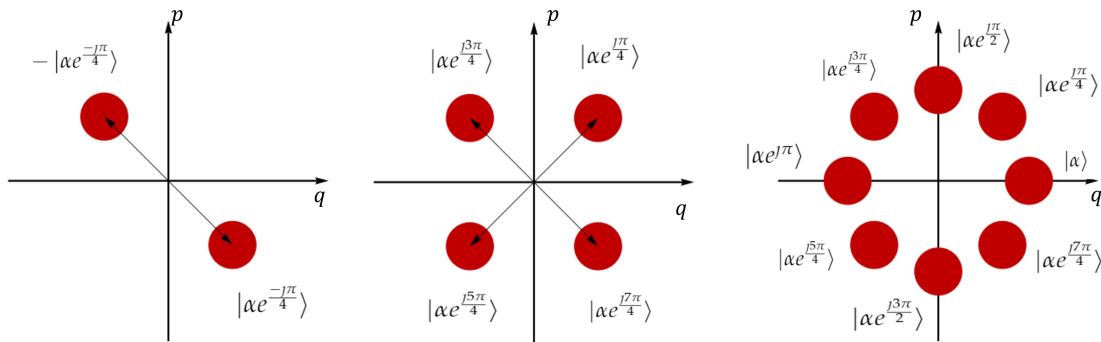


FIGURE 2.4: (a) (b) (c) Encoding schemes used for the two-state, four-state, and eight-state protocol respectively.

Gaussian modulation maximizes the mutual classical correlation Alice can establish with Bob via a quantum channel. The security analysis of implementing CV QKD with Gaussian modulation is also well-developed. But, in real world experiments, it is far more convenient to realize CV QKD with discrete modulation schemes. Figure (2.4) shows the constellation diagrams for three emblematic discrete modulation schemes: 2-PSK, 4-PSK, and 8-PSK respectively,

where PSK stands for phase shift keying. In a phase shift keying modulation scheme, the modulated signals all have the same magnitude, therefore, lie on the same circle on the phasor diagram. But their angular separation is given by $\frac{\pi}{N}$, where $N = 2, 4, 8$ for the two-state (2-PSK), four-state (4-PSK), and the eight-state (8-PSK) protocols, respectively.

The covariance matrix associated with the two-state protocol is given by

$$\Sigma_{2\text{-PSK}} = \begin{pmatrix} (V_A + 1)\mathbb{1}_2 & V_A \frac{1+e^{2V_A}}{\sqrt{1-e^{-2V_A}}} \sigma_z \\ V_A \frac{1+e^{2V_A}}{\sqrt{1-e^{-2V_A}}} \sigma_z & (V_A + 1)\mathbb{1}_2 \end{pmatrix}, \quad (2.97)$$

where V_A stands for Alice's modulation variance. Appendix C contains the relevant derivation. In a similar fashion, the characteristic covariance matrix for 4-PSK [61], and 8-PSK [62] modulation can be deduced. For 4-PSK,

$$\Sigma_{4\text{-PSK}} = \begin{pmatrix} (V_A + 1)\mathbb{1}_2 & Z_4 \sigma_z \\ Z_4 \sigma_z & (V_A + 1)\mathbb{1}_2 \end{pmatrix}, \quad (2.98)$$

where

$$Z_4 = 2\alpha^2 \sum_{k=0}^3 \frac{\lambda_{k-1}^{3/2}}{\lambda_k^{1/2}} = V_A \sum_{k=0}^3 \frac{\lambda_{k-1}^{3/2}}{\lambda_k^{1/2}} \quad (2.99)$$

with k being cyclic: $-k = -k + 4$. Here

$$\lambda_{0,2} = \frac{1}{2} e^{-\alpha^2} (\cosh \alpha^2 \pm \cos \alpha^2), \quad (2.100a)$$

$$\lambda_{1,3} = \frac{1}{2} e^{-\alpha^2} (\sinh \alpha^2 \pm \sin \alpha^2). \quad (2.100b)$$

For 8-PSK

$$\Sigma_{8\text{-PSK}} = \begin{pmatrix} (V_A + 1)\mathbb{1}_2 & Z_8 \sigma_z \\ Z_8 \sigma_z & (V_A + 1)\mathbb{1}_2 \end{pmatrix}, \quad (2.101)$$

where

$$Z_8 = V_A \sum_{k=0}^7 \frac{\lambda_{k-1}^{3/2}}{\lambda_k^{1/2}} \quad (2.102)$$

with k being cyclic: $-k = -k + 8$. Here

$$\lambda_{0(4)} = \frac{1}{4}e^{-\alpha^2}[\cosh \alpha^2 + \cos \alpha^2 \pm 2 \cos \frac{\alpha^2}{\sqrt{2}} \cosh \frac{\alpha^2}{\sqrt{2}}], \quad (2.103a)$$

$$\lambda_{1(5)} = \frac{1}{4}e^{-\alpha^2}[\sinh \alpha^2 + \sin \alpha^2 \pm \sqrt{2} \cos \frac{\alpha^2}{\sqrt{2}} \sinh \frac{\alpha^2}{\sqrt{2}} \pm \sqrt{2} \sin \frac{\alpha^2}{\sqrt{2}} \cosh \frac{\alpha^2}{\sqrt{2}}], \quad (2.103b)$$

$$\lambda_{2(6)} = \frac{1}{4}e^{-\alpha^2}[\cosh \alpha^2 - \cos \alpha^2 \pm 2 \sin \frac{\alpha^2}{\sqrt{2}} \sinh \frac{\alpha^2}{\sqrt{2}}], \quad (2.103c)$$

$$\lambda_{3(7)} = \frac{1}{4}e^{-\alpha^2}[\sinh \alpha^2 - \sin \alpha^2 \mp \sqrt{2} \cos \frac{\alpha^2}{\sqrt{2}} \sinh \frac{\alpha^2}{\sqrt{2}} \pm \sqrt{2} \sin \frac{\alpha^2}{\sqrt{2}} \cosh \frac{\alpha^2}{\sqrt{2}}]. \quad (2.103d)$$

One noteworthy thing is that the covariance matrices Σ_G , Σ_2 , Σ_4 , or Σ_8 only differ in their off-diagonal (covariance) terms. Figure (2.5) plots how the covariance terms for each different protocol varies with Alice's modulation variance. Moreover, Σ_2 , Σ_4 , or Σ_8 transform analogously as Σ_G under the action of a channel with transmittance T and excess noise ξ .

To conclude, in this subsection, we have seen the covariance matrices for the 2-PSK, 4-PSK (QPSK), and 8-PSK protocols in equations (2.97), (2.98), and (2.101) respectively, which will be used in their respective security analysis.

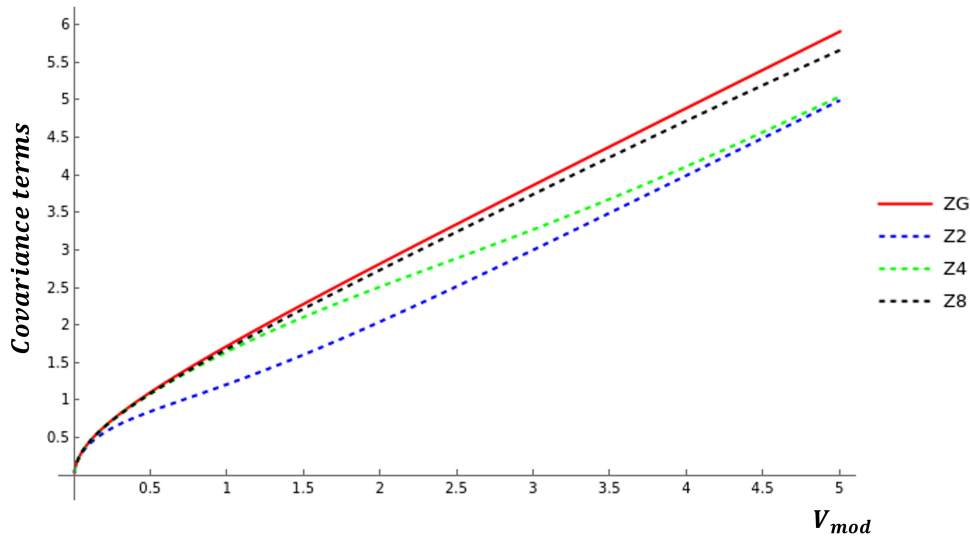


FIGURE 2.5: Comparison between the correlations: Z_G for Gaussian modulation (solid red), Z_2 for 2-state modulation (dotted blue), Z_4 for 4-state modulation (dotted green), Z_8 for 8-state modulation (dotted black). It is to be noted that for variance $V_{\text{mod}} \leq 1$, Z_G and Z_8 are almost indistinguishable, meaning that in this regime, one has the Holevo bound $S_8(B, E) \approx S_G(B, E)$. Therefore, the security bounds for the Gaussian modulation can be applied here as well.

2.7 Transmission and Noise

The performance of a CV-QKD protocol depends on, apart from the modulation alphabet, the particular setup (for example fiber based bulk optics, fiber based integrated optics, free space optics) used to implement the protocol. A quantum channel is characterized by its transmittance T and excess noise ξ . In the entanglement-based covariance matrix analysis, we have seen Bob receives coherent states with variance

$$V_B = TV + (1 - T)N, \quad (2.104)$$

where we have simulated the channel transmittance with a beam splitter of transmittance T which mixes one half of the TMSVS meant for Bob with input noise N . Now, by logic, the states received by Bob should have the following variance

$$V_B = TV_{\text{mod}} + 1 + \xi, \quad (2.105)$$

where V_{mod} is Alice's modulation variance, 1 is due to inherent shot noise, ξ refers to the additional contribution to the noise variance on top of the inherent vacuum noise (hence the name *excess noise*). Here we have referred the noise variance to Bob's side where measurement is performed. The excess noise may have origins both on Alice and on Bob's side

$$\xi = T\xi_A + \xi_B. \quad (2.106)$$

A detailed analysis of excess noise and their quantitative models for the fiber-based bulk optics implementation has been presented in [54]. Some of the results are relevant to our case. From equations (2.104) and (2.105), one can define

$$N = 1 + \frac{\xi}{1 - T}. \quad (2.107)$$

The channel transmittance, detection efficiency, and the coupling losses all can be lumped together in the term T

$$T = T_{\text{ch}} \cdot \eta_{\text{coup}} \cdot \eta_{\text{det}}. \quad (2.108)$$

Inserting the expression for N in equation (2.88)

$$\Sigma_3^{\text{EB}} = \begin{pmatrix} V\mathbb{1}_2 & \sqrt{T}\sqrt{V^2 - 1}\sigma_z \\ \sqrt{T}\sqrt{V^2 - 1}\sigma_z & (TV_{\text{mod}} + 1 + \xi)\mathbb{1}_2 \end{pmatrix}, \quad (2.109)$$

where, for simplicity, we have considered Alice's mode as a whole. The transformation of discrete modulation covariance matrices under the action of channel with transmittance T and excess noise ξ also have similar form

$$\Sigma_{2,4,8\text{-PSK}} = \begin{pmatrix} (V_A + 1)\mathbb{1}_2 & \sqrt{T}Z_{(2,4,8)}\sigma_z \\ \sqrt{T}Z_{(2,4,8)}\sigma_z & (TV_A + 1 + \xi)\mathbb{1}_2 \end{pmatrix}. \quad (2.110)$$

Equation (2.109) for Gaussian modulation, and Equation (2.110) for discrete modulations represent an intermediary step of the CV QKD protocol. In these equations, we have taken into account the effect of channel transmittance and excess noise, but the effect of Bob's measurement process is still to come.

2.8 State Measurement

Bob receives noisy coherent states with randomly modulated quadrature components q and p . He can either employ *homodyne detection*, or *heterodyne detection* to measure the quadratures. With homodyne detection technique the signal mode is mixed with a reference laser, the local oscillator(LO) at a balanced beamsplitter. Bob will measure only one of the components by randomly varying the relative phase θ of the signal mode and the LO. The photon number difference at the output ports of the BS is proportional to either q quadrature for $\theta = 0$ or p quadrature for $\theta = \pi/2$

$$\Delta\hat{n} = |\alpha_{LO}|(\hat{q} \cos \theta + \hat{p} \sin \theta). \quad (2.111)$$

With heterodyne detection, Bob can measure both quadratures of each mode simultaneously. In this approach, Bob will first separate the incoming states with a balanced BS, and will subsequently measure the two quadratures by homodyne detection on each half of the signal.

Let us see the transformation on the EB covariance matrix because of Bob's heterodyning. The input to the balanced beam splitter is a vacuum mode in addition to the Bob's mode

$$\Sigma_4^{\text{EB}} = \Sigma_3^{\text{EB}} \oplus \Sigma_{\text{vac}} = \begin{pmatrix} \frac{V+1}{2}\mathbb{1}_2 & \frac{V-1}{2}\mathbb{1}_2 & \sqrt{\frac{T}{2}}\sqrt{V^2-1}\sigma_z & 0 \\ \frac{V-1}{2}\mathbb{1}_2 & \frac{V+1}{2}\mathbb{1}_2 & \sqrt{\frac{T}{2}}\sqrt{V^2-1}\sigma_z & 0 \\ \sqrt{\frac{T}{2}}\sqrt{V^2-1}\sigma_z & \sqrt{\frac{T}{2}}\sqrt{V^2-1}\sigma_z & (T[V-1] + 1 + \xi)\mathbb{1}_2 & 0 \\ 0 & 0 & 0 & \mathbb{1}_2 \end{pmatrix}. \quad (2.112)$$

Since the balanced beamsplitter leaves Alice's mode unchanged,

$$\text{BS}_{\text{eff}} = \mathbb{1}_A \oplus \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1}_2 & \mathbb{1}_2 \\ -\mathbb{1}_2 & \mathbb{1}_2 \end{pmatrix}_{B, \text{vac}} = \begin{pmatrix} \mathbb{1}_2 & 0 & 0 & 0 \\ 0 & \mathbb{1}_2 & 0 & 0 \\ 0 & 0 & \mathbb{1}_2/\sqrt{2} & \mathbb{1}_2/\sqrt{2} \\ 0 & 0 & -\mathbb{1}_2/\sqrt{2} & \mathbb{1}_2/\sqrt{2} \end{pmatrix}. \quad (2.113)$$

The action of the BS operator on the total state, including Bob's additional mode, yields

$$\Sigma_5^{\text{EB}}(\rho^{A_1 A_2 B_1 B_2}) = \text{BS}_{\text{eff}} \Sigma_4^{\text{EB}} \text{BS}_{\text{eff}}^T = \quad (2.114)$$

$$\begin{pmatrix} \frac{V+1}{2}\mathbb{1}_2 & \frac{V-1}{2}\mathbb{1}_2 & \frac{1}{2}\sqrt{T(V^2-1)}\sigma_z & -\frac{1}{2}\sqrt{T(V^2-1)}\sigma_z \\ \frac{V-1}{2}\mathbb{1}_2 & \frac{V+1}{2}\mathbb{1}_2 & \frac{1}{2}\sqrt{T(V^2-1)}\sigma_z & -\frac{1}{2}\sqrt{T(V^2-1)}\sigma_z \\ \frac{1}{2}\sqrt{T(V^2-1)}\sigma_z & \frac{1}{2}\sqrt{T(V^2-1)}\sigma_z & [\frac{T(V-1)}{2} + 1 + \frac{\xi}{2}]\mathbb{1}_2 & -\frac{1}{2}[T(V-1) + \xi]\mathbb{1}_2 \\ -\frac{1}{2}\sqrt{T(V^2-1)}\sigma_z & -\frac{1}{2}\sqrt{T(V^2-1)}\sigma_z & -\frac{1}{2}[T(V-1) + \xi]\mathbb{1}_2 & [\frac{T(V-1)}{2} + 1 + \frac{\xi}{2}]\mathbb{1}_2 \end{pmatrix}.$$

Bob now has two modes B_1 and B_2 , since he split his signal into half. Considering the covariance matrix between one of Alice's mode and one of Bob's modes

only

$$\Sigma_5^{\text{EB}}(\rho^{A_1, B_1}) = \begin{pmatrix} \frac{V+1}{2} \mathbb{1}_2 & \pm \frac{1}{2} \sqrt{T(V^2-1)} \sigma_z \\ \pm \frac{1}{2} \sqrt{T(V^2-1)} \sigma_z & [\frac{T}{2}(V-1) + 1 + \frac{\xi}{2}] \mathbb{1}_2 \end{pmatrix}, \quad (2.115)$$

which is equivalent to the case of homodyne detection up to a factor of 1/2 at transmission T and excess noise ξ . The covariance matrices for discrete modulation look similar with the off-diagonal terms properly modified.

2.9 Estimating the secret key rate

From [54], for a Gaussian CV-QKD protocol, the asymptotic secret key rate against collective attacks, is lower bounded by ⁴

$$K_{\text{coll}}^{\text{asym}} \geq f_{\text{sym}}(1 - \text{FER})(1 - \nu)(\beta I_{AB} - \chi_{EB}) \quad (2.116)$$

Here, f_{sym} represents the symbol rate: the rate at which optical pulses are sent to Bob from Alice. $\text{FER} \in [0, 1]$ is the frame error rate: often in serial communications message bits are grouped together with additional bits (guard bits, start and end bits) for reliable communication. One such group is called a frame. $\nu \in [0, 1]$ is the fraction of raw key disclosed for parameter estimation. $\beta \in [0, 1]$ is the efficiency of reconciliation at the post-processing step. Modern techniques usually allow us to get as much as $\beta = 0.95$, or more. I_{AB} corresponds to the mutual information between Alice and Bob. According to Shannon's classical information theory, the quantity mutual information is bounded by the channel capacity. For an AWGN channel, the mutual information is given by

$$I_{AB} = \frac{\mu}{2} \log_2(1 + \text{SNR}), \quad (2.117)$$

where SNR denotes signal to noise ratio. The above equation needs to be modified with a factor of μ in the case of coherent optical communication in order to indicate homodyne ($\mu = 1$) or heterodyne ($\mu = 2$) detection. Finally, χ_{EB} refers to the Holevo information: an upper bound on the mutual information between Alice and Bob during reverse reconciliation process. This Holevo upper bound puts a lower bound on the secret key rate. In the subsequent sections, we shall see how the covariance matrix elements can be used in order to calculate various terms in the secret key formulae.

2.10 Signal to Noise Ratio and Mutual Information

The signal-to-noise ratio (SNR) is given by

$$\text{SNR} = \frac{P_S}{P_N}, \quad (2.118)$$

⁴Appendix A provides the tools from quantum Shannon theory to appreciate how one arrives at this formula.

where P_S is the total signal power and P_N is the total noise power arriving at the channel output. The equal variance in Bob's both quadratures is

$$V_B = V(\hat{q}_B) = V(\hat{p}_B) = \underbrace{\frac{T}{\mu} V_{\text{mod}}}_{\text{signal}} + 1 + \underbrace{\frac{\xi}{\mu}}_{\text{noise}}, \quad (2.119)$$

where T is the total transmittance (including coupling, and detection efficiency) and $\mu = 1$ in case of homodyne detection and $\mu = 2$ in case of heterodyne detection. Since a zero mean random variable's power is equal to its variance, hence

$$\text{SNR} = \frac{\frac{1}{\mu} T V_{\text{mod}}}{1 + \frac{1}{\mu} \xi}. \quad (2.120)$$

The mutual information between Alice and Bob, then is given by equation (2.117)

$$I_{AB} = \frac{\mu}{2} \log_2(1 + \text{SNR}) = \frac{\mu}{2} \log_2 \left(1 + \frac{\frac{1}{\mu} T V_{\text{mod}}}{1 + \frac{1}{\mu} \xi} \right). \quad (2.121)$$

The above formula gives one component of the secret key rate, the mutual information, in terms of Alice's modulation variance and channel parameters. But we also need to be able to quantify Eve's or the adversary's maximum knowledge to establish the key rate

2.11 Calculation of Holevo Information

The Holevo information χ_{EB} is the upper bound on an eavesdropper's information. Under reverse reconciliation, this quantity is given by

$$\chi_{EB} = S_E - S_{E|B}, \quad (2.122)$$

where S_E is the von Neumann entropy (ignorance) of the state accessible to Eve for collective measurement, and $S_{E|B}$ is the von Neumann entropy (remaining ignorance) of the same state after a projective (homodyne or heterodyne) measurement has been performed by Bob. The entropies S_E and $S_{E|B}$ are determined by the symplectic eigenvalues of the covariance matrices describing these states. The actual quantum state from which S_E and $S_{E|B}$ are derived depends on what kind of eavesdropping attack is assumed. For a simplified security analysis, it is very convenient to assume that Eve holds a purification⁵ of Alice's and Bob's mutual state ρ_{AB} . Then, by Uhlmann's theorem, for any two ancillary states ρ_{E1} and ρ_{E2} , which purify ρ_{AB} such that ρ_{ABE1} and ρ_{ABE2} are both pure, there exists a unitary transformation U_{E1} acting on ρ_{E1} which transforms ρ_{ABE1} into ρ_{ABE2}

$$(\mathbb{1}_2 \otimes U_{E1}) \rho_{ABE1} (\mathbb{1}_2 \otimes U_{E1}^\dagger) = \rho_{ABE2}. \quad (2.123)$$

⁵A pure state $|\Psi\rangle_{AB}$ is a purification of $\rho_A = \sum_i p_i |\psi_i\rangle \langle \psi_i|$, if for some system B with basis $\{|b\rangle_i\}$ and dimension $d_B = d_A$, $|\Psi\rangle_{AB} = \sum_i \sqrt{p_i} |\psi_i\rangle |b_i\rangle$ from which it follows that $\text{Tr}_B |\Psi\rangle \langle \Psi| = \rho_A$.

Since the von Neumann entropy is invariant under unitary transformations

$$S(\rho) = S(U\rho U^\dagger), \quad (2.124)$$

any purification of ρ_{AB} that Eve may possess will result in the same entropy and hence the same Holevo information χ_{EB} .

2.11.1 Universal Analysis of Purification Attacks

Let us assume Eve is holding a purification of Alice's and Bob's shared quantum state ρ_{AB} , which allows us to write the overall state ρ_{ABE} as

$$\rho_{ABE} = |\psi\rangle\langle\psi|. \quad (2.125)$$

Therefore, tracing out any one part will result in the corresponding mixed state

$$\text{Tr}_E(\rho_{ABE}) = \rho_{AB}, \quad (2.126a)$$

$$\text{Tr}_{AB}(\rho_{ABE}) = \rho_E. \quad (2.126b)$$

The von Neumann entropy of a mixed state $\rho = \sum_i p_i |i\rangle\langle i|$ is defined as

$$S = -\sum_i p_i \log_2(p_i). \quad (2.127)$$

Let us consider the following Schmidt decomposition of $|\psi\rangle$

$$|\psi\rangle = \sum_i \sqrt{\lambda_i} |i\rangle_{AB} \otimes |i\rangle_E. \quad (2.128)$$

Tracing out either of the subsystems then yields

$$\text{Tr}_E(\rho_{ABE}) = \rho_{AB} = \sum_i \lambda_i |i\rangle_{AB} \langle i|_{AB}, \quad (2.129a)$$

$$\text{Tr}_{AB}(\rho_{ABE}) = \rho_E = \sum_i \lambda_i |i\rangle_E \langle i|_E. \quad (2.129b)$$

The von Neumann entropy depends only on the Schmidt coefficients $\sum_i \lambda_i \log_2(\frac{1}{\lambda_i})$. From the equation above, λ_i 's are the same for both ρ_{AB} , and ρ_E . Therefore, the assumption that Eve holds a purification of ρ_{AB} leads to

$$S_E = S_{AB} = -\sum_i \lambda_i \log_2 \lambda_i, \quad (2.130)$$

which drastically simplifies the computation of the Holevo information

$$\chi_{EB} = S_E - S_{E|B} = S_{AB} - S_{A|B}. \quad (2.131)$$

The entropy S_{AB} can be obtained using Alice's and Bob's covariance matrix and its symplectic eigenvalues. The symplectic analysis yielded the following covariance matrix showing relation between Alice and Bob's states

$$\Sigma_{AB}^{\text{EB}} = \begin{pmatrix} V\mathbb{1}_2 & \sqrt{T(V^2-1)}\sigma_z \\ \sqrt{T(V^2-1)}\sigma_z & (T[V-1]+1+\xi)\mathbb{1}_2 \end{pmatrix} =: \begin{pmatrix} a\mathbb{1}_2 & c\sigma_z \\ c\sigma_z & b\mathbb{1}_2 \end{pmatrix}, \quad (2.132)$$

where $a, b, c \in \mathbb{R}$. It has been shown that for a covariance matrix of the form in equation (2.132), the symplectic eigenvalues are given by [63]

$$\nu_{1,2} = \frac{1}{2}[y \pm (b - a)] \quad (2.133)$$

with

$$y = \sqrt{(a + b)^2 - 4c^2}. \quad (2.134)$$

Therefore, $S_{AB} = \sum_{k=1}^2 g(\nu_k) = g(\nu_1) + g(\nu_2)$, from equation (2.66).

The computation of $S_{A|B}$ depends on whether Bob performs a homodyne or heterodyne measurement.

Homodyne Detection

A homodyne measurement by Bob will transform Alice's state as follows

$$\Sigma_{A|B} = \Sigma_A - \frac{1}{V(q_B)} \Sigma_C \Pi_q \Sigma_C^T, \quad (2.135)$$

where $\Sigma_A = a\mathbb{1}_2$, $\Sigma_C = c\sigma_z$, and $V(q_B) = V_B = b$. Therefore

$$\Sigma_{A|B} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} - \frac{1}{b} \begin{pmatrix} c^2 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a - \frac{c^2}{b} & 0 \\ 0 & a \end{pmatrix}. \quad (2.136)$$

But the symplectic eigenvalues of $\Sigma_{A|B}$ are given by the modulus of the eigenvalues of the Hermitian matrix $\tilde{\Sigma}_{A|B} = J\Omega\Sigma_{A|B}$ [54], [55]. We obtain

$$\tilde{\Sigma}_{A|B} = J \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} a - \frac{c^2}{b} & 0 \\ 0 & a \end{pmatrix} = J \begin{pmatrix} 0 & a \\ -a + \frac{c^2}{b} & 0 \end{pmatrix}. \quad (2.137)$$

The eigenvalues ν_3 of this matrix are

$$\begin{aligned} \det(\tilde{\Sigma}_{A|B} - \nu_3 \mathbb{1}_2) &= \det \begin{pmatrix} -\nu_3 & ja \\ j(-a + \frac{c^2}{b}) & -\nu_3 \end{pmatrix} \\ &= \nu_3^2 + a(-a + \frac{c^2}{b}) = 0, \end{aligned} \quad (2.138)$$

and, therefore,

$$\nu_3 = \sqrt{a(a - \frac{c^2}{b})}. \quad (2.139)$$

Heterodyne Detection

In the case of heterodyne detection the relation describing the effect of a partial heterodyne measurement

$$\Sigma_{A|B} = \Sigma_A - \frac{1}{V(\hat{q}_B) + 1} \Sigma_C \Sigma_C^T, \quad (2.140)$$

where $\Sigma_A = a\mathbb{1}_2$, $\Sigma_C = c\sigma_z$, and $V(q_B) = V_B = b$. and a, b and c are defined as in (2.132). Therefore,

$$\Sigma_{A|B} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} - \frac{1}{b+1} \begin{pmatrix} c^2 & 0 \\ 0 & c^2 \end{pmatrix} = \left(a - \frac{c^2}{b+1}\right) \mathbb{1}_2. \quad (2.141)$$

The eigenvalue is computed to be

$$\nu_3 = a - \frac{c^2}{b+1}. \quad (2.142)$$

Therefore, Holevo information, or Eve's maximum accessible information is given by $\chi_{EB} = S_{AB} - S_{A|B} = g(\nu_1) + g(\nu_2) - g(\nu_3)$.

To conclude, in this section, we have seen how to calculate Eve's Holevo information from the final stage (following Bob's measurement) of the covariance matrix. The next section discusses how one can estimate the channel transmittance and the excess noise from the covariance matrix elements. Therefore, from equations (2.116) and (2.121), **one can ultimately establish the key rate a CV QKD protocol from the covariance matrix.**

2.12 Practical Parameter Estimation

Since all the practical parameters of interest such as Holevo information (χ), channel transmittance (T), channel noise (N), signal-to-noise ratio (SNR) can be calculated from the covariance matrix of the final stage, the knowledge of how the covariance matrix transforms at the various stages of the CV QKD protocol is of utmost importance. In the P&M scenario, the experimentally realistic case, the covariance matrix describing correlation between Alice's mode, and one of Bob's mode when he decides to perform heterodyne detection is

$$\Sigma^{\text{PM}} = \begin{pmatrix} V_{\text{mod}} \mathbb{1}_2 & \sqrt{\frac{T}{2}} V_{\text{mod}} \mathbb{1}_2 \\ \sqrt{\frac{T}{2}} V_{\text{mod}} \mathbb{1}_2 & [\frac{TV_{\text{mod}}}{2} + 1 + \frac{\xi}{2}] \mathbb{1}_2 \end{pmatrix} = \begin{pmatrix} a^{\text{PM}} \mathbb{1}_2 & c^{\text{PM}} \mathbb{1}_2 \\ c^{\text{PM}} \mathbb{1}_2 & b^{\text{PM}} \mathbb{1}_2 \end{pmatrix}. \quad (2.143)$$

On the other hand, at the same stage, the covariance matrix between one of Alice's modes and one of Bob's modes following the description of entanglement based version of the protocol is

$$\Sigma^{\text{EB}} = \begin{pmatrix} \frac{V+1}{2} \mathbb{1}_2 & \frac{1}{2} \sqrt{T(V^2-1)} \sigma_z \\ \frac{1}{2} \sqrt{T(V^2-1)} \sigma_z & [\frac{T(V-1)}{2} + 1 + \frac{\xi}{2}] \mathbb{1}_2 \end{pmatrix} = \begin{pmatrix} a^{\text{EB}} \mathbb{1}_2 & c^{\text{EB}} \mathbb{1}_2 \\ c^{\text{EB}} \mathbb{1}_2 & b^{\text{EB}} \mathbb{1}_2 \end{pmatrix} \quad (2.144)$$

As usual, $V = V_{\text{mod}} + 1$. By inspection, from equations (2.143) and (2.144),

$$a^{\text{EB}} = \frac{V+1}{2(V-1)} a^{\text{PM}}, \quad (2.145a)$$

$$c^{\text{EB}} = \sqrt{\frac{V+1}{2(V-1)}} c^{\text{PM}}. \quad (2.145b)$$

Alice and Bob can fill in for their a^{PM} and b^{PM} during the state preparation and measurement steps respectively. Calculating the covariance term (c^{PM}) requires them sharing a fraction of their samples n .

The Holevo information (χ) can be calculated using equations (2.145a) and (2.145b), as described in section (2.11). For transmittance T , from equation (2.145b), one finds

$$(c^{\text{EB}})^2 = \frac{V+1}{2(V-1)}(c^{\text{PM}})^2 \Rightarrow T = 2\left(\frac{c^{\text{PM}}}{V_{\text{mod}}}\right)^2. \quad (2.146)$$

The excess noise ξ can be calculated from

$$V_B = b^{\text{PM}} = \frac{TV_{\text{mod}}}{2} + 1 + \frac{\xi}{2} \Rightarrow \xi = 2b^{\text{PM}} + TV_{\text{mod}} + 2. \quad (2.147)$$

With transmittance T , and excess noise ξ , the SNR can easily be determined from equation (2.120), which, in turn, gives the mutual information according to equation (2.121). Therefore, once the covariance matrix in (2.143) gets populated with elements for a particular implementation of CV-QKD, one can ultimately establish the key rate for that scheme.

In an actual laboratory experiment, Bob reads his measurement outcomes on a Voltmeter in voltages, not in shot noise units. Therefore, he needs to figure out a conversion factor, such that

$$V(U) = \phi V(\hat{q}_B). \quad (2.148)$$

The voltage variance is given by the maximum likelihood estimator (MLE) of the U_i voltages measured during N samples

$$V(U) = \langle U^2 \rangle - \langle U \rangle^2 = \frac{1}{N} \sum_{i=1}^N U_i^2 - \left(\frac{1}{N} \sum_{i=1}^N U_i \right)^2. \quad (2.149)$$

The process of determining ϕ has been described in [54] for their setup, which may also be applicable in our case.

The parameter estimation procedure will allow Alice and Bob to ascertain if their mutual information is greater than Eve's intercepted information. If not, they abort the protocol. Otherwise, they proceed to implement the information reconciliation (error correction) and the privacy amplification steps. Following privacy amplification, Alice and Bob will be in possession of a shared secret key.

2.13 Security Aspect of CV QKD

The security proofs of CV QKD are comparatively harder since one needs to deal with ∞ -dimensional Hilbert space. The security analysis of Gaussian modulated CV QKD protocols is all but complete. On the other hand, the security of discretely-modulated CV QKD, which is more experiment-friendly, is still an area of active research. The security proofs can be sketched either in the *asymptotic limit* in which Alice and Bob use their quantum channel infinite number of

times, or for the *finite size* case (realistic) where the channel use is limited to finite number of times [54]. The asymptotic limit is not used to analyze any realistic system, but it can be derived relative easily. Moreover, it provides an upper bound on the achievable rate for the finite size case. The amount of security which can be extracted from a QKD protocol depends on the eavesdropping strategies of Eve. Eve's possible attacks can be classified in three different types [64], [54]:

- *Individual Attack*: Eve performs an independent and identically distributed (i.i.d.) attack on all signals., that is, she prepares separable ancilla states each of which interacts individually with one sent state in the quantum channel. The states are stored in a quantum memory before public reconciliation and subsequently measured independently from one another.
- *Collective Attack*: Alice performs the same i.i.d. attack with separable ancilla states, stores her states in a quantum memory but performs an optimal collective measurement at any time during Alice and Bob's reconciliation stage.
- *Coherent Attack*: In general, Eve is allowed to perform any unitary collective interaction over the sent states and any measurement strategy at any time she wants. This is the most general type of attack where no i.i.d. assumption is made. In particular, Eve may prepare an optimal global ancilla state whose (possibly mutually dependent) modes interact with the signal pulses and are then stored and collectively measured after the classical post processing.

The lower bound for the asymptotic secret key rate for a CV QKD system in the case of collective attacks is given by the Devetak-Winter formula [65]

$$r_{\text{coll}}^{\text{asympt}} \geq I_{AB} - \chi \quad (2.150)$$

χ is Holevo information which denotes the upper bound on Eve's knowledge, or the maximum information she can access from the quantum states shared between Alice and Bob. Depending on the error correction protocol, the representation of χ will vary. For example, in the case of direct reconciliation, when Bob corrects his data according to Alice's choice of quadratures during the public communication stage, then

$$r_{\text{coll, DR}}^{\text{asympt}} \geq I_{AB} - \chi_{EA}, \quad (2.151)$$

where χ_{EA} represents Eve's maximum information on Alice's raw key. On the other hand, when Alice modifies her data according to the information she receives from Bob during reverse reconciliation, the key rate is given by

$$r_{\text{coll, RR}}^{\text{asympt}} \geq I_{AB} - \chi_{EB}, \quad (2.152)$$

where χ_{EB} represents Eve's maximum information on Bob's raw key. In the finite size limit, it is not quite possible for Alice and Bob to perfectly extract all their mutual information and it is customary to replace I_{AB} with βI_{AB} where β is the reconciliation efficiency. The value of βI_{AB} can be observed during a

given protocol. Bounding the value of χ_{EB} involves optimization over a family of infinite-dimensional quantum channels. It has been shown that the optimal attack Eve can perform is Gaussian [64], [66], [67]. The security of Gaussian modulated CV QKD against coherent attacks (the strongest possible attack) has been aided by the fact that coherent states can be reduced to collective attacks using the de Finetti representation theorem for infinite dimensions [68]. While these proofs only hold in the asymptotic limit, the exploitation of a new kind of Gaussian de Finetti reduction has led to the establishment of security proofs against the general attacks in a realistic finite size regime [69]. With this work, the security analysis for Gaussian modulated CV QKD is almost complete.

In actual implementation, it is not possible to define infinite continuous constellation for Gaussian modulation because optical modulators have a finite precision and range. Therefore, discrete modulated CV QKD is the more realistic scenario and it has the best of appeals for large scale deployment of CV QKD using the existing telecom infrastructure. However, the security analysis for DM CV QKD is still not well established and currently an area of vital research. Numerical bounds have been provided in the case of simple phase shift keying (QPSK) modulations [70], [71]. The strategy in both consists in expressing the asymptotic secret key rate as a convex optimization problem, and more precisely a semidefinite program (SDP). The main difference between the two papers is that [71] considers a linear objective function, while [70] relies on a tighter nonlinear objective function. However, both the processes described in the two references are computationally expensive. A more recent publication purports to establish an analytical lower bound on the asymptotic secret key rate of continuous-variable quantum key distribution with an arbitrary (discrete) modulation of coherent states [72]. The final frontier in the security analysis of CV QKD, which is also the most important case realistically, is to establish a secret key rate for the discrete modulations in the finite size regime.

2.14 Conclusion

In this chapter, a generic theoretical analysis for continuous variable quantum key distribution, including the cases of Gaussian and discrete modulation, has been provided. We have seen that the concept of covariance matrix (defined with respect to a Gaussian quantum state) and its symplectic transformation are essential to the understanding of this kind of analysis. Two key takeaway messages from this chapter are the following. First, within a certain operation range, it is possible to simulate Gaussian modulated CV QKD with discrete modulation without sacrificing much in terms of performance. This is highly desirable, since discrete modulation is more implementation-friendly. Second, the secure key rate that Alice and Bob can establish depends directly upon the channel transmittance and channel noise. Therefore depending upon the realization platform (fiber or free space, bulk or integrated), the secure key rate will vary for the same modulation scheme. In the following chapter, we have proposed a design for a photonic integrated transceiver for the proof-of-principle demonstration of CV QKD. The integrated transmitter and the receiver circuits

take the role of Alice and Bob, respectively. A standard single mode optical fiber (SSMF), which connects these two circuits, works as the quantum channel for encoded key transmission.

Chapter 3

On Chip Implementation of CV QKD

3.1 Introduction

This chapter describes the thoughts and considerations which gave shape to the design of our photonic integrated circuit capable of implementing continuous variable quantum key distribution (CV QKD). As mentioned earlier, from the engineering point of view, implementing CV QKD with coherent states of light is akin to carrying out a coherent optical communication protocol. Therefore, a system-on-chip enabling CV QKD should have all the components required for coherent communication. In this optical communication scenario, Alice and Bob take the form of an optical transmitter, and optical receiver respectively. The encoded quantum states travel from the transmitter to the receiver via optical fiber (quantum channel). The material system with which we choose to implement our transmitter and receiver chips is InGaAsP/InP. The design is flexible in the sense that it enables the encoder to realize different modulation schemes with the same architecture. At first, we performed a simple pen and paper calculation to estimate the link budget - the minimum optical power that should leave the transmitter, with all the propagation and insertion losses taken into account, in order to excite the photodiodes at the receiver for producing a detectable photocurrent. This initial feasibility analysis permitted us to simulate the performance of the transceiver system with a circuit solver. The next step was to generate a mask layout¹ for the transceiver circuits which could be read by the foundry for the chip fabrication. Before submitting the GDSII files to the foundry, a design rule check (DRC) was done to ensure that the layouts are in line with the prescribed guidelines for the foundry specific technology.

3.2 High Level Overview

Figures (3.1) and (3.2) show the block diagrams of our proposed transmitter and receiver architecture to implement CV QKD. Alice generates TE-polarized coherent light pulses with a repetition rate (modulation frequency) of 350 MHz by

¹The universally accepted file format for mask layouts is GDSII, which stands for graphic design system

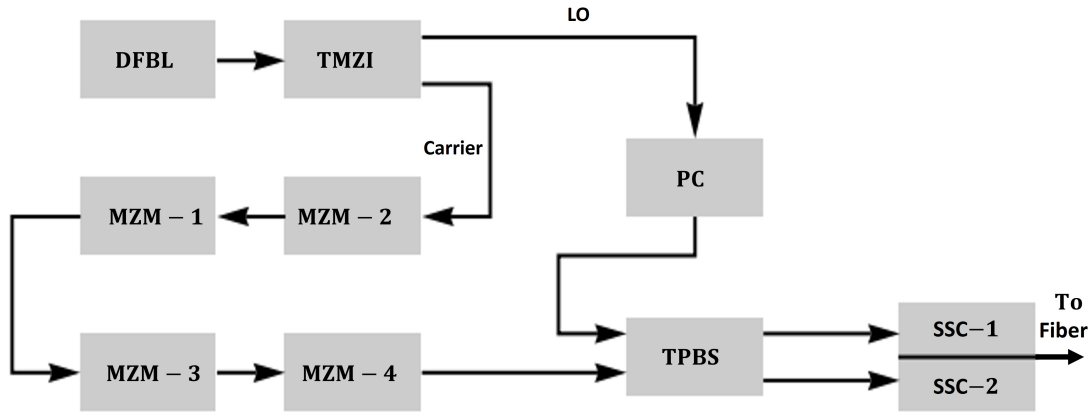


FIGURE 3.1: Block diagram of the transmitter architecture. The acronyms used in the diagram refer to DFBL = Distributed Feedback Laser; TMZI = Tunable Mach Zehnder Interferometer; MZM = Mach Zehnder Modulator; PC = Polarization Converter; TPBS = Tunable Polarization Beam Splitter; SSC = Spot Size Converter.

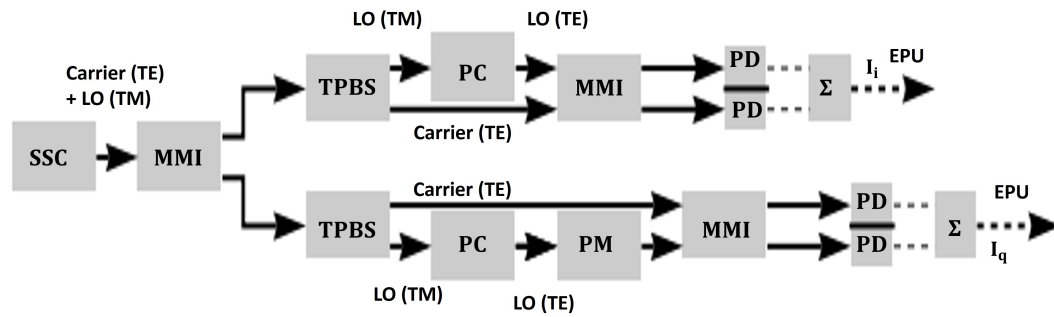


FIGURE 3.2: Block diagram of the receiver architecture. The acronyms used in the diagram refer to MMI = Multimode Interferometer; PM = Phase Modulator; PD = Photo Detector; EPU = Electronic Processing Unit. The Σ symbol refers to the electronic sum-mer/subtractor.

directly modulating a distributed feedback (DFB) laser diode output emitting at 1550 nanometer. The use of a pulsed laser facilitates time synchronization with the receiver. The pulse train is then fed into a highly asymmetric power splitter. It splits each pulse into a weak carrier and a strong local oscillator. The power splitter is realized in the form of a tunable Mach-Zehnder interferometer (MZI) which can achieve arbitrary power splitting ratio based on the interference effect. Alice, then, encodes her information on the amplitude and the phase of the carrier according to a modulation scheme. This step is realized by sending the carrier pulse through a series of Mach Zehnder modulators (MZMs). An MZM is a versatile photonic integrated sub-circuit. Depending on the biasing conditions, they can perform diverse functionalities such as amplitude modulators, phase modulators, and variable optical attenuators. The local oscillator acts as a phase reference for the modulated carrier to enable coherent detection at the receiver. To send the modulated carrier and the local oscillator over the same fiber to Bob, Alice uses polarization multiplexing to minimize cross-talk. To that end, Alice first imposes a 90° rotation on the polarization state of the LO (TE to TM)

with a polarization rotator. Subsequently, she combines the TE polarized carrier and the TM polarized LO with a polarization beam splitter (PBS) which can provide the additional extinction ratio between the two signals as required by the protocol. The PBS is an MZI structure which induces bi-refrindex dependent propagation delays on the guided modes. Finally, the carrier and the LO propagate through the same quantum channel with orthogonal polarizations.

Bob employs a coherent receiver to decode the message from the in-phase (I) and the quadrature-phase (Q) amplitudes of the carrier signal via heterodyne detection. First, he uses a 2×2 MMI to split the transmitted signal into two tributaries. In both the tributaries, the LO and the carrier undergo polarization demultiplexing. The LO needs to be converted back to TE polarization before mixing it with the carrier signal in a second 2×2 MMI. If there is no phase difference between the carrier and the LO, the outputs of the MMI excite the photodiodes to produce an electrical current proportional to the in-phase quadrature of the optical signal upon balanced detection. On the other hand, if there is a $\frac{\pi}{2}$ phase difference between the carrier and the LO, the outputs of the MMI excite the photodiodes to produce an electrical current proportional to the $\pi/2$ -phase quadrature of the optical signal upon balanced detection. With classical post-processing, Alice and Bob can then go onto distill a secret key from the correlations of the encoded and decoded data.

3.3 Description of the Building Blocks

The above discussion suggests the key functionalities required for implementing a coherent optical transmitter and a receiver photonic integrated circuit are: light generation, single mode waveguiding of light, manipulation of optical amplitude and phase, polarization management, and absorption of light. In this regard, the essential building blocks are: passive optical waveguide, optical gain element, optical feedback element, power splitter, amplitude modulator, phase shifter, optical detector, and polarization management devices. For our design, we have used the building blocks available from the generic platform of our partner foundry Fraunhofer-HHI. All the building blocks are composed of InGaAsP/InP material system. The generic PIC platform from HHI creates all the building blocks by growing different epitaxial semiconductor layer stacks (which define different optical properties) on top of InP substrate (Figure 3.3). The waveguide core in the active components always form p-i-n diodes. The epitaxial layers for an active building block may contain a multiple quantum well region (for example, semiconductor optical amplifier, electroabsorption modulator) in addition to the layer stacks for passive components. In this section, we provide a functional description of these building blocks in terms of optical transfer matrices. An optical transfer matrix characterizes an optical component reflecting how it transforms an input light wave to an output. For example, an optical component with two inputs and two outputs is represented

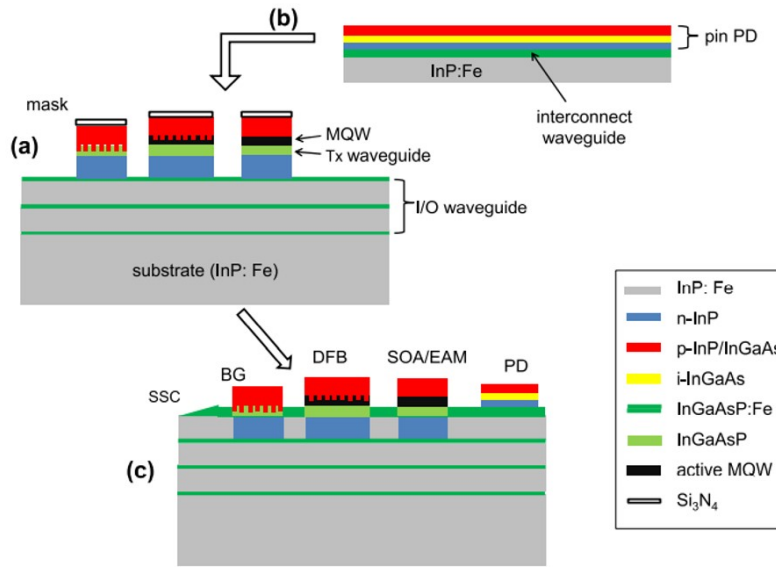


FIGURE 3.3: Schematic of generic InP PIC platform as developed at Fraunhofer HHI (BG = Bragg grating; PD = photodiode; DFB = DFB laser; EAM = electro-absorption modulator; SOA = semiconductor optical amplifier)[73], [74]

by a 2×2 transfer matrix

$$\mathbf{b} = \mathbf{T}\mathbf{a} \Rightarrow \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}. \quad (3.1)$$

The matrix elements are conditioned by the reciprocity theorem and also by unitarity, if the component is assumed to be lossless. The transfer matrix of a composite building block is the product of transfer matrices of the constituent basic building blocks (Figure 3.4).

Waveguide Section: The most fundamental building block in our transmitter and receiver photonic integrated circuits is waveguides. A passive waveguide section directs light from one point to another on the circuit restricting energy transmission only to one direction. In practice, a high refractive index material, called the core, is sandwiched between two lower index materials, called the claddings, to implement a waveguide. In this slab waveguide configuration (Figure 3.5), solution to Maxwell's equations yields guided modes and radiation modes. For guided modes most of the optical energy is confined within the core. The mode with the lowest cutoff frequency, determined by the waveguide geometry for a given optical wavelength, is called the fundamental mode. In our case, the quaternary material InGaAsP forms the core between two InP cladding layers. But, instead of regular slab geometry, Fraunhofer-HHI provides rib/ridge type waveguides. In this rib/ridge structure there is no top cladding layer (air confined) (Figure 3.6). Depending on the etch depth h there are three options-E200, E600, E1700. The deeper the etch, the better the optical confinement because of higher index contrast. All three variants (E200, E600, E1700) have a width of $2\mu\text{m}$ to satisfy optical single mode operation.

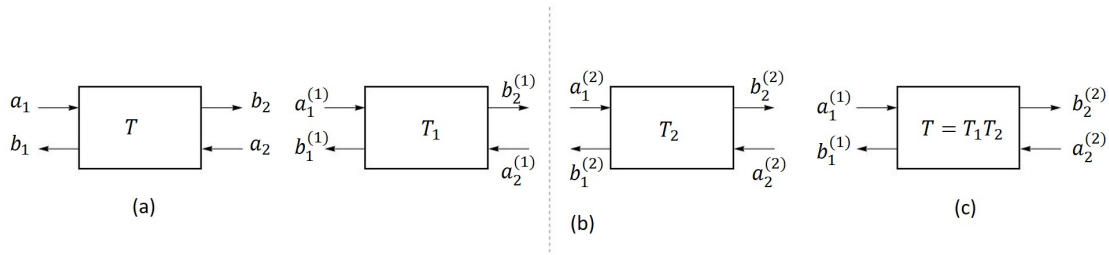


FIGURE 3.4: (a) Transfer matrix of a 2×2 component with (a_1, a_2) input amplitudes and (b_1, b_2) output amplitudes. Two components arranged in cascaded configuration (b) have an overall transfer matrix which is the product of the transfer matrices of both the components (c)[9]

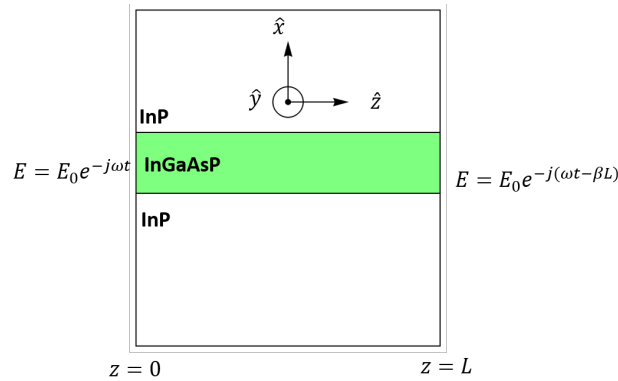


FIGURE 3.5: In a slab waveguide the core and the claddings have rectangular geometry. The cross section is assumed to be in the $x - y$ plane. The step index change is along the $\hat{x}$ axis. The mode propagates along $\hat{z}$ axis. Throughout this document we stick to this convention. The mode picks up a phase factor as it travels from $z = 0$ to $z = L$.

A guided mode suffers loss and accumulates phase as it travels along a waveguide section. The relevant optical transfer matrix captures this effect

$$T_{\text{WG}} = \begin{pmatrix} e^{-j\beta L} & 0 \\ 0 & 0 \end{pmatrix}, \quad (3.2)$$

where L is the length of the waveguide section, and β is the propagation constant of the guided mode along the waveguide. In general, β is a complex quantity. The $\text{Im}(\beta)$ determines propagation loss experienced by the guided mode. A passive waveguide section is characterized by the loss experienced by an optical wave propagating along it. This propagation loss is parameterized by the waveguide length. The passive waveguide loss for straight waveguides from HHI is $< 2 \text{ dB/cm}^2$. A waveguide bend might introduce mode-mismatch loss and radiation loss in addition to the propagation loss. As light turns a corner, a bend can excite other guided modes (contribute to mode mismatch loss) and radiation modes (contribute to radiation loss) which can exchange energy with the fundamental mode. The tighter the bend, the more the chance for light to

²All the foundry-specific numbers have been collected from [75]

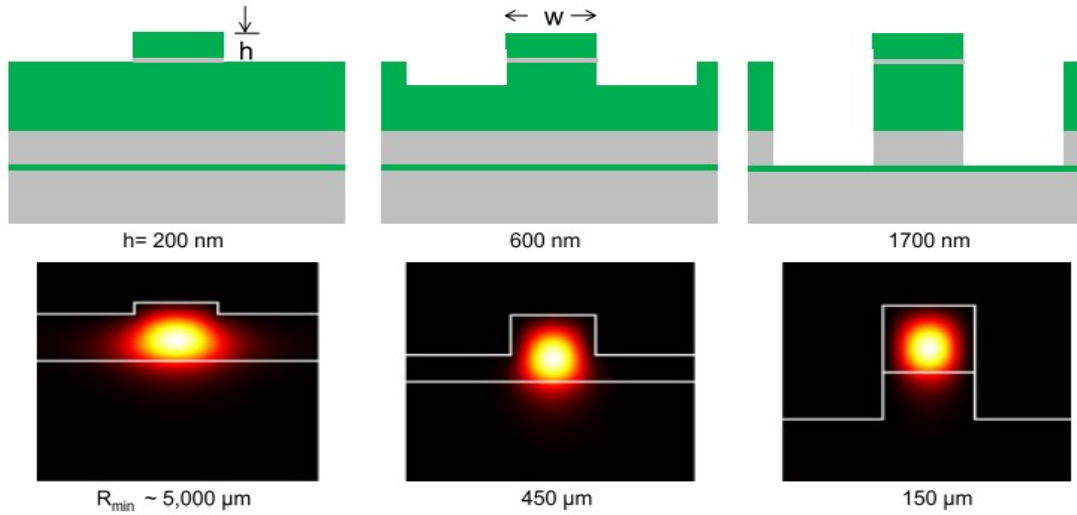


FIGURE 3.6: Different waveguide sections available from Fraunhofer-HHI. The waveguides consist of InP/InGaAsP layers. The rib width w is $2 \mu m$. The rib height h determines the index contrast and correspondingly the minimum bend radius R_{min} [73]

leak out. The suggested minimum bend radius for HHI waveguides is $150 \mu m$. In our design, we have used waveguide bends to route signals, which in some cases had enabled us to make use of the allocated die area more efficiently.

Tunable Phase Shifter: From equation (3.2), the phase accumulated by the guided mode while travelling along a waveguide section of length L is

$$\phi = \beta L = \frac{2\pi n}{\lambda} L. \quad (3.3)$$

For a waveguide section of fixed length L , any change in the value of n will result in a corresponding change in the phase shift. Here n is the effective index of the multilayered structure. It depends on the waveguide geometry, and the refractive indices (n_{core}, n_{clad}) of the core and the cladding materials. The core and the cladding material refractive indices, in turn, can be changed through a number of mechanisms such as temperature, current injection, electro-optic effect, mechanical stress etc. A tunable phase shifter is realized by placing a phase-actuating device on top of an integrated waveguide (Figure 3.7). With the application of an external control signal S over a suitable actuating device of length d , the material refractive index can be changed in that localised region, and thereby the propagation constant of any guided mode.

For our design, we have used a thermo-optic phase modulator (TOPM) and a current injection phase modulator (CIPM) available from HHI. In a TOPM, the refractive index of InGaAsP/InP waveguide structure is varied with temperature. A metal plate (actuator) placed on top the integrated waveguide is heated up by external current flow. The thermo-optic coefficient ($\frac{dn}{dT}$) for InGaAsP based waveguides is found to be 2.15×10^{-4} at $1.55 \mu m$ for TE polarization [76]. Therefore the refractive index change is positive with respect to the applied thermal energy.

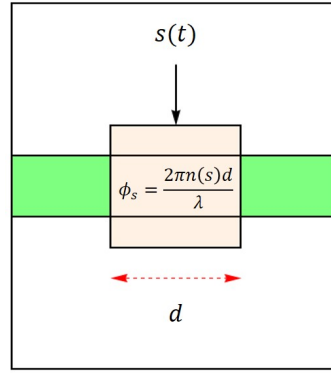


FIGURE 3.7: A tunable phase shifter consists of a phase actuator of length d mounted on top of a waveguide section.

The refractive index n is a complex quantity: $n = n_r + jn_i$. The imaginary part of the refractive index is determined by the absorption or loss coefficient α of the material. In a semiconductor structure such as doped InGaAsP/InP, the concentration of free charges contributes to the loss via absorption. Therefore, changing the concentration of free charges can lead to the change in the refractive index of the material. In practice, a p-i-n InP-diode (actuator) grown on top of the integrated waveguide is injected with current. The equation relating the refractive index variation with respect to the carrier concentration N can be derived from the Drude-Lorenz model [77]

$$\Delta n = -\frac{e^2 \lambda_0^2}{8\pi^2 c^2 \epsilon_0 n} \left[\frac{N_e}{m_{ce}^*} + \frac{N_v}{m_{ve}^*} \right], \quad (3.4)$$

where m_{ce}^* , m_{ve}^* are the effective masses of the conduction band electrons and holes, respectively. The minus ('-') sign indicates the change in refractive index is negative as free carriers are injected. Two important figures of merit for the phase modulators are the insertion loss [78] and the I_π (current required to induce a π phase-shift), both of which are parameterized by the length of the phase actuator.

The transfer matrix of a tunable phase shifter can be written as

$$T_{\text{PS}} = \begin{pmatrix} e^{-j\phi(s)} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} e^{-j\frac{2\pi n(s)d}{\lambda}} & 0 \\ 0 & 0 \end{pmatrix}. \quad (3.5)$$

For two parallel waveguides equipped with tunable phase shifters, the overall transfer matrix is a tensor product of the individual transfer matrices, but in concise form

$$T_{\text{PS}} = \begin{pmatrix} e^{-j\phi(s_1)} & 0 \\ 0 & e^{-j\phi(s_2)} \end{pmatrix}. \quad (3.6)$$

MMI: A multi mode interferometer, or an MMI coupler is an integrated equivalent of a power splitter (/combiner). It consists of single-mode access waveguides, for guiding light in and out of the device, and a multimode region. The

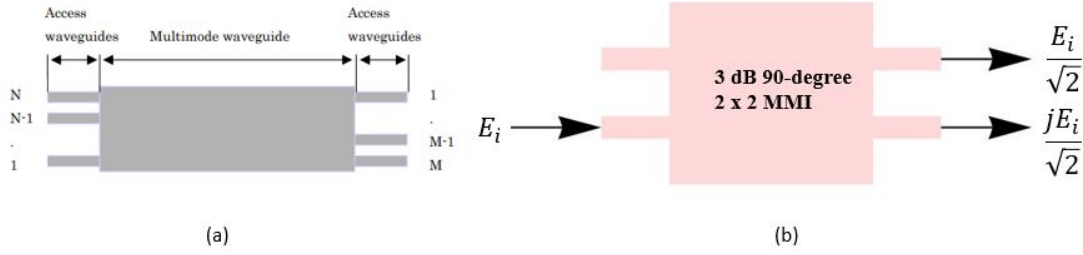


FIGURE 3.8: (a) An MMI coupler with N inputs and M outputs with a multimode region in between. (b) Schematic of the 2×2 MMI from HHI which we have used in our design. This building block is equivalent of a symmetric power splitter with a $\pi/2$ relative phase between the two outputs.

operation of optical MMI devices is based on self-imaging: *a property of multimode waveguides by which an input field profile is reproduced in single or multiple images at periodic intervals along the propagation direction of the guide* [79]. Generally, an MMI coupler can have any number of access waveguides, for example $M \times N$ couplers, where M and N are the number of input and output waveguides (Figure 3.8(a)). For our design, we have used 90° 3dB 2×2 MMIs, which can split an input field evenly into two output fields with $\frac{\pi}{2}$ phase shifts between them (Figure 3.8(b)). Appendix D illustrates the mathematical analysis of such a device.

The transfer matrix of a 90° 3-dB 2×2 -MMI coupler is similar to that of a symmetric lossless beam splitter. Ignoring the global phase terms one can write

$$T_{\text{MMI}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & j \\ j & 1 \end{pmatrix}. \quad (3.7)$$

The derivation can be found in [9], [80]. The matrix elements are determined by imposing the unitarity (energy conservation) and the reciprocity conditions on the MMI operation. The consequence is that one of the outputs will have a phase difference of $\frac{\pi}{2}$ with respect to the other.

Tunable MZI: We have used a tunable Mach Zehnder interferometer (MZI) in our design as an asymmetric power splitter to split the laser output into a weak carrier and a strong local oscillator. It is a compound building block consisting of two 90° 3-dB 2×2 MMIs at the input and at the output. Both the waveguide sections connecting the two MMIs are equipped with thermo-optic tunable phase shifters (Figure 3.9). The transfer matrix of an MZI comprising the "90° 3-dB 2×2 MMI–Tunable phase shifters–90° 3-dB 2×2 MMI" configuration has a transfer matrix

$$\begin{aligned} T_{\text{TMZI}} &= T_{\text{MMI}_1} T_{\text{PS}} T_{\text{MMI}_2} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & j \\ j & 1 \end{pmatrix} \begin{pmatrix} e^{j\phi_1} & 0 \\ 0 & e^{j\phi_2} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & j \\ j & 1 \end{pmatrix} = j e^{j\bar{\phi}} \begin{pmatrix} \sin(\frac{\Delta\phi}{2}) & \cos(\frac{\Delta\phi}{2}) \\ \cos(\frac{\Delta\phi}{2}) & -\sin(\frac{\Delta\phi}{2}) \end{pmatrix} \end{aligned} \quad (3.8)$$

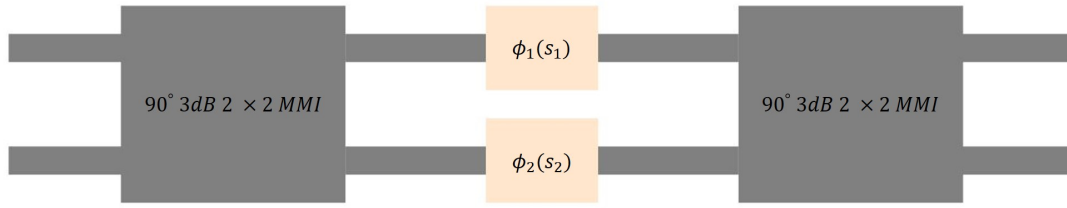


FIGURE 3.9: Schematic of a tunable MZI consisting of two MMIs at the input and the output. The phase shifters can be adjusted to produce outputs with any splitting ratio between 0 to 1.

Here, ϕ_1 and ϕ_2 are the static phase shifts experienced by the optical waves on the upper arms and the lower arms respectively. Moreover,

$$\Delta\phi = \phi_1 - \phi_2, \quad (3.9a)$$

$$\bar{\phi} = \frac{\phi_1 + \phi_2}{2}. \quad (3.9b)$$

It follows that for a field E_{in} at the input of the MZI, the two output fields are $E_{\text{out},1}$ and $E_{\text{out},2}$

$$E_{\text{out},1} = jE_{\text{in}} \exp\left\{j\frac{\phi_1 + \phi_2}{2}\right\} \sin\left(\frac{\phi_1 - \phi_2}{2}\right) \exp\{-j\omega_c t\}, \quad (3.10a)$$

$$E_{\text{out},2} = jE_{\text{in}} \exp\left\{j\frac{\phi_1 + \phi_2}{2}\right\} \cos\left(\frac{\phi_1 - \phi_2}{2}\right) \exp\{-j\omega_c t\}. \quad (3.10b)$$

The two outputs have a 90° phase difference. Disregarding the time harmonic part, one gets

$$E_{\text{out},1} = j e^{j\bar{\phi}} E_{\text{in}} \sin\left(\frac{\Delta\phi}{2}\right) \quad (3.11a)$$

$$E_{\text{out},2} = j e^{j\bar{\phi}} E_{\text{in}} \cos\left(\frac{\Delta\phi}{2}\right), \quad (3.11b)$$

resulting in

$$P_{\text{out},1}^2 = E_{\text{in}}^2 \sin^2\left(\frac{\Delta\phi}{2}\right) = E_{\text{in}}^2 (1 - \cos^2\left(\frac{\Delta\phi}{2}\right)) = 1 - P_{\text{out},2}^2 \quad (3.12)$$

assuming $P_{\text{in}} = 1$. Therefore, by properly adjusting the phase difference, it is possible to achieve any value from $[0, 1]$ as the splitting ratio for the two outputs.

Mach Zehnder Modulator: The building blocks in the Mach Zehnder modulator (MZM) are arranged in exactly the same way as in the tunable MZI (Figure 3.9) except that the phase shifters on the connecting arms are based on current-injection. The current-injection phase shifter is able to induce a dynamic phase shift on the guided mode compared to a static phase shift by the thermo-optic phase shifter. An MZM can act both as an amplitude modulator and a phase modulator depending on the operating conditions of the phase shifters [81]. The weak coherent signal is modulated both in amplitude and in phase. When

both the phase shifters in the MZI are electrically driven, the MZM is said to be operating in *dual drive configuration*. One can write the dynamic phase shifts $\phi_1(t), \phi_2(t)$ undergone by the optical waves in the upper arm and the lower arm of the MZM, respectively, as

$$\phi_j(t) = \gamma S_j(t) \quad j = 1 \text{ (upper arm), } 2 \text{ (lower arm)} \quad (3.13)$$

where $S_1(t)$ and $S_2(t)$ are the driving signals and γ is a proportionality factor. The magnitude of the signal S_π required to introduce a π -phase shift is the half-wave signal and can be found as

$$\phi_j = \gamma S_j \Rightarrow \pi = \gamma S_\pi \Rightarrow \gamma = \frac{\pi}{S_\pi}, \quad (3.14)$$

which allows us to write $\phi_j = \frac{\pi}{S_\pi} S_j$. Therefore, the output electric fields can be rewritten from equations (3.11a) and (3.11b) as

$$E_{\text{out},1} = jE_{\text{in}} \exp\{j\bar{\phi}\} \cos\left(\frac{[S_1(t) - S_2(t)]\pi}{2S_\pi}\right), \quad (3.15a)$$

$$E_{\text{out},2} = jE_{\text{in}} \exp\{j\bar{\phi}\} \sin\left(\frac{[S_1(t) - S_2(t)]\pi}{2S_\pi}\right), \quad (3.15b)$$

with

$$\bar{\phi} = \frac{\phi_1 + \phi_2}{2} = \frac{[S_1(t) + S_2(t)]\pi}{2S_\pi}, \quad (3.16a)$$

$$\Delta\phi = \phi_1 - \phi_2 = \frac{\pi}{S_\pi} [S_1(t) - S_2(t)]. \quad (3.16b)$$

Similarly, the output optical power can be written as

$$P_{\text{out},1} = |E_{\text{out}}|^2 = P_0 \cos^2\left(\frac{[S_1(t) - S_2(t)]\pi}{2S_\pi}\right), \quad (3.17a)$$

$$P_{\text{out},2} = |E_{\text{out}}|^2 = P_0 \sin^2\left(\frac{[S_1(t) - S_2(t)]\pi}{2S_\pi}\right), \quad (3.17b)$$

where $P_0 = |E_{\text{in}}|^2$. The instantaneous frequency shift or frequency chirp is given by

$$\omega_i = -\frac{d\bar{\phi}}{dt} = \frac{\pi}{2S_\pi} \left(\frac{dS_1}{dt} + \frac{dS_2}{dt} \right). \quad (3.18)$$

The frequency chirp combined with the fiber dispersion could lead to a pulse broadening and performance degradation. Therefore, it is desirable to have zero chirp for which the condition is

$$\begin{aligned} \frac{dS_1(t)}{dt} &= -\frac{dS_2(t)}{dt} \\ \Rightarrow S_1(t) &= -S_2(t) + S_{\text{bias}}, \end{aligned} \quad (3.19)$$

where S_{bias} is a constant bias signal. This driving condition is known as a balanced driving or push-pull operation [81]. Substituting equation (3.19) in equations (3.15a) and (3.16a) results in

$$E_{\text{out}} = E_{\text{in}} \cos \left(\left[S_1(t) - \frac{S_{\text{bias}}}{2} \right] \frac{\pi}{S_{\pi}} \right), \quad (3.20a)$$

$$P_{\text{out}} = P_{\text{in}} \cos^2 \left(\left[S_1(t) - \frac{S_{\text{bias}}}{2} \right] \frac{\pi}{S_{\pi}} \right), \quad (3.20b)$$

where we have considered only one of the outputs and ignored the constant phase factors while writing these expressions.

Implementation of Amplitude Shift Keying

From equation (3.20b), the optical power from one of the MZM outputs is

$$P_{\text{out}} = P_{\text{in}} \cos^2 \left(\left[S_1(t) - \frac{S_{\text{bias}}}{2} \right] \frac{\pi}{S_{\pi}} \right).$$

Let $S_1(t)$ be a polar NRZ such that

$$S_1(t) = \begin{cases} +S & \text{for bit '1'} \\ -S & \text{for bit '0'}. \end{cases} \quad (3.21)$$

Let us say we desire the following Mach-Zehnder output power

$$P_{\text{out}} = \begin{cases} P_{\text{in}} & \text{for bit '1'} \\ 0 & \text{for bit '0'} \end{cases} \quad (3.22)$$

Therefore, for bit '1',

$$\begin{aligned} P_{\text{out}} &= P_{\text{in}} \cos^2 \left(\left[S_1(t) - \frac{S_{\text{bias}}}{2} \right] \frac{\pi}{S_{\pi}} \right) = P_{\text{in}} \\ \Rightarrow \frac{\pi S}{S_{\pi}} - \frac{\pi S_{\text{bias}}}{2S_{\pi}} &= k\pi \text{ where } k = 0, \pm 1, \pm 2, \dots \end{aligned} \quad (3.23)$$

Similarly, for bit '0',

$$\begin{aligned} P_{\text{out}} &= P_{\text{in}} \cos^2 \left(\left[S_1(t) - \frac{S_{\text{bias}}}{2} \right] \frac{\pi}{S_{\pi}} \right) = 0 \\ \Rightarrow \frac{-\pi S}{S_{\pi}} - \frac{\pi S_{\text{bias}}}{2S_{\pi}} &= \frac{l\pi}{2} \text{ where } l = \pm 1, \pm 3, \dots \end{aligned} \quad (3.24)$$

Subtraction and addition of the above two equations lead to

$$S = \left(k - \frac{l}{2} \right) \frac{S_{\pi}}{2} \quad (3.25a)$$

$$S_{\text{bias}} = - \left(k + \frac{l}{2} \right) \frac{S_{\pi}}{2}, \quad (3.25b)$$

respectively. If we choose $k = 0$ and $l = -1$, we find

$$S = \frac{S_\pi}{4}, \quad (3.26a)$$

$$S_{\text{bias}} = \frac{S_\pi}{2}. \quad (3.26b)$$

Therefore, one way to perform binary amplitude shift keying or on off keying with a Mach-Zehnder Modulator in push-pull configuration, is to set the bias point at $S_{\text{bias}} = \frac{S_\pi}{2}$ and drive the phase modulators with modulating signal $\frac{S_\pi}{2}$ peak-to-peak. When $S = \frac{S_\pi}{4}$, corresponding to bit '1', constructive interference occurs and the MZM power transmission is at its peak. When $S = \frac{S_\pi}{4}$, destructive interference occurs and the MZM power output is zero.

Implementation of Phase Shift Keying

A Mach Zehnder modulator can be operated in dual drive configuration to generate arbitrary quadrature signals [82], [83]. The mathematical underpinning of the statement is explained in the following discussion. Let us consider the cross output of an MZM which is given by the expression ³

$$E_{\text{out}} = \frac{E_{\text{in}}}{2} [e^{j(\phi_1 - \frac{\pi}{2})} - e^{j(\phi_2 + \frac{\pi}{2})}], \quad (3.27)$$

where ϕ_1, ϕ_2 are the phase shifts induced by the upper and the lower phase modulators respectively (Figure (3.10a)). In terms of the bias voltages, ϕ_1 and ϕ_2 are given by

$$\phi_1 = \frac{\pi V_1}{V_\pi}, \quad \phi_2 = \frac{\pi V_2}{V_\pi}. \quad (3.28)$$

Now, an arbitrary M -ary modulated signal can be represented as a constellation point on the complex plane

$$s_i = r_i e^{j\theta_i}. \quad (3.29)$$

(r_i, θ_i) denote the polar coordinates of the constellation point. θ_i takes value within the range $0 \leq \theta_i \leq 2\pi$. Let us denote r_{max} as the maximum from the set of values

$$\{r_0, r_1, \dots, r_{M-1}\}. \quad (3.30)$$

Our goal is to rewrite equation (3.29) as

$$s_i = \frac{r_{\text{max}}}{2} [e^{j\phi_{i,1}} - e^{j\phi_{i,2}}], \quad (3.31)$$

so that it has a one-to-one correspondence with the MZM output expression written in equation (3.27). To that end, first, referring to Figure (3.10b), we note

³which is derived from equation (3.8) in Appendix D.

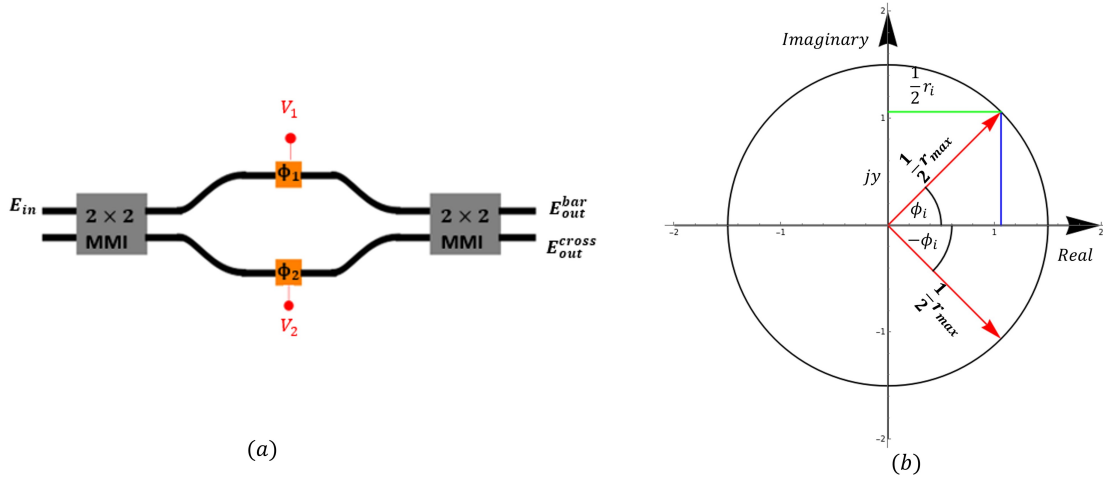


FIGURE 3.10: (a) A diagrammatic view of an MZM showing the control voltages and the outputs. (b) Procedure to find out $\phi_{i,1}, \phi_{i,2}$ for $s_i = r_i e^{j\theta_i}$

that the real number r_i can be calculated in the following way. From the figure, we can equate

$$\frac{1}{2}r_i \pm jy = \frac{1}{2}r_{max}e^{j\pm\phi_i}, \quad (3.32)$$

where

$$\phi_i = \cos^{-1} \left(\frac{r_i}{r_{max}} \right). \quad (3.33)$$

$\frac{1}{2}r_{max}$ is the distance of the point from the center. From equation (3.32), by simple algebra, it follows that

$$r_i = \frac{r_{max}}{2} [e^{j\phi_i} + e^{-j\phi_i}]. \quad (3.34)$$

Multiplying both sides of the equation (3.34) with $e^{j\theta_i}$, one gets

$$r_i e^{j\theta_i} = \frac{r_{max}}{2} [e^{j(\phi_i+\theta_i)} + e^{j(-\phi_i+\theta_i)}] = \frac{r_{max}}{2} [e^{j(\phi_i+\theta_i)} - e^{j(-\phi_i+\theta_i+\pi)}]. \quad (3.35)$$

If we define

$$\phi_{i,1} = \phi_i + \theta_i = \cos^{-1} \left(\frac{r_i}{r_{max}} \right) + \theta_i, \quad (3.36)$$

$$\phi_{i,2} = -\phi_i + \theta_i + \pi = -\cos^{-1} \left(\frac{r_i}{r_{max}} \right) + \theta_i + \pi, \quad (3.37)$$

then s_i is expressed in the form as in equation (3.31). Therefore, comparing with equation (3.27), to implement a particular constellation point $s_i = r_i e^{j\theta_i}$ with an MZM, the phase shifts need to be provided by the phase modulators on the upper arm and the lower arm are respectively given by

$$\phi_1 = \cos^{-1} \left(\frac{r_i}{r_{max}} \right) + \theta_i + \frac{\pi}{2}, \quad (3.38)$$

$$\phi_2 = -\cos^{-1} \left(\frac{r_i}{r_{max}} \right) + \theta_i + \frac{\pi}{2}. \quad (3.39)$$

Given these expressions, one can deduce the corresponding bias voltages from equation (3.28).

Implementing BPSK: Let us consider implementing binary phase shift keying (BPSK) with a single MZM in dual drive operation (Figure (3.11)). In BPSK, the phase modulator randomly switches the phase of the carrier between 0 and π . Table (3.1) lists the phases and the associated bias voltages for the phase modulators required to implement the BPSK constellation points.

Binary phase shift keying		
Constellation points	0	π
ϕ_1	$\frac{\pi}{2}$	$\frac{3\pi}{2}$
ϕ_2	$\frac{\pi}{2}$	$\frac{3\pi}{2}$
V_1	$\frac{V_\pi}{2}$	$\frac{3V_\pi}{2}$
V_2	$\frac{V_\pi}{2}$	$\frac{3V_\pi}{2}$

TABLE 3.1: BPSK implementation parameters for a single dual drive MZM

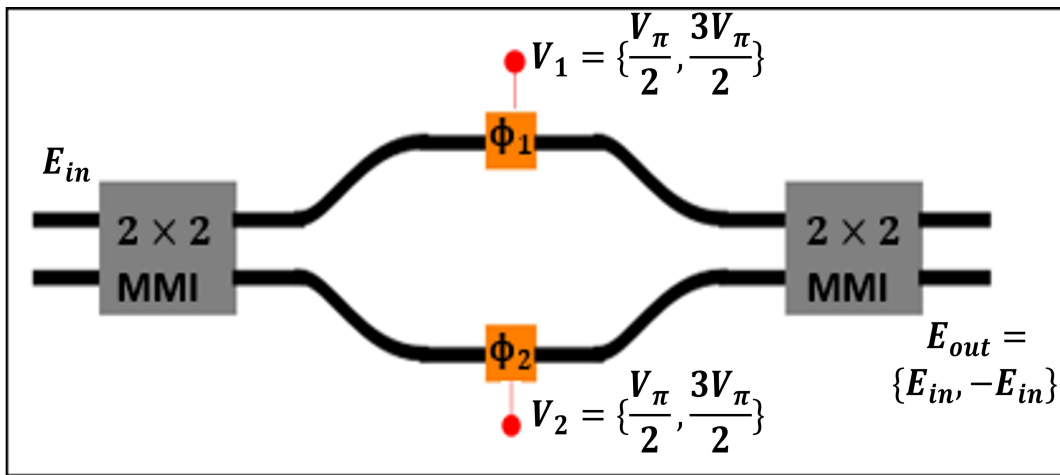


FIGURE 3.11: A diagrammatic view of an MZM with the control voltages and the outputs for binary phase shift keying.

The output waveforms from the MZM corresponding to the constellation points "0" and " π ", respectively, are

$$E_{\text{out,I}} = \frac{E_{\text{in}}}{2}(e^{j \cdot 0} - e^{j \cdot \pi}) = E_{\text{in}} \quad (3.40a)$$

$$E_{\text{out,II}} = \frac{E_{\text{in}}}{2}(e^{j \cdot \pi} - e^{j \cdot 2\pi}) = -E_{\text{in}}. \quad (3.40b)$$

Figure (3.11) gives a visual representation of the preceding discussion. The bias signals for the phase modulators flip flop between two levels $\frac{V_\pi}{2}$ and $\frac{3V_\pi}{2}$.

A stand-alone phase modulator can also implement BPSK, but a MZM-based PSK has been shown to have a more robust performance [84]. This is because, waveform imperfections in a phase modulator leads to the distortion of the optical phase as well, which degrades the device performance. On the other hand, imperfections in driving condition in an MZM only result in optical power variation, leaving the information-bearing optical phase intact.

Implementing QPSK: In quadrature phase shift keying (QPSK), the phase modulator randomly switches the phase of the carrier between four values, for example $\{\frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}\}$. Table (3.2) lists the phases and the associated bias voltages for the phase modulators required to implement the QPSK constellation points.

Quadrature phase shift keying				
Constellation points	$\frac{\pi}{4}$	$\frac{3\pi}{4}$	$\frac{5\pi}{4}$	$\frac{7\pi}{4}$
ϕ_1	$\frac{3\pi}{4}$	$\frac{5\pi}{4}$	$\frac{7\pi}{4}$	$\frac{\pi}{4}$
ϕ_2	$\frac{3\pi}{4}$	$\frac{5\pi}{4}$	$\frac{7\pi}{4}$	$\frac{\pi}{4}$
V_1	$\frac{3V_\pi}{4}$	$\frac{5V_\pi}{4}$	$\frac{7V_\pi}{4}$	$\frac{V_\pi}{4}$
V_2	$\frac{3V_\pi}{4}$	$\frac{5V_\pi}{4}$	$\frac{7V_\pi}{4}$	$\frac{V_\pi}{4}$

TABLE 3.2: QPSK implementation parameters for a single dual drive MZM

It is seen that to implement QPSK with only one MZM, the bias signal needs to switch among four levels. The output waveforms from the MZM corresponding to the constellation points $\{\frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}\}$ are, respectively, given by

$$E_{\text{out,I}} = \frac{E_{\text{in}}}{2} \left(e^{j(\frac{3\pi}{4} - \frac{\pi}{2})} - e^{j(\frac{3\pi}{4} - \frac{\pi}{2})} \right) = E_{\text{in}} e^{j\frac{\pi}{4}}, \quad (3.41a)$$

$$E_{\text{out,II}} = \frac{E_{\text{in}}}{2} \left(e^{j(\frac{5\pi}{4} - \frac{\pi}{2})} - e^{j(\frac{5\pi}{4} - \frac{\pi}{2})} \right) = E_{\text{in}} e^{j\frac{3\pi}{4}}, \quad (3.41b)$$

$$E_{\text{out,III}} = \frac{E_{\text{in}}}{2} \left(e^{j(\frac{7\pi}{4} - \frac{\pi}{2})} - e^{j(\frac{7\pi}{4} - \frac{\pi}{2})} \right) = E_{\text{in}} e^{j\frac{5\pi}{4}}, \quad (3.41c)$$

$$E_{\text{out,IV}} = \frac{E_{\text{in}}}{2} \left(e^{j(\frac{\pi}{4} - \frac{\pi}{2})} - e^{j(\frac{\pi}{4} - \frac{\pi}{2})} \right) = E_{\text{in}} e^{-j\frac{\pi}{4}} \quad (3.41d)$$

Figure (3.12) gives a visual representation of the preceding discussion. However, multilevel biasing leads to performance degradation. Instead of using one single MZM to implement QPSK, one can use two MZMs both in series. Both of the MZMs perform 2-PSK in dual drive configuration. The first one switches carrier phase randomly between 0 and π , while the second one switches carrier phase randomly between 0 and $\pi/2$. The overall output is a QPSK-modulated signal. This can be showed as follows.

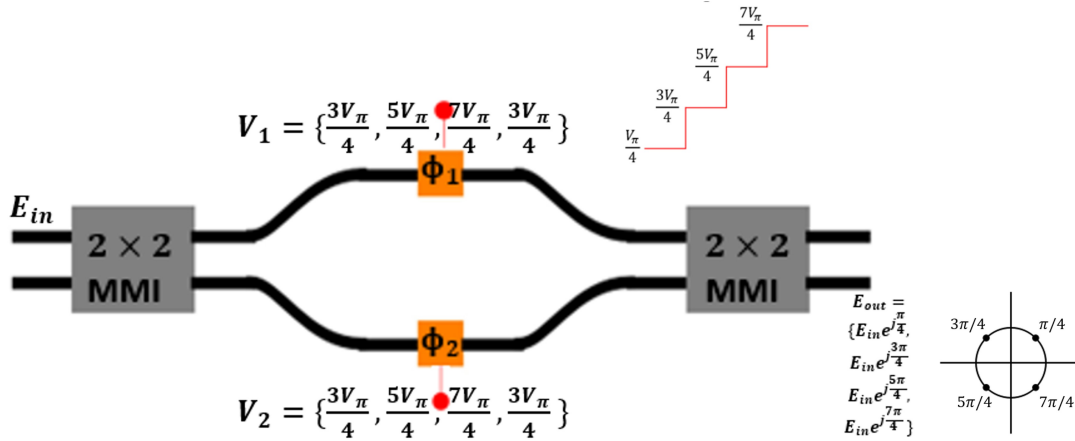


FIGURE 3.12: A diagrammatic view of an MZM with the control voltages and the outputs for quadrature phase shift keying.

For the first MZM, the constellation points are 0 and π . Therefore, from Table (3.1),

$$\phi_1 = \left\{ \frac{\pi}{2}, \frac{3\pi}{2} \right\} \Rightarrow V_1 = \left\{ \frac{V_\pi}{2}, \frac{3V_\pi}{2} \right\} \quad (3.42a)$$

$$\phi_2 = \left\{ \frac{\pi}{2}, \frac{\pi}{2} \right\} \Rightarrow V_2 = \left\{ \frac{V_\pi}{2}, \frac{3V_\pi}{2} \right\}. \quad (3.42b)$$

The output waveforms from the first MZM, are given by equation (3.40a) and (3.40b). For the second MZM, for constellation points 0 and $\frac{\pi}{2}$, from equations (3.38) and (3.39),

$$\phi_{I,1} = 0 + \frac{\pi}{2} = \frac{\pi}{2}, \quad (3.43a)$$

$$\phi_{I,2} = 0 + \frac{\pi}{2} = \frac{\pi}{2}, \quad (3.43b)$$

$$\phi_{II,1} = \frac{\pi}{2} + \frac{\pi}{2} = \pi, \quad (3.43c)$$

$$\phi_{II,2} = \frac{\pi}{2} + \frac{\pi}{2} = \pi. \quad (3.43d)$$

In the above equations, (1,2) correspond to the second MZM's upper arm phase modulator and the lower arm phase modulators respectively. Whereas (I, II) refer to the constellation points 0 and $\frac{\pi}{2}$ respectively. The corresponding bias

voltages according to equation (3.28) are

$$\phi_1 = \left\{ \frac{\pi}{2}, \pi \right\} \Rightarrow V_1 = \left\{ \frac{V_\pi}{2}, V_\pi \right\} \quad (3.44a)$$

$$\phi_2 = \left\{ \frac{\pi}{2}, \pi \right\} \Rightarrow V_2 = \left\{ \frac{V_\pi}{2}, V_\pi \right\}. \quad (3.44b)$$

The input to the second MZM are the outputs from the first MZM which have phases either 0 or π . Upon random phase modulation of this input between 0 and $\frac{\pi}{2}$, the output waveforms from the second MZM may appear with four possible phases.

For modulating with phase 0,

$$E_{\text{out,I}} = \frac{E_{\text{in}}}{2} (e^{j \cdot 0} - e^{j \cdot \pi}) = E_{\text{in}}, \quad (3.45a)$$

$$E_{\text{out,II}} = \frac{E_{\text{in}}}{2} e^{j\pi} (e^{j \cdot 0} - e^{j \cdot \pi}) = E_{\text{in}} e^{j\pi}. \quad (3.45b)$$

For modulating with phase $\frac{\pi}{2}$,

$$E_{\text{out,III}} = \frac{E_{\text{in}}}{2} (e^{j \cdot \frac{\pi}{2}} - e^{j \cdot \frac{3\pi}{2}}) = E_{\text{in}} e^{j \frac{\pi}{2}}, \quad (3.45c)$$

$$E_{\text{out,IV}} = \frac{E_{\text{in}}}{2} e^{j\pi} (e^{j \cdot \frac{\pi}{2}} - e^{j \cdot \frac{3\pi}{2}}) = E_{\text{in}} e^{j \frac{3\pi}{2}}. \quad (3.45d)$$

Figure (3.13) illustrates the idea of implementing QPSK with two consecutive MZMs using different biasing conditions. The concept of using MZMs in series

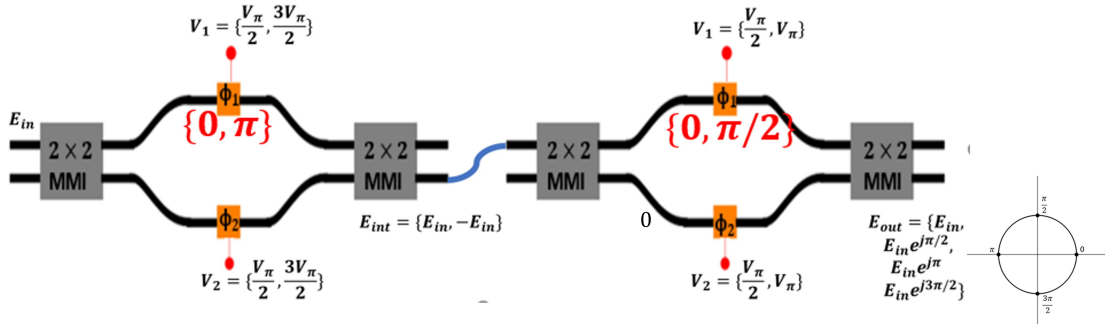


FIGURE 3.13: A diagrammatic view of implementing quadrature phase shift keying with cascaded MZMs.

for realizing higher order modulation formats can be extended to the case of 8-PSK. Three MZMs are placed in series. The first MZM switches carrier phase randomly between 0 and π , the second one switches carrier phase randomly between 0 and $\frac{\pi}{2}$, the last one switches carrier phase randomly between 0 and $\frac{\pi}{4}$. This randomization in modulation is absolutely crucial, otherwise the overall output will not emerge as a 8-PSK modulated signal. The modulator phases and bias voltages corresponding to the constellation points $\{0, \frac{\pi}{4}\}$ are

$$\phi_1 = \left\{ \frac{\pi}{2}, \frac{3\pi}{4} \right\} \Rightarrow V_1 = \left\{ \frac{V_\pi}{2}, \frac{3V_\pi}{4} \right\} \quad (3.46a)$$

$$\phi_2 = \left\{ \frac{\pi}{2}, \frac{3\pi}{4} \right\} \Rightarrow V_2 = \left\{ \frac{V_\pi}{2}, \frac{3V_\pi}{4} \right\}. \quad (3.46b)$$

The input to the last modulator are the waveforms denoted by the equations (3.45a)-(3.45d). Upon random phase modulation of this input between 0 and $\frac{\pi}{4}$, the output waveforms from the last MZM may appear with eight possible phases.

For modulating with phase 0,

$$E_{\text{out,I}} = \frac{E_{\text{in}}}{2}(e^{j \cdot 0} - e^{j \cdot \pi}) = E_{\text{in}} \quad (3.47a)$$

$$E_{\text{out,III}} = \frac{E_{\text{in}}}{2}e^{j\frac{\pi}{2}}(e^{j \cdot 0} - e^{j \cdot \pi}) = E_{\text{in}}e^{j\frac{\pi}{2}}, \quad (3.47b)$$

$$E_{\text{out,II}} = \frac{E_{\text{in}}}{2}e^{j\pi}(e^{j \cdot 0} - e^{j \cdot \pi}) = E_{\text{in}}e^{j\pi}, \quad (3.47c)$$

$$E_{\text{out,IV}} = \frac{E_{\text{in}}}{2}e^{j\frac{3\pi}{2}}(e^{j \cdot 0} - e^{j \cdot \pi}) = E_{\text{in}}e^{j\frac{3\pi}{2}}. \quad (3.47d)$$

For modulating with phase $\frac{\pi}{4}$,

$$E_{\text{out,V}} = \frac{E_{\text{in}}}{2}(e^{j(\frac{3\pi}{4}-\frac{\pi}{2})} - e^{j(\frac{3\pi}{4}-\frac{\pi}{2})}) = E_{\text{in}}e^{j\frac{\pi}{4}} \quad (3.47e)$$

$$E_{\text{out,VI}} = \frac{E_{\text{in}}}{2}e^{j\frac{\pi}{2}}(e^{j(\frac{3\pi}{4}-\frac{\pi}{2})} - e^{j(\frac{3\pi}{4}-\frac{\pi}{2})}) = E_{\text{in}}e^{j\frac{3\pi}{4}}, \quad (3.47f)$$

$$E_{\text{out,VII}} = \frac{E_{\text{in}}}{2}e^{j\pi}(e^{j(\frac{3\pi}{4}-\frac{\pi}{2})} - e^{j(\frac{3\pi}{4}-\frac{\pi}{2})}) = E_{\text{in}}e^{j\frac{5\pi}{4}}, \quad (3.47g)$$

$$E_{\text{out,VIII}} = \frac{E_{\text{in}}}{2}e^{j\frac{3\pi}{2}}(e^{j(\frac{3\pi}{4}-\frac{\pi}{2})} - e^{j(\frac{3\pi}{4}-\frac{\pi}{2})}) = E_{\text{in}}e^{j\frac{7\pi}{4}}. \quad (3.47h)$$

Figure (3.14) illustrates the idea of implementing 8-PSK with three consecutive MZMs using different biasing conditions.

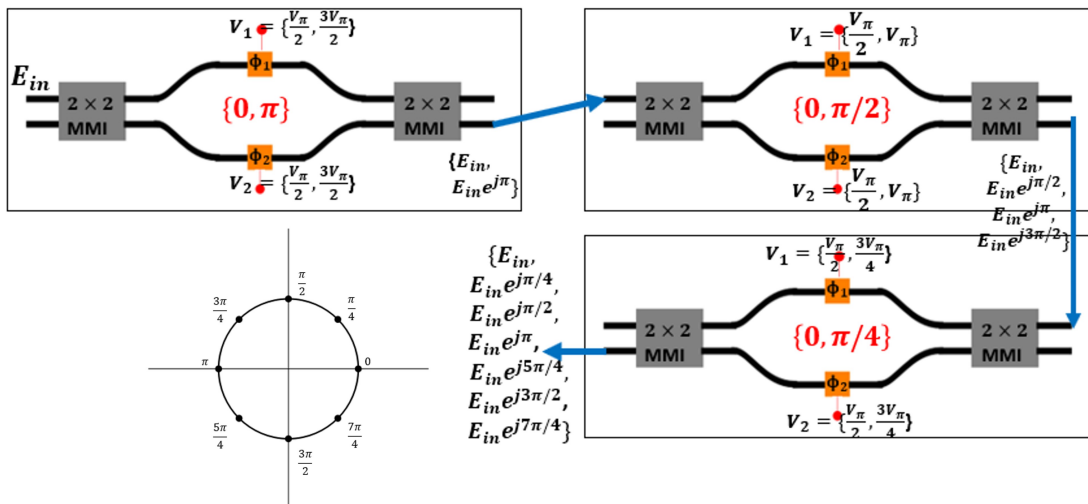


FIGURE 3.14: A diagrammatic view of implementing 8-PSK with cascaded MZMs.

The most well-studied case for implementing CV QKD has been the one based on Gaussian modulation. It is possible to implement Gaussian modulation of

the optical quadratures by two or more MZMs placed in series. Appendix D contains the mathematical reasoning. Since the power of a random signal is related to its variance, one can make use of an MZI-based variable optical attenuator (VOA) to set the target variance for a Gaussian modulation. Thanks to their versatility, MZI's or MZM's are at the heart of our transmitter design. We have placed four MZM blocks on our transmitter chip, which should provide Alice with plenty of control while encoding her data.

Polarization Rotator: The polarization of a guided mode is expressed in terms of its relative energy [85] as

$$\text{Pol}_{\text{TE}} = \frac{\int |E_{\text{TE}}|^2 dx dy}{\left(\int |E_{\text{TE}}|^2 + \int |E_{\text{TM}}|^2 \right) dx dy}; \text{Pol}_{\text{TM}} = \frac{\int |E_{\text{TM}}|^2 dx dy}{\left(\int |E_{\text{TE}}|^2 + \int |E_{\text{TM}}|^2 \right) dx dy}. \quad (3.48)$$

The mode angle is defined as $\alpha = \cos^{-1} \sqrt{\text{Pol}_{\text{TE}}}$. From this definition, TE mode which is parallel to the substrate has a mode angle of 0° . A polarization rotator induces a 90° rotation on the polarization state of the guided mode (TE $\rightarrow$ TM, or TM $\rightarrow$ TE). It can be regarded as the integrated equivalent of a half-wave plate rotated at 45° . In the Jones matrix formalism, the transformation matrix for a half wave plate at that orientation is

$$J_{\text{HWP}}(45^\circ) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (3.49)$$

It transforms $\hat{x}$ polarized light to $\hat{y}$ polarized light and vice versa (Appendix D). The polarization rotator available from the generic platform of HHI structurally employs two different sections of slanted sidewall waveguides (Figure 3.15)[86], [87], [88], [89]. A waveguide with a straight and a slanted (at an angle ε) sidewall can support a modal basis tilted by an angle α because of the boundary conditions⁴. A fundamental TE mode coming from an input straight waveguide to this PR waveguide excites both the tilted modes. The two modes exhibit a birefringence Δn because of the asymmetric geometry of the waveguide. After half a beat length $L_\pi = \frac{\pi}{\beta_1 - \beta_2}$, the modes are completely out of phase and recombine into the orthogonally polarized TM mode in an output straight waveguide. The Jones matrix of this polarization rotator waveguide can be found as follows

$$J_{\text{PR}} = R(-\alpha) T_{\text{biref}} R(\alpha) = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \cdot \begin{pmatrix} e^{j\pi\delta} & 0 \\ 0 & e^{-j\pi\delta} \end{pmatrix} \cdot \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}, \quad (3.50)$$

where we first switch to the tilted basis from the TE-TM basis, then take into account the birefringence effect, and finally switch back to the original coordinate

⁴all the angles are with respect to the substrate plane

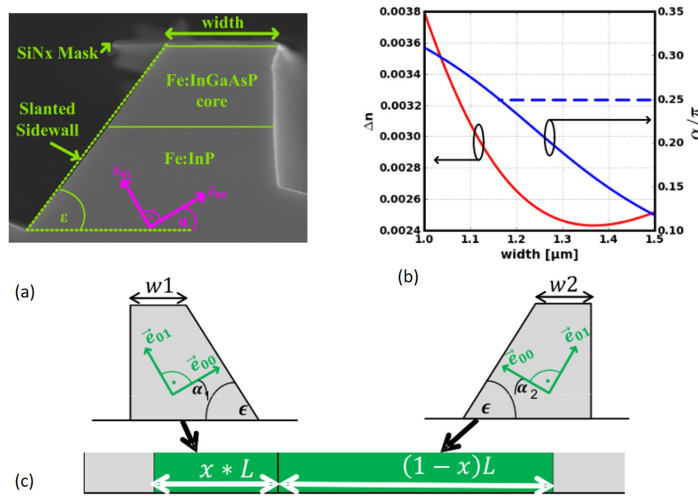


FIGURE 3.15: (a) SEM image of the slanted waveguide cross section from HHI. The waveguide width is defined at the top. $\epsilon = 52^\circ$ in this work [87]. (b) Birefringence and modal tilt as a function of waveguide width. The slope of these curves become smaller for larger widths, making them more favorable for fabrication tolerance [88]. (c) Schematic presentation of the two-section slanted sidewall PRs with relevant parameters [87].

system [90]. In the closed form,

$$J_{\text{PR}}^{(1)} = \begin{pmatrix} e^{j\pi\delta} \sin^2 \alpha + e^{-j\pi\delta} \cos^2 \alpha & \frac{1}{2}(1 - e^{2j\pi\delta})e^{-j\pi\delta} \sin 2\alpha \\ \frac{1}{2}(1 - e^{2j\pi\delta})e^{-j\pi\delta} \sin 2\alpha & e^{j\pi\delta} \sin^2 \alpha + e^{-j\pi\delta} \cos^2 \alpha \end{pmatrix}. \quad (3.51)$$

Here $\delta = \frac{\Delta n L}{\lambda}$ depends on the birefringence and the length of the waveguide. For a fixed length L and slant angle ϵ the Jones matrix for a polarization rotator is parameterized by the waveguide width w (Figure 3.15b). It has been shown that a polarization rotator realized in the form of a single section slanted wall waveguide is very sensitive to fabrication tolerance. In the two-section design, two slanted sidewall waveguides with different widths are placed in series configuration. The geometries of both sections should be the same but the cross-sections are mirror images of one another. As a result the tilt angles are related as $\alpha_1 = -\alpha_2$. The overall Jones matrix for the polarization rotator is

$$J_{\text{PR}}^{(2)} = J_1(w_1, L_1) \cdot J_2(w_2, L_2). \quad (3.52)$$

The two sections of differing widths (w_1, w_2) have a tapered joint. The optimum widths are those for which

$$J_{\text{PR}}^{(2)} \approx \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (3.53)$$

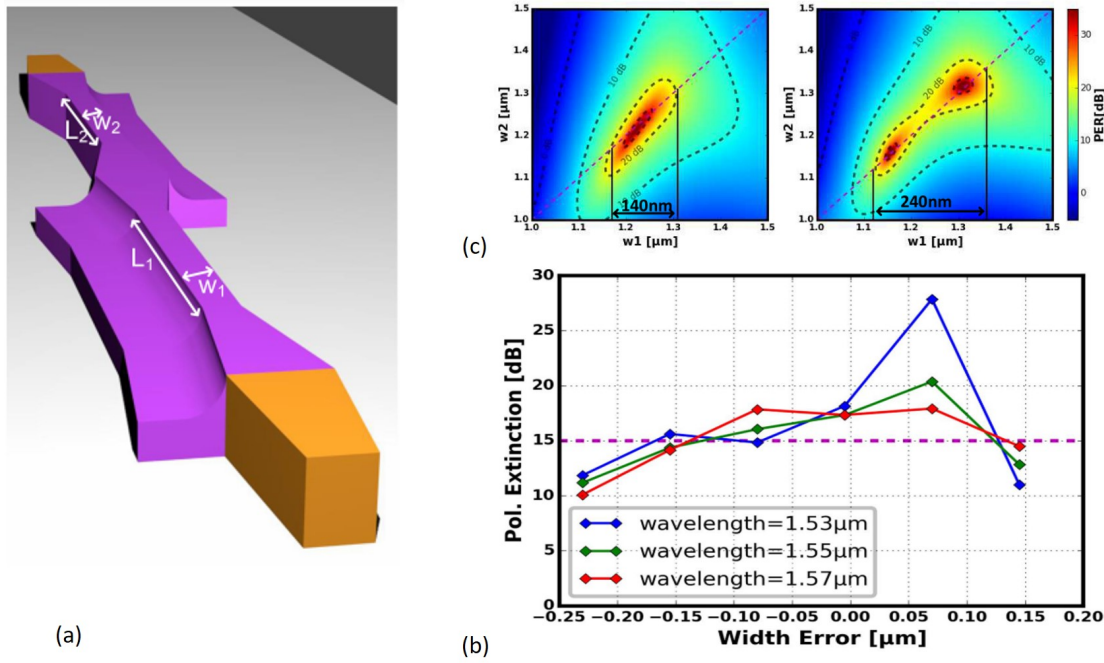


FIGURE 3.16: (a) 3D cartoon of the overall structure of the PR design including input-output waveguides in different colors. The asymmetry is flipped between the two sections. All transitions are tapered to minimize losses and reflection [89]. (c) Performance of a polarization converter building block consisting of a 2-section slanted waveguide structure. Simulated polarization extinction ratio (PER) as a function of waveguide top widths of the two sections. The dashed magenta lines are iso-PER contour lines. The first one for device $\alpha_1 = -\alpha_2 = \pi/4$, $\delta_1 = 0.25$, $\delta_2 = 0.75$ with waveguide width tolerance of 140 nm. The second one for device $\alpha_1 = -\alpha_2 = 0.21\pi$, $\delta_1 = 0.29$, $\delta_2 = 0.74$ with waveguide width tolerance of 240 nm. These parameters are obtained from numerical optimization of the Jones matrix tolerance equation (??). It is assumed that a deviation in width will be equal over both sections [88]. (b) Measurement results of PER as a function of deviation from the nominal width = 1.2 μm [87]

The two section design has been shown to have a decreased sensitivity to fabrication errors. The figure of merit for this building block is the polarization extinction ratio at the output which is defined as

$$\text{PER} = \frac{\text{Pol}_{\text{TE}}}{\text{Pol}_{\text{TM}}} = \frac{\int |E_{\text{TE}}|^2 dx dy}{\int |E_{\text{TM}}|^2 dx dy}. \quad (3.54)$$

Figure (3.16c) & (3.16b) shows the dependence of PER with respect to the design parameter, waveguide width w .

Polarization Beam Splitter: This building block splits the incident light into its TE and TM components. The PBS from HHI is realized in MZI configuration with two 2×2 90° 3-dB MMIs at the input and at the output and two straight waveguide sections equipped with thermo-optic phase shifters in between (Figure 3.17a) [91],[92]. The MZI is unbalanced with respect to the TE and TM polarization. The first MMI equally splits the input light into two outputs with a phase difference of 90° for both polarizations. The widths of the straight

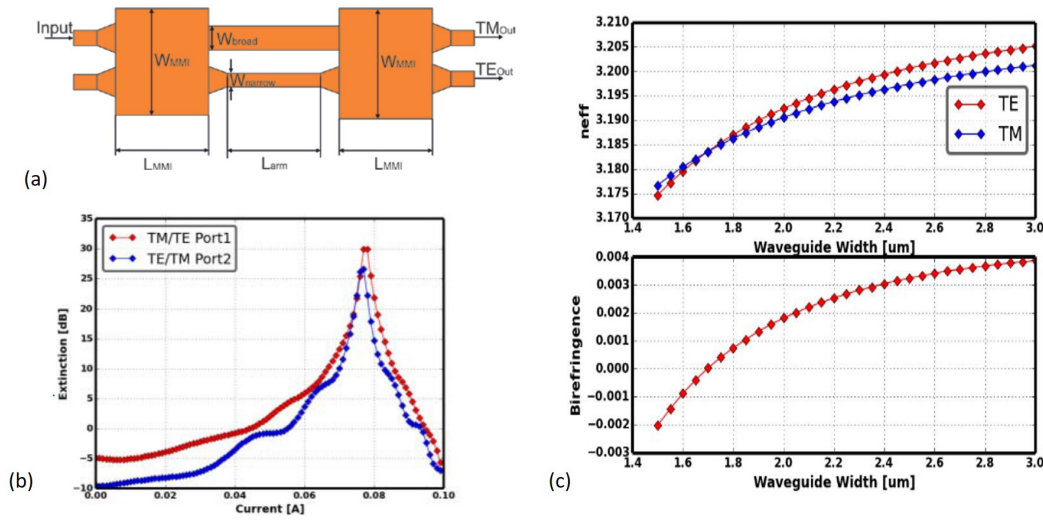


FIGURE 3.17: (a) Schematic view of the MZI-PBS [91] (c) Simulated dependence of birefringence on waveguide widths with etch depth of 1700 nm. At the waveguide width of $1.7\mu m$ the group indices of the two polarization states are equal, and hence no birefringence at this point [92]. (b) PER as a function of heating current for a PBS with arm length of $200\mu m$ [92].

waveguide sections can be designed to introduce birefringence between TE and TM polarizations (Figure 3.17c). The thermo-optic phase shifters enable finer adjustment of the relative phases for both the polarization states. The phases accumulated during propagation per arm is

$$\phi_{TE,TM}^{(i)} = \frac{2\pi n_{TE,TM}^{(i)} L}{\lambda}, \quad (3.55)$$

where $i = 1, 2$ refers to the upper and the lower arms of the PBS respectively, L is the arm length, λ is the optical wavelength. The phase differences can be explicitly written as

$$\Delta\phi_{TE} = \frac{2\pi L}{\lambda} (n_{broad, TE} - n_{narrow, TE}), \quad (3.56a)$$

$$\Delta\phi_{TM} = \frac{2\pi L}{\lambda} (n_{broad, TM} - n_{narrow, TM}), \quad (3.56b)$$

where the subscripts "broad" and "narrow" refer to the upper and lower waveguide sections, respectively. Assuming the PBS to be lossless, and neglecting polarization dependence of the MMIs the power transfer functions at the two output ports of the second MMI are given as

$$P_{TE, TM}^{(1)} = \cos^2 \left(\frac{\Delta\phi_{TE, TM}}{2} + \frac{\pi}{2} \right), \quad (3.57a)$$

$$P_{TE, TM}^{(2)} = \cos^2 \left(\frac{\Delta\phi_{TE, TM}}{2} \right). \quad (3.57b)$$

If one wants the PBS to output TE at port 1 and TM at port 2⁵, the phase differences can be chosen as follows $\Delta\phi_{TE} = \pi$, and $\Delta\phi_{TM} = 0$. Then from the power transfer function

$$P_{TE}^{(1)} = \sin^2\left(\frac{\Delta\phi_{TE}}{2}\right) = 1 \text{ while } P_{TE}^{(2)} = 0, \quad (3.58a)$$

$$P_{TM}^{(2)} = \cos^2\left(\frac{\Delta\phi_{TM}}{2}\right) = 1 \text{ while } P_{TM}^{(1)} = 0 \quad (3.58b)$$

The FOMs for this device is its maximum polarization extinction ratio (Figure 3.17b), and the insertion loss.

Photodetector: A photodetector produces an electric current proportional to the optical energy impinged on it. A p-i-n diode can detect light under reverse bias. The photodetector available from FhG-HHI is a composite building block. The PIN photodiode is placed on top of the integrated waveguide. Some important performance characteristics for this building block are sensitivity (0.8 A/W @1550 nm), 3 dB bandwidth (5 GHz), dark current (2 nA @-2V), and input power threshold for linear regime operation.

DFB Laser: The source of our coherent light for CV QKD at the telecommuni-

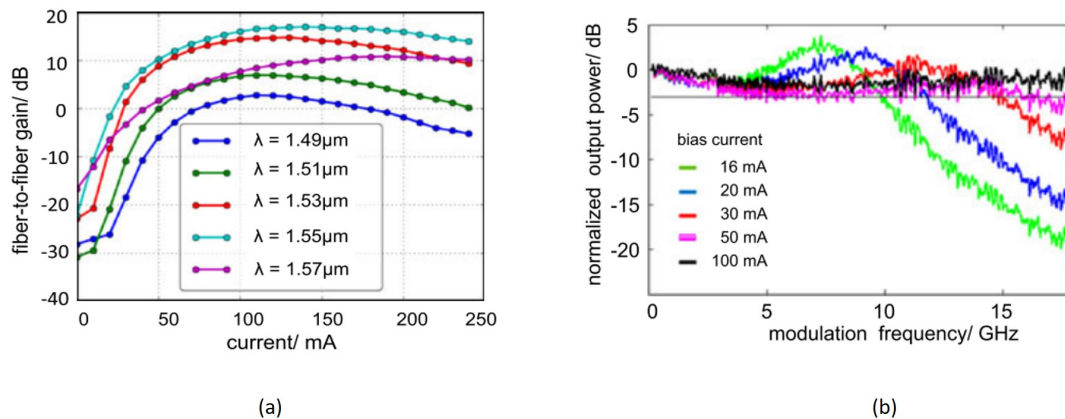


FIGURE 3.18: (a) Optical gain plotted against pump current of a 500 μm long SOA. The SOA had an input and an output waveguide both terminating to single mode fiber via spot size converters (b) Small signal response vs bias current curve for a 200 μm long DFB laser[73]

cation wavelength is a distributed feedback laser, another composite building block available from FhG-HHI. It contains a semiconductor optical amplifier (SOA) which is based on a multiple quantum well (MQW) ridge waveguide structure. The MQW in SOA provides optical gain. Besides there is a grating with complex coupling coefficients. The grating structure is etched on top of the MQW structure, and distributed across the gain region. The task of the grating is to provide selective feedback to the resonant wavelength so that the laser

⁵contrary to what shown in (Figure 3.17a)

outputs a narrowband spectrum which is essential for our application. The DFB laser building block from FhG-HHI only emits light at TE polarization. It has a tuning range of 4 nm, output power of 5 mW (@ 100 mA), and small signal bandwidth of 20 GHz (@ 100 mA). The small signal bandwidth related to the modulation frequency is sufficient for our purpose.

Spot Size Converter: The spot size converter couples light in and out of the chip via optical fiber. It has a tapered cross section which enables the waveguide mode to undergo an adiabatic size conversion. An SSC at the transmitter enlarges the modal area to have the maximum overlap with the cross section of standard single mode fibers. An SSC at the receiver does the opposite. The SSC building block from FhG-HHI has a waveguide to SSMF coupling loss of ~ 1.5 dB.

Miscellaneous: The only active-passive integration technology available from HHI MPW platform is butt joining two different epitaxial layers. The HHI *butt-joint coupler* BB acts as a transition element between active waveguides (lasers, photodiodes, phase modulators etc) to passive waveguides (E200/E1700). The insertion loss is 2 dB. Besides, there are *passive waveguide transitions* for example E200-E1700 (insertion loss 0.15 dB). The *waveguide metal transitions* allow crossings between metal stripes (connecting electrical ports of the active BBs to the bondpads) and waveguides.

3.4 Mathematical description of the whole system

Let the electric field of the optical wave emitted by the DFB laser be denoted by

$$\mathbf{E}(\mathbf{r}, t) = E_0(x, y)p(t)e^{-j(\omega t - \phi)}\hat{y}. \quad (3.59)$$

Here, $p(t)$ denotes the pulse shape underneath which the field is oscillating at angular frequency $\omega = 193.414$ THz. $\phi = \frac{2\pi}{\lambda_0}n_{\text{eff}}$ denotes the phase accumulated by the wave as it travels through out the circuit. And $\hat{y}$ refers to the TE polarization of the emitted wave. Although the exact analysis of the field propagation in a photonic integrated circuit should invoke the formalism of guided wave optics, here we take a simplified approach by replacing the guided modes with plane waves i.e. $E_0(x, y) = E_0$.

The tunable MZI splits the input field wave into a weak carrier signal and a strong local oscillator (LO) signal according to an splitting ratio $\alpha : 1 - \alpha$ defined by thermo-optic tuning

$$\alpha E(t)\hat{y} = p(t)E_C e^{-j(\omega t - \phi_c)}\hat{y}, \quad (3.60a)$$

$$(1 - \alpha)E(t)\hat{y} = p(t)E_{LO} e^{-j(\omega t - \phi_{LO})}\hat{y}, \quad (3.60b)$$

where $E_C = \alpha E_0$ and $E_{LO} = (1 - \alpha)E_0$ are the carrier and LO amplitudes respectively. The carrier signal is the one on which Alice will encode her message with letters from a certain alphabet whereas the local oscillator acts as the phase reference for the modulated signal at the receiver for coherent detection.

The weak coherent pulse signal then goes through a series of Mach Zehnder modulators. As discussed earlier, depending on how an MZM is operated, it can either act as amplitude modulator or a phase modulator. If the carrier wave is modulated both in amplitude and phase, the outgoing field wave can be written as

$$\hat{y}_p(t)E_C e^{-j(\omega t - \phi_C)} \xrightarrow{\text{series of MZMs}} \hat{y}_p(t)m_k e^{j\theta_k} E_C e^{-j(\omega t - \phi_C)}. \quad (3.61)$$

Here, $m_k e^{j\theta_k}$ is the complex envelop which contains the message encoded on the carrier. k is a random variable which can denote different value of $\{m, \theta\}$ at different pulse interval.

To send the local oscillator and the carrier signal over the same fiber from the transmitter to the receiver, they need to be polarization multiplexed to minimize cross-talk. To that end, the local oscillator is converted from TE-polarized light to TM-polarized light by a polarization converter,

$$\hat{y}_{E_{LO}} p(t) e^{-j(\omega t - \phi_{LO})} \xrightarrow{\text{PC}} \hat{x}_{E_{LO}} p(t) e^{-j(\omega t - \phi_{LO})}, \quad (3.62)$$

where $\hat{x}$ denotes the TM polarization state. The TE-polarized carrier and the TM-polarized LO then are combined with a MZI-based 2×2 polarization beam splitter. The outputs of the polarization beam splitter are

$$\hat{y}_{E_C}(t) \sin\left(\frac{\Delta\phi_{TE}}{2}\right) + \hat{x}_{E_{LO}}(t) \cos\left(\frac{\Delta\phi_{TM}}{2}\right) \text{ at output port 1,} \quad (3.63a)$$

$$\hat{y}_{E_C}(t) \cos\left(\frac{\Delta\phi_{TE}}{2}\right) - \hat{x}_{E_{LO}}(t) \sin\left(\frac{\Delta\phi_{TM}}{2}\right) \text{ at output port 2.} \quad (3.63b)$$

Here

$$\Delta\phi_{TE} = \frac{2\pi L}{\lambda} (n_{TE}^{(1)} - n_{TE}^{(2)}) \quad (3.64a)$$

$$\Delta\phi_{TM} = \frac{2\pi L}{\lambda} (n_{TM}^{(1)} - n_{TM}^{(2)}) \quad (3.64b)$$

are the polarization dependent phase shifts experienced by the carrier and LO, respectively on the arms of the MZI. The superscripts (1,2) denote the upper and the lower arm of the PBS-MZI. With suitable bi-refrindex engineering of the MZI-arms and thermo-optic tuning, a desired polarization extinction ratio (PER) can be achieved at any of the outputs. For our case, we want a high extinction ratio between our carrier signal and local oscillator. Therefore if we transmit signal to the fiber from the PBS output port 1, the condition is $\Delta\phi_{TE} \rightarrow 0$, and $\Delta\phi_{TM} = 0$. On the other hand, to transmit signal from output port 2, the condition is $\Delta\phi_{TE} \rightarrow \pi$, and $\Delta\phi_{TM} = \pi$. The polarization-multiplexed signal is then routed from the chip to the fiber via spot size converters.

Considering an idealized situation, where the optical fiber channel introduces neither distortion nor noise, but only the field amplitudes get attenuated, one can write the received signal at the receiver as

$$E_C(t) \xrightarrow{\text{optical fiber}} \tilde{E}_C(t), \quad (3.65a)$$

$$E_{LO}(t) \xrightarrow{\text{optical fiber}} \tilde{E}_{LO}(t). \quad (3.65b)$$

At the receiver, a 2×2 3-dB 90° MMI first acts on the received signal in the following way

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & j \\ j & 1 \end{pmatrix} \begin{pmatrix} \tilde{E}_C(t) \\ 0 \end{pmatrix} \hat{y} = \frac{1}{\sqrt{2}} \begin{pmatrix} \tilde{E}_C(t) \\ j\tilde{E}_C(t) \end{pmatrix} \hat{y}, \quad (3.66a)$$

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & j \\ j & 1 \end{pmatrix} \begin{pmatrix} \tilde{E}_{LO}(t) \\ 0 \end{pmatrix} \hat{x} = \frac{1}{\sqrt{2}} \begin{pmatrix} \tilde{E}_{LO}(t) \\ j\tilde{E}_{LO}(t) \end{pmatrix} \hat{x}. \quad (3.66b)$$

Hence, at the outputs of the MMI, the following signals should exist

$$\frac{1}{\sqrt{2}} (\tilde{E}_C(t) \hat{y} + \tilde{E}_{LO}(t) \hat{x}), \quad (3.67a)$$

$$\frac{j}{\sqrt{2}} (\tilde{E}_C(t) \hat{y} + \tilde{E}_{LO}(t) \hat{x}). \quad (3.67b)$$

For coherent detection, the LO and the carrier need to be mixed with the same polarization, frequency, and phase. To that end, the first output from the MMI i.e. equation (3.67a) is sent to a PBS which produces the following outputs

$$\frac{1}{\sqrt{2}} [\hat{y} \tilde{E}_C(t) \sin\left(\frac{\Delta\phi_{TE}}{2}\right) + \hat{x} \tilde{E}_{LO}(t) \sin\left(\frac{\Delta\phi_{TM}}{2}\right)] \text{ at port 1}, \quad (3.68a)$$

$$\frac{1}{\sqrt{2}} [\hat{y} \tilde{E}_C(t) \cos\left(\frac{\Delta\phi_{TE}}{2}\right) + \hat{x} \tilde{E}_{LO}(t) \cos\left(\frac{\Delta\phi_{TM}}{2}\right)] \text{ at port 2}. \quad (3.68b)$$

The PBS needs to be tuned in such a way that output from the upper arm is only $\hat{x} \frac{1}{\sqrt{2}} \tilde{E}_{LO}(t)$ ⁶ and from the lower arm is $\hat{y} \frac{1}{\sqrt{2}} \tilde{E}_C(t)$. The LO signal is then converted from TM to TE and then mixed with another MMI

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & j \\ j & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} \hat{y} \tilde{E}_{LO}(t) \\ \hat{y} \tilde{E}_C(t) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \tilde{E}_{LO}(t) + j\tilde{E}_C(t) \\ j\tilde{E}_{LO}(t) + \tilde{E}_C(t) \end{pmatrix}. \quad (3.69)$$

The electric fields of the two optical field waves impinge on the photodiodes producing electrical currents

$$I_1 = \mathcal{R} \frac{1}{4} \left| \tilde{E}_{LO}(t) + j\tilde{E}_C(t) \right|^2, \quad (3.70a)$$

$$I_2 = \mathcal{R} \frac{1}{4} \left| j\tilde{E}_{LO}(t) + \tilde{E}_C(t) \right|^2. \quad (3.70b)$$

⁶because in our design the polarization converter is on the upper arm as in Figure (3.2)

Here, $\mathcal{R}$ corresponds to the responsivity of the photodetector. For balanced detection, the resulting photocurrents are subtracted with an electronic adder/ (subtractor), yielding

$$I_2 - I_1 = I_Q = \mathcal{R} \text{Im}\{\tilde{E}_C(t)\tilde{E}_{LO}^*(t)\} \propto m_j \sin(\Delta\phi + \theta_j) \quad (3.71)$$

Therefore, the quadrature-phase component of the complex message envelope can be recovered. In theory, for coherent detection both the polarization and the phase of the LO and the carrier need to be in perfect alignment. With the advancement of digital signal processing, it is possible to have this alignment post detection via phase compensation by setting $\Delta\phi = 0$.

The other output of the first MMI i.e. equation (3.67b) is passed into a second PBS which produces a set of output similar to equation (3.68a) and (3.68b), but with an additional j factor. In this case the polarization converter is tuned in such a way that the output from the upper arm is only $\hat{y}\frac{j}{\sqrt{2}}\tilde{E}_C(t)$ and the the output from the lower arm is only $\hat{x}\frac{j}{\sqrt{2}}\tilde{E}_{LO}(t)$ ⁷. The LO signal is first converted from TM to TE with a polarization converter and then phase-shifted by $\frac{\pi}{2}$

$$\hat{x}\tilde{E}_{LO}(t) \xrightarrow{\text{Pol. Conv.}} \hat{y}\tilde{E}_{LO}(t) \xrightarrow[\text{by } 90^\circ]{\text{phase shift}} \hat{y}j\tilde{E}_{LO}(t). \quad (3.72)$$

The phase shifting of the LO by $\pi/2$ is required to recover both the quadratures. The second MMI then mixes both the LO and the carrier to produce

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & j \\ j & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} j\tilde{E}_C(t) \\ -\tilde{E}_{LO}(t) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} j\tilde{E}_C(t) - j\tilde{E}_{LO}(t) \\ -\tilde{E}_C(t) - \tilde{E}_{LO}(t) \end{pmatrix}. \quad (3.73)$$

The electric fields of the two optical field waves impinge on the photodiodes producing electrical currents

$$I_3 = \mathcal{R} \frac{1}{4} \left| \tilde{E}_C(t) - \tilde{E}_{LO}(t) \right|^2, \quad (3.74a)$$

$$I_4 = \mathcal{R} \frac{1}{4} \left| \tilde{E}_C(t) + \tilde{E}_{LO}(t) \right|^2. \quad (3.74b)$$

For balanced detection, the resulting photocurrents are subtracted with an electronic adder/(subtractor), yielding

$$I_4 - I_3 = I_I = \mathcal{R} \text{Re}\{\tilde{E}_C(t)\tilde{E}_{LO}^*(t)\} \propto m_j \cos(\Delta\phi + \theta_j). \quad (3.75)$$

Therefore the in-phase component of the complex message envelope can be recovered. In theory, for coherent detection both the polarization and the phase of the LO and the carrier need to be in perfect alignment. With the advancement of digital signal processing, it is possible to have this alignment post detection via phase compensation by setting $\Delta\phi = 0$.

⁷because in our design the polarization converter is on the lower arm as in Figure (3.2)

3.5 Circuit Level Schematics

This section illustrates the functional behavior of the building blocks we have described in section (3.3) with the help of a circuit level simulator namely Lumerical Interconnect. We have basically run time domain simulations, where the integrated optical components employed in our CV QKD design are put to virtual tests. The simulation results help us extract useful performance parameters about these components. In a way, the analysis performed in this current section tries to simulate the upcoming experiment with the fabricated chips.

Distributed feedback laser: The schematic for the set-up to characterize a DFB laser can be visualized as shown in the Figure (3.19).

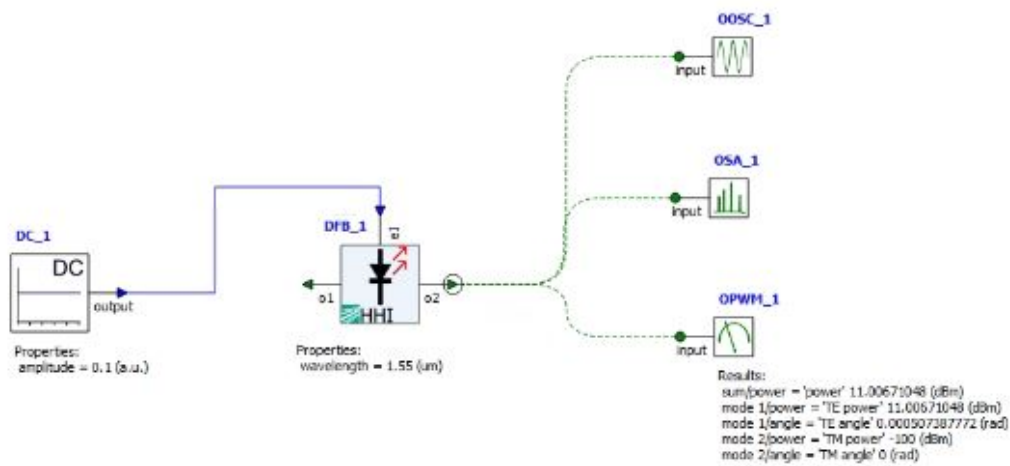


FIGURE 3.19: Schematic set-up for the characterization of an on-chip DFB laser

The DFB laser has two optical output ports and one electrical input port. A dc source is supplying the necessary pump current so that the laser diode can operate beyond the threshold level. The OOSC denotes an optical oscilloscope, the OSA denotes an optical spectrum analyzer, and the OPWM denotes an optical power meter. It is to be noted that, although the optical power meter is collecting optical output only from one side, the reading on it shows the double-sided optical output power. For the particular example shown in the Figure (3.19), the dc bias current is 100 mA. At this bias current, the optical output power at the target wavelength of 1550 nanometer is 11.0067 dBm (≈ 12.608 mW). Therefore, the single-sided output should be approximately 6 milliwatts. Moreover, the emitted light is predominantly TE polarized with barely any TM component.

Figure (3.21) shows the optical waveform of the DFB laser under characterization as seen on the optical oscilloscope. Figure (3.21) shows a parametric sweep: how the optical output power of the laser varies with respect to the increasing bias current. From this figure, the threshold current is approximately 10 mA at 300K temperature. Figure (3.22) shows the power spectral density of the on-chip laser at 300K at two different bias currents namely 80 mA and 100 mA. It

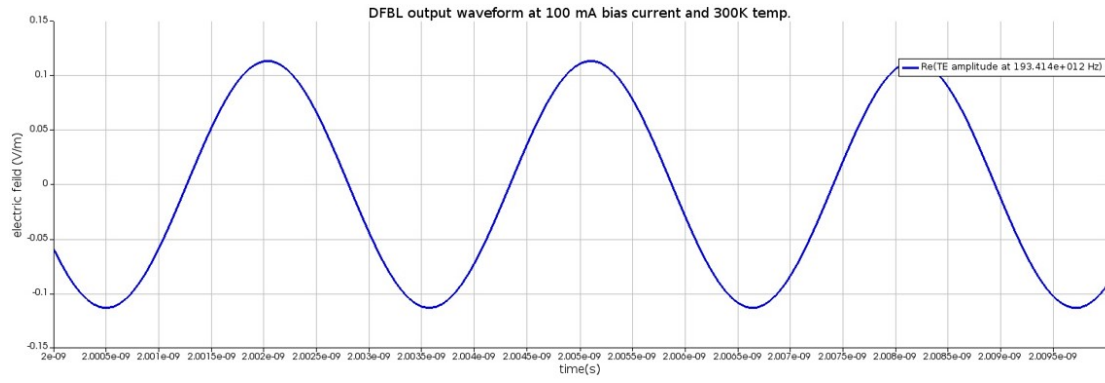


FIGURE 3.20: The output waveform from the on-chip DFBL at the target wavelength of 1550 nanometer at 100 mA bias current at 300K temperature.

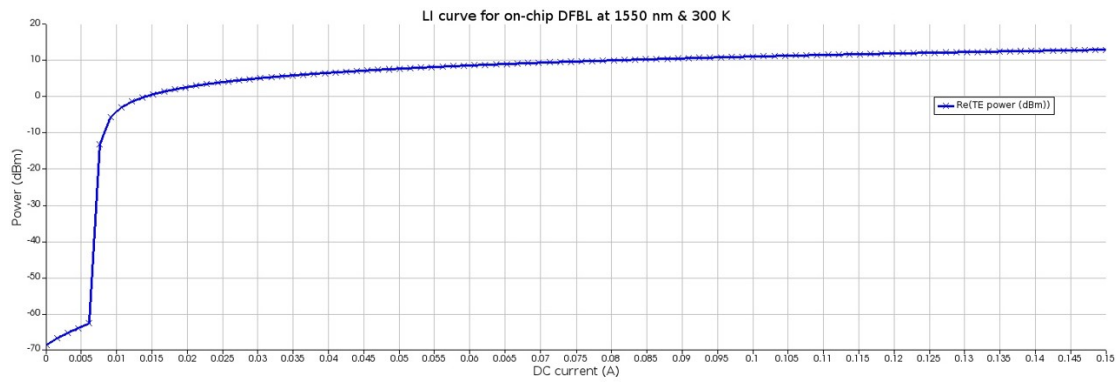


FIGURE 3.21: Output power vs bias current (LI) characteristics of the on-chip DFBL at the target wavelength of 1550 nanometer at 300K temperature.

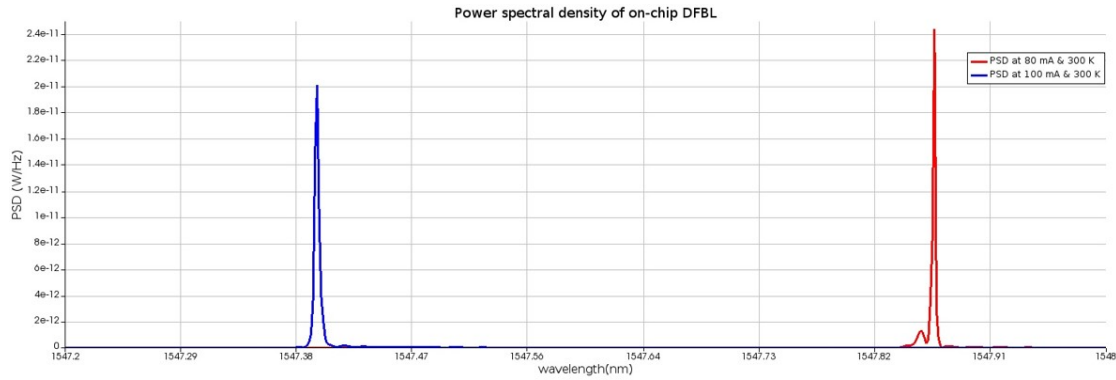


FIGURE 3.22: Power spectral density of the on-chip DFBL at the target wavelength of 1550 nanometer at two different bias currents at 300K temperature.

can be seen that in both the cases the measured peak wavelength has a negative deviation of ~ 2 nm from the target wavelength of $1.55 \mu\text{m}$. In reality, under operating conditions, bias current generates self-heating, which leads to the red-shift of the emitted wavelength i.e. the measured peak wavelength shifts to the right of the target wavelength. In this sense, self-heating can be used as a tuning mechanism. However, self-heating can also contribute to the degradation of lasing (peak at 100 mA is lower than at 80 mA) as seen from the same

figure.

Multi-mode interferometer: Figure (3.23) shows the schematic for characterizing an MMI building block.

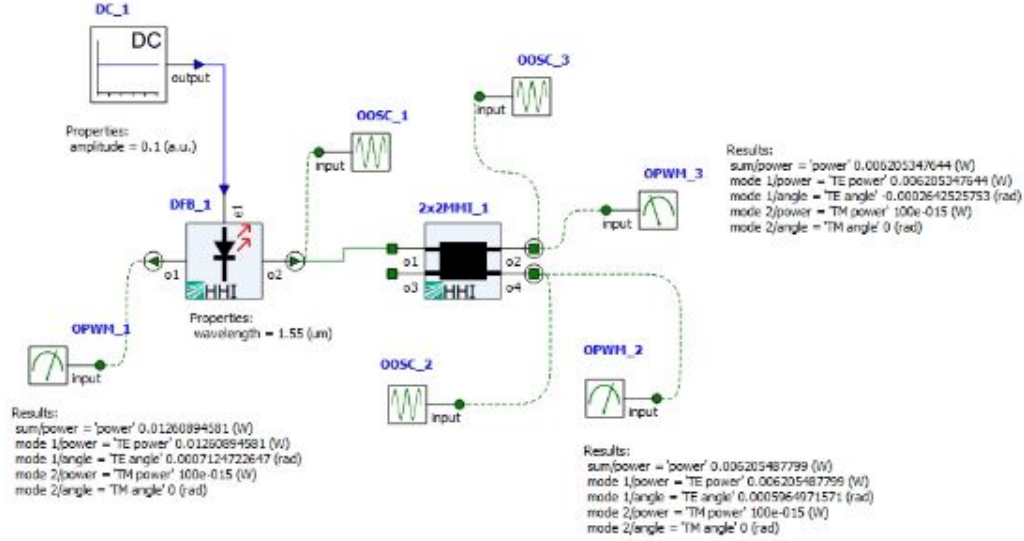


FIGURE 3.23: Schematic set-up for the characterization of a photonic integrated MMI

From the optical power meter readings in Figure (3.23), the optical power input to the MMI is almost equally split at its cross-output and at its bar-output. The imbalance in the input power and the resultant output power can be attributed to the insertion loss of the building block (which is < 1 dB here). Moreover, examining the waveforms from the two output ports of the MMI, as shown in Figure (3.24), reveals that the two outputs are phase-shifted by 90° . Therefore, the MMI is an integrated equivalent of a 2×2 beam splitter.

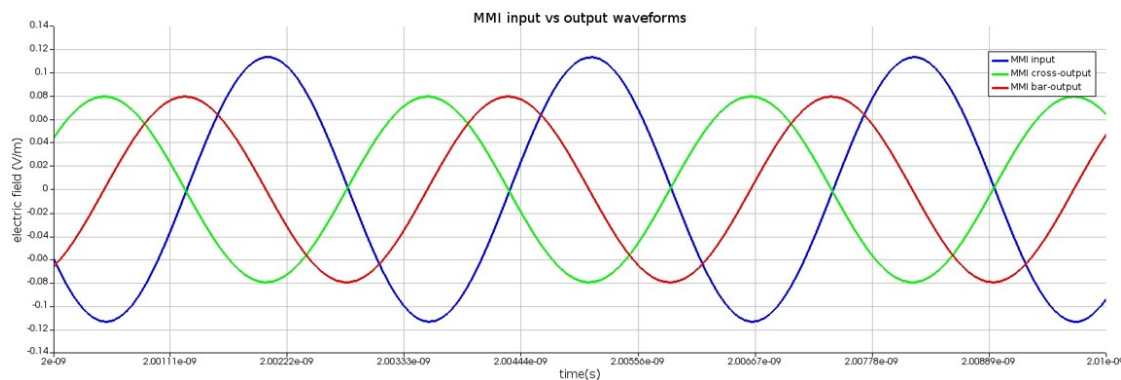


FIGURE 3.24: The two output waveforms of an MMI with respect to its input.

Figure (3.24) also illustrates that the bar-output is leading the cross-output by $\frac{\pi}{2}$ and the cross-output is completely out-of-phase with the MMI's input. Taking these two observations into consideration suggest that the optical transfer

power matrix of a 2×2 integrated MMI should be

$$T_{\text{MMI}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & j \\ j & 1 \end{pmatrix}, \quad (3.76)$$

which is the same as equation (3.7) including the global phase. Consequently, when two MMI's are cascaded, almost all the input power is available at the global cross-output with non-significant power at the global bar-output. Figure (3.25) points to this fact.

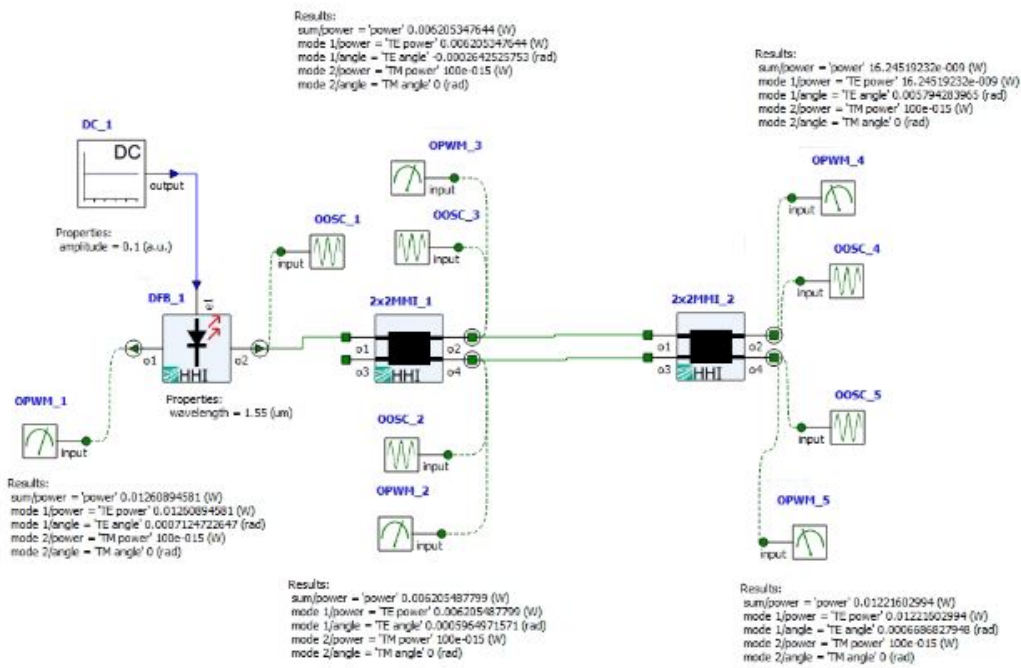


FIGURE 3.25: Schematic set-up for the characterization of two sequential MMIs.

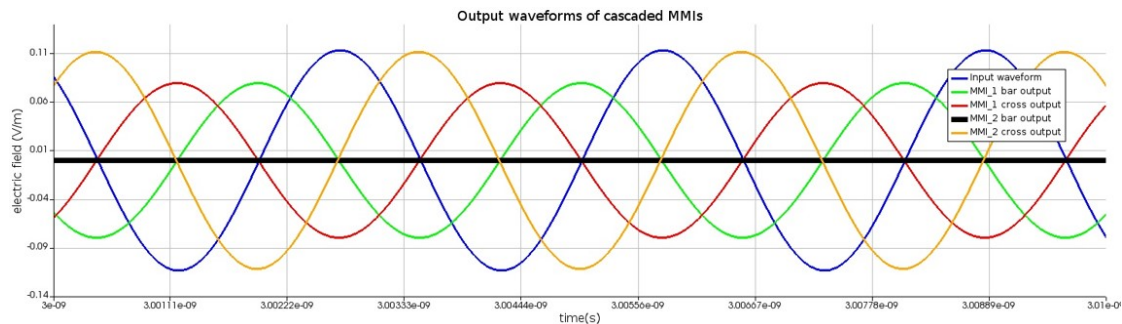


FIGURE 3.26: Input vs output waveforms when two MMI's are placed sequentially

Figure (3.26) shows the input vs output MMI-waveforms in the cascaded configuration. Evidently, the global bar-output (the black curve) almost coincides

with the null power level. On the other hand, the global cross-output (the orange curve) is almost at the same power level with the input but with a relative $\frac{3\pi}{2} (= -90^\circ)$ phase shift. This is a further justification of the expression we wrote in equation (3.76) for the optical transfer matrix of our photonic integrated MMI.

Tunable Beam Splitter: The tunable beam splitter (TBS) is a composite building block, which in its expanded looks like as shown in Figure (3.27). As a whole, the building block has four optical ports and two electrical ports. The relays indicate ports for interfacing the device externally with other devices.

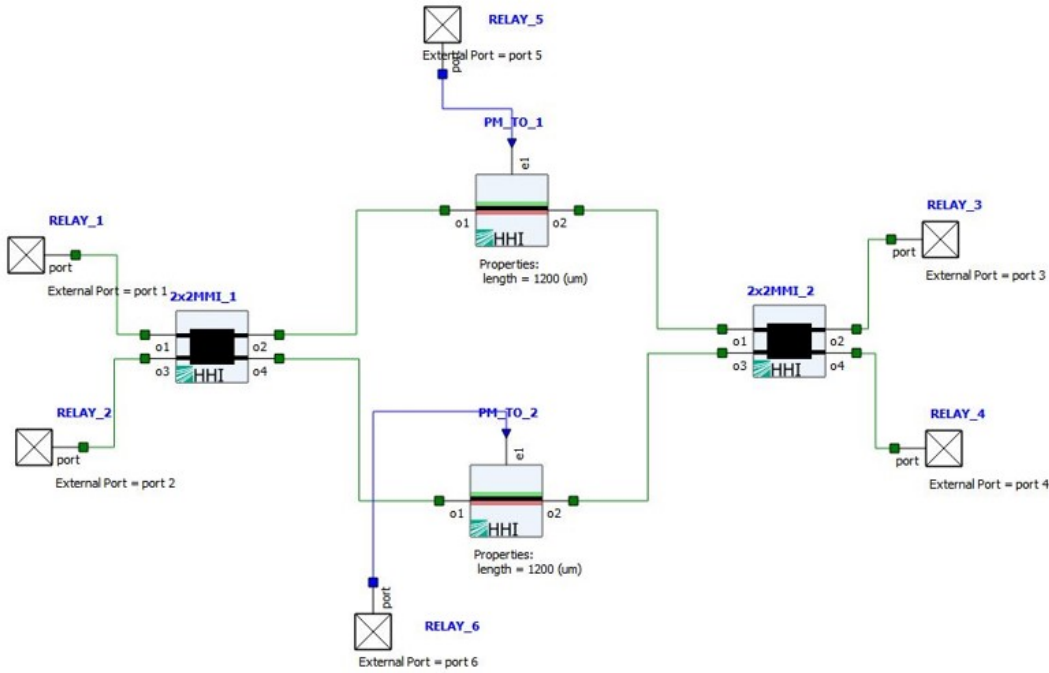


FIGURE 3.27: Schematic view of the tunable beam splitter consisting of an input MMI, an output MMI, and two thermo-optic phase shifters on both the connecting arms.

Figure (3.28) shows the schematic for a TBS characterization. By tuning the phase difference between the two optical waves propagating through the connecting arms, one can, in principle, achieve arbitrary splitting ratio. Since only the phase difference matters, it suffices to operate only one heater.

If none of the heaters is operated, i.e. the phase difference is identical for both the optical waves, the entire input power will appear at the cross-output of the TBS barring some insertion loss. On the other hand, if one of the heaters is operated such that it induces a π phase shift on the wave travelling across it, the situation reverses i.e. all the optical power will appear at the bar-output. At $\frac{\pi}{2}$ phase difference, there is equal transmission. Hence, by controlling the phase difference within the interval $[0, \pi]$, one can adjust the optical transmission to any of the output ports of the TBS between 0 and 1. Figure (3.29) shows the power distribution at the two output ports of the TBS as the control current is gradually increased from 0. To cite three specific examples, at 3.7 mA, the power

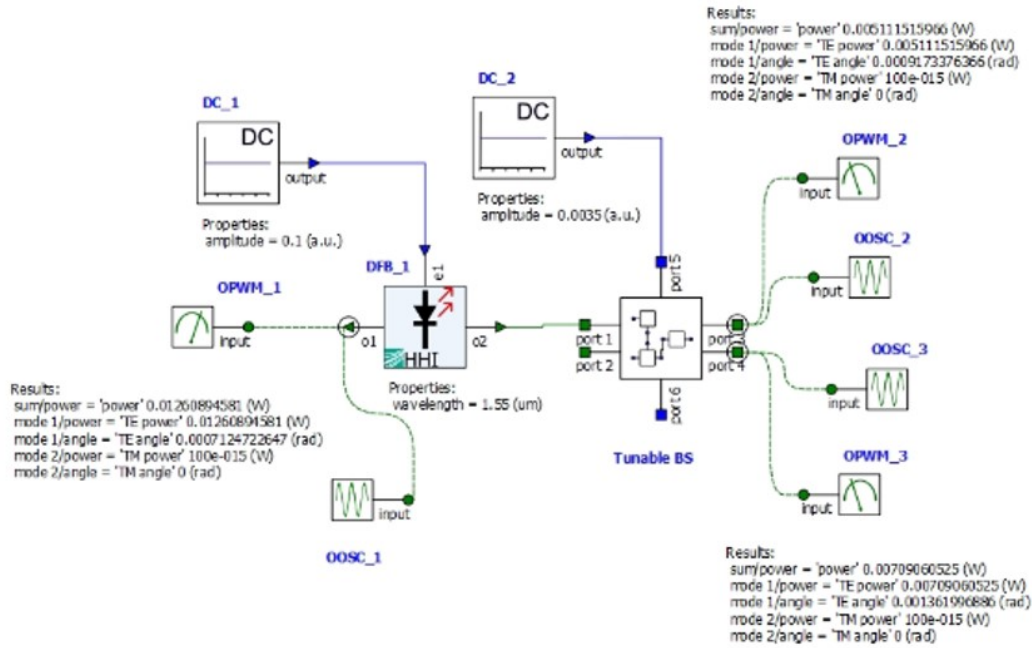


FIGURE 3.28: Schematic set-up for the characterization of a tunable beam splitter. The expanded form of the Tunable BS component is shown in the previous figure.

splitting ratio is $\approx 50 : 50$, at 2 mA the ratio is $\approx 95 : 5$, and at 2.5 mA the ratio is $\approx 87.5 : 12.5$.

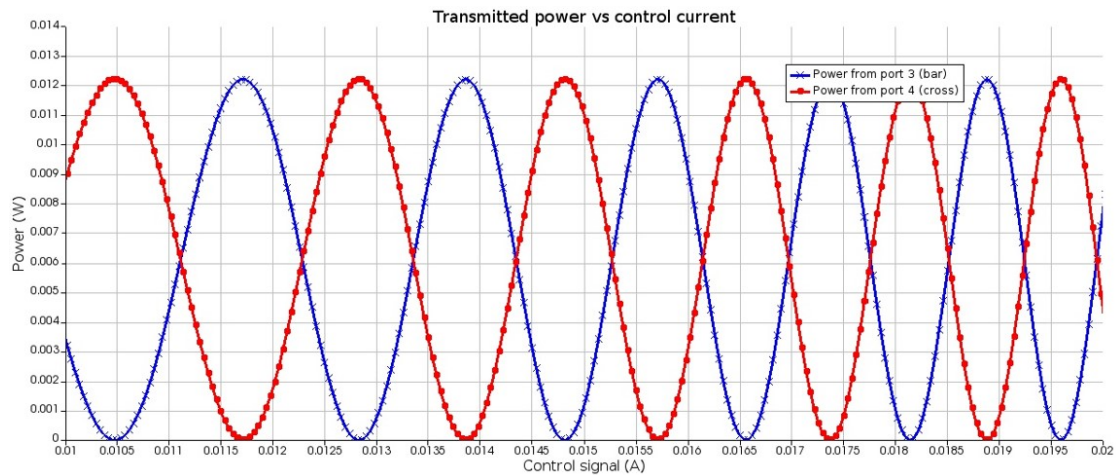


FIGURE 3.29: Interference pattern obtained at the two output ports of the TBS subject to the variation in the control current for the phase shifter on the upper arm.

Polarization Converter: Figure (3.30) shows a possible setup to determine the polarization extinction ratio of a polarization converter (PC). The reading from the optical polarization analyzer allows us to calculate the polarization extinction ratio of our integrated PC (≈ 20 dB) at the particular wavelength of the DFB. However, the building block is not perfect. It cannot quite convert the

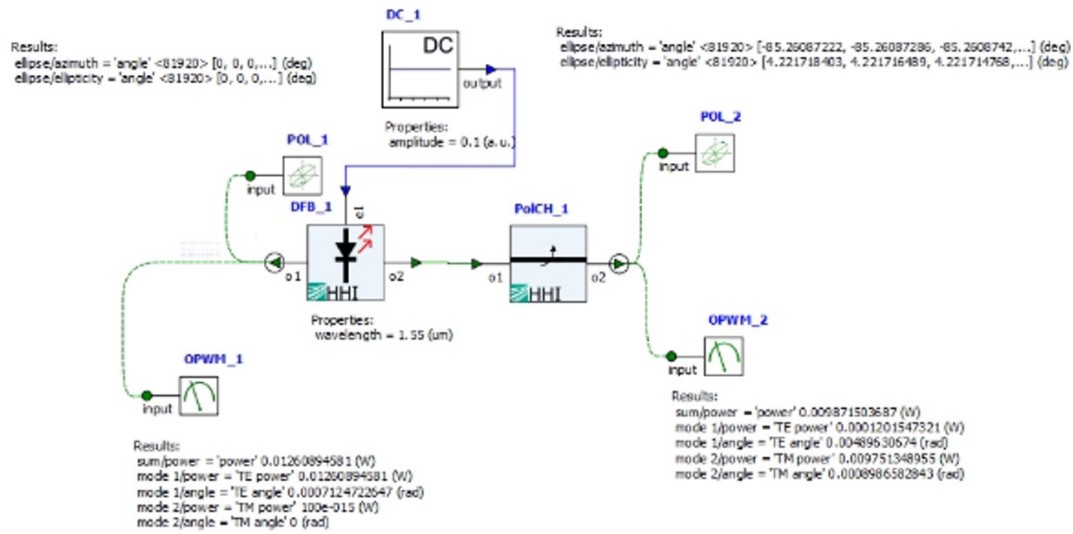


FIGURE 3.30: Schematic set-up for the characterization of a polarization converter. POL stands for optical polarization analyzer.

TE-polarized light emitted from the DFBL into TM, but it comes close. The output from the PC has an azimuth angle of -85.25° (instead of exactly -90°). Figure (3.31) illustrates the performance of our integrated equivalent of half-wave plate within the emitted wavelength range of the DFBL.

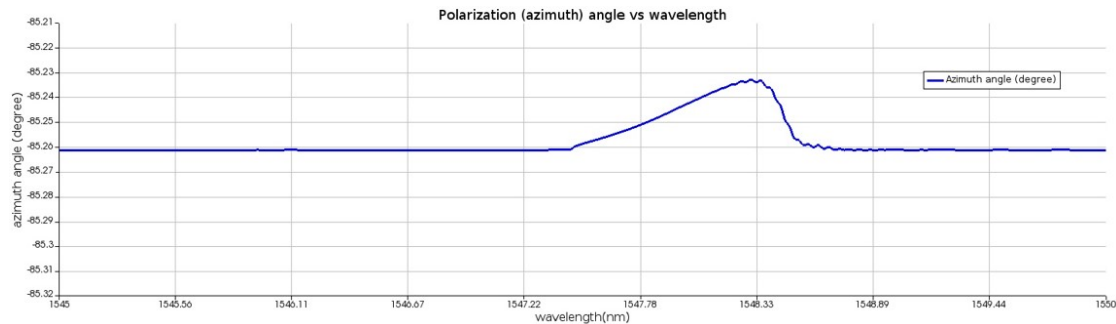


FIGURE 3.31: Polarization rotation by the polarization converter within a certain wavelength range.

Polarization Beam Splitter: Let us consider the set up schematically described in Figure (3.32). A DFB laser is excited to emit TE-polarized light at the target wavelength of 1550 nm. The emitted light is then split in equal halves by a 2×2 MMI. For one half, a polarization converter is placed on its way so that its TE polarization state is switched to TM.

Both the TE-polarized and the TM-polarized light are subsequently combined by the PBS. Structurally, the PBS is similar to a Mach-Zehnder Interferometer as we have seen in section (3.3). The phases of the TE and TM optical waves,

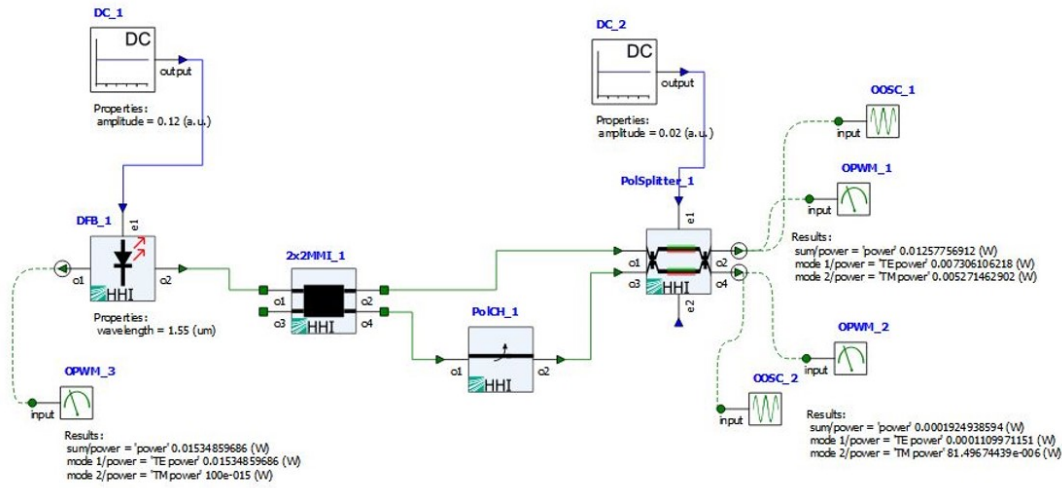


FIGURE 3.32: Schematic set-up to analyze the behavior of a polarization beam splitter. Operation of only one heater suffices to produce the necessary phase differences.

therefore, can be controlled by operating the thermo-optic phase shifters placed on top of the arms connecting the input and output MMIs. We are interested in observing the relative distribution of the TE and TM modes at the output ports of the PBS with respect to this change in phase difference. Figures (3.33), (3.34), (3.35), (3.36) illustrate power distributions obtained for both polarization states with respect to a PBS control current sweep.

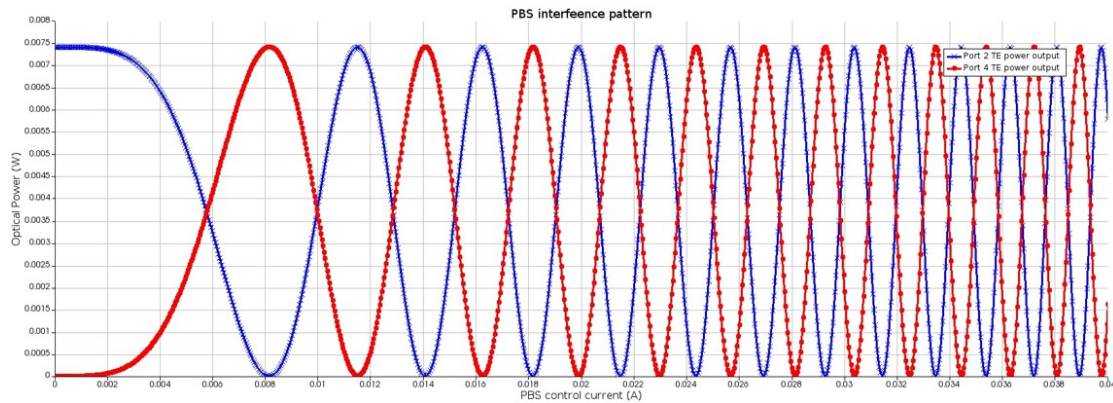


FIGURE 3.33: Interference pattern obtained at the output ports of the PBS for the TE wave subject to the variation in the phase shifter control current.

The above graphical analysis is useful when a PBS is combining (polarization multiplexing) the TE-polarized carrier signal and the TM-polarized LO signal on the transmitter side for sending them together through the fiber. On the receiver side, the PBS is used to separate out the two differently polarized light from the combined signal for the subsequent balanced detection. This case is schematically visualized in Figure (3.37).

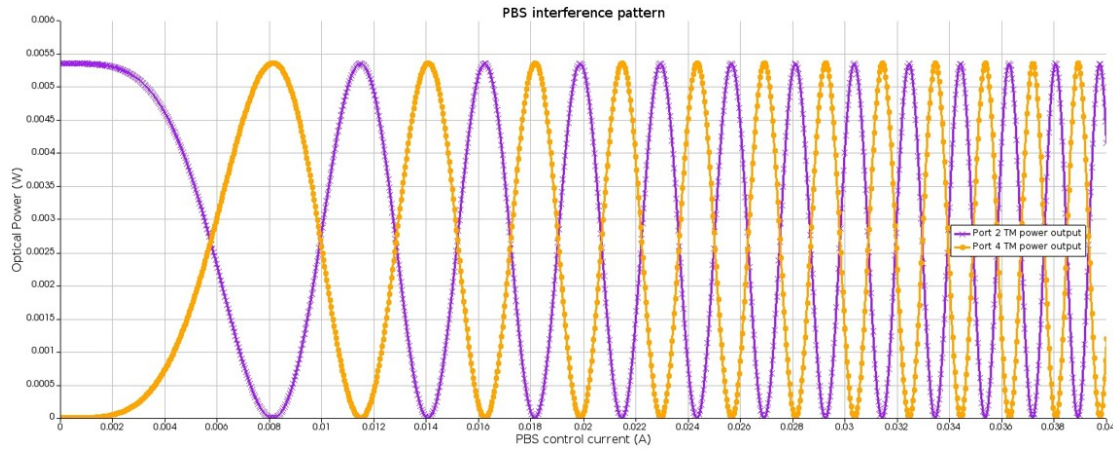


FIGURE 3.34: Interference pattern obtained at the output ports of the PBS for the TM wave subject to the variation in the phase shifter control current.

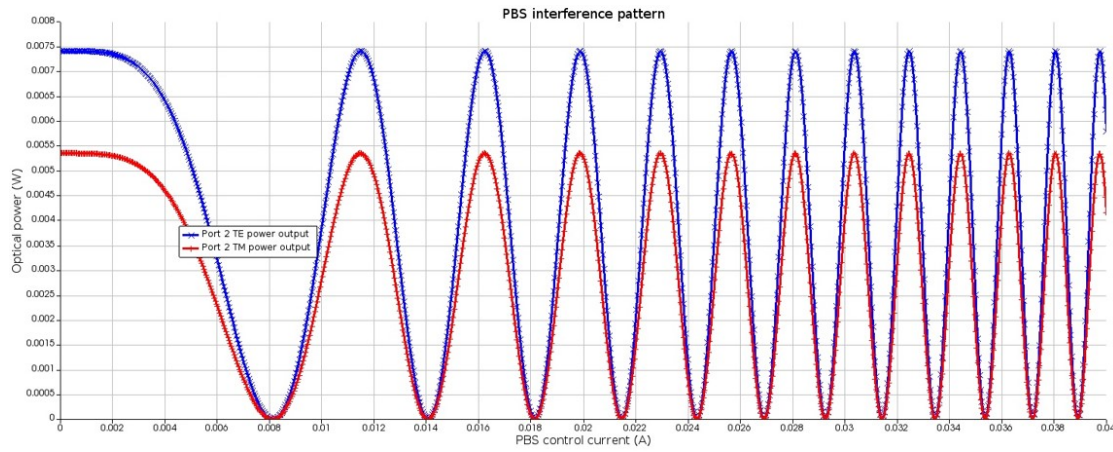


FIGURE 3.35: Relative distribution of the TE and TM waves at Port 2 of the PBS subject to the variation in the phase shifter control current.

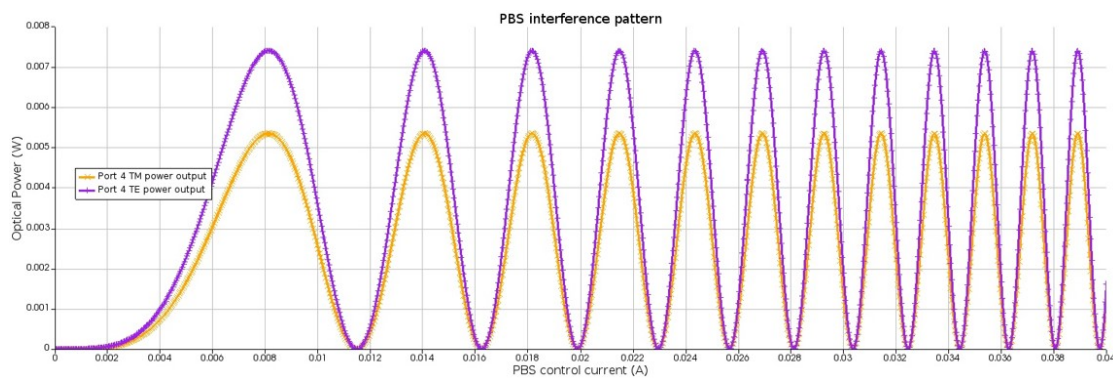


FIGURE 3.36: Relative distribution of the TE and TM waves at Port 4 of the PBS subject to the variation in the phase shifter control current.

In this figure, we have made use of results from Figure (3.36). When the value of the bias current for the first PBS is 8 mA, both TE and TM polarizations arrive at output port 4 with maximum amplitude. Now, if this output signal is input to a second PBS with zero control current, it will split the combined signal into

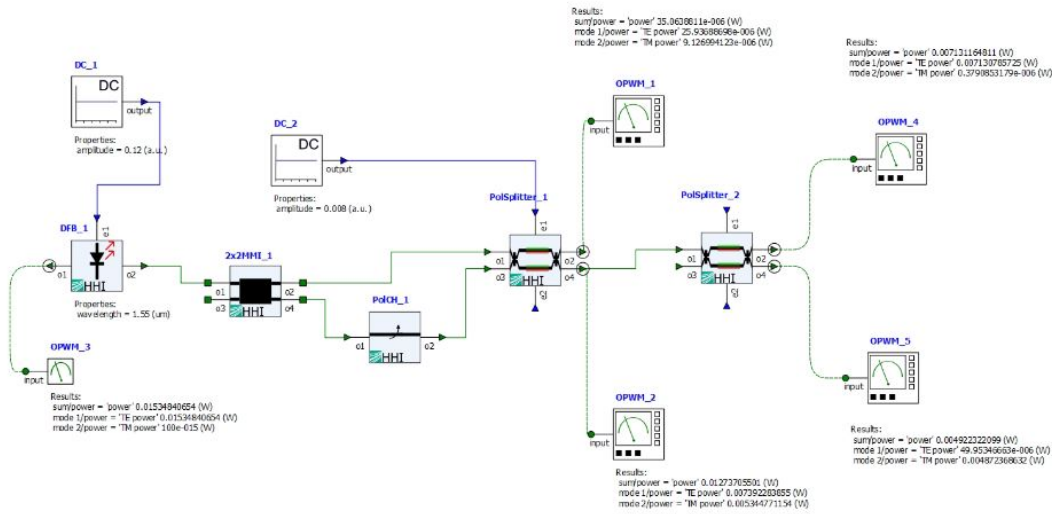


FIGURE 3.37: Schematic showing a PBS acting both as a polarization multiplexer (the first one from the left) and a demultiplexer (the second one).

predominantly a TE component (port 2) and a predominantly TM component (port 4). The discrepancy in the input and the output total power is due to the

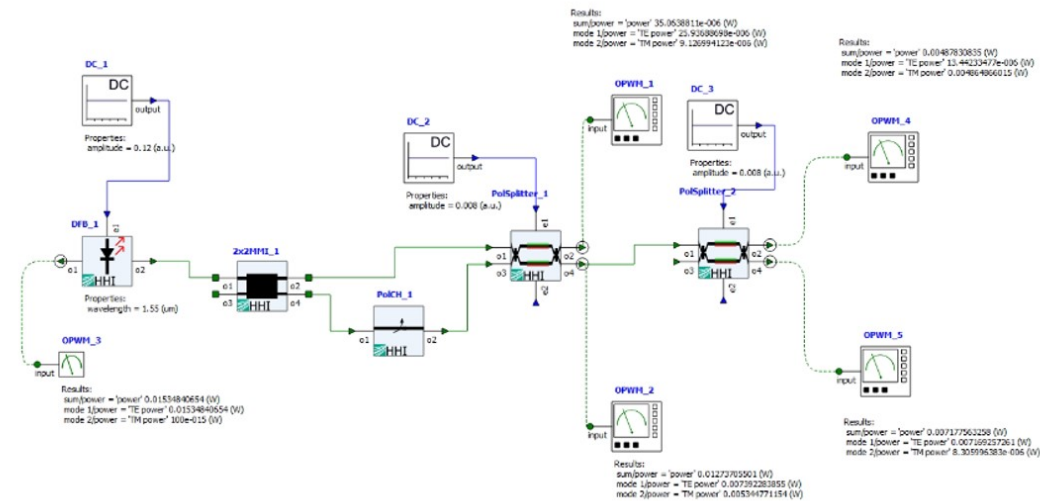


FIGURE 3.38: Schematic showing the reverse situation to the one shown in figure (3.37)

insertion loss (≈ 2 dB) of the device. We could have swapped the arrival points of the TE and TM components by simply changing the bias current of the second PBS from 0 to 8 mA (Figure (3.38)). Figure (3.39) shows the output waveforms from the second PBS as a function of bias current. It is to be noted that for both the output ports, when power in TE mode is maximum, then power in TM mode is minimum. This is in sharp contrast to the the case when the PBS was

acting as a combiner (Figures (3.35), (3.36)) where power in TE and TM modes were peaking at the same time at a particular output port.

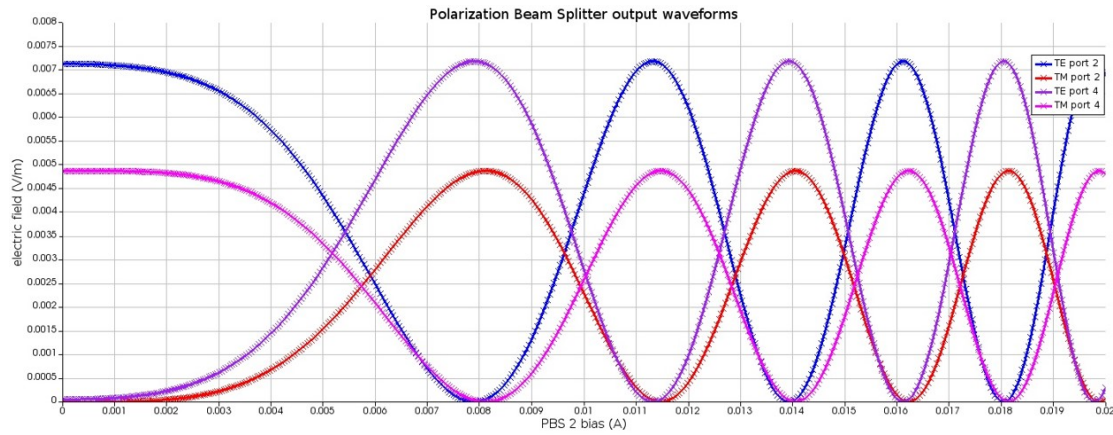


FIGURE 3.39: Relative distribution of the TE and TM waves at at Port 2 and Port 4 of the PBS subject to the variation in the phase shifter control current.

Mach Zehnder Modulator: An integrated Mach Zehnder modulator (MZM) is similar in composition to a tunable beam splitter (Figure (3.40)). In an MZM, however, current injection phase modulators (CIPM) are placed on the connecting arms instead of thermo-optic phase shifters (TOPS), which are too slow.

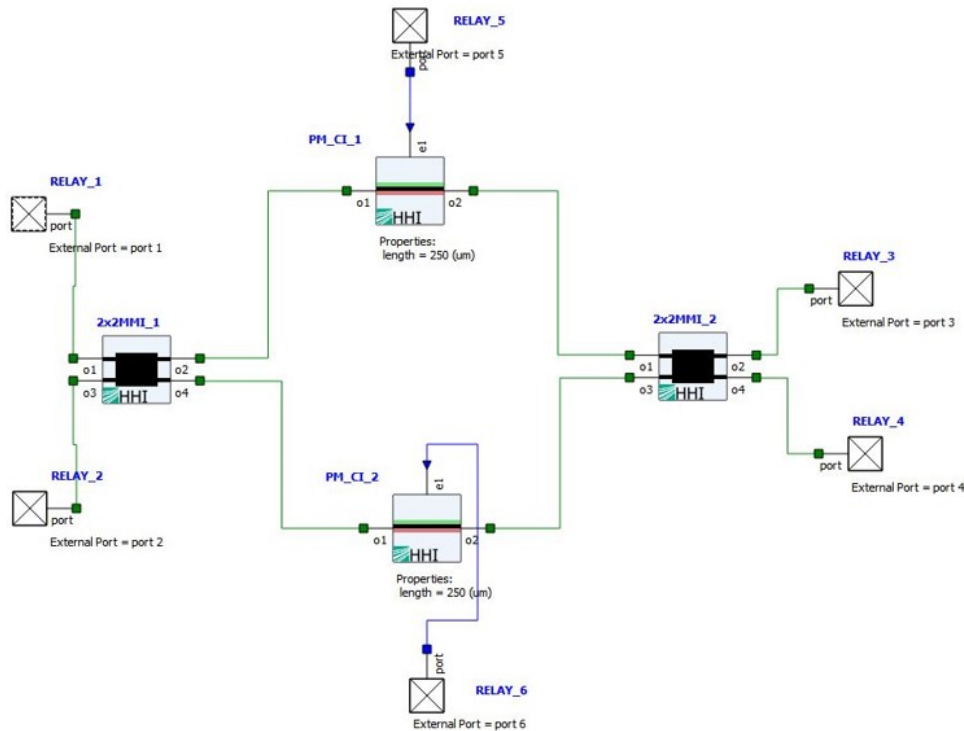


FIGURE 3.40: A Mach Zehnder modulator is composed of an input MMI, an output MMI, and two current injection phase modulators (CIPM) on both the connecting arms.

The characteristic parameter of an MZM is its $V_\pi (/I_\pi)$ voltage (/current). If a CIPM in an MZM is biased at this voltage (/current), it induces a π phase

shift on the optical wave. The implication is that at the global cross-output, the two optical waves propagating through the upper arm and the lower arm interfere destructively. While for zero bias voltage (/current), they interfere constructively.

Figure (3.41) shows the schematic set up to determine the I_π current of an MZM. If the bias current to the MZM is gradually increased from zero, the power distribution at the two output ports will look like in Figure (3.42). In this figure port 3 and port 4 refer to the bar output and cross output respectively. ϕ_1, ϕ_2 are the phase changes experienced by the optical waves on the upper arm and the lower arm respectively. If the CIPM on the lower arm is inactive, then $\phi_2 = 0$. To create the necessary phase difference, it suffices to operate only one CIPM. The point on the graph marked with a star denotes the pi-current with regard to the cross output of our MZM which is $250 \mu\text{m}$ long.

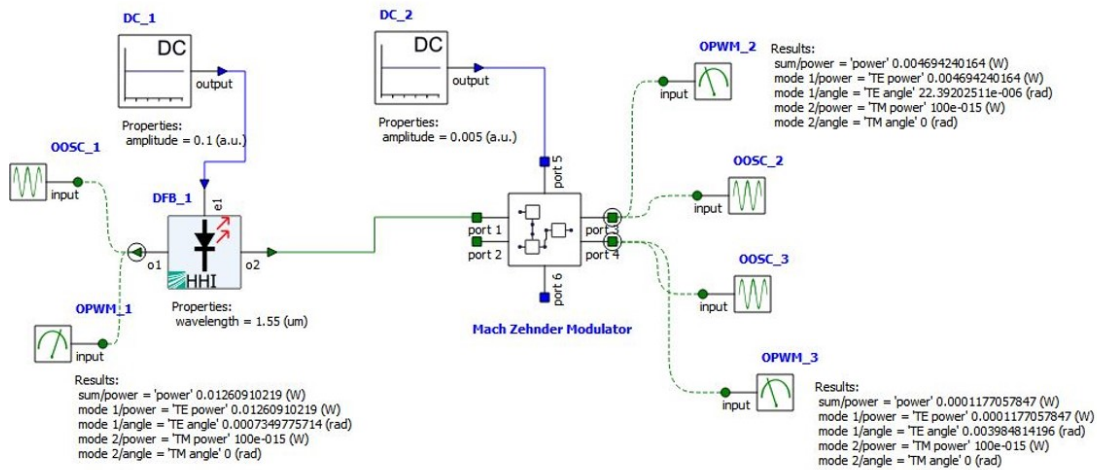


FIGURE 3.41: Schematic set-up for the characterization of an MZM.

From Figure (3.42), it can be clearly seen that the interference pattern obtained from our MZM is not symmetric. This is because our devices are not perfect, the phase modulators in particular.

On-off keying with MZM: One can implement on-off keying by switching the MZM bias current between 0 and I_π . Figure (3.43) shows the relevant circuit schematic. The pseudo random number generator (PRBS) randomly generates sequences of 0's and 1's. The non-return to pulse generator maps these 0's and 1's into electric pulses of 0 mA and I_π mA respectively at a rate of 25 Gbps (bit rate of our system) shown in Figure (3.44).

The optical waveforms obtained from the output ports of the MZM with optical oscilloscopes are shown in Figure (3.45). The red waveform corresponds to the cross output and the blue waveform corresponds to the bar output. The cross outputs and the bar outputs are complementary as they should be. The

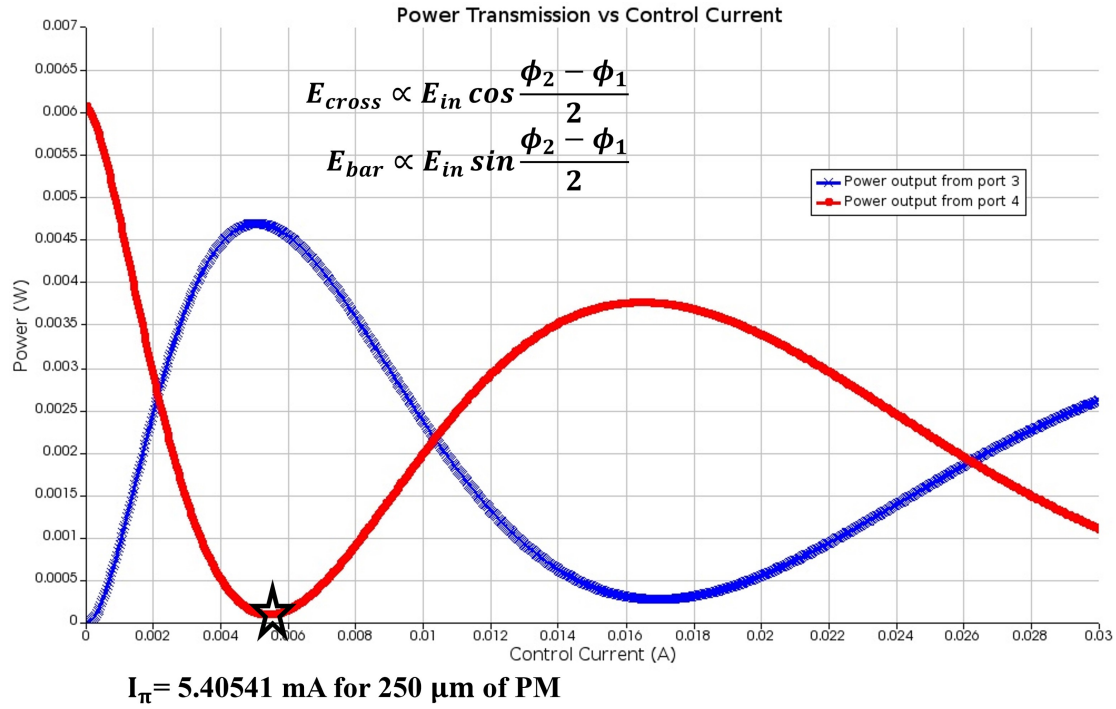


FIGURE 3.42: Interference pattern obtained at the two output ports of the MZM subject to the variation in the control current for CIPM on the upper arm.

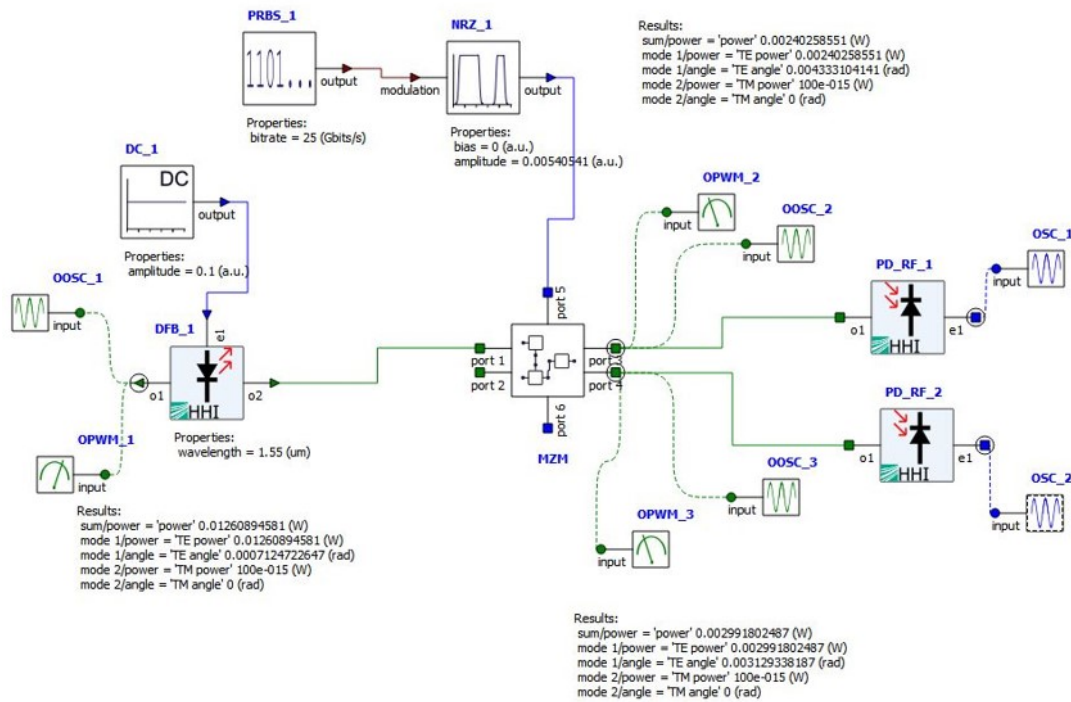


FIGURE 3.43: Schematic showing implementation of intensity modulation and direct detection with an MZM.

extinction ratio is

$$\text{Extinction Ratio} = 10 \log_{10} \left(\frac{P_1}{P_0} \right) = 10 \log_{10} \left(\frac{0.00606631}{9.16058 \times 10^{-05}} \right) = 18.21 \text{ dB} \quad (3.77)$$

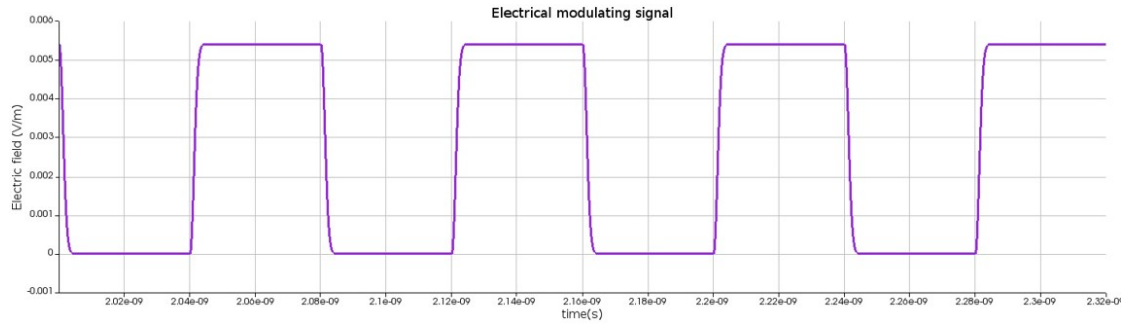


FIGURE 3.44: Non return to zero electrical modulating signal at the rate of 25 Gbps.

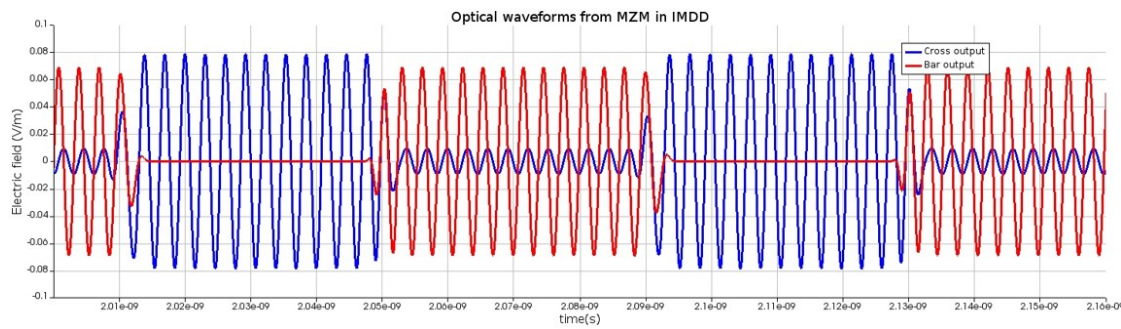


FIGURE 3.45: Optical waveforms observed at the output ports of the MZM corresponding to on off keying operation.

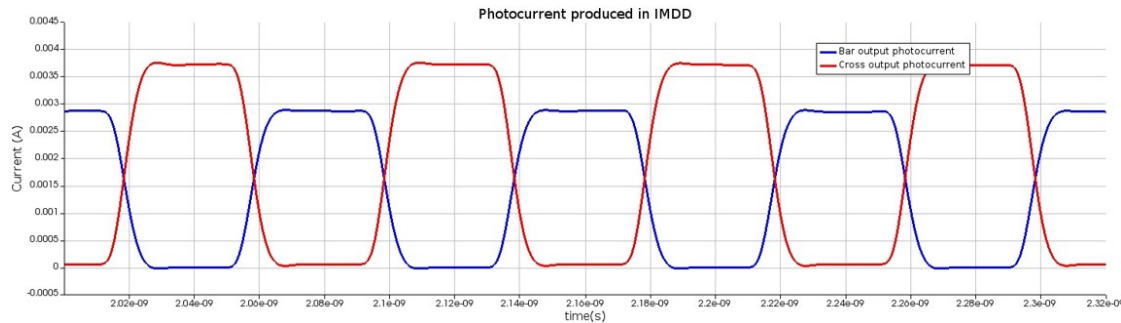


FIGURE 3.46: Photocurrents generated by the photodiodes as observed with an electrical oscilloscope in IMDD.

The photocurrents generated by the photodiodes corresponding to the optical waveforms are shown in the Figure (3.46). Ideally, the complementary waveforms should have been of equal magnitude. But due to imperfection in our device this is not the case.

Current Injection Phase Modulator: Apart from using in the MZI configuration for fast modulation, an integrated phase modulator is also required in the receiver circuit to recover information encoded in both P & Q quadratures of the optical wave. For balanced detection, the information bearing signal has to be mixed with the local oscillator signal. To that end, it is first split into two parts. One part is mixed with the local oscillator as it is. The other part is made to undergo a $\frac{\pi}{2}$ phase shift prior mixing. A phase modulator is employed to impart this $\frac{\pi}{2}$ phase shift. The current required to produce this desired amount

of phase shift can be deduced from equations (3.11a), (3.11b). When $\Delta\phi = \frac{\pi}{2}$, the input power is equally distributed at the two output ports of the MZI. From Figure (3.42), the relevant bias current is found to be 0.00211495 A. Figure (3.47) shows schematically a CIPM being connected with a DFBL. Figure (3.48) shows

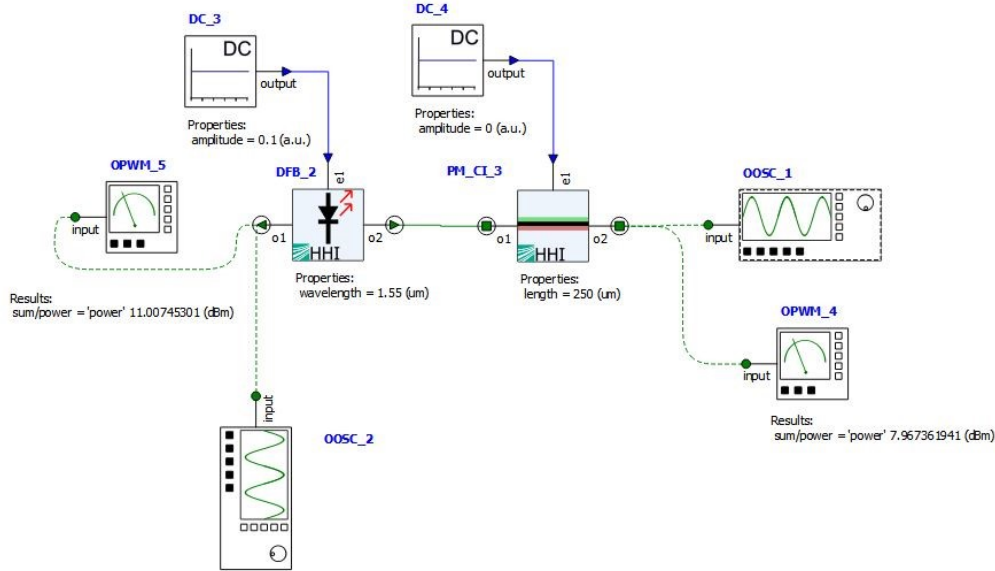


FIGURE 3.47: Schematic showing a current injection phase modulator in conjunction with an on chip distributed feedback laser.

the CIPM output for zero bias current compared to the laser output. From this figure, and also from Figure (3.48), the CIPM has significant insertion loss.

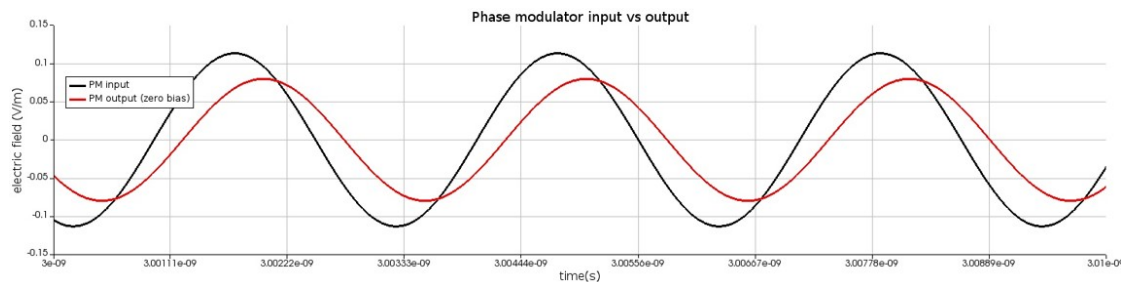


FIGURE 3.48: Input and output waveforms of the CIPM.

Moreover, the CIPM output is phase shifted with respect to its input. This phase shift information might be useful for phase compensation via digital signal processing at the post detection stage. Figure (3.49) shows that the insertion loss of the CIPM increases with the bias current.

Binary phase shift keying with MZM: Figure (3.50) shows the circuit level schematic for BPSK implementation. The black dashed box denotes the MZM with its constituent building blocks. The red dashed box denotes the 2×4 optical hybrid. The optical hybrid is required to mix the local oscillator signal with

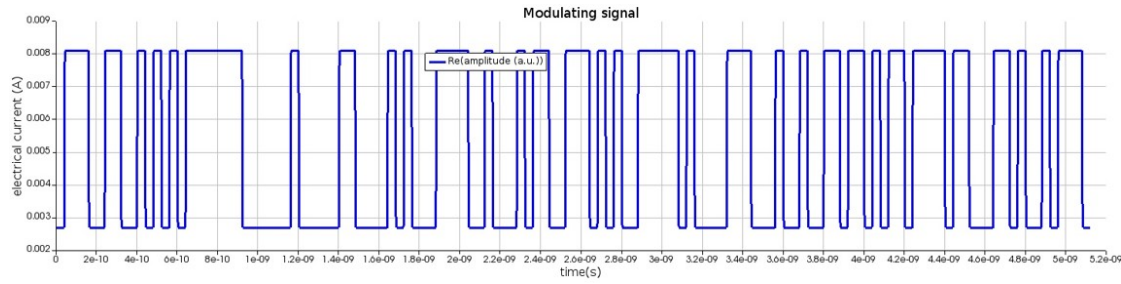


FIGURE 3.51: Non return to zero electrical modulating signal at the rate of 25 Gbps.

The subtracted current from the upper two photodiodes correspond to the Q quadrature, while the subtracted current from the lower two photodiodes correspond to the P quadrature. Figure (3.53) shows the relevant photocurrents. In principle, the Q quadrature components should have been zero. But this is not the case here for two reasons: 1) the optical hybrid is not mixing the LO and the carrier signal with identical phases, 2) the power transmission vs control current graph of our MZM is not symmetric around I_π . The second one is a fundamental limitation while the first one can be taken care of with digital signal processing post detection.

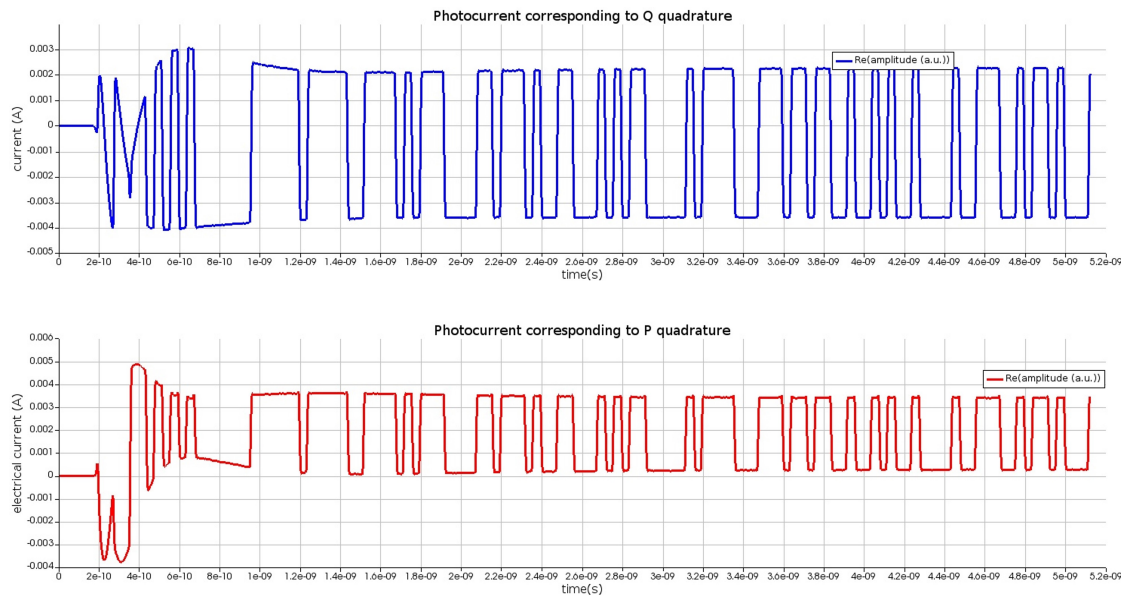


FIGURE 3.52: Photocurrents corresponding to the in-phase and quadrature-phase quadratures of the modulated signal.

The results from the eye diagram analyzers for both the quadratures are shown in Figure (3.53). Eye diagram is a data dependent time domain electrical measurement used to evaluate high speed data quality. Appendix (D) provides an introductory discussion on an eye diagram.

The vector signal analyzer (VSA) provides measurement statistics for both the quadratures. Figure (3.54) shows that the signal statistics (in red) are generally within the decision boundaries (in blue) in the symbol map. The error vector

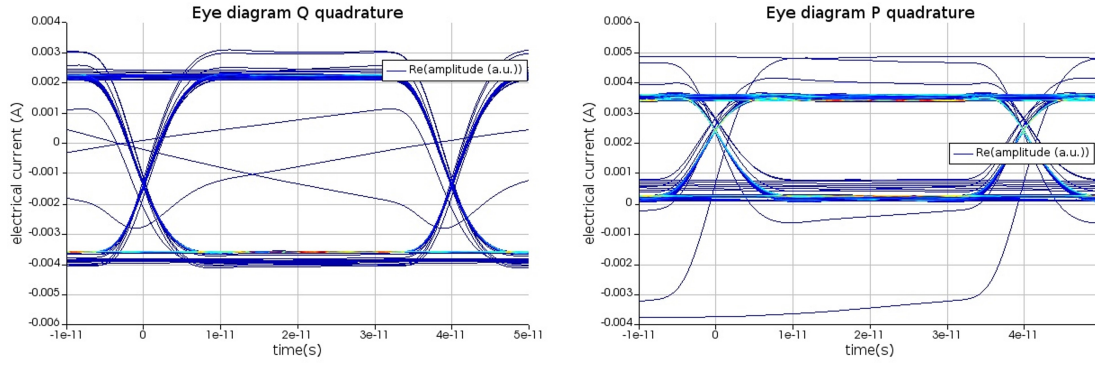


FIGURE 3.53: Eye diagram results for P and Q quadrature in BPSK.

magnitude (EVM) has been found to be $\approx 94\%$, and the symbol error rate (SER) has been found to be 2.828×10^{-36} .

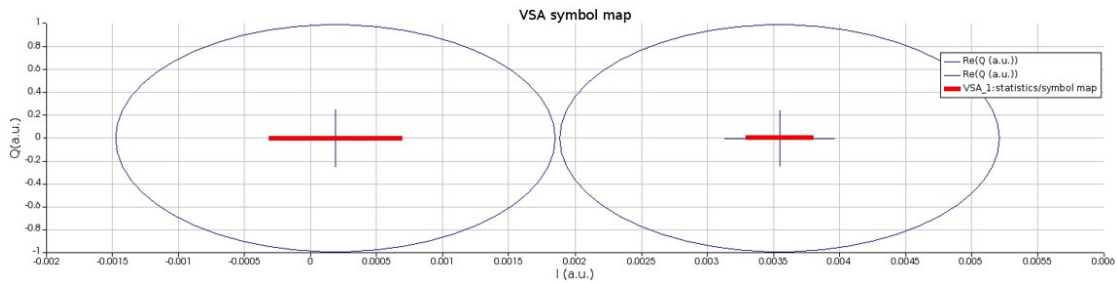


FIGURE 3.54: The signal statistic with respect to the decision boundaries.

Quadrature phase shift keying with MZM: Figure (3.55) shows the circuit level schematic for QPSK implementation. The dashed box denotes the two cascaded MZMs with their constituent building blocks. The first MZM is biased at $\frac{I_\pi}{2}$ with magnitude I_π to provide phase shifts of $\{0, \pi\}$. The second MZM is biased at $\frac{I_\pi}{2}$ with magnitude $\frac{I_\pi}{2}$ to provide phase shifts of $\{0, \frac{\pi}{2}\}$. The overall output is a QPSK-modulated signal.

Figure (3.56) shows the relevant photocurrents. For the constellation points $\{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$, the P and Q quadratures will have three values, for example $(-P, 0, P)$ and $(-Q, 0, Q)$.

The results from the eye diagram analyzers for both the quadratures are shown in Figure (3.57).

Figure (3.58) shows the measurement statistic from the VSA. The detected signals (in red) are generally within the decision boundaries (in blue) in the symbol map. The EVM and SER have been found to be 37.13% and 0.00025, respectively.

Appendix D has the relevant simulation results for 8-PSK.

With the information obtained from the analysis so far, we can now simulate the transmitter and the receiver circuits shown in Figure (3.1) and Figure (3.2). Figure (3.59) shows the circuit level schematic of our transceiver system aimed

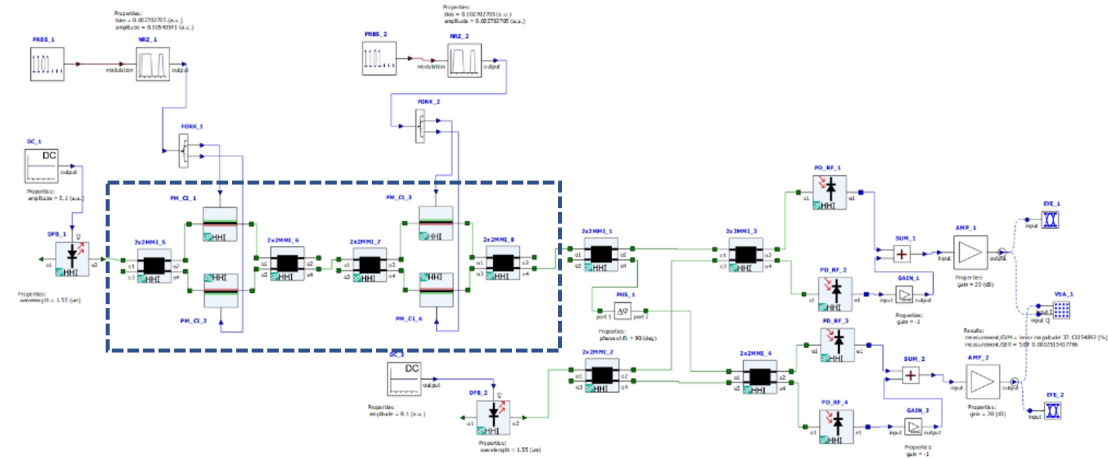


FIGURE 3.55: Schematic illustrating implementation of quadrature phase shift keying with a single MZM in dual drive operation.

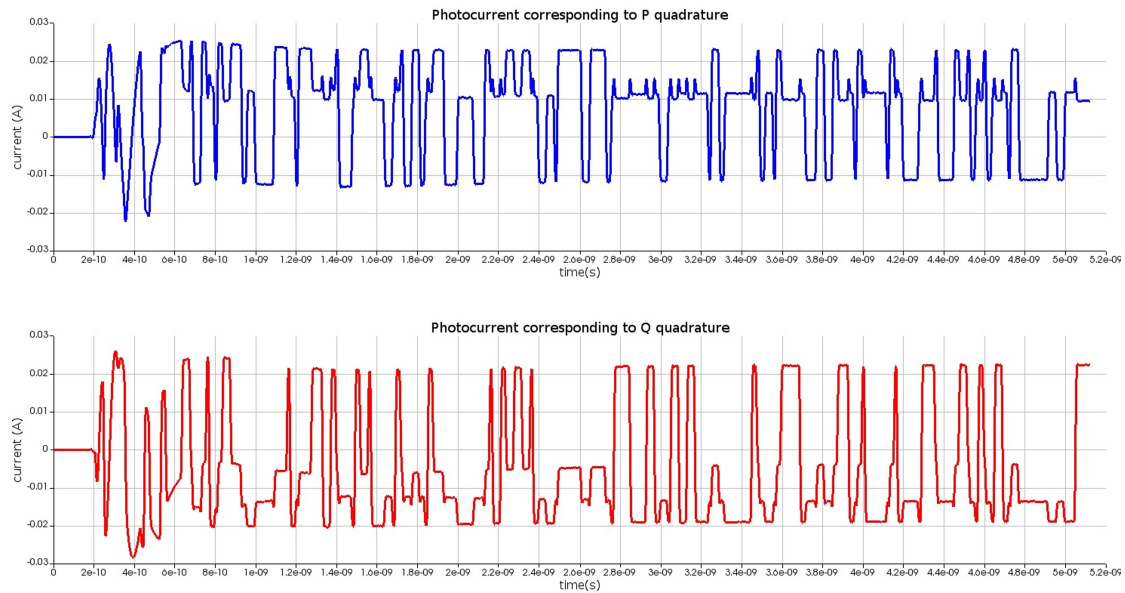


FIGURE 3.56: Photocurrents corresponding to the in-phase and quadrature-phase quadratures of the modulated signal in QPSK.

at implementing CV QKD. This schematic was produced before the mask layout generation. At that stage, it is difficult to foretell the actual orientations and the dimensions of the optical interconnects on the chip. Hence, in this simulation, the building blocks are simply wired without any intermediate waveguide connections. Figure (3.60) shows the circuit level schematic of our transceiver system with all the waveguide dimensions obtained from the mask layout. In both the figures, the modulation scheme is QPSK. The sample rate is 1.6×10^{13} Hz, with 640 samples/bit and a total of 81920 samples. From the optical power meter readings, one can calculate the insertion losses of various components. Needless to say, the average optical power of the carrier signal leaving the transmitter chip in Figure (3.59), will appear more than the one shown in Figure

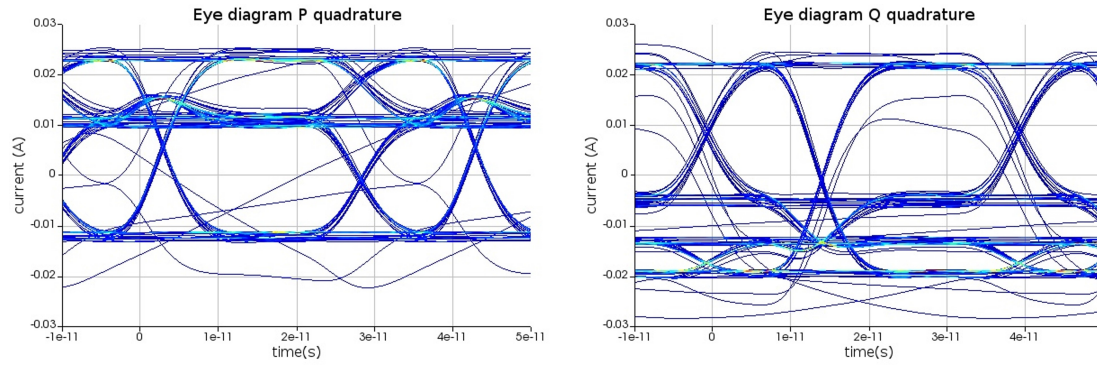


FIGURE 3.57: Eye diagram results for P and Q quadrature in QPSK.

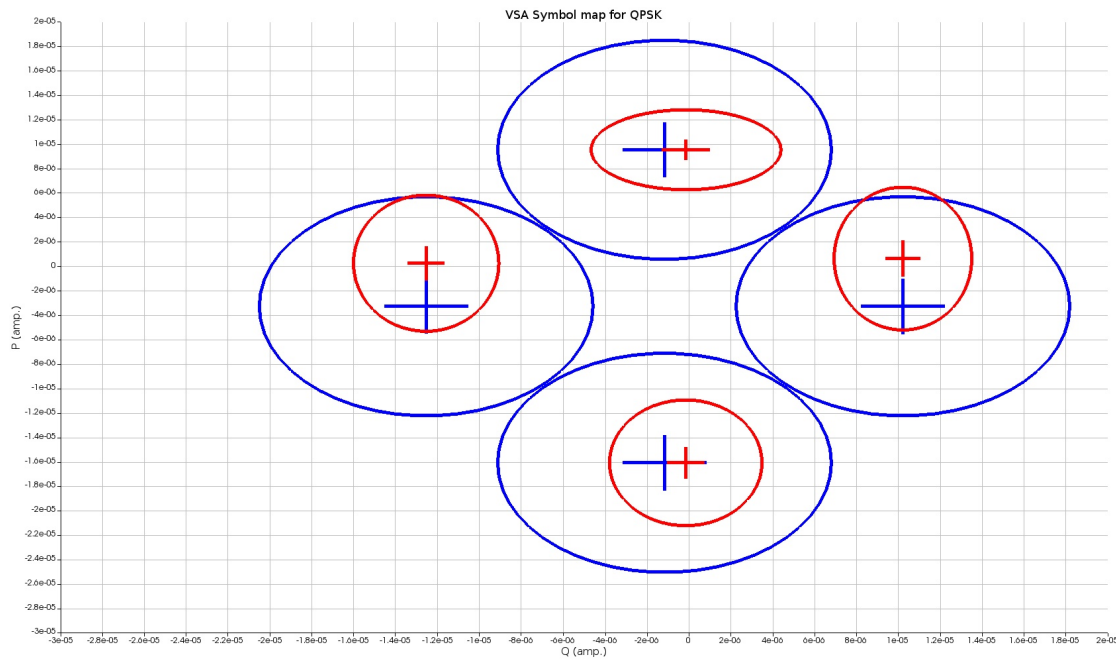


FIGURE 3.58: The signal statistic with respect to the decision boundaries in QPSK.

(3.60) because of not accounting for all the propagation losses incurred due to the waveguides. Hence, the later simulation is more close to the reality.

3.5.1 Power Budget Estimation

In Figure (3.59), on the transmitter side, the emitted power (double-sided) from the DFB laser is 11.86 dBm with TE polarization at 0.12 A bias current. The tunable beam splitter is biased at 2 A. At this operating point, the power splitting ratio between its cross output and bar output is 95:5. The bar output is used as the carrier signal while the cross output is used as the local oscillator. The carrier signal goes through four consecutive MZMs, all of which operate in dual-drive configuration. The first two MZMs are driven with non-return to zero pulses oscillating between $\frac{I_\pi}{2}$ and $\frac{3I_\pi}{2}$ current levels according to the sequences of 0's and 1's coming from a pseudo random number generator. Consequently, these two modulators realize BPSK by imparting either 0 or π phase shifts on

the optical carrier signal. On the other hand, the last two MZMs are driven with non-return to zero pulses oscillating between $\frac{I_\pi}{2}$ and I_π current levels according to the sequences of 0's and 1's coming from a pseudo random number generator. As a consequence, these two modulators realize BPSK by imparting either 0 or $\frac{\pi}{2}$ phase shifts on the optical carrier signal. The overall signal is a QPSK-modulated signal with phase distribution $\{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$. Higher average bias current gives rise to higher doping concentration, which in turn lead to higher absorption in a current injection phase modulator (CIPM). That is why, the insertion loss associated with the last two modulators is higher than the first two. The local oscillator signal is subjected to a polarization rotation, from TE to TM, with a polarization rotator. The modulated carrier signal and the LO signal are combined via a polarization beam splitter (PBS), which is biased at 0 A. As a result, the TE carrier and TM LO predominantly appear at its cross output port. The combined signal is coupled to the fiber from the chip via a spot size converter (SSC). At the point of leaving the transmitter chip, the power levels of the of the TM LO and the TE carrier are 7.43 dBm (5.54 mW) and -24.5 dBm (3.5 μ W), respectively, with extinction ratio of almost 30 dB.

While running our simulation, we did not find any good compact model for an optical fiber available from the circuit simulator. But in an actual experiment, our transmitter chip will be connected with the receiver chip via a standard single mode fiber. Therefore, to take into account the effect of the fiber, we need to consider additional propagation loss of 0.2 dB/km for both the TE and the TM modes. For example, for a 25 km long fiber, the LO and the carrier will arrive at the receiver with additional 5 dB transmission loss other than the coupling and insertion losses from the transmitter chip.

On the receiver side, the transmitted signal is coupled onto the receiver chip by another SSC. The combined signal is, first, halved in two with an MMI. For both the halves, the combined signal is split in its TE and TM components with polarization beam splitters. For one PBS, we want the bar output to be TM, because a polarization converter lies along this path. This PBS is biased at 8 mA. The other PBS is biased at 0 A so that the cross output is TM for the exact same reason. Now that the carrier and the LO are polarization matched, they are mixed in a 2×4 optical hybrid. A standalone CIPM, which is a part of the hybrid, is biased at 0.00211495 A to recover the quadrature phase of the optical carrier. In this regard, a half of either the LO or the carrier can be made to go through a $\pi/2$ phase shift. The four average optical power outputs from the 2×4 hybrid are -2.85 dBm (0.52 mW), -2.9 dBm (0.51 mW), -7.2 dBm (0.19 mW), -7.3 dBm (0.186 mW), respectively. The first two power levels are related to the p quadrature detection, while the last two are related to the q quadrature detection. The extra loss associated with one of the quadratures can be attributed to the presence of a phase modulator on its recovery path, which has an insertion loss of approximately 4 dB. Our PIN photodiodes have a characteristic responsivity of 0.72 A/W. Therefore, the electrical currents generated by the four photodiodes due to the impinging optical power are 374.165 μ A, 368.78 μ A, 137.72 μ A, 135.52 μ A, respectively. In a balanced, coherent, heterodyne detection scheme the first two diodes are connected together to provide the differential current of |5.385|

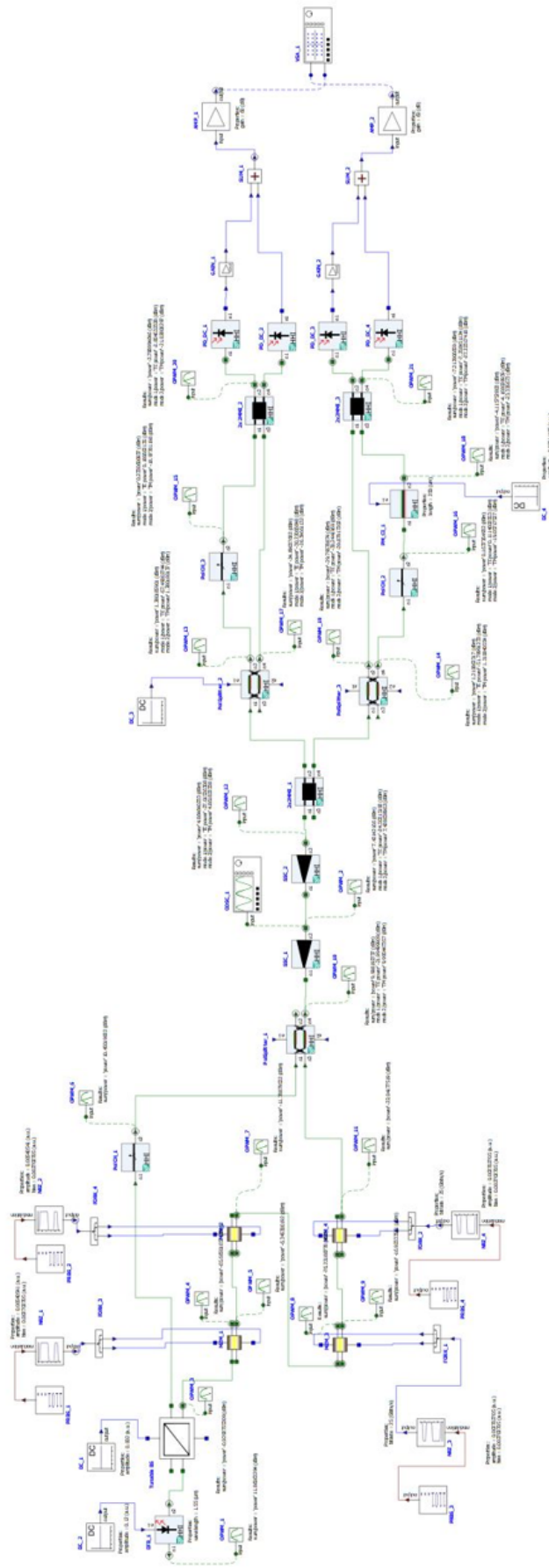


FIGURE 3.59: The circuit level simulation for our transceiver without the interconnects.

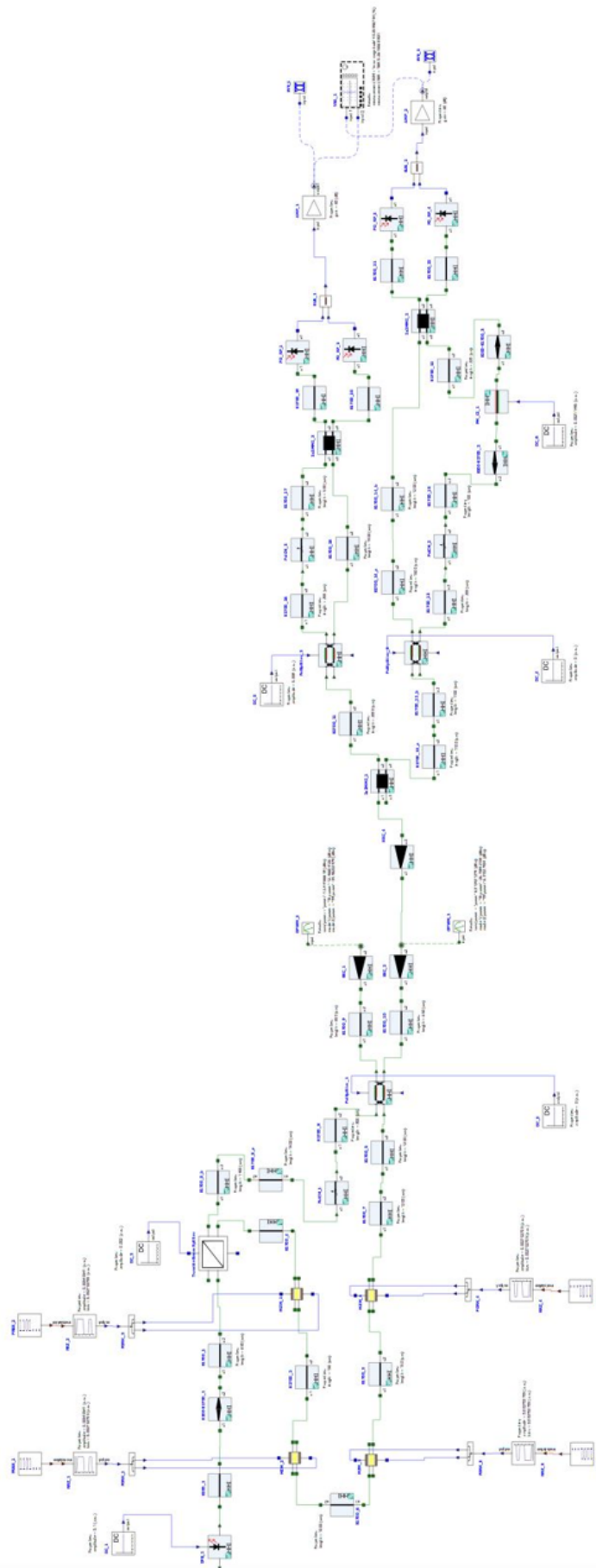


FIGURE 3.60: The circuit level simulation for our transceiver with all the connecting waveguides.

μA , which is proportional to the p quadrature. Similarly, the last two diodes are connected together to provide the differential current of $|2.2| \mu\text{A}$, which is proportional to the q quadrature. To boost the power levels of the differential currents, two transimpedance amplifiers with $Z_{\text{trans}} = 5k\Omega$ are also included in the off-chip electronic circuitry.

Figure (3.61) shows the signal statistic of the simulated system obtained from the vector signal analyzer (VSA). The symbol error rate (SER) is found to be 0.0542, from which one can calculate the signal-to-noise ratio (SNR) for the QPSK system according to the formula

$$P_s = 1 - [1 - Q(\sqrt{2\gamma_b})]^2, \quad (3.78)$$

where γ_b denotes the SNR associated with QPSK, P_s is the probability of symbol error, and $Q(x)$ is the complementary error function given by

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{x^2}{2}} dx. \quad (3.79)$$

Making use of equations (3.78) and (3.79), and looking up the $\text{erfc}(x)$ table, one calculates the SNR to be 1.2168. With this SNR, one can find that the mutual information of the quantum channel, as per equation (2.121), is 0.5742 bits.

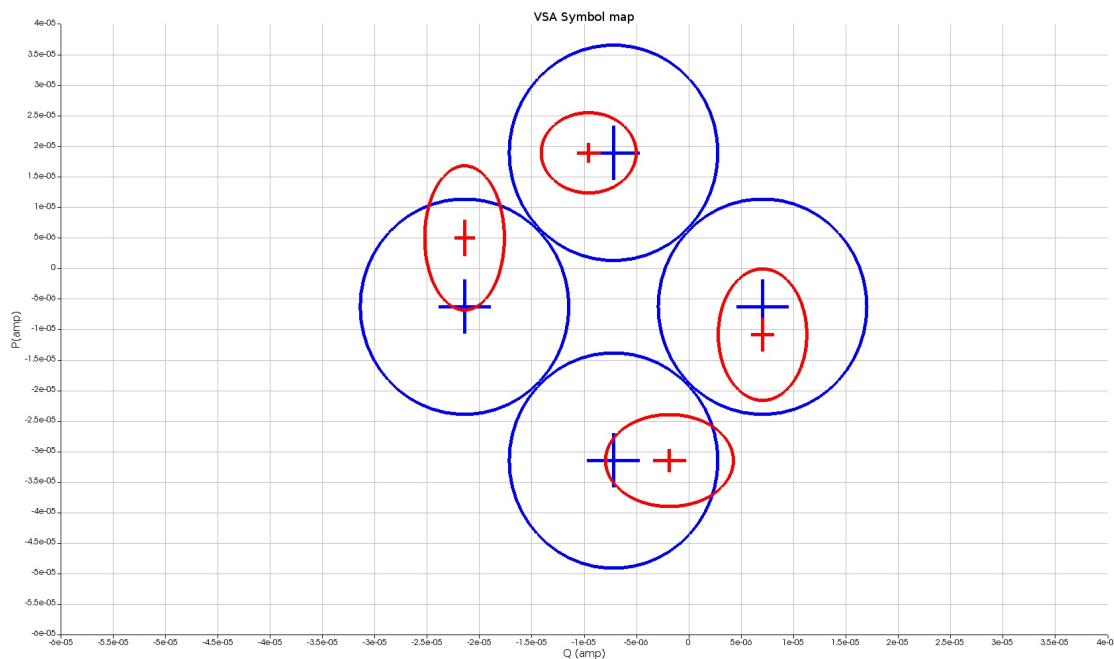


FIGURE 3.61: The signal statistics, against the backdrop of the decision boundaries, of our transceiver system aiming to implement QPSK.

3.6 Mask Layout Generation

We have used the *Nazca Design* framework to generate the GDSII file of our photonic integrated circuit design. The GDSII file contains lithographic information

for the foundry to make sense of the design. *Nazca* is an open source photonic IC design framework from Bright Integrated Photonics. One needs Python3 environment to execute *Nazca* scripts. The generated mask layouts can be viewed with any GDSII file reader such as *KLayout*. During the mask layout generation process I was mentored by Ronald Broeke from Bright Photonics BV.

Figure (3.62) shows the top cell of our photonic integrated transmitter. The sequence of components starts from the top right corner with the DFB laser. The DFB laser is made of ridge waveguides with etch depth $200\text{ }\mu\text{m}$ (e200). But the interconnection (blue lines) among different components is provided with e1700 waveguides. Therefore a waveguide transition element (e200-e1700) is required to bridge the two waveguides of two different etch depths. A tunable MZI splits the TE polarized laser output into LO and carrier. The carrier goes

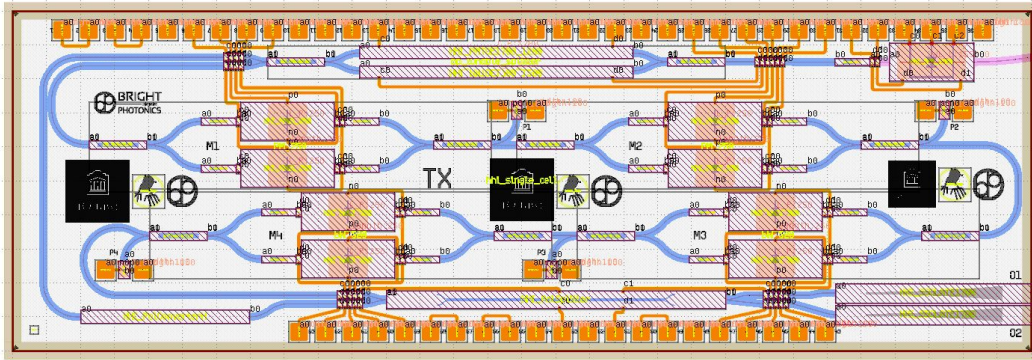


FIGURE 3.62: Mask layout of a photonic integrated transmitter for CV QKD.

through a sequence of four MZMs before mixing with the LO in a polarization beam splitter. One of the outputs from each MZM is guided to monitor photodiodes which helps to analyze the MZM output performance in real time. The LO is converted from TE to TM prior mixing. The output of the PBS is taken out of the chip to the optical fiber by means of a spot size converter (bottom right corner). The laser, the MZI, the MZMs, and the PBS are all active components. They require external electrical driver signals for their proper operation. That is why these active components have electrical ports in addition to optical ports. To interface electrical driver circuits with our photonic chip electrical bond pads are placed at the top and the bottom periphery. The electrical ports of the active components can be connected to these bond pads either directly or via waveguide metal transitions (tiny red rectangular boxes).

Figure (3.63) shows the top cell of our photonic integrated receiver. There is an MMI right next to an SSC (the last one from the top) which receives the transmitted light. The MMI splits both the LO and the carrier so that they can be polarization demultiplexed prior to balanced (homodyne or heterodyne) detection. The photo-currents produced by the photodetectors are routed out of the photonic chip to an electrical circuit board for electrical subtraction followed by amplification. In the receiver cell, in addition to the coherent receiver elements,

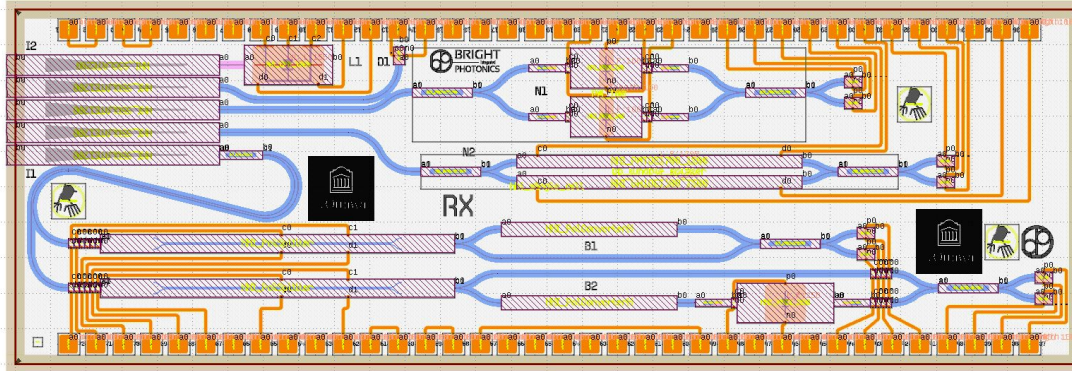


FIGURE 3.63: Mask layout of a photonic integrated receiver for CV QKD. The receiver chip contains coherent receiver components for balanced detection as well as some test structures for characterization.

there was enough room for us to place a laser, a tunable MZI, an MZM along with some photodiodes as test structures. We thought it would be a good idea to characterize these components before going on to perform real experiments with the chips.

3.7 Wafer Validation

The design was submitted on due time for the multi project wafer run HHI17. It took a little over six months to get the chips shipped to Ottawa upon fabrication. The chips were subjected to on-wafer testing by the foundry. Table (3.3) and Figure (3.64) shows measurement results done by the foundry on some of the building blocks.

Performance parameters						
Waveguide loss E1700	Butt-joint loss	PD responsivity	DFB wave-length deviation	DFB power @50 mA	DFB series resistance	SOA gain
2 dB/cm	2 dB	0.4 A/W	-5 nm	0 dBm	14 Ω	4 dB

TABLE 3.3: Overview of the validation results reported by the foundry.

The loss reported for ridge waveguide of etch depth of 1700 nm is the maximum loss over all polarizations and the C-band. The butt-joint loss refers to the active-passive transition loss over all polarizations and the C-band. To measure the responsivity of the photodetector, it was directly attached to a spot size converter.

Optical power is launched into the detector via a standard single mode fiber which is butt-coupled to the SSC. The reported responsivity applies to this set up for all polarizations and the C-band. The DFB wavelength deviation denotes

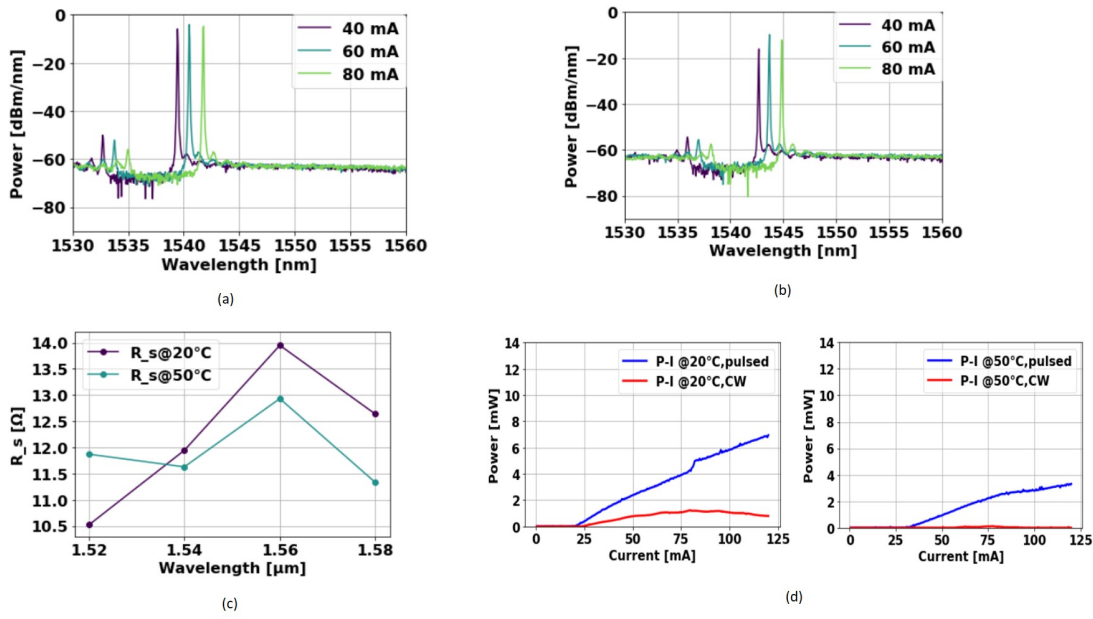


FIGURE 3.64: Spectrum of a DFB laser with a target wavelength of $1.54 \mu\text{m}$, measured at various currents and at (a) 20°C (b) 50°C . (c) DFB series resistance above threshold for various DFB wavelengths. DFB LIV curves obtained using a CW current (red) and a pulsed current (blue) for (d) 20°C (e) 50°C .

the difference between DFB wavelength design target and the peak wavelength measured at 50 mA (also Figure 3.64(a,b)). Positive values mean that the measured wavelength is red-shifted. The DFB power is measured using a CW current of 50 mA on bar level without proper heat sink. The given value is the minimum power across the C-band (also Figure 3.64(d,e)). The measurement is done using an on-chip photodiode with known responsivity.

The DFB series resistance is the differential resistance ($r_{\text{diff}} = \frac{dv}{di}$) of a DFB laser above threshold (also Figure 3.64(c)). The reported SOA gain is the gain of a $900 \mu\text{m}$ long SOA in TE. The given value is the minimum across the C-band and includes coupling loss of the active/passive transition but not any fiber coupling. The measurement is done with a CW current of 100 mA and without special heat management.

3.8 Conclusion

This chapter steps through the design analysis of an integrated photonic transmitter and a receiver capable of implementing continuous variable quantum key distribution. Needless to say, we had to paly around within the confines of technologies which were available from the foundry at the time of the targeted multi-project wafer run schedule. With progress in technology, there is room for improvement in the same design both performance wise and capability wise. For example, our system is constrained by how fast the phase modulator (for CIPM it is 350 MHz) is. Now HHI has introduced an electro-optic phase modulator (EOPM) with travelling wave electrode (TWE) which can operate in the

GHz region. Finally, the design is just a small part of a bigger project– proof-of-principle demonstration of on-chip CV QKD in InP. The photonic chips need to be properly packaged before characterizing it in the lab. The package would contain, among other things, electrical PCBs wire-bonded to the photonic integrated circuits for driving the active components. The experimentally obtained data from the photonic chips subject to classical post processing will eventually establish the key rate of the protocol. All these steps will be performed by the new group members who will continue the project after my graduation.

Chapter 4

Electromagnetic Analysis of a DFB Laser

4.1 Introduction

This chapter introduces a project, which grew out of a collaboration with one of our industrial partners, Integent. The goal of the project is to investigate how the physical parameters of a distributed feedback (DFB) semiconductor laser (emitting at 1310 nm) with a specific design can be optimized for obtaining enhanced output performance. In that regard, my contribution has been the theoretical derivations documented in section (4.4) of this chapter, and the numerical results presented in Chapter 5. This chapter endeavors to lay the theoretical groundwork to better appreciate these works. A DFB laser is conspicuous by its narrowband emission (Figure (4.1b)). While in a regular Fabry-Perot laser resonator discrete feedback is provided at the two ends of the optical gain medium (Figure (4.1c)), in a DFB laser the feedback is distributed over the length of the whole gain region. One can argue that in the case of a DFB laser the phase matching condition becomes more rigorous (Figure (4.1e)) resulting in a more selective spectral condition. In a semiconductor DFB laser, the feedback is fabricated in the form of surface corrugation (Figure (4.1d)), which works as a grating. The electromagnetic analysis of such a laser combines the effects of a slab waveguide (necessary for optical confinement), and a grating. Our upcoming discussion on the transverse mode analysis and the longitudinal mode analysis of a generic DFB laser will reflect on this fact. The discussion synthesizes materials mostly from the following two sources [23], [93]. The important equations we make use of in the next chapter are boxed around to distinguish them from the rest.

4.2 Distributed Feedback Laser

A Fabry-Perot optical cavity formed by two mirrors positioned at $z = 0$, and $z = L$ will support electromagnetic modes such that

$$k_q d = q\pi \Rightarrow \nu_q = q \frac{c}{2L} \quad q \in \mathbb{N}. \quad (4.1)$$

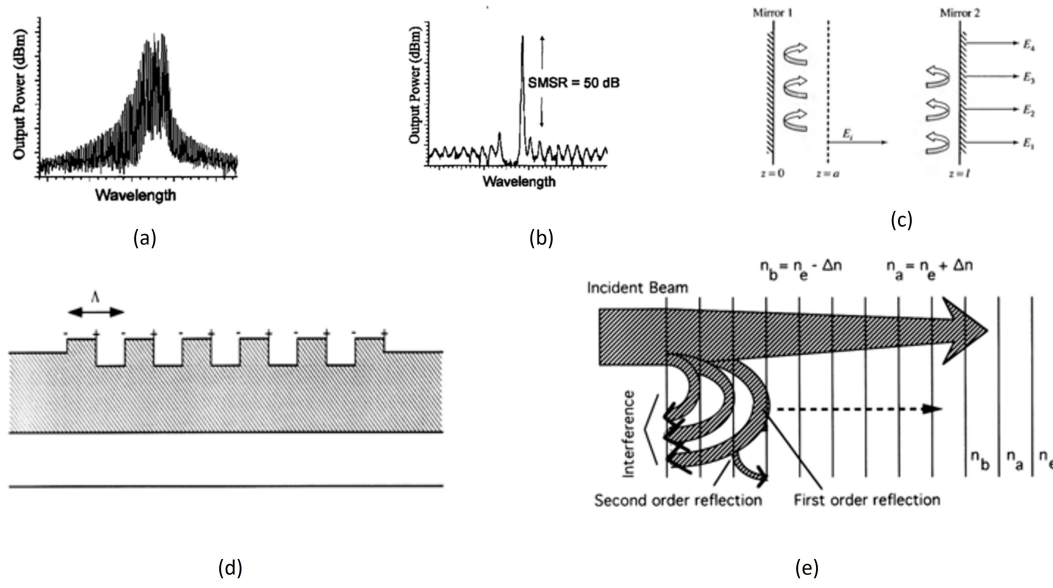


FIGURE 4.1: (a),(b) Transmission spectrum of a Fabry-Perot laser and a DFB laser [24]. (c) Phase matching condition illustrated in a Fabry Perot cavity [94]. (d) Surface corrugation to define periodic grating in a DFB laser [95]. (e) Phase matching condition illustrated in a DFB cavity [95].

In words, the round trip phase of a travelling wave should be an integer multiple of 2π within the cavity. This is known as the phase matching condition. Figure (4.1c) illustrates the idea. ν_q denotes a particular frequency supported by the structure. Figure (4.1a) gives an idea of the associated spectrum. The frequency spacing between two adjacent supported modes is given by

$$\nu_{q+1} - \nu_q = \frac{c}{2L}. \quad (4.2)$$

Therefore, by letting L be small, one can enlarge the frequency spacing, and hence can make the cavity more spectrally selective. A DFB cavity can be visualized as a series of planes (reflectors) with periodicity Λ . For a wave to be supported by the DFB structure, the phase matching condition needs to be satisfied at every interface, i.e.

$$\lambda_g = \frac{2\Lambda}{N}, \quad N = 1, 2, 3, \dots \quad (4.3)$$

which is known as the Bragg condition. Here λ_g is the optical wavelength within the grating region, and N is the grating order which is an integer. Since $\Lambda \ll L$, from equation (4.2), one can infer that the DFB cavity spectrum will be more spaced out. Now if the optical cavity is filled with a gain medium, then only those frequencies within the gain spectrum of the medium will experience amplification. With the strategy discussed above, in theory, one can build an optical resonator with a gain medium and distributed feedback, which will selectively boost a certain mode with a desired frequency.

The above discussion is just a hand-wavy explanation of the working principle of a distributed feedback laser. The theoretical analysis of a first-order, $N = 1$

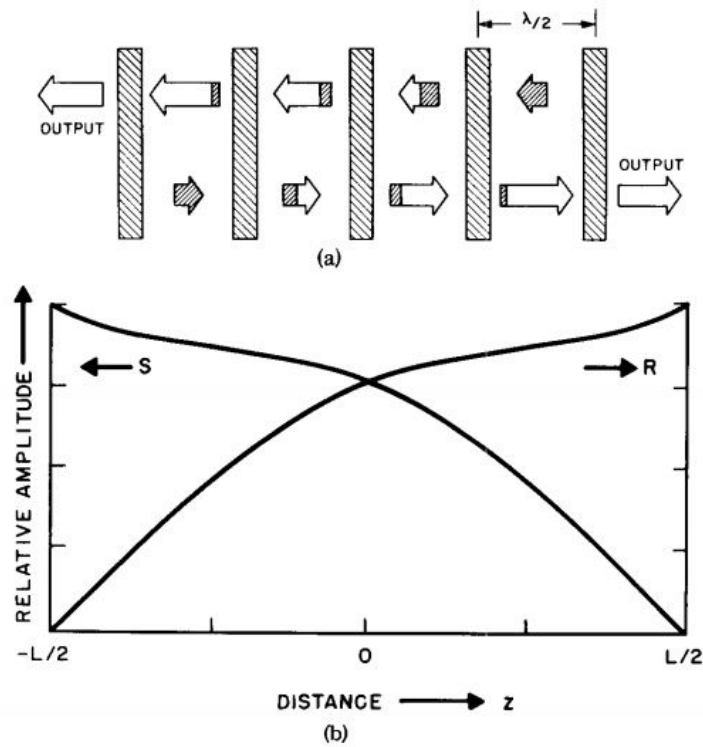


FIGURE 4.2: (a) The contra-propagating Bragg modes in a 1st order DFB laser feed off each other. The grating induces the coupling. (b) Plot of the relative amplitudes of left travelling wave S and right travelling wave R vs distance.

in equation (4.3), DFB laser was first notably performed by Kogelnik and Shank [96]. They applied the coupled mode formalism in formulating the electromagnetic behavior of a DFB laser (Figure 4.2). The essence of the theory is the following: For the first order DFB, the field inside the cavity may be decomposed into a forward propagating wave $R(z)$ and a backward propagating wave $S(z)$ both satisfying equation (4.3). $R(z)$ starts at $z = -\frac{L}{2}$, and grows as it traverses along $+z$. $S(z)$ starts at $z = +\frac{L}{2}$, and grows as it traverses along $-z$. Figure (4.2 b) illustrates this idea. The grating couples the two modes as they propagate. At each grating interface, a fraction of $R(z)$ is coupled with $S(z)$, and the rest is transmitted. At $z \rightarrow \frac{L}{2}$ i.e. at the output end, the amplified $R(z)$ somewhat flattens. The same applies to $S(z)$. The energy exchange with Bragg modes and any other diffraction mode is negligible.

Another approach based on Floquet-Bloch theorem has also been proposed [97]. In this approach, the grating is represented by a periodic modulation of the electrical permittivity (refractive index). The field inside the structure is expanded in an infinite series of partial waves (spatial harmonics). Partial waves with matching phase exchange energy. In DFB lasers these are the forward and the backward propagating harmonics. It has been demonstrated that the truncated form of Floquet-Bloch expansion is equivalent to the coupled wave theory [98].

William Streifer coalesced both the approaches to develop the theory for higher order ($N > 1$) DFB lasers in a consistent way [93]. In a higher order DFB laser, it is necessary to consider energy exchange between Bragg modes and the other

diffraction order modes. Streifer modified the coupled wave equations with some correction terms which are known as the Streifer coefficients. It is his approach that we discuss in great detail below.

4.3 Transverse Mode Analysis

Let us start with the Helmholtz wave equation, which must be satisfied by the TE polarized wave $E_y(x, z)$ in the structure shown in Figure (4.3)

$$\nabla^2 E_y(x, z) + k_0^2 n^2(x, z) E_y(x, z) = 0. \quad (4.4)$$

Since the structure is uniform in the y -direction, we can write

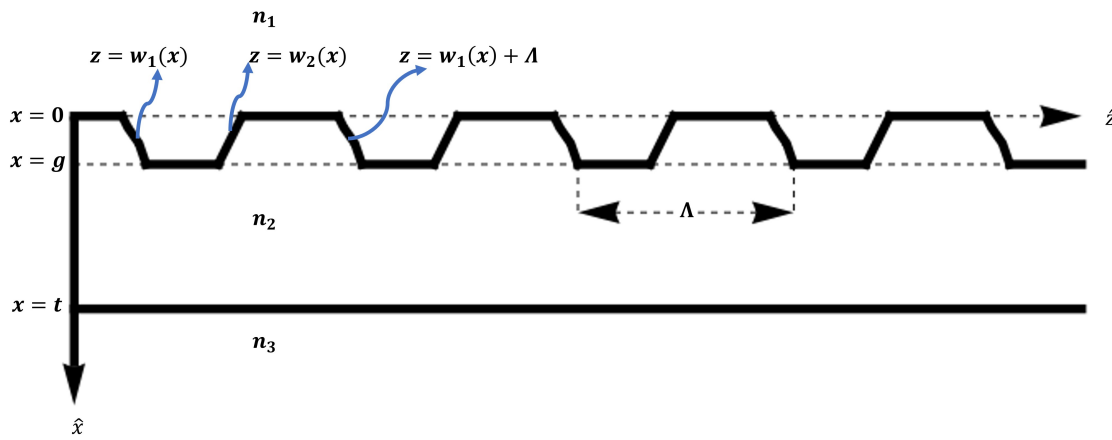


FIGURE 4.3: Side view of a four layer DFB laser structure. The grating is described by its period Λ , and the tooth shape functions $w_1(x)$, and $w_2(x)$

$$\frac{\partial^2 E_y}{\partial z^2} + \frac{\partial^2 E_y}{\partial x^2} + k_0^2 n^2(x, z) E_y = 0, \quad (4.5)$$

where a harmonic time dependence of $e^{j\omega t}$ is assumed; the free space wave number is given by $k_0 = \frac{2\pi}{\lambda_0}$; $n^2(x, z)$ is the refractive index profile of the structure. For our structure of interest, the squared refractive index $n^2(x, z)$ is constant for $x < 0$ and $x > g$. In the grating region $0 < x < g$, it can be written in terms of unit step functions as [23]

$$n^2(x, z) = \sum_{p=-\infty}^{\infty} \left[n_1^2 \{ u[z - w_1(x) + p\Lambda] - u[z - w_2(x) + p\Lambda] \} + n_2^2 \{ u[z - w_2(x) + p\Lambda] - u[z - w_1(x) - \Lambda + p\Lambda] \} \right], \quad (4.6)$$

where

$$u(z) = \begin{cases} 0 & z < 0 \\ 1 & z > 0. \end{cases} \quad (4.7)$$

$w_1(x)$ and $w_2(x)$ describe the tooth shape, and Λ is the grating period. The summation over p mathematically represents the periodicity of the grating structure. Next we expand $n^2(x, z)$ for $0 < x < g$ in a Fourier series

$$n^2(x, z) = \sum_{q=-\infty}^{\infty} A_q(x) e^{j \frac{2\pi q z}{\Lambda}}, \quad (4.8)$$

with

$$A_q(x) = \frac{1}{\Lambda} \int_{z_0(x)}^{z_0(x)+\Lambda} n^2(x, z) \exp\left(-j \frac{2\pi q z}{\Lambda}\right) dz. \quad (4.9)$$

$z_0(x)$ can be chosen arbitrarily because of the periodicity. From equation (4.9), for $q = 0$,

$$\begin{aligned} A_0(x) &= \frac{n_1^2}{\Lambda} \int_{w_1(x)}^{w_2(x)} dz + \frac{n_2^2}{\Lambda} \int_{w_2(x)}^{w_1(x)+\Lambda} dz \\ \Rightarrow \boxed{A_0(x) &= n_2^2 - (n_2^2 - n_1^2)[w_2(x) - w_1(x)]/\Lambda}. \end{aligned} \quad (4.10)$$

The boxed equation represents an 'averaged-over' refractive index with which the grating region can be approximated as a homogeneous region. Similarly, $A_q(x)$ for $q \neq 0$ is

$$\begin{aligned} A_q(x) &= \frac{n_1^2}{\Lambda} \int_{w_1(x)}^{w_2(x)} \exp\{(-2\pi j q z / \Lambda)\} dz + \frac{n_2^2}{\Lambda} \int_{w_2(x)}^{w_1(x)+\Lambda} \exp\{(-2\pi j q z / \Lambda)\} dz \\ \Rightarrow \boxed{A_q(x) &= \frac{(n_2^2 - n_1^2)}{j 2\pi q} \left[\exp\left\{\frac{-2\pi j q w_2(x)}{\Lambda}\right\} - \exp\left\{\frac{-2\pi j q w_1(x)}{\Lambda}\right\} \right]}. \end{aligned} \quad (4.11)$$

From the above equation, it is obvious that $A_{-q}(x) = A_q^*(x)$. Now, the refractive index of the whole structure can be written as

$$n^2(x, z) = \tilde{n}_0^2(x) + \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} A_q(x) \exp\left(j \frac{2\pi q z}{\Lambda}\right). \quad (4.12)$$

The $q = 0$ term of the Fourier Series which does not depend upon z has been subsumed in $\tilde{n}_0(x)$. With the above expression equation (4.5) can be rewritten in the following way

$$\frac{\partial^2 E_y}{\partial z^2} + \frac{\partial^2 E_y}{\partial x^2} + k_0^2 \tilde{n}_0^2(x) E_y + k_0^2 E_y \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} A_q(x) \exp\left(j \frac{2\pi q z}{\Lambda}\right) = 0. \quad (4.13)$$

$\tilde{n}_0(x)$ is a complex quantity, the imaginary part of which accounts for the loss, or the gain of the structure

$$\tilde{n}_0(x) = n_0(x) - j\kappa(x). \quad (4.14)$$

Therefore, the $k_0^2 \tilde{n}_0^2(x)$ term in the equation (4.13) evaluates to

$$k_0^2 \tilde{n}_0^2(x) = k_0^2 \left[n_0(x) - j \frac{\tilde{\alpha}(x)}{k_0} \right]^2 = k_0^2 n_0^2(x) - 2jk_0 n_0(x) \tilde{\alpha}(x) - \tilde{\alpha}^2(x), \quad (4.15)$$

where $\tilde{\alpha}(x)$ represents gain if $\tilde{\alpha} > 0$, or loss if $\tilde{\alpha} < 0$ in each region. Since $|\tilde{\alpha}| \ll k_0 n_0$, one can drop $\tilde{\alpha}^2$, and write

$$k_0^2 \tilde{n}_0^2(x) \approx k_0^2 n_0^2(x) - 2jk_0 n_0(x) \tilde{\alpha}(x). \quad (4.16)$$

The real part of the refractive index is distributed in each region according to the following way

$$n_0^2(x) = \begin{cases} n_1^2 & x \leq 0 \\ A_0(x) & 0 \leq x \leq g \\ n_2^2 & g \leq x \leq t \\ n_3^2 & t \leq x. \end{cases} \quad (4.17)$$

If, for the moment, we ignore the z variation of the refractive index by disregarding all the terms with $q \neq 0$, and consider $\text{Im}(\tilde{n}_0) \approx 0$, equation (4.13) becomes

$$\frac{\partial^2 E_y}{\partial z^2} + \frac{\partial^2 E_y}{\partial x^2} + k_0^2 n_0^2(x) E_y = 0. \quad (4.18)$$

Equation (4.18) is the governing equation for a multi-stack slab waveguide which admits solution of the form $E_y(x, z) = \varepsilon_0(x) \exp(j\beta z)$ because of the assumed uniformity along the propagation direction. In particular, the transverse profile $\varepsilon_0(x)$ satisfies

$$\boxed{\frac{d^2 \varepsilon_0}{dx^2} + [k_0^2 n_0^2(x) - \beta^2] \varepsilon_0(x) = 0}. \quad (4.19)$$

This is an eigenvalue problem which yields a set of allowable values for β , each of which is the propagation constant for a different guided TE mode

$$\beta = \frac{2\pi n_{\text{eff}}}{\lambda_0}. \quad (4.20)$$

Therefore, corresponding to a set of values for β , there is a set of values for n_{eff} .¹ Here n_{eff} represents the refractive index the guided mode effectively sees while propagating through the structure. In order for a mode to be guided, n_{eff} must satisfy $n_2 > n_{\text{eff}} > \max(n_1, n_3)$, where n_2 is the core index where the field is mostly confined, (n_1, n_3) are the cladding indices surrounding the core.

Now with a grating included in the structure, we assume an electric field solution of the following form

$$E_y(x, z) = \sum_{m=-\infty}^{\infty} E_y^{(m)}(x, z) e^{j\beta_m z}, \quad (4.21)$$

¹The mode with the highest n_{eff} , is called the fundamental mode, and denoted as $\beta_0 = \frac{2\pi n_{\text{eff, max}}}{\lambda_0}$

since according to the Floquet-Bloch theorem the electric field in a periodic medium can be expressed as an infinite summation of partial waves. Here $E_y^{(m)}$ is the m th partial field wave (diffracted mode) component; (m denotes the field order), and it slowly varies with z . β_m is the propagation constant for an m th order partial wave with respect to β_0

$$\beta_m = \beta_0 + \frac{2\pi m}{\Lambda}. \quad (4.22)$$

The mode field is assumed to be consisting of a series of standing waves in the z direction with the lowest order mode $m = 0$ denoting the mode that would have existed in the waveguide in the absence of a grating. Substituting equation (4.21) into equation (4.13), results in the following expression upon algebraic manipulation (Appendix E)

$$\begin{aligned} \frac{\partial^2 E_y^{(m)}}{\partial x^2} + \frac{\partial^2 E_y^{(m)}}{\partial z^2} + 2j\beta_m \frac{\partial E_y^{(m)}}{\partial z} - \beta_m^2 E_y^{(m)} + k_0^2 \tilde{n}_0^2(x) E_y^{(m)} \\ = -k_0^2 \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} A_q(x) E_y^{(m-q)} \end{aligned} \quad (4.23)$$

Resonant case: For a DFB laser with an N^{th} order grating with a period of Λ , the propagation constant of the TE mode should satisfy the Bragg condition

$$\beta_0 = \frac{N\pi}{\Lambda} \quad N = 1, 2, 3, \dots \quad (4.24)$$

One of the partial waves generated by the grating of the order $p = -N$ has propagation constant

$$\beta_p = \beta_0 + \frac{2\pi p}{\Lambda} = \beta_0 - \frac{2\pi N}{\Lambda} = \beta_0 - 2\beta_0 = -\beta_0, \quad (4.25)$$

which couples resonantly with the forward propagating wave making $E_y^{(0)}$ and $E_y^{(p)}$ the dominant modes. The Fourier coefficient $A_p(x)$ corresponding to this particular p serves to couple $E_y^{(0)}(x, z)$ to its contradirectional part $E_y^{(p)}(x, z)$. While $A_{-p}(x) = A_p^*(x)$ provides the reverse coupling. In this resonant case, let us assume

$$E_y^{(0)}(x, z) = R(z)\varepsilon_0(x) \quad (4.26a)$$

$$E_y^{(p)}(x, z) = S(z)\varepsilon_0(x). \quad (4.26b)$$

Here $\varepsilon_0(x)$ is the fundamental mode profile, and $R(z)$ and $S(z)$ are the axial dependence of the waves propagating in the $+z$ and $-z$ direction respectively. Substituting (4.26a) into (4.23) for $m = 0$, gives

$$\begin{aligned} R(z) \frac{d^2 \varepsilon_0(x)}{dx^2} + \varepsilon_0(x) \frac{d^2 R(z)}{dz^2} + 2j\beta_0 \varepsilon_0(x) \frac{dR(z)}{dz} + k_0^2 \tilde{n}_0^2(x) R(z) \varepsilon_0(x) - \beta_0^2 \varepsilon_0 R(z) \\ = -k_0^2 \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} A_q(x) E_y^{(-q)}(x, z). \end{aligned} \quad (4.27)$$

Under the slowly varying wave approximation, we can neglect $\frac{d^2 R(z)}{dz^2}$, obtaining

$$2j\beta_0 R' \varepsilon_0 + R \left\{ \frac{d^2 \varepsilon_0(x)}{dx^2} + [k_0^2 \tilde{n}_0^2(x) - \beta_0^2] \varepsilon_0(x) \right\} = -k_0^2 \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} A_q(x) E_y^{(-q)}(x, z). \quad (4.28)$$

Similarly, for $m = p$, with $\beta_p^2 = (-\beta_0)^2 = \beta_0^2$, one gets

$$-2j\beta_0 S' \varepsilon_0 + S \left\{ \frac{d^2 \varepsilon_0(x)}{dx^2} + [k_0^2 \tilde{n}_0^2(x) - \beta_0^2] \varepsilon_0(x) \right\} = -k_0^2 \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} A_q(x) E_y^{(p-q)}(x, z). \quad (4.29)$$

Coupled terms have been placed on the right hand side of both equations (4.28) and (4.29). The expression inside the curly brackets on the left hand side of both the equations can be simplified to

$$-2jk_0 n_0(x) \tilde{\alpha}(x) + (\beta^2 - \beta_0^2) \varepsilon_0(x), \quad (4.30)$$

where we have used equations (4.19) and (4.16). Substituting (4.30) in (4.28) and (4.29), we have, respectively

$$\begin{aligned} & 2j\beta_0 R' \varepsilon_0(x) + R \left[-2j\tilde{\alpha}(x) k_0 n_0(x) + (\beta^2 - \beta_0^2) \right] \varepsilon_0(x) \\ &= \underbrace{-k_0^2 A_{-p}(x) S \varepsilon_0(x)}_{\substack{q=-\infty \\ q \neq 0, p}} - k_0^2 \sum_{\substack{q=-\infty \\ q \neq 0, p}}^{\infty} A_q(x) E_y^{(-q)}(x, z), \end{aligned} \quad (4.31)$$

$$\begin{aligned} \text{and } & -2j\beta_0 S' \varepsilon_0(x) + S \left[-2j\tilde{\alpha}(x) k_0 n_0(x) + (\beta^2 - \beta_0^2) \right] \varepsilon_0(x) \\ &= \underbrace{-k_0^2 A_p(x) R \varepsilon_0(x)}_{\substack{q=-\infty \\ q \neq 0, p}} - k_0^2 \sum_{\substack{q=-\infty \\ q \neq 0, p}}^{\infty} A_q(x) E_y^{(p-q)}(x, z). \end{aligned} \quad (4.32)$$

The curly braces indicate that the $-p$ th and the p th term have been taken out of the summation. We multiply equation (4.31) with $\varepsilon_0(x)$ and integrate over $(-\infty, \infty)$

$$\begin{aligned} & 2j\beta_0 R' \int_{-\infty}^{\infty} \varepsilon_0^2(x) dx + R(-2jk_0) \int_{-\infty}^{\infty} \tilde{\alpha}(x) n_0(x) \varepsilon_0^2(x) dx + R(\beta^2 - \beta_0^2) \int_{-\infty}^{\infty} \varepsilon_0^2(x) \\ &= -k_0^2 S \int_{-\infty}^{\infty} A_{-p}(x) \varepsilon_0^2(x) dx - k_0^2 \sum_{\substack{q=-\infty \\ q \neq 0, p}}^{\infty} \int_0^g A_q(x) \varepsilon_0(x) E_y^{(-q)} dx. \end{aligned} \quad (4.33)$$

Let us define the following quantities

$$\text{Modal power} \quad \boxed{P = \int_{-\infty}^{\infty} \varepsilon_0^2(x) dx}, \quad (4.34)$$

$$\text{Modal gain/loss} \quad \boxed{\alpha = \frac{k_0}{\beta_0 P} \int_{-\infty}^{\infty} n_0(x) \tilde{\alpha}(x) \varepsilon_0^2(x) dx}, \quad (4.35)$$

Coupling Coefficient	$\kappa_{-p} = \frac{k_0^2}{\beta_0 P} \int_0^g A_{-p}(x) \varepsilon_0^2(x) dx,$	(4.36)
----------------------	---	--------

Bragg detuning	$\delta = \frac{\beta^2 - \beta_0^2}{2\beta_0} \approx \beta - \beta_0.$	(4.37)
----------------	--	--------

It is to be noted that the integral for κ_{-p} extends over the region $(0, g)$, since $A_q(x) = 0$ outside that region. Dividing equation (4.33) by $2j\beta_0 P$ and making use of the definitions from (4.34) to (4.37) yields

$$R' + (-\alpha - j\delta)R = j\kappa_{-p}S + \frac{jk_0^2}{2\beta_0 P} \sum_{\substack{q=-\infty \\ q \neq 0, p}}^{\infty} \int A_q(x) \varepsilon_0 E_y^{(-q)} dx. \quad (4.38)$$

Going along the same line from equation (4.32), one would find

$$-S' + (-\alpha - j\delta)S = j\kappa_p R + \frac{jk_0^2}{2\beta_0 P} \sum_{\substack{q=-\infty \\ q \neq 0, p}}^{\infty} \int A_q(x) \varepsilon_0 E_y^{(p-q)} dx. \quad (4.39)$$

To evaluate the right hand side of (4.38) and (4.39), equation (4.23) must be solved for $E_y^{(m)}$ with $m \neq 0, p$. One can go forward with the following assumptions [23]: (i) Only $E_y^{(0)}$ and $E_y^{(p)}$ are important in generating $E_y^{(m)}$, $m \neq 0, p$ (ii) The z derivative terms in (4.23) are negligible compared to others, (iii) $\tilde{\alpha}$ may be neglected. With all these assumptions, we can write from (4.23)

$$\frac{\partial^2 E_y^{(m)}}{\partial x^2} + [k_0^2 n_0^2(x) - \beta_m^2] E_y^{(m)} = -k_0^2 [A_m(x) R(z) + A_{m-p}(x) S(z)], \quad (4.40)$$

which admits a solution of the form

$$E_y^{(m)}(x, z) = R(z) \varepsilon_m^{(0)}(x) + S(z) \varepsilon_m^{(p)}(x), \quad \forall m, m \neq 0, p \quad (4.41)$$

where $\varepsilon_m^{(0)}$ and $\varepsilon_m^{(p)}$ both satisfy

$\frac{d^2 \varepsilon_m^{(i)}(x)}{dx^2} + [k_0^2 n_0^2(x) - \beta_m^2] \varepsilon_m^{(i)}(x) = -k_0^2 A_{m-i}(x) \varepsilon_0(x)$	$\forall m, \quad m \neq i, \quad i=0, p \quad (4.42)$
--	--

Here, $\varepsilon_m^{(0)}$, and $\varepsilon_m^{(p)}$ are generated by the $+z$ wave $E_y^{(0)}$ and $-z$ wave $E_y^{(p)}$ respectively. The next task is to substitute equation (4.41) into the right hand sides of (4.38) and (4.39). From (4.38), it results

$$\begin{aligned} R' + (-\alpha - j\delta)R &= j\kappa_{-p}S + \frac{jk_0^2}{2\beta_0 P} \sum_{\substack{q=-\infty \\ q \neq 0, p}}^{\infty} \int_0^g A_q(x) \varepsilon_0 \{R \varepsilon_{-q}^{(0)}(x) + S \varepsilon_{-q}^{(p)}\} dx, \end{aligned} \quad (4.43)$$

the last term of which can be consolidated in the following way

$$R \sum_{\substack{q=-\infty \\ q \neq 0, p}}^{\infty} \eta_{q,-q}^{(0)} + S \sum_{\substack{q=-\infty \\ q \neq 0, p}}^{\infty} \eta_{q,-q}^{(p)} = R(z)\zeta_1 + S(z)\zeta_2. \quad (4.44)$$

Similarly from (4.39), results

$$\begin{aligned} & -S' + (-\alpha - j\delta)S \\ & = j\kappa_p R + \frac{jk_0^2}{2\beta_0 P} \sum_{\substack{q=-\infty \\ q \neq 0, p}}^{\infty} \int_0^g A_q(x) \varepsilon_0 \{R\varepsilon_{-q}^{(0)}(x) + S\varepsilon_{-q}^{(p)}\} dx, \end{aligned} \quad (4.45)$$

the last term can be consolidated also in this case

$$R \sum_{\substack{q=-\infty \\ q \neq 0, p}}^{\infty} \eta_{q,p-q}^{(0)} + S \sum_{\substack{q=-\infty \\ q \neq 0, p}}^{\infty} \eta_{q,p-q}^{(p)} = R(z)\zeta_4 + S(z)\zeta_3. \quad (4.46)$$

Putting things together, we can write the following pair of equations

$$\boxed{R' + (-\alpha - j\delta - j\zeta_1)R = j(\kappa_{-p} + \zeta_2)S} \quad (4.47)$$

$$\boxed{S' + (-\alpha - j\delta - j\zeta_3)S = j(\kappa_p + \zeta_4)R} \quad (4.48)$$

where $\zeta_1 \dots \zeta_4$ are the Streifer coefficients

$$\zeta_1 = \sum_{\substack{q=-\infty \\ q \neq 0, p}}^{\infty} \eta_{q,-q}^{(0)}, \quad (4.49a)$$

$$\zeta_2 = \sum_{\substack{q=-\infty \\ q \neq 0, p}}^{\infty} \eta_{q,-q}^{(p)}, \quad (4.49b)$$

$$\zeta_3 = \sum_{\substack{q=-\infty \\ q \neq 0, p}}^{\infty} \eta_{q,p-q}^{(p)}, \quad (4.49c)$$

$$\zeta_4 = \sum_{\substack{q=-\infty \\ q \neq 0, p}}^{\infty} \eta_{q,p-q}^{(0)}, \quad (4.49d)$$

$$\text{with } \boxed{\eta_{r,s}^{(i)} = \frac{k_0^2}{2\beta_0 P} \int_G A_r(x) \varepsilon_0(x) \varepsilon_s^{(i)}(x) dx,} \quad (4.49e)$$

with precise physical significance as described below

- ζ_1 represents the reaction of partial waves, be it radiating ($p < -q < 0$), or decaying ($-q < p, 0 < -q$), generated by the forward propagating wave on the forward propagating wave itself.

- ζ_3 represents the reaction of all partial waves generated by the backward waves on the backward propagating wave itself.
- For all gratings, $\zeta_1 = \zeta_3$.
- $\text{Re}(\zeta_1, \zeta_3)$ could be positive or negative, and will affect the deviation from the Bragg frequency.
- $\text{Im}(\zeta_1, \zeta_3)$ represents the additional radiation loss in the transverse direction due to the partial waves, and hence are always positive.
- ζ_2 term accounts for reaction of all the partial waves generated by the backward propagating mode on the forward propagating mode.
- ζ_4 term accounts for reaction of all the partial waves generated by the forward propagating mode on the backward propagating mode.
- For symmetric gratings, $\zeta_2 = \zeta_4$
- The ζ_2, ζ_4 terms act as correction terms to the coupling coefficient as they represent indirect coupling between the forward and the backward propagating mode through the partial waves.
- We can define a new term called effective coupling coefficient

$$\kappa_{\text{eff}} = \sqrt{(\kappa_p^* + \zeta_2)(\kappa_p + \zeta_4)} = |\kappa_{\text{eff}}| e^{j\phi(\kappa_{\text{eff}})}. \quad (4.50)$$

- κ_{eff} is complex with $\text{Re}(\kappa_{\text{eff}})$ denoting index coupling, and $\text{Im}(\kappa_{\text{eff}})$ denoting gain coupling.

4.4 Longitudinal Mode Analysis

As mentioned in section (4.3), at the resonant condition inside the DFB cavity, there are two dominant modes satisfying the Bragg condition, namely equations (4.26a) and (4.26b)

$$\begin{aligned} \varepsilon_0(x)R(z) \exp(j\beta_0 z), \\ \varepsilon_0(x)S(z) \exp(-j\beta_0 z). \end{aligned}$$

The transverse mode profile $\varepsilon_0(x)$ has been assumed to be identical for both the forward, and the backward propagating modes. The longitudinal modes are governed by equations (4.47) and (4.48)

$$\begin{aligned} R' + (-\alpha - j\delta - j\zeta_1)R &= j(\kappa_p^* + \zeta_2)S, \\ -S' + (-\alpha - j\delta - j\zeta_3)S &= j(\kappa_p + \zeta_4)R. \end{aligned}$$

The above two equations always admit solutions of the form

$$R(z) = r_1 \exp(\gamma z) + r_2 \exp(-\gamma z), \quad (4.51a)$$

$$S(z) = s_1 \exp(\gamma z) + s_2 \exp(-\gamma z). \quad (4.51b)$$

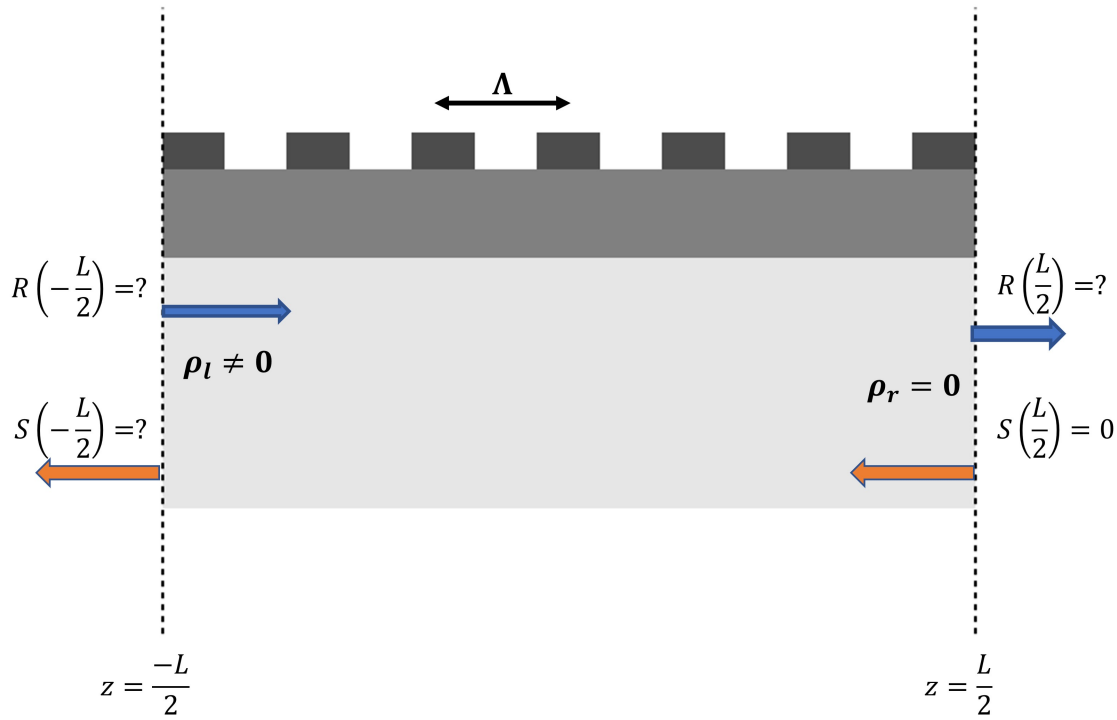


FIGURE 4.4: A generic DFB longitudinal section with additional reflector on one end.

Substituting (4.51a) into (4.47), and (4.51b) into (4.48), and subsequently collecting coefficients for the same exponentials, one arrives at (Appendix E for reference)

$$\hat{\Gamma} r_1 = j(\kappa_p^* + \zeta_2) s_1, \quad (4.52a)$$

$$\Gamma r_2 = j(\kappa_p^* + \zeta_2) s_2, \quad (4.52b)$$

$$\Gamma s_1 = j(\kappa_p + \zeta_4) r_1, \quad (4.52c)$$

$$\hat{\Gamma} s_2 = j(\kappa_p + \zeta_4) r_2, \quad (4.52d)$$

with

$$\hat{\Gamma} = \gamma + (-\alpha - j\delta - j\zeta_1), \quad (4.53a)$$

$$\Gamma = -\gamma + (-\alpha - j\delta - j\zeta_1). \quad (4.53b)$$

Now, from (4.52c),

$$s_1 = \frac{j(\kappa_p + \zeta_4)}{\hat{\Gamma}} r_1. \quad (4.54)$$

Inserting that into equation (4.52a), one obtains the characteristic longitudinal mode equation

$$\hat{\Gamma} r_1 = j(\kappa_p^* + \zeta_2) \frac{j(\kappa_p + \zeta_4)}{\hat{\Gamma}} r_1 \Rightarrow \boxed{\gamma^2 = (\alpha + j\delta + j\zeta_1)^2 + \kappa_{\text{eff}}^2}, \quad (4.55)$$

where we have used the definition of effective coupling coefficient namely equation (4.50). From equation (4.55), if $\gamma, \zeta_1, \kappa_{\text{eff}}$ are known, then α and δ uniquely describe a longitudinal DFB mode. $\zeta_1, \dots, \zeta_4$ and κ_{eff} can be calculated from the transverse mode analysis. In the following, how γ can be solved analytically is discussed.

Making use of boundary conditions:

Figure (4.4) shows a longitudinal section of a distributed feedback laser which can be considered generic. Let us denote

$$\begin{aligned}\hat{\rho}_l &\rightarrow \text{reflection coefficient of the left reflector;} \\ \hat{\rho}_r &\rightarrow \text{reflection coefficient of the right reflector;}\end{aligned}$$

Then, applying these boundary conditions, at the leftmost end

$$\begin{aligned}(i) \quad R(z)e^{j\beta_0 z}|_{z=-L/2} &= \hat{\rho}_l S(z)e^{-j\beta_0 z}|_{z=-L/2} \\ \Rightarrow (r_1 e^{-\gamma L/2} + r_2 e^{\gamma L/2})e^{-j\beta_0 L/2} &= \hat{\rho}_l (s_1 e^{-\gamma L/2} + s_2 e^{\gamma L/2})e^{j\beta_0 L/2}.\end{aligned}\quad (4.56)$$

and at the rightmost end

$$\begin{aligned}(ii) \quad S(z)e^{-j\beta_0 z}|_{z=L/2} &= \hat{\rho}_r R(z)e^{j\beta_0 z}|_{z=L/2} \\ \Rightarrow (s_1 e^{\gamma L/2} + s_2 e^{-\gamma L/2})e^{-j\beta_0 L/2} &= \hat{\rho}_r (r_1 e^{\gamma L/2} + r_2 e^{-\gamma L/2})e^{j\beta_0 L/2}.\end{aligned}\quad (4.57)$$

To simplify algebra let us eliminate r_1 and s_2 . From equations (4.52a), and (4.52d), respectively

$$r_1 = \frac{j(\kappa_p^* + \zeta_2)}{\hat{\Gamma}} s_1, \quad (4.58)$$

$$s_2 = \frac{j(\kappa_p + \zeta_4)}{\hat{\Gamma}} r_2. \quad (4.59)$$

Let us substitute the above expressions in (4.56),

$$\begin{aligned}&\left(\frac{j(\kappa_p^* + \zeta_2)}{\hat{\Gamma}} s_1 e^{-\gamma L/2} + r_2 e^{\gamma L/2} \right) e^{-j\beta_0 L/2} = \hat{\rho}_l \left(s_1 e^{-\gamma L/2} + \frac{j(\kappa_p + \zeta_4)}{\hat{\Gamma}} r_2 e^{\gamma L/2} \right) e^{j\beta_0 L/2} \\ \Rightarrow &\left(\frac{j(\kappa_p^* + \zeta_2)}{\hat{\Gamma}} e^{-\gamma L/2} - \hat{\rho}_l e^{-\gamma L/2} e^{j\beta_0 L} \right) s_1 = \left(\hat{\rho}_l \frac{j(\kappa_p + \zeta_4)}{\hat{\Gamma}} e^{j\beta_0 L} e^{\gamma L/2} - e^{\gamma L/2} \right) r_2.\end{aligned}\quad (4.60)$$

and also in (4.57),

$$\begin{aligned}&\left(s_1 e^{\gamma L/2} + \frac{j(\kappa_p + \zeta_4)}{\hat{\Gamma}} r_2 e^{-\gamma L/2} \right) e^{-j\beta_0 L/2} = \hat{\rho}_r \left(\frac{j(\kappa_p^* + \zeta_2)}{\hat{\Gamma}} s_1 e^{\gamma L/2} + r_2 e^{-\gamma L/2} \right) e^{j\beta_0 L/2} \\ \Rightarrow &\left(e^{\gamma L/2} - \hat{\rho}_r e^{\gamma L/2} e^{j\beta_0 L} \frac{j(\kappa_p^* + \zeta_2)}{\hat{\Gamma}} \right) s_1 = \left(e^{-\gamma L/2} \hat{\rho}_r e^{j\beta_0 L} - \frac{j(\kappa_p + \zeta_4)}{\hat{\Gamma}} e^{-\gamma L/2} \right) r_2.\end{aligned}\quad (4.61)$$

Dividing (4.61) by (4.60), one gets (Appendix E)

$$\frac{\left(1 - \rho_l \frac{\hat{\Gamma}}{j\kappa_{\text{eff}}} \right) e^{-\gamma L}}{\rho_l - \frac{\hat{\Gamma}}{j\kappa_{\text{eff}}}} = \frac{\left(\rho_r - \frac{\hat{\Gamma}}{j\kappa_{\text{eff}}} \right) e^{\gamma L}}{1 - \rho_r \frac{\hat{\Gamma}}{j\kappa_{\text{eff}}}}, \quad (4.62)$$

where we have defined the following quantities

$$\rho_l := \hat{\rho}_l e^{j\beta_0 L} \sqrt{\frac{\kappa_p + \zeta_4}{\kappa_p^* + \zeta_2}}; \quad \rho_r := \hat{\rho}_r e^{j\beta_0 L} \sqrt{\frac{\kappa_p^* + \zeta_2}{\kappa_p + \zeta_4}}. \quad (4.63)$$

We would like to solve equation (4.62) for $\frac{\hat{\Gamma}}{j\kappa_{\text{eff}}}$. Toward that end, from (4.53a) and (4.53b), let us write

$$\hat{\Gamma} - \Gamma = 2\gamma \Rightarrow \Gamma = \hat{\Gamma} - 2\gamma \Rightarrow \hat{\Gamma} = \Gamma + 2\gamma \quad (4.64)$$

Now, from (4.52a) and (4.52b), respectively, we have

$$\frac{\hat{\Gamma}}{j(\kappa_p^* + \zeta_2)} r_1 = s_1; \quad s_1 = \frac{j(\kappa_p + \zeta_4)}{\Gamma} r_1 \quad (4.65)$$

Equating on both sides, and using (4.64), results in

$$\frac{\hat{\Gamma}}{j(\kappa_p^* + \zeta_2)} = \frac{j(\kappa_p + \zeta_4)}{\Gamma} \Rightarrow \boxed{\gamma = \frac{j\kappa_{\text{eff}}}{2} \left(\frac{\hat{\Gamma}}{j\kappa_{\text{eff}}} - \frac{j\kappa_{\text{eff}}}{\hat{\Gamma}} \right)}. \quad (4.66)$$

Let us define $\frac{\hat{\Gamma}}{j\kappa_{\text{eff}}} := x$. Then equation (4.62) becomes an easily recognizable quadratic equation in x

$$\begin{aligned} \frac{(1 - \rho_l x)e^{-\gamma L}}{\rho_l - x} &= \frac{(\rho_r - x)e^{\gamma L}}{1 - \rho_r x} \\ \Rightarrow x^2(\rho^2 e^{-\gamma L} - e^{\gamma L}) - (\rho_l + \rho_r)xe^{-\gamma L} + (\rho_l + \rho_r)xe^{\gamma L} + e^{-\gamma L} - \rho^2 e^{\gamma L} &= 0, \end{aligned} \quad (4.67)$$

where we have used $\rho^2 := \rho_l \rho_r = \hat{\rho}_l \hat{\rho}_r e^{j\beta_0 L}$. Solving the above equation for x allows us to evaluate (Appendix E)

$$\begin{aligned} x - \frac{1}{x} &= \frac{\hat{\Gamma}}{j\kappa_{\text{eff}}} - \frac{j\kappa_{\text{eff}}}{\hat{\Gamma}} \\ &= \frac{\{(\rho_l + \rho_r)(1 - \rho^2)2 \cosh(\gamma L) \pm (1 + \rho^2)2\{(\rho_l - \rho_r)^2 \sinh^2 \gamma L + (1 - \rho^2)^2\}^{1/2} \sinh(\gamma L)\}}{[(1 + \rho^2)^2 - 4\rho^2 \cosh^2(\gamma L)]}. \end{aligned} \quad (4.68)$$

For convenience, let us define

$$\mathcal{D} := (1 + \rho^2)^2 - 4\rho^2 \cosh^2(\gamma L). \quad (4.69)$$

Therefore, from equation (4.66),

$$\begin{aligned} (\gamma L) &= \frac{j(\kappa_{\text{eff}} L)}{2} \times \frac{2 \sinh(\gamma L)}{\mathcal{D}} \left[(\rho_l + \rho_r)(1 - \rho^2) \cosh(\gamma L) \right. \\ &\quad \left. \pm (1 + \rho^2)\{(\rho_l - \rho_r)^2 \sinh^2(\gamma L) + (1 - \rho^2)^2\}^{1/2} \right] \end{aligned} \quad (4.70)$$

Squaring the above expression followed by slight rearrangement leads to

$$\begin{aligned}
&\Rightarrow (\gamma L \mathcal{D})^2 \\
&\quad - 2j(\kappa_{eff}L)(\gamma L \mathcal{D})(\rho_r + \rho_l)(1 - \rho^2) \sinh(\gamma L) \cosh(\gamma L) \\
&\quad - (\kappa_{eff}L)^2(\rho_l + \rho_r)^2(1 - \rho^2)^2 \sinh^2(\gamma L) \cosh^2(\gamma L) \\
&\quad + (\kappa_{eff}L)^2(1 + \rho^2)^2 \sinh^2(\gamma L)(\rho_l - \rho_r)^2 \sinh^2(\gamma L) \\
&\quad + (\kappa_{eff}L)^2(1 + \rho^2)^2 \sinh^2(\gamma L)(1 - \rho^2)^2 = 0.
\end{aligned} \tag{4.71}$$

The last three terms, when added together, evaluate to (Appendix E)

$$(\kappa_{eff}L)^2 \sinh^2(\gamma L) \mathcal{D}(1 - \rho_l^2)(1 - \rho_r^2). \tag{4.72}$$

Substituting back to equation (4.76), and factoring out the common $\mathcal{D}$,

$$\begin{aligned}
&(\gamma L)^2 \mathcal{D} - 2j(\kappa_{eff}L)(\rho_l + \rho_r)(1 - \rho^2) \sinh(\gamma L) \cosh(\gamma L)(\gamma L) + \\
&\quad (\kappa_{eff}L)^2 \sinh^2(\gamma L)(1 - \rho_r^2)(1 - \rho_l^2) = 0.
\end{aligned} \tag{4.73}$$

For our purpose, we consider a special case $\rho_r \approx 0$, which is realistic since light from the laser has to be collected from one side. It results in $\rho^2 = 0$, and $\mathcal{D} = 1$. Then equation (4.73) becomes

$$(\gamma L)^2 + (\kappa_{eff}L)^2 \sinh^2(\gamma L)(1 - \rho_l^2) - 2j(\kappa_{eff}L)\rho_l(\gamma L) \sinh(\gamma L) \cosh(\gamma L) = 0. \tag{4.74}$$

For the convenience of calculation, i.e. not to carry around an extra minus sign, we can introduce

$$-\rho_l = \rho_l(\cos \pi + j \sin \pi) = \rho_l e^{j\pi} = \tilde{\rho}_l \Rightarrow (-\rho_l)^2 = \tilde{\rho}_l^2 = \rho_l^2, \tag{4.75}$$

so that equation (4.74) becomes

$$\boxed{(\gamma L)^2 + (\kappa_{eff}L)^2 \sinh^2(\gamma L)(1 - \rho_l^2) - 2j(\kappa_{eff}L)\tilde{\rho}_l(\gamma L) \sinh(\gamma L) \cosh(\gamma L) = 0.} \tag{4.76}$$

This is the governing equation for a longitudinal mode in a distributed feed-back laser. We solve this equation for γL for particular values of $\kappa_{eff}L$, and $\tilde{\rho}_l$ to obtain a pair of values αL and δL . $\kappa_{eff}L$ comes from the calculation of the overlap between fundamental mode and the partial modes. But this is a complex valued equation, and we need to break it down into real and imaginary parts. To that end, let us first rewrite equation (4.76) as follows.

$$2(\gamma L)^2 + 2(\kappa_{eff}L)^2 \sinh^2(\gamma L)(1 - \rho_l^2) + 2j\tilde{\rho}_l(\kappa_{eff}L)(\gamma L) \sinh(2\gamma L) = 0. \tag{4.77}$$

Let us assign

$$\gamma l = a + jb; \quad \tilde{\rho}_l = -\rho_l = -(c + jd); \quad (\kappa_{eff}L) = q + jp \tag{4.78}$$

After a considerable amount of algebra the real and imaginary parts of equation (4.77) are separated as follows (Appendix E)

Real Part Equation :

$$\begin{aligned}
& 2(a^2 - b^2) \\
& + \cos(2b) \cosh(2a) [m(q^2 - p^2) + 2qpn] \\
& + \sin(2b) \sinh(2a) [-2qpm + n(q^2 - p^2)] \\
& + \cos(2b) \sinh(2a) [2b(qc - pd) + 2a(qd + pc)] \\
& + \sin(2b) \cosh(2a) [2a(qc - pd) - 2b(qd + pc)] \\
& - m(q^2 - p^2) \\
& - 2qpn = 0.
\end{aligned} \tag{4.79}$$

Imaginary Part Equation :

$$\begin{aligned}
& 2(a^2 - b^2) \\
& + \cos(2b) \cosh(2a) [-n(q^2 - p^2) + 2qpm] \\
& + \sin(2b) \sinh(2a) [m(q^2 - p^2) + 2qpn] \\
& + \cos(2b) \sinh(2a) [-2a(qc - pd) + 2b(qd + pc)] \\
& + \sin(2b) \cosh(2a) [2b(qc - pd) + 2a(qd + pc)] \\
& + n(q^2 - p^2) \\
& - 2qpm = 0.
\end{aligned} \tag{4.80}$$

4.4.1 Transmission and Reflection Spectrum

Now we are in a position to derive the expressions for the transmission and the reflection spectrum of our DFB laser. Let us recall equations (4.47) and (4.48)

$$\begin{aligned}
R' + (-\alpha - j\delta - j\zeta_1)R &= j(\kappa_p^* + \zeta_2)S \\
-S' + (-\alpha - j\delta - j\zeta_3)S &= j(\kappa_p + \zeta_4)R
\end{aligned}$$

Taking derivatives of both the equations with respect to z (propagation direction), results in (Appendix E)

$$R'' - \gamma^2 R = 0, \tag{4.81a}$$

$$S'' - \gamma^2 S = 0. \tag{4.81b}$$

As mentioned earlier, these two equations admit solutions of the form (equations (4.51a) and (4.51b))

$$\begin{aligned}
R(z) &= r_1 e^{\gamma z} + r_2 e^{-\gamma z}, \\
S(z) &= s_1 e^{\gamma z} + s_2 e^{-\gamma z}.
\end{aligned}$$

Evaluating the above two equations at the boundaries $z = -L/2$, and $z = L/2$, and following some substantial algebra one arrives at (Appendix E)

$$R(L/2) = \frac{j\gamma(\kappa_p + \zeta_4)}{j(\kappa_p + \zeta_4)[\gamma \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)] + \kappa_{\text{eff}}^2 r \sinh(\gamma L)}, \quad (4.82)$$

if $r = 0$,

$$R(L/2) = \frac{\gamma}{\gamma \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)}, \quad (4.83)$$

which is the same as in derived [23] for zero reflectivity on both the facets. Now, referring to Figure (4.4), if we set $R(-L/2) = 1$, then $|R(L/2)|^2$ will denote the power transmission coefficient of our distributed feedback laser cavity.

Starting from equation (E.27), one can show that

$$\frac{S(-L/2)}{R(-L/2)} = \frac{\gamma r R(L/2) + j(\kappa_p + \zeta_4) \sinh(\gamma L)}{\gamma \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)}. \quad (4.84)$$

For $R(-L/2) = 1$, $|S(-L/2)|^2$ will denote the power reflection coefficient of our distributed feedback laser cavity. Here, also for $r = 0$, the reflection coefficient takes the expected form from [23]

$$S(-L/2) = \frac{j(\kappa_p + \zeta_4) \sinh(\gamma L)}{\gamma \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)}. \quad (4.85)$$

4.5 Conclusion

This chapter presents a thorough electromagnetic analysis of a distributed feedback laser. In principle, the analysis applies to any DFB laser. The following chapter uses the results established here to numerically solve the transverse and the longitudinal modes of a DFB laser which employs a specific cross section and grating profile. The only adjustment we make, as we shall see, is to take into account both the transverse coordinates instead of one, since for some of the layers none of the directions can be assumed to be infinitely uniform.

Chapter 5

Numerical Analysis of a DFB Laser

5.1 Introduction

In this chapter, we study the transmission spectrum of a DFB laser with a specific cross-section and a grating profile. First, we present the transverse profile of the fundamental guided mode obtained by a standard commercial 2-D mode solver. In this step, the grating region is considered homogeneous with an averaged over refractive index. Next, we simulate the transverse profile of the partial modes. The same 2-D mode solver can be used in this case also. But the Helmholtz equation for the structure now has a source term in the form of the fundamental mode from the previous step. From the overlap integral between the fundamental mode and the partial modes, one can calculate all the Streifer coefficients, and hence the modified coupling coefficient. This allows us to solve the complex longitudinal mode equation for a particular value of the mirror reflection coefficient. Finally, we are in a position to plot the transmission and the

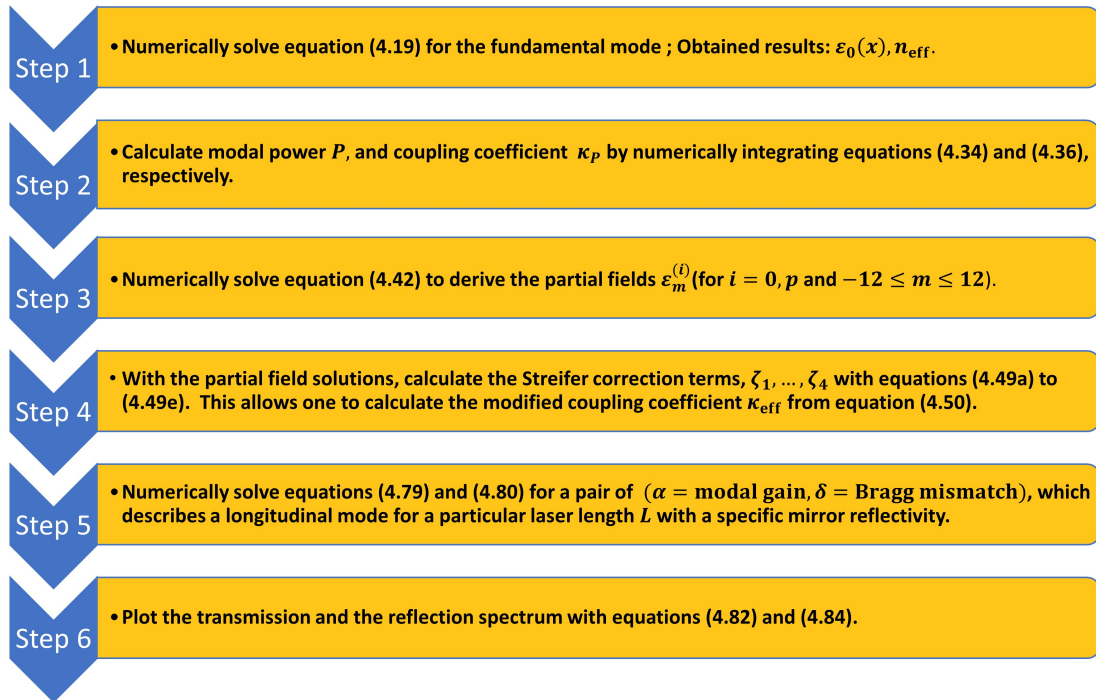


FIGURE 5.1: Workflow related with the numerical analysis of a generic DFB Laser.

reflection spectrum of the laser. Figure (5.1) shows the workflow of the analysis just stated.

5.2 Fundamental Mode Profile

Figure (5.2) shows the cross-section of the distributed feedback laser which is to be used for taper assisted vertical integration at 1310 nm.

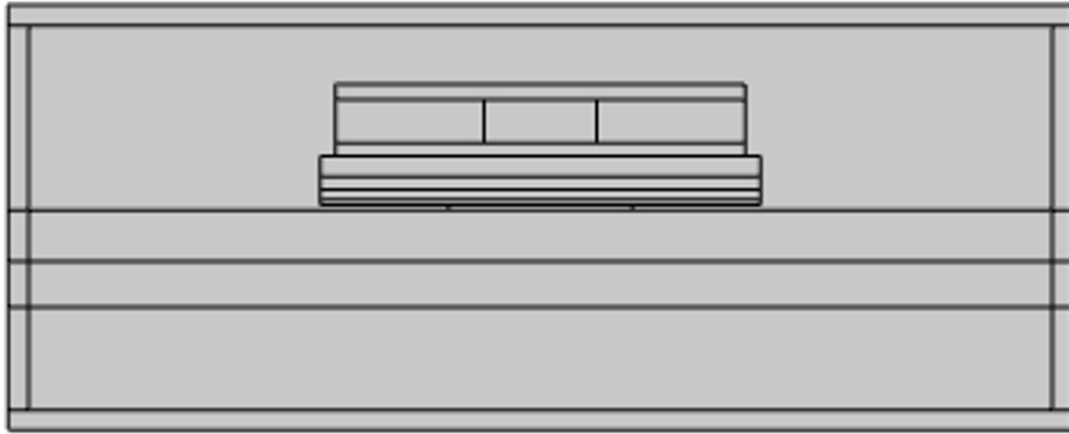


FIGURE 5.2: The cross-section of the TAVI DFBL. For confidential reasons, the layers are not labeled.

Figure (5.3) shows the specific grating profile it employs along with the geometric parameters. The trapezoidal-shaped grating is formed by surface corrugation-

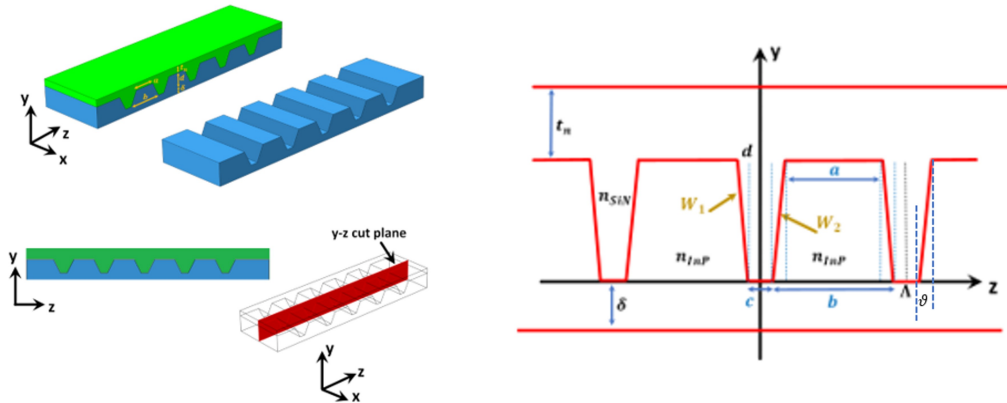


FIGURE 5.3: The grating profile of the TAVI DFBL. The geometric parameters of the trapezoid-shaped grating are illustrated with respect to its y-z cut plane. Image courtesy Sina Aghili

tion with the tooth shape functions

$$w_1 = -w_2 = -\frac{b-a}{2d}y - \frac{\Lambda}{d} + \frac{b}{2}, \quad (5.1)$$

where Λ is the grating period, a and b are the lengths of the bigger and smaller arms of the trapezoid, respectively, d corresponds to the thickness of the grating region (the etch depth), ν corresponds to the side-angle. The ratio $\gamma = \frac{b}{\Lambda}$ is known as the duty cycle. For 2-D modal analysis, the grating region is approximated as a homogeneous region with refractive index

$$n_{\text{seg}} = \sqrt{n_1^2 \left(\gamma - \frac{b-a}{\Lambda d} y \right) + n_2^2 \left((1-\gamma) + \frac{b-a}{\Lambda d} y \right)}, \quad (5.2)$$

which follows from equation (4.10). Here, n_1 and n_2 correspond to the component indices which make up the grating region. If we start with the following values for the grating parameters.

$$\gamma = 0.75, \quad \Lambda = 600\text{nm}, \quad d = 420\text{nm} \quad \nu = 10^\circ, \quad (5.3)$$

and

$$b = \gamma * \Lambda = 0.75 * 600 \text{ nm} = 450 \text{ nm}, \quad (5.4a)$$

$$a = b - 2 * d * \tan \nu = 301.885 \text{ nm}, \quad (5.4b)$$

then it follows from equation (5.2) that $n_{\text{seg}} = 2.909$. With the refractive index

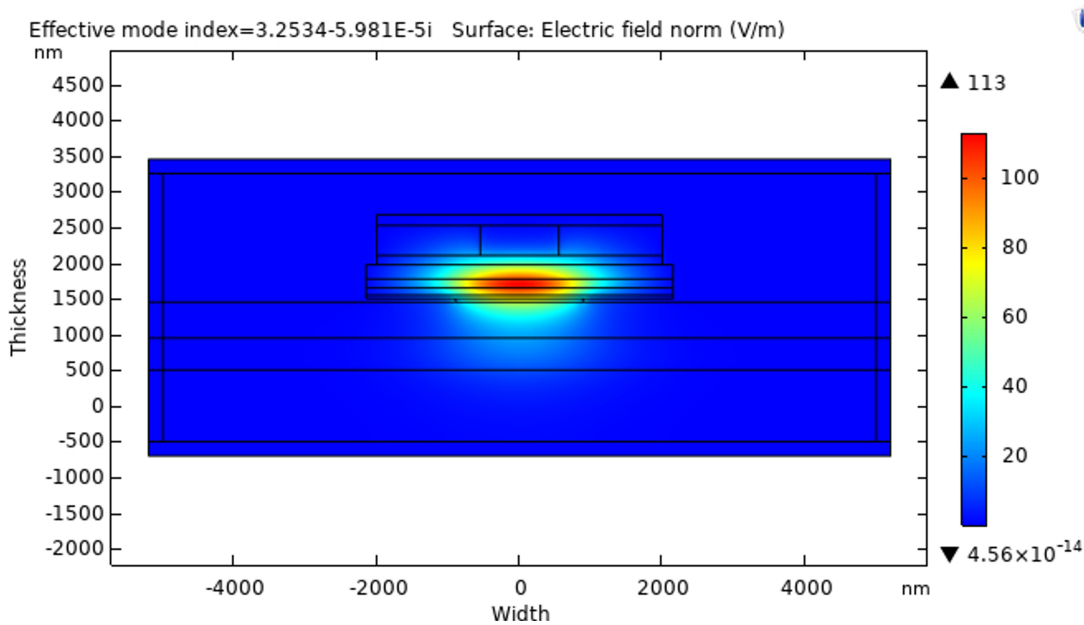


FIGURE 5.4: Fundamental mode of the laser under consideration

of the grating region, we can numerically calculate the fundamental mode of the structure. The governing equation for the fundamental mode analysis, from the previous chapter, is (4.19)

$$\frac{d^2 \varepsilon_0(x, y)}{dx^2} + \frac{d^2 \varepsilon_0(x, y)}{dy^2} + [k_0^2 n_0^2(x, y) - \beta^2] \varepsilon_0(x, y) = 0.$$

This is an eigenvalue problem ($Ax = \lambda x$) with the effective refractive indices (β 's) as the eigenvalues. We have used the RF module from COMSOL Multiphysics. The "Mode Analysis" study was carried out with absorbing boundary

conditions. Figure (5.4) shows the result with the real part of the effective index of the quasi-TE mode found to be 3.25.

From the definition of the Bragg grating, one can calculate

$$\Lambda = N \times \frac{\lambda}{2n_{\text{eff}}} \Rightarrow n_{\text{eff}} = 3 \times \frac{1310 \text{ nm}}{2 \times 600 \text{ nm}} = 3.275, \quad (5.5)$$

which is very close to our numerically found value. Here, N corresponds to the grating order, which in our case equals to 3. One can calculate the coupling coefficient, once the fundamental field mode has been calculated, as per equation (3.36)

$$\kappa_p = \frac{k_0^2}{\beta_0 P} \int_0^g A_p(x) \varepsilon_0^2(x) dx = -1241.6347 + j0.0203 \text{ m}^{-1}$$

5.3 Partial Wave Analysis

The partial waves are calculated according to the equation (4.42)

$$\frac{d^2 \varepsilon_m^{(i)}(x, y)}{dx^2} + \frac{d^2 \varepsilon_m^{(i)}(x, y)}{dx^2} + [k_0^2 n_0^2(x, y) - \beta_m^2] \varepsilon_m^{(i)}(x, y) = -k_0^2 A_{m-i}(x, y) \varepsilon_0(x, y),$$

$$\forall m, \quad \begin{matrix} m \neq i, \\ i=0, p \end{matrix}$$

which is of the form $A\mathbf{x} = \mathbf{b}$. This part of the analysis was carried out by Sina Aghili, a PhD student with our group, by using the PDE (partial differential equation) module from COMSOL for $-12 \leq m \leq 12$. Equation (4.42) was solved for the structure in Figure (5.2) with phase matched layer boundary conditions. The simulation results are shown in Figure (5.5). Although the partial field waves obtained in this particular figure are for the duty cycle $\gamma = 0.655$, they capture the essence which apply to the partial waves associated with any duty cycle.

The results obtained from this step for the duty cycle $\gamma = 0.75$ are the Streifer coefficients $\zeta_1, \dots, \zeta_4$, which are listed in Table (5.1).

Streifer Coefficients	
ζ_1	$-1236.8126 + j 0.0203$
ζ_2	$-1630.8282 + j 236.6659$
ζ_3	equal to ζ_1
ζ_4	$-1630.8277 + j 192.7949$

TABLE 5.1: Streifer coefficients obtained from the partial wave analysis.

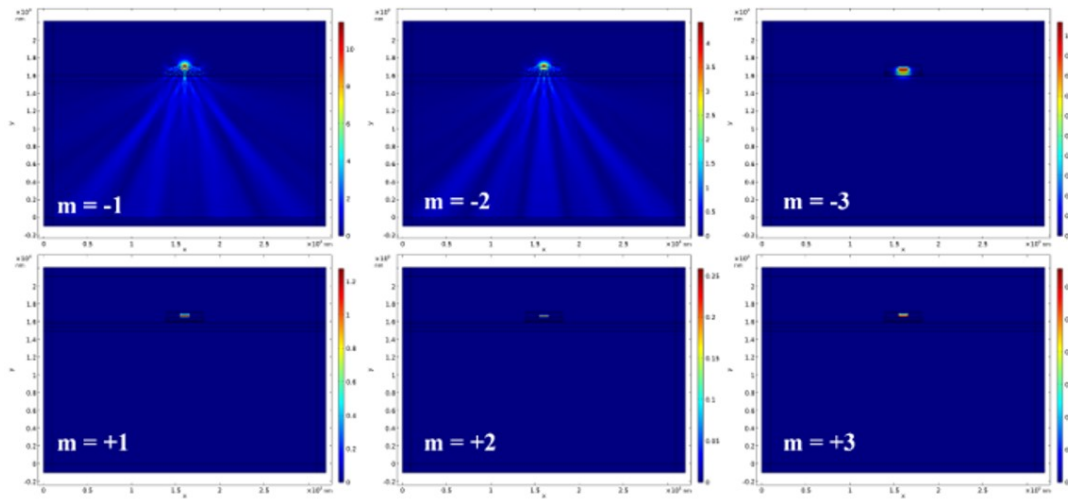


FIGURE 5.5: $m = \pm 1, \pm 2, \pm 3$ partial wave for duty cycle $\gamma = 0.655$. Image courtesy Sina Aghili

With these coefficients calculated, one can determine the effective coupling coefficient as follows

$$\kappa_{\text{eff}} = \sqrt{(\kappa_p^* + \zeta_2)(\kappa_p + \zeta_4)} = 389.19 + j192.79 \text{ m}^{-1}. \quad (5.6)$$

5.4 Longitudinal Mode Profile

The longitudinal mode of a distributed feedback laser is characterized by its modal gain (α) and Bragg detuning (δ) for a particular cavity length. These parameters can be calculated using equation (4.55)

$$\gamma^2 = (\alpha + j\delta + j\zeta_1)^2 + \kappa_{\text{eff}}^2.$$

To that end, first we need to solve equation (4.76) for γ for some cavity length L

$$(\gamma L)^2 + (\kappa_{\text{eff}} L)^2 \sinh^2(\gamma L)(1 - \rho_l^2) - 2j(\kappa_{\text{eff}} L)\tilde{\rho}_l(\gamma L) \sinh(\gamma L) \cosh(\gamma L) = 0.$$

Since this equation is complex valued, one needs to solve the real part equation and the imaginary part equation separately as given in equations (4.84) and (4.85).

Real Part Equation :

$$\begin{aligned} & 2(a^2 - b^2) \\ & + \cos(2b) \cosh(2a) [m(q^2 - p^2) + 2qpn] \\ & + \sin(2b) \sinh(2a) [-2qpm + n(q^2 - p^2)] \\ & + \cos(2b) \sinh(2a) [2b(qc - pd) + 2a(qd + pc)] \\ & + \sin(2b) \cosh(2a) [2a(qc - pd) - 2b(qd + pc)] \\ & - m(q^2 - p^2) \\ & - 2qpn = 0. \end{aligned}$$

Imaginary Part Equation :

$$\begin{aligned}
 & 2(a^2 - b^2) \\
 & + \cos(2b) \cosh(2a) [-n(q^2 - p^2) + 2qpm] \\
 & + \sin(2b) \sinh(2a) [m(q^2 - p^2) + 2qp n] \\
 & + \cos(2b) \sinh(2a) [-2a(qc - pd) + 2b(qd + pc)] \\
 & + \sin(2b) \cosh(2a) [2b(qc - pd) + 2a(qd + pc)] \\
 & + n(q^2 - p^2) \\
 & - 2qpm = 0.
 \end{aligned}$$

The numerical solution of the above two equations results in numerous values for α and δ . In this chapter, we report such pairwise solutions (Figure (5.6)) for the specific case of $L = 1$ mm, and reflectivity $\rho_l = 0.565e^{j\pi}$.

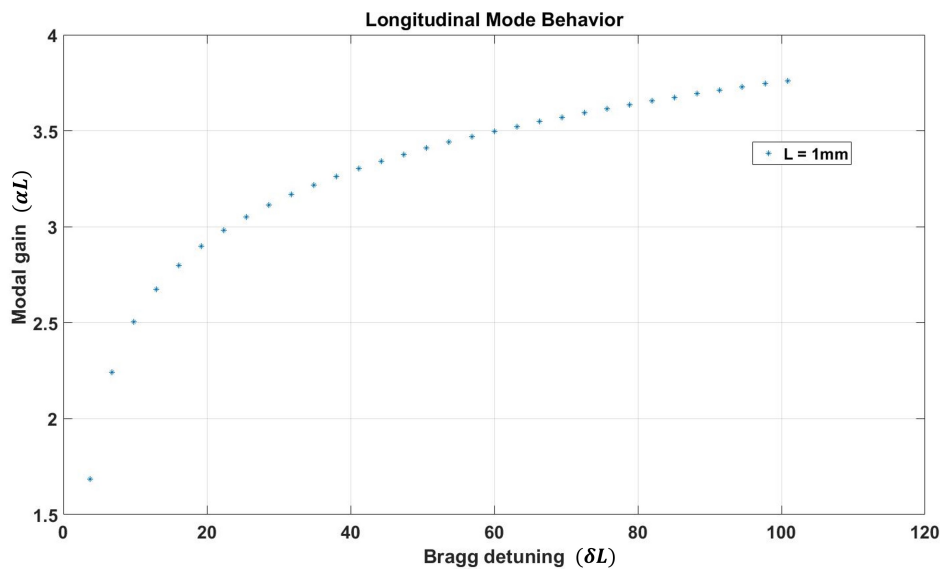


FIGURE 5.6: Pair of solution for the modal gain and the Bragg detuning cavity length = 1 mm, duty cycle = 0.75 and mirror reflectivity = -0.565.

One such pair of solutions is $(\alpha L, \delta L) = (1.6843, 3.72365)$.

5.5 Transmission & Reflection Spectrum

With the results from the preceding sections (5.2, 5.3, 5.4), we are in a position to calculate the transmission and the reflection spectrum of our laser. The transmission coefficient of our DFB laser cavity is given by

$$\text{Transmission} = \left| \frac{R(L/2)}{R(-L/2)} \right|^2, \quad (5.7)$$

where, according to equation (4.82),

$$\frac{R(L/2)}{R(-L/2)} = \frac{j\gamma(\kappa_p + \zeta_4)}{j(\kappa_p + \zeta_4) [\gamma \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)] + \kappa_{\text{eff}}^2 r \sinh(\gamma L)}.$$

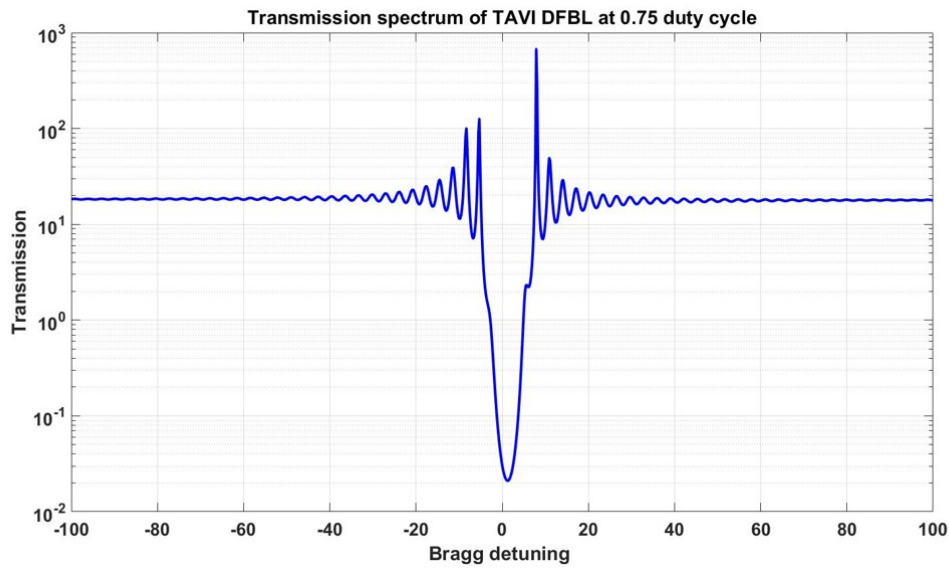


FIGURE 5.7: Transmission spectrum of our TAVI distributed feedback laser.

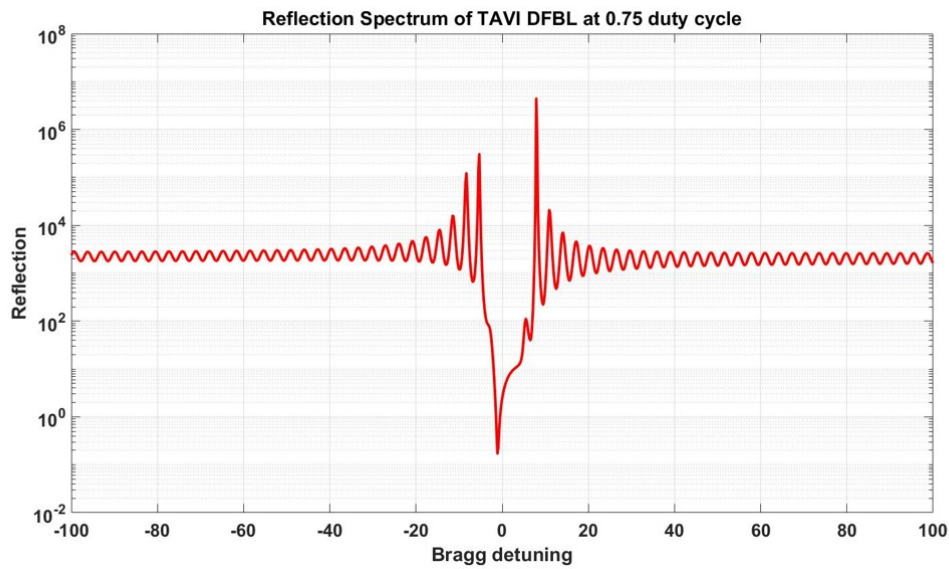


FIGURE 5.8: Reflection spectrum of our TAVI DFB laser.

Figure (5.7) plots the right hand side of the above equation with respect to δL . Here, the cavity length is $L = 1\text{mm}$ and the modal gain is $\alpha L = 1.6843$ from the previous step.

The reflection coefficient of our DFB laser cavity is defined as

$$\text{Reflection} = \left| \frac{S(-L/2)}{R(-L/2)} \right|^2, \quad (5.8)$$

where, according to equation (4.84)

$$\frac{S(-L/2)}{R(-L/2)} = \frac{\gamma r R(L/2) + j(\kappa_p + \zeta_4) \sinh(\gamma L)}{\gamma \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)}.$$

Without loss of generality, we can set $R(-L/2) = 1$, which enables us to plug the expression of $R(L/2)$ from equation (4.82) in the above expression. Similar to the case for transmission, we plot the right hand side of the above equation with respect to δL for cavity length $L = 1\text{mm}$ and modal gain $\alpha L = 1.6843$ (Figure (5.8)).

5.6 Conclusion

In this chapter, we have seen how to numerically investigate the behaviour of a distributed feedback laser. Although the results were illustrated with respect to a particular structure, the procedure for obtaining the results apply to any generic DFB laser. The results we have demonstrated in this chapter are obtained for a particular value of duty cycle of the grating. The logical extension of this work such as how the variation of the duty cycle of the grating may affect the coupling coefficient, Streifer coefficients, the transmission and the reflection spectrum of this particular laser is already being actively pursued by other members in our group.

Chapter 6

Final Remarks

In this dissertation, I have documented my contribution in two different research projects with which I was actively involved over the course of my masters program. The common motivation behind both the projects is to develop photonic integrated technologies at the operating wavelengths of fiber based optical communication (1550 nm and 1310 nm) systems.

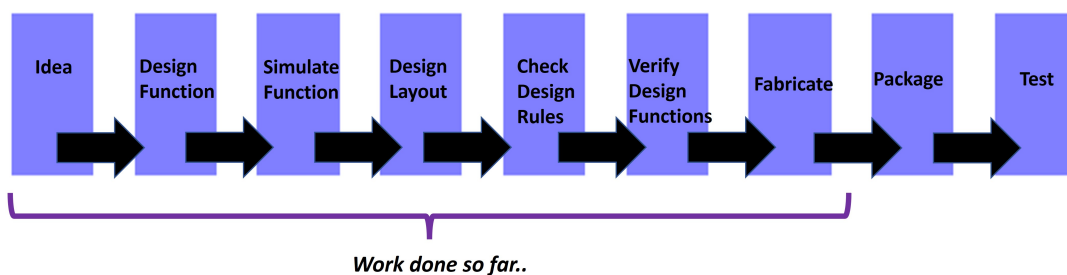


FIGURE 6.1: The workflow involved in a research project which seeks to demonstrate validation of an idea with respect to a photonic integrated circuit.

To reiterate my contribution in the continuous variable quantum key distribution project, I have proposed a design for a photonic integrated transmitter and a photonic integrated receiver. I have mathematically analyzed the whole transceiver system based on the optical transfer matrices of the building blocks available from the foundry Fraunhofer Heinrich Hertz Institute. Based on the mathematical model, I assessed the feasibility of the our scheme by estimating the link budget, and subsequently running circuit level simulations. Finally, I translated the circuit schematic to mask layouts based on which the foundry fabricated our chips. The specialty of our approach is that all the active and passive components are monolithically integrated on an InGaAsP/InP material system. Besides, the transceivers are potentially capable of supporting more than one modulation formats. However, my effort in this project does not wrap up the whole story. Figure (6.1) highlights the various stages of a proof-of-principle demonstration project involving a photonic integrated circuit. It also shows where we stand at the point of writing this thesis. Most importantly, in order to verify our design claims, we still need to characterize the components on the chip following the actual experiment.



FIGURE 6.2: (a) Standard package from Technobis (Tipps) for photonic integrated circuits. (b) A customized package from the same company. Their PDK is available through Phoenix design kit

Testing and characterizing a photonic integrated circuit requires optical probes as well as electrical probes to be connected to the die for the measurement of optical, and electrical DC and RF parameters. While it is possible to test and characterize a bare-die photonic integrated circuit on a probe station, testing a packaged die is much easier since there are already connectorized inputs and outputs defined by electrical pins and optical fiber pigtails [99]. Besides, a durable packaged module also facilitates building a prototype device. Among other things, the packaging module should include the electronic driver circuits for the active components, thermo-electric cooler along with the chips themselves. As of writing this thesis, we are exploring options for the packaging of our photonic chips. The companies which provide packaging services, such as Cordon Electronics (Italy), Technobis (The Netherlands), usually offer either a generic standard package option or a customized package option. Figure (6.2) shows images of a standard and a custom package available from Technobis (Tipps). Most often, to maximize performance, a fabricated PIC requires custom packaging. In that case, it forms the most expensive stage of a project involving PICs, since owing to MPW (multi project wafer) runs the cost of PIC design and fabrication has significantly reduced. Moreover, now-a-days, packaging companies offer their own PDK's which can be integrated with a PIC design environment. For example, Cordon Electronics has a PDK [100], which contains the required layout information to fit a photonic integrated circuit into their standard package or customized package. This PDK is available through Synopsis design kit.

Since our project is about the proof-of-principle demonstration of a quantum key distribution protocol, namely CV QKD, we also need to develop analytic tools for the extensive classical post-processing stage. The post-processing steps enable us to distill a completely secure key from a partially secure key. In this regard, we may have to write states-of-the-art error correcting codes such as multi-edge-type (MET) LDPC codes as mentioned in [54].

Similar to the CV QKD project, my contribution in the distributed feedback laser project constitutes part of a larger effort. The taper-assisted vertical integration (TAVI) distributed feedback laser makes use of a characteristic cross-section, and grating profile. My work embodies full-length mathematical analysis of the distributed feedback laser culminating in the derivation of the expressions for the transmission and the reflection spectrum of the laser. To extract the parameters on which these expressions depended, I performed the relevant transverse mode analysis (fundamental and partial wave calculation) and the longitudinal mode analysis. The transverse mode analysis gave us the Streifer coefficients, and the modified coupling coefficient, while the longitudinal mode analysis provided modal gain, and Bragg detuning. I was amply supported by Boris Rosenstein Levin and Sina Aghili while working on this project, especially in the partial wave calculation step. The future line of work will include observing the effect of the transmission (/reflection) of the laser subject to the variation of the grating parameters. The optimized design will go into fabrication, and the subsequent experimental characterization of the fabricated lasers will verify the design claims.

Appendix A

Private Classical Communication

The following discussion is extracted from [101].

A.1 Classical Shannon Theory

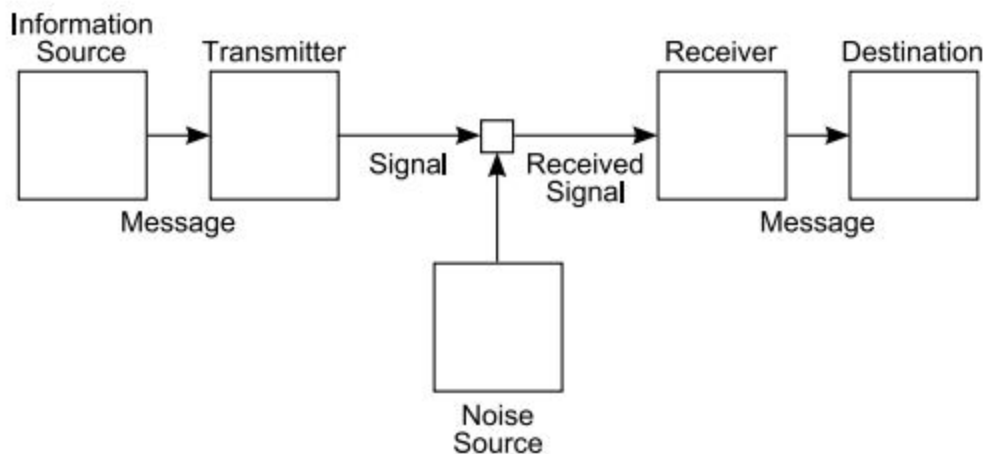


FIGURE A.1: Shannon's model for the transmission of a message

Claude Shannon laid the foundation of modern day (classical) information theory in his seminal 1948 paper [30]. In this paper he mentioned "the fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point." To address this problem he put forward two fundamental theorems, with proofs, in terms of two information-processing tasks.

1. The *source coding theorem*, also called the *noiseless coding theorem*, states at what rate (with what efficiency) a message can be compressed so that it still conveys essentially the same information. According to Shannon the minimum rate is given by *entropy*.
2. The *channel coding theorem*, also called the *noisy coding theorem*, states at what rate (with what efficiency) a message can be transmitted over a noisy channel so that upon detection it can be recovered reliably. According to Shannon the maximum rate is given by (*channel*) *capacity*.

In general, entropy < capacity refers to the feasibility criterion for carrying out a communication protocol. Figure (A.1) illustrates how Shannon envisioned a point to point communication system. This diagram puts the tasks of source coding and channel coding in a better context. As we shall see in the following, the effects of these two tasks in a sense balance each other out. With source coding we are basically getting rid of redundancies from our message, whereas with channel coding we are adding redundancies to our message to protect against errors. Before going into the specifics of Shannon's ideas, some definitions and concepts are reviewed first.

A.1.1 Notions from discrete probability

Entropy: The entropy of a discrete random variable X with realizations x 's according to the probability distribution $p_X(x)$ is

$$H(X) \equiv - \sum_{x \in \mathcal{X}} p_X(x) \log(p_X(x)) \quad (\text{A.1})$$

where $\mathcal{X}$ is the support of X . The logarithm function has a base 2 (throughout this chapter). From the definition, it follows: (i) $H(X) \geq 0$ with equality if and only (iff) X is deterministic; (ii) $H(X) \leq \log|\mathcal{X}|$ with equality iff X is uniformly distributed; (iii) $H(X)$ is concave in $p_X(x)$, with (i) and (ii) constituting the two extreme cases.

Joint Entropy: For two discrete random variables X and Y with pairwise realizations (x, y) according to the joint distribution $p_{X,Y}(x, y)$, the joint entropy is given by

$$H(X, Y) \equiv - \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} p_{X,Y}(x, y) \log p_{X,Y}(x, y) \quad (\text{A.2})$$

where $\mathcal{X}$ and $\mathcal{Y}$ are supports for X and Y respectively. The above definition is generalized to the case of n – number of random variables

$$H(X_1, \dots, X_n) = - \sum_{x_1 \in \mathcal{X}_1} \dots \sum_{x_n \in \mathcal{X}_n} p_{X_1, \dots, X_n}(x_1, \dots, x_n) \log p_{X_1, \dots, X_n}(x_1, \dots, x_n) \quad (\text{A.3})$$

Conditional Entropy: If X and Y are two discrete random variables with joint probability distribution $p_{X,Y}(x, y) = p_{X|Y}(x|y)p_Y(y)$, then their conditional entropy is given by

$$\begin{aligned} H(X|Y) &= \sum_y p_Y(y) H(X|Y=y) = - \sum_y p_Y(y) \sum_x p_{X|Y}(x|y) \log p_{X|Y}(x|y) \\ &= \sum_{x,y} p_{X,Y}(x, y) \log p_{X|Y}(x|y) \quad (\text{A.4}) \end{aligned}$$

Conditioning does not increase entropy: $H(X) \geq H(X|Y)$ with equality iff X and Y are independent. From the rule in probability theory $p_{X,Y}(x, y) = p_{Y|X}(y|x)p_X(x) = p_{X|Y}(x|y)p_Y(y)$ it also follows

$$H(X, Y) = H(X) + H(Y|X) = H(Y) + H(X|Y) \quad (\text{A.5})$$

The above relation can be extended to the n -tuple case

$$H(X_1, \dots, X_n) = H(X_1) + H(X_2|X_1) + \dots + H(X_n|X_{n-1}, \dots, X_1) \quad (\text{A.6})$$

In general, the joint entropy is sub-additive

$$H(X_1, \dots, X_n) \leq \sum_{i=1}^n H(X_i) \quad (\text{A.7})$$

with equality iff $X_1, \dots, X_n$ are independent.

Mutual Information: The mutual information quantifies correlation between two discrete random variables X and Y with joint probability distribution $p_{X,Y}(x, y)$, and is defined as

$$I(X; Y) \equiv H(X) + H(Y) - H(X, Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \quad (\text{A.8})$$

From this definition, it is obvious $I(X; Y)$ is symmetric in X and Y . Moreover, if $p_X(x)$ and $p_Y(y)$ correspond to marginal distributions of X and Y , then one can also write

$$I(X; Y) = \sum_{x,y} p_{X,Y}(x, y) \log \left(\frac{p_{X,Y}}{p_X(x)p_Y(y)} \right) \quad (\text{A.9})$$

Therefore, if X and Y are independent, their mutual information $I(X; Y) = 0$.

A.1.2 Classical Typicality

Two notions of classical typicality ([102], [101]) can be defined: strong typicality (based on the relative frequency of the realizations of the random variables), and weak typicality (based on the convergence of sample entropy toward true entropy). Weak typicality can be derived from strong typicality. However, for the discussions of this dissertation the concept of weak typicality suffices. The theory of typicality is built upon the law of large number from probability theory.

Weak Typicality

Let us consider a random variable X whose realizations x are letters in an alphabet $\mathcal{X} = (a_1, \dots, a_{|\mathcal{X}|})$. The probability with which X realizes any particular letter a is given by the distribution $p_X(a)$. Let x^n be a sequence generated by sampling X independently n -times

$$x^n \equiv x_1 x_2 \dots x_n \quad (\text{A.10})$$

Such a sequence is independent and identically distributed (i.i.d.). The probability of any such sequence occurring is given by the product distribution

$$p_{X^n}(x^n) = p_X(x_1)p_X(x_2)\dots p_X(x_n) = \prod_{i=1}^n p_X(x_i) \quad (\text{A.11})$$

Let $N(a_i|x^n)$ be the number of occurrences of the letter a_i in the sequence x^n . Then

$$x^n \rightarrow \underbrace{a_1 \dots a_1}_{N(a_1|x^n)} \underbrace{a_2 \dots a_2}_{N(a_2|x^n)} \dots \underbrace{a_{|\mathcal{X}|} \dots a_{|\mathcal{X}|}}_{N(a_{|\mathcal{X}|}|x^n)} \quad (\text{A.12})$$

from which it follows that

$$\prod_{i=1}^n p_X(x_i) = \prod_{j=1}^{|\mathcal{X}|} p_X(a_j)^{N(a_j|x^n)} \quad (\text{A.13})$$

Sample Entropy: The sample entropy $\bar{H}(x^n)$ of a sequence x^n with respect to a probability distribution $p_X(x)$ is

$$\bar{H}(x^n) \equiv -\frac{1}{n} \log(p_{X^n}(x^n)) = -\sum_{i=1}^{|\mathcal{X}|} \frac{N(a_i|X^n)}{n} \log(p_X(a_i)) \quad (\text{A.14})$$

Weak Typical Set: The δ - weakly typical set $T_\delta^{X^n}$ is the set of all those sequences the sample entropy for which has a deviation of δ ($\delta > 0$) from their true entropy $H(X)$

$$T_\delta^{X^n} \equiv \{x^n : |\bar{H}(x^n) - H(X)| \leq \delta\} \quad (\text{A.15})$$

These concepts can be extended to the joint random sequence and conditional random sequences.

(Weak) Jointly Typical Set: If two random variables X and Y with joint probability distribution $p_{X,Y}(x,y)$ are sampled n - times independently to realize $x^n = x_1 \dots x_n$ and $y^n = y_1 \dots y_n$, then the δ - jointly typical set $T_\delta^{X^n Y^n}$ consists of all δ - jointly typical sequences

$$T_\delta^{X^n Y^n} \equiv \{(x^n, y^n) : |\bar{H}(x^n, y^n) - H(X, Y)| \leq \delta, x^n \in T_\delta^{X^n}, y^n \in T_\delta^{Y^n}\} \quad (\text{A.16})$$

where the joint sample entropy $\bar{H}(x^n, y^n)$ of these two sequences is

$$\bar{H}(x^n, y^n) \equiv -\frac{1}{n} \log(p_{X^n, Y^n}(x^n, y^n)) \quad (\text{A.17})$$

(Weak) Conditionally Typical Set: Let us consider two random variables X and Y with joint probability distribution of the form $p_{X,Y}(x,y) = p_X(x)p_{Y|X}(y,x)$. Then for n independent realizations of X giving rise to the sequence $x^n = x_1 \dots x_n$, there will be a corresponding sequence for Y such as $y^n = y_1 \dots y_n$ according to $p_{Y|X}(y,x)$. The δ - conditionally typical set $T_\delta^{Y^n|x^n}$ contains all the y^n sequences which would typically arise for a particular $x^n \in \mathcal{X}^n$ in the sense that their conditional sample entropy is within δ - distance of their true entropy

$$T_\delta^{Y^n|x^n} \equiv \{y^n : \bar{H}(y^n|x^n) - H(Y|X) \leq \delta\} \quad (\text{A.18})$$

where

$$\bar{H}(y^n|x^n) = -\frac{1}{n} \log p_{Y^n|X^n}(y^n|x^n) \quad (\text{A.19})$$

with

$$p_{Y^n|X^n}(y^n|x^n) \equiv p_{Y|X}(y_1|x_1) \dots p_{Y|X}(y_n|x_n) \quad (\text{A.20})$$

Strong Typicality

Strong typicality is built on the concept of *type* and *type class* **Type**: The type $t_{x^n}(a)$ of a sequence x^n lists the relative frequency of occurrences for each symbol in $\mathcal{X}$

$$t_{x^n}(a) \equiv \frac{1}{n}N(a|x^n) \quad \forall a \in \mathcal{X} \quad (\text{A.21})$$

Type Class: The type class $T_t^{X^n}$ groups together all the sequences with length n and type t (same empirical distribution of letters)

$$T_t^{X^n} \equiv \{x^n \in \mathcal{X}^n : t_{x^n} = t\} \quad (\text{A.22})$$

Typical Type: A typical type (τ_δ) is the set for all types t for which the empirical distribution of all the letters $\mathcal{X}$ has a maximum deviation of δ ($\delta > 0$) from their true distribution $p_X(a)$

$$\tau_\delta \equiv \{t : \forall x \in \mathcal{X}, |t(a) - p_X(a)| \leq \delta\} \quad (\text{A.23})$$

Strongly Typical Set: The set of all sequences for which the empirical distribution of all the letters in $\mathcal{X}$ has a maximum deviation of δ ($\delta > 0$) from their true distribution p_X is called a δ -strongly typical set.

$$T_\delta^{X^n} \equiv \{x^n : \forall x \in \mathcal{X}, \left| \frac{1}{n}N(a|x^n) - p_X(a) \right| \leq \delta\} \quad (\text{A.24})$$

The δ -strongly typical set consists of all the sequences belonging to the typical type classes

$$T_\delta^{X^n} = \bigcup_{t \in \tau_\delta} T_t^{X^n} \quad (\text{A.25})$$

The notions of strong typicality extend to the case of joint and conditional random variables as well.

Strong Jointly Typical Set Let us consider two sequences $x^n \in \mathcal{X}^n$ and $y^n \in \mathcal{Y}^n$ generated from sampling two random variables X and Y which have joint probability distribution $p_{X,Y}(x,y)$. Then δ -strong jointly typical set $T_\delta^{X^n Y^n}$ collects all the sequences whose empirical distribution (Type) has maximum deviation δ from their true distribution (δ -jointly typical sequences)

$$T_\delta^{X^n Y^n} \equiv \{x^n, y^n : \forall (x,y) \in \mathcal{X} \times \mathcal{Y}, |t_{x^n y^n}(a,b) - p_{X,Y}(a,b)| \leq \delta\} \quad (\text{A.26})$$

where

$$t_{x^n y^n}(x,y) \equiv \frac{1}{n} \log N(a,b|x^n, y^n) \quad (\text{A.27})$$

is the type for joint random sequence $x^n y^n$ and a, b are two representative letters from alphabets $\mathcal{X}$ and $\mathcal{Y}$ respectively.

Strong Conditionally Typical Set Let us consider two random variables X and Y with joint probability distribution $p_{XY}(x,y) = p_X(x)p_{Y|X}(y,x)$. Then for n

independent realizations of X giving rise to the sequence $x^n \in T_\delta^{X^n}$, there will be a corresponding $y^n \in T_\delta^{Y^n}$ according to $p_{Y|X}(y|x)$. The δ - conditionally typical set $T_\delta^{Y^n|x^n}$ contains all the y^n sequences which would typically arise for a particular $x^n \in T_\delta^{X^n}$ in the sense that their conditional sample entropy is within δ - distance of their true entropy

$$T_\delta^{Y^n|x^n} \equiv \{y^n : \forall (a, b) \in \mathcal{X} \times \mathcal{Y}, |t_{y^n|x^n}(b|a) - p_{Y|X}(a, b)| \leq \delta\} \quad (\text{A.28})$$

where

$$t_{y^n|x^n}(b|a) = \frac{t_{x^n y^n}(a, b)}{t_{x^n}(a)} = \frac{N(a, b|x^n, y^n)}{N(a|x^n)} \quad (\text{A.29})$$

is the conditional type (empirical distribution).

Properties of Typical Set:

- Unit Probability: With increasing n , it is highly likely that a randomly sampled sequence will belong to the typical set. In the asymptotic limit $n \rightarrow \infty$, this probability is one.

$$\Pr\{X^n \in T_\delta^{X^n}\} \geq 1 - \epsilon \quad (\text{A.30})$$

$$\Pr\{X^n, Y^n \in T_\delta^{X^n Y^n}\} \geq 1 - \epsilon \quad (\text{A.31})$$

$$\Pr\{Y^n \in T_\delta^{Y^n|x^n}\} \geq 1 - \epsilon \quad (\text{A.32})$$

for all $\epsilon \in (0, 1)$, $\delta > 0$, and sufficiently large. n .

- Exponential Smaller Cardinality: The size of the typical set is exponentially smaller than the set of all the possible random sequences.

$$(1 - \epsilon)2^{n(H(X) - \delta)} \leq |T_\delta^{X^n}| \leq 2^{n(H(X) + \delta)} \quad (\text{A.33})$$

$$(1 - \epsilon)2^{n(H(X, Y) - \delta)} \leq |T_\delta^{X^n Y^n}| \leq 2^{n(H(X, Y) + \delta)} \quad (\text{A.34})$$

$$(1 - \epsilon)2^{n(H(Y|X) - \delta)} \leq |T_\delta^{Y^n|x^n}| \leq 2^{n(H(Y|X) + \delta)} \quad (\text{A.35})$$

for all $\epsilon \in (0, 1)$, $\delta > 0$, and sufficiently large.

- Equipartition: The probability of picking a particular typical sequence from the typical set is uniform.

$$2^{-n(H(X) + \delta)} \leq p_{X^n}(x^n) \leq 2^{-n(H(X) - \delta)} \quad (\text{A.36})$$

$$2^{-n(H(X, Y) + \delta)} \leq p_{X^n Y^n}(x^n y^n) \leq 2^{-n(H(X, Y) - \delta)} \quad (\text{A.37})$$

$$2^{-n(H(Y|X) + \delta)} \leq p_{Y^n|X^n}(y^n|x^n) \leq 2^{-n(H(Y|X) - \delta)} \quad (\text{A.38})$$

for all $\epsilon \in (0, 1)$, $\delta > 0$, and sufficiently large.

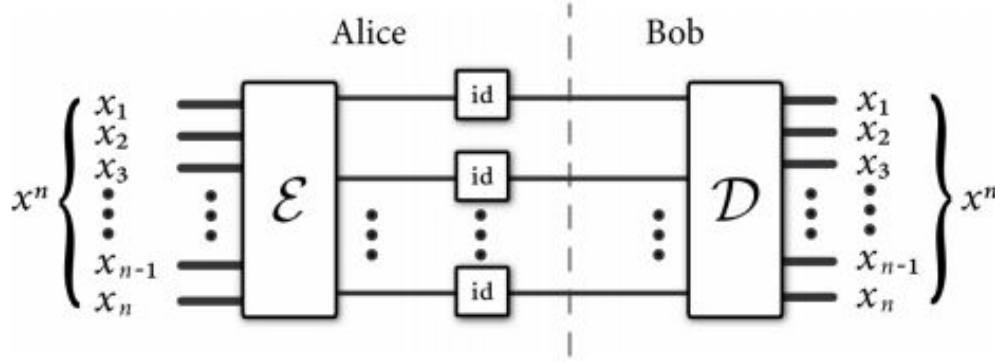


FIGURE A.2: Schematic showing Shannon's idea for a source code. Alice has an encoding channel such that $\mathcal{E} : x^n \mapsto c^{\tilde{n}}$. $\tilde{n} < n$, as indicated by fewer lines following the encoder. The transmission channel is noiseless indicated by "id" for identity. Bob's decoding channel undoes Alice's encoding $\mathcal{D} : c^{\tilde{n}} \mapsto x^n$. A good code will allow Bob to decode Alice's message with arbitrarily low probability of error [101]

A.1.3 Source Coding

Figure (A.2) shows the setting for Shannon's idea for source coding. Let us consider the following scenario of a bipartite communication. Alice sends a message, selected randomly from a set of messages, to Bob over a noiseless channel. A message is a string of letters taken from an alphabet. The goal is to minimize the number of channel uses (bits per channel use).

Alice's random choice of a message can be modeled as an independent and identically distributed (i.i.d.) information source. An i.i.d information source can be represented with a random variable X with realizations x , and corresponding probability distribution $p_X(x)$. The realizations x 's take values which are letters $\{a_1, a_2, \dots, a_{|\mathcal{X}|}\}$ in an alphabet $\mathcal{X}$ with size $|\mathcal{X}|$. An n -letter sequence output by the source is represented as an n -tuple

$$x^n \rightarrow x_1 x_2 x_3 \dots x_n \quad (\text{A.39})$$

The information content of a particular symbol x_i is defined as

$$i(x_i) \equiv \log_2 \left(\frac{1}{p_X(x_i)} \right) = -\log_2(p_X(x_i)) \quad (\text{A.40})$$

In this way, the lower the probability of occurrence for x_i in the sequence, the richer it is in information content. The information content of a particular sequence x^n

$$\begin{aligned} i(x^n) &\equiv -\log(p_{X_1 X_2 \dots X_n}(x_1 x_2 \dots x_n)) \\ &= -\log(p_{X_1}(x_1) p_{X_2}(x_2) \dots p_{X_n}(x_n)) \\ &= -\log(p_{X_1}(x_1)) - \log(p_{X_2}(x_2)) \dots - \log(p_{X_n}(x_n)) \\ &= -\log(p_X(x_1)) - \log(p_X(x_1)) \dots - \log(p_X(x_n)) \\ &= i(x_1) + i(x_2) + \dots + i(x_n) \end{aligned} \quad (\text{A.41})$$

The second equality is possible if the symbols occur independently from each other (the source is *memory less*). The forth equality is due to the symbols are all identically distributed. Therefore, the information content of a sequence output by an i.i.d. information source is additive. The expected information content of a source denoted by a random variable X as above is called the entropy of the source

$$H(X) \equiv - \sum_{x \in \mathcal{X}} p_X(x) \log(p_X(x)) \quad (\text{A.42})$$

Now if X^n denotes the random variable associated with the set of sequences x^n 's, then under the i.i.d. assumption, the probability of any emitted given sequence x^n factors as

$$\begin{aligned} p_{X^n}(x^n) &= p_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) \\ &= p_{X_1}(x_1) p_{X_2}(x_2) \dots p_{X_n}(x_n) \\ &= p_X(x_1) p_X(x_2) \dots p_X(x_n) \\ &= \prod_{i=1}^n p_X(x_i) \end{aligned} \quad (\text{A.43})$$

which is itself a random variable. Now if any random sequence x^n is written in terms of the letters from the alphabet $\mathcal{X}$, then the i.i.d. assumption allows us to permute the sequence x^n as

$$x^n \rightarrow \underbrace{a_1 \dots a_1}_{N(a_1|x^n)} \underbrace{a_2 \dots a_2}_{N(a_2|x^n)} \dots \underbrace{a_{|\mathcal{X}|} \dots a_{|\mathcal{X}|}}_{N(a_{|\mathcal{X}|}|x^n)} \quad (\text{A.44})$$

we can write the probability distribution in equation (A.43) as

$$p_{X^n}(x^n) = \prod_{i=1}^n p_X(x_i) = \prod_{j=1}^{|\mathcal{X}|} p_X(a_j)^{N(a_j|x^n)} \quad (\text{A.45})$$

Sample average of the information content of the random sequence X^n is

$$\begin{aligned} -\frac{1}{n} \log(p_{X^n}(X^n)) &= -\frac{1}{n} \log \left(\prod_{j=1}^{|\mathcal{X}|} p_X(a_j)^{N(a_j|X^n)} \right) \\ &= -\frac{1}{n} \sum_{j=1}^{|\mathcal{X}|} \log(p_X(a_j)^{N(a_j|X^n)}) \\ &= \sum_{j=1}^{|\mathcal{X}|} \frac{N(a_j|X^n)}{n} \log(p_X(a_j)) \end{aligned} \quad (\text{A.46})$$

which is a random quantity by itself.

Shannon's idea of data compression is to let the source emit a large number of realizations and then code the emitted data as a large block of size n , where $n \gg |\mathcal{X}|$, the size of the alphabet. In this technique of coding, called the *block*

coding, the law of large numbers comes into play. The law of large numbers states as the sample size gets bigger, the empirical distribution of the letters in a given sample $\frac{N(a_j|x^n)}{n}$ approaches their true distribution, $p_X(a_j)$. Conversely it is highly unlikely that a random sequence does not satisfy this property. Concisely, it is *typical* for a sequence to satisfy the following condition for all $\delta > 0$ as n becomes large

$$\lim_{n \rightarrow \infty} \Pr \left\{ \left| -\frac{1}{n} \log(p_{X^n}(X^n)) - \underbrace{\sum_{i=1}^{|X|} p_X(a_i) \log \left(\frac{1}{p_X(a_i)} \right)}_{H(X)} \right| \leq \delta \right\} = 1 \quad (\text{A.47})$$

Therefore in a typical sequence x^n of size n the letters would appear reflecting their true distribution within $\pm\delta$ margin of error. The size of a typical sequence can be calculated considering that in an n -letter sequence, different letters $a_1, a_2, \dots, a_{|X|}$ should appear approximately $np(a_1), np(a_2), \dots, np(a_i)$ times respectively. The number of such sequences is therefore

$$\binom{n}{\prod_{a_i} np(a_i)} \quad (\text{A.48})$$

Using Stirling's formula $\log n! = n \log n - n + O(\log n)$ the binomial term can be simplified as

$$\binom{n}{\prod_{a_i} np(a_i)} \approx 2^{nH(X)} \quad (\text{A.49})$$

Shannon's strategy of data compression is only to establish a one-to-one encoding function that maps from the set of all typical sequences (size $2^{nH(X)}$) to the set of all binary strings of length $nH(X)$ (also of size $2^{nH(X)}$). If the source emits an atypical sequence, an error is declared. This coding scheme is reliable in the asymptotic limit because the probability of an error event vanishes as n becomes large, due to the unit probability property of the asymptotic equipartition theorem. The rate (efficiency) of this coding scheme is

$$\text{compression rate} = \frac{\text{number of noiseless channel bits}}{\text{number of source symbols}} = \frac{nH(X)}{n} \quad (\text{A.50})$$

The above formula gives an operational interpretation of Shannon entropy, $H(X)$: The entropy of an information source specified by a discrete random variable X is the smallest achievable rate for compression:

$$\inf\{R : R \text{ is achievable for } X\} = H(X) \quad (\text{A.51})$$

Finally, the statement for source coding theorem, which Shannon proved in two parts, can be expressed as follows

Direct Theorem: If the rate of compression $R > H(X)$ the entropy of the source,, then there exists a coding scheme that can achieve lossless data compression

in the sense that it is possible to make the probability of error for incorrectly decoding arbitrarily small.

Converse Theorem: If there exists a coding scheme that can achieve lossless data compression with arbitrarily small probability of decoding error, then the rate of compression is greater than the entropy of the source.

A.1.4 Channel Coding

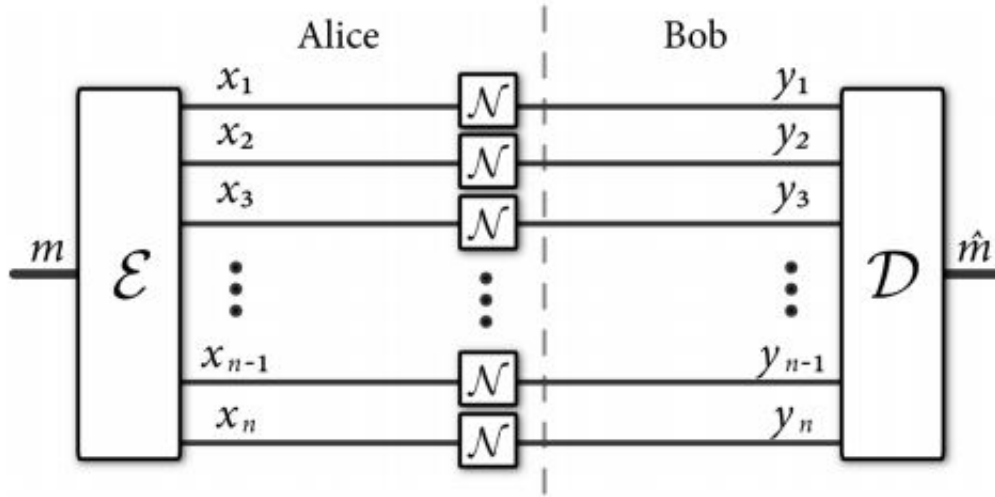


FIGURE A.3: Schematic showing Shannon's idea for a channel code. Alice chooses a message $m \in [M]$. The encoder channel outputs according to $\mathcal{E} : m \mapsto x^n$. The noisy channel induces $\mathcal{N} : x^n \mapsto y^n$. Bob's decoder finds $\mathcal{D} : y^n \mapsto \hat{m}$. A reliable channel code will allow Alice and Bob to communicate at a finite rate with vanishing probability of error $\Pr[m \neq \hat{m}] \rightarrow 0$. [101]

Alice wants to send a message, m from the set $\mathcal{M} \equiv \{1, \dots, M\}$ over a noisy classical channel. Without loss of generality it can be assumed that each message is equally likely, therefore it would require $\log_2 M$ bits to represent any message from the above set. To transmit a particular message with asymptotic perfect reliability, Alice and Bob agree on establishing a codebook $\mathcal{C} \equiv \{x^n(m)\}_{m \in \mathcal{M}}$ for the following encoding operation $\text{Enc} : m \mapsto x^n(m)$, where $x^n(m)$ denotes a particular codeword corresponding to message m generated i.i.d. according to the probability distribution $p_{X^n}(x^n)$:

$$\begin{aligned} \mathcal{C}(1) &= x_1(1), x_2(1), \dots, x_n(1) = x^n(1) \\ \mathcal{C}(2) &= x_1(2), x_2(2), \dots, x_n(2) = x^n(2) \\ &\vdots \\ \mathcal{C}(2^k) &= x_1(2^k), x_2(2^k), \dots, x_n(2^k) = x^n(2^k) \end{aligned} \tag{A.52}$$

Since Alice assigns the codewords randomly to the messages, the probability distribution for choosing a particular codeword

$$P(\text{particular code}) = \prod_{m=1}^{2^k} \prod_{i=1}^n P_X(x_i(m)) \tag{A.53}$$

Alice inputs codeword $x^n(m)$ to the noisy channel to send the message m to Bob. The performance measure of the coding scheme is rate, R which is bits per channel use

$$R = \frac{\text{number of message bits}}{\text{number of channel uses}} = \frac{\log_2 M}{n} = \frac{k}{n} \quad (\text{A.54})$$

From which one can see the number of messages (and hence the number of codewords) can be written $2^k = 2^{nR}$. The classical channel can be generally modeled as a conditional probability distribution $\mathcal{N} \equiv p_{Y|X}(y|x)$ which produces a random output sequence $y^n \sim \text{i.i.d. } p_{Y^n}(y^n)$ conditioned input random sequence $x^n \sim \text{i.i.d. } p_{X^n}(x^n)$. Therefore Bob receives the corrupted codeword y^n

$$P_{Y^n|X^n}(y^n|x^n(m)) = \prod_{i=1}^n p_{Y|X}(y_i|x_i(m)) \quad (\text{A.55})$$

Bob has complete knowledge of the codebook and the channel; based on which he performs the following decoding operation $\text{Dec} : y^n(m) \mapsto \hat{m}$. An error occurs when $\hat{m}(y^n) \neq m$. If Bob performs "jointly typical decoding"¹, there are two events which contribute to the probability of error:

- The channel output y^n is not in the typical set $T_\delta^{Y^n}$ corresponding to the distribution $p_Y(y) \equiv \sum_x p_{Y|X}(y|x)p_X(x)$.
- The channel output y^n is in $T_\delta^{Y^n}$ but it is in the conditional typical set for some other message:

$$\{y^n \in T_\delta^{Y^n}\} \text{ and } \{\exists m' \neq m : y^n \in T_\delta^{Y^n|x^n(m')}\} \quad (\text{A.56})$$

The justification for randomly selecting a code to encode a message before transmission becomes evident as it simplifies the error analysis. The average probability of error for a particular code, $\mathcal{C}_0$

$$\bar{p}_e(\mathcal{C}_0) \equiv \frac{1}{|\mathcal{M}|} \sum_{m=1}^{|\mathcal{M}|} p_e(m, \mathcal{C}_0) \quad (\text{A.57})$$

while the maximal probability of error is

$$p_e^*(\mathcal{C}_0) \equiv \max_{m \in \mathcal{M}} p_e(m, \mathcal{C}_0) \quad (\text{A.58})$$

The ultimate aim of channel coding is to bound the maximal probability of error $p_e^* \leq \varepsilon$ where ε is an arbitrary small number. In fact, a channel code is characterized by a triplet (n, R, ε) , and these numbers are interrelated. The expectation

¹please refer to appendix

of the average probability of error is

$$\begin{aligned}
& \mathbb{E}_{\mathcal{C}}\{\bar{p}_e(\mathcal{C})\} \text{ here the expectation is with respect to the random code, } \mathcal{C} \\
&= \mathbb{E}_{\mathcal{C}}\left\{\frac{1}{|\mathcal{M}|} \sum_{m=1}^{|\mathcal{M}|} p_e(m, \mathcal{C})\right\} \\
&= \frac{1}{|\mathcal{M}|} \sum_{m=1}^{|\mathcal{M}|} \underbrace{\mathbb{E}_{\mathcal{C}}\{p_e(m, \mathcal{C})\}}_{\text{does not depend on any particular } m} \quad \text{since expectation is linear} \\
&= \mathbb{E}_{\mathcal{C}}\{p_e(1, \mathcal{C})\} \text{ choosing a particular message } m = 1 \text{ suffices} \\
&= \sum_{\mathcal{C}} P[\mathcal{C}] p_e(1, \mathcal{C}) \tag{A.59}
\end{aligned}$$

With jointly typical decoding, it is possible to derive an upper bound on the expectation of the average probability of error. With jointly typical decoding, it is possible to derive an upper bound on the expectation of the average probability of error.

1. The first type of error is because $(x^n(1), y^n)$ are not jointly typical. From the first part of jointly typical theorem, for any $\varepsilon > 0$ there exists an n_ε such that

$$\forall n > n_\varepsilon : \text{Prob}\{(x^n(1), y) \notin T_\delta^{X^n Y^n}\} < \varepsilon \tag{A.60}$$

Therefor by letting $n \rightarrow \infty$, one can have $\text{Prob}\{a\} \rightarrow 0$. For large enough n , this type of error is bounded as $\text{Prob}\{a\} \leq \varepsilon$.

2. The second type of error is because $(x^n(m'), y^n)$ is jointly typical for $m' \neq 1$. By the third item of the joint AEP, the probability that this type of error occurs is at most

$$\begin{aligned}
\sum_{m' \neq 1}^{2^{n \cdot R}} 2^{(-nI(X;Y) - 3\delta)} &= (2^{n \cdot R} - 1) 2^{(-nI(X;Y) - 3\delta)} \\
&\leq 2^{n \cdot R} 2^{(-nI(X;Y) - 3\delta)} \\
&= 2^{n(I(X;Y) - R - 3\delta)} \tag{A.61}
\end{aligned}$$

Here $I(X;Y)$ is the *mutual information* between X and Y . As long as $R < I(X;Y) - 3\delta$, there exists an n_ε for each $\varepsilon > 0$ such that

$$\forall n > n_\varepsilon, 2^{-n(I(X;Y) - R - 3\delta)} < \varepsilon \tag{A.62}$$

Therefore for each $\varepsilon > 0$ there exists an $n_\varepsilon = \max\{n_\varepsilon^{(a)}, n_\varepsilon^{(b)}\}$ such that for any $n > n_\varepsilon$ one has

$$\bar{p}_e \leq 2\varepsilon \tag{A.63}$$

Since the above upper bound is for the average probability of error taken with respect to the expectation over a random code $\mathcal{C}$, there must be some deterministic code $\mathcal{C}_0$ whose average probability of error meets this same bound

$$\bar{p}_e(\mathcal{C}_0) \leq \varepsilon \tag{A.64}$$

Now for this specific code $\mathcal{C}_0$, from Markov's inequality, at least half of the messages $m \in \mathcal{M}$, will have maximal probability of error

$$p_e^*(\mathcal{C}_0) \leq 4\epsilon \quad (\text{A.65})$$

otherwise the average probability of error would have exceeded the upper bound derived above leading to a contradiction. To construct a better code, one can get rid of half of the codewords $\frac{2^k}{2} = 2^{n(R-\frac{1}{n})}$. Our new channel code has a rate $R' = R - \frac{1}{n}$ which approaches R , as $n \rightarrow \infty$, and the maximum probability of error $p_e^* \leq 4\epsilon$. Therefor Alice can send information reliably to Bob over a noisy channel with arbitrary accuracy by making n uses of the channel ($n \rightarrow \infty$) as long as the rate of her channel code, R is bounded by the entropic quantity $I(X;Y)$. By definition

$$I(X;Y) \equiv H(Y) - H(Y|X)$$

The operational interpretation of mutual information can be understood as *the largest rate (bits sent on average per use of a channel) at which communication can be done reliably over a noisy channel*. The mutual information also provides insight into how Alice and Bob should structure their code. To minimize decoding error, it is necessary that, for a unique message m , the channel output sequence belongs to a unique $y^n(m) \in T_\delta^{Y^n|x^n(m)}$ conditionally typical set. Therefore, $|\mathcal{M}|$ sets of conditionally typical output sets, each of size $2^{nH(Y|X)}$ should fit inside the set of output typical sequences $T_\delta^{Y^n}$ with no, or negligible overlap.

$$2^{nR} \approx \frac{2^{nH(Y)}}{2^{nH(Y|X)}} = 2^{n(H(Y)-H(Y|X))} = 2^{nI(X;Y)} \quad (\text{A.66})$$

For a fixed conditional channel distribution $p_{Y|X}(y|x)$, the mutual information is a concave function of the channel input distribution $p_X(x)$ meaning there is a distribution $p_X^*(x)$ which maximizes the mutual information. This quantity is called the *capacity* of the channel

$$C(\mathcal{N}) \equiv \max_{p_X(x)} I(X;Y) \quad (\text{A.67})$$

Finally, the statement for channel coding theorem, which Shannon proved in two parts, can be expressed as follows **Direct Theorem:** If the rate of communication, $R > C$, the channel capacity, then it is possible for Alice to communicate reliably to Bob, in the sense that a sequence of codes exists whose maximal probability of error vanishes as the number of channel uses become increasingly large.

Converse Theorem: If a reliable sequence of codes exists, then the rate of this sequence of codes is less than the channel capacity.

A.2 Notions from Quantum Mechanics

State: The most general description of a quantum mechanical state, which also takes into account statistical distribution of states, is the density operator $\rho \in$

$\mathcal{D}(\mathcal{H})$ on the Hilbert space such that ρ is (i) Hermitian $\rho^\dagger = \rho$, (ii) positive $\rho \geq 0$, (iii) unit trace $\text{Tr}(\rho) = 1$.

A pure state corresponds a rank-one density operator such that $\rho^2 = \rho$, with notation $|\psi\rangle\langle\psi|$, for some unit norm vector $|\psi\rangle \in \mathcal{H}$. A mixed state corresponds to a convex sum of all the possible states

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \quad (\text{A.68})$$

with not necessarily $\langle\psi_i|\psi_j\rangle \neq 0$. The above representation is non-unique. Since ρ is Hermitian, according to the spectral decomposition theorem, it can be diagonalized in an orthonormal basis

$$\rho = \sum_x \lambda_x |x\rangle\langle x|, \quad (\text{A.69})$$

where $\{\lambda_x\}$ is the set of eigenvalues. With this representation ρ can be validly interpreted as a classical probability distribution.

Measurement: The most general measurement one can perform on a quantum state to extract classical information (i.e. the results are mutually distinguishable) is called the POVM (positive operator-valued measure), which is a collection $\Lambda = \{\Lambda_x\}$, such that

$$\sum_x \Lambda_x = 1, \quad \forall x : \Lambda_x \geq 0. \quad (\text{A.70})$$

The probability of getting an outcome x is when performing Λ on ρ is

$$p_X(x) = \text{Tr}\{\Lambda_x \rho\}, \quad (\text{A.71})$$

and if the outcome is x , the post-measurement state is given by

$$\rho \rightarrow \frac{\sqrt{\Lambda_x} \rho \sqrt{\Lambda_x}}{\text{Tr}\{\Lambda_x \rho\}}. \quad (\text{A.72})$$

Trace Distance: Given two quantum states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, the closeness between them can be quantified with the trace distance(/norm)

$$\|\rho - \sigma\|_1 = \sum_i |\lambda_i|, \quad (\text{A.73})$$

where $\{\lambda_i\}$'s are the eigenvalues of $\rho - \sigma$. In general, $\|A\|_1 = \text{Tr}\{|A|\} = \text{Tr}\{\sqrt{A^\dagger A}\}$. If $\rho \perp \sigma$, then the trace distance between them is maximized, and they are perfectly distinguishable. On the other hand the trace distance between ρ and σ is minimum, when they are identical. With trace distance one can prove the following lemma

$$\text{Tr}\{\Lambda \rho\} \leq \|\rho - \sigma\|_1 + \text{Tr}\{\Lambda \sigma\}, \quad (\text{A.74})$$

which states the following: if the measurement outcome corresponding to Λ is not known for a state ρ , but known for a state σ , then one can upper bound the

probability $\text{Tr}\{\Lambda\rho\}$ by knowing the trace distance between ρ and σ . The above lemma is also useful in quantum hypothesis testing. Suppose two states ρ and σ are prepared with equal probability. $\{\Lambda_\rho, \Lambda_\sigma\}$ denotes the POVM elements for perfectly distinguishing ρ and σ respectively ($\Lambda_\rho, \Lambda_\sigma = \mathbb{1}$). The probability of error to correctly identify them is given by

$$p_{\text{err}} = \frac{1}{2}(\text{Tr}\{\Lambda_\rho\sigma\} + \text{Tr}\{\Lambda_\sigma\rho\}) = \frac{1}{2}(1 - \frac{1}{2}\|\rho - \sigma\|_1). \quad (\text{A.75})$$

Channel: A channel is a completely-positive, trace-preserving (CPTP) map which corresponds to the most general evolution a quantum state can undergo. A channel is characterized with the Kraus decomposition theorem

$$\rho \mapsto \mathcal{N}(\rho) = \rho' = \sum_{k=1}^K A_k \rho A_k^\dagger, \quad (\text{A.76})$$

for any $\rho \in \mathcal{D}(\mathcal{H})$ and $\rho' \in \mathcal{D}(\mathcal{H}')$. A_k 's are the linear Karus operators $A \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ with completeness relation $\sum_{k=1}^K A_k^\dagger A_k = \mathbb{1}$. The cardinality $K \leq \dim(\mathcal{H})\dim(\mathcal{H}')$. The above representation is used to characterize Gaussian CP maps in the next chapter.

The concept of channel is wide-ranging. The act of quantum state preparation can be thought of as a classical to quantum channel through which a classical variable gets mapped to a quantum state.

$$x \mapsto \rho^x. \quad (\text{A.77})$$

Similarly, a POVM or an act of measurement can be thought of as a quantum to classical channel

$$\rho \mapsto \sum_x \text{Tr}\{\Lambda_x \rho\} \underbrace{|x\rangle\langle x|}_{=\text{orthonormal}}. \quad (\text{A.78})$$

A.3 Quantum Information-theoretic Measures

The information-theoretic measures from classical information theory have analogues in quantum information theory. The entropy of a quantum system A in a state ρ_A is called the Von Neumann entropy, of the quantum system, and is defined as

$$H(\rho_A) \equiv H(A)_\rho \equiv -\text{Tr}\{\rho_A \log \rho_A\}. \quad (\text{A.79})$$

If the density operator ρ has the spectral decomposition as in equation (A.69), then quantum entropy takes the same form as Shannon entropy $H(X)$

$$H(\rho_A) = -p_X(x) \log p_X(x). \quad (\text{A.80})$$

The quantum entropy has the following properties: (i) Non-negative, $H(\rho) \geq 0$; (ii) $\min\{H(\rho)\} = 0$, it occurs when ρ is a pure state; (iii) $\max\{H(\rho)\} = \log d$, where d is the dimension of the quantum system which is in a maximally mixed

state; (iv) concave under mixing operation: $H(\rho) \geq p_X(x)H(\rho_x)$, with $\rho \equiv \sum_x p_X(x)\rho_x$; (v) invariant under isometric operation $H(\rho) = H(U\rho U^\dagger)$.

The joint quantum entropy $H(AB)_\rho$ of the density operator ρ_{AB} denoting the composite state of a bipartite system A and B is $H(AB)_\rho \equiv -\text{Tr}\{\rho_{AB}\log\rho_{AB}\}$. One consequence of quantum entropy being solely dependent on the eigenvalues of a given spectral decomposition of a density matrix is that the marginal entropies $H(A)_\rho$ and $H(B)_\rho$ of a pure bipartite state $|\phi\rangle_{AB}$ are equal: $H(A)_\rho = H(B)_\rho$, while the joint entropy vanishes: $H(AB)_\rho = 0$. In fact, the result applies to bipartite cut of any number of systems. In this respect quantum entropy differs from classical entropy markedly in which case joint entropy can never be less than marginal entropies. Finally, for a special case, the joint entropy $H(XB)_\rho$ of a bipartite state which is a classical-quantum state of the form $\rho_{XB} \equiv \sum_x p_X(x) |x\rangle\langle x| \otimes \rho_B^x$ is

$$H(XB)_\rho = H(X) + \sum_x p_X(x)H(\rho_B^x), \quad (\text{A.81})$$

where $H(X)$ is the entropy of a random variable X with distribution $p_X(x)$.

The conditional quantum entropy of ρ_{AB} is $H(A|B)_\rho \equiv H(AB)_\rho - H(B)_\rho$. As in the case for classical entropy, conditioning does not increase entropy $H(A)_\rho \geq H(A|B)_\rho$. But interestingly, conditional quantum entropy can be negative, for example, if ρ_{AB} is a pure state then $H(AB)_\rho = 0$, and $H(A|B)_\rho = -H(B)_\rho$. The conditional entropy $H(B|X)_\rho$ for a classical quantum state

$$H(B|X)_\rho = H(XB)_\rho - H(X)_\rho = \sum_x p_X(x)H(\rho_B^x). \quad (\text{A.82})$$

The quantum mutual information $I(A;B)_\rho$ of a bipartite state ρ_{AB} is

$$I(A;B)_\rho \equiv H(A)_\rho + H(B)_\rho - H(AB)_\rho. \quad (\text{A.83})$$

A.4 Quantum Typicality

The notions from classical typicality can be extended to the quantum regime. Let us consider a quantum state preparation system. Depending on different realizations of random variable Y according to the probability distribution it prepares states $|\psi_y\rangle$. Therefore, the expected state prepared by the system can be written as

$$\rho_{\text{exp}} = \sum_y p_Y(y) |\psi_y\rangle\langle\psi_y|. \quad (\text{A.84})$$

The above decomposition is non-unique. Of particular importance is the spectral decomposition

$$\rho_{\text{exp}} = \sum_{x \in \mathcal{X}} p_X(x) |x\rangle\langle x|. \quad (\text{A.85})$$

$\{|x\rangle\}$'s are the eigenvectors of ρ which form a complete orthonormal basis for Hilbert space $\mathcal{H}$ where ρ lives. $\{p_X(x)\}$'s are the corresponding eigenvalues.

This particular decomposition allows us to treat the quantum source as a classical source based on Boolean logic. The x 's can be considered taking values from an alphabet $\mathcal{X}$ of distinct letters.

Now if X is sampled n times to realize a sequence $x^n \equiv x_1 x_2 \dots x_n$, the state preparation system would prepare the following tensor-product state correspondingly

$$\rho^n = \rho_1 \otimes \dots \otimes \rho_n = (\rho_{\text{exp}})^{\otimes n}, \quad (\text{A.86})$$

since in each case the expected density operator remains the same. Now

$$(\rho_{\text{exp}})^{\otimes n} = \left(\sum_{x \in \mathcal{X}} p_X(x) |x\rangle \langle x| \right)^{\otimes n} = \sum_{x^n \in \mathcal{X}^n} p_{X^n}(x^n) |x^n\rangle \langle x^n|, \quad (\text{A.87})$$

where following shorthand notations have been used

$$p_{X^n}(x^n) \equiv p_X(x_1) \dots p_X(x_n) \quad |x^n\rangle \equiv |x_1\rangle \dots |x_n\rangle \quad (\text{A.88})$$

Typical Subspace: The quantum analogue of classical typical set is the typical subspace. The δ -typical subspace $T_\delta^{\rho^n}$ is a subspace of the full Hilbert space $\mathcal{H}^{\otimes n}$ associated with the density operator ρ^n is spanned by the states $|x^n\rangle$ whose corresponding classical sequences are x^n are δ -typical.

$$T_\delta^{\rho^n} \equiv \text{span}\{|x^n\rangle : x^n \in T_\delta^{X^n}\}. \quad (\text{A.89})$$

Typical Subspace Projector: The typical subspace projector projects a state onto the typical subspace. For the case discussed above, the typical subspace projector $\Pi_\delta^{\rho^n}$ for the Hilbert space $\mathcal{H}^{\otimes n}$ with respect to the state ρ^n can be expressed as

$$\Pi_\delta^{\rho^n} \equiv \sum_{x^n \in T_\delta^{X^n}} |x^n\rangle \langle x^n|. \quad (\text{A.90})$$

By acting on a state ρ^n , the typical subspace projector $\Pi_\delta^{\rho^n}$ slices out any component that does not belong to the typical subspace $T_\delta^{\rho^n}$

$$\tilde{\rho}^n \equiv \Pi_\delta^{\rho^n} \rho^n \Pi_\delta^{\rho^n} = \sum_{x^n \in T_\delta^{X^n}} p_{X^n}(x^n) |x^n\rangle \langle x^n|. \quad (\text{A.91})$$

The atypical subspace projector $\mathbb{I} - \Pi_\delta^{\rho^n}$ is the complement of the typical subspace projector, and together they form a POVM.

Joint Quantum Typicality: In general, a bipartite state ρ_{AB} prepared by two systems A and B can have the following decomposition as per the spectral theorem

$$\rho_{AB} = \sum_{z \in \mathcal{Z}} p_Z(z) |z\rangle \langle z|_{AB}. \quad (\text{A.92})$$

The states $\{|z\rangle\}$'s form an orthonormal basis for the joint quantum system AB . The n -th extension of the joint system can be expressed as follows

$$\rho_{A^n B^n} \equiv (\rho_{AB})^{\otimes n} = \sum_{z^n \in \mathcal{Z}^n} p_Z^n(z^n) |z^n\rangle \langle z^n|_{A^n B^n}. \quad (\text{A.93})$$

The δ -typical subspace $T_\delta^{A^n B^n}$ of ρ_{AB} is the space spanned by the states $|z^n\rangle_{A^n B^n}$ whose corresponding classical sequence z^n is in the typical set

$$T_\delta^{A^n B^n} \equiv \text{span}\{|z^n\rangle_{A^n B^n} : z^n \in T_\delta^{Z^n}\}. \quad (\text{A.94})$$

The projector onto that subspace is called the jointly typical projector

$$\Pi_{A^n B^n}^\delta \equiv \sum_{z^n \in T_\delta^{Z^n}} |z^n\rangle \langle z^n|_{A^n B^n}. \quad (\text{A.95})$$

Let us consider the case when only the classical correlations hold. Two random variables X and Y , both of which have been sampled n times to generate i.i.d. sequences x^n and y^n according to $p_X(x)$ and $p_Y(y)$ respectively. There are two quantum state preparation systems A and B which can prepare quantum states with mappings $x \mapsto \rho^x$ and $y \mapsto \rho^y$. The joint state can be written as

$$\rho_{AB} = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p_{X,Y}(x,y) \rho^x \otimes \rho^y = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p_{X,Y}(x,y) |x\rangle \langle x|_A \otimes |y\rangle \langle y|_B. \quad (\text{A.96})$$

Considering the n -th extension i.e. $x^n \mapsto \rho^{x^n}$ and $y^n \mapsto \rho^{y^n}$.

$$\rho_{A^n B^n} = \sum_{x^n \in \mathcal{X}^n} \sum_{y^n \in \mathcal{Y}^n} p_{X^n, Y^n}(x^n, y^n) |x^n\rangle \langle x^n|_A \otimes |y^n\rangle \langle y^n|_B. \quad (\text{A.97})$$

The jointly typical subspace in this case

$$T_\delta^{A^n B^n} \equiv \text{span}\{|x^n\rangle_{A^n} |y^n\rangle_{B^n} : x^n y^n \in T_\delta^{X^n Y^n}\} \quad (\text{A.98})$$

and the jointly typical projector

$$\Pi_{A^n B^n}^\delta \equiv \sum_{x^n, y^n \in T_\delta^{X^n Y^n}} |x^n\rangle \langle x^n|_{A^n} \otimes |y^n\rangle \langle y^n|_{B^n}. \quad (\text{A.99})$$

Conditional Quantum Typicality Let us consider the system XB , which refers to the following classical-quantum ensemble

$$\{p_X(x), |x\rangle \langle x|_X \otimes \rho_B^x\}_{x \in \mathcal{X}}. \quad (\text{A.100})$$

X is a random variable with probability distribution $p_X(x)$. With each realization x of X we associate a classical state $|x\rangle \langle x|$ and a quantum state ρ_B^x . The expected density operator of the above classical-quantum state is

$$\rho_{XB} = \sum_{x \in \mathcal{X}} p_X(x) |x\rangle \langle x|. \quad (\text{A.101})$$

The conditional quantum entropy $H(B|X)_\rho$ of the classical quantum state is

$$H(B|X)_\rho = \sum_{x \in \mathcal{X}} p_X(x) H(\rho_B^x). \quad (\text{A.102})$$

The spectral form of ρ_B^x can be written as follows

$$\sum_{y \in \mathcal{Y}} p_{Y|X}(y|x) |y_x\rangle \langle y_x|_B. \quad (\text{A.103})$$

With this expression one can write

$$H(B|X) = \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p_X(x) p_{Y|X}(y|x) \log \frac{1}{p_{Y|X}(y|x)}. \quad (\text{A.104})$$

Let us consider n – such implementations from the ensemble

$$\rho_{X^n B^n} \equiv (\rho_{XB})^{\otimes n} = \sum_{x^n \in \mathcal{X}^n} p_{X^n}(x^n) |x^n\rangle \langle x^n|_{X^n} \otimes \rho_{B^n}^{x^n}. \quad (\text{A.105})$$

The spectral form of $\rho_{B^n}^{x^n}$ is

$$\rho_{B^n}^{x^n} = \sum_{y^n \in \mathcal{Y}^n} p_{Y^n|X^n}(y^n|x^n) |y_{x^n}^n\rangle \langle y_{x^n}^n|_{B^n}. \quad (\text{A.106})$$

Conditionally Typical Subspace: Corresponding to a particular sequence x^n and an ensemble $\{p_X(x), \rho_B^x\}$, the conditionally typical subspace $T_\delta^{B^n|x^n}$ is spanned by the states $|y_{x^n}^n\rangle_{B^n}$ which classically form the δ – typical conditional set $T_\delta^{Y^n|x^n}$

$$T_\delta^{B^n|x^n} \equiv \text{span}\{|y_{x^n}^n\rangle_{B^n} : y_{x^n}^n \in T_\delta^{Y^n|x^n}\}, \quad (\text{A.107})$$

with the typical subspace projector

$$\Pi_{B^n|x^n}^\delta \equiv \sum_{y_{x^n}^n \in T_\delta^{Y^n|x^n}} |y_{x^n}^n\rangle \langle y_{x^n}^n|_{B^n}. \quad (\text{A.108})$$

Properties of Typical Subspace: The properties of quantum typical subspace follow similarly as the properties of typical set in the classical case [101]. In the following A stands for a standalone system, AB stands for a joint system, and $B|X$ refers to a conditional classical-quantum system.

- Unit Probability: With increasing n , it is highly likely that a typical subspace measurement on a random state will give a positive outcome. In the asymptotic limit $n \rightarrow \infty$, this probability is one.

$$\text{Tr}\{\Pi_{A^n}^\delta \rho_{A^n}\} \geq 1 - \epsilon, \quad (\text{A.109})$$

$$\text{Tr}\{\Pi_{A^n B^n}^\delta \rho_{A^n B^n}\} \geq 1 - \epsilon, \quad (\text{A.110})$$

$$\text{Tr}\{\Pi_{B^n|x^n}^\delta \rho_{B^n}^{x^n}\} \geq 1 - \epsilon, \quad (\text{A.111})$$

for all $\epsilon \in (0, 1)$, $\delta > 0$, and n sufficiently large. c is a positive quantity.

- Exponential Smaller Cardinality: The size of the typical subspace is exponentially smaller than the corresponding Hilbert space (except when the state under consideration is not maximally mixed)

$$(1 - \epsilon)2^{n[H(A)+c\delta]} \leq \text{Tr}\{\Pi_{A^n}^\delta \rho_{A^n}\} \leq 2^{n[H(A)+c\delta]}, \quad (\text{A.112})$$

$$(1 - \epsilon)2^{n[H(AB)+c\delta]} \leq \text{Tr}\{\Pi_{A^n B^n}^\delta \rho_{A^n B^n}\} \leq 2^{n[H(AB)+c\delta]}, \quad (\text{A.113})$$

$$(1 - \epsilon)2^{n(H(B|X)-c\delta)} \leq \text{Tr}\{\Pi_{B^n|x^n}^\delta \rho_{B^n}^{x^n}\} \leq 2^{n(H(B|X)+c\delta)}, \quad (\text{A.114})$$

for all $\epsilon \in (0, 1)$, $\delta > 0$, and sufficiently large.

- Equipartition: Upon the typical subspace measurement onto the typical subspace, the projected state looks approximately maximally mixed

$$2^{-n(H(A)+c\delta)} \Pi_{A^n}^\delta \leq \Pi_{A^n}^\delta \rho_{A^n} \Pi_{A^n}^\delta \leq 2^{-n(H(A)-c\delta)} \Pi_{A^n}^\delta, \quad (\text{A.115})$$

$$2^{-n(H(AB)+c\delta)} \Pi_{A^n B^n}^\delta \leq \Pi_{A^n B^n}^\delta \rho_{A^n B^n} \Pi_{A^n B^n}^\delta \leq 2^{-n(H(AB)-c\delta)} \Pi_{A^n B^n}^\delta, \quad (\text{A.116})$$

$$2^{-n(H(B|X)+\delta)} \leq \Pi_{B^n|x^n}^\delta \rho_{B^n|x^n} \Pi_{B^n|x^n}^\delta \leq 2^{-n(H(B|X)-\delta)}, \quad (\text{A.117})$$

for all $\epsilon \in (0, 1)$, $\delta > 0$, and sufficiently large.

A.5 The Classical Capacity Theorem

Alice wants to transmit a message $m \in \{1, 2, \dots, |\mathcal{M}|\}$ to Bob through a quantum channel $\mathcal{N}$. $|\mathcal{M}|$ denotes the size of her set, and M corresponds to her random choice of message. Alice can exploit an ensemble $\{p_X(x), \rho^x\}$ of states, with expected density operator $\rho_{\text{exp}} = \sum_x p_X(x) \rho^x$, to make a random code. The ensemble refers to the fact that as the random variable X takes value $X = x \in \mathcal{X}$ with probability distribution $p_X(x)$, Alice can prepare quantum states according to $x \mapsto \rho^x$. $\mathcal{X}$ is the alphabet from which x takes values. Now let us denote $|\mathcal{X}|^n$ as the size of the set for all n i.i.d. sequential realizations x^n according to $p_{X^n}(x^n) = p_X(x_1) \dots p_X(x_n)$. A subset of $|\mathcal{X}|^n$ will form the set of strongly typical sequences $T_\delta^{X^n}$. Alice picks $|\mathcal{M}|$ sequences belonging to this set as classical codewords $\{x^n(m)\}_{m \in \{1, 2, \dots, |\mathcal{M}|\}}$ completely randomly according to the "pruned" distribution

$$p'_{X^n}(x^n) = \begin{cases} \frac{p_{X^n}(x^n)}{\sum_{x^n \in T_\delta^{X^n}} p_{X^n}(x^n)} & : x^n \in T_\delta^{X^n} \\ 0 & : x^n \notin T_\delta^{X^n} \end{cases}, \quad (\text{A.118})$$

where X^n is the random variable for the new adjusted distribution. $p'_{X^n}(x^n)$ is approximately close to $p_{X^n}(x^n)$ because the probability mass of the typical set is nearly one. Now for any $x^n \in |\mathcal{X}|^n$, Alice can prepare a sequence of quantum states

$$\rho^{x^n} \equiv \rho^{x_1} \otimes \dots \otimes \rho^{x_n} = \rho_{\text{exp}}^{\otimes n} \quad (\text{A.119})$$

employing $\{p_X(x), \rho^x\}$. Therefore the classical codewords $\{x^n(m)\}_{m \in \{1, 2, \dots, |\mathcal{M}|\}}$ chosen by Alice lead to the quantum codewords of the following form

$$\rho^{x^n(m)} \equiv \rho^{x_1(m)} \otimes \dots \otimes \rho^{x_n(m)},$$

as par $\{p'_{X^n}(x^n), \rho^{x^n(m)}\}$. She transmits this state over n independent uses of the channel $\mathcal{N}$. Therefore, the rate of the communication scheme is

$$C \equiv \frac{1}{n} \log |\mathcal{M}| \quad (\text{A.120})$$

The state at Bob's receiving end is

$$\sigma^{x^n(m)} \equiv \sigma^{x_1(m)} \otimes \dots \otimes \sigma^{x_n(m)} \equiv \mathcal{N}(\rho^{x_1(m)}) \otimes \dots \otimes \mathcal{N}(\rho^{x_n(m)}) \quad (\text{A.121})$$

Bob can exploit the *sequential decoding* technique to extract Alice's message from the transmitted codeword. In this technique Bob first asks "Does the received state belong to the typical subspace? If yes, then he sequentially asks "Is it the first codeword?". If no, then "Is it the second codeword?". And so on until he gets the a hit. Operationally, he can perform a typical subspace measurement corresponding to $\{\Pi, \mathbb{I} - \Pi\}$ first. Then for sequential queries he can perform measurements corresponding to the conditionally typical projectors $\{\Pi_i, \mathbb{I} - \Pi_i\}$. This technique works well on average because the transmitted codewords lie on the typical subspace on average

$$\mathbb{E}_{X^n} \{\text{Tr}\{\Pi \rho^{X^n}\}\} = \text{Tr}\{\Pi \mathbb{E}_{X^n} \{\rho^{X^n}\}\} = \text{Tr}\{\Pi \rho_{\text{exp}}^{\otimes n}\} \geq 1 - \epsilon. \quad (\text{A.122})$$

The last inequality results from the unit probability property of the typical subspace. Same logic applies for the conditional typical projectors. Provided a codeword lies in the typical subspace, the conditional typical projector would project it to the corresponding conditional typical subspace with high degree of probability

$$\mathbb{E}_{X^n} \{\text{Tr}\{\Pi_i \rho^{X^n}\}\} \geq 1 - \epsilon. \quad (\text{A.123})$$

Moreover the Gentle Measurement Lemma [103] states that

Gentle Measurement Lemma: Given an ensemble $\{p_X(x), \rho^x\}$ with $\rho_{\text{exp}} \equiv \sum_x p_X(x) \rho^x$, suppose a POVM element Λ exists such that $0 < \Lambda \leq I$ succeeds with high probability on the state ρ_{exp}

$$\text{Tr}\{\Lambda \rho_{\text{exp}}\} = \sum_x p_X(x) \text{Tr}\{\Lambda \rho_x\} \geq 1 - \epsilon. \quad (\text{A.124})$$

Then the sub-normalized state $\sqrt{\Lambda} \rho_x \sqrt{\Lambda}$ is close in expected trace distance to the original state ρ_x

$$\sum_x p_X(x) \left\| \sqrt{\Lambda} \rho_x \sqrt{\Lambda} - \rho_x \right\|_1 \leq 2\sqrt{\epsilon}. \quad (\text{A.125})$$

Therefore, the typical subspace measurement in Bob's case is gentle. The typical subspace projector Π barely disturbs the state on which conditional typical projector Π_i would subsequently act.

Error Analysis: The probability that Bob detects the m -th codeword correctly under the sequential decoding scheme some decoding POVM $\{\Lambda_m\}$ at his disposal to figure out the codeword (message) sent by Alice. If M' denotes the random variable for Bob's estimate of the message, then the probability that he detects the correct message m is as follows

$$\text{Tr}\{\Pi_m \hat{\Pi}_{m-1} \dots \hat{\Pi}_1 \Pi \rho^{x^n(m)} \Pi \hat{\Pi}_1 \dots \hat{\Pi}_{m-1} \Pi_m\}. \quad (\text{A.126})$$

where $\hat{\Pi}_i = \mathbb{I} - \Pi_i$. Therefore, the probability of error associated with the detection of m -th codeword is

$$1 - \text{Tr}\{\Pi_m \hat{\Pi}_{m-1} \dots \hat{\Pi}_1 \Pi \rho^{x^n(m)} \Pi \hat{\Pi}_1 \dots \hat{\Pi}_{m-1} \Pi_m\}, \quad (\text{A.127})$$

from which the average probability of error follows as

$$1 - \frac{1}{|\mathcal{M}|} \sum_m \text{Tr}\{\Pi_m \hat{\Pi}_{m-1} \dots \hat{\Pi}_1 \Pi \rho^{x^n(m)} \Pi \hat{\Pi}_1 \dots \hat{\Pi}_{m-1} \Pi_m\}. \quad (\text{A.128})$$

Since Alice constructs her code randomly, therefore one analyzes the expectation of the average error probability. The expectation is with respect to all possible choices of code

$$1 - \mathbb{E}_{X^n} \left\{ \frac{1}{|\mathcal{M}|} \sum_m \text{Tr}\{\Pi_m \hat{\Pi}_{m-1} \dots \hat{\Pi}_1 \Pi \rho^{x^n(m)} \Pi \hat{\Pi}_1 \dots \hat{\Pi}_{m-1} \Pi_m\} \right\}. \quad (\text{A.129})$$

A bound on this expectation can be found from a result called *non-commutative union bound* [104] which states the following

Non-commutative Union Bound Lemma: For a subnormalized state σ such that $0 \leq \sigma$ and $\text{Tr}\{\sigma\} \leq 1$ with $\Pi_1, \dots, \Pi_N$

$$\text{Tr}\{\sigma\} - \text{Tr}\{\Pi_N \dots \Pi_1 \sigma \Pi_1 \dots \Pi_N\} \leq 2 \sqrt{\sum_{i=1}^N \text{Tr}\{(I - \Pi_i) \sigma\}}. \quad (\text{A.130})$$

The expression in equation (A.129) can be cast in the above form by noting that

$$\begin{aligned} 1 &= \mathbb{E}_{X^n} \left\{ \frac{1}{|\mathcal{M}|} \sum_m \text{Tr}\{\rho^{X^n(m)}\} \right\} \\ &= \mathbb{E}_{X^n} \left\{ \frac{1}{|\mathcal{M}|} \sum_m \text{Tr}\{\Pi \rho^{X^n(m)}\} \right\} + \mathbb{E}_{X^n} \left\{ \frac{1}{|\mathcal{M}|} \sum_m \text{Tr}\{\hat{\Pi} \rho^{X^n(m)}\} \right\} \\ &= \mathbb{E}_{X^n} \left\{ \frac{1}{|\mathcal{M}|} \sum_m \text{Tr}\{\Pi \rho^{X^n(m)}\} \right\} + \left\{ \frac{1}{|\mathcal{M}|} \sum_m \text{Tr}\{\hat{\Pi} \mathbb{E}_{X^n} \rho^{X^n(m)}\} \right\} \\ &= \mathbb{E}_{X^n} \left\{ \frac{1}{|\mathcal{M}|} \sum_m \text{Tr}\{\Pi \rho^{X^n(m)}\} \right\} + \text{Tr}\{\hat{\Pi} \rho_{\text{exp}}^{\otimes n}\} \\ &\leq \mathbb{E}_{X^n} \left\{ \frac{1}{|\mathcal{M}|} \sum_m \text{Tr}\{\Pi \rho^{X^n(m)}\} \right\} + \epsilon. \end{aligned} \quad (\text{A.131})$$

The first equality follows from the uniformity of the expected ensemble density over all messages. The second equality follows from $\Pi + \hat{\Pi} = I$. The third

equality follows from the linearity of the trace and the expectation operator. The forth equality is for the same reason for the first one. The fifth equality is due to the unit probability property of the typical subspace. The small quantity ϵ can be neglected. Hence one can write equation (A.129)

$$\mathbb{E}_{X^n} \left\{ \frac{1}{|\mathcal{M}|} \sum_m \text{Tr} \{ \Pi \rho^{X^n(m)} \Pi \} \right\} - \mathbb{E}_{X^n} \left\{ \frac{1}{|\mathcal{M}|} \sum_m \text{Tr} \{ \Pi_m \hat{\Pi}_{m-1} \dots \hat{\Pi}_1 \Pi \rho^{X^n(m)} \Pi \hat{\Pi}_1 \dots \hat{\Pi}_{m-1} \Pi_m \} \right\}. \quad (\text{A.132})$$

Applying the non-commutative bound for $\sigma = \Pi \rho^{X^n(m)} \Pi$, and sequence of projectors $\Pi_m \hat{\Pi}_{m-1} \dots \hat{\Pi}_1 \Pi \rho^{X^n(m)} \Pi \hat{\Pi}_1 \dots \hat{\Pi}_{m-1} \Pi_m$, one gets

$$\mathbb{E}_{X^n} \left\{ \frac{1}{|\mathcal{M}|} \sum_m 2 \left[\text{Tr} \{ (I - \Pi_m) \Pi \rho^{X^n(m)} \Pi \} + \sum_{i=1}^{m-1} \text{Tr} \{ \Pi_i \Pi \rho^{X^n(m)} \Pi \Pi_i \} \right]^{1/2} \right\}. \quad (\text{A.133})$$

Since a $y = \sqrt{x}$ function is always bounded from above by the $y = x$ function (concavity of square root), the above expression is upper bounded by

$$2 \left[\mathbb{E}_{X^n} \left\{ \left[\frac{1}{|\mathcal{M}|} \sum_m \text{Tr} \{ (I - \Pi_m) \Pi \rho^{X^n(m)} \Pi \} + \frac{1}{|\mathcal{M}|} \sum_m \sum_{i=1}^{m-1} \text{Tr} \{ \Pi_i \Pi \rho^{X^n(m)} \Pi \Pi_i \} \right] \right\} \right]^{1/2}. \quad (\text{A.134})$$

Considering the first term

$$\mathbb{E}_{X^n} \left\{ \left[\frac{1}{|\mathcal{M}|} \sum_m \text{Tr} \{ (I - \Pi_m) \Pi \rho^{X^n(m)} \Pi \} \right] \right\} \quad (\text{A.135})$$

$$\leq \mathbb{E}_{X^n} \left\{ \frac{1}{|\mathcal{M}|} \sum_m \left[\text{Tr} \{ (I - \Pi_m) \rho^{X^n(m)} \} + \left\| \rho^{X^n(m)} - \Pi \rho^{X^n(m)} \Pi \right\|_1 \right] \right\} \quad (\text{A.136})$$

$$\leq \epsilon + 2\sqrt{\epsilon} \quad (\text{A.137})$$

The first inequality follows from the trace distance relation in equation (A.74). The second inequality is possible because of the gentle measurement lemma, and the unit probability property of conditional typicality.

Now considering the second term

$$\begin{aligned} \mathbb{E}_{X^n} \left\{ \frac{1}{|\mathcal{M}|} \sum_m \sum_{i=1}^{m-1} \text{Tr} \{ \Pi_i \Pi \rho^{X^n(m)} \Pi \Pi_i \} \right\} &= \mathbb{E}_{X^n} \left\{ \sum_{i=1}^{m-1} \text{Tr} \{ \Pi_i \Pi \rho^{X^n(m)} \Pi \Pi_i \} \right\} \\ &= \sum_{i \neq m} \text{Tr} \left\{ \mathbb{E}_{X^n} \{ \Pi_i \} \Pi \underbrace{\mathbb{E}_{X^n} \{ \rho^{X^n(m)} \}}_{=\rho_{\text{exp}}^{\otimes n}} \Pi \right\} \leq \sum_{i \neq m} 2^{-n[H(B)-\delta]} \text{Tr} \left\{ \mathbb{E}_{X^n(m)} \{ \Pi_i \} \Pi \right\} \\ &\leq \sum_{i \neq m} 2^{-n[H(B)-\delta]} \text{Tr} \left\{ \mathbb{E}_{X^n(m)} \{ \Pi_i \} \right\} \leq \sum_{i \neq m} 2^{-n[H(B)-\delta]} 2^{n[H(B|X)+\delta]} \\ &= \sum_{i \neq m} 2^{-n[I(X;B)-2\delta]} \leq |\mathcal{M}| 2^{-n[I(X;B)-2\delta]}. \end{aligned}$$

The first and the second equality follow because $X^n(m)$ and $X^n(i)$ are independent. The first inequality is due to the equipartition property of the typical subspace. The second inequality follows from the fact that $\Pi_i \leq I$. The third inequality is because of the exponentially smaller cardinality property of the typical subspace. The third equality exploits the definition of mutual information. The last inequality is straightforward.

Finally, one can put together

$$p_e = 1 - \mathbb{E}_{X^n} \left\{ \frac{1}{|\mathcal{M}|} \sum_m \text{Tr} \{ \Pi_m \hat{\Pi}_{m-1} \dots \hat{\Pi}_1 \Pi \rho^{x^n(m)} \Pi \hat{\Pi}_1 \dots \hat{\Pi}_{m-1} \Pi_m \} \right\} \leq \epsilon + 2[(\epsilon + 2\sqrt{\epsilon}) + |\mathcal{M}| 2^{-n[I(X:B)-2\delta]}]^{1/2}. \quad (\text{A.138})$$

As $n \rightarrow \infty, \epsilon \rightarrow 0$; therefore as the long as the size of the message set is

$$|\mathcal{M}| \approx 2^{n[I(X:B)-2\delta]}, \quad (\text{A.139})$$

there exists an achievable rate. The classical capacity $C(\mathcal{N})$ of a quantum channel $\mathcal{N}$ is equal to the supremum of all achievable rates for classical communication. The classical capacity theorem (also known as HSW theorem after Holevo-Schumacher-Westmoreland [105]) states that the classical capacity of a quantum channel is equal to the regularization of the Holevo information

$$C(\mathcal{N}) = \chi_{\text{reg}}(\mathcal{N}) \equiv \lim_{k \rightarrow \infty} \frac{1}{k} \chi(\mathcal{N}^{\otimes k}), \quad (\text{A.140})$$

where the Holevo information is

$$\chi(\mathcal{N}) \equiv \max_{\rho_{XB}} I(X; B). \quad (\text{A.141})$$

The maximization is with respect to the classical-quantum state

$$\rho_{XB} = \sum_x p_X(x) |x\rangle \langle x|_X \otimes \mathcal{N}_{A' \rightarrow B}(\rho_{A'}^x) \quad (\text{A.142})$$

The proof of the theorem suggests the condition for achieving Holevo information as the capacity Alice should choose the cardinality of her message set as $|\mathcal{M}|$ for arbitrarily low p_e .

A.6 Private Classical Capacity Theorem

The private classical capacity theorem establishes the maximum rate at which Alice can communicate classical information privately to Bob without anyone else in the universe knowing about the content of the information. This theorem is significant in the context of quantum key distribution as it suggests the maximum rate at which two parties can generate a shared secret key.

The set-up for the information processing task is exactly the same as for transmission of information using a public classical code. But in this case the previous quantum channel $\mathcal{N}_{A' \rightarrow B}$ has an isometric extension of the form $U_{A' \rightarrow BE}^{\mathcal{N}}$ so that

$$\hat{\mathcal{N}}_{A' \rightarrow E}(\rho_{A'm}^m) = \text{Tr}_B \{ U_{A' \rightarrow BE}^{\mathcal{N}}(\rho_{A'm}^m) \}. \quad (\text{A.143})$$

For Alice transmitting a message m encoding in $\rho_{A'}^m$, Eve receives

$$\omega_{E^n}^m \equiv \hat{\mathcal{N}}_{A^n \rightarrow E^n}(\rho_{A^n}^m) \quad (\text{A.144})$$

Together with the maximal probability of error, there is a condition for privacy or secrecy in the case for private communication

$$\forall m \in \mathcal{M} : \frac{1}{2} \|\omega_{E^n}^m - \sigma_{E^n}\|_1 \leq \epsilon. \quad (\text{A.145})$$

which states that regardless of which message Alice transmits through the quantum channel Eve's state is always ϵ -close to a constant σ_{E^n} in the trace distance sense. From this condition one can show that as $\epsilon \rightarrow 0$ for $n \rightarrow \infty$, Eve's mutual information with Alice becomes vanishingly small.

The private classical capacity $C_P(\mathcal{N})$ of a quantum channel according to the corresponding theorem is given by

$$C_P(\mathcal{N}) = P_{\text{reg}}(\mathcal{N}) \equiv \lim_{k \rightarrow \infty} \frac{1}{k} P(\mathcal{N}^{\otimes k}). \quad (\text{A.146})$$

The private information $P(\mathcal{N})$ is defined as

$$P(\mathcal{N}) \equiv \max_{\rho_{XA'}} [I(X; B)_\sigma - I(X; E)_\sigma], \quad (\text{A.147})$$

where $\rho_{XA'}$ denotes the state which is the expected density operator for Alice's ensemble

$$\rho_{XA'} \equiv \sum_x p_X(x) |x\rangle \langle x|_X \otimes \rho_{A'}^x, \quad (\text{A.148})$$

and $\sigma_{XBE} \equiv \mathcal{U}_{A' \rightarrow BE}^{\mathcal{N}}(\rho_{XA'})$

To satisfy both the conditions Alice uses codes with two indices $\{x_n(m, k)\}_{m \in \mathcal{M}, k \in \mathcal{K}}$. The message m is randomized over the second index k to cover up the typical subspace of Eve so that Eve is in complete ignorance about the m . For Bob, to distinguish the transmitted states unequivocally, the size of the private code should be

$$|\mathcal{M}| |\mathcal{K}| \approx 2^{nI(X; B)}. \quad (\text{A.149})$$

Moreover in order for Alice to approximately cover Eve's typical subspace, $\mathcal{K}$ at least needs to be of the size $\approx 2^{nI(X; E)}$. Based on these considerations the number of distinguishable messages Alice can send to Bob is

$$|\mathcal{M}| \approx 2^{n[I(X; B) - I(X; E)]}, \quad (\text{A.150})$$

which leads to

$$\boxed{\frac{1}{n} \log |\mathcal{M}| \approx \underbrace{I(X; B) - I(X; E)}_{\text{secret rate}}}. \quad (\text{A.151})$$

Therefore, by maximizing the R.H.S. over Alice's ensemble Alice and Bob can operate at the capacity.

A.7 Entropy with continuous variables

So far all the results described in this chapter are set in the discrete world, where Hilbert spaces are finite dimensional. For continuous variable systems corresponding Hilbert spaces are infinite dimensional. However, the concepts of typical sets and typical subspaces from the finite dimensional Hilbert spaces effortlessly generalizes to the infinite-dimensional Hilbert spaces for the Gaussian states. Therefore, the results are still valid in this case. In the following, we introduce some information measures based on continuous density functions.

Differential Entropy: The differential entropy of a continuous random variable X with probability density function $p_X(x)$ is defined as

$$H(X) = - \int_{S_x} p_X(x) \log_2 p_X(x) dx. \quad (\text{A.152})$$

where S_x is the support set of the random variable. The size of the typical set for X is now a volume $|T_X|$.

Joint and Conditional Entropy: If two continuous random variables X and Y have joint density $p_{X,Y}(x,y)$, then their joint entropy and conditional entropy are given respectively as

$$H(X,Y) = - \int_{S_x} \int_{S_y} p_{X,Y}(x,y) \log_2 p_{X,Y}(x,y) dx dy, \quad (\text{A.153})$$

and

$$H(X|Y) = \int_{S_y} p_Y(y) H(X|Y=y) dy \quad (\text{A.154})$$

$$\stackrel{*}{=} - \int_{S_x} \int_{S_y} p_{X,Y}(x,y) \log_2 p_{X|Y}(x|Y=y) dx dy = H(X,Y) - H(Y), \quad (\text{A.155})$$

where for arriving at the expression in (*) the definition of $H(X|Y=y)$, and the fact that $p_{X|Y}(x|Y=y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$ have been used.

Mutual Information: The mutual information between two continuous random variables are defined similar to the discrete case

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) = H(X) + H(Y) - H(X,Y) \quad (\text{A.156})$$

Gaussian Distribution

Gaussian distribution is the most important continuous distribution from the practical standpoint, and also from the perspective of this thesis.

A continuous random variable X with a Gaussian distribution

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-(x-\mu_X)^2/2\sigma_X^2}. \quad (\text{A.157})$$

is completely characterized by its mean μ_X , and variance σ_X . The entropy for X , which can be found according to the definition in equation (A.152), depends only on the variance

$$H(X) = \frac{1}{2} \log_2(2\pi\sigma_X^2 e). \quad (\text{A.158})$$

The entropy for a multivariate normal distribution involving random variables $X_1, X_2, \dots, X_n$ with mean vector $\bar{\mu}$, and covariance matrix Σ can be shown to be

$$H(X_1, X_2, \dots, X_n) = \frac{1}{2} \log_2(2\pi e)^n |\Sigma|, \quad (\text{A.159})$$

where $|\Sigma|$ is the determinant of Σ . The most relevant example for us is the bivariate case with . The joint Gaussian distribution for two continuous random variables X and Y have the generic form

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left(\frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sqrt{\sigma_X\sigma_Y}} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} \right) \right\} \quad (\text{A.160})$$

with μ_X, μ_Y and σ_X^2, σ_Y^2 denote the respective means and variances for X and Y . The ρ term denotes correlation, and is given by $\frac{\sigma_{XY}}{\sqrt{\sigma_X^2\sigma_Y^2}}$, where σ_{XY} is the covariance term. The variance terms and the covariance terms can be grouped together in the covariance matrix

$$\Sigma_{XY} = \begin{pmatrix} \sigma_X & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y \end{pmatrix} \quad (\text{A.161})$$

According to equation (A.153), the joint entropy for X and Y , are given by

$$H(X, Y) = \frac{1}{2} \log_2(2\pi e)^2 |\Sigma_{XY}| = \frac{1}{2} \log_2(2\pi e)^2 (\sigma_X\sigma_Y - \sigma_{XY}^2) \quad (\text{A.162})$$

The marginals $f_X(x)$ and $f_Y(y)$ of $f_{X,Y}(x, y)$ are also Gaussians. Since $f_{Y|X}(y|X = x) = \frac{f_{XY}(x, y)}{f_X(x)}$, the conditional distribution for Y when X is given

$$f_{Y|X}(y|X = x) = \frac{1}{\sqrt{2\pi(1-\rho^2)\sigma_Y^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left(y - \mu_Y - \frac{\rho\sigma_Y(x - \mu_X)}{\sigma_X} \right)^2 \right\}, \quad (\text{A.163})$$

which is Gaussian with modified mean and variance

$$Y|X = x \sim \mathcal{G} \left(\mu_Y - \frac{\rho\sigma_Y(x - \mu_X)}{\sigma_X}, \underbrace{(1-\rho^2)\sigma_Y^2}_{\sigma_{Y|X}} \right). \quad (\text{A.164})$$

Therefore, one can write the conditional entropy for a bi-variate distribution

$$H_{Y|X} = \frac{1}{2} \log_2(2\pi\sigma_{Y|X}^2 e), \quad (\text{A.165})$$

and the mutual information

$$I(X;Y) = H(Y) - H(Y|X) = \frac{1}{2} \log_2 \left[\frac{\sigma_Y^2}{\sigma_{Y|X}^2} \right], \quad (\text{A.166})$$

$$= H(X) - H(X|Y) = \frac{1}{2} \log_2 \left[\frac{\sigma_X^2}{\sigma_{X|Y}^2} \right], \quad (\text{A.167})$$

$$= H(X) + H(Y) - H(X,Y) = \frac{1}{2} \log_2 \left[\frac{\sigma_X^2 \sigma_Y^2}{|\Sigma_{XY}|} \right]. \quad (\text{A.168})$$

The above results can be contextualized in studying information transmission through an additive white Gaussian (AWGN) channel (Figure A.4). Let the in-

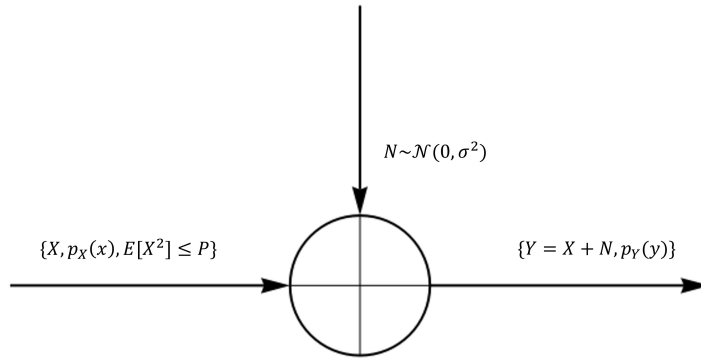


FIGURE A.4: Schematic of information transmission through an AWGN channel. The input signal is power limited since the channel cannot withstand arbitrary large power. The challenge is to enhance correlation between the input and the output in the presence of Gaussian noise while also taking into account the power constraint.

put signal be denoted by the random variable X with probability distribution $p_X(x)$, and maximum power $E[X^2]$ as allowed by the channel. The channel adds noise, which is normally distributed $N \sim \mathcal{N}(0, \sigma^2)$, to the signal. Since the noise is additive, the output will be $Y = X + N$. X and N are uncorrelated i.e. $E[XN] = 0$. Therefore

$$E[Y^2] = E[(X + N)^2] = E[X^2 + 2XN + N^2] = E[X^2] + E[N^2] \leq P + \sigma^2 = \tilde{\sigma}^2, \quad (\text{A.169})$$

Given $X = x$, Y is Gaussian distributed with translated mean $0 + x$ and variance σ^2 i.e. $Y \sim \mathcal{G}(0 + x, \sigma^2)$. Hence the conditional distribution should be of the form

$$Y|X \sim \mathcal{G}(X, \sigma^2), \quad (\text{A.170})$$

which leads to the conditional entropy

$$H(Y|X) = H(\mathcal{N}(X, \sigma^2)) = \frac{1}{2} \log_2(2\pi\sigma^2 e) \quad (\text{A.171})$$

The mutual information between the input and the output

$$I(X; Y) = H(Y) - H(Y|X) = H(Y) - \frac{1}{2} \log_2(2\pi\sigma^2 e). \quad (\text{A.172})$$

To achieve the capacity (maximum information transfer rate at an arbitrary low probability of error) over the channel, one needs to maximize the mutual information over the input probability distribution taking the power constraint into account

$$\begin{aligned} C &= \max_{p_Y(y): E[Y^2] \leq \tilde{\sigma}^2} I(X; Y) \\ &= \max_{p_Y(y): E[Y^2] \leq \tilde{\sigma}^2} H(Y) - H(Y|X) \\ &= \left(\max_{p_Y(y): E[Y^2] \leq \tilde{\sigma}^2} H(Y) \right) - \frac{1}{2} \log_2(2\pi\sigma^2 e). \end{aligned} \quad (\text{A.173})$$

It can be shown that the maximum occurs when $p_Y(y)$ is a Gaussian distribution implying $\max_{p_Y(y): E[Y^2] \leq \tilde{\sigma}^2} H(Y) = \frac{1}{2} \log_2(2\pi\tilde{\sigma}^2 e)$. It leads to

$$C = \frac{1}{2} \log_2\left(\frac{\tilde{\sigma}^2}{\sigma^2}\right) = \frac{1}{2} \log_2\left(1 + \frac{P}{\sigma^2}\right) = \frac{1}{2} \log_2(1 + \text{SNR}). \quad (\text{A.174})$$

Since the expression for $I(X; Y)$ is symmetric in X and Y , one can conclude the input probability distribution $p_X(x)$ for X has to be Gaussian. This is the reason while implementing continuous variable quantum key distribution (CV-QKD) one chooses Gaussian modulation to achieve the maximum possible information transfer rate (hence the highest key rate).

Appendix B

Phase Space Quantum Mechanics

To write this appendix, these sources were principally consulted [43], [46], [55], [53].

B.1 Wigner Characteristic Function

Let us consider a classical random variable X with realizations x . x can take value from a continuous set. The probability distribution which characterizes it is $\rho_X(x)$ such that $\rho_X(x) \geq 0$ and $\int \rho_X(x)dx = 1$. From the probability distribution one can calculate any n th order moment of X

$$\langle x^n \rangle = \int x^n \rho_X(x) dx. \quad (\text{B.1})$$

All the moments are generated from a characteristic function which is defined as the

$$C(k) = \langle e^{jkx} \rangle = \int e^{jkx} \rho_X(x) dx = \sum_{n=0}^{\infty} \frac{(jk)^n}{n!} \langle x^n \rangle, \quad (\text{B.2})$$

where k has the inverse dimension of x . Once $C(k)$ is known, the moments follow from

$$\langle x^n \rangle = \frac{1}{j^n} \left. \frac{d^n C(k)}{dk^n} \right|_{k=0}. \quad (\text{B.3})$$

From the definition of $C(k)$, it is obvious that the probability density function is just the Fourier transform of characteristic function

$$\rho_X(x) = \frac{1}{2\pi} \int dk e^{-jkx} C(k). \quad (\text{B.4})$$

Therefore, if all the moments $\langle x^n \rangle$ of a random variable X are known, then its characteristic function is known, which eventually lets us calculate the transfer function $\rho_X(x)$. The characteristic functions in quantum optics follows similarly with some adjustments to the Fourier kernel e^{jkx} . First, the Fourier integral is

performed over a complex plane. The Fourier conjugate variables u, α are now complex, and related as

$$g(u) = \int f(\alpha) e^{\alpha^* u - \alpha u^*} d^2 \alpha, \quad (\text{B.5a})$$

$$f(\alpha) = \frac{1}{(2\pi)^2} \int g(u) e^{\alpha u^* - \alpha^* u} d^2 u. \quad (\text{B.5b})$$

Upon substituting $\alpha = \alpha_R + j\alpha_I$, and $u = u_R + ju_I$, in (B.5a) and (B.5b) one will get a two-fold Fourier integral for both the equations. In equation (B.5a), one can identify $g(u)$ as the Wigner characteristic function $C_W(u)$, which through Fourier transform produces Wigner distribution function $f(\alpha) = W(\alpha)$ up to some scaling factor. One can manipulate equation (B.5a) in the following way

$$\begin{aligned} \chi_W(u) &= \int f(\alpha) e^{\alpha^* u - \alpha u^*} d^2 \alpha \\ &= \int f(\alpha) e^{\alpha^* u} e^{-\alpha u^*} d^2 \alpha \\ &= \int f(\alpha) e^{\alpha^* u} \langle \alpha | \alpha \rangle e^{-\alpha u^*} d^2 \alpha \\ &= \int f(\alpha) \langle \alpha | e^{\alpha^* u} e^{-\alpha u^*} | \alpha \rangle d^2 \alpha \\ &= \int f(\alpha) \langle \alpha | e^{\hat{a}^\dagger u} e^{-\hat{a} u^*} | \alpha \rangle d^2 \alpha \\ &= \text{Tr} \left[\int f(\alpha) | \alpha \rangle \langle \alpha | e^{\hat{a}^\dagger u} e^{-\hat{a} u^*} d^2 \alpha \right] \\ &= e^{\frac{1}{2}|u|^2} \text{Tr} \left[\int f(\alpha) | \alpha \rangle \langle \alpha | e^{\hat{a}^\dagger u} e^{-\hat{a} u^*} e^{-\frac{1}{2}[\hat{a}^\dagger u, -\hat{a} u^*]} d^2 \alpha \right] \\ &= \text{Tr} \left[\int e^{\frac{1}{2}|u|^2} f(\alpha) | \alpha \rangle \langle \alpha | e^{u \hat{a}^\dagger - u^* \hat{a}} d^2 \alpha \right] \\ &= \text{Tr}[\rho \hat{D}(u)] \\ &= \langle \hat{D}(u) \rangle, \end{aligned} \quad (\text{B.6})$$

where we have used the unit norm condition of coherent states $\langle \alpha | \alpha \rangle = 1$ in the third line, $f(\hat{a}) | \alpha \rangle = f(\alpha) | \alpha \rangle$ in the fifth line, cyclic property of trace operation in the sixth line, disentangling theorem in the seventh line, and the resolution of a quantum state in coherent basis $\int W(\alpha) | \alpha \rangle \langle \alpha | = \rho$ in the eighth line. The Wigner characteristic function results from the symmetrically ordered arrangement of the ladder operators

$$\chi_W(u) = \text{Tr}[\rho \hat{D}(u)], \quad (\text{B.7})$$

which in turn gives the Wigner distribution by Fourier transform

$$W(\alpha) = \frac{1}{(2\pi)^2} \int \chi_W(u) e^{u^* \alpha - u \alpha^*} d^2 u. \quad (\text{B.8})$$

Finally, before concluding, we provide the vectorial form of the displacement operator which is useful for analyzing multimode quantum state

$$\hat{D}(\boldsymbol{\eta}) = e^{i \hat{\mathbf{x}}^T \boldsymbol{\Omega} \boldsymbol{\eta}}, \quad (\text{B.9})$$

where $\boldsymbol{\eta}$ contains the quadrature eigenvalues, and $\hat{\mathbf{x}}^T$ contains the quadrature operators for all the modes. It is easy to check for the single mode case k ,

$$\hat{D}_k \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = e^{j(\eta_2 \hat{q}_k - \eta_1 \hat{p}_k)} \text{ reduces to familiar form with } \boldsymbol{\eta} \equiv \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \sqrt{2} \begin{pmatrix} \mathcal{R}e(\alpha) \\ \mathcal{I}m(\alpha) \end{pmatrix}$$

B.2 Properties of Wigner Function

1. Wigner function produces the quadrature probability distributions as its marginal distributions. Let us consider the state ρ of a single mode electromagnetic field. The corresponding Wigner representation is

$$W_\rho(\mathbf{x}) = \int_{\mathbb{R}^2} \frac{d^2 \boldsymbol{\eta}}{(2\pi)^2} e^{-j\mathbf{x}^T \boldsymbol{\Omega} \boldsymbol{\eta}} \chi(\boldsymbol{\eta}), \quad (\text{B.10})$$

where $\boldsymbol{\eta} = \begin{pmatrix} q' \\ p' \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} q \\ p \end{pmatrix}$. Integrating with respect to the p coordinates

$$\begin{aligned} & \int_{\mathbb{R}} dp W_\rho(\mathbf{x}) \\ &= \int_{\mathbb{R}} dp \int_{\mathbb{R}^2} \frac{d^2 \boldsymbol{\eta}}{(2\pi)^2} e^{-j\mathbf{x}^T \boldsymbol{\Omega} \boldsymbol{\eta}} \chi(\boldsymbol{\eta}) \\ &= \int_{\mathbb{R}^2} \frac{d^2 \boldsymbol{\eta}}{(2\pi)^2} \int_{\mathbb{R}} dp e^{-j(qp' - pq')} \chi_\rho(\boldsymbol{\eta}) \\ &= \int_{\mathbb{R}^2} \frac{d^2 \boldsymbol{\eta}}{(2\pi)^2} \left[\int_{\mathbb{R}} dp e^{jpq'} \right] e^{-jq p'} \chi_\rho(\boldsymbol{\eta}) \\ &= \int_{\mathbb{R}^2} \frac{d^2 \boldsymbol{\eta}}{(2\pi)^2} 2\pi \delta(q') e^{-jq p'} \chi_\rho(\boldsymbol{\eta}) \\ &= \int_{\mathbb{R}} \frac{dp'}{2\pi} e^{-jq p'} \underbrace{\eta_\rho(0, p')}_{\text{Tr}[\rho \hat{D}(0, p')]} \\ &= \int_{\mathbb{R}} \frac{dp'}{2\pi} e^{-jq p'} \int_{\mathbb{R}} dy |y\rangle \langle y| \rho \hat{D}(0, p') \\ &= \int_{\mathbb{R}} dy \int_{\mathbb{R}} \frac{dp'}{2\pi} e^{-jq p'} \langle y| \rho e^{-jp' \hat{q}} |y\rangle \\ &= \int_{\mathbb{R}} dy \int_{\mathbb{R}} \frac{dp'}{2\pi} e^{-j(q-y)p'} \langle y| \rho |y\rangle \\ &= \int_{\mathbb{R}} dy 2\pi \delta(q-y) \langle y| \rho |y\rangle \\ &= \langle q| \rho |q\rangle \end{aligned} \quad (\text{B.11})$$

gives the probability distribution for q -coordinate. In a similar fashion, one can show

$$\int_{\mathbb{R}} dq W_{\rho}(\mathbf{x}) = \langle p | \rho | p \rangle \quad (\text{B.12})$$

2. From the previous derivation, it is trivial to show that the Wigner function is normalized

$$\int_{\mathbb{R}^2} d^2\eta W_{\rho}(\mathbf{x}) = \int_{\mathbb{R}} dq \int_{\mathbb{R}} dp W_{\rho}(\mathbf{x}) = \int_{\mathbb{R}} dq \langle q | \rho | q \rangle = \text{Tr}(\rho) = 1 \quad (\text{B.13})$$

B.3 Admissible condition for a covariance matrix

The elements of a covariance matrix are

$$\Sigma_{j,k} = \frac{1}{2} \langle \delta \hat{x}_j \delta \hat{x}_k + \delta \hat{x}_k \delta \hat{x}_j \rangle.$$

If we define for a single mode

$$\delta \hat{\mathbf{x}} = \begin{pmatrix} \delta \hat{q} \\ \delta \hat{p} \end{pmatrix}, \text{ then } \delta \hat{\mathbf{x}}^T = \begin{pmatrix} \delta \hat{q} & \delta \hat{p} \end{pmatrix} \text{ and } \delta \hat{\mathbf{x}} \delta \hat{\mathbf{x}}^T = \delta \hat{X} = \begin{pmatrix} \delta \hat{q}^2 & \delta \hat{q} \delta \hat{p} \\ \delta \hat{p} \delta \hat{q} & \delta \hat{p}^2 \end{pmatrix}. \quad (\text{B.14})$$

The elements of the covariance for the single mode field are

$$\Sigma_{11} = \frac{1}{2} \langle \{ \hat{x}_1 - \langle \hat{x}_1 \rangle, \hat{x}_1 - \langle \hat{x}_1 \rangle \} \rangle = \langle \delta \hat{q}^2 \rangle \quad (\text{B.15a})$$

$$\Sigma_{12} = \frac{1}{2} \langle \{ \hat{x}_1 - \langle \hat{x}_1 \rangle, \hat{x}_2 - \langle \hat{x}_2 \rangle \} \rangle = \frac{1}{2} \langle (\delta \hat{q} \delta \hat{p} + \delta \hat{p} \delta \hat{q}) \rangle = \langle \delta \hat{q} \delta \hat{p} \rangle - j \quad (\text{B.15b})$$

$$\Sigma_{21} = \frac{1}{2} \langle \{ \hat{x}_2 - \langle \hat{x}_2 \rangle, \hat{x}_1 - \langle \hat{x}_1 \rangle \} \rangle = \frac{1}{2} \langle (\delta \hat{p} \delta \hat{q} + \delta \hat{q} \delta \hat{p}) \rangle = \langle \delta \hat{p} \delta \hat{q} \rangle + j \quad (\text{B.15c})$$

$$\Sigma_{22} = \frac{1}{2} \langle \{ \hat{x}_2 - \langle \hat{x}_2 \rangle, \hat{x}_2 - \langle \hat{x}_2 \rangle \} \rangle = \langle \delta \hat{p}^2 \rangle \quad (\text{B.15d})$$

Comparing the elements of Σ and $\delta \hat{X}$, one finds

$$\Sigma_{jk} = \langle \delta \hat{X}_{jk} \rangle - j \Omega_{jk} \Rightarrow \Sigma = \langle \delta \hat{\mathbf{x}} \delta \hat{\mathbf{x}}^T \rangle - j \Omega \quad (\text{B.16})$$

The determinant of the covariance matrix

$$\Sigma = \begin{pmatrix} \langle \delta \hat{q}^2 \rangle & \frac{1}{2} \langle \{ \delta \hat{p}, \delta \hat{q} \} \rangle \\ \frac{1}{2} \langle \{ \delta \hat{p}, \delta \hat{q} \} \rangle & \langle \delta \hat{p}^2 \rangle \end{pmatrix} \quad (\text{B.17})$$

is

$$\det(\Sigma) = \Sigma_{11}\Sigma_{22} - \Sigma_{12}\Sigma_{21} = \langle \delta \hat{q}^2 \rangle \langle \delta \hat{p}^2 \rangle - \left[\frac{1}{2} \langle \{ \delta \hat{p}, \delta \hat{q} \} \rangle \right]^2 \quad (\text{B.18})$$

From Robertson's uncertainty relation for the canonically conjugate variables

$$V(\hat{q})V(\hat{p}) - \left[\frac{1}{2} \langle \{\delta\hat{p}, \delta\hat{q}\} \rangle \right]^2 \geq \frac{1}{4} \quad (\text{B.19})$$

Therefore

$$\det(\Sigma) \geq \frac{1}{4} \quad (\text{B.20})$$

If Σ is diagonal of the form

$$\begin{pmatrix} \nu & 0 \\ 0 & \nu \end{pmatrix} \quad (\text{B.21})$$

then the uncertainty relation is simply $\nu \geq \frac{1}{2}$. In fact, Williamson's theorem states that for any real symmetric positive definite $2n \times 2n$ matrix Σ , there exists a symplectic transformation $S \in \text{Sp}(2n, \mathbb{R})$ such that the transformed matrix has the canonical scaled diagonal form. Therefore for the N -mode covariance matrix

$$S\Sigma S^T = \Sigma^\oplus, \quad \Sigma^\oplus := \bigoplus_{k=1}^N \nu_k \mathbb{1}_2 \quad (\text{B.22})$$

The diagonal matrix $\Sigma^\oplus$ is the Williamson form of Σ and $\{\nu_k\}_{k=1}^N$ are the symplectic eigenvalues which must satisfy $\nu_k \geq \frac{1}{2}$ due to uncertainty condition. Since S is invertible, if $\Sigma^\oplus$ qualifies as a valid covariance matrix so should Σ . Now let us consider the Hermitian covariance matrix $\Sigma^\oplus + \frac{1}{2}\Omega$. In this case $\hat{q}_j$'s are related to $\hat{p}_j$'s but not to any $\hat{q}_k$'s or $\hat{p}_k$'s. In the simplest case for a single mode

$$\begin{pmatrix} \nu & \frac{1}{2} \\ -\frac{1}{2} & \nu \end{pmatrix} \quad (\text{B.23})$$

In this case the spectrum of eigenvalues λ 's is $\nu \pm \frac{1}{2}$, which suggests $\lambda \geq 0$. This condition of non-zero eigenvalue amounts to demanding $\Sigma^\oplus + \frac{1}{2}\Omega$ is positive semi-definite. Now the symplectic transformation

$$\Sigma^\oplus + \frac{1}{2}\Omega = S \left[\Sigma + \frac{1}{2}\Omega \right] S^T \quad (\text{B.24})$$

preserves semi-definiteness. Therefore the admissibility condition of the covariance matrix Σ ultimately boils down to the positive semi definite condition of the Hermitian matrix $\Sigma + \frac{1}{2}\Omega$

$$\boxed{\Sigma + \frac{1}{2}\Omega \geq 0} \quad (\text{B.25})$$

B.4 Gaussian Unitaries

A unitary transformation in the Hilbert space has the following equivalent effect on the mean vector and the covariance matrix in the phase space. Two different quadratic forms which are equal must be connected via a linear transformation.

$$\begin{aligned}\sqrt{\det V_0} W_{\rho_0}(\mathbf{x}) &= \sqrt{\det V} W_{U\rho_0 U^\dagger}(S\mathbf{x} + \mathbf{a}) \\ \Rightarrow e^{-\frac{1}{2}(\mathbf{x}-\mathbf{d}_0)^T \Sigma_0^{-1}(\mathbf{x}-\mathbf{d}_0)} &= e^{-\frac{1}{2}(S\mathbf{x}+\mathbf{a}-\mathbf{d})^T \Sigma^{-1}(S\mathbf{x}+\mathbf{a}-\mathbf{d})} \\ &= e^{-\frac{1}{2}\{[\mathbf{x}+\Sigma^{-1}(\mathbf{a}-\mathbf{d})]^T \Sigma^T \Sigma^{-1} \Sigma [\mathbf{x}+\Sigma^{-1}(\mathbf{a}-\mathbf{d})]\}}\end{aligned}\quad (\text{B.26})$$

Comparing both sides we find

$$\bar{\mathbf{x}} = S\bar{\mathbf{x}}_0 + \mathbf{a} \quad (\text{B.27a})$$

$$\Sigma = S\Sigma S^T \quad (\text{B.27b})$$

Here $\mathbf{a} \in \mathbb{R}^{2N}$ and $S \in \mathbb{R}^{2N} \times \mathbb{R}^{2N}$. There is no restriction on $\mathbf{a}$ but S needs to satisfy $S\Omega S^T = \Omega$ in addition to being a real, symmetric matrix. In the Heisenberg picture, the unitary induces a linear transformation on the quadrature operators

$$U^\dagger \hat{\mathbf{x}} U = S\hat{\mathbf{x}} + \mathbf{a} \quad (\text{B.28})$$

It can be seen first for the mean vector

$$\text{Tr}[\rho_0 U^\dagger \hat{\mathbf{x}} U] = \text{Tr}[U \rho_0 U^\dagger \hat{\mathbf{x}}] = \text{Tr}[\rho \hat{\mathbf{x}}] = \mathbf{d} = S\mathbf{d}_0 + \mathbf{a} = \text{Tr}[\rho_0 (S\hat{\mathbf{x}} + \mathbf{a})] \quad (\text{B.29})$$

and then for the covariance matrix

$$\begin{aligned}\text{Tr}[\rho_0 (U^\dagger \hat{\mathbf{x}} U - \mathbf{d}_0)(U^\dagger \hat{\mathbf{x}} U - \mathbf{d}_0)^T] - J\Omega &\stackrel{(a)}{=} \text{Tr}[\rho (\hat{\mathbf{x}} - \mathbf{d})(\hat{\mathbf{x}} - \mathbf{d})^T] - J\Omega = \Sigma \\ &= S\Sigma_0 S^T = \text{Tr}[\rho_0 (S\hat{\mathbf{x}} - \underbrace{S\mathbf{d}_0}_{=\mathbf{d}-\mathbf{a}})(S\hat{\mathbf{x}} - S\mathbf{d}_0)^T] - J \underbrace{S\Omega S^T}_{=\Omega}\end{aligned}\quad (\text{B.30})$$

where in (a) we have used the following properties of unitary (i) $(U^\dagger)^T = U^*$, (ii) $UU^* = \mathbb{1}$, (iii) $U = U^T$. Preservation of symplectic invariance

$$\begin{aligned}U^\dagger \underbrace{[\hat{\mathbf{x}}_m, \hat{\mathbf{x}}_n]}_{2J\Omega_{mn}} U &= U^\dagger \left[\sum_{i=1}^{2N} S_{mi} \hat{\mathbf{x}}_i + a_m, \sum_{k=1}^{2N} S_{nk} \hat{\mathbf{x}}_k + a_n \right] U = \sum_{i,k=1}^N S_{mi} S_{kn} \underbrace{[\hat{\mathbf{x}}_i, \hat{\mathbf{x}}_k]}_{2J\Omega_{ik}} \\ &= S\Omega S^T = \Omega\end{aligned}\quad (\text{B.31})$$

B.5 Symplectic Analysis of Gaussian States

Williamson's theorem states that any positive-definite, real, symmetric matrix of even dimension $(2N \times 2N)$ is diagonalizable with a symplectic transformation. Applying this theorem for the covariance matrix of an arbitrary N -mode Gaussian states

$$\Sigma = S\Sigma^\oplus S^T, \quad \Sigma^\oplus := \bigoplus_{k=1}^N \nu_k \mathbb{1}_2 \quad (\text{B.32})$$

where S is a symplectic matrix. $\Sigma^\oplus$ is the diagonal form of Σ . ν_k 's are the symplectic eigenvalues. The symplectic spectrum $\{\nu_k\}_{k=1}^N$ can be computed as the standard spectrum of $|J\Omega V|$, where $|A| = \sqrt{A^\dagger A}$. Since $J\Omega V$ is Hermitian, its $2N$ eigenvalues are real. It implies the N symplectic eigenvalues of Σ are positive quantities. The above symplectic decomposition of covariance matrix has important physical implications for the Gaussian states. First, for example, $\Sigma^\oplus$ can be interpreted as the covariance matrix of N independent modes each in a thermal state with mean photon numbers $\{\bar{n}_k = \frac{\nu_k - 1}{2}\}_{k=1,2,\dots,N}$. Therefore in the Hilbert space picture

$$\hat{\rho}_G(\mathbf{0}, \Sigma^\oplus) \equiv \bigotimes_{k=1}^N \hat{\rho}_{th}\left(\frac{\nu_k - 1}{2}\right) \quad (\text{B.33})$$

Furthermore, we know, corresponding to the symplectic transformation $S\Sigma^\oplus S^T$, there is a corresponding unitary transformation $\hat{U}_S$ such that

$$\rho(\mathbf{0}, \Sigma^\oplus) = \hat{U}_S \rho(\mathbf{0}, \Sigma^\oplus) \hat{U}_S^\dagger \quad (\text{B.34})$$

In general, to center the Gaussian state at any point in the phase space

$$\rho(\bar{\mathbf{x}}, \Sigma) = \hat{D}(\bar{\mathbf{x}}) \hat{U}_S \underbrace{[\rho(\mathbf{0}, \Sigma^\oplus)]}_{\text{thermal decomposition}} \hat{U}_S^\dagger \hat{D}^\dagger(\bar{\mathbf{x}}) \quad (\text{B.35})$$

An arbitrary symplectic matrix $S \in \mathbb{R}^{2N \times 2N}$ is subject to the following decomposition (Euler decomposition)

$$S = \hat{O}_1 \left[\bigoplus_{k=1}^N S(r_k) \right] \hat{O}_2 \quad (\text{B.36})$$

where O_1, O_2 are symplectic, orthogonal matrices, while $S(r_1), \dots, S(r_k)$ is a set of one mode squeezing matrices. As a result, the corresponding canonical unitary $\hat{U}_S$ can be decomposed as

$$\hat{U}_S = \hat{U}_K \underbrace{\left[\bigotimes_{k=1}^N S(r_k) \right]}_{\text{Euler decomposition}} \hat{U}_L \quad (\text{B.37})$$

Any unitary transformation on an N -mode Gaussian state is equivalent to applying a passive transformation in the form of $N \times N$ interferometer ($\hat{U}_L$), followed by a parallel set of N one-mode squeezers ($\bigotimes_k S(r_k)$), followed by another $N \times N$ interferometer. Considering the results from the Euler decomposition and thermal decomposition together, one finds, an arbitrary multimode Gaussian state $\rho(\bar{\mathbf{x}}, \Sigma)$ can be realized from N -mode thermal states $\rho(\mathbf{0}, \bar{\Sigma}^\oplus)$ applying multimode interferometers and one-mode squeezers, and finally displacing them by $\bar{\mathbf{x}}$.

B.5.1 Two mode Gaussian States:

The covariance matrix of a two-mode Gaussian state can be written as

$$\Sigma_2 = \begin{pmatrix} \Sigma_A & \Sigma_C \\ \Sigma_C^T & \Sigma_B \end{pmatrix} \quad (\text{B.38})$$

where $\Sigma_A = \Sigma_A^T$, $\Sigma_B = \Sigma_B^T$, and $\Sigma_A, \Sigma_B, \Sigma_C \in \mathbb{R}^{2 \times 2}$. In the block diagonal form $\Sigma^\oplus = \nu_1 \mathbb{1} \oplus \nu_2 \mathbb{1}$. For two-mode Gaussian covariance matrix, there are two symplectic invariants (i) the determinant $\det \Sigma$, and (ii) the principle minor of order 2 i.e. $\Delta(\Sigma) = \det \Sigma_A + \det \Sigma_B + 2\det \Sigma_C$. The symplectic eigenvalues are related to the symplectic invariants as

$$\det \Sigma = \nu_1^2 \nu_2^2 \quad \Delta(\Sigma) = \nu_1^2 + \nu_2^2 \quad (\text{B.39})$$

from which it follows

$$\nu_{1,2}^2 = \frac{1}{2} \left[\Delta \pm \sqrt{\Delta^2 - 4\det \Sigma} \right] \quad (\text{B.40})$$

The physicality condition for the covariance matrix $\nu_j \geq 1$ can be written in terms of the symplectic invariant

$$\det \Sigma \geq 1 \quad \Delta(\Sigma) \leq 1 + \det \Sigma \quad (\text{B.41})$$

It has been shown that a Σ_2 can be put in block diagonal form

$$\begin{pmatrix} a & 0 & c & 0 \\ 0 & a & 0 & c \\ c & 0 & b & 0 \\ 0 & c & 0 & b \end{pmatrix} \quad (\text{B.42})$$

B.6 Hilbert Space vs Phase Space

Property	Hilbert Space $\mathcal{H}$	Phase Space Γ
dimension	∞	$2N$
structure	$\otimes$	$\oplus$
description	ρ	Σ
<i>bona fide</i>	$\rho \geq 0$	$\Gamma + j\Omega \geq 0$
operations	$\hat{U} : \begin{cases} \hat{U}\hat{U}^\dagger = 1 \\ \rho \rightarrow \hat{U}\rho\hat{U}^\dagger \end{cases}$	$S : \begin{cases} S\Sigma S^T = \Omega \\ \Sigma \rightarrow S\Sigma S^T \end{cases}$
spectra	$\hat{U}\rho\hat{U}^\dagger = \text{diag}\{\lambda_k\}$ $0 \leq \lambda_k \leq 1$	$S\Sigma S^T = \text{diag}\{\nu_k\}$ $1 \leq \nu_k \leq \infty$
pure states	$\lambda_i = 1, \lambda_{j \neq i} = 0$	$\nu_j = 1, \forall j \in \{1, \dots, N\}$
purity	$\text{Tr}[\rho^2] = \sum_k \lambda_k^2$	$1/\sqrt{\det\Sigma} = \prod_k \nu_k^{-1}$

TABLE B.1: Hilbert Space versus phase space representation of Gaussian states

B.7 Conditional Gaussian Dynamics

The n – mode subsystem of an $(n + m)$ – mode Gaussian state, with initial covariance matrix and first moments partitioned as

$$\Sigma = \begin{pmatrix} \Sigma_A & \Sigma_{AB} \\ \Sigma_{AB}^T & \Sigma_B \end{pmatrix}, \quad \bar{\mathbf{r}} = \begin{pmatrix} \bar{\mathbf{r}}_A \\ \bar{\mathbf{r}}_B \end{pmatrix} \quad (\text{B.43})$$

undergoes the following Gaussian mapping upon the general-dyne measurement of the remaining modes

$$\Sigma_A \mapsto \Sigma_A - \Sigma_{AB} \frac{1}{\Sigma_B + \Sigma_m} \Sigma_{AB}^T \quad (\text{B.44a})$$

$$\bar{\mathbf{r}}_A \mapsto \bar{\mathbf{r}}_A + \frac{1}{\Sigma_B + \Sigma_m} (\bar{\mathbf{r}}_m - \bar{\mathbf{r}}_B) \quad (\text{B.44b})$$

while the covariance matrix Σ_m characterises the general-dyne measurement, while the first moments $\bar{\mathbf{r}}_m$ label the measurement outcome. The above two results can be obtained in the following way.

Trace out system B , to observe how system A has changed, given that a Gaussian projective measurement $|\psi_G\rangle \langle \psi_G|$ has been performed on system B

$$\text{Tr}_B[\rho_G \rho_{AB}] = \text{Tr}_B[|\psi_G\rangle \langle \psi_G| \rho_{AB}] = \langle \psi_G | \rho | \psi_G \rangle \quad (\text{B.45})$$

From the Fourier-Weyl relation,

$$\rho = \frac{1}{(2\pi)^N} \int_{\mathbb{R}^{2n}} d\mathbf{r} e^{-\frac{1}{4}\mathbf{r}^T \Omega^T \Sigma \Omega \mathbf{r} + j\mathbf{r}^T \Omega^T \tilde{\mathbf{r}} \hat{D}_{\mathbf{r}}} = \frac{1}{(2\pi)^N} \int_{\mathbb{R}^{2n}} d\tilde{\mathbf{r}} e^{-\frac{1}{4}\tilde{\mathbf{r}}^T \Sigma \tilde{\mathbf{r}} + j\tilde{\mathbf{r}}^T \tilde{\mathbf{r}} \hat{D}_{\Omega^T \tilde{\mathbf{r}}}} \quad (\text{B.46})$$

where $\tilde{\mathbf{r}} = \Omega \mathbf{r}$, one can write

$$\text{Tr}[\rho_G \rho_{AB}] = \frac{1}{(2\pi)^{n+m}} \int_{\mathbb{R}^{2(n+m)}} e^{-\frac{1}{4}\tilde{\mathbf{r}}^T \Sigma \tilde{\mathbf{r}} + j\tilde{\mathbf{r}}^T \tilde{\mathbf{r}} \text{Tr}[\rho_G \hat{D}_{\Omega^T \tilde{\mathbf{r}}}] d\tilde{\mathbf{r}} \quad (\text{B.47})$$

Now, by definition, the characteristic function of a Gaussian state with covariance matrix Σ and vector of first moments $\tilde{\mathbf{r}}$

$$\chi(\tilde{\mathbf{r}}) = \text{Tr}[\hat{D}_{-\tilde{\mathbf{r}}} \rho] = e^{-\frac{1}{4}\tilde{\mathbf{r}}^T \Sigma \tilde{\mathbf{r}} e^{j\tilde{\mathbf{r}}^T \tilde{\mathbf{r}}} \quad (\text{B.48})$$

Therefore $\text{Tr}[\rho_G \hat{D}_{\tilde{\mathbf{r}}_B}]$ is the characteristic function of ρ_G up to a sign flip in the argument.

$$\text{Tr}[\rho_G \rho_{AB}] = \frac{1}{(2\pi)^{(n+m)}} \int_{\mathbb{R}^{2(n+m)}} e^{-\frac{1}{4}\tilde{\mathbf{r}}^T \Sigma \tilde{\mathbf{r}} \hat{D}_{\Omega^T \tilde{\mathbf{r}}_A} \underbrace{e^{-\frac{1}{4}\tilde{\mathbf{r}}_B^T \Sigma_B \tilde{\mathbf{r}}_B - j\tilde{\mathbf{r}}_B^T \tilde{\mathbf{r}}_B}}_{=\chi_G(-\tilde{\mathbf{r}}_B)} d\tilde{\mathbf{r}} \quad (\text{B.49})$$

where we have used

$$\hat{D}_{\tilde{\mathbf{r}}} = \hat{D}_{\tilde{\mathbf{r}}_A \oplus \tilde{\mathbf{r}}_B} = \hat{D}_{\tilde{\mathbf{r}}_A} \hat{D}_{\tilde{\mathbf{r}}_B} \quad (\text{B.50})$$

At this point one observes

$$\begin{aligned} \tilde{\mathbf{r}}^T \Omega \tilde{\mathbf{r}} &= \begin{pmatrix} \tilde{\mathbf{r}}_A^T & \tilde{\mathbf{r}}_B^T \end{pmatrix} \begin{pmatrix} \Sigma_A & \Sigma_{AB} \\ \Sigma_{AB}^T & \Sigma_B \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{r}}_A \\ \tilde{\mathbf{r}}_B \end{pmatrix} \\ &= \tilde{\mathbf{r}}_A^T \Sigma_A \tilde{\mathbf{r}}_A + \underbrace{\tilde{\mathbf{r}}_A^T \Sigma_{AB} \tilde{\mathbf{r}}_B + \tilde{\mathbf{r}}_B^T \Sigma_{AB}^T \tilde{\mathbf{r}}_A}_{=2\tilde{\mathbf{r}}_A^T \Sigma_{AB} \tilde{\mathbf{r}}_B} + \tilde{\mathbf{r}}_B^T \Sigma_B \tilde{\mathbf{r}}_B \end{aligned} \quad (\text{B.51})$$

which is a scalar. And similarly

$$\tilde{\mathbf{r}}^T \tilde{\mathbf{r}} = \begin{pmatrix} \tilde{\mathbf{r}}_A^T & \tilde{\mathbf{r}}_B^T \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{r}}_A \\ \tilde{\mathbf{r}}_B \end{pmatrix} = \tilde{\mathbf{r}}_A^T \tilde{\mathbf{r}}_A + \tilde{\mathbf{r}}_B^T \tilde{\mathbf{r}}_B \quad (\text{B.52})$$

Plugging in the expanded forms and rearranging the integral in (B.48)

$$\begin{aligned} &\frac{1}{(2\pi)^{(n+m)}} \int_{\mathbb{R}^{2n}} d\tilde{\mathbf{r}}_A \hat{D}_{\Omega^T \tilde{\mathbf{r}}_A} e^{-\frac{1}{4}\tilde{\mathbf{r}}_A^T \Sigma_A \tilde{\mathbf{r}}_A} e^{j\tilde{\mathbf{r}}_A^T \tilde{\mathbf{r}}_A} \\ &\left(\int_{\mathbb{R}^{2m}} d\tilde{\mathbf{r}}_B e^{-\frac{1}{4}\tilde{\mathbf{r}}_B^T (\Sigma_B + \Sigma_m) \tilde{\mathbf{r}}_B} e^{\tilde{\mathbf{r}}_B^T (-\frac{1}{2}\Sigma_{AB}^T \tilde{\mathbf{r}}_A + j(\tilde{\mathbf{r}}_B - \tilde{\mathbf{r}}_m))} \right) \end{aligned} \quad (\text{B.53})$$

Now, given a symmetric, real, positive-definite $2n \times 2n$ matrix A , and a $2n$ -dimensional vector $\mathbf{b}$, one has the equality (Gaussian multi-dimensional integral)

$$\int_{\mathbb{R}^{2n}} d\mathbf{r} e^{-\mathbf{r}^T A \mathbf{r} + \mathbf{r}^T \mathbf{b}} = \frac{\pi^n}{\sqrt{\det A}} e^{\frac{1}{4} \mathbf{b}^T A^{-1} \mathbf{b}} \quad (\text{B.54})$$

Comparing equation (B.53) and the integral over $\tilde{\mathbf{r}}_B$ in equation (B.52)

$$A = \Sigma_B + \Sigma_m \quad (\text{B.55a})$$

$$\mathbf{b} = -\frac{1}{2} \Sigma_{AB}^T \tilde{\mathbf{r}}_A + J(\tilde{\mathbf{r}}_B - \tilde{\mathbf{r}}_m) \quad (\text{B.55b})$$

Therefore the integral within the parentheses evaluates to

$$\begin{aligned} & \frac{(4\pi)^m}{\sqrt{\det(\Sigma_B + \Sigma_m)}} e^{\left(\frac{1}{2} \Sigma_{AB}^T \tilde{\mathbf{r}}_A + J(\tilde{\mathbf{r}}_B - \tilde{\mathbf{r}}_m)\right)^T \frac{1}{\Sigma_B + \Sigma_m} \left(\frac{1}{2} \Sigma_{AB}^T \tilde{\mathbf{r}}_A + J(\tilde{\mathbf{r}}_B - \tilde{\mathbf{r}}_m)\right)} \\ &= \frac{(4\pi)^m}{\sqrt{\det(\Sigma_B + \Sigma_m)}} e^{\frac{1}{4} \tilde{\mathbf{r}}_A^T \Sigma_{AB} \Sigma_B^{-1} \Sigma_{AB}^T \tilde{\mathbf{r}}_A + \tilde{\mathbf{r}}_A^T \Sigma_{AB} \Sigma_B^{-1} J(\tilde{\mathbf{r}}_B - \tilde{\mathbf{r}}_m) - (\tilde{\mathbf{r}}_B - \tilde{\mathbf{r}}_m)^T \Sigma_B^{-1} J(\tilde{\mathbf{r}}_B - \tilde{\mathbf{r}}_m)} \end{aligned} \quad (\text{B.56})$$

Plugging in the above expression into equation (B.53)

$$\begin{aligned} \text{Tr}[\rho_G \rho_{AB}] &= \frac{2^m e^{-(\tilde{\mathbf{r}}_B - \tilde{\mathbf{r}}_m)^T \Sigma_B^{-1} (\tilde{\mathbf{r}}_B - \tilde{\mathbf{r}}_m)}}{(2\pi)^n \sqrt{\det(\Sigma_B + \Sigma_m)}} \times \\ & \int_{\mathbb{R}^{2n}} \underbrace{e^{-\frac{\tilde{\mathbf{r}}_A^T}{4} (\Sigma_A - \Sigma_{AB} \Sigma_B^{-1} \Sigma_{AB}^T) \tilde{\mathbf{r}}_A} e^{J \tilde{\mathbf{r}}_A^T (\tilde{\mathbf{r}}_A + \Sigma_{AB} \Sigma_B^{-1} (\tilde{\mathbf{r}}_B - \tilde{\mathbf{r}}_m))}}_{=\chi_A(\tilde{\mathbf{r}}_A)} \hat{D}_{\Omega^T \tilde{\mathbf{r}}_A} d\tilde{\mathbf{r}}_A \end{aligned} \quad (\text{B.57})$$

Therefore the covariance matrix and the mean vector of subsystem A changes as

$$\begin{aligned} \Sigma_A &\mapsto \Sigma_A - \Sigma_{AB} \frac{1}{\Sigma_B + \Sigma_m} \Sigma_{AB}^T \\ \tilde{\mathbf{r}}_A &\mapsto \tilde{\mathbf{r}}_A + \frac{1}{\Sigma_B + \Sigma_m} (\tilde{\mathbf{r}}_m - \tilde{\mathbf{r}}_B) \end{aligned}$$

as stated in (B.44a) and (B.44b).

Appendix C

Discrete Modulation

C.1 Discrete modulation equations derivation

In the prepare and measure version of the two-state protocol, Alice sends the coherent states

$$|\beta\rangle = |\alpha e^{-j\pi/4}\rangle \quad (\text{C.1})$$

$$|-\beta\rangle = -|\alpha e^{-j\pi/4}\rangle \quad (\text{C.2})$$

each with probability $\frac{1}{2}$ to Bob. Hence Bob sees a maximally mixed state ρ_B given by

$$\rho_B = \frac{1}{2}(|\beta\rangle\langle\beta| + |-\beta\rangle\langle-\beta|) \quad (\text{C.3})$$

We want to find a purification of ρ_B . To that end, we want to write ρ_{AB} in the spectral decomposition form

$$\begin{aligned} \rho_B &= \frac{1}{2} |\beta\rangle\langle\beta| + \frac{1}{2} |-\beta\rangle\langle-\beta| \\ &= e^{-|\alpha|^2} \left[\sum_{p,q=0}^{\infty} \frac{|\alpha|^{2(p+q)}}{\sqrt{2p!}\sqrt{2q!}} (e^{-j\pi/2})^{p-q} |2p\rangle\langle 2q| \right. \\ &\quad \left. + \sum_{p,q=0}^{\infty} \frac{|\alpha|^{2(p+q+1)}}{\sqrt{(2p+1)!}\sqrt{(2q+1)!}} (e^{-j\pi/2})^{p-q} |2p+1\rangle\langle 2q+1| \right] \end{aligned} \quad (\text{C.4})$$

At this point let us define

$$|\phi_0\rangle := \frac{1}{\sqrt{\cosh|\alpha|^2}} \sum_{p=0}^{\infty} \frac{|\alpha|^{2p}}{\sqrt{2p!}} (e^{-j\pi/4})^{2p} |2p\rangle \quad (\text{C.5a})$$

$$|\phi_1\rangle := \frac{1}{\sqrt{\sinh|\alpha|^2}} \sum_{p=0}^{\infty} \frac{|\alpha|^{(2p+1)}}{\sqrt{(2p+1)!}} (e^{-j\pi/4})^{2p+1} |2p+1\rangle \quad (\text{C.5b})$$

The hyperbolic prefactors ensure normalization

$$\begin{aligned} \langle \phi_0 | \phi_0 \rangle &= \frac{1}{\cosh |\alpha|^2} \sum_{p=0}^{\infty} \frac{|\alpha|^{2p}}{\sqrt{(2p)!}} (e^{j\pi/4})^{2p} \sum_{q=0}^{\infty} \frac{|\alpha|^{2q}}{\sqrt{(2q)!}} (e^{-j\pi/4})^{2q} \underbrace{\langle 2p | 2q \rangle}_{\delta_{p,q}} \\ &= \frac{1}{\cosh |\alpha|^2} \sum_{p=0}^{\infty} \frac{|\alpha|^{4p}}{(2p)!} = \frac{1}{\cosh |\alpha|^2} \underbrace{\sum_{p=0}^{\infty} \frac{(|\alpha|^2)^{2p}}{(2p)!}}_{\cosh |\alpha|^2} = 1 \quad (\text{C.6}) \end{aligned}$$

Going along the similar direction, $\langle \phi_1 | \phi_1 \rangle = 1$. Setting $\lambda_0 = e^{-|\alpha|^2} \cosh |\alpha|^2$ and $\lambda_1 = e^{-|\alpha|^2} \sinh |\alpha|^2$, one can write the diagonal form of ρ_B

$$\rho_B = \lambda_0 |\phi_0\rangle \langle \phi_0| + \lambda_1 |\phi_1\rangle \langle \phi_1| \quad (\text{C.7})$$

The security analysis, like the Gaussian modulation case, is easier in the entanglement based version of the protocol. In this version, Alice starts with a bipartite pure state $|\Phi_2\rangle$. She performs a projective measurement on one half of the state, and sends the other half to Bob. Depending on the binary result of her measurement, the state sent to Bob is either $|\alpha e^{-j\pi/4}\rangle$ or $|- \alpha e^{-j\pi/4}\rangle$ with equal probabilities. The purification of Bob's state can be found out by following the Schmidt decomposition prescription. Since we already know the spectral decomposition of the reduced density ρ_B we can propose the following purification for it

$$|\Phi_2\rangle = \sqrt{\lambda_0} |\phi_0^*\rangle |\phi_0\rangle + \sqrt{\lambda_1} |\phi_1^*\rangle |\phi_1\rangle \quad (\text{C.8})$$

where

$$|\phi_0^*\rangle := \frac{1}{\sqrt{\cosh |\alpha|^2}} \sum_{p=0}^{\infty} \frac{|\alpha|^{2p}}{\sqrt{2p!}} (e^{j\pi/4})^{2p} |2p\rangle \quad (\text{C.9a})$$

$$|\phi_1^*\rangle := \frac{1}{\sqrt{\sinh |\alpha|^2}} \sum_{p=0}^{\infty} \frac{|\alpha|^{(2p+1)}}{\sqrt{(2p+1)!}} (e^{j\pi/4})^{2p+1} |2p+1\rangle \quad (\text{C.9b})$$

With proper unitary transformation $\hat{U}_A \otimes \hat{U}_B$ performed on the bases of the subsystems, one can write $|\Phi_2\rangle$ as

$$|\Phi_2\rangle = \frac{1}{\sqrt{2}} |\psi_0\rangle |\beta\rangle + \frac{1}{\sqrt{2}} |\psi_1\rangle |-\beta\rangle \quad (\text{C.10})$$

with

$$|\psi_0\rangle = \frac{1}{\sqrt{2}} (|\phi_0^*\rangle + |\phi_1^*\rangle) \quad (\text{C.11a})$$

$$|\psi_1\rangle = \frac{1}{\sqrt{2}} (|\phi_0^*\rangle - |\phi_1^*\rangle) \quad (\text{C.11b})$$

and

$$|\beta\rangle = \sqrt{\lambda_0}(|\phi_0\rangle + |\phi_1\rangle) \quad (\text{C.12a})$$

$$|-\beta\rangle = \sqrt{\lambda_1}(|\phi_0\rangle - |\phi_1\rangle) \quad (\text{C.12b})$$

Therefore, in the entanglement based version of the protocol, Alice needs to apply the projective measurement $\{|\psi_0\rangle\langle\psi_0|, |\psi_1\rangle\langle\psi_1|\}$ to the first half of the state $|\Phi_2\rangle$ by which she can project the second half on the coherent state $|\beta\rangle$ or the coherent state $|-\beta\rangle$ with equal probability.

We can now proceed to calculate the covariance matrix Σ_2 of the state $|\Phi_2\rangle$. The matrix elements are

$$\Sigma_{2\text{-PSK}} = \begin{pmatrix} V(\hat{q}_A) & \text{Cov}(\hat{q}_A, \hat{p}_A) & \text{Cov}(\hat{q}_A, \hat{q}_B) & \text{Cov}(\hat{q}_A, \hat{p}_B) \\ \text{Cov}(\hat{p}_A, \hat{q}_A) & V(\hat{p}_A) & \text{Cov}(\hat{p}_A, \hat{q}_B) & \text{Cov}(\hat{p}_A, \hat{p}_B) \\ \text{Cov}(\hat{q}_B, \hat{q}_A) & \text{Cov}(\hat{q}_B, \hat{p}_A) & V(\hat{q}_B) & \text{Cov}(\hat{q}_B, \hat{p}_B) \\ \text{Cov}(\hat{p}_B, \hat{q}_A) & \text{Cov}(\hat{p}_B, \hat{q}_B) & \text{Cov}(\hat{p}_B, \hat{q}_B) & V(\hat{p}_B) \end{pmatrix} \quad (\text{C.13})$$

From the symmetry arguments, and the admissibility condition of a covariance matrix $\Sigma_{2\text{-PSK}}$ should have the following form

$$\Sigma_{2\text{-PSK}} = \begin{pmatrix} X\mathbb{1}_2 & Z\sigma_z \\ Z\sigma_z & Y\mathbb{1}_2 \end{pmatrix} \quad (\text{C.14})$$

To calculate the matrix elements, first we note from equation (C.8)

$$\begin{aligned} \rho_{AB} &= |\Phi_2\rangle\langle\Phi_2| \\ &= \lambda_0 |\phi_0^*\phi_0\rangle\langle\phi_0^*\phi_0| + \sqrt{\lambda_0\lambda_1} |\phi_0^*\phi_0\rangle\langle\phi_1^*\phi_1| + \sqrt{\lambda_0\lambda_1} |\phi_1^*\phi_1\rangle\langle\phi_0^*\phi_0| + \lambda_1 |\phi_1^*\phi_1\rangle\langle\phi_1^*\phi_1| \end{aligned} \quad (\text{C.15})$$

where

$$|\phi_0^*\phi_0\rangle = \frac{1}{\cosh|\alpha|^2} \sum_{p,q=0}^{\infty} (e^{j\pi/4})^{2(p-q)} \frac{|\alpha|^{2(p+q)}}{\sqrt{(2p)!}\sqrt{(2q)!}} |2p, 2q\rangle \quad (\text{C.16a})$$

$$|\phi_1^*\phi_1\rangle = \frac{1}{\sinh|\alpha|^2} \sum_{p,q=0}^{\infty} (e^{j\pi/4})^{2(p-q)} \frac{|\alpha|^{2(p+q+1)}}{\sqrt{(2p+1)!}\sqrt{(2q+1)!}} |2p+1, 2q+1\rangle \quad (\text{C.16b})$$

With the orthogonality of the Fock states $\langle i, j | k, l \rangle = \delta_{i,j;kl}$ it is easy to show

$$\begin{aligned} \langle\Phi_2| \hat{a}_*^\dagger |\Phi_2\rangle &= \langle\Phi_2| \hat{a}_* |\Phi_2\rangle = \langle\Phi_2| \hat{a}_*^2 |\Phi_2\rangle = \langle\Phi_2| (\hat{a}_*^\dagger)^2 |\Phi_2\rangle \\ &= \langle\Phi_2| \hat{a}_A^\dagger \hat{a}_B |\Phi_2\rangle = \langle\Phi_2| \hat{a}_A \hat{a}_B^\dagger |\Phi_2\rangle = 0 \end{aligned} \quad (\text{C.17})$$

(* stands for A, B). The above relations facilitate the calculation of means and variances of the quadrature operators. For example,

$$\begin{aligned}\langle \hat{q}_* \rangle &= \langle \Phi_2 | \hat{q}_* | \Phi_2 \rangle = \langle \Phi_2 | \hat{a}_* | \Phi_2 \rangle + \langle \Phi_2 | \hat{a}_*^\dagger | \Phi_2 \rangle = 0 \\ \langle \hat{p}_* \rangle &= \langle \Phi_2 | \hat{p}_* | \Phi_2 \rangle = -j \left(\langle \Phi_2 | \hat{a}_* | \Phi_2 \rangle - \langle \Phi_2 | \hat{a}_*^\dagger | \Phi_2 \rangle \right) = 0\end{aligned}\quad (\text{C.18})$$

We are now in a position to calculate the elements in the covariance matrix $\Sigma_{2\text{-PSK}}$. The diagonal elements $X = V(\hat{q}_A) = V(\hat{p}_A)$ are given by

$$V(\hat{q}_A) = \langle \hat{q}_A^2 \rangle - \langle \hat{q}_A \rangle^2 = \langle (\hat{a}_A + \hat{a}_A^\dagger)^2 \rangle = 1 + 2|\alpha|^2 = V_A + 1 \quad (\text{C.19})$$

From symmetry, $\Sigma_{22} = \Sigma_{33} = \Sigma_{44} = 1 + 2|\alpha|^2$. Subsequently, we can calculate the off-diagonal elements.

$$\begin{aligned}\text{Cov}(\hat{q}_A, \hat{q}_B) &= \frac{1}{2} (\langle \hat{q}_A \hat{q}_B \rangle + \langle \hat{q}_B \hat{q}_A \rangle) - \langle \hat{q}_A \rangle \langle \hat{q}_B \rangle = \langle \hat{q}_A \hat{q}_B \rangle \\ &= \langle \Phi_2 | \hat{a}_A^\dagger \hat{a}_B^\dagger | \Phi_2 \rangle + \underbrace{\langle \Phi_2 | \hat{a}_A^\dagger \hat{a}_B | \Phi_2 \rangle}_{=0} + \underbrace{\langle \Phi_2 | \hat{a}_A \hat{a}_B^\dagger | \Phi_2 \rangle}_{=0} + \langle \Phi_2 | \hat{a}_A \hat{a}_B | \Phi_2 \rangle\end{aligned}\quad (\text{C.20})$$

The terms which contribute evaluate to

$$\begin{aligned}& \sqrt{\lambda_0 \lambda_1} \langle \phi_0 \phi_0^* | (\hat{a}_A \hat{a}_B + \hat{a}_A^\dagger \hat{a}_B^\dagger) | \phi_1^* \phi_1 \rangle + \sqrt{\lambda_0 \lambda_1} \langle \phi_1 \phi_1^* | (\hat{a}_A \hat{a}_B + \hat{a}_A^\dagger \hat{a}_B^\dagger) | \phi_0^* \phi_0 \rangle \\ &= 2\Lambda |\alpha|^2 \cosh 2|\alpha|^2 = \frac{2e^{-|\alpha|^2} \cosh 2|\alpha|^2}{\sqrt{\cosh |\alpha|^2 \sinh |\alpha|^2}} = 2|\alpha|^2 \frac{1 + e^{-4|\alpha|^2}}{\sqrt{1 - e^{-4|\alpha|^2}}} = V_A \frac{1 + e^{2V_A}}{\sqrt{1 - e^{-2V_A}}}\end{aligned}\quad (\text{C.21})$$

Putting all the terms together in the matrix in (C.14)

$$\Sigma_{2\text{-PSK}} = \begin{pmatrix} (V_A + 1)\mathbb{1}_2 & V_A \frac{1 + e^{2V_A}}{\sqrt{1 - e^{-2V_A}}} \sigma_z \\ V_A \frac{1 + e^{2V_A}}{\sqrt{1 - e^{-2V_A}}} \sigma_z & (V_A + 1)\mathbb{1}_2 \end{pmatrix} \quad (\text{C.22})$$

Appendix D

Complementary to Chapter 3

D.1 Working Principle of a 2×2 3-dB MMI

The following derivation follows along [106]. Let us consider an optical field input to a 2×2 MMI at $z = 0$, which excites a host of eigenmodes in the multimode region.

The overall field distribution, $f(x)$ can be written as a linear combination of modal field distributions

$$f(x, 0) = \sum_i c_i \psi_i(x) \quad (\text{D.1})$$

where the summation is over all the guided and excited modes, and c_i 's are weights for modal excitations. Considering $f(x)$ to be tightly confined within the waveguide region, the summation can be taken only over the guided modes

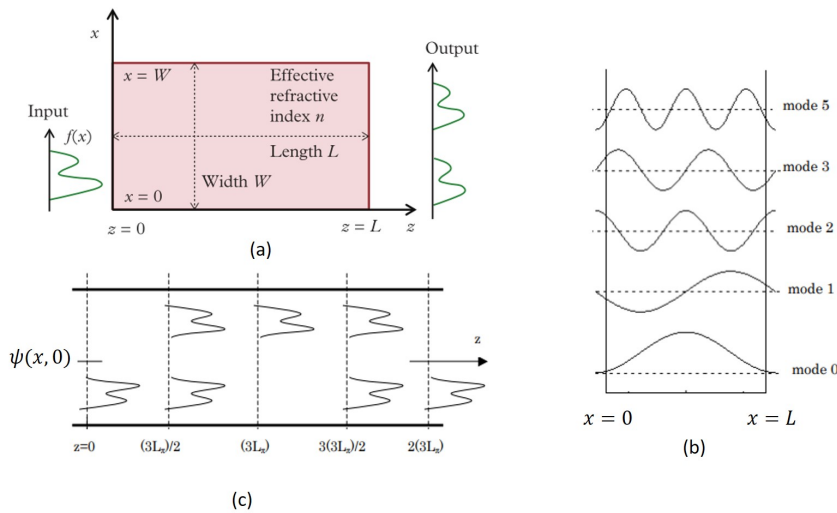


FIGURE D.1: (a) General layout of a MMI coupler implemented by means of a 2D slab waveguide (access waveguides are not shown) [9]. (b) Supported modes in the multimode region [106]. (c) Multimode waveguide showing the input field $\psi(x, 0)$, a mirrored single image at $3L_\pi$, a direct single image at $2(3L_\pi)$, and two-fold images at $(3L_\pi)/2$ and $3 \cdot (3L_\pi)/2$ [106].

$$f(x, 0) = \sum_{i=0}^{m-1} c_i \psi_i(x), \quad (\text{D.2})$$

where m is the total number of guided modes. The eigenmodes are independent and can be written as

$$\psi_i(x) = \sin \left[\frac{(i+1)\pi x}{W} \right], \quad (\text{D.3})$$

reflecting their standard wave field patterns (Figure (D.1b)). For even index i , the eigenmode is even, and for odd index i , the eigenmode is odd. Therefore one can write equation (D.1) as the sum for even and odd parts.

$$f(x, 0) = \sum_{i \text{ even}} c_i \psi_i(x) + \sum_{i \text{ odd}} c_i \psi_i(x), \quad (\text{D.4})$$

with

$$\sum_{i \text{ even}} c_i \psi_i(x) = \frac{f(x, 0) + f(-x, 0)}{2}, \quad (\text{D.5})$$

$$\sum_{i \text{ odd}} c_i \psi_i(x) = \frac{f(x, 0) - f(-x, 0)}{2}. \quad (\text{D.6})$$

From equation (D.3), the transverse wavenumber is

$$k_{x,i} = \frac{(i+1)\pi}{W},$$

where W is the width of the multimode region. From the dispersion relation, one can have the propagation constant for mode, i as

$$\beta_i^2 = n^2 k_0^2 - k_{x,i}^2 \Rightarrow \beta_i = nk_0 \sqrt{1 - \left(\frac{k_{x,i}}{nk_0} \right)^2} \approx nk_0 - \frac{k_{x,i}^2}{2nk_0} \quad (\text{D.7})$$

with $k_0 = \frac{2\pi}{\lambda_0}$ is the free space wave number and n is the effective index of the waveguide structure. In the last line we have made use of binomial approximation. Combining equations (D.7) and (??)

$$\beta_i = nk_0 - \frac{(i+1)^2 \pi^2}{2nk_0 W^2}, \text{ with } \beta_0 = nk_0 - \frac{\pi^2}{2nk_0 W^2} \quad (\text{D.8})$$

Therefore the spacing between propagation constant for mode i , and mode 0 is

$$\beta_0 - \beta_i = \frac{\pi^2 i(i+2)}{2nk_0 W^2} \quad (\text{C.1.9})$$

By definition the beat length of the two lowest order modes is

$$L_\pi = \frac{\pi}{\beta_0 - \beta_1} = \frac{2W^2 nk_0}{3\pi} \quad (\text{D.9})$$

which allows us to write equation (C.1.9) as

$$\beta_0 - \beta_i = \frac{i(i+2)\pi}{3L_\pi} \quad (\text{D.10})$$

The field profile at a distance, z can be written as the superposition of all the guided mode field distributions

$$f(x, z) = \sum_{i=0}^{m-1} c_i \psi_i(x) \exp\{-j\beta_i z\} \quad (\text{D.11})$$

with implicit time harmonic behavior. Rewriting differently

$$f(x, z) = \sum_{i=0}^{m-1} c_i \psi_i(x) \exp\{j(\beta_0 - \beta_i z)\} \quad (\text{D.12})$$

where the $\exp\{-j\beta_0 z\}$ term has been dropped. At $z=L$ the field distribution has the form

$$f(x, L) = \sum_{i=0}^{m-1} c_i \psi_i(x) \exp\left\{j \frac{i(i+2)\pi}{3L_\pi} L\right\} \quad (\text{D.13})$$

Under certain circumstances, the field $f(x, L)$ will be a reproduction (self-imaging) of the input field $f(x, 0)$. The shape and the types of these images formed will be determined by the coefficients, c_i 's and the modal phase factor $\exp\left\{j \frac{i(i+2)\pi}{3L_\pi} L\right\}$. In a generic scenario, the image formation will not be effected by the mode amplitudes. In this case $f(x, L)$ will be an image of $f(x, 0)$ if

$$\exp\left\{j \frac{i(i+2)\pi}{3L_\pi} L\right\} = 1 \text{ or } (-1)^i \quad (\text{D.14})$$

In the first case, all the guided modes have to interfere with the same relative phases as in $z = 0$. The interference produces an image which is a direct replica of the input field. In the second case, the guided modes have to interfere alternatively in even and odd multiples of π , for example, the even modes are all in-phase, and the odd modes are all anti-phase. The interference produces an image mirrored with respect to the plane $x = 0$. Both these conditions essentially mean the cumulative phase factors should be integer multiples of π which is fulfilled when

$$L = p(3L_\pi) \text{ with } p = 0, 1, 2, \dots \quad (\text{D.15})$$

Therefore, direct and mirrored single images of the input field $f(x, 0)$ will therefore be formed at distances which are respectively even and odd multiples of the length $3L_\pi$ along z (Figure (D.1c)). This fact can be exploited to function the MMI as a crossbar coupler.

Now at $L = p(\frac{3L_\pi}{2})$, with $p=1,3,5\dots$ the total field at these lengths is

$$f(x, \frac{p}{2}3L_\pi) = \sum_{i=0}^{m-1} c_i \psi_i(x) \exp\left\{j \frac{i(i+2)p\pi}{2}\right\} \quad (\text{D.16})$$

From the mode field symmetric condition the overall field distribution can be decomposed as

$$\begin{aligned}
 & f(x, \frac{p}{2}3L\pi) \\
 &= \sum_{i \text{ even}} c_i \psi_i(x) + \sum_{i \text{ odd}} (-j)^p c_i \psi_i(x) \\
 &= \frac{1}{2} \sum_{i \text{ even}} c_i \psi_i(x) + \frac{1}{2} \sum_{i \text{ even}} c_i \psi_i(x) + \frac{1}{2} \sum_{i \text{ odd}} (-j)^p c_i \psi_i(x) + \frac{1}{2} \sum_{i \text{ odd}} (-j)^p c_i \psi_i(x) \\
 &+ \frac{1}{2} \sum_{i \text{ odd}} c_i \psi_i(x) - \frac{1}{2} \sum_{i \text{ odd}} c_i \psi_i(x) - \frac{1}{2} \sum_{i \text{ even}} (-j)^p c_i \psi_i(x) + \frac{1}{2} \sum_{i \text{ even}} (-j)^p c_i \psi_i(x) \\
 &= \frac{1}{2} \left(\sum_{i \text{ even}} c_i \psi_i(x) + \sum_{i \text{ odd}} c_i \psi_i(x) \right) + \frac{1}{2} \left(\sum_{i \text{ even}} c_i \psi_i(x) - \sum_{i \text{ odd}} c_i \psi_i(x) \right) \\
 &+ \frac{1}{2} (-j)^p \left(\sum_{i \text{ even}} c_i \psi_i(x) + \sum_{i \text{ odd}} c_i \psi_i(x) \right) + \frac{1}{2} (-j)^p \left(\sum_{i \text{ odd}} c_i \psi_i(x) - \sum_{i \text{ even}} c_i \psi_i(x) \right) \\
 &= \frac{1}{2} \psi(x, 0) + \frac{1}{2} \psi(-x, 0) + \frac{1}{2} (-j)^p \psi(x, 0) - \frac{1}{2} (-j)^p \psi(-x, 0) \\
 &= \frac{1 + (-j)^p}{2} \psi(x, 0) + \frac{1 - (-j)^p}{2} \psi(-x, 0) \tag{D.17}
 \end{aligned}$$

The last expression suggests the field distribution $f(x, 0)$ reproduces itself into a pair of images, at distances $z = \frac{3L\pi}{2}, 3\frac{3L\pi}{2}, 5\frac{3L\pi}{2}, \dots$, which are reduced in amplitudes by $\frac{1}{\sqrt{2}}$ and oscillate in quadrature (Figure (D.1c)). The integrated 90° 3-dB 2×2 MMI achieves the exact same task. To minimize the footprint of the device usually the smallest possible length is chosen.

D.2 Derivation of equation (3.21)

Referring to Figure (D.2), let us calculate the bar and cross outputs of a Mach Zehnder modulator with respect to the input E_{in} . From equation (3.8),

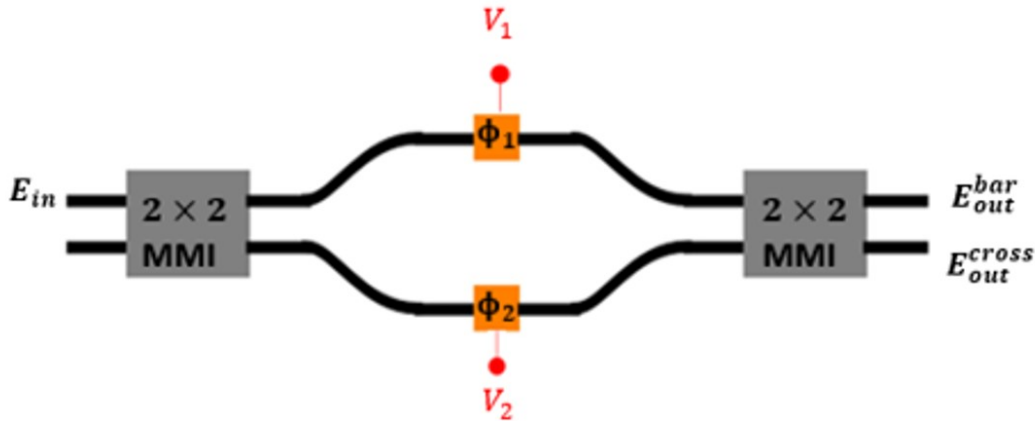


FIGURE D.2: Diagrammatic view of a Mach Zehnder Modulator with only one input.

$$\begin{aligned}
T_{\text{MZM}} \mathbf{E} &= T_{\text{MMI}_1} T_{\text{PM}} T_{\text{MMI}_2} \mathbf{E} \\
&= \frac{j}{\sqrt{2}} \begin{pmatrix} 1 & j \\ j & 1 \end{pmatrix} \begin{pmatrix} e^{j\phi_1} & 0 \\ 0 & e^{j\phi_2} \end{pmatrix} \frac{j}{\sqrt{2}} \begin{pmatrix} 1 & j \\ j & 1 \end{pmatrix} \begin{pmatrix} E_{\text{in}} \\ 0 \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} 1 & j \\ j & 1 \end{pmatrix} \begin{pmatrix} e^{j\phi_1} & 0 \\ 0 & e^{j\phi_2} \end{pmatrix} \begin{pmatrix} E_{\text{in}} \\ jE_{\text{in}} \end{pmatrix} \\
&= -\frac{1}{2} \begin{pmatrix} 1 & j \\ j & 1 \end{pmatrix} \begin{pmatrix} E_{\text{in}} e^{j\phi_1} \\ jE_{\text{in}} e^{j\phi_2} \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} E_{\text{in}} e^{j\phi_1} - E_{\text{in}} e^{j\phi_2} \\ jE_{\text{in}} e^{j\phi_1} + jE_{\text{in}} e^{j\phi_2} \end{pmatrix}. \tag{D.18}
\end{aligned}$$

The cross output, therefore, is

$$E_{\text{out}}^{\text{cross}} = -\frac{j}{2} (E_{\text{in}} e^{j\phi_1} + E_{\text{in}} e^{j\phi_2}) = \frac{E_{\text{in}}}{2} (e^{j(\phi_1 - \frac{\pi}{2})} + e^{j(\phi_2 - \frac{\pi}{2})}) = \frac{E_{\text{in}}}{2} (e^{j(\phi_1 - \frac{\pi}{2})} - e^{j(\phi_2 + \frac{\pi}{2})}). \tag{D.19}$$

D.3 Gaussian Modulation

Gaussian modulation is one form of quadrature amplitude modulation (QAM) where the field quadratures take values from a continuous set. It is possible to realize Gaussian modulation of the signal pulse by using an amplitude modulator and a phase modulator sequentially. The theory follows from the so called *Box Muller Method*. With this method, one can produce a bi variate Gaussian distribution by sampling from two uniform random variables following a coordinate transformation. The marginal distributions are independent of each other. In the Gaussian modulation version of CV-QKD the field quadratures are sampled from two normal distributions $P \sim \mathcal{N}(0, V)$, $Q \sim \mathcal{N}(0, V)$. The key point to note is that the field quadrature observables are related to the field amplitude and phase via the following coordinate transformation (Cartesian to polar).

$$P = E \cos \theta \tag{D.20}$$

$$Q = E \sin \theta \tag{D.21}$$

Let X and Y be independent standard normal random variables $X, Y \sim (0, 1)$ and X and Y are independent $X \perp Y$. The joint probability density function of X and Y is given by

$$\begin{aligned}
f_{XY}(x, y) &= f_X(x) f_Y(y) \\
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \\
&= \underbrace{\frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}}_{\text{target density}} \tag{D.22}
\end{aligned}$$

With $x = r \cos \theta$ and $y = r \sin \theta$, it is possible to have $f_{R\Theta}(r, \theta)$ from $f_{XY}(x, y)$. From multi-variable calculus, the change of variables is given by

$$\begin{aligned} f_{R\Theta}(r, \theta) &= f_{XY}(x, y) \frac{dx dy}{dr d\theta} \\ &= f_{XY}(x, y) \underbrace{\left| \frac{\partial(x, y)}{\partial(r, \theta)} \right|}_{=J} \end{aligned} \quad (\text{D.23})$$

where J is the Jacobian.

$$J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r \cos^2 \theta + r \sin^2 \theta = r \quad (\text{D.24})$$

Therefore for $r \geq 0$ and $\theta \in [0, 2\pi)$, change of variable leads to

$$f_{R\Theta}(r, \theta) dr d\theta = \frac{1}{2\pi} e^{-\frac{r^2}{2}} r dr d\theta \quad (\text{D.25})$$

Applying another change of variables from (r, θ) to (r^2, θ) one has

$$dr^2 = 2r dr \Rightarrow r dr = \frac{1}{2} dr^2 \quad (\text{D.26})$$

The joint PDF $f_{R\Theta}(r, \theta) dr d\theta$, for $r \geq 0$ and $\theta \in [0, 2\pi)$ can be written as

$$\begin{aligned} f_{R\Theta}(r, \theta) dr d\theta &= f_{R^2\Theta}(r^2, \theta) dr^2 d\theta \\ \Rightarrow \frac{1}{2\pi} e^{-\frac{r^2}{2}} r dr d\theta &= \frac{1}{2\pi} e^{-\frac{r^2}{2}} dr^2 d\theta \\ &= \underbrace{\left(\frac{1}{2} e^{-\frac{r^2}{2}} dr^2 \right)}_{\text{exponential dist.}} \underbrace{\left(\frac{1}{2\pi} d\theta \right)}_{\text{uniform dist.}} \\ &= f_{R^2}(r^2) dr^2 f_{\Theta}(\theta) d\theta \end{aligned} \quad (\text{D.27})$$

which shows $R^2 \perp \Theta$. Therefore the recipe for generating two independent normal distributions can be listed as follows:

1. generate $\Theta \sim \text{unif}(0, 2\pi)$
2. generate $V \sim \exp\left\{\lambda = \frac{1}{2}\right\}$ i.e. $V = R^2$, and compute $R = \sqrt{V}$
3. compute $X = R \cos \theta, Y = R \sin \theta$

Then $X, Y \sim \mathcal{N}(0, 1)$. Incidentally, one can see the cumulative distribution function of the PDF $f(x) = \frac{1}{2}e^{-\frac{1}{2}x}$ is

$$\Pr(X < x) = F(x) = \int_0^x \frac{1}{2} e^{-\frac{1}{2}t} dt = 1 - \frac{1}{2} e^{-\frac{1}{2}x} \quad (\text{D.28})$$

Setting $u = 1 - \frac{1}{2}e^{-\frac{1}{2}x}$ leads to $X = F^{-1}(U) = -2\log(1 - U)$ where $U \sim \text{unif}(0, 1)$. This is known as the inverse CDF method which makes use of the fact that a continuous CDF, $F(x)$ is a one-to-one mapping of the domain of the CDF into the interval $[0, 1]$. Therefore if U is a uniform random variable on $[0, 1]$ the $X = F^{-1}(U)$ has the distribution of $F(x)$. Since $U \sim \text{unif}[0, 1]$, so is $1 - U \sim \text{unif}[0, 1]$. Therefore one can write $X = -2\log(U)$. Putting things together

$$\begin{aligned} R^2 &= (-2\log U_1)^{\frac{1}{2}} \\ \Theta &= 2\pi U_2 \end{aligned} \quad (\text{D.29})$$

Finally the way to generate (or simulate) two independent normal random variables X and Y from two uniform distributions $U_1 \sim [0, 1]$, $U_2 \sim [0, 1]$ according to the following formulae

$$\begin{aligned} X &= \sqrt{-2\log U_1} \cos(2\pi U_2) \\ Y &= \sqrt{-2\log U_1} \sin(2\pi U_2) \end{aligned} \quad (\text{D.30})$$

as $X = R \cos \theta$ and $Y = R \sin \theta$.

D.4 Half Wave Plate

In bulk optics, the effect of a half-wave plate is realized with a birefringent crystal. A birefringent crystal shows anisotropy in the real part of the refractive index, n_r . A uniaxial birefringent crystal can be described by two indices n_1 and n_2 along the principal optical axes of the crystal. If $n_1 > n_2$, then light will travel at a faster speed along n_2 compared to n_1 . Hence we can relabel n_1, n_2 as n_s, n_f with f and s standing for fast and slow axis respectively. Since an anisotropic crystal allows optical oscillation only along the crystal optical axes [85], therefore the incoming polarization is first decomposed into components along those two axes. Following a propagation through a distance L , the optical transfer matrix of the optical system can be written as

$$\begin{pmatrix} e^{-jk_0 n_f L} & 0 \\ 0 & e^{-jk_0 n_s L} \end{pmatrix} = e^{-jk_0 n_f L} \begin{pmatrix} 1 & 0 \\ 0 & e^{-jk_0 \Delta n L} \end{pmatrix} \quad (\text{D.31})$$

where $\Delta n = n_s - n_f$ is defined to be the birefringence of the material. The above matrix, called the Jones matrix, is for a general wave plate. By setting $L = \frac{\lambda_0}{2\Delta n}$, which makes $e^{-jk_0 \Delta n L} = e^{-j\pi} = -1$, we get the Jones matrix for a half wave plate.

$$J_{\text{HWP}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (\text{D.32})$$

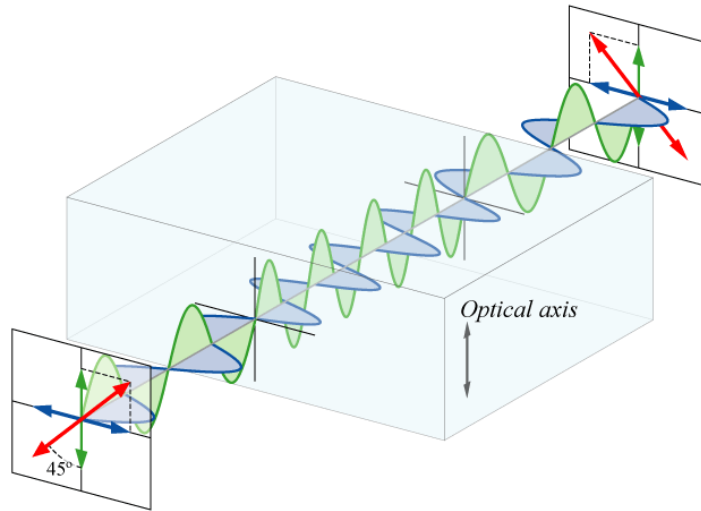


FIGURE D.3: The green arrow and the blue arrow indicate electric field parallel and perpendicular to the optic axis respectively. The red arrow is the resultant field. Linearly polarized light entering a half-wave plate can be resolved into two waves, parallel and perpendicular to the optic axis of the waveplate. In the plate, the parallel wave propagates slightly slower than the perpendicular one. At the far side of the plate, the parallel wave is exactly half of a wavelength delayed relative to the perpendicular wave, and the resulting combination is a mirror-image of the entry polarization state (relative to the optic axis)[107].

where we have ignored the common phase term.

While writing the above matrix we have assumed the fast axis is along $\hat{x}$, and the slow axis is along $\hat{y}$. If the crystal is rotated at 45° with respect to $\hat{x}$, then in the new coordinate system, the Jones matrix for the HWP is

$$J_{\text{HWP}}(45^\circ) = R(-45^\circ)J_{\text{HWP}}R(45^\circ) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (\text{D.33})$$

where $R(\theta)$ is the 2D rotation matrix. Therefore it transforms $\hat{x}$ polarized light to $\hat{y}$ polarized light. Jones formalism applies to plane waves. But, to a good approximation, it can be adapted to the case of guided modes as well. The InGaAsP/InP crystal has a cubic lattice which is optically isotropic. Nonetheless the guided modes can be made to experience birefringence by tweaking the waveguide geometry. This birefringence, therefore, is called waveguide (or geometric) birefringence to differentiate from the material birefringence [85].

D.5 8-PSK with MZM:

Figure (D.4) shows the circuit level schematic for 8-PSK modulation. The series of MZMs along with their electronic driver circuits have been bordered with a dashed box. The first MZM is biased at $\frac{I_\pi}{2}$ with magnitude I_π to provide

phase shifts of $\{0, \pi\}$. The second MZM is biased at $\frac{I_\pi}{2}$ with magnitude $\frac{I_\pi}{2}$ to provide phase shifts of $\{0, \frac{\pi}{2}\}$. The third MZM is biased at $\frac{I_\pi}{2}$ with magnitude $\frac{I_\pi}{4}$ to provide phase shifts of $\{0, \frac{\pi}{4}\}$. The overall output is an 8-PSK modulated signal.

Figure (D.5) shows the relevant photocurrents. For the constellation points shown in Figure (3.14), P and Q quadratures may have five possible values, for example

$$P : \{-P, -P/\sqrt{2}, 0, P/\sqrt{2}, P\},$$

$$Q : \{-Q, -Q/\sqrt{2}, 0, Q/\sqrt{2}, Q\}.$$

The results from the eye diagram analyzers for both the quadratures are shown in Figure (D.6). Figure (D.7) shows the measurement statistic from the VSA. The detected signals (in red) are generally within the decision boundaries (in blue) in the symbol map. The EVM and SER have been found to be 25.35% and

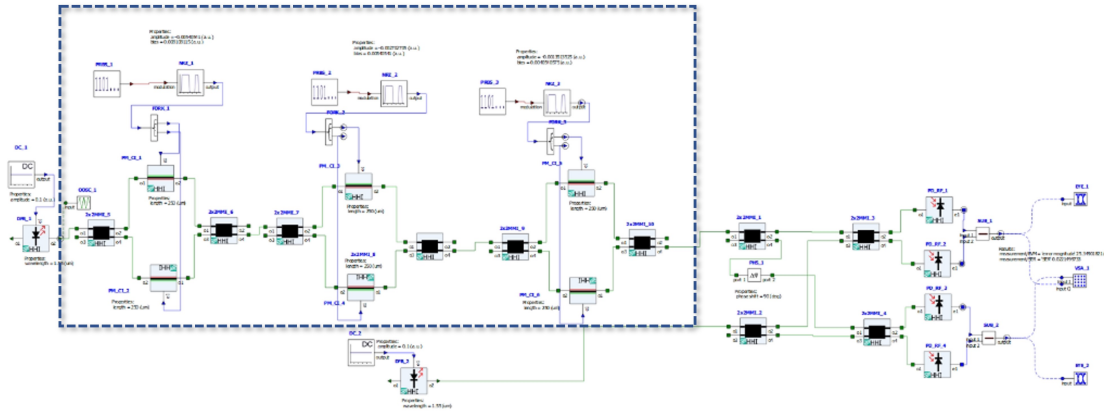


FIGURE D.4: Schematic illustrating implementation of 8 phase shift keying with a single MZM in dual drive operation.

0.02215 respectively. The vector diagram indicates the quadrature and in-phase components mapped to the vertical and horizontal directions.

D.6 Eye Diagram

An eye diagram is measured in the time domain and variation in the measurements will occur with every set of samples. In Figure (D.8), on the left, one can see various waveforms that can be transmitted through the transmission medium in a high speed signalling environment and measured by an oscilloscope. The oscilloscopes will capture these waveforms and overlay them on top of each other each sample period to create an eye diagram. This variation is captured over multiple unit intervals (UIs). An UI, also commonly known as the pulse time, is the amount of time required for two consecutive transitions to occur. By collecting many such different bit transitions, large amounts of information about electrical signal and how it propagates through the transmission

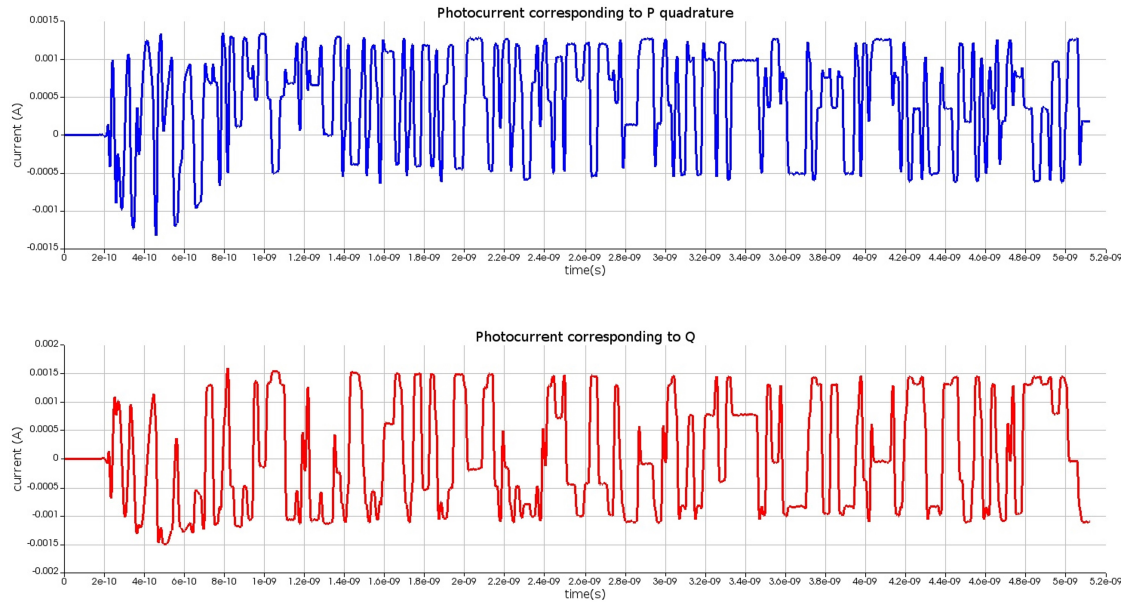


FIGURE D.5: Photocurrents corresponding to the in-phase and quadrature-phase quadratures of the modulated signal in 8PSK.

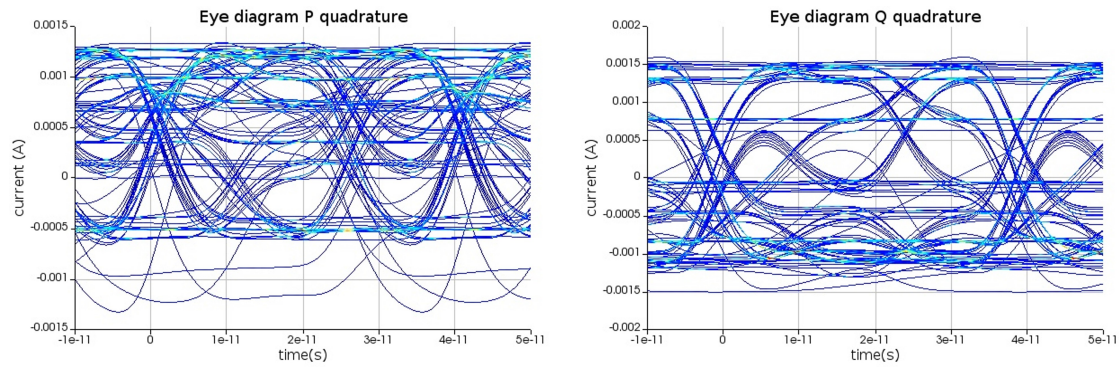


FIGURE D.6: Eye diagram results for P and Q quadrature in 8PSK.

EVM 25.35%; SER 0.02215

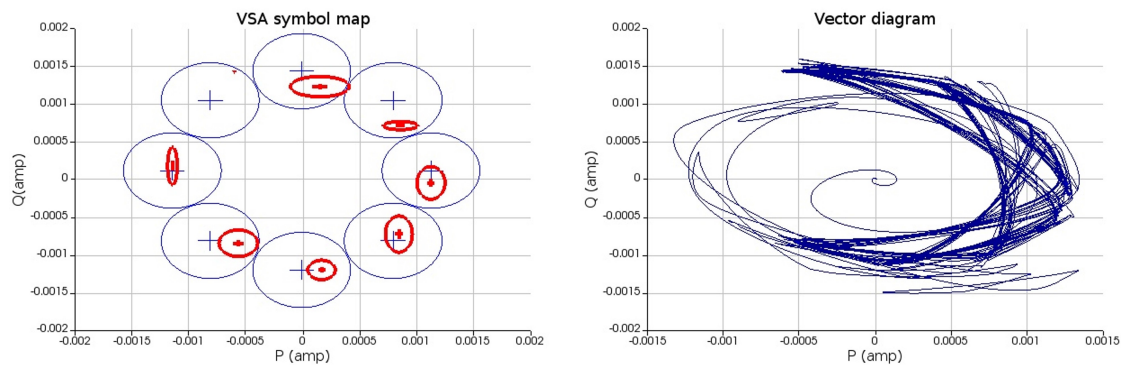


FIGURE D.7: The signal statistic with respect to the decision boundaries in 8PSK.

media can be collected. Eye diagrams should be taken over many samples to ensure the diagram truly represents system performance. Pseudo random bit

sequence patterns are commonly used to generate eye diagrams.

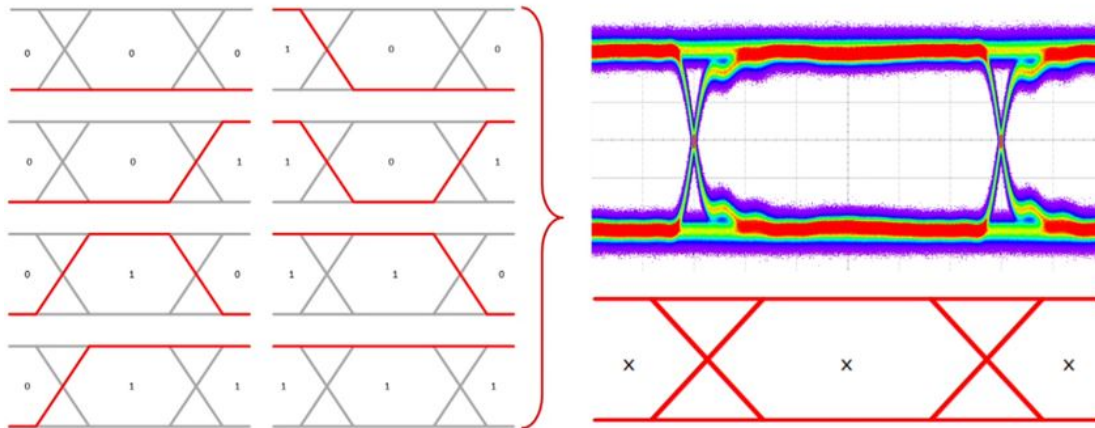


FIGURE D.8: Construction of an eye diagram. [108]

Figure (D.9) shows the anatomy of an eye diagram. V_H , V_L are the two fundamental components in an eye diagram. These two features will indicate if the receiver will see the expected voltage levels corresponding to the correct bit. V_H corresponds to 1, and V_L corresponds to 0. These two voltage levels are typically described in the electrical specification of the transmitter. The three lines shown in green represent a bit transition. Because the bit transition occurs on the left of side of the eye diagram, it is commonly called the leading edge.

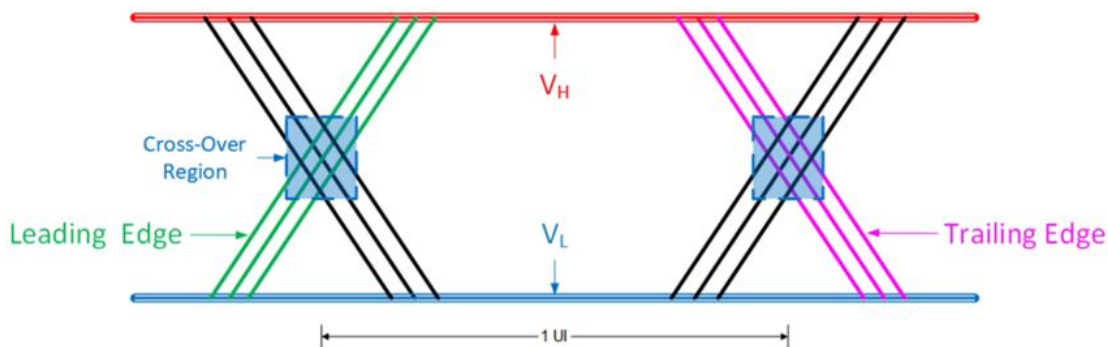


FIGURE D.9: Anatomy of an eye diagram [108]

The bit transitions which occur on the right side are called the trailing edge. These bit transitions contain high frequency content of the signal. When analyzing the eye diagram, the leading and the trailing edge will show the effects of different types of equalization as well as the effects of different types of transmission media loading. The crossover region is another important part of the eye diagram. The cross over regions can show variation in the bit transitions commonly known as jitter.

Appendix E

Complementary to Chapter 4

A. Derivation of equation (4.23)

Plugging equation (4.21) into (4.13), one has

$$\begin{aligned} & \frac{\partial^2}{\partial x^2} \left(\sum_{m=-\infty}^{\infty} E_y^m \exp(j\beta_m z) \right) + \frac{\partial^2}{\partial z^2} \left(\sum_{m=-\infty}^{\infty} E_y^m \exp(j\beta_m z) \right) \\ & + k_0^2 \tilde{n}_0^2(x) \left(\sum_{m=-\infty}^{\infty} E_y^m \exp(j\beta_m z) \right) + k_0^2 \left(\sum_{m=-\infty}^{\infty} E_y^m \exp(j\beta_m z) \right) \\ & \left(\sum_{q=-\infty}^{\infty} A_q(x) \exp\left(j \frac{2\pi q z}{\Lambda}\right) \right) = 0 \end{aligned} \quad (\text{E.1})$$

Let us evaluate the left hand side of (E.1) term by term. From the first term

$$\frac{\partial^2}{\partial x^2} \left(\sum_{m=-\infty}^{\infty} E_y^{(m)} \exp(j\beta_m z) \right) = \sum_{m=-\infty}^{\infty} \exp(j\beta_m z) \frac{\partial^2}{\partial x^2} E_y^{(m)} \quad (\text{E.2})$$

Second term

$$\begin{aligned} & \frac{\partial^2}{\partial z^2} \left(\sum_{m=-\infty}^{\infty} E_y^m \exp(j\beta_m z) \right) \\ & = \frac{\partial}{\partial z} \left[\sum_{m=-\infty}^{\infty} \left(\frac{\partial E_y^m}{\partial z} \exp(j\beta_m z) + (j\beta_m) E_y^m \exp(j\beta_m z) \right) \right] \\ & = \sum_{m=-\infty}^{\infty} \left[\frac{\partial^2 E_y^m}{\partial z^2} \exp(j\beta_m z) + (j\beta_m) \frac{\partial E_y^m}{\partial z} \exp(j\beta_m z) \right. \\ & \quad \left. + (j\beta_m)^2 E_y^m \exp(j\beta_m z) + (j\beta_m) \frac{\partial E_y^m}{\partial z} \exp(j\beta_m z) \right] \end{aligned} \quad (\text{E.3})$$

Fourth term

$$\begin{aligned}
& k_0^2 \sum_{m=-\infty}^{\infty} \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} E_y^m A_q(x) \exp\left(j \frac{2\pi q z}{\Lambda}\right) \exp(j\beta_m z) \\
&= k_0^2 \sum_{m=-\infty}^{\infty} \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} E_y^m A_q(x) \exp\left(j \frac{2\pi q z}{\Lambda}\right) \exp\left(j\left(\beta_0 + \frac{2\pi m}{\Lambda}\right)z\right) \\
&= k_0^2 \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} A_q(x) \sum_{m=-\infty}^{\infty} E_y^m \exp\left[j\left(\beta_0 + \frac{2\pi m}{\Lambda} + \frac{2\pi q}{\Lambda}\right)z\right] \\
&= k_0^2 \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} A_q(x) \sum_{m=-\infty}^{\infty} E_y^m \exp[j\beta_{m+q} z] \\
&= k_0^2 \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} A_q(x) \sum_{m=-\infty}^{\infty} E_y^{(m-q)} \exp(j\beta_m z) \tag{E.4}
\end{aligned}$$

Collecting all the terms multiplied by the same factor $e^{j\beta_m z}$ from (E.1)-(E.4), one arrives at equation (4.23).

B. Derivation of equation (4.27) from (4.23)

Substituting $E_y^m(x, z) = E_y^0$ into equation (4.23)

$$\frac{\partial^2 E_y^{(0)}}{\partial x^2} + \frac{\partial^2 E_y^{(0)}}{\partial z^2} + 2j\beta_0 \frac{\partial E_y^{(0)}}{\partial z} - \beta_0^2 E_y^0 + k_0^2 \tilde{n}_0^2(x) E_y^{(0)} = -k_0^2 \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} A_q(x) E_y^{(-q)}(x, z) \tag{E.5}$$

Evaluating term by term yields

$$\text{1st term:} \quad \frac{\partial^2 E_y^{(0)}}{\partial x^2} = \frac{\partial \varepsilon_0(x)}{\partial x^2} R(z) = R(z) \frac{d^2 \varepsilon_0(x)}{dx^2} \tag{E.6}$$

$$\text{2nd term:} \quad \frac{\partial^2 E_y^{(0)}}{\partial z^2} = \frac{\partial \varepsilon_0(x)}{\partial z^2} R(z) = \varepsilon_0(x) \frac{d^2 R(z)}{dz^2} \tag{E.7}$$

$$\text{3rd term:} \quad 2j\beta_0 \frac{\partial E_y^{(0)}}{\partial z} = 2j\beta_0 \varepsilon_0(x) \frac{dR(z)}{dz} \tag{E.8}$$

$$\text{4th term:} \quad -\beta_0^2 E_y^{(0)} = -\beta_0^2 \varepsilon_0(x) R(z) \tag{E.9}$$

$$\text{5th term:} \quad k_0^2 \tilde{n}_0^2 E_y^{(0)} = k_0^2 \tilde{n}_0^2 \varepsilon_0(x) R(z) \tag{E.10}$$

Collecting everything together gives equation (4.28).

C. Deriving equation (4.30)

$$\begin{aligned}
& \frac{d^2 \varepsilon_0(x)}{dx^2} + [k_0^2 \tilde{n}_0^2 - \beta_0^2] \varepsilon_0 \\
&= \frac{d^2 \varepsilon_0(x)}{dx^2} + [k_0^2 n_0^2(x) - 2jk_0 n_0(x) \tilde{a}(x) - \beta_0^2] \varepsilon_0 \\
&= \frac{d^2 \varepsilon_0(x)}{dx^2} + [k_0^2 n_0^2(x) - \beta^2] \varepsilon_0(x) + \beta^2 \varepsilon_0(x) - 2jk_0 n_0(x) \tilde{a}(x) - \beta_0^2 \varepsilon_0(x) \\
&= -2jk_0 n_0(x) \tilde{a}(x) + (\beta^2 - \beta_0^2) \varepsilon_0(x)
\end{aligned}$$

In the second line and the last line we have made use of equations (4.16) and (4.19) respectively.

D. Simplifying equations (4.43) and (4.45)

The last term of equation (4.43)

$$\begin{aligned}
& \frac{k_0^2}{2\beta_0 P} \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} \int_{-\infty}^{\infty} A_q(x) \varepsilon_0 \{ R \varepsilon_{-q}^{(0)}(x) + S \varepsilon_q^{(p)}(x) \} dx \\
&= R \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} \frac{k_0^2}{2\beta_0 P} \int_0^g A_q(x) \varepsilon_0 \varepsilon_{-q}^{(0)}(x) dx + S \frac{k_0^2}{2\beta_0 P} \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} \int_0^g A_q(x) \varepsilon_0 \varepsilon_{-q}^{(p)}(x) dx \\
&= R \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} \eta_{q,-q}^{(0)} + S \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} \eta_{q,-q}^{(p)} = R(z) \zeta_1 + S(z) \zeta_2
\end{aligned}$$

The last term of equation (4.45)

$$\begin{aligned}
& \frac{k_0^2}{2\beta_0 P} \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} \int_{-\infty}^{\infty} A_q(x) \varepsilon_0 \{ R \varepsilon_{(p-q)}^{(0)}(x) + S \varepsilon_{(p-q)}^{(p)}(x) \} dx \\
&= R \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} \frac{k_0^2}{2\beta_0 P} \int_0^g A_q(x) \varepsilon_0 \varepsilon_{p-q}^{(0)}(x) dx + S \frac{k_0^2}{2\beta_0 P} \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} \int_0^g A_q(x) \varepsilon_0 \varepsilon_{p-q}^{(p)}(x) dx \\
&= R \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} \eta_{q,p-q}^{(0)} + S \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} \eta_{q,p-q}^{(p)} = R(z) \zeta_4 + S(z) \zeta_3
\end{aligned}$$

E. Derivation of equations (4.52a)-(4.52d)

Applying (4.51a) into (4.47)

$$\begin{aligned}
& \gamma(r_1 \exp(\gamma z) - r_2 \exp(-\gamma z)) + (-\alpha - j\delta - j\zeta_1)(r_1 \exp(\gamma z) + r_2 \exp(-\gamma z)) \\
& \quad = j(\kappa_p^* + \zeta_2)(s_1 \exp(\gamma z) + s_2 \exp(-\gamma z)) \\
& \Rightarrow \{ \gamma r_1 + (\alpha - j\delta - j\zeta_1) r_1 \} \exp(\gamma z) + \{ -\gamma r_2 + (\alpha - j\delta - j\zeta_1) r_2 \} \exp(-\gamma z) \\
& \quad = \{ j(\kappa_p^* + \zeta_2) s_1 \} \exp(\gamma z) + \{ j(\kappa_p^* + \zeta_2) s_2 \} \exp(-\gamma z)
\end{aligned}$$

Equating terms with the same exponential

$$\begin{aligned}\gamma r_1 + (-\alpha - j\delta - j\zeta_1)r_1 &= j(\kappa_p^* + \zeta_2)s_1 \\ \Rightarrow \{\gamma + (-\alpha - j\delta - j\zeta_1)\}r_1 &= j(\kappa_p^* + \zeta_2)s_1 \\ \Rightarrow \hat{\Gamma}r_1 &= j(\kappa_p^* + \zeta_2)s_1\end{aligned}$$

and

$$\begin{aligned}-\gamma r_2 + (-\alpha - j\delta - j\zeta_1)r_2 &= j(\kappa_p^* + \zeta_2)s_2 \\ \Rightarrow \{-\gamma + (-\alpha - j\delta - j\zeta_1)\}r_2 &= j(\kappa_p^* + \zeta_2)s_2 \\ \Rightarrow \Gamma r_2 &= j(\kappa_p^* + \zeta_2)s_2\end{aligned}$$

Similarly applying (4.51b) into (4.48)

$$\begin{aligned}-(\gamma s_1 \exp(\gamma z) - \gamma s_2 \exp(-\gamma z)) + (-\alpha - j\delta - j\zeta_1)(s_1 \exp(\gamma z) + s_2 \exp(-\gamma z)) \\ = j(\kappa_p + \zeta_4)(r_1 \exp(\gamma z) + r_2 \exp(-\gamma z)) \\ \Rightarrow \{-\gamma s_1 + (\alpha - j\delta - j\zeta_1)s_1\} \exp(\gamma z) + \{\gamma s_2 + (\alpha - j\delta - j\zeta_1)s_2\} \exp(-\gamma z) \\ = \{j(\kappa_p + \zeta_4)r_1\} \exp(\gamma z) + \{j(\kappa_p + \zeta_4)r_2\} \exp(-\gamma z)\end{aligned}$$

Equating terms with the same exponential

$$\begin{aligned}-\gamma s_1 + (-\alpha - j\delta - j\zeta_1)s_1 &= j(\kappa_p + \zeta_4)r_1 \\ \Rightarrow \{-\gamma + (-\alpha - j\delta - j\zeta_1)\}s_1 &= j(\kappa_p + \zeta_4)r_1 \\ \Rightarrow \Gamma s_1 &= j(\kappa_p + \zeta_4)r_1\end{aligned}$$

and

$$\begin{aligned}\gamma s_2 + (-\alpha - j\delta - j\zeta_1)s_2 &= j(\kappa_p + \zeta_4)r_2 \\ \Rightarrow \{\gamma + (-\alpha - j\delta - j\zeta_1)\}s_2 &= j(\kappa_p + \zeta_4)r_2 \\ \Rightarrow \hat{\Gamma} s_2 &= j(\kappa_p + \zeta_4)r_2\end{aligned}$$

F. Derivation of (4.55)

Replacing equation (4.52a) into (4.52d)

$$\begin{aligned}\hat{\Gamma}r_1 &= j(\kappa_p^* + \zeta_2) \frac{j(\kappa_p + \zeta_4)}{\hat{\Gamma}} r_1 \\ \Rightarrow \hat{\Gamma}\Gamma r_1 &= -(\kappa_p^* + \zeta_2)(\kappa_p + \zeta_4)r_1 \\ \Rightarrow \{-\gamma + (-\alpha - j\delta - j\zeta_1)\}\{\gamma + (-\alpha - j\delta - j\zeta_1)\} &= -\kappa_{\text{eff}}^2 \\ \Rightarrow -\gamma^2 + (\alpha + j\delta + j\zeta_1)^2 + \kappa_{\text{eff}}^2 & \\ \Rightarrow \gamma^2 &= (\alpha + j\delta + j\zeta_1)^2 + \kappa_{\text{eff}}^2\end{aligned}$$

G. Derivation of equation (4.62)

Let us divide equation (4.60) by (4.61),

$$\frac{\frac{j(\kappa_p^* + \zeta_2)}{\hat{\Gamma}} e^{-\gamma L/2} - \hat{\rho}_l e^{-\gamma L/2} e^{j\beta_0 L}}{e^{\gamma L/2} - \hat{\rho}_r e^{\gamma L/2} e^{j\beta_0 L} \frac{j(\kappa_p^* + \zeta_2)}{\Gamma}} = \frac{\hat{\rho}_l \frac{j(\kappa_p + \zeta_4)}{\hat{\Gamma}} e^{j\beta_0 L} e^{\gamma L/2} - e^{\gamma L/2}}{e^{-\gamma L/2} \hat{\rho}_r e^{j\beta_0 L} - \frac{j(\kappa_p + \zeta_4)}{\hat{\Gamma}} e^{-\gamma L/2}}$$

$$\begin{aligned}
& \Rightarrow \frac{\left(\frac{1}{\hat{\Gamma}} \sqrt{\frac{(\kappa_p^* + \zeta_2)(\kappa_p + \zeta_4)(\kappa_p^* + \zeta_2)}{\kappa_p + \zeta_4}} - \hat{\rho}_l e^{j\beta_0 L} \right) e^{-\gamma L}}{1 - \hat{\rho}_r e^{j\beta_0 L} \frac{1}{\hat{\Gamma}} \sqrt{\frac{(\kappa_p^* + \zeta_2)(\kappa_p + \zeta_4)(\kappa_p^* + \zeta_2)}{\kappa_p + \zeta_4}}} = \frac{\left(\hat{\rho}_l \frac{1}{\hat{\Gamma}} \sqrt{\frac{(\kappa_p^* + \zeta_2)(\kappa_p + \zeta_4)(\kappa_p + \zeta_4)}{\kappa_p^* + \zeta_2}} e^{j\beta_0 L} - 1 \right) e^{\gamma L}}{\hat{\rho}_r e^{j\beta_0 L} - \frac{1}{\hat{\Gamma}} \sqrt{\frac{(\kappa_p^* + \zeta_2)(\kappa_p + \zeta_4)(\kappa_p + \zeta_4)}{\kappa_p^* + \zeta_2}}} \\
& \Rightarrow \frac{\left(\frac{j\kappa_{eff}}{\hat{\Gamma}} \sqrt{\frac{\kappa_p^* + \zeta_2}{\kappa_p + \zeta_4}} - \hat{\rho}_l e^{j\beta_0 L} \right) e^{-\gamma L}}{1 - \hat{\rho}_r e^{j\beta_0 L} \frac{j\kappa_{eff}}{\hat{\Gamma}} \sqrt{\frac{\kappa_p^* + \zeta_2}{\kappa_p + \zeta_4}}} = \frac{\left(\hat{\rho}_l \frac{j\kappa_{eff}}{\hat{\Gamma}} \sqrt{\frac{\kappa_p + \zeta_4}{\kappa_p^* + \zeta_2}} e^{j\beta_0 L} - 1 \right) e^{\gamma L}}{\hat{\rho}_r e^{j\beta_0 L} - \frac{j\kappa_{eff}}{\hat{\Gamma}} \sqrt{\frac{\kappa_p + \zeta_4}{\kappa_p^* + \zeta_2}}} \\
& \Rightarrow \frac{\left(\frac{j\kappa_{eff}}{\hat{\Gamma}} - \hat{\rho}_l \sqrt{\frac{\kappa_p + \zeta_4}{\kappa_p^* + \zeta_2}} e^{j\beta_0 L} \right) e^{-\gamma L}}{\hat{\rho}_l \frac{j\kappa_{eff}}{\hat{\Gamma}} \sqrt{\frac{\kappa_p + \zeta_4}{\kappa_p^* + \zeta_2}} e^{j\beta_0 L} - 1} = \frac{\left(1 - \hat{\rho}_r \sqrt{\frac{\kappa_p^* + \zeta_2}{\kappa_p + \zeta_4}} e^{j\beta_0 L} \frac{j\kappa_{eff}}{\hat{\Gamma}} \right) e^{\gamma L}}{\hat{\rho}_r e^{j\beta_0 L} \sqrt{\frac{\kappa_p^* + \zeta_2}{\kappa_p + \zeta_4}} - \frac{j\kappa_{eff}}{\hat{\Gamma}}} \\
& \Rightarrow \frac{\left(\frac{j\kappa_{eff}}{\hat{\Gamma}} - \rho_l \right) e^{-\gamma L}}{1 - \rho_r \frac{j\kappa_{eff}}{\hat{\Gamma}}} = \frac{\left(\rho_l \frac{j\kappa_{eff}}{\hat{\Gamma}} - 1 \right) e^{\gamma L}}{\rho_r - \frac{j\kappa_{eff}}{\hat{\Gamma}}} \\
& \Rightarrow \frac{\left(1 - \rho_l \frac{\hat{\Gamma}}{j\kappa_{eff}} \right) e^{-\gamma L}}{\rho_l - \frac{\hat{\Gamma}}{j\kappa_{eff}}} = \frac{\left(\rho_r - \frac{\hat{\Gamma}}{j\kappa_{eff}} \right) e^{\gamma L}}{1 - \rho_r \frac{\hat{\Gamma}}{j\kappa_{eff}}}
\end{aligned}$$

H. Derivation of equation (4.66)

From equation (4.65)

$$\begin{aligned}
& \frac{\hat{\Gamma}}{j(\kappa_p^* + \zeta_2)} = \frac{j(\kappa_p + \zeta_4)}{\Gamma} \\
& \Rightarrow \frac{j(\kappa_p^* + \zeta_2)}{\hat{\Gamma}} = \frac{\hat{\Gamma} - 2\gamma}{j(\kappa_p + \zeta_4)} \\
& \Rightarrow \frac{2\gamma}{j(\kappa_p + \zeta_4)} = \frac{\hat{\Gamma}}{j(\kappa_p + \zeta_4)} - \frac{j(\kappa_p^* + \zeta_2)}{\hat{\Gamma}} \\
& \Rightarrow \frac{2\gamma}{j(\kappa_p + \zeta_4)} = \frac{\hat{\Gamma}}{j\kappa_{eff}} \sqrt{\frac{\kappa_p^* + \zeta_2}{\kappa_p + \zeta_4}} - \frac{j\kappa_{eff}}{\hat{\Gamma}} \sqrt{\frac{\kappa_p^* + \zeta_2}{\kappa_p + \zeta_4}} \\
& \Rightarrow \frac{2\gamma}{j\kappa_{eff}} \sqrt{\frac{\kappa_p^* + \zeta_2}{\kappa_p + \zeta_4}} = \frac{\hat{\Gamma}}{j\kappa_{eff}} \sqrt{\frac{\kappa_p^* + \zeta_2}{\kappa_p + \zeta_4}} - \frac{j\kappa_{eff}}{\hat{\Gamma}} \sqrt{\frac{\kappa_p^* + \zeta_2}{\kappa_p + \zeta_4}} \\
& \Rightarrow \gamma = \frac{j\kappa_{eff}}{2} \left(\frac{\hat{\Gamma}}{j\kappa_{eff}} - \frac{j\kappa_{eff}}{\hat{\Gamma}} \right)
\end{aligned}$$

I. Derivation of equation (4.68)

Referring to equation (4.67), if we assign the following

$$a = \rho^2 e^{-\gamma L} - e^{\gamma L}, \quad (\text{E.11a})$$

$$b = (\rho_l + \rho_r)(e^{\gamma L} - e^{-\gamma L}) = 2(\rho_l + \rho_r) \sinh \gamma L, \quad (\text{E.11b})$$

$$c = e^{-\gamma L} - \rho^2 e^{\gamma L}. \quad (\text{E.11c})$$

Then the roots of equation (4.67) are

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \quad (\text{E.12})$$

At this point, for equation (4.66), let us calculate

$$\begin{aligned} & x - \frac{1}{x} \\ &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} - \frac{2a}{-b \pm \sqrt{b^2 - 4ac}} \\ &= \frac{(-b \pm \sqrt{b^2 - 4ac})^2 - 4a^2}{2a[-b \pm \sqrt{b^2 - 4ac}]} \\ &= \frac{b^2 + b^2 - 4ac \mp 2b\sqrt{b^2 - 4ac} - 4a^2}{2a[-b \pm \sqrt{b^2 - 4ac}]} \\ &= \frac{b^2 - 2ac \mp b\sqrt{b^2 - 4ac} - 2a^2}{a[-b \pm \sqrt{b^2 - 4ac}]} \\ &= \left[\frac{b(b \mp \sqrt{b^2 - 4ac}) - 2a(a + c)}{a[-b \pm \sqrt{b^2 - 4ac}]} \right] \frac{b \pm \sqrt{b^2 - 4ac}}{b \pm \sqrt{b^2 - 4ac}} \\ &= \frac{b(b^2 - b^2 + 4ac) - 2a(a + c)(b \pm \sqrt{b^2 - 4ac})}{a[b^2 - 4ac - b^2]} \\ &= \frac{4abc - 2a(a + c)(b \pm \sqrt{b^2 - 4ac})}{-4a^2c} \\ &= \frac{(a + c)(b \pm \sqrt{b^2 - 4ac}) - 2bc}{2ac} \\ &= \frac{b(a - c) \pm \sqrt{b^2 - 4ac}(a + c)}{2ac} \end{aligned}$$

We can evaluate the terms on the numerator and denominator from the expressions defined in (E.11a), (E.11b), and (E.11c)

$$a + c = \rho^2 e^{-\gamma L} - e^{\gamma L} + e^{-\gamma L} - \rho^2 e^{\gamma L} = -(1 + \rho^2)2 \sinh \gamma L, \quad (\text{E.13a})$$

$$a - c = \rho^2 e^{-\gamma L} - e^{\gamma L} - e^{-\gamma L} + \rho^2 e^{\gamma L} = 2 \cosh \gamma L(\rho^2 - 1), \quad (\text{E.13b})$$

$$b^2 = 4(\rho_l + \rho_r)^2 \sinh^2 \gamma L, \quad (\text{E.13c})$$

$$\begin{aligned} 4ac &= (a + c)^2 - (a - c)^2 \\ &= 4 \sinh^2 \gamma L(1 + \rho^2)^2 - 4 \cosh^2 \gamma L(\rho^2 - 1)^2 \\ &= 4[\sinh^2 \gamma L(1 + \rho^2)^2 - \cosh^2 \gamma L(\rho^2 - 1)^2] \\ &= 4[(\cosh^2 \gamma L - 1)(1 + \rho^2)^2 - \cosh^2 \gamma L(\rho^2 - 1)^2] \\ &= 4[\cosh^2 \gamma L\{(1 + \rho^2)^2 - (\rho^2 - 1)^2\} - (1 + \rho^2)^2] \end{aligned}$$

$$\begin{aligned}
&=4[\cosh^2 \gamma L \{1 + 2\rho^2 + \rho^4 - (\rho^4 - 2\rho^2 + 1)\} - (1 + \rho^2)^2] \\
&=4[\cosh^2 \gamma L 4\rho^2 - (1 + \rho^2)^2] \\
&=4[4\rho^2 \sinh^2 \gamma L - (\rho^2 - 1)^2] \tag{E.13d}
\end{aligned}$$

$$\begin{aligned}
(b^2 - 4ac)^{1/2} &=2\{[(\rho_l + \rho_r)^2 \sinh^2 \gamma L - 4\rho^2 \sinh^2 \gamma L + (\rho^2 - 1)^2]\}^{1/2} \\
&=2\{[(\rho_l^2 + \rho_r^2 + 2\rho_l \rho_r) \sinh^2 \gamma L - 4\rho^2 \sinh^2 \gamma L - (\rho^2 - 1)^2]\} \\
&=2\{(\rho_l - \rho_r)^2 \sinh^2 \gamma L + (1 - \rho^2)^2\}^{1/2}. \tag{E.13e}
\end{aligned}$$

Putting everything together in equation (4.70), we obtain

$$\begin{aligned}
x - \frac{1}{x} &= \frac{b(a - c) \pm \sqrt{b^2 - 4ac}(a + c)}{2ac} \\
&= \frac{2(\rho_l + \rho_r) \sinh(\gamma L) 2 \cosh(\gamma L) (\rho^2 - 1) \mp \sqrt{b^2 - 4ac} (1 + \rho^2) 2 \sinh(\gamma L)}{2[4\rho^2 \cosh^2(\gamma L) - (\rho^2 + 1)^2]} \\
&= \frac{\{(\rho_l + \rho_r)(1 - \rho^2) 2 \cosh(\gamma L) \pm (1 + \rho^2) \sqrt{b^2 - 4ac}\} \sinh(\gamma L)}{[(1 + \rho^2)^2 - 4\rho^2 \cosh^2(\gamma L)]}.
\end{aligned}$$

J. Derivation of equation (4.72)

The last three terms of equation (E.13) can be simplified as follows

$$\begin{aligned}
&(\kappa_{\text{eff}} L)^2 \sinh^2(\gamma L) \\
&\quad [(1 + \rho^2)^2 (\rho_l - \rho_r)^2 \sinh^2(\gamma L) (1 + \rho^2)^2 (1 - \rho^2)^2 - (\rho_l + \rho_r)^2 (1 - \rho^2)^2 \cosh^2(\gamma L)] \\
&= (\kappa_{\text{eff}} L)^2 \sinh^2(\gamma L) \\
&\quad [(1 + \rho^2)^2 (\rho_l - \rho_r)^2 (\cosh^2(\gamma L) - 1) (1 + \rho^2)^2 (1 - \rho^2)^2 - (\rho_l + \rho_r)^2 (1 - \rho^2)^2 \cosh^2(\gamma L)] \\
&= (\kappa_{\text{eff}} L)^2 \sinh^2(\gamma L) \\
&\quad [\cosh^2(\gamma L) \{ (1 + \rho^2)^2 (\rho_l - \rho_r)^2 - (\rho_l + \rho_r)^2 (1 - \rho^2)^2 \} + (1 + \rho^2)^2 (1 - \rho^2)^2 \\
&\quad \quad - (1 + \rho^2)^2 (\rho_l - \rho_r)^2] \\
&= (\kappa_{\text{eff}} L)^2 \sinh^2(\gamma L) \\
&\quad [\cosh^2(\gamma L) \{ (1 + \rho^2)^2 (\rho_l - \rho_r)^2 - (\rho_l + \rho_r)^2 (1 - \rho^2)^2 \} \\
&\quad \quad + (1 + \rho^2)^2 (1 - \rho^2)^2 (1 + \rho^2)^2 (\rho_l - \rho_r)^2] \\
&= (\kappa_{\text{eff}} L)^2 \sinh^2(\gamma L) \\
&\quad [\cosh^2(\gamma L) \{ (1 + 2\rho^2 + \rho^4) (\rho_l + \rho_r^2 - 2\rho^2) - (\rho_l^2 + \rho_r^2 + 2\rho^2) (1 - 2\rho^2 + \rho^4) \}
\end{aligned}$$

$$\begin{aligned}
& + (1 + \rho^2)^2 \{ (1 - \rho^2)^2 - (\rho_l - \rho_r)^2 \}] \\
& = (\kappa_{\text{eff}} L)^2 \sinh^2(\gamma L) \\
& \quad [\cosh^2(\gamma L) \{ \cancel{\rho_l^2} + \cancel{\rho_r^2} - 2\rho^2 + 2\rho^2 \cancel{\rho_l^2} + 2\rho^2 \cancel{\rho_r^2} - \cancel{4\rho^4} + \cancel{\rho_l^2 \rho^4} + \cancel{\rho_r^2 \rho^4} } \\
& \quad - 2\rho^6 - (\cancel{\rho_l^2} + \cancel{\rho_r^2} + 2\rho^2 - 2\rho^2 \cancel{\rho_l^2} - 2\rho^2 \cancel{\rho_r^2} - \cancel{4\rho^4} + \cancel{\rho_l^2 \rho^4} + \cancel{\rho_r^2 \rho^4} + 2\rho^6) \} \\
& \quad + (1 + \rho^2)^2 \{ 1 - \cancel{2\rho^2} + \rho^4 - \rho_l^2 - \rho_r^2 + \cancel{2\rho^2} \}] \\
& = (\kappa_{\text{eff}} L)^2 \sinh^2(\gamma L) \\
& \quad [\cosh^2(\gamma L) \{ 4\rho^2 \rho_l^2 + 4\rho^2 \rho_r^2 - 4\rho^2 - 4\rho^6 \} + (1 + \rho^2) \{ (1 - \rho_l^2) - \rho_r^2 + \rho_l^2 \rho_r^2 \}] \\
& = (\kappa_{\text{eff}} L)^2 \sinh^2(\gamma L) [4\rho^2 \cosh^2(\gamma L) \{ -1 + \rho_l^2 + \rho_r^2 - \rho^4 \} + (1 + \rho^2)^2 \{ (1 - \rho_l^2) - \rho_r^2 (1 - \rho_l^2) \}] \\
& = (\kappa_{\text{eff}} L)^2 \sinh^2(\gamma L) (1 - \rho_l^2) (1 - \rho_r^2) [(1 + \rho^2)^2 - 4\rho^2 \cosh^2(\gamma L)] \\
& = (\kappa_{\text{eff}} L)^2 \sinh^2(\gamma L) \mathcal{D} (1 - \rho_l^2) (1 - \rho_r^2)
\end{aligned}$$

K. Derivation of equations (4.79) and (4.80)

We start from equation (4.77)

$$2(\gamma L)^2 + 2(\kappa_{\text{eff}} L)^2 \sinh^2(\gamma L) (1 - \rho_l^2) + 2j\tilde{\rho}_l(\kappa_{\text{eff}} L)(\gamma L) \sinh(2\gamma L) = 0$$

From the assignment

$$(\gamma l) = a + jb; \quad \tilde{\rho}_l = -\rho_l = -(c + jd); \quad (\kappa_{\text{eff}} L) = q + jp$$

it follows

$$\begin{aligned}
& (1 - \rho_l^2) \\
& = 1 - \{ -(c + jd)^2 \} \\
& = 1 - (c + jd)^2 \\
& = 1 - (c^2 - d^2 + j2cd) \\
& = \underbrace{(1 - c^2 + d^2)}_m - j \underbrace{2cd}_n
\end{aligned} \tag{E.14}$$

From the first term we get

$$\begin{aligned}
& 2(\gamma L)^2 \\
& = 2(a + jb)^2 \\
& = 2(a^2 - b^2 + j2ab) \\
& = 2(a^2 - b^2) + j4ab
\end{aligned} \tag{E.15}$$

From the second term

$$\begin{aligned}
& (\kappa_{\text{eff}}L)^2(1 - \rho_l^2)2 \sinh^2(\gamma L) \\
&= (\kappa_{\text{eff}}L)^2(1 - \rho_l^2)[\cosh(2\gamma L) - 1] \\
&= (q + jp)^2(m - jn) \left[\frac{e^{2\gamma L} + e^{-2\gamma L}}{2} - 1 \right] \\
&= (q^2 - p^2 + 2jqpm)(m - jn) [e^{2\gamma L} + e^{-2\gamma L} - 2] \frac{1}{2} \\
&= [m(q^2 - p^2) + j2qpm - jn(q^2 - p^2) + 2qpn] \frac{1}{2} [e^{2(a+jb)} + e^{-2(a+jb)} - 2] \\
&= [m(q^2 - p^2) + j2qpm - jn(q^2 - p^2) + 2qpn] \frac{1}{2} [e^{2a}e^{j2b} + e^{-2a}e^{-j2b} - 2] \\
&= [m(q^2 - p^2) + j2qpm - jn(q^2 - p^2) + 2qpn] \\
&\quad \frac{1}{2} [e^{2a}(\cos 2b + j \sin 2b) + e^{-2a}(\cos(2b) - j \sin(2b)) - 2] \\
&= [m(q^2 - p^2) + 2qpn + j\{2qpm - n(q^2 - p^2)\}] \\
&\quad \frac{1}{2} [\cos(2b)(e^{2a} + e^{-2a}) + j \sin(2b)(e^{2a} - e^{-2a}) - 2] \\
&= [m(q^2 - p^2) \cos(2b) \cosh(2a) + 2qpn \cos(2b) \cosh(2a) \\
&\quad - 2qpm \sin(2b) \sinh(2a) + n(q^2 - p^2) \sin(2b) \sinh(2a) - m(q^2 - p^2) - 2qpn] \\
&\quad + j[m(q^2 - p^2) \sin(2b) \sinh(2a) + 2qpn \sin(2b) \sinh(2a) \\
&\quad + 2qpm \cos(2b) \cosh(2a) - n(q^2 - p^2) \cos(2b) \cosh(2a) - 2qpm + n(q^2 - p^2)] \\
&\hspace{15em} \text{(E.16)}
\end{aligned}$$

From the third term

$$\begin{aligned}
& 2j\tilde{\rho}_l(\kappa_{\text{eff}}L)(\gamma L) \sinh(2\gamma L) \\
&= -2j(c + jd)(q + jp)(a + jb) \sinh(2\gamma L) \\
&= -2j(qc + jqd + jpc - pd)(a + jb) \left[\frac{e^{2\gamma L} - e^{-2\gamma L}}{2} \right] \\
&= -2j\{(qc - pd) + j(qd + pc)\}(a + jb) \frac{1}{2} [e^{2(a+jb)} - e^{-2(a+jb)}] \\
&= -2j\{a(qc - pd) + jb(qc - pd) + ja(qd + pc) - b(qd + pc)\} \frac{1}{2} [e^{2a}e^{jb} - e^{-2a}e^{-j2b}] \\
&= -2j\{a(qc - pd) - b(qd + pc) - jb(qc - pd) + ja(qd + pc)\} \\
&\quad \frac{1}{2} [e^{2a}(\cos(2b) + j \sin(2b)) - e^{-2a}(\cos(2b) - j \sin(2b))] \\
&= -2j\{a(qc - pd) - b(qd + pc) + jb(qc - pd) + ja(qd + pc)\} \\
&\quad [\cos(2b)(e^{2a} - e^{-2a}) + j \sin(2b)(e^{2a} + e^{-2a})] \\
&= -2j\{a(qc - pd) - b(qd + pc) + jb(qc - pd) + ja(qd + pc)\} \\
&\quad [\cos(2b) \sinh(2a) + j \sin(2b) \cosh(2a)] \\
&= \{-2ja(qc - pd) + 2jb(qd + pc) + 2b(qc - pd) + 2a(qd + pc)\} \\
&\quad [\cos(2b) \sinh(2a) + j \sin(2b) \cosh(2a)]
\end{aligned}$$

$$\begin{aligned}
&= \cos(2b) \sinh(2a) [2b(qc - pd) + 2a(qd + pc)] \\
&+ \sin(2b) \cosh(2a) [2a(qc - pd) - 2b(qd + pc)] \\
&+ j \cos(2b) \sinh(2a) [-2a(qc - pd) + 2b(qd + pc)] \\
&+ j \sin(2b) \cosh(2a) [2a(qd + pc) + 2b(qc - pd)]
\end{aligned} \tag{E.17}$$

Summing up (E.15), (E.16), and (E.17) and equating the real and the imaginary parts one arrives at equations (4.79) and (4.80).

L. Derivation of equations (4.81a) and (4.81b)

Taking derivative of equation (4.47) with respect to z

$$\begin{aligned}
R'' - (\alpha + j\delta + j\zeta_1)R' &= j(\kappa_p^* + \zeta_2)S' \\
\Rightarrow R'' - (\alpha + j\delta + j\zeta_1)((\alpha + j\delta + j\zeta_1)R + j(\kappa_p^* + \zeta_2)S) &= \\
&\quad j(\kappa_p^* + \zeta_2)(-(\alpha + j\delta + j\zeta_1)S - j(\kappa_p + \zeta_4)R) \\
\Rightarrow R'' - (\alpha + j\delta + j\zeta_1)^2 R - j(\kappa_p^* + \zeta_2)(\alpha + j\delta + j\zeta_1)S &= \\
&\quad -j(\kappa_p^* + \zeta_2)(\alpha + j\delta + j\zeta_1)S + (\kappa_p + \zeta_4)(\kappa_p^* + \zeta_2)R \\
\Rightarrow R'' - (\alpha + j\delta + j\zeta_1)^2 R - \kappa_{\text{eff}}^2 R &= 0 \\
\Rightarrow R'' - ((\alpha + j\delta + j\zeta_1)^2 + \kappa_{\text{eff}}^2)R &= 0 \Rightarrow R'' - \gamma^2 R = 0
\end{aligned}$$

Going on the exact similar direction from (4.48) we get

$$S'' - \gamma^2 S = 0$$

M. Derivation of equations (4.82) and (4.84)

From (4.51a),

$$\begin{aligned}
R(z) &= r_1(\cosh(\gamma z) + \sinh(\gamma z)) + r_2(\cosh(\gamma z) - \sinh(\gamma z)) \\
\Rightarrow R(z) &= (r_1 + r_2)\cosh(\gamma z) + (r_1 - r_2)\sinh(\gamma z) \\
\Rightarrow R(z) &= R_1 \cosh(\gamma z) + R_2 \sinh(\gamma z).
\end{aligned} \tag{E.18}$$

From (4.51b),

$$\begin{aligned}
S(z) &= s_1(\cosh(\gamma z) + \sinh(\gamma z)) + s_2(\cosh(\gamma z) - \sinh(\gamma z)) \\
\Rightarrow S(z) &= (s_1 + s_2)\cosh(\gamma z) + (s_1 - s_2)\sinh(\gamma z) \\
\Rightarrow S(z) &= S_1 \cosh(\gamma z) + S_2 \sinh(\gamma z).
\end{aligned} \tag{E.19}$$

Taking derivative of (4.51a) again w.r.t. z ,

$$R'(z) = \gamma r_1 e^{\gamma z} - \gamma r_2 e^{-\gamma z}. \tag{E.20}$$

At $z = -L/2$, from (4.51a) and (E.20),

$$R(-L/2) = r_1 e^{-\gamma L/2} + r_2 e^{\gamma L/2}, \tag{E.21a}$$

$$R'(-L/2) = \gamma r_1 e^{-\gamma L/2} - \gamma r_2 e^{\gamma L/2}. \tag{E.21b}$$

Now, $\gamma \times$ (E.21a) + (E.21b) $\Rightarrow$

$$\begin{aligned} \gamma R(-L/2) + R'(-L/2) &= 2\gamma r_1 e^{-\gamma L/2} \\ \Rightarrow r_1 &= \frac{e^{\gamma L/2}}{2} \left(R(-L/2) + \frac{R'(-L/2)}{\gamma} \right), \end{aligned} \quad (\text{E.22})$$

and $\gamma \times$ (E.21a) - (E.21b) $\Rightarrow$

$$\begin{aligned} \gamma R(-L/2) - R'(-L/2) &= 2\gamma r_2 e^{-\gamma L/2} \\ \Rightarrow r_2 &= \frac{e^{-\gamma L/2}}{2} \left(R(-L/2) - \frac{R'(-L/2)}{\gamma} \right). \end{aligned} \quad (\text{E.23})$$

(E.22) + (E.23) $\Rightarrow$

$$\begin{aligned} r_1 + r_2 &= R(-L/2) \left[\frac{e^{\gamma L/2} + e^{-\gamma L/2}}{2} \right] + \frac{R'(-L/2)}{\gamma} \left[\frac{e^{\gamma L/2} - e^{-\gamma L/2}}{2} \right] \\ \Rightarrow R_1 &= R(-L/2) \cosh(\gamma L/2) + \frac{R'(-L/2)}{\gamma} \sinh(\gamma L/2). \end{aligned} \quad (\text{E.24})$$

Similarly, (E.22) - (E.23) $\Rightarrow$

$$r_1 - r_2 = R_2 = R(-L/2) \sinh(\gamma L/2) + \frac{R'(-L/2)}{\gamma} \cosh(\gamma L/2). \quad (\text{E.25})$$

Replacing R_1, R_2 with their expressions in equation (E.18), we obtain

$$\begin{aligned} R(z) &= \left\{ R(-L/2) \cosh(\gamma L/2) + \frac{R'(-L/2)}{\gamma} \sinh(\gamma L/2) \right\} \cosh(\gamma z) \\ &\quad + \left\{ R(-L/2) \sinh(\gamma L/2) + \frac{R'(-L/2)}{\gamma} \cosh(\gamma L/2) \right\} \sinh(\gamma z) \\ &= R(-L/2) \cosh[\gamma(z + L/2)] + \frac{R'(-L/2)}{\gamma} \sinh[\gamma(z + L/2)], \end{aligned} \quad (\text{E.26})$$

where we have used the following hyperbolic identities

$$\begin{aligned} \sinh(x + y) &= \sinh x \cosh y + \cosh x \sinh y, \\ \cosh(x + y) &= \cosh x \cosh y + \sinh x \sinh y. \end{aligned}$$

Following the exact same procedure, and making use of the boundary conditions at $z = L/2$, one finds from equation (E.19)

$$S(z) = S(L/2) \cosh[\gamma(z - L/2)] + \frac{S'(L/2)}{\gamma} \sinh[\gamma(z - L/2)]. \quad (\text{E.27})$$

For our laser, one of the facets is reflective, while the other is not. In this scenario

$$R(z)|_{z=-L/2} = R(-L/2) = 1, \quad (\text{E.28a})$$

$$S(z)|_{z=L/2} = S(L/2) = rR(L/2), \quad (\text{E.28b})$$

where r is the reflectivity of one facet. The derivative of equation (E.27) is

$$S'(z) = \frac{d}{dz} \left[S(L/2) \cosh(\gamma(z - L/2)) + \frac{S'(L/2)}{\gamma} \sinh(\gamma(z - L/2)) \right]. \quad (\text{E.29})$$

Let us assume

$$S(L/2) = A_1 = rR(L/2), \quad \frac{S'(L/2)}{\gamma} = A_2. \quad (\text{E.30})$$

Then, one can write equation (E.29) as

$$S'(z) = \gamma A_1 \sinh(\gamma(z - L/2)) + \gamma A_2 \cosh(\gamma(z - L/2)). \quad (\text{E.31})$$

Now, from equation (4.48),

$$\begin{aligned} S' + (\alpha + j\delta + j\zeta_1)S &= -j(\kappa_p + \zeta_4)R \\ \Rightarrow \gamma A_1 \sinh(\gamma(z - L/2)) + \gamma A_2 \cosh(\gamma(z - L/2)) + (\alpha + j\delta + j\zeta_1) \\ &\quad [A_1 \cosh(\gamma(z - L/2)) + A_2 \sinh(\gamma(z - L/2))] + j(\kappa_p + \zeta_4)R(z) = 0. \end{aligned} \quad (\text{E.32})$$

at $z = L/2$ from equation (E.32)

$$\begin{aligned} R(L/2) &= \frac{-1}{j(\kappa_p + \zeta_4)} \left[\gamma A_2 + (\alpha + j\delta + j\zeta_1)S(L/2) \right] \\ \Rightarrow R(L/2) &= \frac{-1}{j(\kappa_p + \zeta_4)} \left[\gamma A_2 + (\alpha + j\delta + j\zeta_1)rR(L/2) \right] \\ \Rightarrow R(L/2) &= \frac{-\gamma A_2}{j(\kappa_p + \zeta_4) + (\alpha + j\delta + j\zeta_1)r} \end{aligned} \quad (\text{E.33})$$

at $z = -L/2$ from equation (E.32)

$$\begin{aligned} -j(\kappa_p + \zeta_4)R(-L/2) &= -\gamma A_1 \sinh(\gamma L) + \gamma A_2 \cosh(\gamma L) + (\alpha + j\delta + j\zeta_1) \\ &\quad [A_1 \cosh(\gamma(z - L/2)) - A_2 \sinh(\gamma(z - L/2))] \\ \Rightarrow -j(\kappa_p + \zeta_4)R(-L/2) &= A_1 [-\gamma \sinh(\gamma L) + (\alpha + j\delta + j\zeta_1) \cosh(\gamma L)] \\ &\quad + A_2 [\gamma \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)] \\ \Rightarrow -j(\kappa_p + \zeta_4) + rR(L/2) [\gamma \sinh(\gamma L) - (\alpha + j\delta + j\zeta_1) \cosh(\gamma L)] \\ &= A_2 [\gamma \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)] \\ \Rightarrow A_2 &= \frac{-j(\kappa_p + \zeta_4) + rR(L/2) [\gamma \sinh(\gamma L) - (\alpha + j\delta + j\zeta_1) \cosh(\gamma L)]}{[\gamma \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)]} \end{aligned} \quad (\text{E.34})$$

Therefore

$$\begin{aligned}
 R(L/2) &= \frac{-\gamma}{j(\kappa_p + \zeta_4)(\alpha + j\delta + j\zeta_1)} \times \\
 &\quad \left[\frac{-j(\kappa_p + \zeta_4) + rR(L/2)\{\gamma \sinh(\gamma L) - (\alpha + j\delta + j\zeta_1) \cosh(\gamma L)\}}{\gamma \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)} \right] \\
 \Rightarrow R(L/2) &= \frac{\gamma}{j(\kappa_p + \zeta_4)(\alpha + j\delta + j\zeta_1)} \\
 &\quad \times \left[\frac{j(\kappa_p + \zeta_4) - rR(L/2)\{\gamma \sinh(\gamma L) - (\alpha + j\delta + j\zeta_1) \cosh(\gamma L)\}}{\gamma \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)} \right] \\
 \Rightarrow R(L/2) &\left[1 + \frac{r\{\gamma \sinh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)\}}{\gamma \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)} \times \frac{\gamma}{j(\kappa_p + \zeta_4)(\alpha + j\delta + j\zeta_1)} \right] \\
 &= \frac{j\gamma(\kappa_p + \zeta_4)}{\gamma \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)} \times \frac{\gamma}{j(\kappa_p + \zeta_4) + r(\alpha + j\delta + j\zeta_1)} \quad (E.35)
 \end{aligned}$$

Let be

$$\{j(\kappa_p + \zeta_4) + r(\alpha + j\delta + j\zeta_1)\} \{\gamma \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)\} = \chi. \quad (E.36)$$

Then

$$\begin{aligned}
 R(L/2) &\left[\frac{\chi + \gamma r\{\gamma \sinh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)\}}{\chi} \right] = \frac{j(\kappa_p + \zeta_4)}{\chi} \\
 \Rightarrow R(L/2) &= \frac{j\gamma(\kappa_p + \zeta_4)}{\chi + \gamma r\{\gamma \sinh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)\}} \quad (E.37)
 \end{aligned}$$

The denominator can be consolidated as

$$\begin{aligned}
 &j(\kappa_p + \zeta_4)\gamma \cosh(\gamma L) - j(\kappa_p + \zeta_4)(\alpha + j\delta + j\zeta_1) \sinh(\gamma L) \\
 &+ \gamma r(\alpha + j\delta + j\zeta_1) \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1)^2 r \sinh(\gamma L) \\
 &- \gamma r(\alpha + j\delta + j\zeta_1) \cosh(\gamma L) + \gamma^2 r \sinh(\gamma L) \\
 &= j(\kappa_p + \zeta_4) [\gamma \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)] - r \sinh(\gamma L) [(\alpha + j\delta + j\zeta_1)^2 - \gamma^2] \\
 &= j(\kappa_p + \zeta_4) [\gamma \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)] + \kappa_{\text{eff}}^2 r \sinh(\gamma L)
 \end{aligned}$$

Putting everything together, we arrive at equation (4.82).

From equation (E.27), at $z = -L/2$,

$$\begin{aligned}
 S(-L/2) &= rR(L/2) \cosh(\gamma L) - A_2 \sinh(\gamma L) \\
 &= rR(L/2) \cosh \gamma L - \\
 &\frac{-j(\kappa_p + \zeta_4) \sinh(\gamma L) + rR(L/2) \sinh(\gamma L) [\gamma \sinh(\gamma L) - (\alpha + j\delta + j\zeta_1) \cosh(\gamma L)]}{[\gamma \cosh(\gamma L) - (\alpha + j\delta + j\zeta_1) \sinh(\gamma L)]}
 \end{aligned} \tag{E.38}$$

Upon subtraction, the numerator simplifies as following

$$\begin{aligned}
 &rR(L/2) [\cosh^2(\gamma L) - \sinh^2(\gamma L)] - rR(L/2) \cosh(\gamma L) \sinh(\gamma L) + \\
 &rR(L/2) \cosh(\gamma L) \sinh(\gamma L) + j(\kappa_p + \zeta_4) \sinh(\gamma L) = \gamma rR(L/2) + j(\kappa_p + \zeta_4) \sinh(\gamma L)
 \end{aligned} \tag{E.39}$$

Combining the numerator and the denominator, we find the equation (4.84).

Bibliography

- [1] P. K. Tien, "Integrated optics and new wave phenomena in optical waveguides," *Reviews of Modern Physics*, vol. 49, no. 2, pp. 361–420, 1977.
- [2] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*, ser. Wiley Series in Pure and Applied Optics. New York, USA: John Wiley & Sons, Inc., 1991.
- [3] [Online]. Available: <http://nonlinearphotonics.ca/>.
- [4] E. Knill, R. Laflamme, and G. J. Milburn, "A scheme for efficient quantum computation with linear optics," *Nature*, vol. 409, no. 6816, pp. 46–52, 2001.
- [5] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, "Linear optical quantum computing with photonic qubits," *Reviews of Modern Physics*, vol. 79, no. 1, pp. 135–174, 2007.
- [6] (). "Performance summary table." Accessed Dec. 09, 2020, JePPiX, [Online]. Available: <https://www.jeppix.eu/mpw-services/get-started/performance-summary-table/>.
- [7] J. E. Bowers and A. Y. Liu, "A comparison of four approaches to photonic integration," in *2017 Optical Fiber Communications Conference and Exhibition (OFC)*, 2017, pp. 1–3.
- [8] R. Nagarajan, C. Doerr, and F. Kish, "Semiconductor Photonic Integrated Circuit Transmitters and Receivers," in *Optical Fiber Telecommunications*, Elsevier, 2013, pp. 25–98.
- [9] J. Capmany and D. Pérez, "Introduction to Programmable Integrated Photonics," in *Programmable Integrated Photonics*, Oxford University Press, 2020, pp. 1–37. DOI: 10.1093/oso/9780198844402.003.0001. [Online]. Available: <https://oxford.universitypressscholarship.com/view/10.1093/oso/9780198844402.001.0001/oso-9780198844402-chapter-1>.
- [10] J. Wang, F. Sciarrino, A. Laing, and M. G. Thompson, "Integrated photonic quantum technologies," *Nature Photonics*, vol. 14, no. 5, pp. 273–284, 2020, ISSN: 1749-4893. DOI: 10.1038/s41566-019-0532-1. [Online]. Available: <https://doi.org/10.1038/s41566-019-0532-1>.
- [11] A. W. Elshaari, W. Pernice, K. Srinivasan, O. Benson, and V. Zwiller, "Hybrid integrated quantum photonic circuits," *Nature Photonics*, vol. 14, no. 5, pp. 285–298, 2020. DOI: 10.1038/s41566-020-0609-x. [Online]. Available: <https://doi.org/10.1038/s41566-020-0609-x>.
- [12] Telefonica, *Telefónica, huawei and upm perform a groundbreaking field trial applying quantum cryptography on commercial optical networks to provide secure communication services*, 2021 [Online]. [Online]. Available: <https://www.telefonica.com/en/web/press-office/-/telefonica-huawei->

- and-upm-perform-a-grounbreaking-field-trial-applying-quantum-cryptography-on-commercial-optical-networks-to-provide-secure-communicat.
- [13] CiViQ, *Develop quantum-enhanced physical layer security services, combined with modern cryptographic techniques, to enable unparalleled applications and services*. 2021 [Online]. [Online]. Available: <https://civiquantum.eu/>.
 - [14] UNIQORN, *Uniqorn: Affordable quantum communication for everyone*, 2021 [Online]. [Online]. Available: <https://quantum-uniqorn.eu/>.
 - [15] G. Zhang, J. Y. Haw, H. Cai, F. Xu, S. M. Assad, J. F. Fitzsimons, X. Zhou, Y. Zhang, S. Yu, J. Wu, W. Ser, L. C. Kwek, and A. Q. Liu, "An integrated silicon photonic chip platform for continuous-variable quantum key distribution," *Nature Photonics*, vol. 13, no. 12, pp. 839–842, 2019, ISSN: 1749-4893. DOI: 10.1038/s41566-019-0504-5. [Online]. Available: <https://doi.org/10.1038/s41566-019-0504-5>.
 - [16] Y. Shen, L. Cao, X. Y. Wang, J. Zou, W. Luo, Y. X. Wang, H. Cai, B. Dong, X. S. Luo, W. J. Fan, L. C. Kwek, and A. Q. Liu, "On-chip continuous-variable quantum key distribution(cv-qkd) and homodyne detection," in *Optical Fiber Communication Conference (OFC) 2020*, Optical Society of America, 2020, W2A.53. DOI: 10.1364/OFC.2020.W2A.53. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2020-W2A.53>.
 - [17] V. I. Tolstikhin, R. Moore, K. Pimenov, Y. Logvin, F. Wu, and C. D. Watson, "One-step growth optical transceiver pic in inp," in *2009 35th European Conference on Optical Communication*, 2009, pp. 1–2.
 - [18] V. Tolstikhin, "Regrowth-free multi-guide vertical integration in inp for optical communications," in *IPRM 2011 - 23rd International Conference on Indium Phosphide and Related Materials*, 2011, pp. 1–4.
 - [19] S. Kuntze, V. Tolstikhin, F. Wu, Y. Logvin, C. Watson, K. Pimenov, R. Moore, A. Moore, H. Wang, and T. Oogarah, "Transmitter and receiver solutions for regrowth-free multi-guide vertical integration in inp," in *Integrated Photonics Research, Silicon and Nanophotonics and Photonics in Switching*, Optical Society of America, 2010, ITuC5. DOI: 10.1364/IPRSN.2010.ITuC5. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=IPRSN-2010-ITuC5>.
 - [20] V. Tolstikhin, "Distributed feedback lasers for regrowth-free multi-guide vertical integration in inp," in *2010 23rd Annual Meeting of the IEEE Photonics Society*, 2010, pp. 521–522. DOI: 10.1109/PHOTONICS.2010.5698990.
 - [21] L. Miller, J. Verdeyen, J. Coleman, R. Bryan, J. Alwan, K. Beernink, J. Hughes, and T. Cockerill, "A distributed feedback ridge waveguide quantum well heterostructure laser," *IEEE Photonics Technology Letters*, vol. 3, no. 1, pp. 6–8, 1991. DOI: 10.1109/68.68030.
 - [22] A. G. Bell, "Vertically-coupled surface-etched grating dfb laser," pat. WO2012058742A1, May 2012. [Online]. Available: <https://www.google.com/patents/US174465>.
 - [23] W. Streifer, D. Scifres, and R. Burnham, "Coupled wave analysis of dfb and dbr lasers," *IEEE Journal of Quantum Electronics*, vol. 13, no. 4, pp. 134–141, 1977. DOI: 10.1109/JQE.1977.1069328.

- [24] R. Millett, "Modeling fabrication, and measurement of laterally-coupled distributed feedback lasers," Ph.D. dissertation, Dept. of Elec. Engg., Univ. of Ottawa, Ottawa, ON, CANADA, 2009.
- [25] C. Mead and L. Conway, *Introduction to VLSI Systems*, 1st ed. Boston, MA, USA: Addison-Wesley, 1979.
- [26] M. Smit, K. Williams, and J. van der Tol, "Past, present, and future of InP-based photonic integration," *APL Photonics*, vol. 4, no. 5, p. 050 901, 2019.
- [27] M. Wale, J. Van der Tol, X. Leijters, N. Grote, H. Ambrosius, M. Schell, D. Robbins, M. Smit, and E. Bente, "Generic foundry model for InP-based photonics," *IET Optoelectronics*, vol. 5, no. 5, pp. 187–194, 2011.
- [28] M. Smit, X. Leijters, H. Ambrosius, E. Bente, J. van der Tol, B. Smalbrugge, T. de Vries, E.-J. Geluk, J. Bolk, R. van Veldhoven, L. Augustin, P. Thijs, D. D'Agostino, H. Rabbani, K. Lawniczuk, S. Stopinski, S. Tahvili, A. Corradi, E. Kleijn, D. Dzibrou, M. Felicetti, E. Bitincka, V. Moskalenko, J. Zhao, R. Santos, G. Gilardi, W. Yao, K. Williams, P. Stabile, P. Kuindersma, J. Pello, S. Bhat, Y. Jiao, D. Heiss, G. Roelkens, M. Wale, P. Firth, F. Soares, N. Grote, M. Schell, H. Debregeas, M. Achouche, J.-L. Gentner, A. Bakker, T. Korthorst, D. Gallagher, A. Dabbs, A. Melloni, F. Morichetti, D. Melati, A. Wonfor, R. Pentty, R. Broeke, B. Musk, and D. Robbins, "An introduction to InP-based generic integration technology," *Semiconductor Science and Technology*, vol. 29, no. 8, p. 083 001, 2014. [Online]. Available: <https://iopscience.iop.org/article/10.1088/0268-1242/29/8/083001>.
- [29] M. U. Khan, Y. Xing, Y. Ye, and W. Bogaerts, "Photonic Integrated Circuit Design in a Foundry+Fabless Ecosystem," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 25, no. 5, pp. 1–14, 2019.
- [30] C. E. Shannon, "A mathematical theory of communication.," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, 1948. [Online]. Available: <http://dblp.uni-trier.de/db/journals/bstj/bstj27.html#Shannon48>.
- [31] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," *Theoretical Computer Science*, vol. 560, 7–11, 2014, ISSN: 0304-3975. DOI: 10.1016/j.tcs.2014.05.025. [Online]. Available: <http://dx.doi.org/10.1016/j.tcs.2014.05.025>.
- [32] ID Quantique, Accessed Nov. 18, 2021 [Online]. [Online]. Available: <https://www.idquantique.com/>.
- [33] E. Diamanti and A. Leverrier, "Distributing secret keys with quantum continuous variables: Principle, security and implementations," *Entropy*, vol. 17, no. 9, pp. 6072–6092, 2015, ISSN: 1099-4300. DOI: 10.3390/e17096072. [Online]. Available: <https://www.mdpi.com/1099-4300/17/9/6072>.
- [34] F. Grosshans and P. Grangier, "Continuous variable quantum cryptography using coherent states," *Physical Review Letters*, vol. 88, no. 5, 2002, ISSN: 1079-7114. DOI: 10.1103/physrevlett.88.057902. [Online]. Available: <http://dx.doi.org/10.1103/PhysRevLett.88.057902>.
- [35] N. J. Cerf, M. Lévy, and G. V. Assche, "Quantum distribution of gaussian keys using squeezed states," *Physical Review A*, vol. 63, no. 5, 2001,

- ISSN: 1094-1622. DOI: 10.1103/physreva.63.052311. [Online]. Available: <http://dx.doi.org/10.1103/PhysRevA.63.052311>.
- [36] C. Weedbrook, A. M. Lance, W. P. Bowen, T. Symul, T. C. Ralph, and P. K. Lam, "Quantum cryptography without switching," *Physical Review Letters*, vol. 93, no. 17, 2004, ISSN: 1079-7114. DOI: 10.1103/physrevlett.93.170504. [Online]. Available: <http://dx.doi.org/10.1103/PhysRevLett.93.170504>.
- [37] V. C. Usenko and F. Grosshans, "Unidimensional continuous-variable quantum key distribution," *Physical Review A*, vol. 92, no. 6, 2015, ISSN: 1094-1622. DOI: 10.1103/physreva.92.062337. [Online]. Available: <http://dx.doi.org/10.1103/PhysRevA.92.062337>.
- [38] R. García-Patrón and N. J. Cerf, "Continuous-variable quantum key distribution protocols over noisy channels," *Phys. Rev. Lett.*, vol. 102, p. 130501, 13 2009. DOI: 10.1103/PhysRevLett.102.130501. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.102.130501>.
- [39] R. Filip, "Continuous-variable quantum key distribution with noisy coherent states," *Phys. Rev. A*, vol. 77, p. 022310, 2 2008. DOI: 10.1103/PhysRevA.77.022310. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.77.022310>.
- [40] L. S. Madsen, V. C. Usenko, M. Lassen, R. Filip, and U. L. Andersen, "Continuous variable quantum key distribution with modulated entangled states," *Nature Communications*, vol. 3, no. 1, p. 1083, 2012, ISSN: 2041-1723. DOI: 10.1038/ncomms2097. [Online]. Available: <https://doi.org/10.1038/ncomms2097>.
- [41] ek, familyi=v., given=Jaromír, giveni=J., „ and N. J. Cerf, "Gaussian post-selection and virtual noiseless amplification in continuous-variable quantum key distribution," *Phys. Rev. A*, vol. 86, p. 060302, 6 2012. DOI: 10.1103/PhysRevA.86.060302. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.86.060302>.
- [42] C. H. Bennett, "Quantum cryptography using any two nonorthogonal states," *Phys. Rev. Lett.*, vol. 68, pp. 3121–3124, 21 1992. DOI: 10.1103/PhysRevLett.68.3121. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.68.3121>.
- [43] C. Gerry and P. Knight, *Introductory Quantum Optics*. Cambridge University Press, 2004. DOI: 10.1017/CB09780511791239.
- [44] G. Grynberg, A. Aspect, C. Fabre, and C. Cohen-Tannoudji, *Introduction to Quantum Optics: From the Semi-classical Approach to Quantized Light*. Cambridge University Press, 2010. DOI: 10.1017/CB09780511778261.
- [45] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, "Gaussian quantum information," *Reviews of Modern Physics*, vol. 84, no. 2, 621–669, 2012, ISSN: 1539-0756. DOI: 10.1103/revmodphys.84.621. [Online]. Available: <http://dx.doi.org/10.1103/RevModPhys.84.621>.
- [46] C. Navarrete-Benlloch, *An Introduction to the Formalism of Quantum Information with Continuous Variables*, ser. 2053-2571. Morgan Claypool Publishers, 2015, ISBN: 978-1-6817-4405-6. DOI: 10.1088/978-1-6817-4405-

6. [Online]. Available: <https://dx.doi.org/10.1088/978-1-6817-4405-6>.
- [47] E. Wigner, "On the quantum correction for thermodynamic equilibrium," *Phys. Rev.*, vol. 40, pp. 749–759, 5 1932. DOI: 10.1103/PhysRev.40.749. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRev.40.749>.
- [48] "Quantum states in phase space," in *Quantum Optics in Phase Space*. John Wiley Sons, Ltd, 2001, ch. 4, pp. 99–151, ISBN: 9783527602971. DOI: <https://doi.org/10.1002/3527602976.ch4>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/3527602976.ch4>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/3527602976.ch4>.
- [49] H.-W. Lee, "Theory and application of the quantum phase-space distribution functions," *Physics Reports*, vol. 259, no. 3, pp. 147–211, 1995, ISSN: 0370-1573. DOI: [https://doi.org/10.1016/0370-1573\(95\)00007-4](https://doi.org/10.1016/0370-1573(95)00007-4). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0370157395000074>.
- [50] W. B. Case, "Wigner functions and weyl transforms for pedestrians," *American Journal of Physics*, vol. 76, no. 10, pp. 937–946, 2008. DOI: 10.1119/1.2957889. eprint: <https://doi.org/10.1119/1.2957889>. [Online]. Available: <https://doi.org/10.1119/1.2957889>.
- [51] R. Simon, N. Mukunda, and B. Dutta, "Quantum-noise matrix for multimode systems: U(n) invariance, squeezing, and normal forms," *Phys. Rev. A*, vol. 49, pp. 1567–1583, 3 1994. DOI: 10.1103/PhysRevA.49.1567. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.49.1567>.
- [52] A. Holevo, "Some statistical problems for quantum gaussian states," *IEEE Transactions on Information Theory*, vol. 21, no. 5, pp. 533–543, 1975. DOI: 10.1109/TIT.1975.1055441.
- [53] A. Serafini, *Quantum Continuous Variables: A Primer of Theoretical Methods*, 1st ed. Boca Raton: CRC Press, 2017.
- [54] F. Laudenbach, C. Pacher, C.-H. F. Fung, A. Poppe, M. Peev, B. Schrenk, M. Hentschel, P. Walther, and H. Hübel, "Continuous-variable quantum key distribution with gaussian modulation—the theory of practical implementations," *Advanced Quantum Technologies*, vol. 1, no. 1, p. 1800011, 2018. DOI: <https://doi.org/10.1002/qute.201800011>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/qute.201800011>.
- [55] Arvind, B Dutta, N Mukunda, and R Simon, "The real symplectic groups in quantum mechanics and optics," *Pramana*, vol. 45, no. 6, 471–497, 1995, ISSN: 0973-7111. DOI: 10.1007/bf02848172. [Online]. Available: <http://dx.doi.org/10.1007/BF02848172>.
- [56] S. L. Braunstein and P. van Loock, "Quantum information with continuous variables," *Rev. Mod. Phys.*, vol. 77, pp. 513–577, 2 2005. DOI: 10.1103/RevModPhys.77.513. [Online]. Available: <https://link.aps.org/doi/10.1103/RevModPhys.77.513>.
- [57] M. G. Genoni, L. Lami, and A. Serafini, "Conditional and unconditional gaussian quantum dynamics," *Contemporary Physics*, vol. 57, no. 3, 331–349,

- 2016, ISSN: 1366-5812. DOI: 10.1080/00107514.2015.1125624. [Online]. Available: <http://dx.doi.org/10.1080/00107514.2015.1125624>.
- [58] J. Williamson, "On the algebraic problem concerning the normal forms of linear dynamical systems," *American Journal of Mathematics*, vol. 58, no. 1, pp. 141–163, 1936, ISSN: 00029327, 10806377. [Online]. Available: <http://www.jstor.org/stable/2371062>.
- [59] J. Lodewyck, M. Bloch, R. García-Patrón, S. Fossier, E. Karpov, E. Diamanti, T. Debuisschert, N. J. Cerf, R. Tualle-Brouri, S. W. McLaughlin, and et al., "Quantum key distribution over 25 km with an all-fiber continuous-variable system," *Physical Review A*, vol. 76, no. 4, 2007, ISSN: 1094-1622. DOI: 10.1103/PhysRevA.76.042305. [Online]. Available: <http://dx.doi.org/10.1103/PhysRevA.76.042305>.
- [60] F. Grosshans, N. J. Cerf, J. Wenger, R. Tualle-Brouri, and P. Grangier, "Virtual entanglement and reconciliation protocols for quantum cryptography with continuous variables," vol. 3, no. 7, 535–552, Oct. 2003, ISSN: 1533-7146.
- [61] A. Leverrier and P. Grangier, *Continuous-variable quantum key distribution protocols with a discrete modulation*, 2011. arXiv: 1002.4083 [quant-ph].
- [62] A. Becir, F. A. A. El-Orany, and M. R. B. Wahiddin, *Continuous-variable quantum key distribution protocols with eight-state discrete modulation*, 2010. arXiv: 1006.4216 [quant-ph].
- [63] A. Serafini, F. Illuminati, and S. D. Siena, "Symplectic invariants, entropic measures and correlations of gaussian states," *Journal of Physics B: Atomic, Molecular and Optical Physics*, vol. 37, no. 2, pp. L21–L28, 2003. DOI: 10.1088/0953-4075/37/2/102. [Online]. Available: <https://doi.org/10.1088/0953-4075/37/2/102>.
- [64] M. Navascués, F. Grosshans, and A. Acín, "Optimality of gaussian attacks in continuous-variable quantum cryptography," *Phys. Rev. Lett.*, vol. 97, p. 190502, 19 2006. DOI: 10.1103/PhysRevLett.97.190502. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.97.190502>.
- [65] I. Devetak and A. Winter, "Distillation of secret key and entanglement from quantum states," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 461, no. 2053, pp. 207–235, 2005. DOI: 10.1098/rspa.2004.1372. eprint: <https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.2004.1372>. [Online]. Available: <https://royalsocietypublishing.org/doi/abs/10.1098/rspa.2004.1372>.
- [66] R. García-Patrón and N. J. Cerf, "Unconditional optimality of gaussian attacks against continuous-variable quantum key distribution," *Phys. Rev. Lett.*, vol. 97, p. 190503, 19 2006. DOI: 10.1103/PhysRevLett.97.190503. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.97.190503>.
- [67] A. Leverrier and P. Grangier, "Simple proof that gaussian attacks are optimal among collective attacks against continuous-variable quantum key distribution with a gaussian modulation," *Phys. Rev. A*, vol. 81, p. 062314, 6 2010. DOI: 10.1103/PhysRevA.81.062314. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.81.062314>.

- [68] R. Renner and J. I. Cirac, "De finetti representation theorem for infinite-dimensional quantum systems and applications to quantum cryptography," *Phys. Rev. Lett.*, vol. 102, p. 110504, 11 2009. DOI: 10.1103/PhysRevLett.102.110504. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.102.110504>.
- [69] A. Leverrier, "Security of continuous-variable quantum key distribution via a gaussian de finetti reduction," *Phys. Rev. Lett.*, vol. 118, p. 200501, 20 2017. DOI: 10.1103/PhysRevLett.118.200501. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.118.200501>.
- [70] J. Lin, T. Upadhyaya, and N. Lütkenhaus, "Asymptotic security analysis of discrete-modulated continuous-variable quantum key distribution," *Phys. Rev. X*, vol. 9, p. 041064, 4 2019. DOI: 10.1103/PhysRevX.9.041064. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevX.9.041064>.
- [71] S. Ghorai, P. Grangier, E. Diamanti, and A. Leverrier, "Asymptotic security of continuous-variable quantum key distribution with a discrete modulation," *Phys. Rev. X*, vol. 9, p. 021059, 2 2019. DOI: 10.1103/PhysRevX.9.021059. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevX.9.021059>.
- [72] A. Denys, P. Brown, and A. Leverrier, "Explicit asymptotic secret key rate of continuous-variable quantum key distribution with an arbitrary modulation," *Quantum*, vol. 5, p. 540, 2021. DOI: 10.22331/q-2021-09-13-540.
- [73] N. Grote, M. Baier, and F. Soares, "Photonic integrated circuits on inp," in *Fibre Optic Communication: Key Devices*, H. Venghaus and N. Grote, Eds. Cham: Springer International Publishing, 2017, pp. 799–840.
- [74] Victor DOLORES-CALZADILLA*, Francisco M. SOARES, Moritz BAIER, Tom GAERTNER, Mike FEYER, Martin MÖHRLE, Norbert GROTE, Martin SCHELL, "Inp-based photonic integration platform: Status and prospects," in *European Conference on Integrated Optics*, 2016. [Online]. [Online]. Available: <https://www.ecio-conference.org/wp-content/uploads/2016/06/ECIO-o-37.pdf>.
- [75] JePPiX. (). "Performance Summary Table," apr. 30, 2021, [Online]. Available: <https://www.jeppix.eu/mpw-services/get-started/performance-summary-table/>.
- [76] D. Melati, A. Waqas, A. Alippi, and A. Melloni, "Wavelength and composition dependence of the thermo-optic coefficient for ingaasp-based integrated waveguides," *Journal of Applied Physics*, vol. 120, no. 21, p. 213102, 2016. DOI: 10.1063/1.4970937.
- [77] F. Y. Gardes, G. T. Reed, G. Z. Mashanovich, and C. E. Png, "Optical modulators in silicon photonic circuits," in *Silicon Photonics*. John Wiley Sons, Ltd, 2008, ch. 4, pp. 95–145, ISBN: 9780470994535. DOI: <https://doi.org/10.1002/9780470994535.ch4>.
- [78] Microwaves101. (). "Insertion Loss," apr. 30, 2021, [Online]. Available: <https://www.microwaves101.com/encyclopedias/insertion-loss>.

- [79] L. Soldano and E. Pennings, "Optical multi-mode interference devices based on self-imaging: Principles and applications," *Journal of Lightwave Technology*, vol. 13, no. 4, pp. 615–627, 1995. DOI: 10.1109/50.372474.
- [80] M. Bachmann, P. A. Besse, and H. Melchior, "General self-imaging properties in $n \times n$ multimode interference couplers including phase relations," *Appl. Opt.*, vol. 33, no. 18, pp. 3905–3911, 1994. DOI: 10.1364/AO.33.003905.
- [81] "Optical modulators and modulation schemes," in *Fiber Optic Communications*. John Wiley Sons, Ltd, 2014, ch. 4, pp. 139–187, ISBN: 9781118684207. DOI: <https://doi.org/10.1002/9781118684207.ch4>.
- [82] K.-P. Ho and H.-W. Cui, "Generation of arbitrary quadrature signals using one dual-drive modulator," *Journal of Lightwave Technology*, vol. 23, no. 2, pp. 764–770, 2005. DOI: 10.1109/JLT.2004.838855.
- [83] D. Pérez-Galacho, L. Bramerie, C. Baudot, M. Chaibi, S. Messaoudène, N. Vulliet, L. Vivien, F. Boeuf, C. Peucheret, and D. Marris-Morini, "Qpsk modulation in the o-band using a single dual-drive mach-zehnder silicon modulator," *Journal of Lightwave Technology*, vol. 36, no. 18, pp. 3935–3940, 2018. DOI: 10.1109/JLT.2018.2851370.
- [84] A. H. Gnauck and P. J. Winzer, "Optical phase-shift-keyed transmission," *J. Lightwave Technol.*, vol. 23, no. 1, p. 115, 2005.
- [85] M. Baier, "Polarization multiplexed photonic integrated circuits for 100 gbit/s and beyond," M.S. thesis, Dept. of Math. and Phy. Sci., Berlin, Germany, 2018.
- [86] D. O. Dzibrou, J. J. G. M. van der Tol, and M. K. Smit, "Extremely efficient two-section polarization converter for InGaAsP-InP photonic integrated circuits," in *2013 Conference on Lasers & Electro-Optics Europe & International Quantum Electronics Conference CLEO EUROPE/IQEC*, IEEE, 2013, pp. 1–1, ISBN: 978-1-4799-0594-2. DOI: 10.1109/CLEOE-IQEC.2013.6801452. [Online]. Available: <http://ieeexplore.ieee.org/document/6801452/>.
- [87] M. Baier, F. Soares, T. Gaertner, R. Weiser, M. Moehrle, N. Grote, and M. Schell, "Highly fabrication tolerant polarization converter for generic photonic integration technology," in *2016 Compound Semiconductor Week (CSW) [Includes 28th International Conference on Indium Phosphide & Related Materials (IPRM) & 43rd International Symposium on Compound Semiconductors (ISCS)*, IEEE, 2016. DOI: 10.1109/iciprm.2016.7528537. [Online]. Available: <https://doi.org/10.1109%2Ficiprm.2016.7528537>.
- [88] Moritz BAIER, Francisco M. SOARES, Martin MOEHRLE, Norbert GROTE and Martin SCHELL, "A new approach to designing polarization rotating waveguides," in *European Conference on Integrated Optics*, 2016. [Online]. [Online]. Available: <http://www.ecio-conference.org/wp-content/uploads/2016/06/ECIO-o-37.pdf>.
- [89] M. Baier, F. M. Soares, T. Gartner, M. Moehrle, and M. Schell, "Fabrication Tolerant Integrated Polarization Rotator Design Using the Jones Calculus," *Journal of Lightwave Technology*, vol. 37, no. 13, pp. 3106–3112, 2019, ISSN: 0733-8724. DOI: 10.1109/JLT.2019.2910931. [Online]. Available: <https://ieeexplore.ieee.org/document/8691446/>.

- [90] L. Augustin, "Polarization handling in photonic integrated circuits," Ph.D. thesis, Dept. of Elec. Engg., TU Eindhoven, Netherlands, 2008.
- [91] Ruiyong Zhang, Heinz-Gunter Bach, Reinhard Kunkel, Klemens Janiak, Patrick Runge, Gan Zhou, Angela Seeger, S. Schubert, Iñigo Molina Fernandez, Alejandro Ortega-Moñux, Robert Halir, Diego Pérez Galacho, "A polarization beam splitter based on mzi-structure," in *European Conference on Integrated Optics*, 2012. [Online]. Available: <http://www.ecio-conference.org/wp-content/uploads/2016/06/ECIO-o-37.pdf>.
- [92] M. F. Baier, F. M. Soares, W. Passenberg, T. Gaertner, M. Moehrlle, and N. Grote, "Polarization beam splitter building block for InP based generic photonic integrated circuits," in *26th International Conference on Indium Phosphide and Related Materials (IPRM)*, IEEE, 2014. DOI: 10.1109/iciprm.2014.6880576. [Online]. Available: <https://doi.org/10.1109%2Ficiprm.2014.6880576>.
- [93] W. Streifer, R. Burnham, and D. Scifres, "Effect of external reflectors on longitudinal modes of distributed feedback lasers," *IEEE Journal of Quantum Electronics*, vol. 11, no. 4, pp. 154–161, 1975. DOI: 10.1109/JQE.1975.1068581.
- [94] A. Yariv, *Photonics: Optical Electronics in Modern Communication*, 6th ed. New York, NY, USA: Oxford University Press, 2006.
- [95] G. Morthier and P. Vankwikelberge. 2013.
- [96] H. Kogelnik and C. V. Shank, "Coupled-wave theory of distributed feedback lasers," *Journal of Applied Physics*, vol. 43, no. 5, pp. 2327–2335, 1972. DOI: 10.1063/1.1661499.
- [97] G. Hadjicostas, J. Butler, G. Evans, N. Carlson, and R. Amantea, "A numerical investigation of wave interactions in dielectric waveguides with periodic surface corrugations," *IEEE Journal of Quantum Electronics*, vol. 26, no. 5, pp. 893–902, 1990. DOI: 10.1109/3.55531.
- [98] A. Yariv and A. Gover, "Equivalence of the coupled-mode and floquet-bloch formalisms in periodic optical waveguides," *Applied Physics Letters*, vol. 26, no. 9, pp. 537–539, 1975. DOI: 10.1063/1.88247.
- [99] L. Carroll, J.-S. Lee, C. Scarcella, K. Gradkowski, M. Duperron, H. Lu, Y. Zhao, C. Eason, P. Morrissey, M. Rensing, S. Collins, H. Y. Hwang, and P. O'Brien, "Photonic packaging: Transforming silicon photonic integrated circuits into photonic devices," *Applied Sciences*, vol. 6, no. 12, 2016, ISSN: 2076-3417. DOI: 10.3390/app6120426. [Online]. Available: <https://www.mdpi.com/2076-3417/6/12/426>.
- [100] Cordon Electronics, *Cordon Electronics PDK*, Accessed Sep. 19, 2021 [Online]. Available: <https://www.cordongroup.it/Wp/wp-content/uploads/2018/08/CORDON-design-rules-for-photonic-packaging-August-2018.pdf>.
- [101] M. M. Wilde, "Preface to the second edition," *Quantum Information Theory*, xi–xii, DOI: 10.1017/9781316809976.001. [Online]. Available: <http://dx.doi.org/10.1017/9781316809976.001>.

- [102] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. USA: Wiley-Interscience, 2006, ISBN: 0471241954.
- [103] A. Winter, “Extrinsic and intrinsic data in quantum measurements: Asymptotic convex decomposition of positive operator valued measures,” *Communications in Mathematical Physics*, vol. 244, no. 1, 157–185, 2004, ISSN: 1432-0916. DOI: 10.1007/s00220-003-0989-z. [Online]. Available: <http://dx.doi.org/10.1007/s00220-003-0989-z>.
- [104] P. Sen, “Achieving the han-kobayashi inner bound for the quantum interference channel,” in *2012 IEEE International Symposium on Information Theory Proceedings*, 2012, pp. 736–740. DOI: 10.1109/ISIT.2012.6284656.
- [105] B. Schumacher and M. D. Westmoreland, “Sending classical information via noisy quantum channels,” *Phys. Rev. A*, vol. 56, pp. 131–138, 1 1997. DOI: 10.1103/PhysRevA.56.131. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.56.131>.
- [106] W. Zhigang, “Study on low-loss si waveguides and photonic integrated circuit devices using multi-mode interference waveguides for optical networks,” Ph.D. thesis, Dept. of Information and Telecom. studies, Waseda University, Japan, 2005.
- [107] Wikipedia, “waveplate” apr. 6, 2020. [Online]. Available: <https://en.wikipedia.org/wiki/Waveplate>.
- [108] Texas Instruments, *What is an eye diagram*. Accessed Apr. 1, 2022 [Online]. [Online]. Available: https://www.youtube.com/watch?v=tZiKRfH2yZ4&t=355s&ab_channel=TexasInstruments.